The Diophantine moment problem and the analytic structure in the activity of the ferromagnetic Ising model

M. Barnsley, D. Bessis, and P. Moussa

Cen Saclay, BP n° 2, 91190 Gif-Sur-Yvette, France

(Received 18 July 1978)

We show that the intensity of magnetization $I(z,x)$ where $z = e^{-2\beta H}$ and $x = e^{-2\beta J}$, for the ferromagnetic Ising model in arbitrary dimension, reduces, for rational values of x, to a Diophantine moment problem

$I(z) = \Sigma_0^\infty n_k z^k$,

where

$n_k = \int_0^\Lambda \sigma(\lambda)\lambda^k d\lambda$,

$\sigma(\lambda)$ is a positive measure, $n_0 = 1/2$, and n_k is integer for $k \neq 0$. The fact that the n_k are positive integers puts very stringent constraints on the measure $\sigma(\lambda)$. One of the simplest results we obtain is that for $\Lambda < 4$, $\sigma(\lambda)$ is necessarily a finite sum of Dirac δ functions whose support is of the form $4\cos^2(p\pi/m)$, $p = 0,1,2,\ldots,m-1$, with m a finite integer. For $\Lambda = 4$, which correspond to the one-dimensional Ising model, we have the result that either $I(z)$ is a rational fraction belonging to the previous class $\Lambda < 4$, or $I(z) = (1/2)(1-4z)^{-1/2}$ which corresponds precisely to the exact answer for dimension 1. For $\Lambda > 4$, which is associated with Ising models in dimension $d \geq 2$ we show that all cases are reducible to $\Lambda = 6$, by a quadratic transformation which transforms integers into integers and positive measures into positive measures. The fixed point of this type of transformation is analyzed in great detail and is shown to provide a devil's staircase measure. Various other results are also discussed as well as conjectures.

INTRODUCTION

While the ferromagnetic Ising model in two dimensions and zero field is well understood,[1] comparatively little is known in the presence of a magnetic field. The Lee–Yang representation[2] although very interesting has been little exploited. In particular the analytic nature of the singularities in the complex activity plane on the circle $|z| = 1$ is not known, and the possibility that below the critical temperature T_c when the circle is closed there exists a natural frontier has not been excluded. It thus seems important to know how to characterize the class of analytic functions to which the thermodynamic quantities belong when considered as functions over the complex activity plane.

An associated question is that of the critical indices which appear, at least for those which are exactly computable in two dimensions, to be rational numbers. Is this a general feature of the Ising model, and via the universality principle a characteristic of classes of physical processes? If it is the case, then it would be interesting to be able to classify these processes and to understand how such rational indices are generated.

In this paper we shall only be concerned with the ordinary ferromagnetic Ising model on a lattice of dimension d with c nearest neighbors, with interaction limited to the nearest neighbors.

In this case, the perturbative expansion of thermodynamical quantities are of the form

$$\tau(z,x) = \sum_0^\infty \bar{\mu}_l(x)\, z^l,$$

where x is the usual temperature variable $x = \exp(-2\beta J)$ and z is the activity variable, $z = \exp(-2\beta H)$.

The Mayer–Yvon coefficients which arise in the z expansion of the free energy are *polynomials in x with integer coefficients*. Furthermore, from the Lee–Yang theorem, we know that these coefficients are moments of a positive measure according to

$$\bar{\mu}_l(x) = \int_0^\pi g(\theta,x)\cos l\theta\, d\theta, \quad g(\theta,x) > 0.$$

We are thus led to the consideration of a trigonometrical moment problem on the ring of polynomials with integer coefficients.

Such a problem defines a specific class of analytic functions. It is the aim of this paper to set in motion an investigation of the content of such a class. Unexpectedly, the fact that the moments must belong to a ring introduces in many cases stringent constraints on the measure $g(\theta,x)$, as will be seen in the sequel.

In Sec. I we show how the Ising model is associated with a moment problem on a ring and how in the case of rational values of x this reduces to the following ordinary moment problem on the ring of integers: *Find a positive measure $\sigma(\lambda)$ such that*

$$n_k = \int_0^\Lambda \sigma(\lambda)\,\lambda^k\, d\lambda, \quad k = 0,1,2,\cdots,$$

where each n_k is an integer. The nature of $\sigma(\lambda)$ depends crucially on Λ. The problem is explicitly solved in Sec. II for the case $0 \leq \Lambda \leq 4$, and over this range the measures $\sigma_\Lambda(\lambda)$ are found to be quantized.

In Sec. III we consider the case $\Lambda > 4$ and show that it can always be reduced to the situation when $\Lambda = 6$. Associated with $\Lambda = 6$ is an interesting transformation—its fixed point belongs to a measure whose support is contained in and naturally associated with a Cantor set.[3] The analogous transformation in the case $\Lambda = 4$ is associated with the solution of the one-dimensional Ising model.

 0022-2488/79/040535-12$01.00

In Sec. IV we discuss various examples which cast light on the unsolved part of the problem ($4 < \Lambda \leq 6$), and in conclusion we examine several conjectures.

1. THE ISING MODEL VIEWED AS A MOMENT PROBLEM IN A RING

The ferromagnetic Ising model on a lattice of dimension d with c nearest neighbors is described by the Hamiltonian

$$H = -J\sum_{i,j} \sigma_i\sigma_j - H\sum_i \sigma_i, \quad J > 0, \tag{1.1}$$

the first sum being performed over nearest neighbor pairs and the second over all sites of the lattice. We introduce the notations

$$x = \exp(-2\beta J), \quad 0 \leq x \leq 1, \quad \text{and } z = \exp(-2\beta H), \tag{1.2}$$

where H is the magnetic field. Then, following Lee and Yang,[2] the intensity of magnetization per site is found to be

$$I(z,x) = 2(1-z^2)\int_{\theta_0(x)}^{\pi} \frac{g(\theta,x)}{1-2z\cos\theta+z^2}\, d\theta, \tag{1.3}$$

where $\theta_0(x)$ is the Lee–Yang angle which vanishes for $x < x_c$ where x_c corresponds to the critical temperature, see Fig. 1. $g(\theta, x)$ is a positive measure, being the density of zeros of the grand partition function on the circle $|z| = 1$ in the complex activity plane in the thermodynamic limit. The representation (1.3) is also valid when the lattice has only finitely many sites, in which case the measure simply consists of a finite sum of delta functions with positive weights.

The measure g is normalized according to

$$\int_{\theta_0(x)}^{\pi} g(\theta,x)\, d\theta = \tfrac{1}{2}, \tag{1.4}$$

and $I(z,x)$ has the property that

$$I(z,x) = -I(1/z,x) \tag{1.5}$$

which refers to the symmetry of the system under reversal of the magnetic field. In Fig. 2 we represent the domain of analyticity of $I(z,x)$ in z.

Developing $I(z,x)$ around $z=0$ we obtain the Mayer–Yvon expansion

$$I(z,x) = 1 - 2\sum_{l\geq 1} lM_l(x)\, z^l, \tag{1.6}$$

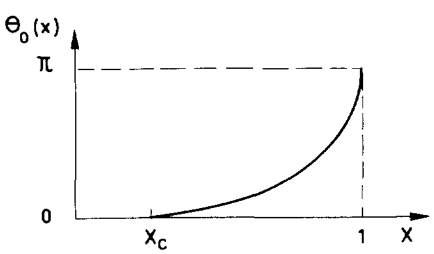


FIG. 1. The Lee–Yang angle $\theta_0(x)$ as a function of x. It vanishes when $x < x_c$.

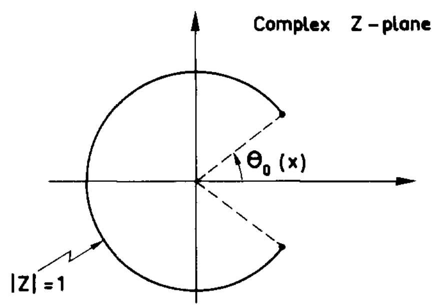


FIG. 2. When $\theta_0(x) > 0$, $I(z,x)$ is regular both inside and outside the unit circle, and also for $|z|=1$ when $|\arg z| < \theta_0(x)$. When $\theta_0(x) = 0$ its singularities may be dense on the unit circle.

with

$$lM_l(x) = -2\int_{\theta_0(x)}^{\pi} g(\theta,x)\cos(l\theta)\, d\theta, \quad l = 1, 2, 3, \cdots. \tag{1.7}$$

This defines the trigonometrical moment problem, whose moments have the following properties:

(i) $M_1(x) = x^c$ where c is the number of nearest neighbors.

(ii) $lM_l(x)$ is a polynomial of degree lc in x whose parity is that of the highest degree term.

(iii) All coefficients of $lM_l(x)$ are integers.[4]

The proper moment problem associated with (1.7) is

$$\bar{\mu}_l(x) = \int_{\theta_0(x)}^{\pi} g(\theta,x)[\cos^2(\theta/2)]^l\, d\theta, \tag{1.8}$$

where $\bar{\mu}_l(x)$ and $lM_l(x)$ are connected by

$$4^l\bar{\mu}_l(x) = \binom{2l}{l} M_0 - \sum_{p=1}^{l}\binom{2l}{l-p} pM_p(x), \tag{1.9}$$

$$lM_l(x) = (-1)^{l+1}2\bar{\mu}_0 + \sum_{p=1}^{l}(-1)^{l-p-1}\frac{l}{p}\binom{l+p-1}{l-p}4^p\bar{\mu}_p(x), \tag{1.10}$$

and, for $l = 0$,

$$\bar{\mu}_0 = M_0 = \tfrac{1}{2}. \tag{1.11}$$

A consequence of (i), (ii), and (iii) is that $4^l\bar{\mu}_l(x)$ is a polynomial with integer coefficients. A deeper result, proved in Ref. 5, is that in fact this polynomial is exactly divisible by $(1-x)^l$ and we have

$$4^l\bar{\mu}_l(x) = (1-x)^l P_l(x), \tag{1.12}$$

where $P_l(x)$ is of degree $l(c-1)$ and has integer coefficients. *When c is even* the $\bar{\mu}_l(x)$'s are even polynomials so that setting $u = x^2$ we have

$$4^l\bar{\mu}_l(x) = (1-u)^l P_l(u), \tag{1.13}$$

where $P_l(u)$ is of degree $l(c/2-1)$. For instance, for both the square and diamond lattices $c=4$, the degree of P_l is exactly l. Thus from now onwards *we will restrict attention to the case where c is even.*

We introduce the variable

$$v = 4z/(1+z)^2 = 1 - \tanh^2\beta H. \tag{1.14}$$

Then the generating function for the $\bar{\mu}_l(x)$'s is given by the intensity of magnetization[6] in the new variable, namely

$$I(v,x)=2\sqrt{1-v}\sum_{l=0}^{\infty} v^l\mu_l(x)=2\sqrt{1-v}$$

$$\times\int_{\theta_0(x)}^{\pi}\frac{g(\theta,x)}{1-v\cos^2(\theta/2)}\,d\theta. \qquad (1.15)$$

Setting

$$w=(1-u)\,v/4=z(1-u)/(1+z)^2 \quad\text{and}\quad \xi=\frac{4}{1-u}\circ\cos^2(\theta/2), \qquad (1.16)$$

we obtain

$$\frac{I(w,u)}{2(1-4w/(1-u))^{1/2}}=\sum_{l=0}^{\infty} w^l P_l(u)=\int_0^L\frac{\bar g(\xi,u)\,d\xi}{1-\xi w}, \qquad (1.17)$$

where

$$\bar g(\xi,u)=\frac{g(2\mathrm{Arc}\cos(\tfrac12\sqrt{\xi(1-u)}),u)}{\sqrt{\xi(4/(1-u)-\xi)}}\quad,$$

and (1.18)

$$L=\frac{4}{1-u}\cos^2(\theta_0/2).$$

This shows that the $P_l(u)$ are moments of the positive measure $\bar g(\xi,u)$ over its support $0\leq\xi\leq L$, that is

$$P_l(u)=\int_0^L\bar g(\xi,u)\,\xi^l\,d\xi,\quad l=1,2,3,\cdots. \qquad (1.19)$$

When $u<u_c$ (where u_c is the value of u corresponding to the critical temperature), θ_0 is zero. When $u\to 1$, $\theta_0\to\pi$, but we have the asymptotic bound[7]

$$\frac{4}{1-u}\cos^2(\theta_0/2)\big|_{u=1}\leq c^2, \qquad (1.20)$$

so that L remains finite for all $0\leq u\leq 1$. On the other hand, explicit calculations up to the highest available order have indicated that all of the coefficients of $P_l(u)$ are positive.[5] If this is true for all orders, then it follows that $L=\overline{\lim}[P_l(u)]^{1/l}$ is a monotone increasing function of $u>0$, as is illustrated in Fig. 3 for the case $c=4$. Note that for $c=2$, $P_l(u)$ is simply a positive constant so that L is independent of u and equal to 4 as seen from (1.20).

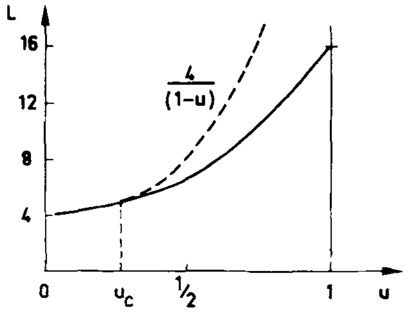


FIG. 3. L as a function of u for $c=4$. The dotted curve is $4/(1-u)$ which equals L just up to u_c.

We come now to the problem we are interested in. Consider a positive measure $\bar g(\xi,u)$ for which the moments

$$P_l(u)=\int_0^L\bar g(\xi,u)\,\xi^l\,d\xi,\quad l=0,1,2,\cdots \qquad (1.21)$$

are known to belong to a ring, for example the ring of polynomials with integer coefficients. Then is it possible to characterize, or define in a precise way, the analytic structure of the generating function

$$G(w)=\sum_{l=0}^{\infty} w^l P_l(u). \qquad (1.22)$$

One can envisage eventually imposing the additional constraint that the coefficient of the $P_l(u)$'s are positive.

For simplicity in this first attempt we shall consider the reduced problem obtained by choosing u to be a rational number,

$$u=p/r,\quad\text{where } p \text{ and } r \text{ are positive integers, with } p\leq r. \qquad (1.23)$$

Then $P_l(p/r)$ takes the form of an integer n_l divided by $r^{l(c/2-1)}$, so that

$$n_l=r^{l(c/2-1)}\int_0^L\bar g(\xi,p/r)\,\xi^l\,d\xi$$

$$=\int_0^{Lr^{(c/2-1)}} r^{1-c/2}\bar g(\lambda r^{1-c/2},p/r)\lambda^l d\lambda, \qquad (1.24)$$

where we have set $\lambda=\xi r^{c/2-1}$. We are thus led to consider the *Diophantine moment problem*

$$n_l=\int_0^{\Lambda}\sigma(\lambda)\lambda^l d\lambda,\quad l=0,1,2,\cdots, \qquad (1.25)$$

where $n_0=\frac12$, n_l is a positive integer for $l\geq 1$, $\sigma(\lambda)$ is a positive measure with support $0\leq\lambda\leq\Lambda$, and where Λ can be chosen to be a positive integer, without loss of generality. Notice that by the trivial modification $\sigma(\lambda)\to\sigma(\lambda)+\frac12\delta(\lambda)$ we can if we like take $n_0=1$. Our new moment problem is still posed over a ring—this time it's the ring of integers.

The generating function associated with (1.25) is

$$G(w)=\sum_{l=0}^{\infty} n_l w^l=\int_0^{\Lambda}\frac{\sigma(\lambda)d\lambda}{1-w\lambda}. \qquad (1.26)$$

This is in fact a Stieltjes function, see Ref. 8, and is holomorphic in the w plane cut from $1/\Lambda$ to $+\infty$, as shown in Fig. 4. The important point which we will demonstrate and explore in the rest of this paper is that the positivity of the measure in (1.26) combined with the fact that the n_l's are integers imposes stringent constraints on the nature of $\sigma(\lambda)$ and hence on the analytic character of $G(w)$.

We mention two practical situations where results along the above lines would be directly applicable.

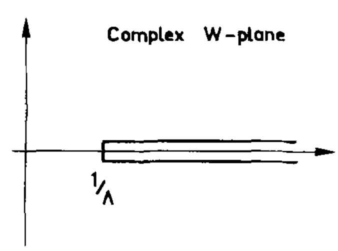


FIG. 4. The location of the cut for the Stieltjes function $G(w)$; elsewhere this function is holomorphic.

First, for the triangular planar lattice which has $d=2$ and $c=6$, the value of u_c is $\frac{1}{3}$.[4] Therefore, on the critical isotherm we have a Diophantine moment problem (D. M. P.) with $L=6$, and $r=3$ so that $\Lambda=54$. Second, for $u=1$ the generating function $G(w,u)$ reduces to the Monomer—Dimer partition function[9,10] which appears as the solution of a D.M.P. with $\Lambda=c^2$.

2. THE DIOPHANTINE MOMENT PROBLEM WITH FINITE SUPPORT

A. The problem

Given that $\sigma(\lambda)$ is a positive measure defined on $0\leqslant\lambda\leqslant\Lambda<\infty$, such that all of the moments

$$n_k=\int_0^\Lambda \lambda^k\sigma(\lambda)\,d\lambda,\quad k=0,1,2,\cdots \tag{2.1}$$

are finite integers (with the occasional exception that n_0 is half-integer) we ask what can be said about the generating function

$$G(w)=\int_0^\Lambda \frac{\sigma(\lambda)d\lambda}{1-\lambda w},\quad w\in\mathbb{C}. \tag{2.2}$$

In all that follows the support is taken to be the closed interval $[0,\Lambda]$ so that if $\sigma(\lambda)$ has delta function contributions at either end point then these must be included in the evaluation of such integrals as (2.1) and (2.2).

The nature of $G(\omega)$ depends critically on Λ, as we shall see, and we begin by examining the simplest cases.

B. The problem when Λ = 1

The fact that $\{\lambda^k\}_{k=0}^\infty$ forms a monotone nonincreasing sequence of functions defined on $[0,1]$, together with the positivity of $\sigma(\lambda)$ implies in this case the inequalities

$$n_0\geqslant n_1\geqslant n_2\geqslant\cdots\geqslant 0. \tag{2.3}$$

It follows that the sequence $\{n_k\}_{k=0}^\infty$ converges after finitely many steps to a limiting value n_{k_0} such that

$$n_{k_0}=n_{k_0+1}=n_{k_0+2}=\cdots. \tag{2.4}$$

Hence

$$n_{k_0}-n_{k_0+1}=\int_0^1\lambda^{k_0}(1-\lambda)\sigma(\lambda)d\lambda=0, \tag{2.5}$$

and since $\lambda^{k_0}(1-\lambda)$ is strictly positive when $\lambda\in(0,1)$ we discover that

$$\sigma(\lambda)=\sigma_0\delta(\lambda)+\sigma_1\delta(\lambda-1), \tag{2.6}$$

where σ_0 and σ_1 are both nonnegative. Using (2.1) we now find

$$\sigma_0=n_0-n_1\geqslant 0,\quad \sigma_1=n_1\geqslant 0, \tag{2.7}$$

and

$$G(w)=(n_0-n_1)+n_1/(1-w). \tag{2.8}$$

We conclude that the Diophantine moment problem on $[0,1]$ has a solution if and only if $n_0\geqslant n_1=n_2=n_3=\cdots\geqslant 0$ and in this case $G(w)$ is a $[1/1]$ rational fraction which is uniquely specified by the first two terms in its expansion about zero. Note that if $0<\Lambda<1$ then the problem has a solution if and only if $n_0\geqslant n_1=n_2=\cdots=0$ and in this case $G(w)$ is a constant.

C. The problem when Λ = 2

Noting that $\{\lambda^k(2-\lambda)^k\}_{k=0}^\infty$ forms a monotone nonincreasing sequence of functions defined on $[0,2]$, the sequence of integers

$$m_k=\int_0^2\lambda^k(2-\lambda)^k\sigma(\lambda)\,d\lambda,\quad k=0,1,2,\cdots \tag{2.9}$$

must be nonincreasing. Thus we again find that there exists an integer k_0 such that

$$m_{k_0}=m_{k_0+1}=m_{k_0+2}=\cdots, \tag{2.10}$$

so that

$$m_{k_0}-m_{k_0+1}=\int_0^2\lambda^{k_0}(2-\lambda)^{k_0}(1-\lambda)^2\sigma(\lambda)d\lambda=0, \tag{2.11}$$

and hence

$$\sigma(\lambda)=\sigma_0\delta(\lambda)+\sigma_1\delta(\lambda-1)+\sigma_2\delta(\lambda-2), \tag{2.12}$$

where the c_i's are nonnegative. Identification with the first three moments now provides

$$\sigma_0=n_0-\frac{3n_1}{2}+\frac{n_2}{2},\quad \sigma_1=2n_1-n_2,\quad \sigma_2=\frac{(n_2-n_1)}{2} \tag{2.13}$$

and

$$G(w)=\frac{n_0-(3n_0-n_1)w+(2n_0-3n_1+n_2)w^2}{1-3w+2w^2}. \tag{2.14}$$

We conclude that the Diophantine moment problem on $[0,2]$ has a solution if and only if $2n_1\geqslant n_2\geqslant \mathrm{Max}\{n_1, 3n_1-2n_0\}$, $n_0\geqslant 0$, and $n_k=(2n_1-n_2)+(n_2-n_1)2^{k-1}$ for $k\geqslant 2$; and in this case $G(w)$ is a $[2/2]$ rational fraction which is uniquely specified by the first three terms in its expansion about zero.

D. The problem when $\Lambda<4$

On setting $\lambda=4\cos^2(\theta/2)$ and defining $\Lambda=4\cos^2(\theta_0/2)$ we obtain

$$2n_k/4^k=\int_{\theta_0}^\pi[\cos^2(\theta/2)]^k g(\theta)\,d\theta,\quad k=0,1,2,\cdots, \tag{2.15}$$

where

$$\begin{aligned}g(\theta)&=8\cos(\theta/2)\sin(\theta/2)\,\sigma(4\cos^2(\theta/2)),\\ \sigma(\lambda)&=2[\lambda(4-\lambda)]^{-1/2}g(2\mathrm{Arc}\cos(\sqrt{\lambda}/2)).\end{aligned} \tag{2.16}$$

The positivity of the measure $\sigma(\lambda)$ implies that $g(\theta)$ is also nonnegative over its support. If we now introduce the trigonometrical moments

$$c_k=\int_{\theta_0}^\pi g(\theta)\cos k\theta\,d\theta,\quad k=0,1,2,\cdots, \tag{2.17}$$

and use the formulas

$$\cos k\theta=T_{2k}(\cos(\theta/2))=\overline{T}_k(\cos^2(\theta/2)), \tag{2.18}$$

where $T_{2k}(x)$ is the (even) Tchebycheff polynomial of order $2k$ so that

$$\overline{T}_k(\cos^2(\theta/2))=\sum_{p=0}^k \overline{T}_k^p(\cos^2(\theta/2)^p) \tag{2.19}$$

with

$$\overline{T}_k^0=(-1)^k\quad\text{and}\quad \overline{T}_k^p=(-1)^{k-p}\binom{k+p-1}{k-p}\frac{k}{p}\frac{4^p}{2}\quad\text{when } p\neq 0, \tag{2.20}$$

then we obtain

$$c_k=\sum_{p=0}^{k}\overline{T}_k^p\,\frac{2n_p}{4^p}=(-1)^k 2n_0+\sum_{p=1}^{k}(-1)^{k-p}\frac{k}{p}\binom{k+p-1}{k-p}n_p,\qquad k=0,1,2,\cdots. \tag{2.21}$$

We now note that

$$\frac{k}{p}\binom{k+p-1}{k-p}=2\binom{k+p}{k-p}-\binom{k+p-1}{k-p} \tag{2.22}$$

is an integer for all $p=1,2,\ldots,k$, and so all of the c_k's are themselves integers even in the physical case which corresponds to $n_0=\frac{1}{2}$ and for which $c_0=1$. Observe that the inverse of the formulas (2.20) is

$$2n_k=\binom{2k}{k}c_0+2\sum_{p=1}^{k}\binom{2k}{k-p}c_p,\quad k=0,1,2,\cdots. \tag{2.23}$$

The generating function $G(w)$ for the n_k's is

$$G(w)=\sum_{k=0}^{\infty}n_k w^k=\int_0^{\Lambda}\frac{\sigma(\lambda)\,d\lambda}{(1-\lambda w)}$$
$$=\int_0^{4\cos^2(\theta_0/2)}\frac{\frac{1}{2}g(2\mathrm{Arc}\cos(\sqrt{\lambda}/2))\,d\lambda}{(1-\lambda w)\sqrt{\lambda(4-\lambda)}} \tag{2.24}$$

while the generating function for the c_k's is

$$I(z)=\sum_{k=0}^{\infty}c_k z^k=\sum_{k=0}^{\infty}\tfrac{1}{2}z^k\int_{\theta_0}^{\pi}(\exp(ik\theta)+\exp(-ik\theta))\,g(\theta)\,d\theta$$
$$=\int_{\theta_0}^{\pi}\frac{(1-z\cos\theta)g(\theta)\,d\theta}{1-2z\cos\theta+z^2}=\frac{c_0}{2}+\frac{(1-z^2)}{2}\times\int_{\theta_0}^{\pi}\frac{g(\theta)\,d\theta}{1-2z\cos\theta+z^2}. \tag{2.25}$$

The two generating functions $G(w)$ and $I(z)$ are connected by *de la Vallée Poussin's* transformation

$$w=z/(1+z)^2 \tag{2.26}$$

according to

$$I(z)=\frac{c_0}{4}+\frac{1-z}{1+z}\,G\left(\frac{z}{(1+z)^2}\right)$$

or (2.27)

$$G(w)=\frac{1}{\sqrt{1-4w}}\left[I\left(\frac{1-\sqrt{1-4w}}{1+\sqrt{1-4w}}\right)-\frac{1}{2}I(0)\right]$$

When $\Lambda<4$, $I(z)$ is holomorphic in the z plane less an arc of the circle of radius 1, just as in the case of the function $I(z,x)$ shown in Fig. 2. The relationship between its values when $|z|<1$ and its values when $|z|>1$ is

$$I(1/z)+I(z)=c_0, \tag{2.28}$$

which can be verified with the aid of (2.27). In particular, in this situation where $I(z)$ can be analytically continued from inside the unit circle to outside it, we must have

$$I(\infty)=0\quad\text{when }\Lambda<4. \tag{2.29}$$

We will next make key use of Szegö's Theorem[11]: If among the coefficients c_k of a Taylor series $I(z)=\sum_{k=0}^{\infty}c_k z^k$ there appear only a finite number of different values, then either $I(z)=P(z)/(1-z^m)$, where $P(z)$ is a polynomial of finite degree and m is a nonnegative integer, or else $I(z)$ cannot be analytically continued beyond the unit circle.

Now from (2.17) it follows that

$$|c_k|\le c_0\quad\text{for all }k, \tag{2.30}$$

and since the c_k's are all integers, there appear only finitely many different coefficients in the expansion of $I(z)$. Furthermore when $\Lambda<4$, $I(z)$ can certainly be continued beyond the unit circle and so Szegö's theorem provides

$$I(z)=\frac{P(z)}{(1-z^m)}=P(z)+z^m P(z)+z^{2m}P(z)+\cdots, \tag{2.31}$$

where in view of (2.29) we must have

$$P(z)=c_0+c_1 z+\cdots+c_{m-1}z^{m-1}\quad\text{when }\Lambda<4. \tag{2.32}$$

Thus we have deduced that when $\Lambda<4$, $I(z)$ is a rational fraction whose only singularities are poles located at various roots of unity.

Notice that the polynomial $P(z)$ in (2.32) necessarily possesses the factor $(1-z)$ because $\Lambda<4$ corresponds to $\theta_0>0$ in (2.25) which means $I(z)$ cannot have a pole at $z=1$. More generally, there can occur many cancellations between the numerator and denominator in (2.31) as can be seen by forming a superposition of two such generating functions—the result is a new generating function corresponding to a D. M. P. for which there will in general be many cancellations when expressed in the form (2.31). This means that there is no straightforward way of expressing m as a function of θ_0.

In general, when $\Lambda\le 4$, we observe that in view of (2.30) the set of numbers

$$\tilde{c}_k=c_k+c_0,\quad k=0,1,2,\cdots \tag{2.33}$$

are all positive integers bounded above by $\tilde{c}_0$ and hence the number

$$C=\tilde{c}_0\cdot\tilde{c}_1\tilde{c}_2\tilde{c}_3\cdots \tag{2.34}$$

considered as a representation in the base $\tilde{c}_0$, is either rational or irrational, and we have the following proposition.

Proposition: The generating function $I(z)$ is a rational function if and only if the number C is rational.

Proof: Suppose C is rational. Then the sequence $\{c_k\}$ must ultimately be periodic so that there exists integers m and n with $c_k=c_{k+m}$ for all k whenever $k>n$. But this implies

$$I(z)=c_0+c_1 z+\cdots+c_n z^n+z^{n+1}\frac{(c_{n+1}+c_{n+2}z+\cdots+c_{n+m}z^{m-1})}{(1-z^m)} \tag{2.35}$$

which is rational.

Conversely, suppose $I(z)=P(z)/Q(z)$ where $P(z)$ and $Q(z)$ are polynomials of finite degree with, say,

$$Q(z)=q_0+q_1 z+\cdots+q_m z^m,\quad q_0\ne 0,\quad q_m\ne 0. \tag{2.36a}$$

Then for all k greater than some k_0 we must have

$$q_0 c_{k+m} + q_1 c_{k+m-1} + \cdots + q_m c_k = 0. \tag{2.36b}$$

Now consider the set of vectors

$$\nu_k = (c_k, c_{k+1}, \ldots, c_{k+m-1}), \quad \text{where } k = 0, 1, 2, \cdots. \tag{2.37}$$

Since only finitely many such vectors can be constructed out of the c_k's there must exist at least one such vector which reappears infinitely many times as k varies. It now follows that C is rational, for if $\nu_{k_1} = \nu_{k_2}$, then using (2.36b) we have $c_{k_1+m} = c_{k_2+m}$ and so on, . . . , which completes the proof of the proposition.

In the case where $\Lambda < 4$ we have already seen that $I(z)$ is necessarily rational, and in fact takes the special form (2.31) and (2.32). However in the case $\Lambda = 4$ the number C may be either rational or irrational. If it is rational, then the proposition provides that $I(z)$ is itself rational and can be expressed in the form (2.35), but where we no longer have any assurance that $I(z)$ tends to zero at infinity. That is, $I(z)$ is the sum of a polynomial and a component of the form (2.31) and (2.32). If, on the other hand C is irrational, then $I(z)$ cannot be a rational function and Szegö's theorem provides that it possesses a natural frontier on the unit circle $|z| = 1$.

Up to now the only use we have made of the positivity of $g(\theta)$ is embodied in (2.30). The full positivity constraint is most conveniently expressed in terms of the Toeplitz determinants[12]

$$T_k(c_0, c_1, \ldots, c_k) = \begin{vmatrix} c_0 & c_1 & c_2 & \cdots & c_k \\ c_1 & c_0 & c_1 & \cdots & c_{k-1} \\ c_2 & c_1 & c_0 & \cdots & c_{k-2} \\ \vdots & \vdots & & & \\ c_k & c_{k-1} & & \cdots & c_0 \end{vmatrix}. \tag{2.38}$$

A necessary and sufficient condition that $g(\theta)$ is nonnegative for $0 \leq \theta \leq \pi$ is that

$$T_k(c_0, c_1, \ldots, c_k) \geq 0 \quad \text{for all } k = 0, 1, 2, \cdots, \tag{2.39}$$

while $g(\theta)$ is nonnegative for $\theta_0 \leq \theta \leq \pi$ and vanishes on $0 \leq \theta \leq \theta_0$ if and only if we have both (2.39) and

$$T_k(d_0, d_1, \ldots, d_k) \geq 0 \quad \text{for all } k = 0, 1, 2, \cdots, \tag{2.40}$$

where for each k

$$d_k = 2c_k \cos\theta_0 - c_{|k+1|} - c_{|k-1|}. \tag{2.41}$$

Furthermore, if $T_{k_0}(c_0, c_1, \ldots, c_{k_0}) = 0$, then $T_k(c_0, c_1, \ldots, c_k) = 0$ for all $k \geq k_0$ and this is the case if and only if $I(z)$ takes the form (2.31) and (2.32) for some $m \leq k_0$.

We believe, but have so far been unable to prove in general, that the positivity constraints (1.39) are in fact such as to ensure that the number C is rational and hence that there cannot exist a natural frontier on the unit circle in the case $\Lambda = 4$. This belief is supported by the fact that it is certainly true in the physical case mentioned in Sec. 1 where $n_0 = \frac{1}{2}$ so that $c_0 = 1$. In this situation, either

$$0 = c_1 = c_2 = c_3 = \cdots \tag{2.42}$$

in which case $I(z) = c_0 = 1$, or else there is some first c_{k_0} such that $|c_{k_0}| = 1$. If the latter is true, then $T_{k_0} = 0$ from which it follows that $I(z)$ is a rational function of the special form (2.31) and (2.32).

We now return to the w plane and convert our results concerning $I(z)$ into statements about the generating function $G(w)$. When $\Lambda < 4$ we see from (2.31) together with (2.27) that $G(w)$ is a rational fraction with its poles located at various points $w = w_k$, $k \in \{1, 2, \ldots, m-1\}$, where

$$w_k = \frac{1}{4\cos^2(k\pi/m)}. \tag{2.43}$$

When $\Lambda = 4$ and $n_0 = \frac{1}{2}$, we either have

$$G(w) = \frac{1}{\sqrt{1-4w}} \left\{ \left[1 \pm \left(\frac{1-\sqrt{1-4w}}{1+\sqrt{1-4w}} \right)^m \right]^{-1} - \frac{1}{2} \right\} \quad \text{for some } m, \tag{2.44}$$

which must always reduce to a rational fraction with its poles at all of the points w_k. For example, when $m = 7$

$$G(w) = \frac{1 - 7w + 14w^2 - 7w^3}{2(1-4w)(1-5w+6w^2-w^3)}, \tag{2.45}$$

or else corresponding to (2.42) we have

$$G_4(w) = \frac{1}{2} \cdot \frac{1}{\sqrt{1-4w}}. \tag{2.46}$$

It is interesting to note that $G_4(w)$ here is precisely the function obtained for the one-dimensional Ising model in the thermodynamic limit. Rational fractions are obtained in the case of a finite Ising chain with periodic boundary conditions.

3. THE PROBLEM WHEN $\Lambda > 4$

This situation is the interesting one for physical problems with the exception of the one-dimensional Ising model which corresponds to $\Lambda = 4$.

Proposition: Any Diophantine moment problem corresponding to $\Lambda > 4$ can be reduced to one with $\Lambda = 6$ by repeated application of the transformation

$$G_{2\sqrt{\Lambda}}(w) = \frac{1}{1 - w\sqrt{\Lambda}} \, G_\Lambda\left(\frac{w^2}{(1 - w\sqrt{\Lambda})^2} \right), \tag{3.1}$$

where $G_\Lambda(w)$ denotes a Diophantine moment generating function associated with the support $[0, \Lambda]$ and $\sqrt{\Lambda}$ is integer.

To prove this result we need to use various transformations which take D.M.P.'s into D.M.P.'s. These are of considerable importance and we begin by discussing the simplest ones.

Starting with

$$n_k = \int_a^b \lambda^k \sigma(\lambda)\, d\lambda, \quad k = 0, 1, 2, \cdots, \tag{3.2}$$

where a and b are real numbers, σ is nonnegative, and the n_k's are integers, let us make a translation

$$\lambda = x - q, \quad \text{where } q \text{ is an integer greater than } -a. \tag{3.3}$$

Then we obtain

$$n_k = \int_{a+q}^{b+q} (x-q)^k \sigma(x-q)\, dx, \tag{3.4}$$

and putting $\bar{\sigma}(x) = \sigma(x-q)$, which is also nonnegative, we define a new set of moments by

$$\bar{n}_k = \int_{a+q}^{b+q} x^k \bar{\sigma}(x)\, dx, \quad k = 0, 1, 2, \cdots. \tag{3.5}$$

These new moments associated with the translated measure $\bar{\sigma}(x)$ are also integers because

$$\bar{n}_k = \int_a^b (\lambda + q)^k \sigma(\lambda)\, d\lambda = \sum_{p=0}^{k} \binom{k}{p} q^{k-p} n_p, \tag{3.6}$$

and the two generating functions

$$G(w) = \sum_{k=0}^{\infty} n_k w^k \text{ and } G_{[q]}(w) = \sum_{k=0}^{\infty} \bar{n}_k w^k \tag{3.7}$$

are connected by the relation

$$G_{[q]}(w) = \frac{1}{1-qw}\, G\left(\frac{w}{(1-qw)}\right). \tag{3.8}$$

The second simple transformation is nonlinear. Setting

$$\lambda = y^2 \tag{3.9}$$

in (3.2) and supposing $0 < a < b$, we get

$$n_k = \int_{\sqrt{a}}^{\sqrt{b}} 2y^{2k+1} \sigma(y^2)\, dy \tag{3.10}$$

and

$$G(w) = \sum_{k=0}^{\infty} n_k w^k = \int_a^b \frac{\sigma(\lambda)\, d\lambda}{1-\lambda w}$$

$$= \int_{\sqrt{a}}^{\sqrt{b}} \sigma(y^2)\, y \left(\frac{1}{1-y\sqrt{w}} + \frac{1}{1+y\sqrt{w}}\right) dy$$

$$= \int_{-\sqrt{b}}^{+\sqrt{b}} \frac{|y|\sigma(y^2)}{1-y\sqrt{w}}\, dy,$$

$$\sigma(y^2) \text{ being zero on } [-\sqrt{a}, +\sqrt{a}]. \tag{3.11}$$

If we now introduce the moments

$$m_k = \int_{-\sqrt{b}}^{+\sqrt{b}} |y|\, \sigma(y^2)\, y^k\, dy, \quad k = 0, 1, 2, \cdots, \tag{3.12}$$

then we see that the generating function

$$G_{\sqcap}(w) = \sum_{k=0}^{\infty} m_k w^k \tag{3.13}$$

is even and

$$m_{2k} = n_k, \quad m_{2k+1} = 0, \quad k = 0, 1, 2, \cdots. \tag{3.14}$$

The two generating functions are related by

$$G_{\sqcap}(w) = G(w^2). \tag{3.15}$$

Next we combine the two transformations (3.8) and (3.15). Starting with

$$n_k = \int_0^{\Lambda} \lambda^k \sigma(\lambda)\, d\lambda \text{ and } G(w) = \sum_{k=0}^{\infty} n_k w^k, \tag{3.16}$$

and applying the quadratic transformation followed by a translation, we arrive at

$$G_{\sqcap[q]}(w) = \frac{1}{1-qw}\, G_{\sqcap}\left(\frac{w}{(1-qw)}\right) = \frac{1}{1-qw}\, G\left(\frac{w^2}{(1-qw)^2}\right), \tag{3.17}$$

where the support of $G_{\sqcap[q]}(w)$ is $[-\sqrt{\Lambda}+q, \sqrt{\Lambda}+q]$ and its moments are

$$\bar{m}_k = \sum_{p=0}^{[k/2]} \binom{k}{2p} q^{k-2p} n_p, \quad k = 0, 1, 2, \cdots, \tag{3.18}$$

$[k/2]$ denoting the integer part of $k/2$. If we insist that $\sqrt{\Lambda}$ is an integer, which we can always do by enlarging the support in (3.16), then we arrive at the D.M.P. preserving transformation (3.1), and we see that the original problem with support $[0, \Lambda]$ is transformed into a problem whose support is $[0, 2\sqrt{\Lambda}]$.

The proposition is now proved by repeated application of this transformation as follows. Begin with any value of $\Lambda > 4$. Increase Λ until it becomes a perfect integer square. Apply the transformation. Repeat the process. It is readily seen that in this way we can always arrive at a problem whose support is $[0, 6]$. For example, for $\Lambda = 54$, which corresponds to the critical isotherm of the planar triangular Ising model, we have the sequence

$$\Lambda = 54 \xrightarrow{\text{enlarged to}} 8^2 = 64 \xrightarrow{\text{trans}^n} 16$$
$$= 4^2 \xrightarrow{\text{trans}^n} 8 \xrightarrow{\text{enlarged to}} 3^2 = 9 \xrightarrow{\text{trans}^n} 6.$$

We have not been able to classify rigorously the classes of solutions that are admitted when $\Lambda = 6$, although we have a fair idea of the types of things to be expected as will be seen from the examples given in the next section. The situation to date is summarized in Table I.

4. EXAMPLES

A. A class of algebraic generating functions

We consider the solutions $G_i(w)$, $i = 1, 2, \ldots, N$, of the algebraic equation

$$+\ p - G(w) + wG(w)^N = 0, \tag{4.1}$$

where p and N are positive integers.

Proposition: The solution of (4.1) which is regular at $w = 0$ is the generating function for a D.M.P. with

$$\Lambda = \frac{N-1}{p}\left[\frac{pN}{N-1}\right]^N \text{ when } N > 1, \text{ and } \Lambda = 1 \text{ when } N = 1.$$

Proof: Note that the only possible singular points for any $G_i(w)$ are $w = \infty$, $w = 0$, and

$$w = w_{\text{sing}} = \left(\frac{N-1}{pN}\right)^N \frac{p}{(N-1)} \quad (= 1 \text{ when } N = 1).$$

The solution with no singularity at zero, $G_1(z)$, has the expansion

$$G_1(w) = p + p^N w + \cdots, \tag{4.2}$$

wherein all the coefficients are integer. Observing directly from (4.1) that the imaginary part of any $G_i(w)$ cannot vanish when $\mathrm{Im}w \neq 0$, we deduce from (4.2) that

$$\mathrm{Im}G_1(w) < 0 \quad \text{when } \mathrm{Im}w > 0. \tag{4.3}$$

Since the only singularities of $G_1(w)$ are at w_{sing} and infinity, we have by Cauchy's theorem

TABLE I. Summary of what is known about $G(w)$ as a function of the length of its support.

Support	Nature of $G(w)$	Remarks
$\Lambda=1$	[1/1] rational fraction	
$\Lambda=2$	[2/2] rational fraction	Only rational fractions are allowed.
$\Lambda<4$	rational fraction whose poles have various locations $w=w_k$ $k\in\{1,2,\ldots,m-1\}$, some m, where $w_k=1/4\cos^2(k\pi/m)$.	
$\Lambda=4$	If $n_0=\frac{1}{2}$ while other moments are integer, then either $G(w)$ is a rational fraction with poles at all of the points $w=w_k$, $k=1,2,\ldots,m-1$, where $w_k=1/4\cos^2(k\pi/m)$, some m; or $G(w)=\frac{1}{2}/\sqrt{1-4w}$.	Algebraic functions with branch points of order two are admitted.
	In general, $G(w)$ may be a superposition of a rational fraction, an analytic function with branch points of order two at $w=\frac{1}{4}$ and $w=\infty$, and a Stieltjes function with a natural frontier on $\frac{1}{4}\le w<\infty$.	Is the possibility of a natural frontier excluded by the positivity constraints?
$4<\Lambda<6$	Very little known	No-man's land
$\Lambda=6$	Examples show that $G(w)$ can have a natural frontier on its second sheet, and that certain hypergeometric functions possessing logarithmic singularities occur. $G(w)$ can also be an algebraic function with high order branch points.	Algebraic functions with high order branch points admitted. Functions with logarithmic singularities admitted.
$\Lambda>6$	$G_\Lambda(w)$ is related to $G_6(w)$ by a sequence of purely algebraic transformations.	

$$G_1(w)=\frac{1}{2\pi i}\oint\frac{G_1(\xi)\,d\xi}{(\xi-w)} \tag{4.4}$$

for any w in the complex plane cut from w_c to ∞, and C is any contour in this cut plane which encloses w. Choosing C as in Fig. 5, and letting the circular part tend to infinity it is readily found that the only contribution to (4.4) comes from the integration back and forth along the cut. Since $G_1(w)$ is real on the real axis for $w<w_{\text{sing}}$, the discontinuity along the cut is

$$2\lim_{\epsilon\to0^+}G(x+i\epsilon)=-2\pi i\theta(x),\quad x\ge w_{\text{sing}}, \tag{4.5}$$

where $\theta(x)$ is positive because of (4.3). Thus we have

$$G_1(w)=\int_{w_{\text{sing}}}^{\infty}\frac{\theta(x)}{x-w}\,dx. \tag{4.6}$$

Defining $\sigma(\lambda)=\theta(1/\lambda)/\lambda$ on $0\le\lambda\le1/w_{\text{sing}}$ we finally

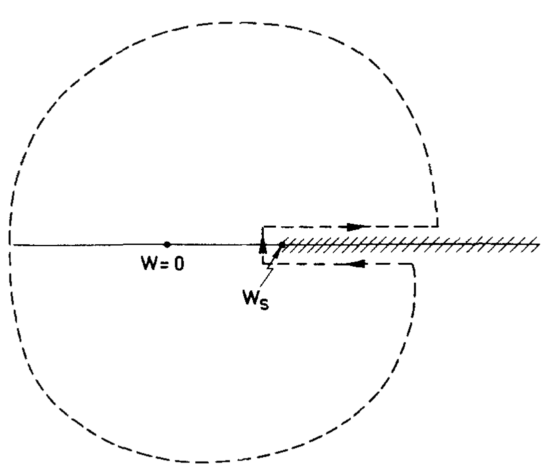


FIG. 5. The contour C used in (4.4) to reduce $G_1(w)$ to the form (4.6). The discontinuity on the cut is positive.

have

$$G_1(w)=\int_0^{1/w_{\text{sing}}}\frac{\sigma(\lambda)\,d\lambda}{1-\lambda w},\quad\text{for all } w \text{ in the cut plane,} \tag{4.7}$$

which completes the proof of the proposition.

Notice that for $p=1$ and N large, $\Lambda\sim e^N$. In this example we have a sequence of algebraic functions which are solutions of D. M. P.'s and whose order increases with increasing Λ.

B. A generating function with logarithmic singularities

Suppose we have two generating functions

$$G_\Lambda(w)=\sum_{k=0}^{\infty}n_kw^k\quad\text{and}\quad G_{\Lambda'}(w)=\sum_{k=0}^{\infty}n'_kw^k \tag{4.8}$$

belonging to supports $[0,\Lambda]$ and $[0,\Lambda']$, respectively. Then the Hadamard product is defined by

$$G_{\Lambda\Lambda'}(w)=G_\Lambda*G_{\Lambda'}(w)=\sum_{k=0}^{\infty}n_kn'_kw^k. \tag{4.9}$$

$G_{\Lambda\Lambda'}(w)$ is a D. M. P. generating function with support $[0,\Lambda\Lambda']$ because

$$n_kn'_k=\int_0^{\Lambda}\int_0^{\Lambda'}(\lambda\lambda')^k\sigma(\lambda)\sigma'(\lambda')\,d\lambda\,d\lambda'$$
$$=\int_0^{\Lambda\Lambda'}t^k\,dt\int\sigma'(t/\lambda)\sigma(\lambda)\frac{d\lambda}{\lambda}. \tag{4.10}$$

Choosing

$$G_4(w)=\sum_{k=0}^{\infty}\binom{2k}{k}w^k=1/\sqrt{1-4w}, \tag{4.11}$$

we find

$$G_{16}(w)=(G_4*G_4)(w)=\sum_{k=0}^{\infty}\binom{2n}{n}^2w^n=F(\tfrac{1}{2},\tfrac{1}{2},1;16w) \tag{4.12}$$

which is the Gauss hypergeometric function. This has a logarithmic singularity at $w=\frac{1}{16}$.

Other hypergeometrics involving rational indices and the variable scaled by a suitable integer also give rise to D. M. P. generating functions.[13]

C. A generating function with a natural boundary on the second sheet[14]

Consider the case $\Lambda=8$ so that

$$n_k=\int_0^8 \lambda^k \sigma(\lambda)\,d\lambda, \quad k=0,1,2,\cdots. \tag{4.13}$$

Then defining $\tilde{\sigma}(\lambda)=2\sigma(2\lambda)$ on $0\le\lambda\le 4$ we find

$$\tilde{n}_k=\int_0^4 \lambda^k d\tilde{\sigma}(\lambda)=n_k/2^k, \quad k=0,1,2,\cdots. \tag{4.14}$$

If we now convert this into a trigonometrical moment problem by applying the transformation (2.26) and (2.27), we find that the trigonometrical moments are of the form

$$\tilde{c}_k=c_k/2^k, \quad k=0,1,2,\cdots, \tag{4.15}$$

where c_k's are integer. Now suppose that the $\tilde{c}_k$'s are such that the associated generating function

$$\tilde{I}(z)=\sum_{k=0}^{\infty}\frac{c_k}{2^k}z^k \text{ is holomorphic for } |z|<2 \tag{4.16}$$

and moreover

$$\mathrm{Re}\tilde{I}(z)\ge c_0/2 \quad \text{when} \quad |z|<1, \tag{4.17}$$

so that the associated trigonometrical measure $\tilde{g}(\theta)$ is nonnegative. Then on inverting the transformation (2.26) and (2.27) we are led to a solution of the D. M. P. (4.13).

Now choose

$$c_0=2, \quad c_k=1 \text{ if } k \text{ is prime, and } c_k=0, \text{ otherwise.} \tag{4.18}$$

Then (4.16) is satisfied, and moreover

$$\mathrm{Re}\tilde{I}(\exp(i\theta))=2+\sum_{k\ \text{prime}}(\cos k\theta)/2^k$$
$$\ge 2-\sum_{k=1}^{\infty}1/2^k\ge 1=c_0/2 \tag{4.19}$$

which is condition (4.17). Thus (4.18) corresponds to a solution of the D. M. P. (4.13), where $\Lambda=8$. Moreover, $\tilde{I}(z)$ clearly has a natural frontier on $|z|=2$ so the solution of the D. M. P. has a natural frontier on its second sheet.

D. A generating function associated with a Cantor set

We consider the effect of repeated application of the D. M. P. preserving transformation (3.1), starting with $\Lambda=6$ and any $G_6^{(0)}(w)$ whose support is the whole interval $[0,6]$. On increasing Λ to 9 and applying the transformation we obtain $G_6^{(1)}(w)$ whose support is contained in $[3-\sqrt{6},\ 3+\sqrt{6}]$. Again increasing $\Lambda=3+\sqrt{6}$ to 9, and again applying the transformation, we now obtain a D. M. P. generating function $G_6^{(2)}(w)$ whose support is contained in $[3-(3+\sqrt{6})^{1/2},\ 3-(3-\sqrt{6})^{1/2}]$ $\cup[3+(3-\sqrt{6})^{1/2},\ 3+(3+\sqrt{6})^{1/2}]$. At the nth iteration we obtain a D. M. P. generating function $G_6^{(n)}(w)$ whose support is contained in the union of the set of intervals

$$I_j^{(n)}=[a_{2j-1}^{(n)},a_{2j}^{(n)}], \quad j=1,2,\ldots,2^{n-1}. \tag{4.20}$$

Here the set $\{a_k^{(n)}\}_{k=1}^{2n}$ consists of all the numbers which are expressible in the form

$$3\pm(3\pm(3\pm\cdots(3\pm\sqrt{6})^{1/2}\cdots)^{1/2})^{1/2}, \tag{4.21}$$

where the number of square roots involved is n. These numbers are ordered so that

$$a_1^{(n)}<a_2^{(n)}<\cdots<a_{2^n}^{(n)}, \tag{4.22}$$

and the only difference between $a_{2j-1}^{(n)}$ and $a_{2j}^{(n)}$ is the sign of the $\sqrt{6}$ term in their representations in the form (4.21). The following points, which we state here without proof, can now be established. Firstly

$$\bigcup_{j=1}^{2^{n-1}} I_j^{(n)}\supset\bigcup_{j=1}^{2^n} I_j^{(n+1)} \text{ for } n=1,2,3,\cdots, \tag{4.23}$$

so that each set of intervals forms a covering to all of its successors. Secondly,

$$\lim_{n\to\infty}\text{meas}\bigcup_{j=1}^{2^{n-1}} I_j^{(n)}=\lim_{n\to\infty}\sum_{j=1}^{2n}|a_{2j}-a_{2j-1}|=0. \tag{4.24}$$

Thus the limiting support, as n tends to infinity, is the set S of all points expressible in the form

$$3\pm\sqrt{(3\pm\sqrt{(3\pm\cdots ad\ infinum)}}: \tag{4.25}$$

S is uncountable and of measure zero, and is therefore a Cantor set.

The above consideration leads us to seek the answers to the following questions. Does the sequence $\{G_6^{(n)}(w)\}_{n=1}^{\infty}$ converge to a well defined limiting function $G_6^{(\infty)}(w)$? Is $G_6^{(\infty)}(w)$ a genuine D. M. P. generating function? What is the nature of the measure of $G_6^{(\infty)}(w)$, this being associated with the Cantor set S?

Proposition: Let $G_6^{(0)}(w)$ be any given D. M. P. generating function with support on $[0,6]$. Then the sequence $\{G_6^{(n)}(w)\}_1^{\infty}$ converges to a function $G_6^{(\infty)}(w)$, the convergence being uniform for $w\in\Omega$, where Ω is any closed bounded region interior to the complex plane cut from $w=\frac{1}{6}$ to $w=\infty$.

Furthermore, $G_6^{(\infty)}(w)$ is a fixed point of the transformation (3.1) corresponding to $\Lambda=9$. It is a bona fide D. M. P. generating function with support $[0,6]$ and is unique up to multiplication by a positive integer n_0, its Taylor series expansion about the origin being

$$G_6^{(\infty)}(w)=n_0\sum_{k=0}^{\infty}g_k w^k, \tag{4.26}$$

where the g_k's are integers given recursively by the relations

$$g_0=1,$$
$$\sum_{p=0}^{2s+1}g_p(-3)^{-p}\binom{2s+1}{p}=0, \tag{4.27}$$

and

$$\sum_{p=0}^{2s+2}g_p(-3)^{-p}\binom{2s+2}{p}=(-3)^{-2(s+1)}g_{s+1}, \quad s=0,1,2,\cdots.$$

The series (4.26) converges for $|w|<\frac{1}{6}$.

Proof: First we note that if (3.1) has a fixed point

$G_6^{(\infty)}(w)$ which is regular in some neighborhood of $w=0$, *then* its Taylor series expansion about $w=0$ would have to be (4.26) where the g_n's are given by (4.27). This can be seen by substituting (4.26) into both sides of (3.1) wherein $\Lambda=9$, and equating coefficients. Furthermore, it is easy to see that the g_k's defined by (4.27) are all integers.

Now let us start with any D.M.P. generating function $G_6^{(0)}(w)$ with $G_6^{(0)}(0)=n_0$ and $\Lambda=6$, and let us apply (3.1) n times to obtain a new D.M.P. generating function $G_6^{(n)}(w)$ belonging to $\Lambda=6$. Set

$$G_6^{(n)}(w)=\sum_{k=0}^{\infty} g_k^{(n)} w^n. \tag{4.28}$$

Then a little algebra provides that

$$g_k^{(n)}=n_0 g_k \quad \text{for} \quad k=0,1,\ldots,(2n-1), \tag{4.29}$$

where the g_k's are those defined by (4.27). Moreover, we know from the previous proposition that $G_6^{(n)}(w)$ is a D.M.P. generating function with support $[0,6]$, and hence

$$g_k^{(n)}=\int_0^6 \lambda^n \sigma^{(n)}(\lambda)\,d\lambda \quad \text{for} \quad k=0,1,2,\cdots, \tag{4.30}$$

where $\sigma^{(n)}(\lambda)$ is a nonnegative measure on $[0,6]$. In addition, from (4.29) it follows that

$$g_0^{(n)}=n_0 \quad \text{and}\ g_1^{(n)}=3n_0 \quad \text{for}\ n\geq 1. \tag{4.31}$$

Applying the theory of upper and lower principal representations[16] to (4.30) and (4.31) we obtain

$$3^k n_0 \leq g_k^{(n)} \leq \tfrac{1}{2}\cdot 6^k n_0 \quad \text{for all}\ k>0\ \text{and}\ n\geq 1. \tag{4.32}$$

Combining this with (4.29) we now have

$$3^k \leq g_k \leq \tfrac{1}{2}\cdot 6^k, \quad \text{for all}\ k>0, \tag{4.33}$$

which says in particular that the series (4.26) is absolutely convergent for $|w|<\frac{1}{6}$. Let us denote the function thus defined by $G_6^{(\infty)}(w)$.

Finally we prove that

$$G_6^{(\infty)}(w)=\int_0^6 \frac{\sigma(\lambda)\,d\lambda}{1-\lambda w} \quad \text{for some nonnegative}\ \sigma(\lambda), \tag{4.34}$$

and that $\{G_6^{(n)}(w)\}_{n=0}^{\infty}$ converges uniformly to $G_6^{(\infty)}(w)$ for all $w\in\Omega$, where Ω is any closed bounded region interior to the complex plane cut from $w=\frac{1}{6}$ to ∞.

To prove (4.34), we recall that

$$F(w)=\int_0^{\infty} \frac{\theta(\lambda)\,d\lambda}{(1-w\lambda)} \quad \text{for some nonnegative}\ \theta(\lambda) \tag{4.35}$$

if and only if[17] $D_{0,k}(F_0,F_1,\ldots,F_{2k})\geq 0$ and $D_{1,k}(F_1,F_2,\ldots,F_{2k+1})\geq 0$ for all k where

$$D_{l,k}(F_l,F_{l+1},\ldots,F_{2k+1}) = \mathrm{Det}\begin{vmatrix} F_l & F_{l+1} & \cdots & F_{l+k} \\ F_{l+1} & F_{l+2} & \cdots & F_{l+k+1} \\ \vdots & & & \\ F_{l+k} & F_{l+k+1} & \cdots & F_{l+k+2} \end{vmatrix} \quad \text{for}\ l=0\ \text{and}\ 1, \tag{4.36}$$

and

$$k!\,F_k=\lim_{\substack{w\to 0\\ w<0}} \frac{d^kF(w)}{dw^k} \quad \left(=\int_0^{\infty} \lambda^k\theta(\lambda)\,d\lambda\right). \tag{4.37}$$

But $D_{l,k}(n_0g_l, n_0g_{l+1},\ldots,n_0g_{l+k})=D_{l,k}(g_l^{(2k+1)}, g_{l+1}^{(2k+1)},\ldots,g_{l+k}^{(2k+1)})\geq 0$ for $l=0$ and 1 because $G_6^{(2k+1)}(w)$ is certainly of the form (4.35), and we have (4.29). It now follows that $G_6^{(\infty)}(w)$ is expressible in the form (4.35). In addition (4.33) implies that the measure is zero for $w>6$, and thus we have (4.34). In particular, the Stieltjes moment problem associated with the set of moments $\{n_0g_k\}_{k=0}^{N}$ is determinate: Thus[18] the set of moments $\{n_0g_k\}_{k=0}^{\infty}$ defines a lense-shaped region $\omega_N(n_0g_0,\ldots,n_0g_N;w)$ in the complex plane, for each $w\in\Omega$, such that for any function of the form (4.35) whose first $(N+1)$ moments are precisely $n_0g_0, n_0g_1,\ldots,n_0g_N$, we have

$$F(w)\in\omega_N(n_0g_0,\ldots,n_0g_N;w) \quad \text{for all}\ w\in\Omega, \tag{4.38}$$

where

$$\mathrm{Size}\,\omega_N(n_0g_0,\ldots,n_0g_N;w)\to 0 \quad \text{as} \quad N\to\infty, \quad \text{uniformly for}\ w\in\Omega. \tag{4.39}$$

The size of a set S in the complex plane is simply $\mathrm{Max}\{|w_1-w_2|: w_1\in S,\ w_2\in S\}$. Let $\epsilon>0$ be given. Choose N so large that $\mathrm{Size}\,\omega_N(n_0g_0,\ldots,n_0g_N;w)<\epsilon$ for all $w\in\Omega$. Then we clearly have

$$|G_6^{(\infty)}(w)-G_6^{(n)}(w)|<\epsilon \quad \text{for all}\ w\in\Omega \tag{4.40}$$

for all $n>N$, which completes the proof.

We are now in a position to specify the measure $\sigma_6^{(\infty)}(\lambda)$ of $G_6^{(\infty)}(w)$, and to show how this is defined over the Cantor set S. Let the measure associated with $G_6^{(n)}(w)$ be $\sigma_6^{(n)}(\lambda)$, and define

$$\theta_6^{(n)}(\lambda)=\int_0^{\lambda} \sigma_6^{(n)}(x)\,dx \quad \text{for all}\ n=0,1,2,\cdots, \tag{4.41}$$

so that

$$G_6^{(n)}(w)=\int_0^6 \frac{d\theta_6^{(n)}(\lambda)}{(1-w\lambda)}. \tag{4.42}$$

Then using the uniform convergence of the sequence $\{G_6^{(n)}(w)\}$, we have from the theory of moments that

$$\lim_{n\to\infty}\theta_6^{(n)}(\lambda)=\theta_6^{(\infty)}(\lambda), \quad \text{uniformly for}\ \lambda\in[0,6]. \tag{4.43}$$

Moreover, from the proposition, the result must be independent of the choice of starting function $G_6^{(0)}(w)$, subject to the normalization condition

$$G_6^{(0)}(0)=\theta_6^{(0)}(6)=1. \tag{4.44}$$

Now, the relation between the successive $\sigma_6^{(n)}(\lambda)$'s is

$$\sigma_6^{(n+1)}(\lambda)=|\lambda-3|\,\sigma_6^{(n)}((\lambda-3)^2), \quad n=0,1,2,\cdots. \tag{4.45}$$

Choosing for simplicity

$$\sigma_6^{(0)}(\lambda)=\delta(\lambda-6) \tag{4.46}$$

we find

$$\sigma_6^{(1)}(\lambda)=\tfrac{1}{2}[\delta(\lambda-(3-\sqrt{6}))+\delta(\lambda-(3+\sqrt{6}))]. \tag{4.47}$$

Iterating, we thus have

$$\sigma_6^{(n)}(\lambda)=\frac{1}{2^n}\sum_{k=1}^{2n}\delta(\lambda-a_k^{(n)}), \tag{4.48}$$

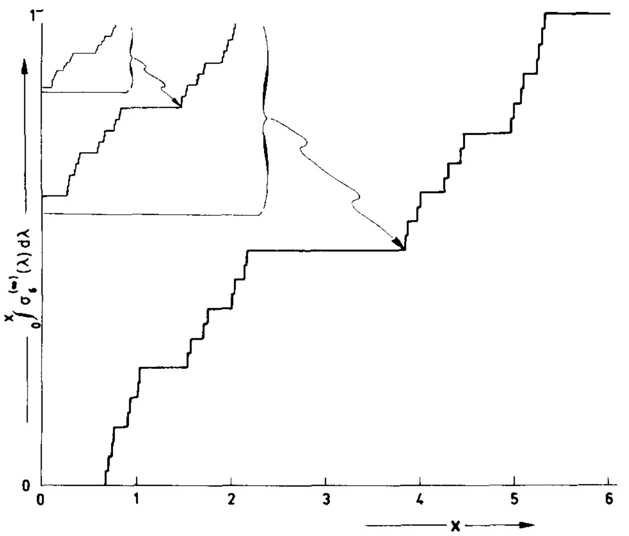


FIG. 6. The devil's staircase. This represents the integrated measure $\int_0^x \sigma_6^{(\infty)}(\lambda)\,d\lambda$ associated with the D.M.P. generating function $G_6^{(\infty)}(w)$. Each "vertical" segment has the detailed structure shown in the inset.

where the $a_k^{(n)}$'s are defined in (4.21). Hence

$$\theta_6^{(n)}(\lambda) = \frac{1}{2^n} \times \{\text{The number of } a_k^{(n)}\text{'s less than or equal to } \lambda\}. \tag{4.49}$$

Taking the limit as $n \to \infty$, we obtain $\theta_6^{(\infty)}(\lambda)$ which is an example of the infamous devil's staircase, illustrated in Fig. 6. It is a continuous bounded monotone nondecreasing function whose derivative exists and is zero at every point except on the Cantor set S.

More generally we can consider the fixed point, $G_{2q}^{(\infty)}(w)$, of the transformation (3.1) corresponding to an initial D.M.P. generating function $G_{2q}^{(0)}(w)$ with support on $[0, 2q]$, where q is an integer ≥ 2. Starting with $\Lambda = 2q$, one increases Λ to q^2 and applies (3.1) to obtain a new D.M.P. generating function $G_{2q}^{(1)}(w)$ with $\Lambda = 2q, \cdots$, and so on. The analysis proceeds just as before except that the role of 3 is played by q throughout. $G_{2q}^{(\infty)}(w)$ is a D.M.P. generating function which has for support the set of all numbers of the form

$$q \pm \sqrt{(q \pm \sqrt{(q \pm \sqrt{(q \pm \cdots ad\ infinum)} \cdots)))}. \tag{4.50}$$

This ensemble appears to be a Cantor set of measure zero for all $q > 2$. When $q = 2$ the ensemble is dense in $[0, 4]$ and we find

$$G_4^{(\infty)}(w) = \frac{n_0}{\sqrt{1-4w}}, \tag{4.51}$$

where n_0 is a positive integer, and the corresponding measure is

$$\sigma_4^{(\infty)}(\lambda) = \frac{(n_0/2\pi)}{\sqrt{\lambda(4-\lambda)}}. \tag{4.52}$$

What happens in this limiting case is that all of the gaps in the support of the measure become filled up and the resulting measure is very smooth.

We note that the support of $G_{2q}^{(\infty)}(w)$ is in fact contained in $[q - \sqrt{\Lambda_0}, \Lambda_0]$, where

$$\Lambda_0 = [2q + \sqrt{(4q+1)}]/2. \tag{4.53}$$

In particular, on supposing that near Λ_0^{-1}

$$G_{2q}^{(\infty)}(\Lambda_0^{-1} - \epsilon) \sim \epsilon^\alpha + \text{higher order in } \epsilon, \tag{4.54}$$

and substituting into the fixed point equation

$$G_{2q}^{(\infty)}(w) = \frac{1}{1-qw}\, G_{2q}\left(\frac{w^2}{(1-qw)^2}\right) \tag{4.55}$$

we obtain after some calculation the consistent conclusion that

$$\alpha = \ln \Lambda_0^{1/2} / [\ln \Lambda_0 - \ln 2(\Lambda_0 + q\Lambda_0^{1/2})]. \tag{4.56}$$

Choosing n, not necessarily an integer, so that $\Lambda_0 = 4^n$, and the integer $q = 4^n - 2^n$, we find the index

$$\alpha = -n/(n+1), \tag{4.57}$$

which agrees with (4.51) where $n = 1$.

These functions have many interesting properties—analyticity in a cut plane, positive discontinuity, expansion with integer coefficients, and behavior like $(w_s - w)^\alpha$ near the first singular point $w_s = \Lambda_0^{-1}$. The occurrence of such functions in our problem was unexpected and a physical interpretation of them can only be for the moment hypothetical. However, functions of this type occur elsewhere in a physical context—in particular, reference are made to such functions in recent studies of crystallographic structures.[19,20]

5. CONCLUSION AND OUTLOOK

An interesting result of our analysis is the quantification of the D.M.P. when the length of the support is less than four. The fact that all supports greater than four can be reduced to a length six is a mathematical equivalence which may correspond to some physical property such as universality. We have not been able to classify the families of solutions which appear when $4 < \Lambda \leq 6$, and suspect that such a classification must depend not only on Λ but on other parameters as well, in distinction from the case $\Lambda \leq 4$.

The nature of families of solutions admitted when $\Lambda > 4$ might be classifiable in terms of the smoothness of the measure $\sigma(\lambda)$. Various solutions of the D.M.P. with continuous $\sigma(\lambda)$ are provided by the hypergeometric functions $F(a, b, c; Tz)$ with a and b rational but not both integers, c integer, and T an integer which depends upon a, b, and c. Similarly, solutions are also provided by functions which are algebraic over hypergeometrics, $F(a_1, a_2, \ldots, a_\nu; zT)$.[13]

An insight into the arithmetical nature of the critical indices and an arithmetical interpretation of their universality will only be revealed when the D.M.P. has been completely classified. However, we anticipate that the class of transformations introduced in Sec. 3 may have surprising consequences for the physical Ising model considered at the outset. These transformations admit fixed points, the simplest of which corresponds exactly to the known solution for the one-dimensional Ising chain. For the moment we can only guess that this transformation may be interpreted as a scale transformation in both the magnetic field and temperature variables. In general we ask if the strange devil's staircase type functions play a role related to the behavior of the thermodynamical functions in a purely imaginary field. The latter has received attention only very recently,[10] being treated from the renormalization point of view. For us, this provides a motivation, over and above mathematical interest, to extend the investigation to the case of polynomial moments and to see if similar transformations prevail.

More generally, the study of the moment problem when the moments belong to special classes, such as integers, rings, discrete sets, etc., may provide a complete new insight into statistical mechanics models as well as quantum systems such as field theory, bearing in mind that most of the physical problems of this sort can be formulated as moment problems on abstract fields.

ACKNOWLEDGMENTS

We are pleased to thank Professor P. Deligne and Professors G. and D. Choodnovsky for their interest in this work and for providing us with examples. We also acknowledge the help of Dr. B. Derrida who proposed the example which is explored in Sec. 4 D. Finally we would like to thank Professor M. Froissart and Professor A. Martin for their special interest in this problem.

[1]L. Onsager, Phys. Rev. **65**, 117 (1944); B. M. McCoy and T. T. Wu, *The Two-Dimension Ising Model* (Harvard U. P., Cambridge, Mass., 1973).

[2]T. D. Lee and C. N. Yang, Phys. Rev. **87**, 410 (1952); D. Ruelle, Commun. Math. Phys. **31**, 265 (1973).

[3]F. Hausdorff, *Set Theory* (Chelsea, New York, 1962).

[4]C. Domb, *Phase Transition and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London, 1974), Vol. 3, p. 357.

[5]J. D. Bessis, J. M. Drouffe, and P. Moussa, J. Phys. A **9**, 2105 (1976).

[6]Analogous Stieltjes representation for the intensity of magnetization was introduced in G. A. Baker, Jr., Phys. Rev. Lett. **20**, 990 (1968).

[7]See Eq. (3.18) in Ref. 5, which contains a misprint and must be read: $\pi - \Theta_0 \leq (c/2)(\pi - \psi)$, where $x = \sin(\psi/2)$.

[8]See G. A. Baker, Jr., Adv. Theor. Phys. **1**, 1 (1965).

[9]D. S. Gaunt, Phys. Rev. **179**, 174 (1969); G. A. Baker, Jr., and P. Moussa, J. Appl. Phys. **43**, 1360 (1978).

[10]M. E. Fisher, Phys. Rev. Lett. **40**, 1610 (1978).

[11]P. Dienes, *The Taylor series* (Dover, New York, 1957), p. 324.

[12]U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications* (University of California Press, Berkeley and Los Angeles, 1958).

[13]D. and G. Choodnovsky (private communication).

[14]P. Deligne (private communication).

[15]Some of the results in this paragraph have been obtained in collaboration with B. Derrida.

[16]S. Karlin and W. J. Studden, *Tchebycheff Systems* (Interscience, New York, London, Sydney, 1966).

[17]J. A. Shohat and J. D. Tamarkin, *The Problem of Moments* (American Mathematical Society, Providence, Rhode Island, 1950).

[18]G. A. Baker, Jr., J. Math. Phys. **10**, 814 (1969).

[19]S. Aubry, "The new concept of transitions by breaking of analyticity in a crystallographic model," communication at the Symposium on Non-Linear Soliton Structure and the Dynamic of Condensed Matter, Oxford, England, June 1978.

[20]R. J. Rubin, J. Math. Phys. **11**, 1857 (1970).

Direct determination of the Langlands decompositions for the parabolic subalgebras of noncompact semisimple real Lie algebras

J. F. Cornwell

Department of Theoretical Physics, University of St Andrews, St Andrews, Fife, Scotland
(Received 23 May 1978)

A direct method for the determination of the Langlands decompositions for the parabolic subalgebras of any noncompact semisimple real Lie algebra is described in detail. The method is based on the canonical form of the Lie algebra. The physically important Lie algebras so(3,2) and so(4,2) are treated as illustrative examples.

I. INTRODUCTION

In this paper it will be shown that the method given earlier (Cornwell,[1] hereafter referred to as paper I) for the direct determination of the Iwasawa decomposition of a noncompact semisimple real Lie algebra can be extended very easily to give directly the Langlands decompositions for all the parabolic subalgebras. These decompositions form an essential part of the construction of various series of unitary irreducible representations of the corresponding semisimple Lie groups (cf. Lipsman[2]).

The definition of a general parabolic subalgebra and its Langlands decomposition is given in some detail in Sec. II. Also included there are the definitions of the "minimal" parabolic subalgebra and of "cuspidal" parabolic subalgebras, together with an outline of the role that all these subalgebras play in representation theory. (Further details may be found in the work of Warner.[3])

The direct method of construction of Langlands decompositions of the parabolic subalgebras is described in Sec. III. This involves an automorphic mapping V between two Cartan subalgebras that appear. This automorphism is examined in Secs. IV and V for the cases in which the noncompact real Lie algebras are generated by inner and outer involutive automorphisms respectively. The physically important Lie algebras so(3, 2) and so(4, 2) are treated as illustrative examples in Sec. IV.

II. PARABOLIC SUBALGEBRAS

As the definition and construction of the parabolic subalgebras of a noncompact semisimple real Lie algebra $\mathcal{L}$ are based on the Iwasawa decomposition of $\mathcal{L}$, it is necessary first to recall certain facts concerning the Iwasawa decomposition. The most important is the existence of a Cartan subalgebra $\mathcal{H}'$ of the complex extension $\tilde{\mathcal{L}}$ of $\mathcal{L}$ which does not coincide with the Cartan subalgebra $\mathcal{H}$ of $\tilde{\mathcal{L}}$ that arises in the canonical construction of $\mathcal{L}$ itself. (A more complete discussion was given in I, the conventions of which will be employed here without modification.)

In the canonical constructions of $\mathcal{L}$ of Cartan[4] and of Gantmacher,[5] $\mathcal{L}$ may be generated from its compact real form $\mathcal{L}_c$ by a chief involutive automorphism Z, which is defined[6] with respect to the Cartan subalgebra $\mathcal{H}$. More explicitly, $\mathcal{L} = \sqrt{Z}\,\mathcal{L}_c$, where $\sqrt{Z} = \frac{1}{2}(1-i)Z + \frac{1}{2}(1+i)I$ (I being the identity), and where for a chief inner automorphism $Z = \exp(\mathrm{ad}h)$, where $h \in \mathcal{H}$, while for a chief outer automorphism $Z = Z_0 \exp(\mathrm{ad}h)$, $h \in \mathcal{H}$, Z_0 being an outer automorphism associated with a "rotation" in $\mathcal{H}$. It will be assumed (as in I) that $\tilde{\mathcal{L}}$ and $\mathcal{L}_c$ are specified in the usual way by the roots defined with respect to *this* Cartan subalgebra $\mathcal{H}$.

The Cartan subalgebra $\mathcal{H}'$ of the Iwasawa decomposition appears in the following way. Let $\mathcal{K}$ be the maximal compact subalgebra of $\mathcal{L}$ defined such that $a \in \mathcal{K}$ if and only if $a \in \mathcal{L}$ and $Za = a$. Let $\mathcal{P}$ be the subspace of $\mathcal{L}$ such that $a \in \mathcal{P}$ if and only if $a \in \mathcal{L}$ and $Za = -a$. Let $\mathcal{A}_I$ be the maximal Abelian subalgebra of $\mathcal{P}$, and $\mathcal{M}_I$ the centralizer of $\mathcal{A}_I$ in $\mathcal{K}$. Suppose $\dim \mathcal{A}_I = m_I$, the so-called "split rank" of $\mathcal{L}$, and that H_j', $j = 1, 2, \ldots, m_I$, form a basis for $\mathcal{A}_I$. Then[7] the maximal Abelian subalgebra of $\mathcal{M}_I$ has dimension $(l - m_I)$, where l is the rank of $\tilde{\mathcal{L}}$, and may be taken to have as its basis $-iH_j'$, $j = m_I + 1, \ldots, l$. Then the complex Lie algebra having H_j', $j = 1, 2, \ldots, l$, as its basis is the Cartan subalgebra $\mathcal{H}'$ of $\tilde{\mathcal{L}}$ associated with the Iwasawa decomposition.

The final stage in the construction of the Iwasawa decomposition involves the determination of a nilpotent subalgebra $\mathcal{N}_I$ of $\mathcal{L}$ whose elements are linear combinations of those of both $\mathcal{K}$ and $\mathcal{P}$. This requires the roots $\alpha'(h')$ of $\tilde{\mathcal{L}}$ defined with respect to $\mathcal{H}'$ and the corresponding elements $e_{\alpha'}'$ of $\tilde{\mathcal{L}}$ defined such that

$$[e_{\alpha'}', h'] = \alpha'(h') \tag{1}$$

(for all $h' \in \mathcal{H}'$) and

$$B(e_{\alpha'}', e_{-\alpha'}') = -1.$$

Moreover, let $h_{\alpha'}' \in \mathcal{H}'$ be such that

$$B(h', h_{\alpha'}') = \alpha'(h')$$

for all $h' \in \mathcal{H}'$. Then

$$h_{\alpha'}' = \sum_{j=1}^{l} c_j(\alpha') H_j'$$

[where the coefficients $c_j(\alpha')$ are all real] and the root α' (defined with respect to $\mathcal{H}'$) is said to be a member of the set P' if and only if $c_j(\alpha') > 0$, where j is the least index such that $c_j(\alpha') \neq 0$. (That is, P' is the set of "positive" roots of $\tilde{\mathcal{L}}$ with respect to $\mathcal{H}'$). The set P' may be divided into two disjoint subsets P_-' and P_+' by the requirement that $\alpha' \in P_-'$ if and only if $\alpha' \in P'$ and

 0022-2488/79/040547-09$01.00

$$\alpha'(Zh') = \alpha'(h')$$

for all $h' \in H'$, and by the requirement that $\alpha' \in P'_+$ if and only if $\alpha' \in P'$ and $\alpha' \notin P'_-$. As $Zh' = -h'$ for all $h' \in A_I$, this implies that $\alpha'(h') = 0$ for all $\alpha' \in P'_-$ and all $h' \in A_I$, and also that if $\alpha' \in P'_+$, then there exists an $h' \in A_I$ such that $\alpha'(h') \neq 0$. If $\tilde{N}_I$ is the subspace of $\tilde{L}$ spanned by the elements $e'_{\alpha'}$ for all $\alpha' \in P'_+$, then Iwasawa[7] has shown that $\tilde{N}_I$ is a nilpotent subalgebra of $\tilde{L}$. The required subalgebra N_I of L is finally defined by $N_I = \tilde{N}_I \cap L$. The Iwasawa[7] decomposition of L is then

$$L = K \oplus A_I \oplus N_I,$$

and any subalgebra that is conjugate to

$$P_I = M_I \oplus A_I \oplus N_I$$

is called a "minimal parabolic subalgebra of L." (In I, A_I, M_I, N_I, and m_I were denoted by A, $C(A)$, N, and m respectively).

A general "*parabolic subalgebra*" of L may now be defined as any subalgebra of L that contains a minimal parabolic subalgebra of L. There exist 2^{m_I} conjugacy classes of parabolic subalgebras of L. In each such class there is a "*standard parabolic subalgebra* P_θ", which may be set up in the following way. (Further details may be found in the comprehensive account of Warner.[3])

First let Σ'_+ be the set of roots λ' that are the restrictions to A_I of the roots α' of the set P'_+, and put $\Sigma' = \Sigma'_+ \cup (-\Sigma'_+)$. Let $\Psi' = \{\lambda'_1, \lambda'_2, \ldots\}$ be the set of m_I roots that are the distinct restrictions to A_I of the simple roots $\{\alpha'_1, \alpha'_2, \ldots, \alpha'_l\}$ of $\tilde{L}$ defined with respect to H' which are contained in $P'_+ \cup (-P'_+)$. Then Ψ' forms a fundamental system of roots for Σ'.

Now let θ' be any subset of Ψ'. Clearly there are 2^{m_I} such subsets, including Ψ' itself and the empty set $\emptyset$. Let $\langle\theta'\rangle$ be the subset of Σ' that consists of linear combinations of the roots of θ', and let $\langle\theta'\rangle_+ = \Sigma'_+ \cap \langle\theta'\rangle$ and $\langle\theta'\rangle_- = (-\Sigma'_+) \cap \langle\theta'\rangle$.

The subalgebra A_θ of L may be defined as the set of elements h' of A_I such that $\lambda'(h') = 0$ for all $\lambda' \in \theta'$. Obviously A_θ is an Abelian subalgebra of A_I.

The set $A(\theta)$ may then be defined to be the subspace of A_I having basis elements $Q_{\lambda'}$ for all $\lambda' \in \theta'$, where $Q_{\lambda'}$ is itself defined as the element of A_I such that

$$B(h', Q_{\lambda'}) = \lambda'(h')$$

for all $h' \in A_I$. Obviously $A(\theta)$ is the orthogonal complement of A_θ in A_I relative to the Killing form.

Now let $L(\lambda')$ be the subspace of L consisting of elements a of L such that

$$[a, h'] = \lambda'(h')a \tag{2}$$

for all $h' \in A_I$. The subspaces $N_+(\theta)$, $N_-(\theta)$, and N_θ of L are then defined as the direct sums of the subspaces $L(\lambda')$ for all λ' of $\langle\theta'\rangle_+$, $\langle\theta'\rangle_-$, and $\{\Sigma'_+ - \langle\theta'\rangle_+\}$, respectively.

Finally the subalgebra M_θ of L is defined by

$$M_\theta = M_I \oplus N_+(\theta) \oplus N_-(\theta) \oplus A(\theta),$$

where again $\oplus$ merely denotes a vector space direct sum, so that $M_\theta \oplus A_\theta$ is the centralizer of A_θ in L. Then

$$P_\theta = M_\theta \oplus A_\theta \oplus N_\theta \tag{3}$$

is a standard parabolic subalgebra of L in which the Langlands decomposition has been explicitly displayed. [Again in (3) $\oplus$ denotes merely the vector space direct sum and does not imply that M_θ, A_θ, and N_θ mutually commute. For an account by Langlands of the significance of this decomposition and its generalization to algebraic groups, see Ref. 8.]

Clearly the parabolic subalgebra P_θ depends on the choice of the subset θ', the 2^{m_I} possible choices of θ' giving 2^{m_I} nonconjugate parabolic subalgebras. It is obvious that

$$A_I = A(\theta) \oplus A_\theta$$

and

$$N_I = N_+(\theta) \oplus N_\theta,$$

implying that

$$P_\theta = P_I \oplus N_-(\theta),$$

so that P_I is certainly a subalgebra of P_θ.

In the extreme case in which θ' is the empty set $\emptyset$, $\langle\theta'\rangle = \emptyset$ so that $A_\theta = A_I$. Moreover $A(\theta) = N_+(\theta) = N_-(\theta) = \{0\}$, the zero element of L, so that $M_\theta = M_I$. As $\{\Sigma'_+ - \langle\theta'\rangle_+\} = \Sigma'_+ = P'_+$, it follows that $N_\theta = N_I$. Thus the extreme choice $\theta' = \emptyset$ corresponds to P_θ being the minimal parabolic subalgebra P_I.

The other extreme choice is $\theta' = \Psi'$, for which $\langle\theta'\rangle = \Sigma'$. In this case $A_\theta = \{0\}$, so that $M_\theta = L$. Finally $\{\Sigma'_+ - \langle\theta'\rangle_+\} = \emptyset$ so that $N_\theta = \{0\}$. Thus with the extreme choice $\theta' = \Psi'$, the parabolic subalgebra P_θ is L itself.

These considerations imply that there are no nontrivial nonminimal parabolic subalgebras when $m_I = 1$. Two physically important examples for which this is the case are $L = \mathrm{so}(4,1)$ and $L = \mathrm{so}(3,1) \sim \mathrm{sl}(2,c)$.

As in the case of the Iwasawa decomposition, it is the determination of the subspaces $N_+(\theta)$, $N_-(\theta)$, and N_θ that present the most difficulties, for the usual method involves solving the set of eigenvalue equations (2) to determine the subspaces $L(\lambda')$. As this essentially implies redetermining the structure of $\tilde{L}$, it is totally unnecessary. In fact the required subspaces are merely generated by the automorphic images of the e_α of the original canonical basis for $\tilde{L}$, as will be demonstrated explicitly in Sec. III.

A *real* Cartan subalgebra H'_r of L may be defined as a real subalgebra of L whose complex extension H' is a Cartan subalgebra of $\tilde{L}$. It is well known[2,3] that although all the complex Cartan subalgebras of $\tilde{L}$ are conjugate, there exists in general more than one conjugacy class of real Cartan subalgebras of L, but the number of such classes is always finite. Moreover within each such conjugacy class a real Cartan subalgebra H'_r may be chosen so that

$$H'_r = (H'_r \cap K) \oplus (H'_r \cap P),$$

such a real Cartan subalgebra being said to be "Z-invariant" or "Z-stable." H'_r is described as being "maximally split" if $H'_r \cap P = A_I$, which is obviously the case for the real Cartan subalgebra with basis H'_j, $j = 1, 2, \ldots, m_I$, and $-iH'_j$, $j = m_I + 1, \ldots, l$, associated with the Iwasawa decomposition. While there is only one conjugacy class of real Cartan subalgebras that are maximally split, there may exist nonconjugate Z-invariant real Cartan subalgebras $\mathcal{H}'_{r1}$ and $\mathcal{H}'_{r2}$ such that $\dim(\mathcal{H}'_{r1}\cap\mathcal{K})=\dim(\mathcal{H}'_{r2}\cap\mathcal{K})$ and $\dim(\mathcal{H}'_{r1}\cap\mathcal{P})=\dim(\mathcal{H}'_{r2}\cap\mathcal{P})$. A parabolic subalgebra $\mathcal{P}_\theta$ is said to be *"cuspidal"* if and only if there exists a Z-invariant real Cartan subalgebra $\mathcal{H}'_r$ such that

$$\mathcal{A}_\theta=\mathcal{H}'_r\cap\mathcal{P}.$$

Clearly the minimal parabolic subalgebra is cuspidal. However, at the other extreme, $\mathcal{L}$ itself is cuspidal if and only if $\mathcal{K}$ contains a real Cartan subalgebra of $\mathcal{L}$, which is the case if and only if $\mathcal{L}$ is generated by an *inner* involutive automorphism (cf. Secs. IV and V). Obviously if $\mathcal{P}_{\theta_1}$ and $\mathcal{P}_{\theta_2}$ are two cuspidal parabolic subalgebras that are nonconjugate and $\mathcal{H}'_{rj}\cap\mathcal{P}=\mathcal{A}_{\theta_j}$, $j=1,2$, then the real Cartan subalgebras $\mathcal{H}'_{r1}$ and $\mathcal{H}'_{r2}$ must also be nonconjugate. Thus the number of conjugacy classes of parabolic subalgebras of $\mathcal{L}$ containing cuspidal parabolic subalgebras cannot exceed the number of conjugacy classes of real subalgebras, this latter number being known for every $\mathcal{L}$ from the investigations of Sugiura.[9]

One major application of the parabolic subalgebras lies in the construction of unitary representations of noncompact semisimple Lie groups, which are, of course, all necessarily infinite-dimensional (except for the trivial identity representation). Suppose that $\mathcal{G}$ is such a group, $\mathcal{L}$ is its corresponding real Lie algebra, and $\mathcal{B}_\theta$ is the "parabolic subgroup" of $\mathcal{G}$ that has $\mathcal{P}_\theta$ as its real Lie algebra. Then various series of unitary representations of $\mathcal{G}$ can be induced[10] from the representations of the various parabolic subgroups $\mathcal{B}_\theta$. For example the "principal series" or "principal P-series"[3] representations are induced from the minimal parabolic subgroup. The other cuspidal parabolic subgroups yield the "principal nondegenerate" series of representations, while the noncuspidal parabolic subgroups yield the "degenerate" series. Further details may be found in the review of Lipsman.[2]

III. DIRECT DETERMINATION OF THE PARABOLIC SUBALGEBRAS AND THEIR LANGLANDS DECOMPOSITIONS

The direct determination of the Iwasawa decomposition given in I was based on the fact that there exists an inner automorphism V of $\tilde{\mathcal{L}}$ that maps $\mathcal{H}'$ into $\mathcal{H}$. Let $V\mathcal{A}_I$, $V\mathcal{A}_\theta$, and $V\mathcal{A}(\theta)$ denote the images under the automorphism V of $\mathcal{A}_I$, $\mathcal{A}_\theta$, and $\mathcal{A}(\theta)$, respectively. Suppose that $Vh'=h$, where $h'\in\mathcal{H}'$ and $h\in\mathcal{H}$. On applying V to (1), one obtains

$$[Ve'_{\alpha'},h]=\alpha'(V^{-1}h)(Ve'_{\alpha'}),$$

which implies that $\alpha'(V^{-1}h)$ is a root of $\tilde{\mathcal{L}}$ with respect to $\mathcal{H}$, and $Ve'_{\alpha'}$ is the corresponding element of $\tilde{\mathcal{L}}$. Let $\alpha'(V^{-1}h)$ be denoted by $\alpha(h)$; that is, let

$$\alpha(h)=\alpha'(V^{-1}h) \tag{4}$$

for all $h\in\mathcal{H}$. Equation (4) establishes a one-to-one correspondence between the roots defined with respect to $\mathcal{H}'$ and the roots defined with respect to $\mathcal{H}$. In this correspondence a fundamental root system with respect to $\mathcal{H}'$ is mapped into a fundamental root system with respect to $\mathcal{H}$, and similarly a fundamental system of roots restricted with respect to $\mathcal{A}_I$ is mapped into a fundamental system of roots restricted with respect to $V\mathcal{A}_I$. Moreover

$$Ve'_{\alpha'}=e_\alpha,$$

and conversely

$$e'_{\alpha'}=V^{-1}e_\alpha.$$

Similarly,

$$Vh'_{\alpha'}=h_\alpha.$$

If $H_1,H_2,\ldots,H_l$ of $\mathcal{H}$ are defined by

$$H_j=VH'_j,\quad j=1,2,\ldots,l,$$

where $H'_1,H'_2,\ldots,H'_l$ are as given in Sec. II, then, on writing

$$h_\alpha=\sum_{j=1}^{l}b_j(\alpha)H_j,$$

the sets P, P_-, and P_+ of roots specified with respect to $\mathcal{H}$ may be defined as follows:

(a) $\alpha\in P$ if and only if $b_j(\alpha)>0$, where j is the least index such that $b_j(\alpha)\neq 0$.

(b) $\alpha\in P_-$ if and only if $\alpha\in P$ and $\alpha(H_j)=0$ for $j=1,2,\ldots,m_I$.

(c) $\alpha\in P_+$ if and only if $\alpha\in P$ and $\alpha\notin P_-$.

Clearly P, P_-, and P_+ are the images under the correspondence (4) of P', P'_-, and P'_+, respectively.

Now define Σ_+ to be the set of roots λ that are the restrictions to $V\mathcal{A}_I$ of the roots α of the set P_+, and put $\Sigma=\Sigma_+\cup(-\Sigma_+)$. Similarly let $\Psi=\{\lambda_1,\lambda_2,\cdots\}$ be the set of m_I roots that are the restrictions to $V\mathcal{A}_I$ of the simple roots $\{\alpha_1,\alpha_2,\ldots,\alpha_l\}$ of $\tilde{\mathcal{L}}$ defined with respect to $\mathcal{H}$ which are contained in $P_+\cup(-P_+)$, so that Ψ forms a fundamental system of roots for Σ. Moreover let θ be any subset of Ψ, and $\langle\theta\rangle$ the subset of Σ that consists of linear combinations of the roots of θ. Finally let $\langle\theta\rangle_+=\Sigma_+\cap\langle\theta\rangle$ and $\langle\theta\rangle_-=(-\Sigma_+)\cap\langle\theta\rangle$. Obviously Σ_+, Σ, Ψ, θ, $\langle\theta\rangle$, $\langle\theta\rangle_+$, and $\langle\theta\rangle_-$ are the images of Σ'_+, Σ', Ψ', θ', $\langle\theta'\rangle$, $\langle\theta'\rangle_+$, and $\langle\theta'\rangle_-$ under the mapping (4).

It follows immediately that $V\mathcal{A}_\theta$ is the subalgebra of $V\mathcal{A}_I$ consisting of elements h of $V\mathcal{A}_I$ such that $\lambda(h)=0$ for all $\lambda\in\theta$. Similarly, $V\mathcal{A}(\theta)$ is the subalgebra of $V\mathcal{A}_I$ having basis elements Q_λ for all $\lambda\in\theta$, where Q_λ is itself defined as the element of $V\mathcal{A}_I$ such that

$$B(h,Q_\lambda)=\lambda(h)$$

for all $h\in V\mathcal{A}_I$. Thus Q_λ must be a linear combination of the elements h_α for which the restriction of α to $V\mathcal{A}_I$ is λ. The condition that Q_λ must be a member of $V\mathcal{A}_I$ then completely determines Q_λ up to a real multiplicative factor.

From (2) it follows that $\mathcal{L}(\lambda')=\tilde{\mathcal{L}}(\lambda')\cap\mathcal{L}$, where $\tilde{\mathcal{L}}(\lambda')$ is the subspace of $\tilde{\mathcal{L}}$ that has as its basis elements the set of all elements $e'_{\alpha'}$ such that α' is a root of $\tilde{\mathcal{L}}$ defined with respect to $\mathcal{H}'$ whose restriction to $\mathcal{A}_I$ is λ'. Thus $\tilde{\mathcal{L}}(\lambda')$ has basis elements $V^{-1}e_\alpha$, where α is the root of $\tilde{\mathcal{L}}$ defined with respect to $\mathcal{H}$ whose restriction λ to $V\mathcal{A}_I$ is the image of λ' under the mapping (4). Consequently $\mathcal{N}_+(\theta)$, $\mathcal{N}_-(\theta)$, and $\mathcal{N}_\theta$ are given by $\mathcal{N}_+(\theta)=\tilde{\mathcal{N}}_+(\theta)\cap\mathcal{L}$, $\mathcal{N}_-(\theta)=\tilde{\mathcal{N}}_-(\theta)\cap\mathcal{L}$, and $\mathcal{N}_\theta=\tilde{\mathcal{N}}_\theta\cap\mathcal{L}$, where $\tilde{\mathcal{N}}_+(\theta)$, $\tilde{\mathcal{N}}_-(\theta)$, and $\tilde{\mathcal{N}}_\theta$ are subspaces of $\tilde{\mathcal{L}}$ that are generated by the elements $V^{-1}e_\alpha$ for all roots α of $\mathcal{L}$ defined with respect to $\mathcal{H}$ whose restrictions to $V\mathcal{A}_I$ are contained in $\langle\theta\rangle_+$, $\langle\theta\rangle_-$, and $\{\Sigma_+-\langle\theta\rangle_+\}$, respectively. [It should be noted that the elements $V^{-1}e_\alpha$ of $\tilde{\mathcal{N}}_+(\theta)$ and $\tilde{\mathcal{N}}_\theta$ are precisely the set of elements that appear in the direct determination of the Iwasawa decomposition as generators of $\tilde{\mathcal{N}}_I$, where $\mathcal{N}_I=\tilde{\mathcal{N}}_I\cap\mathcal{L}$. (In I, $\tilde{\mathcal{N}}_I$ was denoted by $\tilde{\mathcal{N}}$.) Thus there is no effort involved in the determination of $\tilde{\mathcal{N}}_+(\theta)$ and $\tilde{\mathcal{N}}_\theta$.]

It is obvious that the choice $\theta=\emptyset$, the empty set, yields $\mathcal{P}_\theta=\mathcal{P}_I$, and that $\theta=\Psi$ gives $\mathcal{P}_\theta=\mathcal{L}$, so the only interest attaches to the cases in which θ is a nonempty proper subset of Ψ.

With the Iwasawa decomposition of $\mathcal{L}$ constructed as in I, the procedure for determining $\mathcal{A}_\theta$, $\mathcal{N}_\theta$, and $\mathcal{M}_\theta$ may be summarized by listing the ten stages involved:

(a) Construct the fundamental root system Ψ.

(b) Choose a nonempty proper subset θ of Ψ.

(c) Construct $V\mathcal{A}_\theta$ by the criteria that $h\in V\mathcal{A}_\theta$ if $h\in V\mathcal{A}_I$ and $\lambda(h)=0$ for all $\lambda\in\theta$.

(d) The elements of $\mathcal{A}_\theta$ are then given by $V^{-1}h$, where $h\in V\mathcal{A}_\theta$, using the properties of the automorphism V described in more detail in Secs. IV and V.

(e) For each $\lambda\in\theta$, construct Q_λ from the criteria that Q_λ is a member of $V\mathcal{A}_I$ and is a linear combination of all the elements $h_\alpha\in\mathcal{H}$ for which the restriction of α to $V\mathcal{A}_I$ is λ.

(f) Evaluate $V^{-1}Q_\lambda$ for all $\lambda\in\theta$, these being the basis elements of $\mathcal{A}(\theta)$.

(g) Evaluate $V^{-1}e_\alpha$ for all α whose restriction to $V\mathcal{A}_I$ is contained in $\langle\theta\rangle_-$, these being the basis elements of $\tilde{\mathcal{N}}_-(\theta)$.

(h) Divide the subalgebra $\tilde{\mathcal{N}}_I$ associated with the Iwasawa decomposition into its two complementary subspaces $\tilde{\mathcal{N}}_+(\theta)$ and $\tilde{\mathcal{N}}_\theta$ by the criteria that $V^{-1}e_\alpha\in\tilde{\mathcal{N}}_+(\theta)$ if the restriction of α to $V\mathcal{A}_I$ is contained in $\langle\theta\rangle_+$ while $V^{-1}e_\alpha\in\tilde{\mathcal{N}}_\theta$ if this restriction is contained in $\{\Sigma_+-\langle\theta\rangle_+\}$.

(i) Determine $\mathcal{N}_+(\theta)$, $\mathcal{N}_-(\theta)$, and $\mathcal{N}_\theta$ by $\mathcal{N}_+(\theta)=\tilde{\mathcal{N}}_+(\theta)\cap\mathcal{L}$, $\mathcal{N}_-(\theta)=\tilde{\mathcal{N}}_-(\theta)\cap\mathcal{L}$, and $\mathcal{N}_\theta=\tilde{\mathcal{N}}_\theta\cap\mathcal{L}$, respectively.

(j) Determine $\mathcal{M}_\theta$ from the definition

$$\mathcal{M}_\theta=\mathcal{M}_I\oplus\mathcal{N}_+(\theta)\oplus\mathcal{N}_-(\theta)\oplus\mathcal{A}(\theta).$$

(It should be emphasized that of these stages only (g) involves anything other than the most simple algebraic manipulation.)

IV. LANGLANDS DECOMPOSITIONS FOR PARABOLIC SUBALGEBRAS OF REAL LIE ALGEBRAS GENERATED BY CHIEF INNER INVOLUTIVE AUTOMORPHISMS

A. General theory

The chief inner involutive automorphism $Z=\exp(\mathrm{ad}h)$ is diagonal with respect to the canonical basis of $\mathcal{L}_c$. $\mathcal{K}$ has a basis consisting of ih_α for $\alpha=\alpha_j$, $j=1,2,\ldots,l$, together with $(e_\alpha+e_{-\alpha})$ and $i(e_\alpha-e_{-\alpha})$ for all α such that $\exp[\alpha(h)]=1$, while the basis of $\mathcal{P}$ consists of $i(e_\alpha+e_{-\alpha})$ and $(e_\alpha-e_{-\alpha})$ for all α such that $\exp[\alpha(h)]=-1$. Consequently $\mathcal{A}_I$ may be taken to have a basis consisting of elements of the form $i(e_\alpha+e_{-\alpha})$. Let $\mathcal{R}(\mathcal{A}_I)$ denote the set of r_I positive roots α that appear in this way in $\mathcal{A}_I$. Obviously in this case $r_I=m_I$. [In order to ensure that $\mathcal{A}_I$ is Abelian, $\mathcal{R}(\mathcal{A}_I)$ must be chosen so that if β and γ are any two roots in $\mathcal{R}(\mathcal{A}_I)$, then neither $\beta+\gamma$ nor $\beta-\gamma$ is a root of $\tilde{\mathcal{L}}$.]

At this point it is worth noting that a small but useful simplification is possible in the method described in I, for the maximal Abelian subalgebra of $\mathcal{M}_I$ (which was denoted by $\mathcal{M}$ in I) can *always* be chosen to lie *entirely* within $\mathcal{H}\cap\mathcal{L}$. This means that $-iH'_j$, $j=m_I+1,\ldots,l$, can be chosen to be certain linear combinations of the basis elements ih_α of $\mathcal{H}\cap\mathcal{L}$. Correspondingly the set $\mathcal{R}_\mathcal{M}$ of Sec. IV A of I can be taken to be empty. Consequently the automorphism V of $\mathcal{L}$ that maps $\mathcal{H}'$ into $\mathcal{H}$ [and which in I was given by Eq. (15)] is now given by

$$V=\prod_\alpha V_\alpha, \tag{5}$$

where

$$V_\alpha=\exp[\mathrm{ad}\{ia_\alpha(e_\alpha-e_{-\alpha})\}],$$
$$a_\alpha=\pi/[8(\alpha,\alpha)]^{1/2},$$

and where the product in (5) is now over the roots of $\mathcal{R}(\mathcal{A}_I)$ alone. {This simplification is always possible because if $\beta\in\mathcal{R}_\mathcal{M}$ and $\alpha\in\mathcal{R}(\mathcal{A}_I)$ then $[e_{\pm\beta},e_{\pm\alpha}]=0$. As $h_\beta=[e_\beta,e_{-\beta}]$ this implies that $[h_\beta,i(e_\alpha+e_{-\alpha})]=0$, so that ih_β may be used in place of $(e_\beta+e_{-\beta})$ as a basis element of the maximal Abelian subalgebra of $\mathcal{M}_I$.} All the properties of the inner automorphisms V_α and their inverses that are needed for the various stages of the procedure of Sec. III are summarized in the Appendix of I.

To facilitate the application of the above method, a complete specification of the maximal compact subalgebra $\mathcal{K}$ and the subspace $\mathcal{P}$, together with a convenient choice of $\mathcal{R}(\mathcal{A}_I)$ and the maximal Abelian subalgebra of $\mathcal{M}_I$ is given in the Appendix for each of the classical noncompact simple Lie algebras $\mathcal{L}$ that are generated by inner involutive automorphisms.

B. Examples

1. $\mathcal{L}=\mathrm{so}(3,2)$

The positive roots of $\tilde{\mathcal{L}}=B_2$ are α_1, α_2, $\alpha_1+\alpha_2$, and $\alpha_1+2\alpha_2$. As noted in the Appendix, the chief inner involutive automorphism $Z=\exp(\mathrm{ad}h)$ generating $\mathcal{L}=\mathrm{so}(3,2)$ may be chosen[11] so that

$$\exp\{\alpha(h)\}=\begin{cases}1, & \alpha=\alpha_1+\alpha_2,\\ -1, & \alpha=\alpha_1,\ \alpha_2,\ \alpha_1+2\alpha_2.\end{cases}$$

Thus $\mathcal{K}$ has basis ih_α with $\alpha=\alpha_1$, α_2, together with $(e_\alpha+e_{-\alpha})$ and $i(e_\alpha-e_{-\alpha})$ with $\alpha=\alpha_1+\alpha_2$, while $\mathcal{P}$ has basis $i(e_\alpha+e_{-\alpha})$ and $(e_\alpha-e_{-\alpha})$ with $\alpha=\alpha_1$, α_2, $\alpha_1+2\alpha_2$. In this case $m_I=r_I=2$, and (as noted in the Appendix) a convenient choice for $\mathcal{R}(\mathcal{A}_I)$ is

$$\mathcal{R}(\mathcal{A}_I)=\{\alpha_1,\alpha_1+2\alpha_2\},$$

with $\mathcal{M}_I=\{0\}$. As $m_I=l=2$, there is no restriction involved in this case in considering the subalgebra $V\mathcal{A}_I$ of $\mathcal{H}$. Thus $\Psi=\{\lambda_1,\lambda_2\}$, where $\lambda_1=\alpha_1$ and $\lambda_2=\alpha_2$. Moreover, as shown in Sec. IV of I,

$$P_+=P=\{-\alpha_1,\alpha_2,-(\alpha_1+\alpha_2),-(\alpha_1+2\alpha_2)\},$$

so that

$$\Sigma_+=\{-\lambda_1,\lambda_2,-(\lambda_1+\lambda_2),-(\lambda_1+2\lambda_2)\}.$$

$\mathcal{L}$ = so(3, 2) has four standard parabolic subalgebras, namely the minimal parabolic subalgebra (found in Sec. IV of I), $\mathcal{L}$ itself, and two others which will now be determined.

(i) Choose $\theta=\{\lambda_1\}$. Then $V\mathcal{A}_\theta$ has generator h_α with $\alpha=\alpha_1+2\alpha_2$, and so [on using Eq. (A2) of I] $\mathcal{A}_\theta$ has generator $i(e_\alpha+e_{-\alpha})$ with $\alpha=\alpha_1+2\alpha_2$. As $Q_{\lambda_1}=h_{\alpha_1}$, Eq. (A2) of I also shows that $\mathcal{A}(\theta)$ has generator $i(e_{\alpha_1}+e_{-\alpha_1})$.

Because $\langle\theta\rangle_-=\{\lambda_1\}$, $\tilde{\mathcal{N}}_-(\theta)$ is generated by

$$V^{-1}e_{\alpha_1}=\tfrac{1}{2}(e_{\alpha_1}-e_{-\alpha_1})-\tfrac{1}{2}i6^{1/2}h_{\alpha_1}$$

[where Eq. (A3) of I has been used]. As $\langle\theta\rangle_+=\{-\lambda_1\}$ and $\{\Sigma_+-\langle\theta\rangle_+\}=\{\lambda_2,-(\lambda_1+\lambda_2),-(\lambda_1+2\lambda_2)\}$, $\tilde{\mathcal{N}}_+(\theta)$ is generated by

$$V^{-1}e_{-\alpha_1}=-\tfrac{1}{2}(e_{\alpha_1}-e_{-\alpha_1})-\tfrac{1}{2}ih^{1/2}h_{\alpha_1} \tag{6}$$

and $\tilde{\mathcal{N}}_\theta$ is generated by

$$V^{-1}e_{\alpha_2}=\tfrac{1}{2}(e_{\alpha_2}-se_{-\alpha_2})-\tfrac{1}{2}is'(e_{\alpha_1+\alpha_2}-se_{-(\alpha_1+\alpha_2)}),$$
$$V^{-1}e_{-(\alpha_1+\alpha_2)}=\tfrac{1}{2}iss'(e_{\alpha_2}+se_{-\alpha_2})+\tfrac{1}{2}s(e_{\alpha_1+\alpha_2}+se_{-(\alpha_1+\alpha_2)}),$$
$$V^{-1}e_{-(\alpha_1+2\alpha_2)}=-\tfrac{1}{2}(e_{\alpha_1+2\alpha_2}-e_{-(\alpha_1+2\alpha_2)})-\tfrac{1}{2}i6^{1/2}\{h_{\alpha_1}+2h_{\alpha_2}\}, \tag{7}$$

where $s=\operatorname{sgn}(N_{\alpha_1,\alpha_2}N_{\alpha_1+2\alpha_2,-\alpha_2})$, $s'=\operatorname{sgn}N_{\alpha_1,\alpha_2}$.

Consequently the basis elements of $\mathcal{N}_\theta$ are

$$-\tfrac{1}{2}(e_{\alpha_1+2\alpha_2}-e_{-(\alpha_1+2\alpha_2)})-\tfrac{1}{2}i6^{1/2}(h_{\alpha_1}+2h_{\alpha_2}),$$
$$\tfrac{1}{2}(e_{\alpha_2}-e_{-\alpha_2})-\tfrac{1}{2}is'(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)}),$$
$$\tfrac{1}{2}i(e_{\alpha_2}+e_{-\alpha_2})+\tfrac{1}{2}s'(e_{\alpha_1+\alpha_2}+e_{-(\alpha_1+\alpha_2)}),$$

and the basis elements of $\mathcal{M}_\theta$ may be taken to be $i(e_{\alpha_1}+e_{-\alpha_1})$, $(e_{\alpha_1}-e_{-\alpha_1})$, and ih_{α_1}. Clearly ih_{α_1} and $i(e_{\alpha_1+2\alpha_2}+e_{-(\alpha_1+2\alpha_2)})$ are generators of a real Z-invariant Cartan subalgebra $\mathcal{H}_r'$ for which $\mathcal{H}_r'\cap\mathcal{P}=\mathcal{A}_\theta$, so this parabolic subalgebra $\mathcal{P}_\theta$ is cuspidal.

(ii) Choose $\theta=\{\lambda_2\}$. In this case $V\mathcal{A}_\theta$ has generator $\{h_{\alpha_1}+h_{\alpha_1+2\alpha_2}\}$, so that [by Eq. (A2) of I] $\mathcal{A}_\theta$ has generator $\{i(e_{\alpha_1}+e_{-\alpha_1})+i(e_{\alpha_1+2\alpha_2}+e_{-(\alpha_1+2\alpha_2)})\}$. As $Q_{\lambda_2}=h_{\alpha_2}$ $(=\tfrac{1}{2}h_{\alpha_1+2\alpha_2}-\tfrac{1}{2}h_{\alpha_1})$, $\mathcal{A}(\theta)$ has generator $\{i(e_{\alpha_1+2\alpha_2}+e_{-(\alpha_1+2\alpha_2)})-i(e_{\alpha_1}+e_{-\alpha_1})\}$.

Moreover, as $\langle\theta\rangle_-=\{-\lambda_2\}$, Eqs. (A6) and (A7) of I show that $\tilde{\mathcal{N}}_-(\theta)$ is generated by

$$V^{-1}e_{-\alpha_2}=(-s)\{\tfrac{1}{2}(e_{\alpha_2}-se_{-\alpha_2})+\tfrac{1}{2}is'(e_{\alpha_1+\alpha_2}-se_{-(\alpha_1+\alpha_2)})\}.$$

As $\langle\theta\rangle_+=\{\lambda_2\}$ and $\{\Sigma_+-\langle\theta\rangle_+\}=\{-\lambda_1,-(\lambda_1+\lambda_2),-(\lambda_1+2\lambda_2)\}$, $\tilde{\mathcal{N}}_+(\theta)$ is generated by $V^{-1}e_{\alpha_2}$ and $\tilde{\mathcal{N}}_\theta$ is generated by $V^{-1}e_{-\alpha_1}$, $V^{-1}e_{-(\alpha_1+\alpha_2)}$, and $V^{-1}e_{-(\alpha_1+2\alpha_2)}$, all of which are given in Eqs. (6) and (7). Consequently the basis elements of $\mathcal{N}_\theta$ may be taken to be $V^{-1}e_{-\alpha_1}$, $V^{-1}e_{-(\alpha_1+2\alpha_2)}$, and

TABLE I. Correspondence between the generators of the canonical basis of so(3, 2) and the "physical" generators of the de Sitter algebra.

ih_{α_1}	$\frac{1}{6}(L_{12}+L_{45})$
ih_{α_2}	$-\frac{1}{3}L_{45}$
$i(e_{\alpha_1}+e_{-\alpha_1})$	$-6^{-1/2}(L_{14}+L_{25})$
$(e_{\alpha_1}-e_{-\alpha_1})$	$-6^{-1/2}(L_{15}-L_{24})$
$i(e_{\alpha_2}+e_{-\alpha_2})$	$3^{-1/2}L_{35}$
$(e_{\alpha_2}-e_{-\alpha_2})$	$3^{-1/2}L_{34}$
$(e_{\alpha_1+\alpha_2}+e_{-(\alpha_1+\alpha_2)})s'$	$3^{-1/2}L_{23}$
$i(e_{\alpha_1\alpha_2}-e_{-(\alpha_1\alpha_2)})s'$	$-3^{-1/2}L_{13}$
$i(e_{\alpha_1+2\alpha_2}+e_{-(\alpha_1+2\alpha_2)})s'$	$6^{-1/2}(L_{14}-L_{25})$
$(e_{\alpha_1+2\alpha_2}-e_{-(\alpha_1+2\alpha_2)})s$	$-6^{-1/2}(L_{15}+L_{24})$

$$i(e_{\alpha_2}+e_{-\alpha_2})+s'(e_{\alpha_1+\alpha_2}+e_{-(\alpha_1+\alpha_2)}),\quad \text{if } s=+1,$$
$$(e_{\alpha_2}-e_{-\alpha_2})-is'(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)}),\quad \text{if } s=-1.$$

Similarly the basis elements of $\mathcal{M}_\theta$ may be taken to be

$$\{i(e_{\alpha_1+2\alpha_2}+e_{-(\alpha_1+2\alpha_2)})-i(e_{\alpha_1}+e_{-\alpha_1})\},$$

together with

$$(e_{\alpha_2}-e_{-\alpha_2})\ \text{and}\ i(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)}),\quad \text{if } s=+1,$$
$$i(e_{\alpha_2}+e_{-\alpha_2})\ \text{and}\ (e_{\alpha_1+\alpha_2}+e_{-(\alpha_1+\alpha_2)}),\quad \text{if } s=-1.$$

Clearly $\{i(e_{\alpha_1+2\alpha_2}+e_{-(\alpha_1+2\alpha_2)})+i(e_{\alpha_1}+e_{-\alpha_1})\}$ together with $i(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)})$ (if $s=+1$) or $(e_{\alpha_1+\alpha_2}+e_{-(\alpha_1+\alpha_2)})$ (if $s=-1$) are generators of a real Z-invariant Cartan subalgebra $\mathcal{H}_r'$ for which $\mathcal{H}_r'\cap\mathcal{P}=\mathcal{A}_\theta$, so this parabolic subalgebra $\mathcal{P}_\theta$ is also cuspidal.

It is well known (cf. Inönü and Wigner[12]) that the Lie algebra of the Poincaré group can be obtained from the Lie algebra of the de Sitter group SO(3, 2) by "contraction." In fact if the metric tensor of SO(3, 2) is chosen so that $g_{mn}=\operatorname{diag}\{1,1,1,-1,-1\}$, then the ten generators of so(3, 2) may be taken to be L_{mn} ($m,n=1,2,3,4,5$, $m<n$) with $[L_{mn},L_{rs}]=-g_{mr}L_{ns}+g_{ms}L_{nr}+g_{nr}L_{ms}-g_{ns}L_{mr}$. Then $L_{\mu\nu}$ ($\mu,\nu=1,2,3,4$, $\mu<\nu$) are the generators of homogeneous Lorentz transformations and, after contraction, $L_{\mu5}$ ($\mu=1,2,3,4$) become the generators of space–time translations. The correspondence between the generators of the canonical basis and the $L_{\mu\nu}$ [obtained using a mapping derived elsewhere (Cornwell[13])] is given in Table I. It follows from I that the generators of the subalgebra of the Iwasawa decomposition for so(3, 2) are:

$\mathcal{A}_I$: $\{L_{14},L_{25}\}$, $\mathcal{M}_I$: zero-dimensional,

$\mathcal{N}_I$: $\{L_{15}-L_{45},L_{12}+L_{24},L_{13}+L_{34},L_{23}+L_{35}\}$,

where the convention has been adopted that $s=-1$. [In the choice of the Iwasawa decomposition for so(3, 2) made by Ehrman,[14] $\mathcal{A}_I$ is exactly as given here but $\mathcal{N}_I$ is slightly different.] Similarly, the nonminimal nontrivial parabolic subalgebras obtained above have the following "physical" generators:

(i) For $\theta=\{\lambda_1\}$,

$$\mathcal{A}_\theta=\{L_{14}-L_{25}\},\quad \mathcal{M}_\theta=\{L_{12}+L_{45}, L_{14}+L_{25}, -L_{15}+L_{24}\},$$
$$\mathcal{N}_\theta=\{L_{13}+L_{34}, L_{12}+L_{24}, L_{45}-L_{12}-L_{24}-L_{15}\}.$$

(ii) For $\theta=\{\lambda_2\}$,

$$\mathcal{A}_\theta=\{L_{14}\},\quad \mathcal{M}_\theta=\{L_{23}, L_{25}, L_{35}\},$$
$$\mathcal{N}_\theta=\{L_{12}+L_{24}, -L_{15}+L_{45}, L_{13}+L_{34}\}.$$

2. $\mathcal{L}=\mathrm{so}(4,2)$ [$\sim$ su(2,2)]

In this case $\tilde{\mathcal{L}}=D_3(\sim A_3)$, the positive roots of which are $\alpha_1, \alpha_2, \alpha_3, \alpha_1+\alpha_2, \alpha_1+\alpha_3, \alpha_1+\alpha_2+\alpha_3$. The chief inner involutive automorphism $Z=\exp(\mathrm{ad}h)$ generating $\mathcal{L}=\mathrm{so}(4,2)$ may be chosen[11] so that

$$\exp[\alpha(h)]=\begin{cases}1, & \alpha=\alpha_2,\alpha_3,\\ -1, & \alpha=\alpha_1, \alpha_1+\alpha_2, \alpha_1+\alpha_3, \alpha_1+\alpha_2+\alpha_3.\end{cases}$$

This implies that $\mathcal{K}$ has basis ih_α with $\alpha=\alpha_1, \alpha_2, \alpha_3$, together with $(e_\alpha+e_{-\alpha})$ and $i(e_\alpha-e_{-\alpha})$ with $\alpha=\alpha_2, \alpha_3$, while $\mathcal{P}$ has basis $i(e_\alpha+e_{-\alpha})$ and $(e_\alpha-e_{-\alpha})$ with $\alpha=\alpha_1$, $\alpha_1+\alpha_2, \alpha_1+\alpha_3, \alpha_1+\alpha_2+\alpha_3$. Then, as noted in the Appendix, $m_I=r_I=2$ and a convenient choice of $\mathcal{R}(\mathcal{A}_I)$ is

$$\mathcal{R}(\mathcal{A}_I)=\{\alpha_1, \alpha_1+\alpha_2+\alpha_3\},$$

so that, as in Sec. IV of I, the generators of $V\mathcal{A}_I$ may be taken to be

$$H_1=-\{2/(\alpha_1,\alpha_1)\}^{1/2}h_{\alpha_1} \tag{8}$$

and

$$H_2=-\{2/(\alpha_1+\alpha_2+\alpha_3, \alpha_1+\alpha_2+\alpha_3)\}^{1/2}(h_{\alpha_1}+h_{\alpha_2}+h_{\alpha_3}), \tag{9}$$

where $(\alpha_1,\alpha_1)=(\alpha_1+\alpha_2+\alpha_3, \alpha_1+\alpha_2+\alpha_3)=\frac{1}{4}$. Also $\mathcal{M}_I$ is one-dimensional and may be chosen to have the generator

$$-iH_3^{\xi}=i(h_{\alpha_2}-h_{\alpha_3}),$$

so that $H_3=H_3^{\xi}$.

It follows that λ_1 is the restriction of α_1 to $V\mathcal{A}_I$, whereas λ_2 is the restriction of α_2 and α_3. Consequently $\lambda_1+\lambda_2$ is the restriction of $\alpha_1+\alpha_2$ and $\alpha_1+\alpha_3$, while $\lambda_1+2\lambda_2$ is the restriction of $\alpha_1+\alpha_2+\alpha_3$. As in Sec. IV of I,

$$P_+=P$$
$$=\{-\alpha_1, \alpha_2, \alpha_3, -(\alpha_1+\alpha_2), -(\alpha_1+\alpha_3), -(\alpha_1+\alpha_2+\alpha_3)\}$$

so that

$$\Sigma_+=\{-\lambda_1, \lambda_2, -(\lambda_1+\lambda_2), -(\lambda_1+2\lambda_2)\}.$$

$\mathcal{L}=\mathrm{so}(4,2)$ [$\sim$su(2,2)] has four standard parabolic subalgebras, namely the minimal parabolic subalgebra (found in Sec. IV of I), $\mathcal{L}$ itself, and two others which will now be constructed.

(i) Choose $\theta=\{\lambda_1\}$. Then $V\mathcal{A}_\theta$ has generator $h_{\alpha_1}+h_{\alpha_2}+h_{\alpha_3}$, so [on using Eq. (A2) of I] $\mathcal{A}_\theta$ has generator

$$i[e_{\alpha_1+\alpha_2+\alpha_3}+e_{-(\alpha_1+\alpha_2+\alpha_3)}]$$

As $Q_{\lambda_1}=h_{\alpha_1}$, Eq. (A2) of I also shows that $\mathcal{A}(\theta)$ has generator $i(e_{\alpha_1}+e_{-\alpha_1})$.

Moreover, as $\langle\theta\rangle_-=\{\lambda_1\}$, $\tilde{\mathcal{N}}_-(\theta)$ is generated by

$$V^{-1}e_{\alpha_1}=\tfrac{1}{2}(e_{\alpha_1}-e_{-\alpha_1})-i2^{1/2}h_{\alpha_1}$$

[where Eq. (A3) of I has been used]. As $\langle\theta\rangle_+=\{-\lambda_1\}$ and $\{\Sigma_+-\langle\theta\rangle_+\}=\{\lambda_2, -(\lambda_1+\lambda_2), -(\lambda_1+2\lambda_2)\}$, $\tilde{\mathcal{N}}_+(\theta)$ is generated by

$$V^{-1}e_{-\alpha_1}=-\tfrac{1}{2}(e_{\alpha_1}-e_{-\alpha_1})-i2^{1/2}h_{\alpha_1}, \tag{10}$$

and $\tilde{\mathcal{N}}_\theta$ is generated by

$$V^{-1}e_{-(\alpha_1+\alpha_2+\alpha_3)}=-\tfrac{1}{2}(e_{\alpha_1+\alpha_2+\alpha_3}-e_{-(\alpha_1+\alpha_2+\alpha_3)})-i2^{1/2}(h_{\alpha_1}+h_{\alpha_2}+h_{\alpha_3}), \tag{11}$$

together with

$$V^{-1}e_{\alpha_2}$$
$$=\tfrac{1}{2}[e_{\alpha_2}-i\{\mathrm{sgn}\,N_{\alpha_1+\alpha_2+\alpha_3,-(\alpha_1+\alpha_3)}\}e_{-(\alpha_1+\alpha_3)}$$
$$-i\{\mathrm{sgn}\,N_{\alpha_1,\alpha_2}\}e_{\alpha_1+\alpha_2}-\{\mathrm{sgn}(N_{\alpha_1,\alpha_2}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_3})\}e_{-\alpha_3}],$$
$$V^{-1}e_{-(\alpha_1+\alpha_2)}$$
$$=\tfrac{1}{2}[e_{-(\alpha_1+\alpha_2)}-i\{\mathrm{sgn}\,N_{\alpha_1+\alpha_2+\alpha_3,-(\alpha_1+\alpha_2)}\}e_{\alpha_3}$$
$$-i\{\mathrm{sgn}\,N_{\alpha_1,-(\alpha_1+\alpha_2)}\}e_{-\alpha_2}-\{\mathrm{sgn}(N_{\alpha_1,-(\alpha_1+\alpha_2)}$$
$$\times N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2})\}e_{\alpha_1+\alpha_3}],$$

and $V^{-1}e_{\alpha_3}$, $V^{-1}e_{-(\alpha_1+\alpha_3)}$, which are given by similar expressions with α_2 and α_3 interchanged. By virtue of the identities

$$N_{\alpha_1+\alpha_2+\alpha_3,-(\alpha_1+\alpha_3)}=-N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2},\quad N_{\alpha_1,-(\alpha_1+\alpha_2)}=-N_{\alpha_1,\alpha_2}$$

(and similar identities with α_2 and α_3 interchanged), and

$$\mathrm{sgn}\{N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_3}\}=\mathrm{sgn}\{N_{\alpha_1,\alpha_2}N_{\alpha_1,\alpha_3}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2}\},$$

it follows that the five basis elements of $\mathcal{N}_\theta$ may be taken to be (11) together with

$$\tfrac{1}{2}(e_{\alpha_2}+e_{-\alpha_2})-\tfrac{1}{2}i\{\mathrm{sgn}\,N_{\alpha_1,\alpha_2}\}(e_{\alpha_1+\alpha_2}+e_{-(\alpha_1+\alpha_2)}),$$
$$\tfrac{1}{2}i(e_{\alpha_2}-e_{-\alpha_2})+\tfrac{1}{2}\{\mathrm{sgn}\,N_{\alpha_1,\alpha_2}\}(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)}),$$

and two similar expressions, with α_2 replaced by α_3.

The basis elements of $\mathcal{M}_\theta$ may be taken to be $i(h_{\alpha_2}-h_{\alpha_3})$, $i(e_{\alpha_1}+e_{-\alpha_1})$, $(e_{\alpha_1}-e_{-\alpha_1})$, and ih_{α_1}. Clearly $i(h_{\alpha_2}-h_{\alpha_3})$, ih_{α_1}, and $i(e_{\alpha_1+\alpha_2+\alpha_3}+e_{-(\alpha_1+\alpha_2+\alpha_3)})$ are generators of a real Z-invariant Cartan subalgebra $\mathcal{H}_r'$ for which $\mathcal{H}_r'\cap\mathcal{P}=\mathcal{A}_\theta$, so this parabolic subalgebra $\mathcal{P}_\theta$ is cuspidal.

(ii) Choose $\theta=\{\lambda_2\}$. In this case $V\mathcal{A}_\theta$ has generator $\{h_{\alpha_1}+h_{\alpha_1+\alpha_2+\alpha_3}\}$, so that [by Eq. (A2) of I] $\mathcal{A}_\theta$ has generator $\{i(e_{\alpha_1}+e_{-\alpha_1})+i(e_{\alpha_1+\alpha_2+\alpha_3}+e_{-(\alpha_1+\alpha_2+\alpha_3)})\}$. As Q_{λ_2} is a linear combination of h_{α_2} and h_{α_3} and of H_1 and H_2 of (8) and (9), one may take

$$Q_{\lambda_2}=h_{\alpha_2}+h_{\alpha_3}\ (=\{h_{\alpha_1+\alpha_2+\alpha_3}-h_{\alpha_1}\}).$$

Consequently Eq. (A2) of I shows that the generator of $\mathcal{A}(\theta)$ may be taken to be

$$\{i(e_{\alpha_1+\alpha_2+\alpha_3}+e_{-(\alpha_1+\alpha_2+\alpha_3)})-i(e_{\alpha_1}+e_{-\alpha_1})\}. \tag{12}$$

As $\langle\theta\rangle_-=\{-\lambda_2\}$, Eqs. (A6) and (A7) of I show that $\tilde{\mathcal{N}}_-(\theta)$ is generated by $V^{-1}e_{-\alpha_2}=\tfrac{1}{2}[e_{-\alpha_2}+i\,\mathrm{sgn}\,N_{\alpha_1,\alpha_2}e_{-(\alpha_1+\alpha_2)}$
$-\mathrm{sgn}\{N_{\alpha_1,\alpha_3}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2}\}e_{\alpha_3}-i\,\mathrm{sgn}\,N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2}e_{\alpha_1+\alpha_3}]$

and $V^{-1}e_{-\alpha_3}$, which is given by a similar expression with α_2 and α_3 interchanged. Moreover, as $\langle\theta\rangle_+=\{\lambda_2\}$ and

$\{\Sigma_+ - \langle\theta\rangle_+\} = \{-\lambda_1, -(\lambda_1+\lambda_2), -(\lambda_1+2\lambda_2)\}$, $\tilde{N}_+(\theta)$ is generated by $V^{-1}e_{\alpha_2}$ and $V^{-1}e_{\alpha_3}$ while $\tilde{N}_\theta$ is generated by $V^{-1}e_{-\alpha_1}$, $V^{-1}e_{-(\alpha_1+\alpha_2)}$, $V^{-1}e_{-(\alpha_1+\alpha_3)}$, $V^{-1}e_{-(\alpha_1+\alpha_2+\alpha_3)}$, all of which are specified in (i) above. It follows that the four basis elements of N_θ may be taken to be (10) and (11) together with

$$[(e_{\alpha_2}+e_{-\alpha_2}) + \operatorname{sgn}\{N_{\alpha_1,\alpha_3}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2}\}(e_{\alpha_3}+e_{-\alpha_3} - i\operatorname{sgn}N_{\alpha_1,\alpha_2}(e_{\alpha_1+\alpha_2} + e_{-(\alpha_1+\alpha_2)}) - i\operatorname{sgn}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2} \times (e_{\alpha_1+\alpha_3}-e_{-(\alpha_1+\alpha_3)})],$$

and

$$[i(e_{\alpha_2}-e_{-\alpha_2}) - i\operatorname{sgn}\{N_{\alpha_1,\alpha_3}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2}\}(e_{\alpha_3}-e_{-\alpha_3}) + \operatorname{sgn}N_{\alpha_1,\alpha_2}(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)}) - \operatorname{sgn}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_2} \times (e_{\alpha_1+\alpha_3}-e_{-(\alpha_1+\alpha_3)})].$$

Similarly the generators of M_θ may be taken to be (12) and $i(h_{\alpha_2}-h_{\alpha_3})$, together with

$$(e_{\alpha_2}+e_{-\alpha_2}) - \operatorname{sgn}\{N_{\alpha_1,\alpha_2}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_3}\}(e_{\alpha_3}+e_{-\alpha_3}),$$
$$i(e_{\alpha_2}-e_{-\alpha_2}) + \operatorname{sgn}\{N_{\alpha_1,\alpha_2}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_3}\}i(e_{\alpha_3}-e_{-\alpha_3}),$$
$$(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)}) + \operatorname{sgn}\{N_{\alpha_1,\alpha_3}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_3}\} \times (e_{\alpha_1+\alpha_3}-e_{-(\alpha_1+\alpha_3)}),$$
$$i(e_{\alpha_1+\alpha_2}+e_{-(\alpha_1+\alpha_2)}) - \operatorname{sgn}\{N_{\alpha_1,\alpha_3}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_3}\}i \times (e_{\alpha_1+\alpha_3}+e_{-(\alpha_1+\alpha_3)}).$$

As there do not exist two mutually commuting linearly independent generators in $M_\theta \cap K$, this parabolic subalgebra P_θ is not cuspidal.

It is interesting to express these results in terms of the physical generators of the conformal group of Minkowski space—time whose Lie algebra is isomorphic to so(4, 2). As shown for example by Murai[15] and Klink,[16] the basis of this Lie algebra consists of six generators of homogeneous Lorentz transformations $L_{\mu\nu}$ ($\mu,\nu=1,2,3,4$, $\mu<\nu$), four generators of space–time translations P_μ ($\mu=1,2,3,4$), four acceleration generators A_μ ($\mu=1,2,3,4$), and a dilatation generator Δ. The correspondence between the generators of the canonical basis of so(4, 2) and these physical generators, obtained using a mapping derived previously (Cornwell[13]), is given in Table II. It then follows from I that the generators of the subalgebras of the Iwasawa decomposition may be taken to be:

$$A_I:\{\Delta, L_{14}\},\quad M_I:\{L_{23}\},$$
$$N_I:\{P_1, P_2, P_3, P_4, L_{13}+L_{34}, L_{12}+L_{24}\},$$

where the convention has been adopted that $\operatorname{sgn}\{N_{\alpha_1,\alpha_2}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_3}\}=-1$. (This Iwasawa decomposition essentially coincides with that derived by Klink[16] by another method.) Similarly the nonminimal nontrivial parabolic sugalgebras obtained above have the following physical generators:

(i) For $\theta=\{\lambda_1\}$,

$$A_\theta:\{\Delta+L_{14}\},\quad M_\theta:\{L_{23}, -\Delta+L_{14}, P_4+P_1, A_4-A_1\},$$
$$N_\theta:\{P_2, P_3, P_4-P_1, L_{13}+L_{34}, L_{12}+L_{24}\}.$$

(ii) For $\theta=\{\lambda_2\}$,

$$A_\theta:\{\Delta\},\quad M_\theta:\{L_{12}, L_{13}, L_{14}, L_{23}, L_{24}, L_{34}\},$$
$$N_\theta:\{P_1, P_2, P_3, P_4\},$$

(where it is interesting to note that $M_\theta \oplus N_\theta$ is the Poincaré subalgebra).

V. LANGLANDS DECOMPOSITIONS FOR PARABOLIC SUBALGEBRAS OF REAL LIE ALGEBRAS GENERATED BY CHIEF OUTER INVOLUTIVE AUTOMORPHISMS

The theory given in Sec. V A of I shows that, although the chief outer automorphism Z is *not* diagonal in the canonical basis for L_c, the basis of A_I may be taken to consist of only two types of element, namely:

(i) r_I elements of the form $H_j^I = i(e_\alpha+e_{-\alpha})$, $j=1,2,\ldots,r_I$, $\alpha\in R(A_I)$, and

(ii) (m_I-r_I) elements H_j^I, $j=r_I+1,\ldots,m_I$, that are each real linear combinations of h_α, $\alpha=\alpha_1,\alpha_2,\ldots,\alpha_l$.

Moreover in every such case $m_I>r_I$, so that the second set is nonempty. [Explicit expressions for the roots of $R(A_I)$ and the basis elements of type (ii) were given in I for all simple Lie algebras generated by outer involutive automorphisms.]

As shown in I, $\dim M_I=0$ for $L=\mathrm{sl}(l+1,R)$ and NE_6^4, while for $L=q_{(l+1)/2}$ (l odd), $\mathrm{so}(2l-1,1)$ ($l\geq 2$), NE_6^3, $\mathrm{sl}(n,C)$, $\mathrm{so}(n,C)$, and $\mathrm{sp}(n,C)$, $r_I=0$ and consequently the entire basis of H' lies in H. Thus only for $L=\mathrm{so}(2l-2p-1,2p+1)$, $1\leq p<\frac{1}{2}l, l\geq 3$, are $\dim M_I$ and r_I

TABLE II. Correspondence between the generators of the canonical basis of so(4, 2) and the "physical" generators of the conformal group of Minkowski space—time.

ih_{α_1}	$\frac{1}{8}2^{-1/2}(P_1+P_4-A_1+A_4)$
ih_{α_2}	$\frac{1}{8}2^{-1/2}(-P_1+A_1)+\frac{1}{8}L_{23}$
ih_{α_3}	$\frac{1}{8}2^{-1/2}(-P_1+A_1)+\frac{1}{8}L_{23}$
$i(e_{\alpha_1}+e_{-\alpha_1})$	$\frac{1}{2}2^{-1/2}(\Delta-L_{14})$
$(e_{\alpha_1}-e_{-\alpha_1})$	$\frac{1}{4}(P_1+P_4+A_1-A_4)$
$(e_{\alpha_2}+e_{-\alpha_2})$	$\frac{1}{4}(-P_2+A_2)+\frac{1}{4}2^{-1/2}L_{13}$
$i(e_{\alpha_2}-e_{-\alpha_2})$	$\frac{1}{4}(P_3-A_3)+\frac{1}{4}2^{-1/2}L_{12}$
$(e_{\alpha_3}+e_{-\alpha_3})$	$\frac{1}{4}(P_2-A_2)+\frac{1}{4}2^{-1/2}L_{13}$
$i(e_{\alpha_3}-e_{-\alpha_3})$	$\frac{1}{4}(P_3-A_3)-\frac{1}{4}2^{-1/2}L_{12}$
$i(e_{\alpha_1+\alpha_2}+e_{-(\alpha_1+\alpha_2)})\operatorname{sgn}N_{\alpha_1,\alpha_2}$	$\frac{1}{4}(-P_2-A_2)-\frac{1}{4}2^{-1/2}L_{34}$
$(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)})\operatorname{sgn}N_{\alpha_1,\alpha_2}$	$\frac{1}{4}(P_3+A_3)+\frac{1}{4}2^{-1/2}L_{24}$
$i(e_{\alpha_1+\alpha_3}+e_{-(\alpha_1+\alpha_3)})\operatorname{sgn}N_{\alpha_1,\alpha_3}$	$\frac{1}{4}(-P_2-A_2)-\frac{1}{4}2^{-1/2}L_{34}$
$(e_{\alpha_1+\alpha_3}-e_{-(\alpha_1+\alpha_3)})\operatorname{sgn}N_{\alpha_1,\alpha_3}$	$\frac{1}{4}(P_3+A_3)-\frac{1}{4}2^{-1/2}L_{24}$
$i(e_{\alpha_1+\alpha_2+\alpha_3}+e_{-(\alpha_1+\alpha_2+\alpha_3)}) \times\operatorname{sgn}\{N_{\alpha_1,\alpha_2}N_{\alpha_1,\alpha_2\ \alpha_3,-\alpha_3}\}$	$\frac{1}{2}2^{-1/2}(-\Delta-L_{14})$
$(e_{\alpha_1+\alpha_2+\alpha_3}-e_{-(\alpha_1+\alpha_2+\alpha_3)}) \times\operatorname{sgn}\{N_{\alpha_1,\alpha_2}N_{\alpha_1+\alpha_2+\alpha_3,-\alpha_3}\}$	$\frac{1}{4}(P_1-P_4+A_1+A_4)$

$-2p-1, 2p+1)$, $1 \le p < \frac{1}{2}l, l \ge 3$, are $\dim \mathcal{M}_I$ and r_I both nonzero. For such an $\mathcal{L}$ a choice was considered in I in which the maximal Abelian subalgebra of $\mathcal{M}_I$ did not lie entirely in $\mathcal{H} \cap \mathcal{L}$. However it is useful to note that in this case as well it is possible to make a more convenient choice in which this subalgebra has basis elements ih_α, where $\alpha = \alpha_{p+1}, \alpha_{p+2}, \ldots, \alpha_{l-p-2}$, and $i(h_{\alpha_{l-p-1}} + \cdots + h_{\alpha_{l-2}} + \frac{1}{2}h_{\alpha_{l-1}} + \frac{1}{2}h_{\alpha_l})$, all of which lie in $\mathcal{H} \cap \mathcal{L}$. (Here $\mathcal{A}_I$ is taken as in I.)

For *any* real Cartan subalgebra $\mathcal{H}_r'$ the maximal value of $\dim(\mathcal{H}_r' \cap \mathcal{K})$ is equal to the rank of $\mathcal{K}$, which itself is equal to $(l - m_I + r_I)$. As the minimal value of $\dim(\mathcal{H}_r' \cap \mathcal{K})$ is $(l - m_I)$ (attained for a "maximally-split" Cartan subalgebra), the only possible values of $\dim(\mathcal{H}_r' \cap \mathcal{K})$ are $l - m_I, l - m_I + 1, \ldots, l - m_I + r_I$.

Clearly when $r_I = 0$, as it is for $\mathcal{L} = q_{(l+1)/2}$ (l odd), $so(2l-1, 1)$ ($l \ge 2$), NF_6^3, $sl(n, C)$, $so(n, C)$, and $sp(n, C)$, the only possible value of $\dim(\mathcal{H}_r' \cap \mathcal{K})$ is $(l - m_I)$, and $\mathcal{H}_r'$, being "maximally split," is unique up to conjugacy. Thus for each of these cases the only *cuspidal* parabolic subalgebra is the minimal parabolic subalgebra.

If $\mathcal{R}(\mathcal{A}_I)$ is nonempty the automorphism V is given by (5), where the product is again over the roots of $\mathcal{R}(\mathcal{A}_I)$ alone. When $\mathcal{R}(\mathcal{A}_I)$ is empty, V may be taken to be the identity.

APPENDIX: $\mathcal{K}, \mathcal{P}, \mathcal{A}_I$ AND THE MAXIMAL ABELIAN SUBALGEBRA OF $\mathcal{M}_I$ FOR THE CLASSICAL NONCOMPACT SIMPLE LIE ALGEBRAS GENERATED BY CHIEF INNER INVOLUTIVE AUTOMORPHISMS

As noted in Sec. IV A, the basis of $\mathcal{K}$ consists of ih_α for $\alpha = \alpha_j, j = 1, 2, \ldots, l$, together with $(e_\alpha + e_{-\alpha})$ and $i(e_\alpha - e_{-\alpha})$ for all α such that $\exp\{\alpha(h)\} = 1$, while the basis of $\mathcal{P}$ consists of $i(e_\alpha + e_{-\alpha})$ and $(e_\alpha - e_{-\alpha})$ for all α such that $\exp\{\alpha(h)\} = -1$. Accordingly for the specification of $\mathcal{K}$ and $\mathcal{P}$ it is sufficient to list the values of $\exp\{\alpha(h)\}$, where the chief inner automorphism $Z = \exp(\mathrm{ad}h)$ is that employed earlier (Cornwell[11]). The standard recursive procedure for finding the positive roots from the simple roots using the Cartan matrix (as described for example by Jacobson[17]) does not suggest a regular pattern for the positive roots. The underlying regularity only becomes obvious when some further linear functionals are introduced. (The expression $[a]$ will always denote the largest integer not greater than a.)

1. $\tilde{\mathcal{L}} = A_l$, $\mathcal{L} = su(l+1-p, p)$, $p = 1, 2, \ldots, [\frac{1}{2}(l+1)]$, $l \ge 1$.

Here there exist $(l+1)$ linear functionals $\epsilon_j(h)$, $j = 1, 2, \ldots, l+1$, $h \in \mathcal{H}$, such that the positive roots are $\epsilon_j(h) - \epsilon_k(h)$, $1 \le j < k \le l+1$, and the simple roots are $\alpha_j(h) = \epsilon_j(h) - \epsilon_{j+1}(h)$, $j = 1, 2, \ldots, l$ (cf. Gantmacher[6]). For the present purpose $\epsilon_{l+1}(h)$ can be taken to be identically zero so that

$$\epsilon_j(h) = \sum_{k=j}^{l} \alpha_k(h).$$

With the choice of $Z = \exp(\mathrm{ad}h)$ given earlier (Cornwell[11])

$$\exp\{\epsilon_j(h) - \epsilon_k(h)\} = \begin{cases} -1, & 1 \le j \le p < k, \\ +1, & \text{otherwise.} \end{cases}$$

Then $r_I = m_I = p$ and a convenient choice of $\mathcal{R}(\mathcal{A}_I)$ is

$$\mathcal{R}(\mathcal{A}_I) = \{\epsilon_j - \epsilon_{j+p}; j = 1, 2, \ldots, p\},$$

and correspondingly the basis of the maximal Abelian subalgebra of $\mathcal{M}_I$ may be taken to consist of the three sets:

(i) $ih_{\alpha_j} + ih_{\alpha_{j+p}}$, $j = 1, 2, \ldots, p-1$;

(ii) $\sum_{j=0}^{p-1} ih_{\alpha_{j+p}} + 2ih_{\alpha_{2p}}$;

(iii) ih_{α_j}, $j = 2p+1, \ldots, l$.

[For $p = \frac{1}{2}(l+1)$ only the set (i) appears, and for $p = \frac{1}{2}l$ only the sets (i) and (ii) appear.]

2. $\tilde{\mathcal{L}} = B_l$, $\mathcal{L} = so(2l+1-2p, 2p)$, $p = 1, 2, \ldots, l$, $l \ge 2$

In this case there exist[6] l linear functions $\epsilon_j(h)$, $j = 1, 2, \ldots, l$, $h \in \mathcal{H}$, such that the positive roots are $\epsilon_j(h)$, $j = 1, 2, \ldots, l$, and $\epsilon_j(h) \pm \epsilon_k(h)$, $1 \le j < k \le l$, the simple roots being $\alpha_j(h) = \epsilon_j(h) - \epsilon_{j+1}(h)$, $j = 1, 2, \ldots, l-1$, and $\alpha_l(h) = \epsilon_l(h)$, so that

$$\epsilon_j(h) = \sum_{k=j}^{l} \alpha_k(h).$$

With earlier choice (Cornwell[11]) of $Z = \exp(\mathrm{ad}h)$, for $l > p$,

$$\exp\{\epsilon_j(h)\} = \begin{cases} -1, & l - p < j \le l, \\ +1, & 1 \le j \le l - p, \end{cases}$$

and

$$\exp\{\epsilon_j(h) \pm \epsilon_k(h)\} = \begin{cases} -1, & 1 \le j \le l - p < k \le l, \\ +1, & \text{otherwise,} \end{cases}$$

whereas for $l = p$,

$$\exp\{\epsilon_j(h)\} = -1, \quad j = 1, 2, \ldots, l$$

and

$$\exp\{\epsilon_j(h) \pm \epsilon_k(h)\} = +1, \quad 1 \le j < k \le l.$$

There are three subcases to be considered:

(i) $2p > l > p$. In this subcase $r_I = m_I = 2l - 2p + 1$ and a convenient choice of $\mathcal{R}(\mathcal{A}_I)$ is the set $\epsilon_j \pm \epsilon_{l-p+j}$, $j = 1, 2, \ldots, l-p$, together with $\epsilon_{2l-2p+1}$. For $2p = l+1$, $\mathcal{M}_I$ is zero dimensional. For $2p > l+1$ the maximal Abelian subalgebra of $\mathcal{M}_I$ may be chosen to have basis ih_{α_j}, $j = 2l - 2p + 2, 2l - 2p + 3, \ldots, l$.

(ii) $2p \le l$. In this subcase $r_I = m_I = 2p$ and a convenient choice of $\mathcal{R}(\mathcal{A}_I)$ is the set $\epsilon_j \pm \epsilon_{2l-2p+1-j}$, $j = l - 2p + 1, l - 2p + 2, \ldots, l - p$. For $2p = l$, $\mathcal{M}_I$ is zero dimensional. For $2p \le l-1$ the basis of the maximal Abelian subalgebra of $\mathcal{M}_I$ may be taken to consist of the two sets:

(a) $i \sum_{j=2}^{2p} (j-1) h_{\alpha_{l-j}} + ih_{\alpha_{l-1}} + ih_{\alpha_l}$,

(b) ih_{α_j}, $j = 1, 2, \ldots, l - 2p - 1$, (which only appears for $2p < l - 1$).

(iii) $l = p$. In this subcase $r_I = m_I = 1$ and it is convenient to let ϵ_l be the root of $\mathcal{R}(\mathcal{A}_I)$. The basis of the maximal Abelian subalgebra of $\mathcal{M}_I$ may be taken to be ih_{α_j}, $j = 1, 2, \ldots, l-2$, together with $ih_{\alpha_{l-1}} + ih_{\alpha_l}$.

3. $\mathcal{L}=C_l,\ l\geqslant 2$

In this case there exist[6] l linear functionals $\epsilon_j(h)$, $j=1,2,\ldots,l$, $h\in\mathcal{H}$, such that the positive roots are $2\epsilon_j(h)$, $j=1,2,\ldots,l$, and $\epsilon_j(h)\pm\epsilon_k(h)$, $1\leqslant j<k\leqslant l$, the simple roots being $\alpha_j(h)=\epsilon_j(h)-\epsilon_{j+1}(h)$, $j=1,2,\ldots,l-1$, and $\alpha_l(h)=2\epsilon_l(h)$, so that

$$\epsilon_j(h)=\begin{cases}\sum_{k=j}^{l-1}\alpha_k(h)+\frac{1}{2}\alpha_l(h), & j=1,2,\ldots,l-1,\\ \frac{1}{2}\alpha_l(h) & , j=l.\end{cases}$$

(a) $\mathcal{L}=\mathrm{nsp}^{2p}_{2l}$, $p=1,2,\ldots,[\frac{1}{2}l]$

With the choice of $Z=\exp(\mathrm{ad}h)$ given in Ref. 11,

$$\exp\{\epsilon_j(h)\pm\epsilon_k(h)\}=\begin{cases}-1, & 1\leqslant j\leqslant p<k\leqslant l,\\ +1, & \text{otherwise},\end{cases}$$

and

$$\exp\{2\epsilon_j(h)\}=1,\quad j=1,2,\ldots,l.$$

Then $r_l=m_l=p$ and a convenient choice of $\mathcal{R}(\mathcal{A}_l)$ is $\epsilon_j-\epsilon_{j+p}$, $j=1,2,\ldots,p$. For $2p<l$ the basis of the maximal Abelian subalgebra of $\mathcal{M}_l$ may be taken to be the three sets:

(i) ih_{α_j}, $j=2p+1,\ldots,l$;

(ii) $i\sum_{j=0}^{p-1}h_{\alpha_{p+j}}+2ih_{\alpha_{2p}}$;

(iii) $ih_{\alpha_j}+ih_{\alpha_{p+j}}$, $j=1,2,\ldots,p-1$ (which only appears for $p>1$).

For $2p=l$ the basis may be taken to consist of the set (iii) together with

$$i\sum_{j=0}^{p}h_{\alpha_{p+j}}.$$

(b) $\mathcal{L}=\mathrm{sp}(l,\mathcal{R})$

With the choice of $Z=\exp(\mathrm{ad}h)$ given earlier,[11]

$$\exp[\epsilon_j(h)+\epsilon_k(h)]=1,\quad 1\leqslant j<k\leqslant l,$$
$$\exp[\epsilon_j(h)-\epsilon_k(h)]=-1,\quad 1\leqslant j<k\leqslant l,$$

and

$$\exp[2\epsilon_j(h)]=-1,\quad 1\leqslant j\leqslant l.$$

Then $r_l=m_l=l$ and a convenient choice of $\mathcal{R}(\mathcal{A}_l)$ is the set $2\epsilon_j$, $j=1,2,\ldots,l$. As $m_l=l$, $\mathcal{M}_l$ is zero dimensional.

4. $\tilde{\mathcal{L}}=D_l,\ l\geqslant 3$

In this case there exist l linear functionals $\epsilon_j(h)$, $j=1,2,\ldots,l$, $h\in\mathcal{H}$, such that the positive roots are $\epsilon_j(h)\pm\epsilon_k(h)$, $1\leqslant j<k\leqslant l$, the simple roots being $\alpha_j(h)=\epsilon_j(h)-\epsilon_{j+1}(h)$, $j=1,2,\ldots,l-1$, and $\alpha_l(h)=\epsilon_{l-1}(h)+\epsilon_l(h)$, so that

$$\epsilon_j(h)=\begin{cases}\sum_{k=j}^{l-2}\alpha_k(h)+\frac{1}{2}\alpha_{l-1}(h)+\frac{1}{2}\alpha_l(h), & j=1,2,\ldots,l-2,\\ \frac{1}{2}\alpha_{l-1}(h)+\frac{1}{2}\alpha_l(h), & j=l-1,\\ -\frac{1}{2}\alpha_{l-1}(h)+\frac{1}{2}\alpha_l(h), & j=l.\end{cases}$$

(a) $\mathcal{L}=\mathrm{so}(2l-2p,\ 2p)$, $p=1,2,\ldots,[\frac{1}{2}l]$

With the previous[11] choice of $Z=\exp(\mathrm{ad}h)$

$$\exp[\epsilon_j(h)\pm\epsilon_k(h)]=\begin{cases}-1, & 1\leqslant j\leqslant p<k\leqslant l,\\ +1, & \text{otherwise}.\end{cases}$$

Thus $r_l=m_l=2p$ and a convenient choice of $\mathcal{R}(\mathcal{A}_l)$ is $\epsilon_j\pm\epsilon_{p+j}$, $j=1,2,\ldots,p$. For $2p=l$, $\mathcal{M}_l$ is zero dimensional. For $2p=l-1$ the maximal Abelian subalgebra of $\mathcal{M}_l$ may be taken to have basis element $ih_{\alpha_{l-1}}-ih_{\alpha_l}$, while for $2p\leqslant l-2$ the basis may be taken to be ih_{α_j}, $j=2p+1, 2p+2,\ldots,l$.

(b) $\mathcal{L}=ND_{2l}$

With the previous[11] choice of $Z=\exp(\mathrm{ad}h)$

$$\exp[\epsilon_j(h)+\epsilon_k(h)]=-1,\quad 1\leqslant j<k\leqslant l,$$

and

$$\exp[\epsilon_j(h)-\epsilon_k(h)]=+1,\quad 1\leqslant j<k\leqslant l.$$

(i) l even: In this case $r_l=m_l=\frac{1}{2}l$, $\mathcal{R}(\mathcal{A}_l)$ may be chosen to be the set $\epsilon_{2j+1}+\epsilon_{2j+2}$, $j=0,1,\ldots,\frac{1}{2}l-1$, and the maximal Abelian subalgebra of $\mathcal{M}_l$ may be taken to have basis $ih_{\alpha_{2j+1}}$, $j=0,1,\ldots,\frac{1}{2}l-1$.

(ii) l odd: In this case $r_l=m_l=\frac{1}{2}(l-1)$, $\mathcal{R}(\mathcal{A}_l)$ may be chosen to be the set $\epsilon_{2j+1}+\epsilon_{2j+2}$, $j=0,1,\ldots,\frac{1}{2}(l-3)$, and the maximal Abelian subalgebra of $\mathcal{M}_l$ may be taken to have basis $ih_{\alpha_{2j+1}}$, $j=0,1,\ldots,\frac{1}{2}(l-3)$, together with $ih_{\alpha_{l-1}}-ih_{\alpha_l}$.

[1]J.F. Cornwell, J. Math. Phys. 16, 1992 (1975).
[2]R.L. Lipsman, *Group Representations*, Lecture Notes in Mathematics, Vol. 388 (Springer, Berlin, 1974).
[3]G. Warner, *Harmonic Analysis on Semi-Simple Lie Groups* (Springer, Berlin, 1972), Vol. I.
[4]M.E. Cartan, J. Math. Pure Appl. 8, 1 (1929).
[5]F. Gantmacher, Rec. Math. (Mat. Sbornik) N.S. 5, 217 (1939).
[6]F. Gantmacher, Rec. Math. (Mat. Sbornik) N.S. 5, 101 (1939).
[7]K. Iwasawa, Ann. Math. 50, 507 (1949).
[8]R.P. Langlands, *Problems in the theory of automorphic forms*, Lectures Notes in Mathematics, Vol. 170 (Springer, Berlin, 1970).
[9]M. Sugiura, J. Math. Soc. Jpn. 11, 374 (1959).
[10]G.W. Mackey, Bull. Am. Math. Soc. 69, 628 (1963).
[11]J.F. Cornwell, Rep. Math. Phys. 2, 239 (1971).
[12]F. Inönü and E.P. Wigner, Proc. Nat. Acad. Sci. 39, 510 (1953).
[13]J.F. Cornwell, Int. J. Theor. Phys. 12, 333 (1975).
[14]J.B. Ehrman, Proc. Camb. Philos. Soc. 53, 290 (1957).
[15]Y. Murai, Progr. Theor. Phys. 11, 441 (1954).
[16]W.H. Klink, J. Math. Phys. 10, 606 (1969).
[17]N. Jacobson, *Lie Algebras* (Interscience, New York, 1962).

Geometrical theory of contractions of groups and representations

Didier Arnal and Jean-Claude Cortet

Physique-Mathématique, Université de Dijon B.P. 138, 21004 Dijon, France
(Received 30 May 1978)

The contractions of Lie groups and Lie algebras and their representations are studied geometrically. We prove they can be defined by deformations in Poisson algebras of symplectic manifolds on which the groups act. These deformations are given by Dirac constraints which induce on C^∞ functions on the deformed manifold an associative twisted product, characterizing the contracted group or its representations. We treat the contractions of SO(n) to E(n) and apply this theory to thermodynamical limits in spin systems.

INTRODUCTION

Some mathematical theories define the contractions of Lie groups or algebras in the Lie algebra structure and connect these deformations to the geometrical action of the corresponding Lie group on suitable manifolds. Nevertheless, only representations of these Lie groups or algebras are relevant for physical applications and in this matter the mathematical theories are somewhat confused and unsatisfactory (passage from finite-dimensional representations to infinite-dimensional ones, carrying spaces not defined in all steps, $\cdots$), and in any case the geometrical interpretation does not exist.

A detailed study including physical applications of general deformations of the Poisson brackets and of the associative algebra of functions on symplectic manifolds (phase spaces) was recently done in Refs. 1 and 2. The representations of Lie groups are described in a natural way by means of twisted products (*-products) on symplectic manifolds. Therefore an inviting subject is to write contractions of the representations in terms of deformations of *-products so as to enlighten the role of the geometry of the phase space. We give some explicit deformations in Sec. 1 and develop in Sec. 2 the formalism of the *-products for the enveloping algebras of the Lie algebras of the orthogonal and Euclidean groups SO(n) and E(n) realized by functions on $\mathbb{R}^{2n}$ with the usual symplectic structure. The diagonalization of Casimir elements is performed in this formalism.

The Dirac bracket usually utilized for quantization of systems for which position and momentum are submitted to some constraints, can be expressed as a deformation of the original Poisson bracket and be considered as the Poisson bracket for the symplectic submanifold defined by the constraints (an example is given in Sec. 3). In particular the contraction of the Lie algebra so$(n+1)$ to $\mathcal{E}(n)$ can be defined via Dirac constraints on $\mathbb{R}^{2n+2}$: so$(n+1)$ (resp. $\mathcal{E}(n)$) is a Lie subalgebra of the Poisson algebra of C^∞-functions on $\mathbb{R}^{2n+2}$ (resp. $\mathbb{R}^{2n}$); the contraction map from so$(n+1)$ to $\mathcal{E}(n)$ is thus the Dirac restriction to $\mathbb{R}^{2n}$. Using *-products, we extend this procedure to the enveloping algebras, $\mathcal{U}(\text{so}(n+1))$ and $\mathcal{U}(\mathcal{E}(n))$. This result leads to an intrinsic and rigorous definition of contractions with a natural geometric interpretation on the phase space.

In order to reduce the obtained representation of E(n), we use in Sec. 5 the above method on the cotangent bundle of an "irreducible" manifold, i.e., a manifold on which the group acts transitively (in our case the cotangent bundle of the $n-1$ dimensional sphere S_{n-1}). We obtain in the case $n=3$ a complete equivalence with the usual Hilbert space realization.

On the other hand quantum mechanics appears naturally as a deformation of classical mechanics via the *-products. (The link between this formulation and the usual one in Hilbert spaces is provided by the Weyl application in the flat case). Observables are functions f, states are distributions ρ on the phase space and the expectation value of f in the state ρ is

$$\langle f \rangle = \int f * \rho \, dw,$$

where dw is the Liouville measure,[2] In quantum statistical mechanics the thermodynamic limit of Gibbs states can be performed as in the classical case, substituting *-products for usual product. For systems described in terms of the spin algebra, the usual computation of the thermodynamical quantities proceeds by calculating them for finite tensor-product power $\otimes^n M_2$ of the complex 2×2-matrices algebra M_2 and taking the limit $n\to\infty$. We proved in Ref. 3 that local Gibbs states ρ_n defined by a symmetrical Hamiltonian H_n are associated with unitary representations of U(2) and the decomposition of these representations in irreducible components corresponds to an ergodic decomposition of these states. The states ρ_n converge when $n\to\infty$ if and only if the sequence of measures of the above decompositions converges and the thermodynamical limit procedure is nothing but a contraction of representations of U(2) to representations of E(2)$\times\mathbb{R}$ for which the decomposition into irreducible components gives the ergodic decomposition of the limit state in symmetric and phase invariant states on the spin algebra.

In Sec. 7 we treat this contraction in the formalism of Dirac brackets and twisted products on a suitable manifold and derive the previous results obtained in Hilbert space formalism.

1. TWISTED PRODUCTS ON SYMPLECTIC MANIFOLD

The definitions and notations are those of Refs. 1 and 2. Let W be a differentiable symplectic manifold of dimension $2n$ with symplectic 2-form F and $T(W)$ [resp. $T^*(W)$] the tangent (resp. cotangent) bundle of W.

 0022-2488/79/040556-08$01.00

Definition 1.1: By extension to tensor bundles of the vector bundles isomorphism μ between TW and T^*W given by $\mu(X) = -i(X)F$ (i: interior product), the Poisson bracket on $N = C^\infty(W)$ is defined by

$$P(u, v) \equiv \{u, v\} = i(\Lambda)(du \wedge dv),$$
$$\Lambda = \mu^{-1}(F), \quad u, v \in C^\infty(W, \mathbb{R}). \tag{1.1}$$

On a symplectic chart

$$U = (x_j, j = 1, 2, \ldots, 2n) = ((x_k = p_k, x_{n+k} = q_k), k = 1, 2, \ldots, n),$$

$$\{u, v\} = \sum_{k=1}^{n} \left(\frac{\partial u}{\partial p_k} \frac{\partial v}{\partial q_k} - \frac{\partial u}{\partial q_k} \frac{\partial v}{\partial p_k} \right).$$

Using a symplectic connection with covariant differentiation operator ∇, the order r (in each argument) bidifferential operators P^r are given by

$$P^r(u, v)\big|_U = \sum \Lambda^{i_1 j_1} \cdots \Lambda^{i_r j_r} \nabla_{i_1 \ldots i_r} u \nabla_{j_1 \ldots j_r} v.$$

Let $E(N, \lambda)$ denote the space of formal series in a parameter λ with coefficients in N.

Definition 1.2: A formal deformation of the Lie algebra N is a Lie algebra law

$$N \times N \ni (u, v) \to [u, v]_\lambda = \sum_{r=0}^{\infty} \lambda^r C^r(u, v) \in E(N, \lambda), \tag{1.2}$$

where the C^r are 2-cochains on N, $C^0(u, v) = P(u, v)$, satisfying formally the Jacobi identity.

Definition 1.3: A $*_\lambda$-product on N is defined by a bilinear map:

$$N \times N \ni (u, v) \to u *_\lambda v = \sum_{r=0}^{\infty} \lambda^r (r!)^{-1} Q^r(u, v) \in E(N, \lambda), \tag{1.3}$$

where:

(i) Q^r is a bidifferential operator on N of maximum order $r (r > 1)$ in each argument, null on the constants, such that the principal symbol of Q^r coincides with the principal symbol of P^r.

(ii) $Q^0(u, v) = u \cdot v$, $\quad Q^1(u, v) = P(u, v)$.

(iii) Q^r is symmetric (resp. skew-symmetric) in (u, v) if r is even (resp. odd).

(iv) $\sum_{r+s=t} (r!s!)^{-1} Q^r(Q^s(u, v), w) = \sum_{r+s=t} (r!s!)^{-1} \times Q^r(u, Q^s(v, w))$

$(r, s > 0, \ t = 1, 2, \cdots)$ (associativity condition).

If $W = \mathbb{R}^{2n}$ (with the usual symplectic structure), the only formal functions $\sum a_r P^r$ defining an associative (resp. Lie) algebra law on N are (up to a factor) given by

$$u *_\lambda v = \exp(i\lambda)(u, v) = u \cdot v + \sum_{r=1}^{\infty} (i\lambda)^r (r!)^{-1} P^r(u, v), \tag{1.4}$$

$$[u, v]_\lambda = (2i\lambda)^{-1}(u *_\lambda v - v *_\lambda u)$$
$$= \lambda^{-1} \sin(\lambda P)(u, v)$$
$$= \sum_{r=0}^{\infty} (-1)^r \lambda^r [(2r+1)!]^{-1} P^{(2r+1)}(u, v). \tag{1.5}$$

In this case, the relevance to quantum mechanics comes from the fact that (1.5) for $\lambda = \hbar/2$ is the Moyal bracket corresponding to the commutator $(i\hbar)^{-1}[\Omega(u), \Omega(v)]$ of the operators $\Omega(u)$, $\Omega(v)$ defined by the Weyl quantization procedure:

$$\Omega : u(p_i, q_i) \to \Omega(u) = \int \tilde{a}(\xi^\alpha, \eta^\alpha) \exp[i(\xi^\alpha P_\alpha + \eta^\alpha Q_\alpha)] \, d\xi \, d\eta, \tag{1.6}$$

where $\tilde{a}$ is the inverse Fourier transform of the function or distribution a and P_α and Q_α the usual operators in the Heisenberg representation. We denote the twisted product (1.4) with $\lambda = \hbar/2$ by $*$ (without index) and refer to it in so far as the Moyal $*$-product on $\mathbb{R}^{2n}$.

A procedure for constructing a $*$-product for cotangent bundles of classical groups is given in Ref. 4. For instance, let $W = T^*(S_n)$ be the cotangent bundle to the sphere of dimension n imbedded in $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ by the two constraints:

$$|p|^2 = \sum_{i=1}^{n+1} p_i^2 = R^2$$

and

$$p \cdot q = \sum_{i=1}^{n+1} p_i q_i = 0, \quad (p, q) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}. \tag{1.7}$$

Consider the group G defined by the symplectic action on $E = (\mathbb{R}^{n+1} - \{0\}) \times \mathbb{R}^{n+1}$ by

$$(p, q) \to (\rho p, \rho^{-1} q + \sigma p), \quad \rho > 0, \quad \sigma \in \mathbb{R}. \tag{1.8}$$

The space of the orbits of E by G is isomorphic to $T^*(S_n)$. Hence an isomorphism between $N(T^*(S_n))$ and the space $\hat{N}(E)$ of C^∞ functions on E, invariant under the group G is defined by

$$\hat{u}(p, q) = u\left(\frac{p}{|p|}, |p| q - \frac{pq}{|p|} p\right), \quad (p, q) \in E, \quad u \in N(T^*(S_n)). \tag{1.9}$$

A $*$-product for $T^*(S_n)$ is then defined by (1.4) with

$$Q^r(u, v) = P^r(\hat{u}, \hat{v})\big|_W \quad \text{and} \quad \lambda = \hbar/2. \tag{1.10}$$

The symplectic structure of E gives by quotient the canonical symplectic structure of $T^*(S_n)$ and one has

$$\hat{P}(\hat{u}, \hat{v})\big|_W = P(u, v).$$

2. LIE AND *-SUBALGEBRAS OF $N(\mathbb{R}^n \times \mathbb{R}^n)$

Consider the following functions in $N(\mathbb{R}^n \times \mathbb{R}^n)$:

$$M_{jk} = q_j p_k - q_k p_j \quad (1 \le j < k \le n), \tag{2.1}$$

$$K_i = c p_i \quad (1 \le i \le n, \ c \in \mathbb{R}). \tag{2.2}$$

The commutation relations (for Poisson or Moyal bracket) are

$$\{M_{jk}, M_{il}\} = \delta_{ki} M_{jl},$$
$$\{K_i, K_j\} = 0, \tag{2.3}$$
$$\{M_{jk}, K_i\} = \delta_{ki} K_j - \delta_{ij} K_k.$$

We have a relization of the Lie algebra $\mathfrak{A} = \mathrm{so}(n)$ (resp. $\mathfrak{A} = \mathcal{E}(n)$) with generators M_{jk}, $1 \le j < k \le n$ (resp. M_{jk}, K_i, $1 \le j < k \le n$, $1 \le i \le n$). It is easily shown that every (Moyal) $*$-polynomial in the elements M_{jk} and K_i is equal to a usual polynomial in the elements M_{jk} and K_i. Moreover, the Moyal $*$-product is invariant with respect to $\mathfrak{A}$, i.e.,[2]

$$\left.\begin{array}{l}\lambda * u = u * \lambda = \lambda u \\ a * b - b * a = i\hbar\{a, b\} \\ \{a, u * v\} = \{a, u\} * v + u * \{a, v\}\end{array}\right\} \begin{array}{l}\lambda \in \mathbb{C}, \\ u, v \in N, \\ a, b \in \mathfrak{A}.\end{array} \qquad (2.4)$$

Proposition 2.1: The enveloping algebra $\mathcal{U}(\mathfrak{A})$ of the Lie algebra $\mathfrak{A}$ is isomorphic to a subalgebra of the algebra $(\mathbb{C}[p,q], *)$ of the polynomials in p_i, q_j $(i,j = 1, 2, \ldots, n)$ with the Moyal $*$-product.

Proof: The above realisation of $\mathfrak{A}$ is an α-morphism in the sense of Ref. 5. Therefore, there exists a morphism Φ from $\mathcal{U}(\mathfrak{A})$ into $(\mathbb{C})\,[p,q], *)$. If $u \in \mathcal{U}(\mathfrak{A})$, $\Phi(u)$ is a $*$-polynomial P^* (and hence a polynomial P) in the elements M_{jk} and p_i. P^* and P have the same homogeneous term Q of highest degree:

If we denote by X_k $(k = 1, 2, \ldots, m)$ a basis of $\mathfrak{A}$ and $\tilde{X}_k = \Phi(X_k)$, one has

$$u = \sum \lambda_{i_1, i_2, \ldots, i_m} X_1^{i_1} X_2^{i_2} \cdots X_m^{i_m},$$

$$\Phi(u) = P^* = \sum \lambda_{i_1, i_2, \ldots, i_m} \tilde{X}_1^{*i_1} * \tilde{X}_2^{*i_2} * \cdots * \tilde{X}_m^{*i_m}$$

$$= P = \sum \lambda_{i_1, i_2, \ldots, i_m} \tilde{X}_1^{i_1} \tilde{X}_2^{i_2} \cdots \tilde{X}_m^{i_m} + R \quad (d^0 R < d^0 P).$$

Considering P as a polynomial in p_i's with coefficients in $\mathbb{C}[q]$, its homogeneous part of highest degree is $Q(p,q)$.

If $\Phi(u) \equiv 0$, $Q(p,q) \equiv 0$ and then $Q(X_k) \equiv 0$ and $u = 0$.

The Weyl transformation (1.6) is defined on $\mathbb{C}[p,q]$ and its restriction to $\Phi(\mathcal{U}(\mathfrak{A}))$ defines a representation of $\mathfrak{A}$ in $L^2(\mathbb{R}^n)$ which is integrable to the quasiregular representation of the corresponding Lie group SO(n) or E(n). The decomposition of this representation in unitary irreducible components is usually performed via the diagonalization of the Casimir elements of $\mathcal{U}(\mathfrak{A})$. This diagonalization can be done directly in $(N, *)$.[2]

Definition 2.2: We call $*$-exponential the function defined on $\mathcal{U}(\mathfrak{A})$ by the formal series

$$\overset{*}{\mathrm{Exp}}(a) = \sum_{k=0}^{\infty} (k!)^{-1} (i\hbar)^{-k} (a*)^k. \qquad (2.5)$$

Suppose that $\overset{*}{\mathrm{Exp}}(at)$ converges for $|t| < \rho$ and that $\overset{*}{\mathrm{Exp}}(at)$ considered as a distribution over $\mathbb{R}^{2n}$ for fixed $t \in \mathbb{C}$ has a Fourier—Dirichlet expansion:

$$\overset{*}{\mathrm{Exp}}(at) = \sum_{j=0}^{\infty} \pi_j \exp(\lambda_j t / i\hbar).$$

This expansion is analogous to the spectral decomposition of an unitary operator in Hilbert space with discrete spectrum, thus we can define the spectrum of a.

Proposition 2.3[2] Let C be the first Casimir element in $\mathcal{U}(\mathrm{so}(n))$:

$$C = \sum_{1 \le j < k \le n} (M_{jk}^*)^2 = p^2 \circ q^2 - (p \circ q)^2 - n(n-1)\hbar^2/4$$

and

$$\gamma = (p^2 q^2 - (pq)^2)^{1/2} - (n-2)\hbar^2 / (4p^2 q^2 - (pq)^2)^{1/2}.$$

Then

$$\overset{*}{\mathrm{Exp}}(\gamma s) = \sum_{k=0}^{\infty} \pi_k \exp\{-is[k + \tfrac{1}{2}(n-2)]\} \quad \text{in } \mathcal{S}'(\mathbb{R}^{2n}) \quad \text{for } \mathrm{Im}(s) < 0.$$

The functions π_k satisfy

$$\pi_k = \overline{\pi}_k, \quad \gamma * \pi_k = \pi_k * \gamma = [k + \tfrac{1}{2}(n-2)]\pi_k, \quad \pi_l * \pi_k = \delta_{l,k} \pi_k.$$

The spectrum of C is $\{k(k+n-2)\hbar^2, \ k \in \mathbb{N}\}$.

Proposition 2.4: Let C be the Casimir element $\sum_{1 \le i \le n} (K_i^*)^2$ of $\mathcal{U}(\mathcal{E}(n))$. Then

$$\overset{*}{\mathrm{Exp}}(-tC) = \int_0^{+\infty} \exp(-t\lambda)\, \delta[\lambda - C]\, d\lambda \quad \text{for } \mathrm{Re}(t) > 0. \qquad (2.6)$$

The spectral resolution of C is given by the Heaviside functions

$$Y_\lambda(p) = 1 \quad \text{if } c^2 p^2 \le \lambda$$
$$= 0 \quad \text{if } c^2 p^2 > \lambda.$$

The functions Y_λ satisfy

$$\overline{Y}_\lambda = Y_\lambda, \quad Y_0 = 0, \quad \lim_{\lambda \to \infty} Y_\lambda = 1, \quad Y_\lambda * Y_\mu = Y_{\inf(\lambda, \mu)}.$$

Proof: $C = \sum_{1 \le i \le n} (K_i^*)^2 = \sum_{1 \le i \le n} c^2 p_i^2$ and (2.6) is the Laplace transform in $\mathcal{D}'^+$ of the distribution $\delta[\lambda - C] \in \mathcal{D}'^+(\lambda)$. If f is a Borel function of the variable $p \in \mathbb{R}^n$ we define by $\Omega(f) = f(P_i)(P_i = i\, d/dq_i)$ an operator in in $L^2(\mathbb{R}^n)$. This extension of the Weyl transform to the Borel functions in the variables p_i gives an extension of the Moyal $*$-product to these functions by the usual product. Denote by $\mathcal{F}$ the Fourier transform. If $\Phi \in \mathcal{S}(\mathbb{R}^n)$, we have

$$\frac{d}{d\lambda}[Y_\lambda(p)] = \delta(\lambda - C) \quad \text{and} \quad [\Omega(Y_\lambda)\Phi](q) = \mathcal{F}^{-1}[\mathcal{F}(\Phi) Y_\lambda](q),$$

$$\Omega(Y_\lambda)\Phi \in L^2(\mathbb{R}^n) \quad \text{and} \quad \|\Omega(Y_\lambda)\phi\|_{L^2} \le \|\phi\|_{L^2}.$$

Therefore, $\Omega(Y_\lambda)$ has an unique extension to a bounded operator E_λ in $L^2(\mathbb{R}^n)$. It is easily shown that $E_0 = 0$, $\text{w-}\lim_{\lambda \to +\infty}(E_\lambda) = 1$, $E_\lambda \cdot E_\mu = E_{\inf(\lambda,\mu)}$ (in fact $E_\lambda = \mathcal{F}^{-1} 1_{\{p^2 \le \lambda\}}$) and the corresponding results for Y_λ by inverse Weyl transform.

3. DIRAC BRACKETS AND DEFORMATIONS

The interpretation of the Dirac bracket as an instant of a formal deformation of the Poisson bracket is established in Ref. 6. We use these results in some special cases.

Let $k_i \in N$, $i = 1, 2$, two second-class constraints on W $[\det(\{k_i, k_j\}) \ne 0]$. The Dirac bracket relative to the constraints k_1, k_2 is defined by

$$[u, v] = \{u, v\} + C(u, v), \quad u, v \in N, \qquad (3.1)$$

with the 2-cochain $C(u,v)$ given by

$$C(u,v) = (\{u, k_1\}\{k_2, v\} - \{u, k_2\}\{k_1, v\})\{k_1, k_2\}^{-1}. \qquad (3.2)$$

If $\tilde{W}$ is a second-class submanifold of W, of codimension 2, $\tilde{W}$ is defined by two conjugate constraints. In this case,

$$[u, v]_\lambda = \{u, v\} + \lambda C(u, v), \quad \lambda \in [0, 1],$$

is a Lie algebra law on N.

For instance, if $W = \mathbb{R}^n \times \mathbb{R}^n$, with $k_1 = q_{n+1} + c$ $(c \in \mathbb{R})$ and $k_2 = p_{n+1}$

$$C(u,v) = \{u, v\}_{n+1} = \frac{\partial u}{\partial p_{n+1}} \frac{\partial v}{\partial q_{n+1}} - \frac{\partial u}{\partial q_{n+1}} \frac{\partial v}{\partial p_{n+1}} \qquad (3.3)$$

and the Dirac bracket (3.1) is the Poisson bracket relative to the submanifold $\mathbb{R}^{n-1}\times\mathbb{R}^{n-1}$ with coordinates $(p_i, q_i; i\leqslant n)$.

Lemma 3.1: On $W=T^*(S_n)$ with coordinates defined by (1.7), the Poisson bracket is given by

$$\{u,v\}=\sum_{i=1}^{n+1}\{u,v\}_{i,i}+R^{-2}\sum_{i,k=1}^{n+1}p_i[-p_k\{u,v\}_{k,i}+q_k\{u,v\}^{k,i}], \tag{3.4}$$

where $\{u,v\}_{i,j}=\dfrac{\partial u}{\partial p_i}\dfrac{\partial v}{\partial q_j}-\dfrac{\partial u}{\partial q_j}\dfrac{\partial v}{\partial p_i}$,

$$\{u,v\}^{i,j}=\frac{\partial u}{\partial q_i}\frac{\partial v}{\partial q_j}-\frac{\partial u}{\partial q_j}\frac{\partial v}{\partial q_i}, \qquad u,v\in N.$$

Proof: If $u\in N$,

$$du=\sum_{i=1}^{n+1}\left(\frac{\partial u}{\partial p_i}dp_i+\frac{\partial u}{\partial q_i}dq_i\right),$$

$$\operatorname{grad}(u)=\sum_{i=1}^{n+1}\left(\lambda_i\frac{\partial}{\partial p_i}+\mu_i\frac{\partial}{\partial q_i}\right)$$

is determined by the equation

$$i(\Lambda)(X)(\operatorname{grad}u)=du(X),\ X\in \text{tangent space to } T^*(S_n). \tag{3.5}$$

In a chart (p_i,q_i) such that $p_{n+1}\neq 0$, (3.5) is equivalent to

$$\left.\begin{aligned} &p_{n+1}\frac{\partial u}{\partial p_i}-p_i\frac{\partial u}{\partial p_{n+1}}=p_{n+1}\mu_i-p_i\mu_{n+1}\\ &p_{n+1}\frac{\partial u}{\partial q_i}-p_i\frac{\partial u}{\partial q_{n+1}}=-p_{n+1}\lambda_i+p_i\lambda_{n+1}\\ &0=\sum_{j=1}^{n+1}(p_j\mu_j-q_j\lambda_j)=\sum_{j=1}^{n+1}p_j\lambda_j \end{aligned}\right\}, \quad i=1,2,\ldots,n.$$

The solutions are

$$\mu_i=R^{-2}\left[\frac{\partial u}{\partial p_i}(R^2-p_i^2)-p_i\sum_{j\neq i}\frac{\partial u}{\partial p_j}p_j+\sum_j\frac{\partial u}{\partial q_j}(p_jq_i-p_iq_j)\right.$$

$$\lambda_i=R^{-2}\left[p_i\sum_{j\neq i}\frac{\partial u}{\partial q_j}p_j-\frac{\partial u}{\partial q_i}(R^2-p_i^2)\right], \quad i=1,2,\ldots,n+1$$

and $\{u,v\}=i(\Lambda)\ (\operatorname{grad}u)(\operatorname{grad}v)$ gives (3.4)

On the submanifold of $T^*(S_n)$ defined by $p_{n+1}^2\neq R^2$, consider the conjugate constraints $k_1=q_{n+1}+c$ *and* $k_2=p_{n+1}$. Then the 2-cochain (3.2) is

$$C(u,v)=-\{u,v\}_{n+1,n+1}+R^{-2}\sum_{i,j=1}^{n+1}p_i[p_j\{u,v\}_{i,j}+q_j\{u,v\}^{i,j}]$$

$$-R'^{-2}\sum_{i,j=1}^{n}p_i[p_j\{u,v\}_{i,j}+q_j\{u,v\}^{i,j}] \tag{3.6}$$

and the Dirac bracket (3.1) is

$$[u,v]=\sum_{i=1}^{n}\{u,v\}_{i,i}+R'^{-2}\sum_{i,k=1}^{n}p_i[-p_k\{u,v\}_{k,i}+q_k\{u,v\}^{k,i}], \tag{3.7}$$

where $R'^2=\sum_{i=1}^n p_i^2$. Denote by $\bar u$, the restriction of $u\in N(T^*(S_n))$ to the manifold $T^*(S_{n-1})$ defined by the constraints k_1,k_2, we have

$$\overline{[u,v]}=\{\bar u,\bar v\}.$$

Proposition 3.2: The Poisson bracket on $T^*(S_{n-1})$ is the restriction of the Dirac bracket on $T^*(S_n)$.

4. CONTRACTION OF so(n +1) TO $\mathcal{E}(n)$

A. Contraction of the Lie algebra so($n+1$) to $\mathcal{E}(n)$

Consider the realizations (2.1) and (2.2) of the Lie algebras so$(n+1)$ and $\mathcal{E}(n)$ with the following notations;

generators of so$(n+1)$ in $N(\mathbb{R}^{2n+2})$;

$$M_{j,k}=q_jp_k-q_kp_j,\quad 1\leqslant j<k\leqslant n+1;$$

generators of $\mathcal{E}(n)$ in $N(\mathbb{R}^{2n})$;

$$\overline{M_{j,k}}=q_jp_k-q_kp_j,\quad 1\leqslant j<k\leqslant n,$$

$$\overline{K}_i=cp_i,\qquad 1\leqslant i\leqslant n,\ c\in\mathbb{R}.$$

The generators $\overline{M_{j,k}}$ and $\overline{K}_i$ are formally obtained from the generators $M_{j,k}$ of so$(n+1)$ by the constraints $k_1=q_{n+1}+c=0$, $k_2=p_{n+1}=0$. Since by (3.3) these constraints lead to a deformation of the Poisson bracket on $\mathbb{R}^{2n+2}$ to the Poisson bracket on $\mathbb{R}^{2n}$, we have

Proposition 4.1: The contraction of the Lie algebra so$(n+1)$ to $\mathcal{E}(n)$ is given by a formal deformation of the Poisson bracket on $\mathbb{R}^{2n+2}$ to the Dirac bracket relative to the constraints $k_1=q_{n+1}+c$ and $k_2=p_{n+1}$ $(c\in\mathbb{R})$.

B. Contraction of the enveloping algebra $\mathcal{U}$(so (n + 1)) to $\mathcal{U}$($\mathcal{E}$ (n))

Let φ_λ $(\lambda\in[0,1[)$ be the diffeomorphism of $\mathbb{R}^{2n+2}$ defined by

$$\varphi_\lambda(p_i)=p_i,\quad \varphi_\lambda(q_i)=q_i,\quad 1\leqslant i\leqslant n,$$

$$\varphi_\lambda(p_{n+1})=(1-\lambda)^{1/2}p_{n+1},$$

$$\varphi_\lambda(q_{n+1})=(1-\lambda)^{1/2}q_{n+1}-c.$$

We have $\{u\circ\varphi_\lambda, v\circ\varphi_\lambda\}(p,q)=[u,v]_\lambda\ (\varphi_\lambda(p,q))$ and $[u,v]_\lambda(p,q)=\{u\circ\varphi_\lambda, v\circ\varphi_\lambda\}(\varphi^{-1}(p,q))$, $\lambda\in[0,1[$. Therefore, a $*$-product $*_\lambda$ is defined on $\mathbb{R}^{2n+2}$ by

$$(u*_\lambda v)(p,q)=(u\circ\varphi_\lambda * v\circ\varphi_\lambda)(\varphi_\lambda^{-1}(p,q)),\quad u,v\in N(\mathbb{R}^{2n+2}). \tag{4.1}$$

If $u\in N(\mathbb{R}^{2n})$, let $\tilde u$ be its canonical extension on $\mathbb{R}^{2n+2}$ defined by

$$\tilde u(p_1,p_2,\ldots,p_{n+1},q_1,q_2,\ldots,q_{n+1})$$

$$=u(p_1,p_2,\ldots,p_n,q_1,q_2,\ldots,q_n).$$

$\lim_{\lambda\to1}(u*_\lambda v)(p,q)$ exists iff $(p,q)\in\varphi_1\mathbb{R}^{2n+2})=\mathbb{R}^{2n}$; hence $\lim_{\lambda\to1}u*_\lambda\tilde v=\tilde u*_1\tilde v$ exists on $\mathbb{C}[p,q]\subset N(\mathbb{R}^{2n})$. Since $[\tilde u,\tilde v]=\{u,\tilde v\}$, we have $u*_1v=u*v$ $u,v\in\mathbb{C}[p,q]$, $*$ is the Moyal $*$-product on $\mathbb{R}^{2n}$. This proves the following proposition:

Proposition 4.2: The map $u\to\tilde u|_{\mathbb{R}^{2n}}$ is a deformation of the enveloping algebra $\mathcal{U}(\text{so}(n+1),*)$ to $\mathcal{U}(\mathcal{E}(n),*)$ which extends the contraction map defined by Proposition 4.1.

Of course, the well-known contractions

$$\mathrm{SO}(p,q)\to \begin{cases}\mathrm{ISO}(p,q-1)\\ \mathrm{ISO}(p-1,q),\ p\neq q\end{cases}$$

can be treated exactly in a similar way.

5. CONTRACTIONS OF THE REPRESENTATIONS

The diagonalization (2.6) of the Casimir element $C=\sum_{1\le i\le n}(K_i^*)^2$ of $\mathcal{E}(n)$ reduces the representation (2.1) and (2.2) of $\mathcal{E}(n)$ in $N(\mathbb{R}^{2n})$ because it corresponds by the Weyl transform to its diagonalization in $L^2(\mathbb{R}^n)$ (Prop. 2.4) which reduces the quasiregular representation.

In order to obtain representations of $\mathcal{E}(n)$ with scalar Casimir element C, we have to consider the symplectic submanifold $T^*(S_{n-1})$ of $\mathbb{R}^{2n}$ on which the function $C=\sum_{1\le i\le n}p_i^2$ is constant.

In order to use the Dirac bracket and associated procedure of contraction, we introduce a $*$-representation of $\mathrm{so}(n+1)$ on $T^*(S_n)$ with the $*$-product defined by (1.10):

$$u*v=\hat{u}\hat{*}\hat{v}, \tag{5.1}$$

where $u,v\in N(T^*(S_n))$, $\hat{u}$ and $\hat{v}$ defined by (1.9), and $\hat{*}$ denote the Moyal $*$-product on $\mathbb{R}^{2n+2}$.

Consider the functions on $T^*(S_n)$:

$$N_{j,k}=q_jp_k-q_kp_j,\quad 1\le j<k\le n+1,$$
$$L_i=cp_i,\qquad 1\le i\le n+1.$$

Then,

$$\hat{N}_{j,k}=M_{j,k}\ \text{ and }\ \hat{L}_i=K_i/|p|.$$

Obviously, we have the proposition:

Proposition 5.1: The functions $N_{j,k}$ (resp. N_{jk} and L_i) are generators with Moyal or Poisson bracket of a representation of $\mathrm{so}(n+1)$ (resp. $\mathcal{E}(n+1)$) on $T^*(S_n)$. Moreover, the $*$-product (5.1) is invariant with respect to $\mathrm{so}(n+1)$, (2.4).

With respect to Dirac bracket (3.7), the following commutation relations hold:

$$\begin{aligned}
&[N_{j,n+1},N_{i,n+1}]=R'^{-2}p_{n+1}N_{j,i}\quad 1\le i<j\le n,\\
&[N_{j,k},N_{i,l}]=R^2R'^{-2}\delta_{ki}N_{j,l},\quad 1\le j<k\le n,\ 1\le i<l\le n,\\
&[N_{j,n+1},N_{i,l}]=0\qquad 1<i<l\le n\\
&\qquad\qquad\qquad\qquad 1\le i<l<j\le n,\\
&[N_{j,n+1},N_{j,l}]=0\qquad 1\le j<l\le n,\\
&[N_{j,n+1},N_{i,j}]=0\qquad 1\le i<j\le n.
\end{aligned}\tag{5.2}$$

Proposition 5.2: The restriction of the Dirac deformation from $T^*(S_n)$ to the submanifold $T^*(S_{n-1})$ defined by the constraints $k_1=q_{n+1}+c$ and $k_2=p_{n+1}$ realizes the contraction of the representation of $\mathrm{so}(n+1)$ on $T^*(S_n)$ to a representation of $\mathcal{E}(n)$ on $T^*(S_{n-1})$ with scalar Casimir element C.

Proof: Impose $p_{n+1}=0$ and $q_{n+1}=-c$ in the relations (5.2); we get $R'=R$ and, if $\overline{N}_{j,k}$ (resp. $\overline{K}_i$) are the restrictions of $N_{j,k}$ (resp. $N_{i,n+1}$) to $T^*(S_{n-1})$,

$$[\overline{N}_{jk},\overline{N}_{il}]=\delta_{ki}\overline{N}_{jl}=\{\overline{N}_{jk},\overline{N}_{il}\},\qquad \begin{cases}1\le j<k\le n,\\ 1\le i<l\le n,\end{cases}$$

$$[\overline{K}_i,\overline{K}_j]=0=\{\overline{K}_i,\overline{K}_j\}\quad (i,j=1,2,\dots,n),$$
$$[\overline{N}_{jk},\overline{K}_i]=\delta_{ik}\overline{K}_j-\delta_{ij}\overline{K}_k=\{\overline{N}_{jk},\overline{K}_i\},\ \begin{cases}i=1,2,\dots,n\\ 1\le j<k\le n.\end{cases}$$

We obtain a representation of $\mathcal{E}(n)$ on $T^*(S_{n-1})$ and, moreover, the Casimir element C takes the scalar value:

$$C=\sum_{p=1}^{n}\overline{K}_i^2=c^2R^2$$

[The same calculations hold for the contractions

$$\mathrm{so}(p,q)\to \begin{cases}\mathrm{ISO}(p-1,q)\\ \mathrm{ISO}(p,q-1)\end{cases}$$

on the suitable manifold.]

In the case $n=2$, we give the explicit form of the representation of $\mathcal{E}(2)$ by diagonalization of $J_3=\overline{N}_{12}=q_1p_2-q_2p_1$.

Proposition 5.3: Let $K^{\pm}=1/2(\overline{K}_1\mp i\overline{K}_2)$ if $c\ne 0$

$=1/2(p_1\mp ip_2)$ if $c=0$.

Then:

(i) $\overset{*}{\mathrm{Exp}}(it\,J_3)=\left(\cos\frac{\hbar}{2}t\right)^{-2}\exp\left(-\frac{2i}{\hbar}J_3\tan\frac{\hbar}{2}t\right)$,

(ii) $J_3=\sum_{k\in Z}k\hbar\pi_k$, where:

$$\pi_0=\frac{\hbar}{2J_3}\sinh\left(\frac{2}{\hbar}J_3\right),\quad \pi_k=(K^{+*})^k*\pi_0*(K^{-*})^k\quad (k>0),$$
$$\pi_k=(K^{-*})^{-k}*\pi_0*(K^{+*})^{-k}\quad (k<0),$$
$$J_3*\pi_k=\pi_k*J_3=k\pi_k,$$

π_k is the unique solution of the equation $J_3*f=f*J_3=kf$ in N.

$$\overline{\pi}_k=\pi_k\ \text{ and }\ \pi_k*\pi_l=\pi_l*\pi_k=\delta_{lk}\pi_l$$

Proof: With coordinates $z=J_3$ and $\theta\in[0,2\pi]$ defined by $p_1=R\cos\theta$ and $p_2=R\sin\theta$, we have

$$\begin{aligned}f*J_3&=\hat{f}\hat{*}\hat{z}_{|T^*(S_1)}\\
&=f\cdot z+\frac{i\hbar}{2}\{f,z\}_{T^*(S_1)}-\left(\frac{\hbar}{2}\right)^2\\
&\quad\times\left(\frac{\partial^2 f}{\partial q_1\partial p_2}-\frac{\partial^2 f}{\partial q_2\partial p_1}\right)_{\substack{|p|=R\\ p.q=0}},\end{aligned}$$

i.e.,

$$f(\theta,z)*z=f(\theta,z)\cdot z-\frac{i\hbar}{2}\frac{\partial f}{\partial\theta}-\left(\frac{\hbar}{2}\right)^2\left(z\frac{\partial^2 f}{\partial z^2}+2\frac{\partial f}{\partial z}\right)$$

The equations $J_3*f=f*J_3=0$ are

$$zf(z)-\left(\frac{\hbar}{2}\right)^2[zf''(z)+2f'(z)]=0=\frac{\partial f}{\partial\theta}. \tag{5.3}$$

Let ψ be the function

$$\psi(t,z)=\left(\cos\frac{\hbar}{2}t\right)^{-2}\exp\left(-\frac{2i}{\hbar}z\tan\frac{\hbar}{2}t\right).$$

Then,

$$i\frac{\partial\psi}{\partial t}=z\cdot\psi-\left(\frac{\hbar}{2}\right)^2\left(z\frac{\partial^2\psi}{\partial z^2}+2\frac{\partial\psi}{\partial z}\right)\ \text{ and }\ \psi(0,z)=1.$$

and therefore $\psi = \mathrm{Exp}(itJ_3)$ and (*i*). Moreover, ψ is the derivative of the distribution defined by the function

$$\varphi(t,z) = \exp\left[-\frac{2i}{\hbar} z \tan\left(\frac{\hbar}{2}\right)t\right].$$

φ has the period $2\pi/\hbar$ and belongs to $L^1(0, 2\pi/\hbar)$. φ (and thus ψ) admits a Fourier development (in the distribution sense). From there we deduce that the spectrum of $\mathrm{Exp}(tJ_3)$ is contained in $\{\exp(k\hbar t),\ k\in \mathbf{Z}\}$.

By the substitution $g(z) = zf(z)$, (5.3) becomes

$$-(\hbar/2)^2 g''(z) + g(z) = 0. \tag{5.4}$$

The solutions of (5.4) in N are $g(z) = \lambda \sinh[(2/\hbar)z]$.

Using the condition $\pi_0 * \pi_0 = \pi_0$, we find a unique solution

$$\pi_0 = \frac{\hbar}{2z} \sinh[(2/\hbar)z].$$

In order to describe the complete spectrum of J_3, we compute

$$J_3 * \bar{K}_i - \bar{K}_i * J_3 = i\hbar\{J_3, \bar{K}_i\},$$

and then

$$J_3 * K^+ - K^+ * J_3 = \hbar K^+,$$

$$J_3 * K^- - K^- * J_3 = -\hbar K^-.$$

For $u \in N(T^*(S_1))$ we define $T^+(u)$ and $T^-(u)$ by

$$T^+(u) = K^+ * u * K^- \quad\text{and}\quad T^-(u) = K^- * u * K^+.$$

We obtain

$$\left.\begin{aligned} J_3 * (T^+)^n(\pi_0) &= (T^+)^n(\pi_0) * J_3 = n\hbar (T^+)^n(\pi_0) \\ J_3 * (T^-)^n(\pi_0) &= (T^-)^n(\pi_0) * J_3 = -n\hbar (T^-)^n(\pi_0) \end{aligned}\right\},\quad n \in \mathbb{N}.$$

Then the spectrum of J_3 is exactly $\{k\hbar,\ k\in\mathbf{Z}\}$. Moreover,

$$T^+[T^-(u)] = T^-[T^+(u)] = \frac{R^4c^4}{16}u \quad \text{if } c \neq 0$$

$$= \frac{R^4}{16}u \quad \text{if } c = 0.$$

It follows that the unique solution of the equation:

$$J_3 * f = f * J_3 = nf \quad (\text{resp. } -nf) \quad (n \in \mathbb{N})$$

is $(T^+)^n(\pi_0)$ (resp. $(T^-)^n(\pi_0)$). The spectrum of J_3 is simple.

Lemma 5.4: In the set of the formal series in the variable z with coefficients in $C^\infty(S_1)$, the solutions of the equation

$$f(\theta, z) * z = 0,$$

are

$$f(\theta,z) = \varphi(\theta) * \pi_0(z), \quad \varphi \in C^\infty(S_1)$$

Proof: If $f(\theta,z) = \sum_{n=0}^\infty a_n(\theta) z^n\ [a_n \in C^\infty(S_1)]$, then

$$f(\theta,z) * \pi_0(z) = f(\theta,z) * \left(\frac{\hbar}{2}\right)\left(\frac{\sinh(2/\hbar)z}{z}\right)$$

$$= f(\theta,z) * \left[\sum_{n=0}^\infty \left(\frac{2}{\hbar}z\right)^{2n} \frac{1}{(2n+1)!}\right] = f(\theta,z).$$

Moreover,

$$f(\theta,z) * \pi_0(z) = \left(\sum_{n=0}^\infty a_n(\theta) z^n * \pi_0(z)\right) = a_0(\theta) * \pi_0(z)$$

and thus $f(\theta,z) = a_0(\theta) * \pi_0(z)$.

Let $\mathfrak{S}$ be the set of these formal series. It is easy to compute the action of the generators of $\mathcal{E}(2)$ on $\mathfrak{S}$:

Proposition 5.5: On $\mathfrak{S}$, the representation of $\mathcal{E}(2)$ defined by Proposition 5.2 is given by

$$J_3 * [\varphi(\theta) * \pi_0] = \left(i\hbar \frac{\partial\varphi}{\partial\theta}\right) * \pi_0,$$

$$K_1 * [\varphi(\theta) * \pi_0] = [cR\cos\theta\, \varphi(\theta)] * \pi_0,$$

$$K_2 * [\varphi(\theta) * \pi_0] = [cR\sin\theta\, \varphi(\theta)] * \pi_0,$$

$$C = (K_1^*)^2 + (K_2^*)^2 = c^2R^2.$$

We find the usual form of the irreducible representation of the Euclidean group E(2) with Casimir element equal to c^2R^2 in $L^2([0,2\pi])$ and obtained by the usual contraction of the representations of SO(3) (including the case $c=0$).

6. THERMODYNAMICAL LIMIT AND GROUP CONTRACTION

In the so-called "algebraic formulation" of quantum statistical mechanics, a class of lattice systems is described in terms of the spin algebra. For each $n \in \mathbb{N}$, let $\mathcal{A}_n$ (with identity I_n) be the nth tensor power of the algebra M_2 of complex 2×2 matrices (with identity I_2). Considering the canonical embedding $\varphi_{n,m}$ $(n<m)$ from $\mathcal{A}_n$ into $\mathcal{A}_m$ given by

$$\varphi_{n,m}(A) = A \otimes I_{m\setminus n}, \quad A \in \mathcal{A}_n.$$

The spin algebra $\mathcal{A}$ is defined by the inductive limit C*-algebra:

$$\mathcal{A} = \varinjlim\{\varphi_{n,m}(\mathcal{A}_n);\ m,n\in\mathbb{N},\ m>n\}.$$

We shall consider systems with symmetric local Hamiltonian H_n, i.e., for each $n\in\mathbb{N}$, H_n is an element in $\mathcal{A}_n$ which is invariant with respect to the automorphism group $\mathfrak{S}_n$ of $\mathcal{A}_n$ induced by permutations s of the indices $\{1,2,\ldots,n\}$:

$$s\left(\sum \bigotimes_{p=1}^n A_p\right) = \sum \bigotimes_{p=1}^n A_{s(p)}, \quad s\in\mathfrak{S}_n,\ A_p \in M_2.$$

The usual computations of the thermodynamical quantities on $\mathcal{A}$ proceed by calculating them for finite $n\in\mathbb{N}$ and taking the limit $n\to\infty$. In particular the global Gibbs state is given by the limit (if it exists) of the local Gibbs states ρ_n with density matrix $\exp(-\beta H_n)$.

Denote by σ^i $(i=1,2,3)$ the usual Pauli matrices; let

$$J_i = \frac{1}{n}\sum_{p=1}^n \left(\bigotimes_{k=1}^{p-1} I_2\right) \otimes \sigma^i \left(\bigotimes_{k=p+1}^n I_2\right), \quad i=1,2,3,$$

$$= \frac{1}{n}\sum_{p=1}^n \sigma_p^i. \tag{6.1}$$

Because H_n is symmetric, it can be written entirely in terms of the elements J_i and I_n.[7] Then the state ρ_n is determined by its values on the symmetric elements of $\mathcal{A}_n$ and appears as a continuous function of positive type on the group $\mathrm{U}(2) = T\cdot\mathrm{SU}(2)$ [semidirect product of the torus T by SU(2)]. Using this aspect, one deduces the following result:

Theorem 6.1: If H_n is an element of $\mathcal{A}_n$ such that

(i) H_n is invariant with respect to the permutation symmetry group $\mathfrak{S}_n$,

(ii) $[H_n, J_3]=0$,

and if the local Gibbs state ρ_n with density matrix $\exp(-\beta H_n)$ has a limit ρ when $n\to\infty$, then:

(1) ρ is invariant with respect to the group $\mathfrak{S}$ of finite permutations of $\mathbb{N}$.

(2) ρ is a continuous function of positive type on the group $E(2)\times\mathbb{R}$ (direct product of the two-dimensional Euclidean group by $\mathbb{R}$).

(3) The representation of $E(2)\times\mathbb{R}$ determined by ρ is obtained by contraction of the representation of U(2) determined by ρ_n.

This theorem is a particular case of the results given in Ref. 3, where the symmetric states with torus invariance are investigated. More precisely, the one-parameter group $\exp[i\frac{1}{2}\theta\,\mathrm{ad}(1-\sigma^3)]$ of automorphisms of M_2 defines a group τ_θ ($\theta\in[0,2\pi]$) of automorphisms of $\mathcal{A}$. We denote by $\mathcal{G}$ (resp. $\mathcal{G}_0$) the set of invariant (resp. ergodic) symmetric states on $\mathcal{A}$ (with respect to τ_θ), U_θ the corresponding group in the G.N.S. representation.

Proposition 6.2 [3]: (1) There exists a homeomorphism $\mathbf{v}\to\rho_{\mathbf{v}}$ between

$$B=\{\mathbf{v}=(\lambda^1,\lambda^2,\lambda^3)\in\mathbb{R}^3,\quad \|\mathbf{v}\|\le 1\}$$

and the set of extremal symmetric states on $\mathcal{A}$,

$(\rho_{\mathbf{v}}=\otimes\omega,\ \omega(\sigma^i)=\lambda^i)$.

(2) $\mathcal{G}_0$ is homeomorphic to the compact set Δ:

$$\Delta=\{(y,w)\in\mathbb{R}^2,\ 0\le|w|\le y\le 1\}.$$

(3) If $\rho\in\mathcal{G}$, there exists a unique probability measure ν_ρ on Δ such that

$$\rho=\int\rho_{yw}\,d\nu_\rho(y,w),\quad \rho_{yw}=(2\pi)^{-1}\int_0^{2\pi}\rho_{\mathbf{v}}d\theta,$$

where $\mathbf{v}=((y^2-w^2)^{1/2}\cos\theta,\ (y^2-w^2)^{1/2}\sin\theta, w)$.

Let $\mathfrak{A}=(K_0,K_1,K_2,K_3)$ be the Lie algebra of $E(2)\times\mathbb{R}$ with the following commutation relations:

$$[K_0,K_1]=K_2,\quad [K_0,K_2]=-K_1,\ [K_1,K_2]=0,\ [K_3,K_j]=0,\quad j=1,2,3. \tag{6.2}$$

Proposition 6.3 [3]: Let $\rho\in\mathcal{G}$. Then

(1) In the GNS representation $(\mathfrak{H}_\rho,\pi_\rho,\Omega_\rho)$ associated with ρ,

$$dV_\rho(K_0)=\frac{d}{d\theta}(U_\theta),$$

$$dV_\rho(K_j)=\operatorname*{s-lim}_{n\to\infty} n^{-1}\sum_{p=1}^{n}\pi_\rho(i\sigma_p^j),\quad j=1,2,3,$$

defines a unitary representation V_ρ of $E(2)\times\mathbb{R}$ in $\mathfrak{H}_\rho$.

(2) $V_\rho=\aleph_0 W_\rho$, where:

(i) W_ρ is a unitary representation of $E(2)\times\mathbb{R}$ without multiplicity,

(ii) $-dW_\rho(K_1^2+K_2^2+K_3^2)\le 1$,

(iii) $dW_\rho(K_1^2+K_2^2)|\Omega\rangle=0 \Rightarrow dW_\rho(K_0)|\Omega\rangle=0$.

(3) $\rho\to W_\rho$ is a continuous bijective map and a Borelian isomorphism from $\mathcal{G}_0$ onto the set of unitary irreducible representations (UIR) of $E(2)\times\mathbb{R}$ satisfying conditions (i), (ii), (iii).

The UIR of $E(2)\times\mathbb{R}$ are either characters or the UIR

$$[W_{yw}(\theta_0,t_1,t_2,t_3)f](\theta)=\exp(iwt_3)\exp[i(y^2-w^2)^{1/2}\times(t_1\cos\theta-t_2\sin\theta)]\cdot f(\theta-\theta_0), \tag{6.3}$$

where $f\in L^2([0,2\pi])$, $(w,y)\in\mathbb{R}^2$, $|w|\le y$. We denote by $W_{|w|,w}$ the character $\exp(iwt_3)$. With this parametrization, if $\rho\in\mathcal{G}$ the decomposition of W_ρ in UIR is given by the measure ν_ρ of Prop. 6.2 (measure of the ergodic decomposition of ρ).

For even finite integer n, the group $U(2)=T\cdot SU(2)$ is realized as the product of $\Pi_{j=1}^3\exp(it_j\sum_{p=1}^n\sigma_p^j)$ by the torus defined by τ_θ. The corresponding Lie algebra $u(2)$ is generated by (J_0, J_j, $j=1,2,3$) with commutation relations

$$\begin{aligned}&[J_0,J_3]=0,\ [J_1,J_2]=J_3,\ [J_0,J_1]=[\tfrac12 J_3,J_1]=J_2,\\&[J_0,J_2]=[\tfrac12 J_3,J_2]=-J_1.\end{aligned} \tag{6.4}$$

$J_0-\frac12 J_3$ is a central element.

The UIR of U(2) are V_{ywn} defined by

$$\begin{aligned}&dV_{ywn}(J_0-\tfrac12 J_3)=-im,\ dV_{ywn|su(2)}=D(j)\\&y=2j/n,\ w=2m/n,\ m\in\mathbb{Z},\ j\in\tfrac12\mathbb{N}.\end{aligned} \tag{6.5}$$

Proposition 6.4 [3]: (1) Let ρ_n be a symmetric (with respect to $\mathfrak{S}_n$) and invariant (with respect to τ_θ) state on $\mathcal{A}_n$. Then

$$dV_{\rho_n}(J_0)=\frac{d}{d\theta}U_\theta,\ dV_{\rho_n}(J_j)=i\sum_{p=1}^{n}\pi_{\rho_n}(\sigma_p^j)\quad(j=1,2,3)$$

defines a unitary representation V_{ρ_n} of U(2) in $\mathfrak{H}_{\rho_n}$ and $V_{\rho_n}=\oplus M_{ywn}V_{ywn}$, where V_{ywn} are UIR and $M_{ywn}\in\mathbb{N}$.

(2) If V_{ywn} is a UIR of U(2) with $(y,w)\in\Delta$, then the normalized vector ω of weights 0 for $dV_{ywn}(J_0)$ defines a unique symmetric state ρ_n on $\mathcal{A}_n$ by

$$\rho_n\Big[\prod_{k=1}^{3}\Big(\sum_{p=1}^{n}i\sigma_p^k\Big)^{j_k}\Big]=\langle\omega|dV_{ywn}(J_1^{j_1}J_2^{j_2}J_3^{j_3})\omega\rangle$$

Moreover, $V_{\rho_n}=M_{ywn}V_{ywn}$ and ρ_n is a pure continuous function of positive type on U(2).

Theorem 6.1 is a particular case of the following consequence of Prop. 6.3 and 6.4. We denote by ν_{ρ_n} the measure $\sum_{(y,w)\in\Delta}\|\Omega_{ywn}\|^2\,\delta(y,w)$, where Ω_{ywn} is the component of Ω_{ρ_n} in $M_{ywn}V_{ywn}$.

Proposition 5.5 [3] (1) Let $(y,w)\in\Delta$ and $(y_n=2j/n,\ w_n=2m/n)\in\Delta$ be a sequence converging to (y,w); then the UIR W_{yw} of $E(2)\times\mathbb{R}$ is obtained by contraction of the UIR $V_{y_nw_n}$ of U(2) when $n\to\infty$.

(2) Let (ρ_n) be a sequence of states on $\mathcal{A}_n$; then (ρ_n) converges to a state ρ on $\mathcal{A}$ if and only if the sequence (ν_{ρ_n}) converges to a measure ν on Δ. Moreover,

$$\rho=\int_\Delta\rho_{yw}\,d\nu(y,w)$$

is the ergodic decomposition of ρ.

We do not discuss the convergence of the measures and limit ourself to the states ρ_n associated by proposition 6.4 with the UIR of U(2). In fact, if ρ is a local Gibbs state with Hamiltonian H_n satisfying conditions (i) and (ii) of Theorem 6.1, then H_n can be diagonalized on a base of eigenvectors of J_3 in $\mathbb{C}^{2n}$ and thus ρ can be written as a mixture of states ρ_n associated with the UIR of U(2).

7. *-FORMALISM

The algebra of symmetric observables in $\mathcal{A}_n$ is a quotient of the enveloping algebra $\mathcal{U}(\mathrm{u}(2))$ and thus can be realized as a *-subalgebra of $N(W)$, where W is a suitable symplectic manifold:

Proposition 7.1: (1)[4] Let $*_3$ (resp. $*_1$) the *-product on $T^*(S_3)$ (resp. $T^*(S_1)$) defined by (1.10). There exists a natural *-product on $W=T^*(S_3)\times T^*(S_1)$ satisfying

$$(u_1\cdot v_1)*(v_1\cdot u_2)=(u_1 *_3 u_2)\cdot(v_1 *_1 v_2),$$
$$u_1,u_2\in N(T^*(S_3)),\quad v_1,\ v_2\in N(T^*(S_1)).$$

(2) Let $(p_i,\ q_i,\ i=1,2,3,4)$ [resp. $(p_i,q_i,\ i=5,6)$] the coordinates on $T^*(S_3)$ [resp. $T^*(S_1)$] defined by (1.7) with $\sum_{i=1}^4 p_i^2=1$ and $p_5^2+p_6^2=y^2-w^2$. The functions

$$J_1=2^{-1/2}(N_{23}+N_{14}),\quad J_2=2^{-1/2}(N_{13}-N_{24}),$$
$$J_3=N_{12}+N_{34},\quad J_0=N_{56}+N_{34} \tag{7.1}$$

satisfy the commutation relations (6.4) for Moyal or Poisson brackets and thus are generators of the Lie algebra u(2).

(2) follows from Proposition 5.1.

A symmetric state on $\mathcal{A}_n$ appears as a state on the symmetric observables and thus as a function ρ on W. The expectation value of the observable $A\in N(W)$ in the state ρ is given by

$$\langle A\rangle_\rho=\int_W A*\rho\, d\omega,$$

where $d\omega$ is the Liouville measure on W and $\int_W\rho\, d\omega=1$. This can be derived from usual quantum formalism by means of the Weyl correspondence Ω. For instance, if $\rho^\wedge=\Omega(\rho)$ is a normalized density matrix

$$\rho^\wedge=\exp(-\beta H_n^\wedge)\cdot[\mathrm{Tr}\exp(-\beta H_n^\wedge)]^{-1},$$

and if $A^\wedge=\Omega(A)$, then

$$\langle A\rangle_\rho=\mathrm{tr}\rho^\wedge A^\wedge.$$

Performing the simultaneous diagonalization of H_n and J_3 in $N(W)$,

$$H_n=\sum\lambda(m,j,n)\pi_{m,j,n},$$
$$J_3=\sum M(m,j,n)\pi_{m,j,n},$$

the Gibbs state ρ is written

$$\rho=\{\sum\exp[-\beta\lambda(m,j,n)]\pi_{m,j,n}\}\cdot\{\sum\exp[-\beta\lambda(m,j,n)]\}^{-1}.$$

Then we are concerned with the states $\rho_n=\pi_{m,j,n}$ and with the associated representations $V_{y,w,n}$ defined by (6.5). In order to take the limit $n\to\infty$, we have to perform the contraction of this representation on the analogy of part 5.

Proposition 7.2: (1) Let us consider the following constraints on W:

$$k_1=p_1p_5+p_2p_6-q_3,\quad k_2=p_3,$$
$$k_3=p_1p_6-p_2p_5-q_4,\quad k_4=p_4.$$

Then the restriction of the Dirac bracket to the submanifold $\overline{W}=T^*(S_1)\times T^*(S_1)$ defined by the constraints k_i, $i=1,2,3,4$, is given by

$$\bar J_1=p_6,\quad \bar J_2=-p_5,\quad \bar J_3=\bar N_{12},\quad \bar J_0=\bar N_{56},$$
$$[\bar J_1,\bar J_2]=[\bar J_3,\bar J_i]=0,\quad i=0,\ 1,\ 2, \tag{7.2}$$
$$[\bar J_0,\bar J_1]=\bar J_2,\quad [\bar J_0,\ \bar J_2]=-\bar J_1,$$

where $\bar J_i$ $(i=0,\ 1,\ 2,\ 3)$ is the restriction of J_i to $\overline{W}$. $(\bar J_i,\ i=0,\ 1,\ 2,\ 3)$ is a set of generators of the Lie algebra $\mathcal{E}(2)+\mathbb{R}$.

(2) Moreover, if we impose the following constraints on $\overline{W}$,

$$k_5=q_1-w,\quad k_6=p_1,$$

then the restriction of the Dirac bracket to the submanifold $\overline{\overline{W}}=T^*(S_1)$ defined by k_5 and k_6 gives the representation of $\mathrm{E}(2)\times\mathbb{R}$ defined by the generators:

$$\bar{\bar J}_1=-p_6,\quad \bar{\bar J}_2=-p_5,\quad \bar{\bar J}_3=w,\quad \bar{\bar J}_0=\bar{\bar N}_{56}$$

on the space $\mathfrak{S}$ (defined by Prop. 5.6), i.e., the UIR defined by (6.3) in the *-formalism which is the contraction of the representation V_{ywn} of U(2) associated with ρ_n.

Proof: (1) The 2-cochain $C(u,v)$ corresponding to the constraints k_1, k_2, k_3, k_4 is

$$C(u,v)=\{u,\ k_1\}\{k_2,\ v\}-\{u,\ k_2\}\{k_1,\ v\}+\{u,\ k_3\}\{k_4,\ v\}$$
$$-\{u,\ k_4\}\{k_3,\ v\}$$

by restriction to $\overline{W}$:

$$\bar C(J_1,\ J_2)=2M_{12},$$
$$\bar C(J_1,\ J_3)=M_{24}-M_{13}+p_5,$$
$$\bar C(J_1,\ J_0-\tfrac{1}{2}J_3)=0.$$

An easy computation (analogous to Prop. 5.2 and 5.5) gives the commutation relations (7.2) and the representation (2).

ACKNOWLEDGMENT

We want to thank M. Flato for helpful advice and suggestions.

[1] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, Lett. Math. Phys. 1, 521—30 (1977).
[2] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, Ann. Phys. (N.Y.) 111, 61—110, 111—51 (1978).
[3] D. Arnal and J.C. Cortet, Lett. Math. Phys. 1, 505—12 (1977).
[4] A. Lichnerowicz, Lett. Math. Phys. 2, 133—43 (1977).
[5] N. Bourbaki, *Eléments de Mathématiques, Fasc. XXVI* (Hermann, Paris, 1971).
[6] M. Flato, A. Lichnerowicz, and D. Sternheimer, J. Math. Phys. 17, 1754—62 (1976).
[7] D.A. Dubin, *Solvable Models in Algebraic Statistical Mechanics* (Clarendon, Oxford, 1974).

Formulation of the linearized Vlasov-fluid model for a sharp-boundary screw pinch

H. Ralph Lewis and Leaf Turner

Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87545
(Received 5 June 1978)

A theoretical formulation is derived for analyzing linearized equations appropriate to a straight, cylindrical, sharp-boundary screw pinch within the framework of the Vlasov-fluid model. Surrounding the plasma is a cylindrical conducting wall, and there is a nonconducting vacuum between the plasma and the wall. By introducing a perturbation-dependent transformation of the phase space and linearizing about a zeroth-order state which depends on the perturbation, the linearized equations of Freidberg's Vlasov-fluid model are put into a form which would be correct for a hypothetical problem in which the plasma boundary is a rigid cylinder. The effects of the impulsive electric field at the actual perturbed boundary are taken into account in the zeroth-order state. The boundary conditions at the perturbed plasma boundary are continuity of the normal component of **B** and vanishing of the normal component of the net current density.

I. INTRODUCTION

Kinetic effects on the stability properties of hot plasmas due to finite ion gyroradii are of current experimental interest.[1,2] A useful model for studying these effects is the Vlasov-fluid model, a low-frequency model in which the ions are treated as collisionless and the electrons are treated as a massless, pressureless fluid.[3] Most of the detailed application of the Vlasov-fluid model to screw pinch configurations has been for sharp-boundary equilibrium profiles[3-5]; this is because some of the analysis can be carried out analytically and there is the possibility of comparing results with results of magnetohydrodynamics calculations. However, because of the kinetic nature of the ions, the moving sharp boundary presents some difficulties which are not encountered in a fluid calculation. The solution must take into account the fact that the ion density outside the moving boundary is always zero and that each ion that impinges on the boundary is reflected specularly in a frame of reference *moving* with the boundary. In this paper, taking the motion of the perturbed boundary into proper account, we derive a set of first-order equations for the perturbations. This is done by introducing a time-dependent transformation of the phase space to a set of variables in terms of which the time-dependent perturbed boundary appears as if it were a rigid cylinder at rest. Writing the basic equations of the Vlasov-fluid model in terms of the new variables, we then consider perturbations about a state which is a modification of the equilibrium and includes the specular reflection of the ions from the perturbed boundary. Instead of expressing the final equations in terms of the perturbation ion distribution function, it is more convenient to introduce an auxiliary function which contains boundary effects. In terms of the auxiliary function and the new phase-space variables, the equations for the linearized problem are *formally* the same as those that would describe perturbations of a system with a hypothetical, rigid cylindrical plasma–vacuum interface.

This formulation has been applied to a numerical study of finite-ion-gyroradius stabilization of $m = 2$ modes in a near-theta-pinch configuration and the results compared with results of earlier theoretical work and some experimental data.[4] Further applications are in progress. The mathematical techniques which were used in deriving the final equations of this formulation also could be applied to kinetic problems in neutral gas dynamics in which there are deformable sharp boundaries.

In Sec. II we present the basic ideas of the Vlasov-fluid model for a sharp-boundary pinch and define the equilibria that we consider. In Sec. III we define the time-dependent transformation of the phase space and derive the linearized equations. Finally, in Sec. IV, we discuss the boundary conditions that must be satisfied.

II. THE MODEL AND THE EQUILIBRIUM

In the Vlasov-fluid model proposed by Freidberg,[3] the ions are treated as collisionless, the electrons are treated as a massless, pressureless fluid, charge neutrality is assumed, and the displacement current is neglected. The governing equation of the model are

$$\mathbf{E} + \frac{1}{c}\mathbf{u}_e \times \mathbf{B} = 0, \tag{1}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{2}$$

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla_r f + \frac{Q}{M}\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right)\cdot\nabla_v f = 0, \tag{3}$$

and

$$(\nabla \times \mathbf{B})\times\mathbf{B} = 4\pi Q\int d^3\mathbf{v}\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right)f, \tag{4}$$

where **E** and **B** are the electric and magnetic fields, $\mathbf{u}_e$ is the electron fluid velocity perpendicular to **B**, f is the ion distribution function, and Q and M are the ionic charge and mass, respectively. ∇_r is the gradient operator holding the velocity **v** fixed, ∇_v is the gradient operator holding the position vector **r** fixed, and ∇ is the usual gradient operator for functions that are independent of **v**.

The electrons in this model are massless and cold ($T_e = 0$). They are treated as a fluid tied perfectly to the magnetic field lines. The electron motion parallel to the magnetic field lines is such as to preserve charge neutrality. The

0022-2488/79/040564-08$01.00

plasma systems which we consider in this paper have sharp boundaries; that is, we assume that the plasma can be adequately described in terms of an infinitely thin plasma–vacuum interface, outside of which the ion density always vanishes, regardless of how the interface moves or changes shape. Of course, the infinitely thin boundary is an approximation to a sheath region whose thickness is of the order of the ion Debye length, λ_{Di}; the ions are restrained by the attraction of the electrons, which are tied to the field lines. The time required to establish the sheath region is of the order of $2\pi/\omega_{\mathrm{pi}}$, where ω_{pi} is the ion plasma frequency. For the sharp-boundary approximation to be valid, any frequency ω and length l of importance to the dynamics of the system must satisfy

$$\omega/\omega_{\mathrm{pi}} \ll 1, \quad l/\lambda_{\mathrm{Di}} \gg 1. \tag{5}$$

(See the note added in proof at the end of the paper.)

The equilibria that we consider are axially and translationally symmetric with respect to the z axis of a cylindrical coordinate system (coordinates r, θ, z, and unit vectors $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{z}}$). The equilibrium ion distribution function, $f_0(\mathbf{r},\mathbf{v})$, is a function only of the total particle energy ϵ:

$$f_0(\mathbf{r},\mathbf{v}) = \mathscr{F}(\epsilon), \tag{6}$$

$$\epsilon = \tfrac{1}{2}Mv^2 + Q\phi_0(r), \tag{7}$$

where the equilibrium scalar potential, $\phi_0(r)$, depends only on r. The equilibrium number density and pressure as functions of r are given by

$$n_0(r) = \int d^3\mathbf{v}\,\mathscr{F}(\epsilon) \tag{8}$$

and

$$p_0(r) = \frac{M}{3}\int d^3\mathbf{v}\,v^2\mathscr{F}(\epsilon). \tag{9}$$

The equilibrium electric field, $\mathbf{E}_0(\mathbf{r})$, is related to $p_0(r)$ by

$$\nabla p_0(r) = Qn_0(r)\mathbf{E}_0(\mathbf{r}), \tag{10}$$

as may be verified by taking the first velocity moment of Eq. (3), or by using ϵ as a variable of integration in Eqs. (8) and (9). The constraint on the equilibrium magnetic field, $\mathbf{B}_0(\mathbf{r})$, is obtained by substituting Eq. (6) into Eq. (4) and using Eq. (10). The result is

$$\nabla p_0(r) = (1/4\pi)[\nabla \times \mathbf{B}_0(\mathbf{r})] \times \mathbf{B}_0(\mathbf{r}), \tag{11}$$

which is the same constraint as in ideal magnetohydrodynamics.

Introducing the sharp-boundary approximation, we take

$$\mathscr{F}(\epsilon) = 0 \quad \text{for } \epsilon \geqslant \epsilon_m, \tag{12}$$

and

$$\phi_0(r) = \begin{cases} 0, & r < r_0, \\ \epsilon_m/Q, & r > r_0, \end{cases} \tag{13}$$

where ϵ_m is the maximum energy of any particle in equilibrium. (For a Maxwellian, ϵ_m would be infinite.) Then $n_0(r)$ and $p_0(r)$ are discontinuous:

$$n_0(r) = \begin{cases} n_0 = \text{const}, & r < r_0, \\ 0, & r > r_0, \end{cases} \tag{14}$$

$$p_0(r) = \begin{cases} p_0 = \text{const}, & r < r_0, \\ 0, & r > r_0. \end{cases} \tag{15}$$

A uniform background of electron charge for $r < r_0$ provides the charge neutrality required by the model. Surrounding the cylindrical plasma column of radius r_0 is a rigid, perfectly conducting, coaxial cylindrical wall of radius $R > r_0$. The enclosed annular region is a (nonconducting) vacuum.

The equilibrium electric field vanishes inside and outside the plasma column, and is impulsive at the plasma–vacuum interface. The impulsive electric field at the plasma–vacuum interface is also present when the system is perturbed and it is responsible for confining all ions to the interior of the plasma column. At any given point on the interface, in a frame of reference moving with the instantaneous velocity of that point, the impulsive electric field is conservative, normal to the boundary, and sufficient to reflect any impinging ion back into the plasma column. Because the impulsive electric field is conservative, the reflection of an ion is specular with respect to the frame of reference moving with the instantaneous velocity of the interface at the point of impact.

We specify the equilibrium magnetic field by

$$\mathbf{B}_0(\mathbf{r}) = \begin{cases} b_i\hat{\mathbf{z}}, & r < r_0, \\ b_z\hat{\mathbf{z}} + b_\theta(r_0/r)\hat{\boldsymbol{\theta}}, & r > r_0, \end{cases} \tag{16}$$

where the constants b_i, b_z, and b_θ are, respectively, the interior z component of $\mathbf{B}_0$, the exterior z component of $\mathbf{B}_0$, and the exterior θ component of $\mathbf{B}_0$ evaluated at the plasma–vacuum interface. The pressure balance constraint on these constants, obtained by integrating Eq. (11) across the boundary at $r = r_0$, is

$$p_0 + b_i^2/8\pi = (1/8\pi)(b_\theta^2 + b_z^2). \tag{17}$$

III. THE LINEARIZED EQUATIONS

Because the plasma boundary can move, it is incorrect to linearize the equations in terms of perturbations about the equilibrium configuration. When the boundary crosses any fixed point, the distribution function at that point changes from a finite value to zero, or *vice versa*; in either case, the change of the distribution function at that point is not small. In order to have perturbation quantities that are indeed small at all points, we shall introduce a new set of independent variables whose definition depends on the displacement of the boundary, and we shall linearize the equations of motion about a zeroth-order state which also depends on the boundary displacement, but which reduces to the equilibrium state in the absence of a perturbation.

We define the new independent variables in terms of a perturbation-dependent scale transformation. When the plasma column suffers a perturbation, its boundary is affected. We let the function $\rho(\theta,z,t)$ be the radial distance at time t from the equilibrium axis to the point on the perturbed plasma–vacuum interface with azimuthal and axial coordinates

θ and z. The new independent variables, which we denote by a prime, are determined by a scale transformation of the radial coordinate and corresponding transformations of v_r and v_θ:

$$\begin{aligned} r' &= \frac{r_0}{\rho(\theta,z,t)}r, \\ \theta' &= \theta, \\ z' &= z, \\ t' &= t, \\ v_r' &= \frac{r_0}{\rho}\left(v_r - \frac{r}{\rho}\dot{\rho}\right), \\ v_\theta' &= \frac{r_0}{\rho}v_\theta, \\ v_z' &= v_z. \end{aligned} \tag{18}$$

(A dot over a symbol denotes partial differentiation with respect to time.) Partial derivatives with respect to the new variables are given in terms of partial derivatives with respect to the old variables by:

$$\begin{aligned} \frac{\partial}{\partial r} &= \frac{r_0}{\rho}\left(\frac{\partial}{\partial r'} - \frac{\dot{\rho}}{\rho}\frac{\partial}{\partial v_r'}\right), \\ \frac{\partial}{\partial q} &= \frac{\partial}{\partial q'} - \frac{1}{\rho}\frac{\partial \rho}{\partial q'} \\ &\quad \times\left(r'\frac{\partial}{\partial r'} + v_r'\frac{\partial}{\partial v_r'} + v_\theta'\frac{\partial}{\partial v_\theta'}\right) \\ &\quad - r'\left(\frac{\partial}{\partial q'}\frac{\dot{\rho}}{\rho}\right)\frac{\partial}{\partial v_r'}, \\ \frac{\partial}{\partial v_r} &= \frac{r_0}{\rho}\frac{\partial}{\partial v_r'}, \\ \frac{\partial}{\partial v_\theta} &= \frac{r_0}{\rho}\frac{\partial}{\partial v_\theta'}, \\ \frac{\partial}{\partial v_z} &= \frac{\partial}{\partial v_z'}, \end{aligned} \tag{19}$$

where q refers to either θ, z, or t. Note that

$$\rho(\theta,z,t,) = \rho(\theta',z',t'),$$

so that

$$\dot{\rho} = \frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial t'}.$$

The gradient operators $\nabla_{\mathbf{r}}$ and $\nabla_{\mathbf{v}}$ written in terms of cylindrical coordinates are

$$\nabla_{\mathbf{r}} = \hat{\mathbf{r}}\frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}}\left[\frac{1}{r}\left(\frac{\partial}{\partial \theta} + v_\theta\frac{\partial}{\partial v_r} - v_r\frac{\partial}{\partial v_\theta}\right)\right] + \hat{\mathbf{z}}\frac{\partial}{\partial z}, \tag{20}$$

$$\nabla_{\mathbf{v}} = \hat{\mathbf{r}}\frac{\partial}{\partial v_r} + \hat{\boldsymbol{\theta}}\frac{\partial}{\partial v_\theta} + \hat{\mathbf{z}}\frac{\partial}{\partial v_z}. \tag{21}$$

When $\nabla_{\mathbf{r}}$ acts on a function that is independent of $\mathbf{v}$, we find from Eqs. (19) and (20)

$$\begin{aligned} \nabla_{\mathbf{r}} = \nabla &= \frac{r_0}{\rho}\left[\hat{\mathbf{r}}\frac{\partial}{\partial r'} + \hat{\boldsymbol{\theta}}\left(\frac{1}{r'}\frac{\partial}{\partial \theta'} - \frac{1}{\rho}\frac{\partial \rho}{\partial \theta'}\frac{\partial}{\partial r'}\right)\right] \\ &\quad + \hat{\mathbf{z}}\left(\frac{\partial}{\partial z'} - \frac{r'}{\rho}\frac{\partial \rho}{\partial z'}\frac{\partial}{\partial r'}\right). \end{aligned} \tag{22}$$

Now, using Eqs. (19) and (22), we can write the streaming part of the total time derivative operator as

$$\begin{aligned} \frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{r}} &= \frac{\partial}{\partial t'} + \mathbf{v}'\cdot\nabla_{\mathbf{r}'}' - \left(\frac{v_\theta'}{\rho r'}\frac{\partial \rho}{\partial \theta'} + \frac{v_z'}{\rho}\frac{\partial \rho}{\partial z'}\right) \\ &\quad \times\left(r'\frac{\partial}{\partial r'} + v_r'\frac{\partial}{\partial v_r'} + v_\theta'\frac{\partial}{\partial v_\theta'}\right) \\ &\quad - \frac{2\dot{\rho}}{\rho}\left(v_r'\frac{\partial}{\partial v_r'} + v_\theta'\frac{\partial}{\partial v_\theta'}\right) - r'\frac{\ddot{\rho}}{\rho}\frac{\partial}{\partial v_r'} \\ &\quad - r'\left[\frac{v_\theta'}{r'}\left(\frac{\partial}{\partial \theta'}\frac{\dot{\rho}}{\rho}\right) + v_z'\left(\frac{\partial}{\partial z'}\frac{\dot{\rho}}{\rho}\right)\right]\frac{\partial}{\partial v_r'}, \end{aligned} \tag{23}$$

where

$$\mathbf{r}' = r'\hat{\mathbf{r}} + z'\hat{\mathbf{z}}, \tag{24}$$

$$\mathbf{v}' = v_r'\hat{\mathbf{r}} + v_\theta'\hat{\boldsymbol{\theta}} + v_z'\hat{\mathbf{z}}, \tag{25}$$

and

$$\nabla_{\mathbf{r}'}' = \hat{\mathbf{r}}\frac{\partial}{\partial r'} + \hat{\boldsymbol{\theta}}\left[\frac{1}{r'}\left(\frac{\partial}{\partial \theta'} + v_\theta'\frac{\partial}{\partial v_r'} - v_r'\frac{\partial}{\partial v_\theta'}\right)\right] + \hat{\mathbf{z}}\frac{\partial}{\partial z_\theta}. \tag{26}$$

(Note that the subscript $\mathbf{r}'$ on $\nabla_{\mathbf{r}'}'$ is superfluous when the operand is independent of $\mathbf{v}'$. In that situation, the subscript will be suppressed.)

The zeroth-order quantities that we shall use for linearizing the equations of the model will be denoted by a superscript zero. As mentioned previously, they depend on the perturbation from equilibrium, but reduce to the equilibrium quantities when the perturbation vanishes. The condition $\nabla\cdot\mathbf{B} = 0$ and the $\nabla\times\mathbf{E}$ Maxwell equation, Eq. (2), will be satisfied to first order by introducing suitable scalar and vector potentials:

$$\mathbf{B} = \mathbf{B}^{(0)} + \mathbf{B}^{(1)} = \nabla\times\mathbf{A}, \tag{27a}$$

$$\mathbf{E} = \mathbf{E}^{(0)} + \mathbf{E}^{(1)} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \tag{27b}$$

$$\phi = \phi^{(0)} + \phi^{(1)}, \tag{27c}$$

$$\mathbf{A} = \mathbf{A}^{(0)} + \mathbf{A}^{(1)}. \tag{27d}$$

The zeroth-order fields and potentials at a given point are the analytic continuations of the interior equilibrium fields and potentials if the point is inside the *perturbed* boundary; if the point is outside the perturbed boundary, then the zeroth-order fields and potentials are the analytic continuations of the corresponding exterior equilibrium quantities.

The zeroth-order scalar potential is

$$\phi^{0} = (\epsilon_m/Q)\Theta(r' - r_0), \tag{28}$$

where Θ is the unit step function defined by

$$\Theta(x) = \begin{cases} 0, & x < 0, \\ 1, & x > 0. \end{cases} \tag{29}$$

Whereas the equilibrium scalar potential, Eq. (13), vanishes for $r < r_0$ and equals the constant ϵ_m/Q for $r_0 < r < R$, our zeroth-order scalar potential vanishes inside the *perturbed* plasma column and equals ϵ_m/Q in the vacuum region outside the *perturbed* column. As a result, $-\nabla\phi^{(0)}$ is everywhere locally normal to the perturbed plasma–vacuum interface. Using Eq. (22), we find

$$-\nabla\phi^{(0)} = -\frac{\epsilon_m}{Q}\frac{r_0}{\rho}\delta(r' - r_0)\left(\hat{\mathbf{r}} - \hat{\boldsymbol{\theta}}\frac{1}{\rho}\frac{\partial\rho}{\partial\theta'} - \hat{\mathbf{z}}\frac{\partial\rho}{\partial z'}\right). \tag{30}$$

The zeroth-order magnetic field is defined by

$$\mathbf{B}^{(0)} = \begin{cases} b_i\hat{\mathbf{z}}, & r < \rho, \\ b_z\hat{\mathbf{z}} + b_\theta(r_0/r)\hat{\boldsymbol{\theta}}, & r > \rho, \end{cases}$$

$$= b_i\hat{\mathbf{z}} + \Theta(r - \rho)[(b_z - b_i)\hat{\mathbf{z}} + b_\theta(r_0/r)\hat{\boldsymbol{\theta}}]. \tag{31}$$

In the case of no perturbation ($\rho \equiv r_0$), the normal component of $\mathbf{B}^{(0)}$ is continuous across the plasma–vacuum interface ($r = r_0$). However, when there is a perturbation, the component of $\mathbf{B}^{(0)}$ normal to the perturbed boundary is usually discontinuous, so that $\mathbf{B}^{(0)}$ is not properly solenoidal at the plasma–vacuum interface. However, one of the conditions to be imposed on $\mathbf{B}^{(1)}$ is continuity to first order of the normal component of the *total* magnetic field, $\mathbf{B}^{(0)} + \mathbf{B}^{(1)}$, at the perturbed boundary; this is achievable because $\mathbf{B}^{(0)}$ is the analytic continuation of the interior or exterior equilibrium field, so that the difference between $\mathbf{B}^{(0)}$ and the total field is of first order. It is convenient to choose $\mathbf{B}^{(0)}$ as we have done, and it is physically irrelevant that $\mathbf{B}^{(0)}$ is not solenoidal at the boundary. For the zeroth-order vector potential we take

$$\mathbf{A}^{(0)} = \tfrac{1}{2}b_i r\hat{\boldsymbol{\theta}} + \Theta(r - \rho)\{\tfrac{1}{2}(b_z - b_i)[(r^2 - r_0^2)/r]\hat{\boldsymbol{\theta}} - b_\theta r_0 \ln(r/r_0)\hat{\mathbf{z}}\}. \tag{32}$$

It is readily verified that

$$\mathbf{B}^{(0)} = \nabla\times\mathbf{A}^{(0)} \quad \text{for } r \neq \rho, \tag{33}$$

and that $\mathbf{A}^{(0)}$ is continuous at the boundary in the case of equilibrium ($\rho \equiv r_0$). The time derivative of this vector potential, $\partial\mathbf{A}^{(0)}/\partial t$, is $\delta(r - \rho)$ times a coefficient which is *second order* in the perturbation quantities. As a result, $(1/c)(\partial\mathbf{A}^{(0)}/\partial t)$ is negligible compared to $\nabla\phi^{(0)}$, and we may write

$$\mathbf{E}^{(0)} \equiv -\nabla\phi^{(0)} - \frac{1}{c}\frac{\partial\mathbf{A}^{(0)}}{\partial t} = -\nabla\phi^{(0)} \tag{34}$$

to first order.

We now use Eq. (1) to determine $\phi^{(1)}$ and to express the radius of the perturbed boundary, $\rho(\theta,z,t)$, in terms of $\mathbf{A}^{(1)}$. The field $\mathbf{E}^{(1)}$ can be written correct to first order for all values of r as

$$\mathbf{E}^{(1)} = -\nabla'\phi^{(1)} - \frac{1}{c}\frac{\partial\mathbf{A}^{(1)}}{\partial t'}; \tag{35a}$$

the field $\mathbf{B}^{(1)}$ for $r \neq \rho$ can be written as

$$\mathbf{B}^{(1)} = \nabla'\times\mathbf{A}^{(1)}. \tag{35b}$$

The total magnetic field can be written correct to first order for all values of r as

$$\mathbf{B} = \nabla\times(\mathbf{A}^{(0)} + \mathbf{A}^{(1)}). \tag{35c}$$

This is true because of the condition to be imposed on $\mathbf{A}^{(1)}$ that the normal component of $\mathbf{B}$ be continuous. As a result, $\mathbf{B}^{(1)}$ is expressible correct to first order for all values of r as

$$\mathbf{B}^{(1)} = \nabla'\times\mathbf{A}^{(1)} + \delta(r' - r_0)(\rho - r_0)[(b_z - b_i)\hat{\mathbf{z}} + b_\theta\hat{\boldsymbol{\theta}}]. \tag{35d}$$

In addition to the zeroth-order quantities already defined, it is useful to define

$$\mathbf{E}^{(0)\prime} = -(\epsilon_m/Q)\delta(r' - r_0)\hat{\mathbf{r}}, \tag{36a}$$

$$\mathbf{B}^{(0)\prime} = b_i\hat{\mathbf{z}} + \Theta(r' - r_0)[(b_z - b_i)\hat{\mathbf{z}} + b_\theta(r_0/r')\hat{\boldsymbol{\theta}}], \tag{36b}$$

$$\mathbf{u}_e^{(0)\prime} = (c\mathbf{E}^{(0)\prime}\times\mathbf{B}^{(0)\prime})/B^{(0)\prime 2}. \tag{36c}$$

These definitions imply

$$\mathbf{E}^{(0)\prime} + (1/c)\mathbf{u}_e^{(0)\prime}\times\mathbf{B}^{(0)\prime} = 0. \tag{37}$$

Writing the scalar product of Eq. (1) with $\mathbf{B}$, correct to first order, we find

$$\mathbf{E}\cdot\mathbf{B} = \mathbf{E}^{(0)}\cdot\mathbf{B}^{(0)} + \mathbf{E}^{(0)\prime}\cdot\mathbf{B}^{(1)} + \mathbf{E}^{(1)}\cdot\mathbf{B}^{(0)\prime} = 0. \tag{38}$$

Correct to first order, the individual terms can be expressed as

$$\mathbf{E}^{(0)}\cdot\mathbf{B}^{(0)} = (\epsilon_m/Q)\delta(r' - r_0)\mathbf{B}^{(0)\prime}\cdot\nabla'(\rho - r_0) = \mathbf{B}^{(0)\prime}\cdot\nabla'[(\epsilon_m/Q)\delta(r' - r_0)(\rho - r_0)],$$

$$\mathbf{E}^{(0)\prime}\cdot\mathbf{B}^{(1)} = \mathbf{B}^{(0)\prime}\cdot\nabla'\left(\mathbf{A}^{(1)}\cdot\frac{1}{c}\mathbf{u}_e^{(0)\prime}\right) - \nabla'\cdot\left((\mathbf{A}^{(1)}\cdot\mathbf{B}^{(0)\prime})\frac{1}{c}\mathbf{u}_e^{(0)\prime}\right),$$

$$\mathbf{E}^{(1)}\cdot\mathbf{B}^{(0)\prime} = -\mathbf{B}^{(0)\prime}\cdot\nabla'\phi^{(1)} - \frac{1}{c}\frac{\partial}{\partial t'}(\mathbf{A}^{(1)}\cdot\mathbf{B}^{(0)\prime}),$$

so that Eq. (38) can be rewritten as

$$\mathbf{B}^{(0)\prime}\cdot\nabla'\left[\frac{\epsilon_m}{Q}\delta(r' - r_0)(\rho - r_0) + \frac{1}{c}\mathbf{A}^{(1)}\cdot\mathbf{u}_e^{(0)\prime} - \phi^{(1)}\right] - \frac{1}{c}\left\{\nabla'\cdot[(\mathbf{A}^{(1)}\cdot\mathbf{B}^{(0)\prime})\mathbf{u}_e^{(0)\prime}] + \frac{\partial}{\partial t'}(\mathbf{A}^{(1)}\cdot\mathbf{B}^{(0)\prime})\right\} = 0. \tag{39}$$

We now choose a gauge such that

$$\mathbf{A}^{(1)}\cdot\mathbf{B}^{(0)\prime} = 0, \tag{40}$$

and we define a vector $\boldsymbol{\xi}$ *perpendicular* to $\mathbf{B}^{(0)\prime}$ by

$$\boldsymbol{\xi} = (\mathbf{B}^{(0)\prime}\times\mathbf{A}^{(1)})/B^{(0)\prime 2}, \tag{41a}$$

so that

$$\mathbf{A}^{(1)} = \boldsymbol{\xi}\times\mathbf{B}^{(0)\prime}. \tag{41b}$$

With these definitions, Eq. (39) becomes

$$\mathbf{B}^{(0)\prime}\cdot\nabla'\{[\boldsymbol{\xi} - (\rho - r_0)\hat{\mathbf{r}}]\cdot\mathbf{E}^{(0)\prime} - \phi^{(1)}\} = 0. \tag{42}$$

We want to use Eq. (42) to determine $\phi^{(1)}$ within the plasma, including the value at the boundary. Because the equilibrium and zeroth-order quantities are independent of θ' and z', neither θ' nor z' will appear in any linearized equation of the model except in the differential operators $\partial/\partial\theta'$ and $\partial/\partial z'$. Therefore, all perturbation quantities can be expanded as a double Fourier series in θ' and z', and we can analyze each term separately. That is, we can let the complete θ' and z' dependence of each perturbation quantity occur in a factor $\exp[i(m\theta' + kz')]$, where m is an integer. We now restrict our attention to the case

$$[k\hat{\mathbf{z}} + (m/r')\hat{\boldsymbol{\theta}}]\cdot\mathbf{B}^{(0)\prime}(\mathbf{r}')\neq 0, \tag{43}$$

in which case the only solution of Eq. (42) is

$$\phi^{(1)} = [\boldsymbol{\xi} - (\rho - r_0)\hat{\mathbf{r}}]\cdot\mathbf{E}^{(0)\prime}. \tag{44}$$

This implies

$$\phi^{(1)} = 0 \quad \text{for } r<\rho. \tag{45}$$

In order to use Eq. (44) for determining $\phi^{(1)}$ when $r=\rho$, we first use Eq. (1) to find the connection between $\boldsymbol{\xi}$ and $(\rho - r_0)$. Let

$$\mathbf{u}_e = \mathbf{u}_e^{(0)} + \mathbf{u}_e^{(1)}, \tag{46a}$$

where

$$\mathbf{u}_e^{(0)} = (c\mathbf{E}^{(0)}\times\mathbf{B}^{(0)})/B^{(0)2}. \tag{46b}$$

For $r<\rho$, both $\mathbf{E}^{(0)}$ and $\mathbf{u}_e^{(0)}$ vanish. Therefore, correct to first order, we can write

$$\mathbf{E}^{(1)} + (1/c)\mathbf{u}_e^{(1)}\times\mathbf{B}^{(0)\prime} = 0 \quad \text{for } r<\rho. \tag{47}$$

Using Eqs. (35a), (41b), and (45), we then obtain

$$\frac{1}{c}\left(\mathbf{u}_e^{(1)} - \frac{\partial\boldsymbol{\xi}}{\partial t}\right)\times\mathbf{B}^{(0)\prime} = 0$$

or, equivalently,

$$\frac{\partial\boldsymbol{\xi}}{\partial t} = \mathbf{u}_e^{(1)}. \tag{48}$$

Thus, within the plasma column, $\boldsymbol{\xi}$ is the displacement from equilibrium of the electron fluid perpendicular to the magnetic field. Because the electron fluid is frozen to the magnetic field lines, $\boldsymbol{\xi}$ is also the displacement of the magnetic field lines within the plasma column. Our assumption, made in Sec. II, that electrons neither enter nor leave the plasma-vacuum interface now allows us to express the radius of the perturbed boundary as

$$\rho(\theta',z',t') = r_0 + \lim_{r'\to r_0}\xi_r(r',\theta',z',t'), \tag{49}$$

where the limit is taken from inside the boundary. In fact, the limit is unnecessary, because continuity of the normal component of $\mathbf{B}$ at the perturbed boundary implies continuity of $\xi_r(r',\theta',z',t')$ for $r'=r_0$. This will be demonstrated in Sec. IV, but it should be borne in mind during the remainder of this section. Without the continuity of ξ_r, some subsequent expressions in this section would be ill defined. Returning to Eq. (44), we now conclude that

$$\phi^{(1)} = 0 \quad \text{for } r\leqslant\rho. \tag{50}$$

The fact that $\phi^{(1)}$ vanishes everywhere within the perturbed plasma column indicates clearly that our choice of the zeroth-order scalar potential $\phi^{(0)}$ indeed does account for the impulsive electric field up to first order.

We shall not need to know $\phi^{(1)}$ in the vacuum region surrounding the plasma, but it is interesting to note that it does not generally vanish. In fact, $\phi^{(1)}$ in the vacuum region is governed by the condition

$$\nabla^2\phi^{(1)} = -\frac{1}{c}\nabla\cdot\frac{\partial\mathbf{A}^{(1)}}{\partial t}, \tag{51}$$

which expresses the fact that the vacuum region is free of electric charge. If $\partial\mathbf{A}^{(1)}/\partial t$ were known, then $\phi^{(1)}$ could be determined from Eq. (51).

We choose the zeroth-order ion distribution function, $f^{(0)}$, such that it reduces to the equilibrium distribution function in the absence of perturbation and satisfies the condition that

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{r}} + \frac{Q}{M}\mathbf{E}^{(0)}\cdot\nabla_{\mathbf{v}}\right)f^{(0)}$$

is nonsingular at the perturbed boundary. Writing

$$f = f^{(0)} + f^{(1)}, \tag{52}$$

we take

$$f^{(0)}(\mathbf{r},\mathbf{v},t) = \mathscr{F}(\epsilon_0), \tag{53a}$$

where

$$\epsilon_0 = \tfrac{1}{2}M\{[v_r - (r/\rho)\dot{\rho}]^2 + v_\theta^2 + v_z^2\} + Q\phi^{(0)}(\mathbf{r},t)$$

$$= \tfrac{1}{2}M[(\rho/r_0)^2 v_r'^2 + (\rho/r_0)^2 v_\theta'^2 + v_z'^2]$$

$$+ \epsilon_m\Theta(r' - r_0). \tag{53b}$$

We also define another energy, ϵ', which has the same form in terms of the primed variables that ϵ, given by Eq. (7), has in terms of the unprimed variables:

$$\epsilon' = \tfrac{1}{2}Mv'^2 + \epsilon_m\Theta(r' - r_0). \tag{54}$$

The linearized form of Eq. (3) within the perturbed plasma column can be written as

$$\left(\frac{\partial}{\partial t'} + \mathscr{L}'\right)\tilde{g}(\mathbf{r}',\mathbf{v}',t')$$

$$= \mathscr{F}'(\epsilon')M\mathbf{v}'\cdot\left(\frac{\partial}{\partial t'} + \mathbf{v}'\cdot\nabla'_{\mathbf{r}'}\right)\frac{\partial\boldsymbol{\xi}}{\partial t'}, \tag{55}$$

where

$$\tilde{g} = f^{(1)} + \mathscr{F}'(\epsilon')\left[M\mathbf{v}'\cdot\frac{\partial\boldsymbol{\xi}}{\partial t'} - Mv_r'\left(\frac{r'}{r_0}\right)\dot{\rho}\right], \tag{56}$$

$$\mathscr{L}' = \mathbf{v}'\cdot\nabla'_{\mathbf{r}'} + \frac{Q}{M}\left(\mathbf{E}^{(0)\prime} + \frac{1}{c}\mathbf{v}'\times\mathbf{B}^{(0)\prime}\right)\cdot\nabla'_{\mathbf{v}'}, \tag{57}$$

$$\nabla'_{\mathbf{v}'} = \hat{\mathbf{r}}\frac{\partial}{\partial v_r'} + \hat{\boldsymbol{\theta}}\frac{\partial}{\partial v_\theta'} + \hat{\mathbf{z}}\frac{\partial}{\partial v_z'}, \tag{58}$$

and

$$\mathscr{F}'(\epsilon')=\frac{d}{d\epsilon'}\mathscr{F}(\epsilon').$$

Note that $\mathscr{F}'$ is an equilibrium Liouville operator for the primed variables. Also note that Eq. (55) is *formally* the same as would have been obtained in terms of the *unprimed* variables had the plasma–vacuum interface been a *rigid* cylinder, in which case the term in Eq. (56) involving $\dot{\rho}$ would have been absent. In order to derive Eq. (55), the following relations which are valid to first order are useful:

$$\left[\frac{\partial}{\partial t}+\mathbf{v}\cdot\nabla_{\mathbf{r}}+\frac{Q}{M}\left(\mathbf{E}^{(0)}+\frac{1}{c}\mathbf{v}\times\mathbf{B}^{(0)}\right)\cdot\nabla_{\mathbf{v}}\right]\mathscr{F}(\epsilon_0)$$
$$=-\left(\frac{\partial}{\partial t'}+\mathscr{L}'\right)\left[\left(\frac{r'}{r_0}\right)\rho(\theta',z',t')\hat{\mathbf{r}}\cdot\nabla'_{\mathbf{v}}\mathscr{F}(\epsilon')\right]$$
$$+Q\mathscr{F}'(\epsilon')\dot{\rho}\hat{\mathbf{r}}\cdot\mathbf{E}^{(0)\prime},\qquad(59)$$

$$\frac{Q}{M}\left(\mathbf{E}^{(1)}+\frac{1}{c}\mathbf{v}\times\mathbf{B}^{(1)}\right)\cdot\nabla_{\mathbf{v}}\mathscr{F}(\epsilon_0)=Q\,\mathbf{E}^{(1)}\cdot\mathbf{v}'\mathscr{F}'(\epsilon'),\qquad(60)$$

$$\frac{Q}{M}\left(\mathbf{E}^{(1)}+\frac{1}{c}\mathbf{v}\times\mathbf{B}^{(1)}\right)\cdot\nabla_{\mathbf{v}}\mathscr{F}(\epsilon_0)$$
$$=\mathscr{F}'(\epsilon')\left[M\left(\frac{\partial}{\partial t'}+\mathscr{L}'\right)\left(\mathbf{v}'\cdot\frac{\partial\boldsymbol{\xi}}{\partial t'}\right)\right.$$
$$\left.-M\mathbf{v}'\cdot\left(\frac{\partial}{\partial t'}+\mathbf{v}'\cdot\nabla'\right)\frac{\partial\boldsymbol{\xi}}{\partial t'}-Q\dot{\rho}\hat{\mathbf{r}}\cdot\mathbf{E}^{(0)\prime}\right].\qquad(61)$$

We now turn to the linearization of Eq. (4) with respect to the zeroth-order quantities that have been defined. The $(\nabla\times\mathbf{B})\times\mathbf{B}$ term is singular at the plasma–vacuum interface because of the discontinuity of $\mathbf{B}$. Writing

$$(\nabla\times\mathbf{B})\times\mathbf{B}=\mathbf{B}\cdot\nabla\mathbf{B}-\tfrac{1}{2}\nabla B^2,$$

and noting that $\mathbf{B}\cdot\nabla\mathbf{B}$ is a derivative of $\mathbf{B}$ tangential to the boundary, we see that the singularity is associated exclusively with the term $-\frac{1}{2}\nabla B^2$. Correct to first order, we have

$$-\tfrac{1}{2}\nabla B^2=-\nabla(\tfrac{1}{2}B^{(0)2}+\mathbf{B}^{(0)}\cdot\mathbf{B}^{(1)})$$
$$=-\nabla\left[B^{(0)\prime 2}+\mathbf{B}^{(0)\prime}\cdot\mathbf{B}^{(1)}\right.$$
$$\left.-\Theta(r'-r_0)b_\theta^2\left(\frac{r_0}{r'}\right)^2\left(\frac{\rho-r_0}{r_0}\right)\right].$$

Thus, the singular part of $(\nabla\times\mathbf{B})\times\mathbf{B}$ can be written correct to first order as

$$\tfrac{1}{2}(b_i^2-b_z^2-b_\theta^2)\nabla\Theta(r-\rho)+\psi(\theta',z',t')\nabla'\Theta(r'-r_0),$$

where

$$\psi(\theta',z',t')=\lim_{\sigma\to 0^+}(\mathbf{B}^{(0)\prime}\cdot\mathbf{B}^{(1)}|_{r'=r_0-\sigma}$$
$$-\mathbf{B}^{(0)\prime}\cdot\mathbf{B}^{(1)}|_{r'=r_0+\sigma})+b_\theta^2\left(\frac{\rho-r_0}{r_0}\right).\qquad(62)$$

We define a vector differential operator $\tilde{\mathbf{F}}$ by

$$(1/4\pi)(\nabla\times\mathbf{B})\times\mathbf{B}-(1/8\pi)(b_i^2-b_z^2-b_\theta^2)\nabla\Theta(r-\rho)$$
$$=\tilde{\mathbf{F}}(\boldsymbol{\xi})+(1/4\pi)\psi(\theta',z',t')\nabla'\Theta(r'-r_0).\qquad(63a)$$

For $r'\neq r_0$, $\tilde{\mathbf{F}}$ is simply

$$\tilde{\mathbf{F}}(\boldsymbol{\xi})=\frac{1}{4\pi}(\nabla'\times\mathbf{B}^{(1)})\times\mathbf{B}^{(0)\prime}.\qquad(63b)$$

For $r'=r_0$, $\tilde{\mathbf{F}}(\boldsymbol{\xi})$ is *finite*, and can be taken equal to the limit of $\tilde{\mathbf{F}}(\boldsymbol{\xi})$ as r' approaches r_0 from within; it is negligible compared to the singular terms in Eq. (63a).

The right-hand side of Eq. (4), divided by $4\pi Q$, can be written correct to first order as the sum of seven terms as follows:

$$\int d^3\mathbf{v}\left(\mathbf{E}+\frac{1}{c}\mathbf{v}\times\mathbf{B}\right)f$$
$$=\int d^3\mathbf{v}\mathbf{E}^{(0)}\mathscr{F}(\epsilon_0)+\int d^3\mathbf{v}\frac{1}{c}\mathbf{v}\times\mathbf{B}^{(0)}\mathscr{F}(\epsilon_0)$$
$$+\int d^3\mathbf{v}'\mathbf{E}^{(1)}\mathscr{F}(\epsilon')+\int d^3\mathbf{v}'\frac{1}{c}\mathbf{v}'\times\mathbf{B}^{(1)}\mathscr{F}(\epsilon')$$
$$-\int d^3\mathbf{v}'\mathscr{F}'(\epsilon')\left(\mathbf{E}^{(0)\prime}+\frac{1}{c}\mathbf{v}'\times\mathbf{B}^{(0)\prime}\right)M\mathbf{v}'\cdot\frac{\partial\boldsymbol{\xi}}{\partial t'}$$
$$+\int d^3\mathbf{v}'\mathscr{F}'(\epsilon')\left(\mathbf{E}^{(0)\prime}+\frac{1}{c}\mathbf{v}'\times\mathbf{B}^{(0)\prime}\right)Mv'_r\left(\frac{r'}{r_0}\right)\dot{\rho}$$
$$+\int d^3\mathbf{v}'\left(\mathbf{E}^{(0)\prime}+\frac{1}{c}\mathbf{v}'\times\mathbf{B}^{(0)\prime}\right)\tilde{g}.\qquad(64)$$

The first term can be expressed as

$$\int d^3\mathbf{v}\mathbf{E}^{(0)}\mathscr{F}(\epsilon_0)=\frac{1}{Q}\nabla p^{(0)}=-\frac{p_0}{Q}\nabla\Theta(r-\rho),$$

where $p^{(0)}$ is a pressure function defined by

$$p^{(0)}=\frac{M}{3}\int d^3\mathbf{v}v^2\mathscr{F}(\epsilon_0).\qquad(65)$$

The second integral can be expressed in terms of the primed variables as

$$\int d^3\mathbf{v}\frac{1}{c}\mathbf{v}\times\mathbf{B}^{(0)}\mathscr{F}(\epsilon_0)$$
$$=\frac{1}{c}\int d^3\mathbf{v}'\mathscr{F}(\epsilon')\left(\frac{r'}{r_0}\right)\dot{\rho}\left[\hat{\mathbf{z}}\Theta(r'-r_0)b_\theta\left(\frac{r_0}{r'}\right)\right.$$
$$\left.-\hat{\boldsymbol{\theta}}[b_i+\Theta(r'-r_0)(b_z-b_i)]\right].$$

The fourth integral vanishes because ϵ' is only a function of v'. By use of Eqs. (47) and (48) and the identity

$$M\int d^3\mathbf{v}'\mathbf{v}'\mathbf{v}'\mathscr{F}'(\epsilon')=-\mathbf{1}\int d^3\mathbf{v}'\mathscr{F}(\epsilon'),$$

where $\mathbf{1}$ is the unit dyad, it may be verified that the fifth term cancels the third term exactly. By using the same identity, we also find that the sixth integral cancels the second integral exactly. Finally, by using the equilibrium constraint, Eq. (17), we can write the linearized form of Eq. (4) as

$$\tilde{\mathbf{F}}(\boldsymbol{\xi}) + (1/4\pi)\psi(\theta',z',t')\nabla'\Theta(r'-r_0) = Q\int d^3\mathbf{v}'\left(\mathbf{E}^{(0)\prime} + \frac{1}{c}\mathbf{v}'\times\mathbf{B}^{(0)\prime}\right)\tilde{g}(\mathbf{r}',\mathbf{v}',t'). \quad (66)$$

Equations (55) and (66) are formally the same as would be obtained by linearizing Eqs. (3) and (4) in terms of the unprimed variables for a hypothetical problem in which the plasma–vacuum interface is a rigid cylinder of radius r_0.

Equations (55) and (66) are satisfactory linearized forms of Eqs. (3) and (4) if an equation for $\boldsymbol{\xi}$ alone is to be obtained by substituting into Eq. (66) a solution of Eq. (55) which gives $\tilde{g}$ as a functional of $\boldsymbol{\xi}$. However, it may be convenient, particularly for numerical purposes, to replace the auxiliary function $\tilde{g}$ by another auxiliary function, g, in order to eliminate the singular term from the left-hand side of Eq. (66). The function g is defined by

$$g(\mathbf{r}',\mathbf{v}',t') = \tilde{g}(\mathbf{r}',\mathbf{v}',t') + (1/4\pi p_0)\psi(\theta',z',t')\mathscr{F}(\epsilon'). \quad (67)$$

We note that the pressure function $p^{(0)}$, defined by Eq. (65), can be written correct to first order as

$$p^{(0)} = \frac{M}{3}\int d^3\mathbf{v}' v'^2 \mathscr{F}(\epsilon'), \quad (68)$$

so that the relation

$$Q\int d^3\mathbf{v}'\mathbf{E}^{(0)\prime}\mathscr{F}(\epsilon') = \nabla' p^{(0)}$$

holds. Now it is easy to verify that Eqs. (55) and (66) can be written in terms of g and $\boldsymbol{\xi}$ as

$$\begin{aligned}\left(\frac{\partial}{\partial t'} + \mathscr{L}'\right)g(\mathbf{r}',\mathbf{v}',t') &= \mathscr{F}'(\epsilon')M\mathbf{v}'\cdot\left(\frac{\partial}{\partial t'} + \mathbf{v}'\cdot\nabla'\right)\frac{\partial\boldsymbol{\xi}}{\partial t'} + \frac{\mathscr{F}(\epsilon')}{4\pi p_0}\left(\frac{\partial}{\partial t'} + \mathscr{L}'\right)\psi(\theta',z',t') \\ &= \mathscr{F}'(\epsilon')M\mathbf{v}'\cdot\left(\frac{\partial}{\partial t'} + \mathbf{v}'\cdot\nabla'\right)\frac{\partial\boldsymbol{\xi}}{\partial t'} + \frac{\mathscr{F}(\epsilon')}{4\pi p_0}\left(\frac{\partial}{\partial t'} + \mathbf{v}'\cdot\nabla'\right)\psi(\theta',z',t'), \quad (69)\end{aligned}$$

$$\tilde{\mathbf{F}}(\boldsymbol{\xi}) = Q\int d^3\mathbf{v}'\left(\mathbf{E}^{(0)\prime} + \frac{1}{c}\mathbf{v}'\times\mathbf{B}^{(0)\prime}\right)g(\mathbf{r}',\mathbf{v}',t'). \quad (70)$$

Equations (69) and (70) have been used for numerical studies of the stability of a near-theta-pinch equilibrium.[4] It is interesting to note that the singular term in the operator on the left-hand side of Eq. (66) prevents that operator from having a nonsingular eigenfunction with a nonzero eigenvalue unless ψ vanishes identically or the eigenfunction vanishes for $r = r_0$.

IV. BOUNDARY CONDITIONS

At the surface of the perfectly conducting cylindrical boundary at $r = R$, the only condition to be imposed is continuity of the normal component of **B**. Since **B** vanishes within the conductor, the condition is

$$\hat{\mathbf{r}}\cdot\mathbf{B}|_{r=R} = 0, \quad (71)$$

where **B** is the vacuum field. Since $\mathbf{B}^{(0)}$ is perpendicular to $\hat{\mathbf{r}}$, to first order the condition is

$$\hat{\mathbf{r}}\cdot\mathbf{B}^{(1)}|_{r=R} = 0.$$

However, since $\hat{\mathbf{r}}\cdot\mathbf{B}^{(1)}$ is itself of first order, the condition at the conductor also can be written correct to first order as

$$\hat{\mathbf{r}}\cdot\mathbf{B}^{(1)}|_{r'=R} = 0. \quad (72)$$

The conducting surface is not a circular cylinder in terms of the primed variables. Nevertheless, Eq. (72) tells us that the continuity condition may be applied to first order as if the conducting surface were a circular cylinder of radius R in terms of the primed variables.

At the perturbed plasma-vacuum interface, there are two conditions to be imposed to first order:

(a) continuity of the normal component of **B**, and

(b) vanishing of the normal component of $\nabla\times\mathbf{B}$ within the plasma column at $r = \rho$.

The second of these conditions is equivalent to requiring that there be no net normal current density within the plasma at the boundary. The reason for this requirement is the following. The electrons contribute no normal current density at any given point on the boundary, with respect to a frame of reference moving with the instantaneous velocity of that point, because they neither enter nor leave the boundary. The ions contribute no normal current density at any given point on the boundary, again with respect to a frame of reference moving with the instantaneous velocity of that point, because they are reflected elastically with respect to the moving frame of reference by the impulsive electric field. Thus there is also no *net* normal current density with respect to the moving frame of reference. Because of the assumed charge neutrality within the plasma, this implies that there is no net normal current density within the plasma at the boundary with respect to any Galilean frame of reference, including the laboratory frame.

In order to apply the conditions at the plasma boundary, we need an expression for the unit vector normal to the perturbed boundary. The cylindrical equilibrium boundary is defined by

$$\hat{\mathbf{r}}(\mathbf{r}_e)\cdot\mathbf{r}_e = r_0, \quad (73)$$

where $\mathbf{r}_e$ is the position vector to a point on the equilibrium boundary. (This equation assumes that the origin of the coordinate system lies on the symmetry axis of the equilibrium configuration.) Let **r** be the point on the perturbed boundary associated with $\mathbf{r}_e$. According to the discussion preceding Eq. (49), **r** and $\mathbf{r}_e$ are related to first order by

$$\mathbf{r} = \mathbf{r}_e + \boldsymbol{\xi}(\mathbf{r}_e),$$

which, again to first order, implies

$$\mathbf{r}_e = \mathbf{r} - \boldsymbol{\xi}(\mathbf{r}). \tag{74}$$

(The function $\boldsymbol{\xi}$ is to be evaluated *inside* the plasma column.) Substituting Eq. (74) into Eq. (73), we obtain to first order

$$\begin{aligned} r_0 &= \{\hat{\mathbf{r}}[\mathbf{r} - \boldsymbol{\xi}(\mathbf{r})]\}\cdot[\mathbf{r} - \boldsymbol{\xi}(\mathbf{r})] \\ &= [\hat{\mathbf{r}}(\mathbf{r}) - \boldsymbol{\xi}(\mathbf{r})\cdot\nabla\hat{\mathbf{r}}(\mathbf{r})]\cdot[\mathbf{r} - \boldsymbol{\xi}(\mathbf{r})] \\ &= \hat{\mathbf{r}}(\mathbf{r})\cdot[\mathbf{r} - \boldsymbol{\xi}(\mathbf{r})]. \end{aligned}$$

Therefore, the unit normal to the perturbed boundary is in the direction of $\nabla[r - \xi_r(\mathbf{r})]$. Letting the θ' and z' dependence of ξ_r be contained in a factor $\exp[(i(m\theta' + kz')]$, as discussed following Eq. (42), we then can write the unit normal correct to first order as

$$\hat{\mathbf{n}} = \hat{\mathbf{r}} - \mathbf{n}^{(1)}, \tag{75a}$$

where

$$\mathbf{n}^{(1)} = [(im/r_0)\hat{\boldsymbol{\theta}} + ik\hat{\mathbf{z}}]\xi_r(\mathbf{r}')|_{r'=r_0} \tag{75b}$$

The normal component of $\mathbf{B}$ now can be written correct to first order as

$$\begin{aligned} \hat{\mathbf{n}}\cdot\mathbf{B} &= (\hat{\mathbf{r}} - \mathbf{n}^{(1)})\cdot(\mathbf{B}^{(0)} + \mathbf{B}^{(1)}) \\ &= \hat{\mathbf{r}}\cdot\mathbf{B}^{(1)} - \mathbf{n}^{(1)}\cdot\mathbf{B}^{(0)\prime} \\ &= \hat{\mathbf{r}}\cdot\nabla'\times(\boldsymbol{\xi}\times\mathbf{B}^{(0)\prime}) - \mathbf{n}^{(1)}\cdot\mathbf{B}^{(0)\prime}, \end{aligned} \tag{76}$$

from which it is straightforward to verify that $\hat{\mathbf{n}}\cdot\mathbf{B}$ vanishes to first order inside the plasma at the plasma–vacuum interface. Then using Eq. (76) to calculate $\hat{\mathbf{n}}\cdot\mathbf{B}$ to first order outside the plasma at the plasma–vacuum interface, we find that continuity of $\hat{\mathbf{n}}\cdot\mathbf{B}$ is equivalent to continuity of ξ_r,

$$\xi_r(\mathbf{r}')|_{r'=r_0} = \xi_r(\mathbf{r}')|_{r'=r_0^+}. \tag{77}$$

We can write the normal component of $\nabla\times\mathbf{B}$ to first order inside the plasma as

$$\begin{aligned} \hat{\mathbf{n}}\cdot\nabla\times\mathbf{B} &= (\hat{\mathbf{r}} - \mathbf{n}^{(1)})\cdot\nabla\times(\mathbf{B}^{(0)} + \mathbf{B}^{(1)}) \\ &= (\hat{\mathbf{r}} - \mathbf{n}^{(1)})\cdot\nabla\times\mathbf{B}^{(1)} \\ &= \hat{\mathbf{r}}\cdot\nabla'\times\mathbf{B}^{(1)}. \end{aligned}$$

Therefore, the condition of no net normal current can be used to first order in the form

$$\hat{\mathbf{r}}\cdot\nabla'\times\mathbf{B}^{(1)}|_{r'=r_0} = 0. \tag{78}$$

In conclusion, we note that Eq. (66) or (70) is a statement of pressure balance, and it must be satisfied everywhere, including across the plasma boundary. However, in contrast to the situation in ideal magnetohydrodynamics, the equation cannot be used to provide an *a priori* boundary condition because g and $\tilde{g}$ (like $f^{(1)}$) are discontinuous at the boundary. Thus, the term involving $\mathbf{E}^{(0)\prime}$ cannot be evaluated *a priori*.

In conclusion, we note that Eq. (66) or (70) is a statement of pressure balance, and it must be satisfied everywhere, including across the plasma boundary. However, in contrast to the situation in ideal magnetohydrodynamics, the equation cannot be used to provide an *a priori* boundary condition because g and $\tilde{g}$ (like $f^{(1)}$) are discontinuous at the boundary. Thus, the term involving $\mathbf{E}^{(0)\prime}$ cannot be evaluated *a priori*.

Note added in proof: We assume that electrons neither enter nor leave the plasma-vacuum interface; electrons on the plasma boundary remain on the boundary forever. As is shown later in the paper, this assumption is closely related to a requirement that there be no electrical current normal to the plasma-vacuum interface. The assumption allows us to express the displacement of the boundary in terms of the perturbation vector potential. [See. Eqs. (41) and (49).] If plasma motions which violate this assumption are possible within the context of the sharp-boundary Vlasov-fluid model, it is not known how to describe them.

ACKNOWLEDGMENT

This work was supported by the United States Energy Research and Development Administration.

[1]R.R. Bartsch, C.J. Buchenauer, E.L. Cantrell, J.N. Downing, B.L. Freeman, K.B. Freese, R.F. Gribble, I. Henins, A.R. Jacobson, F.C. Jahoda, T.R. Jarboe, R. Kristal, G. Miller, W.E. Quinn, F.L. Ribe, A.R. Sherwood, R.E. Siemon, and K.S. Thomas, *Proceedings of the Sixth International Conference on Plasma Physics and Controlled Nuclear Fusion Research*, Berchtesgaden, 1976 (IAEA, Vienna, 1977) [Nucl. Fusion Suppl. (1977)], Vol. II, p. 193; F. Herrnegger, M. Kaufmann, D. Lortz, J. Neuhauser, J. Nührenberg, W. Schneider, and G. Schramm, *ibid.*, Vol. II, p. 183.

[2]D.A. Baker, L.C. Burkhardt, J.N. DiMarco, A. Haberstich, R.L. Hagenson, R.B. Howell, H.J. Karr, and A.E. Schofield, *Proceedings of the Sixth International Conference on Plasma Physics and Controlled Nuclear Fusion Research*, Berchtesgaden, 1976 (IAEA, Vienna, 1977) [Nucl. Fusion Suppl. (1977)], Vol. I, p. 419; C.W. Gowers, D.C. Robinson, A. Sykes, A. J.L. Verhage, J.A. Wesson, M.R.C. Watts, and H.A.B. Bodin, *ibid.*, Vol. I, p. 429; A. Buffa, S. Costa, R. Giannella, G. Malesani, G.F. Nalesso, and S. Ortolani, *ibid.*, Vol. I, p. 447.

[3]J.P. Freidberg, Phys. Fluids **15**, 1102 (1972).

[4]H.R. Lewis and L. Turner, Nucl. Fusion **16**, 993 (1976).

[5]L. Turner, Phys. Fluids **20**, 654, 662 (1977).

Semiclassical quantization of nonseparable systems

Abraham Klein and Ching-teh Li

Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19174
(Received 8 May 1978; revised manuscript received 14 August 1978)

The problem of semiclassical quantization of nonseparable systems with a finite number of degrees of freedom is studied within the framework of Heisenberg matrix mechanics, in extension of previous work on one-dimensional systems. The relationship between the quantum theory and multiply-periodic classical motions is derived anew. A suitably averaged Lagrangian provides a variational basis not only for the Fourier components of the semiclassical equations of motion, but also for the general definition of action variables. A Legendre transformation to the Hamiltonian verifies that these have been properly chosen and therefore provide a basis for the quantization of nonseparable systems. The problem of connection formulas is discussed by a method integral to the present approach. The action variables are shown to be adiabatic invariants of the classical system. An elementary application of the method is given. The methods of this paper are applicable to nondegenerate systems only.

I. INTRODUCTION

This paper is intended as an extension of previous studies[1,2] of the WKB approximation within the framework of Heisenberg matrix mechanics. Previous work was confined to bound systems with spectra specified by a single quantum number, but included some field theories with soliton solutions.[2] In the present work we study nonseparable problems with bound spectra specified by two or more quantum numbers.

Such problems have been attacked with considerable vigor in recent years, mainly through the vehicle of various representations of the Feynman path integral. We cite only some representative examples of a growing literature[3-7] from which the interested reader can start his own investigation. The essence of a full solution for the bound state problem was first provided within the framework of the many-body Schrodinger equation[8-10] and later extended to reaction processes.[11,12]

Here we present a formally complete solution for nondegenerate systems within the framework of Heisenberg matrix mechanics. The approach is a variant of the one used in our previous work: The matrix elements of the equations of motion are studied in the large quantum number limit and seen to yield equations for the Fourier components of the dynamical variables appropriate to the description of a multiply-periodic system. The correspondence of matrix element to Fourier component can be chosen in such a way as to include the first quantum corrections. The semiclassical equations are then associated with a variational principle in which the Lagrangian is averaged independently over each period of the system, in imitation of the double-timing technique of ever widening application in nonlinear problems.[13] These developments are described in Sec. II.

In Sec. III, we seize upon the properties of the average Lagrangian and of the equations of motion to suggest a formally exact definition of a complete set of action variables for the system studied. This definition is ultimately seen to be physically obvious, and indeed a Legendre transformation to the Hamiltonian establishes that the definition is justified.[14] The dual significance of the calculations, that they are superficially classical but also represent quantum mechanics through the first quantum corrections, immediately yields a formally complete set of WKB quantization conditions.

In Sec. IV, the discussion of connection formulas given in our previous work[1,2] is both simplified and extended. Our aim is to provide a treatment integral to the whole approach. We show by example how this may be done.

In Sec. IV we prove that the action variables defined in Sec. III are classical adiabatic invariants, adapting a method of Witham.[13] Finally in Sec. VI we record an elementary application of the method implied by our approach, which is to base all calculations on multiple Fourier series. An Appendix is included on the connection formula for a particle confined to a box.

II. SEMICLASSICAL LIMIT AND AVERAGE VARIATIONAL PRINCIPLE

We study a system with N degrees of freedom described by the Lagrangian

$$L(x^{(1)},...x^{(N)},\dot{x}^{(1)},...,\dot{x}^{(N)}) \equiv L(\mathbf{x},\dot{\mathbf{x}}) = \tfrac{1}{2}\dot{\mathbf{x}}\cdot\dot{\mathbf{x}} - V(\mathbf{x}), \tag{2.1}$$

where the kinetic energy is written as a scalar product. With $\hbar = 1$, we have the commutation relations

$$[x^{(i)},p^{(j)}] = i\delta^{ij}, \tag{2.2}$$

or

$$[x^{(i)},[H,x^{(j)}]] = \delta^{ij}, \tag{2.3}$$

with

$$H = \mathbf{p}\cdot\dot{\mathbf{x}} - L. \tag{2.4}$$

To be definite we assume that the quantum-mechanical system described by (2.1)–(2.4) has only bound states[15] labeled by integers $n_1...n_N$. Thus

$$H|n_1\cdots n_N\rangle = E(n_1,...n_N)|n_1\cdots n_N\rangle. \tag{2.5}$$

The parital derivatives

$$\omega_i = \frac{\partial E}{\partial n_i}(\mathbf{n}) \tag{2.6}$$

[a]Supported in part by the U.S. Department of Energy.

coincide, for large **n** with the frequencies of the associated classical system. We have shown previously for $N = 1$ that the Heisenberg equation of motion

$$[E(n) - E(n+k)]^2\langle n|x|n+k\rangle = \left\langle n\left|\frac{\partial V}{\partial x}\right|n+k\right\rangle, \tag{2.7}$$

when expanded for large n in powers of n^{-1} about the reference matrix element

$$x_k(n) \equiv \langle n - \tfrac{1}{2}k\,|x|n + \tfrac{1}{2}k\rangle, \tag{2.8}$$

reduces in the *first two orders* to the equation

$$(k\omega)^2 x_k(n) = \left(\frac{\partial V}{\partial x}\right)_k \equiv -[F(x)]_k, \tag{2.9}$$

where ω is given by (2.6); if $x_k(n)$ is identified as the Fourier component of a classical dynamical variable,

$$x(t,n) = \sum_{-\infty}^{\infty} x_k(n)e^{-ik\omega t}, \tag{2.10}$$

then in the right-hand side F_k is the Fourier component of the force. Thus (2.9) is a Fourier component of Newton's law.

The same elementary procedure applies without change to the multidimensional case. Thus in terms of the definitions

$$x_{\mathbf{k}}^{(i)}(\mathbf{n}) \equiv \langle n_1 - \tfrac{1}{2}k_1,\dots n_N - \tfrac{1}{2}k_N|x^{(i)}|n_1 + \tfrac{1}{2}k_1,\dots,n_N + \tfrac{1}{2}k_N\rangle \tag{2.11}$$

we arrive at the semiclassical equations

$$(\mathbf{k}\cdot\boldsymbol{\omega})^2 x_{\mathbf{k}}^{(i)}(\mathbf{n}) = \left(\frac{\partial V}{\partial x^{(i)}}\right)_{\mathbf{k}}. \tag{2.12}$$

The most natural way to view Eq. (2.12) is as the N-dimensional Fourier transform of a function of N "times"

$$x^{(i)}(\omega_1 t_1,\dots\omega_N t_N,\mathbf{n}) = \sum_{k_1\cdots k_N} x_{\mathbf{k}}^{(i)}(\mathbf{n})\exp\Big(-\sum_j k_j\omega_j t_j\Big), \tag{2.13}$$

or equivalently

$$x^{(i)}(\boldsymbol{\theta},\mathbf{n}) = \sum_{\mathbf{k}} x_{\mathbf{k}}^{(i)}(\mathbf{n})\exp(-i\mathbf{k}\cdot\boldsymbol{\theta}). \tag{2.14}$$

Newton's law follows if we subsequently equate all the times [see Eq. (2.19) below]. Adhering in the succeeding discussion to the multiple-time formalism emphasizes our exclusive interest in multiply-periodic solutions.

Taking Eq. (2.12) as fundamental, we show that there is an associated variational principle[16] which will prove useful in the sequel. We employ the time average of the Lagrangian with respect to N phases,

$$\bar{L} = \frac{1}{(2\pi)^N}\int_0^{2\pi} d\,\boldsymbol{\theta} L\,[\mathbf{x}(\boldsymbol{\theta}),\boldsymbol{\omega}\cdot\nabla_\theta\mathbf{x}(\boldsymbol{\theta})], \tag{2.15}$$

$$\text{where, e.g., } \boldsymbol{\omega}\cdot\nabla_\theta = \sum_{i=1}^{N}\frac{\partial}{\partial t_i} \equiv \frac{d}{dt} \tag{2.16}$$

on functions of the form (2.14). The requirement

$$\delta\bar{L} = 0 \tag{2.17}$$

for all trial functions and therefore variations satisfying the periodicity condition

$$\delta x^{(i)}(\theta_1\cdots\theta_j + 2\pi,\dots\theta_N) = \delta x^{(i)}(\theta_1,\dots\theta_N), \quad \text{all } j, \tag{2.18}$$

yields, by the standard manipulations,

$$(\boldsymbol{\omega}\cdot\nabla_\theta)\frac{\partial L}{\partial \dot{x}^{(i)}} = \frac{\partial L}{\partial x^{(i)}}, \tag{2.19}$$

where

$$\dot{x}^{(i)} = (\boldsymbol{\omega}\cdot\nabla_\theta)x^{(i)}. \tag{2.20}$$

Equation (2.19) is the configuration space transform of (2.12). Altenatively, we can insert (2.14) directly into (2.15) and carry out the average directly. It then follows that the equations

$$\frac{\partial\bar{L}}{\partial \mathbf{x}_{-\mathbf{k}}^{(i)}} = 0 \tag{2.21}$$

accord with Eqs. (2.12).

To prove this, note that $\bar{L}$, by definition, is the constant term in the multiple Fourier series for the Lagrangian L. Thus, from (2.1) we have

$$\bar{L} = \tfrac{1}{2}\sum_{\mathbf{k}}(\boldsymbol{\omega}\cdot\mathbf{k})^2\mathbf{x}_{\mathbf{k}}\cdot\mathbf{x}_{-\mathbf{k}} - V_0. \tag{2.22}$$

Here $V_0 \equiv \bar{V}$ is the constant term in the Fourier series for the interaction. Therefore,

$$\frac{\partial\bar{L}}{\partial x_{-\mathbf{k}}^{(i)}} = (\boldsymbol{\omega}\cdot\mathbf{k})^2 x_{\mathbf{k}}^{(i)} - \frac{\partial V_0}{\partial x_{-\mathbf{k}}^{(i)}}. \tag{2.23}$$

But

$$\begin{aligned}\frac{\partial V_0}{\partial x_{-\mathbf{k}}^{(i)}} &= \frac{\partial}{\partial x_{-\mathbf{k}}^{(i)}}\frac{1}{(2\pi)^N}\int d\,\boldsymbol{\theta}\,V(x^{(i)})\\ &= \frac{\partial}{\partial x^{(i)}_{\mathbf{k}}}\frac{1}{(2\pi)^N}\int d\,\boldsymbol{\theta}\,V\Big(\sum_{\mathbf{p}} x_{\mathbf{p}}^{(i)}e^{-i\mathbf{p}\cdot\boldsymbol{\theta}}\Big)\\ &= \frac{1}{(2\pi)^N}\int d\,\boldsymbol{\theta}e^{i\mathbf{k}\cdot\boldsymbol{\theta}}\frac{\partial V}{\partial x^{(i)}} \equiv \left(\frac{\partial V}{\partial x^i}\right)_{\mathbf{k}}.\end{aligned} \tag{2.24}$$

By inserting (2.24) into (2.23), we see that (2.21) is identical to (2.12).

III. EXACT ACTION VARIABLES

Further analysis is based on the equations of motion (2.12) and (2.21), (the variational principle) and their interplay. It is essential here to recognize a number of properties of the equations of motion (2.12) which are completely evident as soon as one writes them out in detail for a representative interaction. One may, for example, choose for V a general multinomial for which the present approach is particularly appropriate.

There are then two natural ways in which to view the equations of motion. The first commends itself when interest focuses on computing anharmonic corrections about the harmonic limit. One would then tend to specify a solution by assigning values to N of the Fourier components, most naturally to the set $\mathbf{x}_1 \equiv (x^{(1)}_{1,0\dots0}\cdots x^{(N)}_{0,\dots1})$. Since these are, in general, complex numbers we have specified $2N$ integration constants and thus a unique solution to the equations of motion. Upon solution the latter now yield the equations

$$\boldsymbol{\omega} = \boldsymbol{\omega}(\mathbf{x}_1), \tag{3.1}$$

$$\mathbf{x}_\mathbf{k} = \mathbf{x}_\mathbf{k}(\mathbf{x}_1), \quad \mathbf{k}\neq\mathbf{1}. \tag{3.2}$$

From the structure of the equations, we recognize that (i) special solutions exist with $\mathbf{x}_1$ a real vector. (ii) let $\boldsymbol{\delta}$ be the real (phase) vector

$$\boldsymbol{\delta} = (\delta_1, \ldots \delta_N). \tag{3.3}$$

Then the general solution is obtained from the special solution with real x_1 by the replacement

$$\mathbf{x}_\mathbf{k} \rightarrow \mathbf{x}_\mathbf{k} \exp(i\mathbf{k}\cdot\boldsymbol{\delta}). \tag{3.4}$$

(iii) The time averages $\overline{L}$ and $\overline{H}$ are invariant under the replacement (3.4). This trivial observation will prove essential when coupled with the time independence of H.

Since we shall be concerned in the sequel with $\overline{L}, \overline{H}$ and their relation, we henceforth consider only real $\mathbf{x}_\mathbf{k}$ satisfying

$$\mathbf{x}_\mathbf{K} = \mathbf{x}_{-\mathbf{k}}. \tag{3.5}$$

We are then ready to consider the second natural way of viewing the equations of motion which is more appropriate for current needs. For a real solution we need specify only N real constants and we do so by "choosing" the vector $\boldsymbol{\omega}$, i.e., we suppose that (3.1) and (3.2) together can be written as

$$\mathbf{x}_\mathbf{k} = \mathbf{x}_\mathbf{k}(\boldsymbol{\omega}), \quad \text{all } \mathbf{k}. \tag{3.6}$$

This excludes the harmonic limit where, of course, the amplitude is independent of frequency, but, although we shall not show this in detail, the results of the procedure to follow include that limit. Inserting (3.6) into $\overline{L}$ we have [cf. (2.22)]

$$\overline{L} = \overline{L}[\boldsymbol{\omega}, \mathbf{x}_\mathbf{k}(\boldsymbol{\omega})] = \tfrac{1}{2}\sum_\mathbf{k} (\boldsymbol{\omega}\cdot\mathbf{k})^2 \mathbf{x}_\mathbf{k}\cdot\mathbf{x}_\mathbf{k} - \overline{V}, \tag{3.7}$$

where the sum is over all independent k vectors whose components are either positive or negative integers.

We are finally ready for the essential part of the exercise. We define the action J_i by the formula

$$J_i \equiv 2\pi \frac{\delta \overline{L}}{\delta \omega_i}, \tag{3.8}$$

where the δ derivative means keep ω_j $(j\neq i)$ and $\mathbf{x}_\mathbf{k}$ fixed. The notation ∂ will then mean partial derivative keeping only the other ω_j fixed. Thus

$$\frac{\partial \overline{L}}{\partial \omega_i} = \frac{\partial \overline{L}}{\delta \omega_i} + \sum_\mathbf{k} \frac{\partial \overline{L}}{\partial \mathbf{x}_\mathbf{k}} \frac{\partial \mathbf{x}_\mathbf{k}}{\partial \omega_i} = \frac{\delta \overline{L}}{\delta \omega_i} \tag{3.9}$$

in consequence of the equations of motion (2.21). On the other hand if we calculate J_i from (3.7) and (3.8), we have

$$J_i = 2\pi \sum_\mathbf{k} k_i (\mathbf{k}\cdot\boldsymbol{\omega}) \mathbf{x}_\mathbf{k}\cdot\mathbf{x}_\mathbf{k}$$

$$= (2\pi)^{1-N} \int_0^{2\pi} d\,\boldsymbol{\theta}\; \mathbf{p}\cdot\left(\frac{\partial}{\partial \theta_i}\mathbf{x}\right). \tag{3.10}$$

If we take note of (2.16), Eq. (3.10) implies that

$$\sum_i \frac{1}{2\pi} J_i \omega_i \equiv (2\pi)^{-1} \mathbf{J}\cdot\boldsymbol{\omega} = \overline{\mathbf{p}\cdot\dot{\mathbf{x}}}. \tag{3.11}$$

Now substitute (3.11) into the time average of (2.4). This yields

$$\overline{H} = (2\pi)^{-1} \boldsymbol{\omega}\cdot\mathbf{J} - \overline{L}(\boldsymbol{\omega}). \tag{3.12}$$

But (3.12) has the classic form of a Legendre transformation from the function $\overline{L}(\boldsymbol{\omega})$ to the function $\overline{H}(\mathbf{J})$. It follows that

$$\omega_i = 2\pi \frac{\partial \overline{H}(\mathbf{J})}{\partial J_i}. \tag{3.13}$$

It is well to pause a moment, before drawing the inevitable conclusion from this equation, in order to consider carefully its dual significance in the present discussion. (For it is this double meaning which renders the consequences ineluctable.) On the one hand, we have along the way quietly dropped the quantum vector $\mathbf{n}$ as a label, since the average variational principle can be studied on a purely classical level. Therefore, Eq. (3.13), if viewed at that level, identifies $\overline{H}$ as the classical Hamiltonian expressed in terms of action variables. Essential to this view is the fact that because H is a constant of the motion, it can be identified with $\overline{H}$! From Eq. (3.13), $\mathbf{J}$ is identified as an exact action vector.

At the same time (3.13) is a quantum equation. To see this, note that $\overline{L}$ bears the same relation to $\langle \mathbf{n}|L|\mathbf{n}\rangle$, the expectation value of the quantum operator in the state $|\mathbf{n}\rangle$, that Eqs. (2.12) bear to the exact Heisenberg equations of motion.[1,2] In the same spirit

$$\overline{H} = \langle \mathbf{n}|H|\mathbf{n}\rangle = E(\mathbf{n}). \tag{3.14}$$

Thus Eq. (3.13) must be identical with (2.6), which requires that we quantize the action variables J_i according to the equation

$$J_i = 2\pi(n_i + c_i) \rightarrow (n_i + c_i)h, \tag{3.15}$$

when we use $\hbar = 1$, and c_i is a constant discussed in the next section.

IV. CONNECTION FORMULAS

The purpose of this section is to derive the possible values of the quantum correction constants c_i by means which are integral to our approach. In our previous work,[1] we explained how to do this in the one-dimensional case. We shall start the present discussion with a more rigorous account of the previous argument.

Let us examine the quantity J. From (3.10), for the one-dimensional case

$$J(n) = 2\pi \sum_k k^2 \omega |\langle n - \tfrac{1}{2}k\,|x|\,n + \tfrac{1}{2}k\rangle|^2$$

$$= 2\pi(n + c). \tag{4.1}$$

The previous statement was imply (though we did not say it so simply) that the structure of the Hilbert space implies that

$$J(-\tfrac{1}{2}) = 0, \tag{4.2}$$

which further implies $c = \frac{1}{2}$. As we are (and were) well aware $c = \frac{1}{2}$ does not apply in all cases. It applies only to two-sided potentials where $V(x)$ is analytic in x. In the usual language, we must have bounded motion between two linear turning points.

We attempt to translate this language of configuration space into an equivalent formulation which yields (4.2). We assume that the exact eigenstates $|n\rangle$, and energies $E(n)$ can be mapped onto a representation

$$|n\rangle \to \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle, \tag{4.3}$$

$$H \to E(\hat{n}) = E(a^\dagger a), \tag{4.4}$$

$$[a,a^\dagger] = 1. \tag{4.5}$$

In this representation the original coordinate x is represented in normal form by the series

$$x = \sum_{q_1 q_2} (a^\dagger)^{q_1} a^{q_2} A_{q_1 q_2}, \tag{4.6}$$

or, more usefully for our purposes, the equivalent form

$$x = \sum_{\lambda,q} [B_{\lambda,q}(a^\dagger a)^\lambda a^q + \text{h.c.}]$$

$$\equiv \alpha + \alpha^\dagger \tag{4.7}$$

Now evaluate ($k > 0$), using (4.2), (4.5)

$$\langle n|x|n+k\rangle$$

$$= \langle n|\alpha|n+k\rangle$$

$$= \sum_\lambda B_{\lambda,k} n^\lambda [(n+1)(n+2)\cdots(n+k)]^{1/2}. \tag{4.8}$$

Replacing $n \to n - \frac{1}{2}k$ and comparing with (4.1), we see that (4.2) is correct, provided that k is odd in (4.8).

The restriction to k being odd is possible only if the original potential $V(x)$ is symmetric under reflection. For a nonsymmetric potential we must examine the hypotheses (4.3)–(4.7). The fault can be only with the representation (4.7). This raises a question of great interest, but one outside the scope of the present investigation (see next paragraph, however). First we suggest another mode of thought which may suffice for present purposes. Let us suppose that in the "correct" version of (4.7) the factor $(a^\dagger a)^\lambda$ is replaced by some function $f_\lambda(a^\dagger a) = f_l(\hat{n})$ which is at least finite at $n = -\frac{1}{2}$. Then we can still prove that $x_k(n = -\frac{1}{2}) = 0$ for odd k, as we have just done. In particular $x_1(-\frac{1}{2}) = 0$. Now consider Eq. (3.2). Since we are dealing, in the usual terminology, with an autonomous system, i.e., one with no external sources, if $x_1(-\frac{1}{2}) = 0$, then $x_k(-\frac{1}{2})$, $k \neq 1$, must also be zero *from the equation of motion* which they jointly satisfy.

To produce a more rigorous discussion than the one just given, we would have to start with the canonical formalism from the first, defining the Lagrangian $L(x,\dot{x})$ as the form

$$L(x,\dot{x}), = p\dot{x} - H(p,x), \tag{4.9}$$

where $H(p,x)$ is a general Hermitian operator in p and x. If H is even in p and x, the discussion based on (4.3)–(4.7) is applicable, with minor additions. If H is not even in p and x, we must construct a unitary transformation which brings this about, namely one which maps the successive eigenstates onto states which are successfully even and odd under inversion. This transformation is the one which cannot be of the form (4.7). A second step of bringing this Hamiltonian to diagonal form is the one covered by our previous remarks.

To extend these remarks, incomplete as they are, to the multidimensional case would appear to require nothing other than an extended notation, provided we insist that the $x^{(i)},p^{(i)}$ form a true representation of the canonical commutation relations, i.e., their eigenvalues range only the entire real line. Thus, in this case we expect to the quantization conditions

$$J_i = 2\pi(n_i + \tfrac{1}{2}). \tag{4.10}$$

Of course many applications of interest are not included in the previous considerations. One exception is furnished when there are infinite potential barriers. For one-sided potentials, in one dimension, where x is confined to the half-space, we have already given an elementary argument for the well-known result which replaces $(n+\frac{1}{2}) \to (n+\frac{3}{4})$. The example of a particle confined to a finite segment, where $(n+\frac{1}{2}) \to (n+1)$ is reviewed in the Appendix. A second class of examples not covered by (4.9) is important if the system separates or nearly separates in a coordinate system with cyclic (angle) variables. We shall postpone discussion of such cases to future applications (see Refs. 9 and 10, however).

V. ADIABATIC INVARIANCE

We demonstrate anew, again using a method integral to the present approach, that the action variables defined in this paper are adiabatic invariants. The method of proof is an adaptation of that given by Witham.[13] Many other approaches (and more emphasis on rigor than interests us here) may be found in the literature.[17–21]

We illustrate the approach using the classic example of the simple harmonic oscillator with variable frequency, as described by the Lagrangian

$$L = \tfrac{1}{2}\dot{x}^2 - \tfrac{1}{2}k(\epsilon t)x^2. \tag{5.1}$$

Let $\epsilon t = \tau$, $k(\tau) = k(\tau + 2\pi)$,

$$k_0 = \frac{1}{2\pi}\int_0^{2\pi} k(\tau)\,d\tau, \tag{5.2}$$

and

$$\epsilon \ll \sqrt{k_0}. \tag{5.3}$$

We seek a solution of the form[17]

$$x(t) = 2x_1(\tau)\cos\theta, \tag{5.4}$$

where

$$\frac{d\theta}{dt} = \theta_t \equiv \omega(\tau). \tag{5.5}$$

By introducing (5.4) into the equation of motion, which for solutions of this type may be written

$$\mathscr{L}_x \equiv [\epsilon\partial_\tau + \theta_t\partial_\theta]^2 x + k(\tau)x = 0, \tag{5.6}$$

and equating to zero separately the coefficients of $\cos\theta$ and $\sin\theta$, we find the equations

$$-\theta_t^2 x_1(\tau) + \epsilon^2\partial_\tau^2 x_1(\tau) = -k(\tau)x_1(\tau), \tag{5.7}$$

$$\partial_\tau[\theta_t x(\tau)] + \theta_t\partial_\tau x(\tau) = 0. \quad (5.8)$$

Equation (5.8) implies that if

$$J(\tau) \equiv 4\pi\theta_t x_1^2(\tau), \quad (5.9)$$

then

$$\frac{dJ(\tau)}{d\tau} = 0. \quad (5.10)$$

The definition in (5.9) of the (possibly) time-dependent action variable is the natural extension of the definition (3.10), but we shall deal fully with this point in the subsequent discussion.

We show next that Eqs. (5.7) and (5.10) can be derived from a variational principle which involves double averaging. We consider

$$\bar{L} = \frac{1}{(2\pi)^2}\int_0^{2\pi} d\tau \int_0^{2\pi} d\theta \,\{\tfrac{1}{2}[\epsilon\partial_\tau + \theta_t\partial_\theta)x]^2 - \tfrac{1}{2}k(\tau)x^2\}, \quad (5.11)$$

since it is implied that we consider trial functions only of the form (5.4) and variations of the form

$$\delta x = 2\delta x_1(\tau)\cos\theta - 2x_1(\tau)(\sin\theta)\delta\theta(\tau), \quad (5.12)$$

which are periodic in θ for fixed τ and conversely. This suffices to permit us to carry out the integrations by parts and to find [cf. (5.6)]

$$\delta\bar{L} = \frac{1}{(2\pi)^2}\int d\tau \int d\theta\,(-\mathscr{L}_x)\delta x = 0. \quad (5.13)$$

Alternatively, we may introduce (5.4) directly into (5.11) and carry out the average over θ. This yields

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} d\tau\{\theta_t^2 x_1^2(\tau) + \epsilon^2[\partial_\tau x_1(\tau)]^2 - k(\tau)x_1^2(\tau)\}. \quad (5.14)$$

Varying with respect to $x_1(\tau)$ yields (5.7), which reduces for the time-independent case to the usual equations of motion. Varying with respect to θ_t yields (5.10) directly. In the time-independent case, this variation does not yield an equation of motion, but reduces rather to the definition of the action variable.

Turning now to the problem studied in this paper, we consider the Lagrangian (2.1), except that buried in $V(\mathbf{x})$ is a coupling constant, K which now reverts to a function $K(\epsilon t)$ which we take as describing a period that is long compared to any of the natural periods of the system. We take $x^{(i)}(t)$ in the form

$$x^{(i)}(t) = \sum_{\mathbf{k}} x_{\mathbf{k}}^{(i)}(\tau)\exp(-i\mathbf{k}\cdot\boldsymbol{\theta}). \quad (5.15)$$

The $(N+1)$ times averaged Lagrangian finally assumes the form, analogous to (5.14),

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} d\tau\Big\{\sum_k(\boldsymbol{\theta}_t\cdot\mathbf{k})^2\mathbf{x}_{\mathbf{k}}(\tau)\cdot\mathbf{x}_{\mathbf{k}}(\tau) + \sum_k \epsilon^2\partial_\tau\mathbf{x}_{\mathbf{k}}\cdot\partial_\tau\mathbf{x}_{\mathbf{k}} - \bar{V}[\mathbf{x}_{\mathbf{k}},K(\tau)]\Big\}. \quad (5.16)$$

Varying with respect to $\mathbf{x}_{\mathbf{k}}(\tau)$ yields the generalization of (5.7). Varying with respect to $(\boldsymbol{\theta}_t)_i$ yields

$$\frac{d}{d\tau}J_i(\tau) = 0, \quad (5.17)$$

where

$$J_i(\tau) = 2\pi\frac{\delta\bar{L}}{\delta(\boldsymbol{\theta}_t)_i} = \sum_k 4\pi k_i(\mathbf{k}\cdot\boldsymbol{\theta}_t)\mathbf{x}_{\mathbf{k}}(\tau)\cdot\mathbf{x}_{\mathbf{k}}(\tau) \quad (5.18)$$

is the generalization of the previously defined action variable.

VI. AN ELEMENTARY EXAMPLE

We examine the problem of coupled oscillators as described by the Lagrangian

$$L = \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2) - \tfrac{1}{4}\kappa_1 x^4 - \tfrac{1}{4}\kappa_2 y^4 - \tfrac{1}{2}\lambda x^2y^2. \quad (6.1)$$

For $\lambda = 0$, the large eigenvalues of the quantum problem are given by the usual WKB integrals for each degree of freedom. At the same time an excellent approximation is provided by the simplest possible trial functions

$$x(t) \cong 2x_{10}\cos(\omega_1 t_1), \quad (6.2)$$

$$y(t) \cong 2y_{01}\cos(\omega_2 t_2), \quad (6.3)$$

introduced into the average variation principle (3.7). In fact we have ($\lambda = 0$)

$$\bar{L} \cong \bar{L}_0 = x_{10}^2\omega_1^2 + y_{01}^2\omega_2^2 - \tfrac{3}{2}\kappa_1 x_{10}^4 - \tfrac{3}{2}\kappa_2 y_{01}^4, \quad (6.4)$$

which yields the variational equations

$$x_{10}^2 = \omega_1^2/3\kappa_1, \quad (6.5)$$

$$y_{01}^2 = \omega_2^2/3\kappa_2. \quad (6.6)$$

Also

$$J_1 = 4\pi\omega_1 x_{10}^2 = \frac{4\pi}{3\kappa_1}\omega_1^3, \quad (6.7)$$

$$J_2 = 4\pi\omega_2 y_{01}^2 = \frac{4\pi}{3\kappa_2}\omega_2^3. \quad (6.8)$$

Thus

$$\omega_1 = 2\pi\frac{\partial E(J_1,J_2)}{\partial J_1} = \left(\frac{3\kappa_1}{4\pi}\right)^{1/3} J_1^{1/3}, \quad (6.9)$$

$$\omega_2 = \left(\frac{3\kappa_2}{4\pi}\right) J_2^{1/3}, \quad (6.10)$$

and

$$E(J_1,J_2) = \tfrac{1}{2}\left(\frac{3}{4\pi}\right)^{4/3}[\kappa_1^{1/3}J_1^{4/3} + \kappa_2^{1/3}J_2^{4/3}]. \quad (6.11)$$

The quantum replacement is $J_i \to (n_i + \tfrac{1}{2})h$. By comparison with the exact WKB integral, it can be shown that the numerical coefficients in (6.11) are within 2% of the exact coefficient and most of the discrepancy is removed by adding the third harmonic for each degree of freedom.[22–26]

If $\lambda \ll \kappa_{1,2}$ its effects can be treated in first-order perturbation. To (6.4) we add, using the same trial solutions (6.2), (6.3),

$$\bar{L}_1 = -2\lambda x_{10}^2 y_{01}^2. \quad (6.12)$$

Equations (6.5) and (6.6) become

$$x_{10}^2 = \frac{\omega_1^2}{3\kappa_1} - \frac{2\lambda}{3\kappa_1} y_{01}^2, \tag{6.13}$$

$$y_{01}^2 = \frac{\omega_2^2}{3\kappa_2} - \frac{2\lambda}{3\kappa_2} x_{10}^2. \tag{6.14}$$

On the other hand the first equalities in (6.7) and (6.8) remain valid. Repeating the previous step and always working to the first order in λ we eventually find

$$E(J_1,J_2) = \left(\frac{3}{4\pi}\right)^{4/3} \frac{1}{2}\{(\kappa_1^{1/3} J_1^{4/3} + \kappa_2^{1/3} J_2^{4/3}$$

$$+ \frac{4}{9}\frac{\lambda}{\kappa_1\kappa_2}(\kappa_1 J_1)^{2/3}(\kappa_2 J_2)^{2/3})\}. \tag{6.15}$$

The quantum replacement is again $J_i \to (n_i + \frac{1}{2})h$.

A more elaborate discussion of this and other examples must await a future occasion.

ACKNOWLEDGMENT

This work was supported in part by the U.S. Department of Energy. One of the authors (A.K.) wishes to thank Frank Halpern for a stimulating discussion.

APPENDIX

We study the problem of a particle in a one-dimensional box described by the Hamiltonian

$$H = \tfrac{1}{2}p^2 + V(x), \tag{A1}$$

with

$$V(x) = \begin{cases} 0, & |x| < a \\ V_0, & |x| \geqslant a \end{cases}. \tag{A2}$$

From the equations of motion

$$[x,H] = ip,$$

$$[p,H] = -i\frac{dV}{dx} = -i[\delta(x-a) - \delta(x+a)]V_0, \tag{A3}$$

we derive the matrix elements

$$(E_n - E_m)x_{mn} = -ip_{mn}.$$

$$(E_n - E_m)p_{mn} = -2iV_0\psi_m(a)\psi_n(a)\tfrac{1}{2}[1-(-)^{m+n}], \tag{A4}$$

where $\psi_m(a)$ is the mth eigenfunction evaluated at $x = a$.

For further progress we utilize (reluctantly) the Schrodinger equation

$$-\tfrac{1}{2}\psi_n''(x) + V_0\psi_n(x) = E_n\psi_n(x); \tag{A5}$$

we are interested in the limit $V_0 \gg E_n$. We thereby derive

$$\frac{d}{dx}[\psi_n'(x)\psi_m'(x) - 2V_0\psi_n(x)\psi_n(x)] = 0, \tag{A6}$$

or, for $x \to a + 0$,

$$2V_0\psi_n(a)\psi_m(a) = \psi_n'(a)\psi_m'(a). \tag{A7}$$

Thus (A4b) becomes

$$(E_n - E_m)p_{mn} = -i\psi_n'(a)\psi_m'(a)\tfrac{1}{2}[1-(-)^{m+n}] \tag{A8}$$

and we may now take the limit $V_0 \to \infty$.

In this limit ($p_{mn} = -p_{nm}$)

$$E_m = -\tfrac{1}{2}\sum_n p_{mn}^2 = \tfrac{1}{2}\psi_m'(a)\sum_n x_{mn}\psi_n'(a)$$

$$= \tfrac{1}{2}a[\psi_n'(a)]^2, \tag{A9}$$

combining (A4a) and (A8), utilizing the expression

$$\sum_n x_{mn}\psi_n'(a) = a\psi_n'(a), \tag{A10}$$

which follows from differentiating with respect to a the expression

$$\sum_n x_{mn}\psi_n(a) = a\psi_n(a), \tag{A11}$$

and then taking the limit $V_0 \to \infty$, $\psi_n(a) \to 0$. Inserting the "answer"

$$\psi_m'(a) = \frac{1}{\sqrt{a}}\frac{(m+1)\pi}{2a} \tag{A12}$$

verifies that

$$E_m = \frac{(m+1)^2\pi^2}{8a^2}. \tag{A13}$$

What interests us particularly and was the stimulus for the exercise above is the computation of the semiclassical phase integral. A form particularly convenient for present purposes and equivalent to (3.10) is

$$J = -\sum_k \frac{[p_k(n)]^2}{\omega(n)}, \tag{A14}$$

where (k is odd)

$$p_k(n) = \langle n - \tfrac{1}{2}k|p|n + \tfrac{1}{2}k\rangle. \tag{A15}$$

Inserting (A12) and (A13) into (A8) and evaluating for m and n given the values required by (A15), we find

$$p_k(n) = -\frac{i}{a}\frac{[(n+1)^2 - \frac{1}{4}k^2]}{k(n+1)}$$

$$\cong -\frac{i}{ak}(n+1) \quad \text{(WKB)} \tag{A16}$$

This yields for (A14), summing standard series[27]

$$J = 2\pi(n+1)\left(\sum_{k\,(\text{odd}) = -\infty}^{\infty} \frac{4}{\pi^2}\frac{1}{k^2}\right) = 2\pi(n+1). \tag{A17}$$

The conclusion is that to apply WKB to the particle in a box, we must use the formula (A17). It is a bonus that (A17) yields the exact eigenvalues.

We add a second, more elegant derivation of (A17). The general method of this paper gives in place of (A17)

$$J = \oint p\,dx = 2\pi(n+c), \tag{A18}$$

c to be determined. Consider next the same problem, except that $0 \leqslant x \leqslant a$. By reflecting the potential with respect to the origin as we did for the one-sided potential, we obtain

$$2\oint p dx = 2J = 2\pi(2n + 1 + c), \tag{A19}$$

where J is the action for the unreflected problem and the right-hand side records the fact that we want only the odd wavefunctions of the reflected potential. (A19) reads

$$J = 2\pi[n + \tfrac{1}{2}(c + 1)]. \tag{A20}$$

However, the two problems referred to in (A18) and (A19) are related by a scale transformation which, as a canonical transformation leaves the action invariant. Thus comparing (A18) and (A20), we have

$$c = \tfrac{1}{2}(1 + c) \rightarrow c = 1. \tag{A21}$$

[1]A. Klein, J. Math. Phys. **19**, 1 (1978).
[2]A. Klein and H.A. Weldon, Phys. Rev. D **17**, 1009 (1978).
[3]K.E. Freed, J. Chem. Phys. **56**, 692 (1972).
[4]R.F. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D **10**, 4114 (1974).
[5]C. DeWitt-Morette, Ann. Phys. (N.Y.) **97**, 367 (1976).
[6]S. Levit and V. Smilansky, Ann. Phys. (N.Y.) **103**, 198 (1976).
[7]S. Levit, K. Möhring, V. Smilansky, and T. Dreyfus, Weizmann Institute Report WIS-77/23-Ph.
[8]J.B. Keller, Ann. Phys. (N.Y.) **4**, 180 (1958).
[9]J.B. Keller and S.I. Rubinow, Ann. Phys. (N.Y.) **9**, 24 (1960).
[10]D.J. Vezzetti and S.I. Rubinow, Ann. Phys. (N.Y.) **35**, 373 (1965).
[11]R.A. Marcus, J. Chem. Phys. **54**, 3965 (1971).
[12]W.H. Miller, Adv. Chem. Phys. **25**, 69 (1974).
[13]G.B. Witham, *Linear and Non-Linear Waves* (Wiley, New York, 1974); A.H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973). We are particularly indebted to the first of these references.
[14]It agrees with Eq. (22) of Sec. 15 of M. Born, *The Mechanics of the Atom* (Ungar, New York, 1960). No use of the formula is made in this classic work.
[15]The quantum numbers need not be Cartesian. Even if the natural quantum numbers are not integral, we can obviously replace them by integers by a suitable mapping. The physical nature of the quantum numbers becomes essential when "connection formulas" are considered. Also the restriction to systems which have no continuum states is unnecessary. It is only required that a domain of large quantum numbers exists.
[16]Such multiple averaging is widely used in canonical perturbation theory. In the present context see Refs. 13 and 14.
[17]R.M. Kulsrud, Phys. Rev. **106**, 205 (1957).
[18]A. Lenard, Ann. Phys. (N.Y.) **6**, 261 (1959).
[19]C.S. Gardner, Phys. Rev. **115**, 791 (1959).
[20]M. Kruskal, J. Math. Phys. **3**, 806 (1962).
[21]T.P. Coffey, Jr., J. Math. Phys. **10**, 426 (1969).
[22]The very rapid convergence of these Fourier series is implied by the previous work cited in Refs. 23–26.
[23]A.Z. Capri, Lett. Nuovo Cimento **3**, 351 (1972).
[24]Francis R. Halpern, J. Math. Phys. **14**, 219 (1973); **15**, 733 (1974); **16**, 2469 (1975).
[25]C.A. Uzes, Phys. Rev. D **14**, 3362 (1976) and previous work cited here.
[26]C.T. Li, A. Klein, and F. Krejs, Phys. Rev. D **12**, 2311 (1975).
[27]I.S. Gradshteyn and I.M. Ryzhik, *Tables of Integrals, Series, and Products* (Academic, New York, 1965), p. 39.

Gribov degenracies: Coulomb gauge conditions and initial value constraints[a)]

Vincent Moncrief

Department of Physics, Yale University, New Haven, Connecticut 06520
(Received 21 July 1978)

We discuss, using suitable function spaces, several features of the Gribov degeneracies of non-Abelian gauge theory. We show that the set of degenerate transverse potentials can be expected to fill entire neighborhoods in the space of transverse potentials. Specially we show that if a transverse potential $\tilde{a}_1$ sufficiently near $\tilde{a} = 0$ has a Gribov copy $\tilde{a}_0$ then in fact there is a whole neighborhood of $\tilde{a}_0$ (in the transverse subspace) filled with Gribov copies of transverse potentials near $\tilde{a}_1$. This means that degenerate potentials can be expected to have nonvanishing measure in path integral quantization. We also show how the breakdown of the canonical technique for solving the initial value constraint equations can be circumvented by using a covariant, noncanonical decomposition of the space of electric fields. We prove that the constraint subset of phase space is in fact a submanifold and establish a potentially useful orthogonal decomposition of its tangent space at any point.

I. INTRODUCTION

Gribov[1,2] has recently discussed two problems associated with the quantization of non-Abelian gauge fields:

(i) degeneracies of the Coulomb gauge conditions and

(ii) breakdown of the canonical technique for solving the constraint equations.

He displayed examples of (i) by finding pairs of distinct transverse potentials which are gauge equivalent. He also showed that an orbit of the gauge group may intersect the transverse subspace nontransversally—by having one or more dimensions of tangency to the transverse space at a point of intersection. This second (local) degeneracy problem occurs whenever a certain elliptic differential operator, defined for each transverse potential, admits a nontrivial kernel. Its occurence signals the onset of problem (ii) since the canonical procedure of solving the constraints for the longitudinal part of the electric field requires inversion of this same elliptic operator. In this paper we shall discuss both problems (i) and (ii) within the setting of suitably chosen function spaces (the weighted Sobolev spaces of Nirenberg and Walker[3] and Cantor[4-6]). These spaces have already been used by Cantor[6] and by Choquet-Bruhat, Fischer, and Marsden[7] to solve several outstanding problems in general relativity.

In path integral quantization the Gribov degeneracies are probably important only if the degenerate potentials fill out some open set in the space of transverse potentials. Only then could one reasonably expect them to have nonzero measure in the functional integration. This point has been emphasized by MacDowell[8] whose remarks motivated much of the present work. We show that if the classical vacuum (zero potential) or any other transverse potential $\tilde{a}_1$ sufficiently near the vacuum is gauge equivalent to some distinct transverse $\tilde{a}_0$, then there is a whole neighborhood of $\tilde{a}_0$ (in the transverse subspace) filled with Gribov copies of transverse potentials near $\tilde{a}_1$. The particular examples discussed by Gribov do not suffice to prove the existence of such a neighborhood of degenerate potentials since Gribov's potentials $\tilde{a}_0$ decay too slowly to lie in the function spaces considered here. Lacking a particular example, we can nevertheless conclude that either such degenerate pairs $(\tilde{a}_0,\tilde{a}_1)$ do occur, in which case whole neighborhoods of Gribov copies occur, or else there is a whole neighborhood of transverse potentials near the classical vacuum which admit no Gribov copies at all (even outside the neighborhood).

The difficulty in solving the constraints [problem (ii) above] is related to the Coulomb gauge conditions only indirectly. The standard transverse–longitudinal decomposition of the electric field $\tilde{e}$, which is motivated by considerations of the Coulomb gauge, leads immediately to the elliptic operator mentioned above. However, the solution set of the constraint equations can be discussed more conveniently with a covariant (but noncanonical) decomposition of $\tilde{e}$. We discuss this decomposition and establish the existence (for all $\tilde{a}$) of the operator $\mathscr{P}_{\tilde{a}}$, which projects $\tilde{e}$ to its covariantly transverse component (i.e., to a solution of the constraints). By extending this approach slightly we show that the solution set of the constraint equations, regarded as a subset of phase space, is in fact a submanifold. We give a potentially useful L_2 orthogonal decomposition of the tangent space of phase space at any point $(\tilde{a},\tilde{e})$ of the constraint submanifold. This splitting entails a subspace orthogonal to the constraint submanifold and two mutually orthogonal subspaces tangent to the constraint submanifold, one of which is tangent to the gauge group orbit through $(\tilde{a},\tilde{e})$. This result extends to Yang–Mills theory a decomposition given elsewhere for Einstein's equations.[9] We briefly discuss some possible extension of this geometrical approach to Yang–Mills theory.

Our discussion of gauge degeneracies is limited to the Coulomb gauge conditions as originally discussed by Gribov. A recent paper by Singer[10] shows that, at least for those potentials which can be conformally mapped to S^3, there are Gribov degeneracies for any conceivable set of gauge conditions. Singer shows that there is no regular choice of gauge

a)Research supported in part by NSF Grant No. PHY 76-82353.

 0022-2488/79/040579-07$01.00

possible at all. Thus for fields with suitable asymptotic conditions one cannot simply discard the Coulomb conditions and expect an alternative set of gauge conditions to solve the Gribov problem. Other recent studies of the Gribov problem have been given by Bender, Eguchi and Pagels,[11] Peccei,[12] Jackiw, Muzinich, and Rebbi,[13] and Ademollo, Napolitano, and Sciuto.[14]

II. MATHEMATICAL PRELIMINARIES

The function spaces we shall consider are the weighted Sobolev spaces of Nirenberg and Walker[3] and Cantor.[4-6] These are Banach spaces of functions from $\mathbb{R}^m$ to $\mathbb{R}^n$ (where $m = 3$ in our case) with the norm

$$\|f\|_{p,s,\delta} = \sum_{0 \le |\alpha| \le s} \|\sigma^{|\alpha| + \delta} D^\alpha f\|_{L_p}$$

where $1 < p < \infty$, $\delta \in \mathbb{R}$, s is a nonnegative integer, and $\sigma(\mathbf{x}) = (1 + |\mathbf{x}|^2)^{1/2}$ and $\| \ \|_{L_p}$ is the usual L_p norm (on $\mathbb{R}^3$),

$$\|f\|_{L_p} = \left(\int_{\mathbb{R}^3} d^3x |f|^p \right)^{1/p}.$$

Following Cantor, we shall designate these spaces as $\mathscr{M}^p_{s,\delta}(\mathbb{R}^3,\mathbb{R}^n)$ or simply $\mathscr{M}^p_{s,\delta}$. With a suitable choice of n, we can define $\mathscr{M}^p_{s,\delta}$ spaces of potentials $\tilde{a} = \{a_i^{(a)}\}$ for any desired gauge group. Similarly we can define spaces of conjugate potentials (i.e., electric fields) $\tilde{e} = \{e^i_{(a)}\}$ and model spaces for the gauge transformations $U(\tilde{\omega})$. Some specific choices for these spaces are discussed below.

The elliptic operators that we shall encounter are all related to the Laplacian Δ (and in fact reduce to it in special cases). In three dimensions the Laplacian can be shown[4,5] to define an isomorphism from $\mathscr{M}^p_{s,\delta}$ to $\mathscr{M}^p_{s-2,\delta+2}$ for $s \geqslant 2$, $p > 3$, and $0 \leqslant \delta < 1 - 3/p$. The space $\mathscr{M}^p_{s,\delta}$ contains functions f with the asymptotic behavior

$$f \sim \frac{1}{|\mathbf{x}|^{\delta+\epsilon+3/p}}, \quad Df \sim \frac{1}{|\mathbf{x}|^{\delta+1+\epsilon+3/p}}, \ldots,$$
$$D^s f \sim \frac{1}{|\mathbf{x}|^{\delta+s+\epsilon+3/p}}$$

for any number $\epsilon > 0$. Thus for all $1 < p < \infty$ and all $\delta \geqslant 0$ the functions in $\mathscr{M}^p_{s,\delta}$ vanish at infinity.

The gauge transformations are most naturally described as cross sections of a principal fiber bundle defined over a spacelike hypersurface of spacetime with fibers which are isomorphic to the gauge group (assumed here to be compact and semisimple). The set of such cross sections with suitable differential and asymptotic properties can probably be given the structure of a (nontrivial) Banach manifold modeled on an $\mathscr{M}^p_{s,\delta}$ space. Rather than attempt such a construction we shall take a more naive approach and parametrize the gauge transformations (nonuniquely) by elements in a Banach space of (Lie algebra valued) functions by adopting the exponential from

$$U(\tilde{\omega}) = \exp(i\omega^{(a)}\theta_a),$$

where θ_a are Hermitian generators of the gauge group and $\omega^{(a)}$ are functions in a suitable $\mathscr{M}^p_{s,\delta}$ space.

We let $\mathscr{G}$ be a g-dimensional compact, semisimple Lie group and $\{\theta_a\}$, $a = 1, \ldots g$, be Hermitian generators of a unitary representation of $\mathscr{G}$.

$$[\theta_a, \theta_b] = iC^c_{ab}\theta_c.$$

The gauge potentials may be regarded as Lie algebra valued 1-forms defined over a flat, spacelike hypersurface S and written as

$$\tilde{a} = a_i^{(a)} dx^i \theta_a.$$

The gauge transformations of $\tilde{a}$ have the standard form

$$\tilde{a}' = U(\tilde{\omega})\tilde{a}U^{-1}(\tilde{\omega}) + iU(\tilde{\omega})d(U^{-1}(\tilde{\omega})),$$

where

$$U(\tilde{\omega}) = \exp(i\omega^{(a)}\theta_a)$$

and

$$d(U^{-1}(\tilde{\omega})) = [\partial_i U^{-1}(\tilde{\omega})]dx^i.$$

The electric potentials conjugate to $\tilde{a}$ can be regarded as a Lie algebra valued vector field over S and written as

$$\tilde{e} = e^{i(a)}\theta_a \frac{\partial}{\partial x^i} = e^i_{(a)}\theta_a \frac{\partial}{\partial x^i}$$

(we assume a basis for the Lie algebra in which C^a_{bc} is completely antisymmetric; the Cartan metric is δ_{ab} so that group indices can be freely raised and lowered). The covariant from of $\tilde{e}$ is

$$\tilde{e} = e_i^{(a)}\theta_a dx^i = e_{i(a)}\theta_a dx^i,$$

where $e_i^{(a)} = e^{i(a)}$ in the Cartesian spacial coordinates on $\mathbb{R}^3$ which we shall use throughout.

III. GAUGE DEGENERACIES

For a fixed choice of $s \geqslant 2$, $\delta \geqslant 0$, and $p > 3$ let $\mathscr{A}$ be the space of $\mathscr{M}^p_{s-1,\delta+1}$ gauge potentials $\tilde{a}$ over S ($\mathbb{R}^3$ with Euclidean metric δ_{ij}) and let Ω be the space of $\mathscr{M}^p_{s,\delta}$ (Lie algebra-valued) functions $\tilde{\omega}$ over S. The gauge transformations of elements $\tilde{a}$ of $\mathscr{A}$ induced by elements $\tilde{\omega}$ of Ω define a map

$$\mathscr{F}: \mathscr{A} \times \Omega \rightarrow \mathscr{A},$$
$$(\tilde{a},\tilde{\omega}) \rightarrow U(\tilde{\omega})\tilde{a}U^{-1}(\tilde{\omega}) + iU(\tilde{\omega})[dU^{-1}(\tilde{\omega})], \tag{3.1}$$

where

$$U(\tilde{\omega}) = \exp(i\tilde{\omega}). \tag{3.2}$$

Thus $\mathscr{F}(\tilde{a},\tilde{\omega})$ is just the gauge transform of $\tilde{a}$ by $U(\tilde{\omega})$. For convenience we also define for any fixed $\tilde{\omega}_0 \in \Omega$ the map

$$\mathscr{F}_{\tilde{\omega}_0}: \mathscr{A} \times \Omega \rightarrow \mathscr{A},$$
$$(\tilde{a},\tilde{\omega}) \rightarrow U(\tilde{\omega})U(\tilde{\omega}_0)\tilde{a}U^{-1}(\tilde{\omega}_0)U^{-1}(\tilde{\omega}) + iU(\tilde{\omega})U(\tilde{\omega}_0)[d(U^{-1}(\tilde{\omega}_0)U^{-1}(\tilde{\omega}))], \tag{3.3}$$

which is equivalent to

$$\mathscr{F}_{\tilde{\omega}_0}(\tilde{a},\tilde{\omega}) = \mathscr{F}(\mathscr{F}(\tilde{a},\tilde{\omega}_0),\tilde{\omega}) \tag{3.4}$$

by the group property of gauge transformations.

Now let Σ be the space of $\mathscr{M}^p_{s-2,\delta+2}$ (Lie algebra valued) functions on S and define the map

$$\delta\mathscr{F}_{\tilde{\omega}_0}: \mathscr{A} \times \Omega \rightarrow \Sigma, \quad (\tilde{a},\tilde{\omega}) \rightarrow \delta \cdot [\mathscr{F}_{\tilde{\omega}_0}(\tilde{a},\tilde{\omega})], \tag{3.5}$$

where $\delta\cdot$ is the usual Euclidean divergence operator (i.e., in Cartesian coordinates $\delta\cdot\tilde{b} = \partial_i b_i$).

Define the transverse space $\mathscr{A}^T$ by the standard Coulomb gauge conditions

$$\mathscr{A}^T = \{\tilde{a}\in\mathscr{A} \,|\, \delta\cdot\tilde{a} = 0\}. \qquad (3.6)$$

For $p > 3$, $s \geqslant 2$, and $0 \leqslant \delta \leqslant 1 - 3/p$ (which we shall henceforth assume) a decomposition result of Cantor (Theorem 3 of Ref. 4) shows that $\mathscr{A}^T$ is a closed subspace and thus a submanifold of $\mathscr{A}$. One might propose, by analogy with the Abelian case, that $\mathscr{A}$ is something like a principle fiber bundle with fibers given by the orbits of the gauge group action (3.1) and admitting $\mathscr{A}^T$ as a global cross section. Then $\mathscr{A}^T$ would be a suitable reduced configuration space for the dynamics. Gribov's examples show, however, that this picture cannot be right since they reveal some orbits which intersect the transverse space more than once. The simplest examples considered by Gribov decay too slowly to lie in the spaces considered above. He argues, however, that more rapidly decaying examples also exist and outlines a method for constructing some spherically symmetric ones.

Let $\tilde{a}_0$ and $\tilde{a}_1$ be two distinct elements of $\mathscr{A}^T$ such that there exists an $\tilde{\omega}_0\in\Omega$ for which $\tilde{a}_1 = \mathscr{F}(\tilde{a}_0,\tilde{\omega}_0)$. Thus $\tilde{a}_0$ and $\tilde{a}_1$ are by assumption gauge degenerate transverse fields, and we can also write

$$\delta\cdot(\mathscr{F}(\tilde{a}_0,\tilde{\omega}_0)) = \delta\mathscr{F}_{\tilde{\omega}_0}(\tilde{a}_0,0) = 0. \qquad (3.7)$$

We wish to determine whether there is a neighborhood $\mathscr{N}(\tilde{a}_0)$ of $\tilde{a}_0$ in $\mathscr{A}$ such that for each $\tilde{a}\in\mathscr{N}(\tilde{a}_0)$ there is a solution $\tilde{\omega}(\tilde{a})$ (near $\tilde{\omega} = 0$) of

$$\delta\mathscr{F}_{\tilde{\omega}_0}(\tilde{a},\tilde{\omega}(\tilde{a})) = 0 \qquad (3.8)$$

[i.e., whether each $\tilde{a}$ sufficiently near $\tilde{a}_0$ is gauge equivalent to a transverse potential near $\tilde{a}_1 = \mathscr{F}_{\tilde{\omega}_0}(\tilde{a}_0,0)$]. If such an $\mathscr{N}(\tilde{a}_0)$ exists then, in particular, every transverse potential in $\mathscr{N}(\tilde{a}_0)\cap\mathscr{A}^T$ is "copied" by a gauge equivalent transverse potential near $\tilde{a}_1$.

The problem formulated here is thus to determine whether the equation

$$\delta\mathscr{F}_{\tilde{\omega}_0}(\tilde{a},\tilde{\omega}) = 0 \qquad (3.9)$$

defines $\tilde{\omega}(\tilde{a})$ implicitly in some neighborhood of the particular solution $(\tilde{a}_0,\tilde{\omega}(\tilde{a}_0) = 0)$. The implicit function theorem assures the existence and uniqueness of such a function $\tilde{\omega}(\tilde{a})$ [on some neighborhood of $(\tilde{a}_0,0)\in\mathscr{A}\times\Omega$] provided the linear operator

$$\begin{aligned} &D_{\tilde{\omega}}\delta\mathscr{F}_{\tilde{\omega}_0}(\tilde{a}_0,0): \mathscr{M}^p_{s,\delta}\to\mathscr{M}^p_{s-2,\delta+2},\\ &\tilde{\omega}'\to\delta\cdot\{d\tilde{\omega}' + i[\tilde{\omega}',\mathscr{F}(\tilde{a}_0,\tilde{\omega}_0)]\}\\ &\quad = \Delta\tilde{\omega}' + i\delta\cdot([\tilde{\omega}',\tilde{a}_1]) \end{aligned} \qquad (3.10)$$

is an isomorphism. This operator is the derivative of the function $\delta\mathscr{F}_{\tilde{\omega}_0}(\tilde{a},\tilde{\omega})$ with respect to $\tilde{\omega}$ at $(\tilde{a},\tilde{\omega}) = (\tilde{a}_0,0)$ and represents the divergence of the infinitesimal gauge transformation of $\tilde{a}_1 = \mathscr{F}(\tilde{a}_0,\tilde{\omega}_0)$ induced by $\tilde{\omega}'$. It is the same operator considered by Gribov in a different connection. (Gribov argued that if $\Delta + i\delta\cdot[\ ,\tilde{a}_1]$ has a nontrivial kernel at some $\tilde{a}_1\in\mathscr{A}^T$, then the gauge group orbit through $\tilde{a}_1$ has some directions of tangency to $\mathscr{A}^T$ at $\tilde{a}_1$. This suggests a local breakdown of the Coulomb gauge conditions near $\tilde{a}_1$.)

From Cantor's result[4,5] one knowns that Δ is an isomorphism between the spaces considered. Thus, if $\tilde{a}_1 = 0$ is gauge equivalent to some nonzero $\tilde{a}_0\in\mathscr{A}^T$, there is a full neighborhood of $\tilde{a}_0$ in $\mathscr{A}^T$ of transverse degenerate potentials (i.e., gauge equivalent copies of transverse potentials near $\tilde{a}_1 = 0$). We remark, however, that the explicit example of Gribov of a transverse potential degenerate with $\tilde{a}_1 = 0$ decays too slowly to lie in the function spaces used here.

We can extend the above result by showing that the operator $\Delta + i\delta\cdot([\ ,\tilde{a}])$ is an isomorphism on some neighborhood of $\tilde{a}_1 = 0$ in $\mathscr{A}^T$. First note that for $\tilde{a}\in\mathscr{A}^T$ the coordinate expression for this operator is

$$\{\Delta\tilde{\omega} + i\delta\cdot([\tilde{\omega},\tilde{a}])\}^\alpha = \Delta\omega^\alpha - C^\alpha_{\beta\gamma}a^\gamma_i\partial_i\omega^\beta. \qquad (3.11)$$

Since Δ is a homogeneous second order elliptic operator with constant coefficients and since the coefficient of the first order term is contained in $\mathscr{M}^p_{s-1,\delta+1}$ (and therefore in $\mathscr{M}^p_{s-2,1}$ since $\delta\geqslant 0$) we can apply a result of Cantor (Theorem 2.8 for Ref. 5) to conclude that for $s > (3/p) + 2$ the operator

$$\mathscr{D}_{\tilde{a}}: \mathscr{M}^p_{s,\delta}\to\mathscr{M}^p_{s-2,\delta+2} \qquad (3.12)$$

given by

$$\mathscr{D}_{\tilde{a}} = \Delta + i\delta\cdot([\ ,\tilde{a}]) \qquad (3.13)$$

is, for each $\tilde{a}\in\mathscr{A}^T$, a continuous map with closed range and finite dimensional kernel. Furthermore, there is an $\epsilon > 0$ such that whenever

$$\|\tilde{a}\|_{p,s-1,1} < \epsilon \qquad (3.14)$$

then $\mathscr{D}_{\tilde{a}}$ is an isomorphism of $\mathscr{M}^p_{s,\delta}$ and $\mathscr{M}^p_{s-2,\delta+2}$. Another condition which ensures that $\mathscr{D}_{\tilde{a}}$ is an isomorphism is the existence of a continuous curve c on $[0,1]$ into the space of bounded linear operators between $\mathscr{M}^p_{s,\delta}$ and $\mathscr{M}^p_{s-2,\delta+2}$ such that $c(0) = \Delta$ and $c(1) = \mathscr{D}_{\tilde{a}}$ and $c(t)$ is an injection for all $t\in[0,1]$. A potential application of the second criterion would be to the curve $\mathscr{D}_{\tilde{a}(\lambda)}$ with $\tilde{a}(\lambda) = \lambda\tilde{a}$. If $\mathscr{D}_{\tilde{a}(\lambda)}$ has trivial kernel for $\lambda\in[0,1]$, then $\mathscr{D}_{\tilde{a}}$ is an isomorphism by Cantor's theorem. The first criterion assures us that, for all $\tilde{a}\in\mathscr{A}^T$ sufficiently near $\tilde{a} = 0$, $\mathscr{D}_{\tilde{a}}$ is in fact an isomorphism.

We conclude that if either $\tilde{a} = 0$ or some transverse $\tilde{a}_1$ sufficiently near $\tilde{a} = 0$ has a Gribov copy $\tilde{a}_0\in\mathscr{A}^T$ (where $\tilde{a}_0\neq\tilde{a}_1$), then in fact there is a full neighborhood $\mathscr{N}(\tilde{a}_0)$ *of* $\tilde{a}_0$ such that every element in $\mathscr{N}(\tilde{a}_0)\cap\mathscr{A}^T$ is copied by a transverse potential near $\tilde{a}_1$. To show explicitly that this phenomenon occurs, one could try to exhibit a transverse copy of $\tilde{a} = 0$ (lying in $\mathscr{M}^p_{s-1,\delta+1}$) or to show that potentials arbitrarily near $\tilde{a} = 0$ admit Gribov degeneracies. Alternatively one could try to find an $\tilde{a}_1\in\mathscr{A}^T$ which admits a Gribov copy and for which there is a continuous curve of operators $\mathscr{D}_{\tilde{a}(\lambda)}$ from $\tilde{a}(0) = 0$ to $\tilde{a}(1) = \tilde{a}_1$ which is injective for each $\lambda\in[0,1]$. Any example of this sort would extend to a full neighborhood of degenerate potentials. The degeneracies found explicitly by Gribov all involved highly symmetrical gauge transformations. To find a transverse potential in $\mathscr{M}^p_{s-1,\delta+1}$ which is degenerate with $\tilde{a} = 0$ (the simplest choice) might require a nonsymmetrical gauge transformation lying outside the scope of Gribov's ansatz.

We can summarize our results as follows. There is a neighborhood $\mathscr{Q}$ of $\tilde{a}=0$ in $\mathscr{A}^T$ such that either each gauge group orbit intersecting $\mathscr{Q}$ intersects $\mathscr{A}^T$ at one and only one point or else there is an $\tilde{a}_0 \notin \mathscr{Q}$ and a neighborhood $\mathscr{N}(\tilde{a}_0)$ of $\tilde{a}_0$ such that each $\tilde{a} \in \mathscr{N}(\tilde{a}_0) \cap \mathscr{A}^T$ is degenerate with some transverse $\tilde{a}$ near $\tilde{a}=0$. One can think of the gauge group orbits spraying out of $\mathscr{Q}$ and either

(i) never reintersecting $\mathscr{A}^T$ or else

(ii) covering a whole neighborhood $\mathscr{N}(\tilde{a}_0) \cap \mathscr{A}^T$ of some transverse $\tilde{a}_0 \notin \mathscr{Q}$.

IV. GEOMETRY OF THE CONSTRAINT SUBSET

A standard technique used to solve the constraint equations

$$\nabla_{\tilde{a}} \cdot \tilde{e} = [\partial_i e_i^{(\alpha)} - C^{\alpha}_{\beta\gamma} e_i^{(\beta)} a_i^{(\gamma)}] \Theta_\alpha$$

$$= 0 \tag{4.1}$$

is to decompose $\tilde{e}$ into a divergence free term $\tilde{e}_*$ and a gradient. One substitutes $\tilde{e}=\tilde{e}_* + d\tilde{\varphi}$ (where $\delta \cdot \tilde{e}_* = 0$) into Eq. (4.1) to get

$$\Delta \varphi^{(\alpha)} - C^{\alpha}_{\beta\gamma}(\partial_i \varphi^{(\beta)}) a_i^{(\gamma)} = C^{\alpha}_{\beta\gamma}(e_*)_i^{(\beta)} a_i^{(\gamma)} \tag{4.2}$$

and attempts to solve for $\tilde{\varphi}$. However, as Gribov showed, the operator on the left hand side is not invertible for arbitrary $\tilde{a}$. In particular, if $\tilde{a} \in \mathscr{A}^T$, then the gauge group orbit through $\tilde{a}$ may have one or more dimensions of tangency to $\mathscr{A}^T$ and this implies a nontrivial kernel for the operator in question. Such local degeneracies of the Coulomb gauge condition have also been discussed by Bender, Eguchi, and Pagels,[11] by Peccei,[12] and by Jackiw, Muzinich, and Rebbi.[13]

The above difficulty in solving the constraint arises only because of an inconvenient choice of decomposition for $\tilde{e}$. A more natural decomposition which leads to a globally valid solution of the constraints may be obtained as follows. Let $\mathscr{L}$ be the $\mathscr{M}^p_{s,\delta}$ space of Lie algebra valued functions on $\mathbb{R}^3$ with the restrictions

$$p>3, \quad \delta + 3/p \geqslant \tfrac{1}{2},$$
$$s \geqslant 3, \quad 0 \leqslant \delta < 1 - 3/p. \tag{4.3}$$

Let $\mathscr{E}$ be the space of $\mathscr{M}^p_{s-1,\delta+1}$ electric fields $\tilde{e}$ and $\mathscr{A}$ be the space of $\mathscr{M}^p_{s-1,\delta+1}$ vector potentials $\tilde{a}$ on $\mathbb{R}^3$. The conditions (4.3) are more restrictive (for $\mathscr{A}$) than those considered in Sec. II. The present restrictions require both $\tilde{a}$ and $\tilde{e}$ to decay as $1/|\mathbf{x}|^{3/2+\epsilon}$ and thus to be square integrable. Aside from mathematical convenience there is some physical justification for the added restriction. It assures the convergence of all the Poincaré group generators which thus become well-defined real valued functions on $\mathscr{A} \times \mathscr{E}$. It also assures the convergence of the Yang-Mills charges defined on the constraint subset $\mathscr{C}^{-1}(0)$ of $\mathscr{A} \times \mathscr{E}$. These charges $Q^{(\alpha)}$ are given by the Gaussian flux integrals

$$Q^{(\alpha)} = \int_{\mathbb{R}^3} d^3x (\partial_i e_i^{(\alpha)}), \tag{4.4}$$

which on $\mathscr{C}^{-1}(0)$ are expressible as

$$Q^{(\alpha)} = \int_{\mathbb{R}^3} d^3x (C^{\alpha}_{\beta\gamma} e_i^{(\beta)} a_i^{(\gamma)}). \tag{4.5}$$

These integrals converge for all $(\tilde{a},\tilde{e}) \in \mathscr{A} \times \mathscr{E}$.

For any $\tilde{a} \in \mathscr{A}$ we define

$$\mathscr{B}_{\tilde{a}}: \mathscr{M}^p_{s,\delta} \to \mathscr{M}^p_{s-1,\delta+1},$$
$$\tilde{\varphi} \to \mathscr{B}_{\tilde{a}} \cdot \tilde{\varphi} = d\tilde{\varphi} + i[\tilde{\varphi},\tilde{a}]. \tag{4.6}$$

[Lemmas (3) and (4) of Ref. 6 are particularly useful in verifying properties of the maps considered in this section.] The formal L_2 adjoint operator $\mathscr{B}^*_{\tilde{a}}$ is given by

$$\mathscr{B}^*_{\tilde{a}} \cdot \tilde{f} = -\delta \cdot \tilde{f} - i[\tilde{f} \cdot, \tilde{a}]$$

$$= [-\partial_i f_i^{(\alpha)} + C^{\alpha}_{\beta\gamma} f_i^{(\beta)} a_i^{(\gamma)}] \Theta_\alpha, \tag{4.7}$$

and we note that the constraint equations are equivalent to

$$\mathscr{B}^*_{\tilde{a}} \cdot \tilde{e} = 0. \tag{4.8}$$

We attempt to decompose an arbitrary $\tilde{e} \in \mathscr{M}^p_{s-1,\delta+1}$ as

$$\tilde{e} = \tilde{e}^T + \mathscr{B}_{\tilde{a}} \cdot \tilde{\varphi}, \tag{4.9}$$

where $\tilde{e}^T$ is *covariantly* transverse

$$\nabla_{\tilde{a}} \cdot \tilde{e}^T = -\mathscr{B}^*_{\tilde{a}} \cdot \tilde{e}^T = 0 \tag{4.10}$$

(i.e., satisfies the constraints).

Applying $\mathscr{B}^*_{\tilde{a}}$ to Eq. (4.9), we get

$$\mathscr{B}^*_{\tilde{a}} \cdot \tilde{e} = (\mathscr{B}^*_{\tilde{a}} \cdot \mathscr{B}_{\tilde{a}}) \cdot \tilde{\varphi} \tag{4.11}$$

and wish to show that

$$\Delta_{\tilde{a}} \equiv -\mathscr{B}^*_{\tilde{a}} \cdot \mathscr{B}_{\tilde{a}} \tag{4.12}$$

is an isomorphism of $\mathscr{M}^p_{s,\delta}$ and $\mathscr{M}^p_{s-2,\delta+2}$ for all $\tilde{a} \in \mathscr{M}^p_{s-1,\delta+1}$. This will ensure that

$$\mathscr{P}_{\tilde{a}} \equiv I + \mathscr{B}_{\tilde{a}} \cdot \Delta_{\tilde{a}}^{-1} \cdot \mathscr{B}^*_{\tilde{a}} \tag{4.13}$$

is a globally defined projection operator onto the (convariantly) transverse space for each $\tilde{a}$. Given such an operator, we can project an arbitrary pair $(\tilde{a},\tilde{e}) \in \mathscr{A} \times \mathscr{E}$ to a solution $(\tilde{a},\tilde{e}^T)$ of the constraints by setting

$$(\tilde{a},\tilde{e}^T) = (\tilde{a}, \mathscr{P}_{\tilde{a}} \tilde{e}). \tag{4.14}$$

To prove that $\Delta_{\tilde{a}}$ is an isomorphism for every $\tilde{a} \in \mathscr{M}^p_{s-1,\delta+1}$, we appeal to the elliptic results of Cantor.[4,5] Written explicitly, $\Delta_{\tilde{a}} \tilde{\varphi}$ is

$$\Delta_{\tilde{a}} \tilde{\varphi} = [+\Delta \varphi^{(\alpha)} - C^{\alpha}_{\beta\gamma} \partial_i (\varphi^{(\beta)} a_i^{(\gamma)})$$

$$- C^{\alpha}_{\beta\gamma} a_i^{(\gamma)} \partial_i \varphi^{(\beta)} + C^{\alpha}_{\beta\gamma} C^{\beta}_{\mu\nu} a_i^{(\gamma)} a_i^{(\nu)} \varphi^{(\mu)}] \Theta_\alpha. \tag{4.15}$$

This is a second order linear elliptic operator with a highest order term which is elliptic with constant coefficients (namely the Laplacian). The coefficients of the first order terms are contained in $\mathscr{M}^p_{s-1,\delta+1}$ and thus in $\mathscr{M}^p_{s-2,1}$ for $\delta \geqslant 0$. The coefficients of the zeroth order terms are contained in $\mathscr{M}^p_{s-2,2}$. It follows [from Theorem (1.4) of Ref. 6] that $\Delta_{\tilde{a}}\mathscr{M}^p_{s,\delta} \rightarrow \mathscr{M}^p_{s-2,\delta+2}$ is a continuous map with closed range and finite dimensional kernel. Furthermore, $\Delta_{\tilde{a}}$ will be an isomorphism if there is a continuous curve $c(\lambda)$, $\lambda \in [0,1]$, in the space of bounded linear operators from $\mathscr{M}^p_{s,\delta}$ to $\mathscr{M}^p_{s-2,\delta+2}$ for which

$$c(0) = \Delta, \quad c(1) = \Delta_{\tilde{a}} \tag{4.16}$$

and $c(\lambda)$ is an injection for all $\lambda \in [0,1]$. Here $\Delta = \Delta_0$ is the Laplacian which is known to be an isomorphism of the given spaces by Theorem (2) of Ref. 4. To display such a curve, we take, for any $\tilde{a} \in \mathscr{A}$,

$$c(\lambda) = \Delta_{\tilde{a}(\lambda)} \tag{4.17}$$

with $\tilde{a}(\lambda) = \lambda\tilde{a}$. We must show that $\Delta_{\tilde{a}}$ is injective for any $\tilde{a} \in \mathscr{A}$.

Suppose $\tilde{\varphi} \in \ker\Delta_{\tilde{a}}$. Then

$$\Delta_{\tilde{a}}\tilde{\varphi} = -\mathscr{B}^*_{\tilde{a}}\cdot\mathscr{B}_{\tilde{a}}\tilde{\varphi} = 0 \tag{4.18}$$

and thus

$$0 = \int_{\mathbf{R}^3} d^3x[\tilde{\varphi}\cdot(\mathscr{B}^*_{\tilde{a}}\cdot\mathscr{B}_{\tilde{a}}\tilde{\varphi})]$$

$$= \int_{\mathbf{R}^3} d^3x[-\delta\cdot(\tilde{\varphi}\cdot\mathscr{B}_{\tilde{a}}\tilde{\varphi}) + (\mathscr{B}_{\tilde{a}}\tilde{\varphi})\cdot(\mathscr{B}_{\tilde{a}}\tilde{\varphi})]. \tag{4.19}$$

The first term in the last integral may be reexpressed as a surface integral using Gausss' theorem. However, since $\tilde{\varphi} \in \mathscr{L}$ decays faster than $1/|\mathbf{x}|^{1/2}$ and $\mathscr{B}_{\tilde{a}}\tilde{\varphi} \in \mathscr{M}^p_{s-1,\delta+1}$ decays faster than $1/|\mathbf{x}|^{3/2}$, the vector $(\tilde{\varphi}\cdot\mathscr{B}_{\tilde{a}}\tilde{\varphi})$ decay faster than $1/|\mathbf{x}|^2$ and so has vanishing flux over the sphere at infinity. Therefore, any $\tilde{\varphi} \in \ker\Delta_{\tilde{a}}$ must satisfy

$$\mathscr{B}_{\tilde{a}}\tilde{\varphi} = 0 \tag{4.20}$$

or

$$\partial_i\varphi^{(\alpha)} - C^{\alpha}_{\beta\gamma}\varphi^{(\beta)}a^{(\gamma)}_i = 0. \tag{4.21}$$

Contracting this equation with $\varphi^{(\alpha)} = \varphi_{(\alpha)}$ and using the total antisymmetry of $C^{\alpha}_{\beta\gamma}$, we get

$$0 = \varphi_{(\alpha)}\partial_i\varphi^{(\alpha)} = \tfrac{1}{2}\partial_i(\varphi_{(\alpha)}\varphi^{(\alpha)}). \tag{4.22}$$

Thus $\varphi^{(\alpha)}\varphi_{(\alpha)}$ is a constant which, since $\tilde{\varphi} \in \mathscr{M}^p_{s,\delta}$ must vanish asymptotically and therefore everywhere. It follows that $\tilde{\varphi} = 0$ and thus that $\Delta_{\tilde{a}}$ has trivial kernel. It follows from Cantor's theorem that $\Delta_{\tilde{a}}$ is an isomorphism for all $\tilde{a}$ and thus that the projection operator $\mathscr{P}_{\tilde{a}}$ is well defined for all $\tilde{a} \in \mathscr{A}$. The decomposition (4.9) is quite analogous to one used by York[15] and O'Murchadha[16] and Cantor[5] to solve the momentum constraints in general relativity. The existence of nontrivial solutions to the equation $\mathscr{B}_{\tilde{a}}\tilde{\varphi} = 0$ is equivalent to the existence of a gauge symmetry of $\tilde{a}$ (i.e., a one-parameter family of gauge transformations which leave $\tilde{a}$ fixed).[17] The above argument shows that nontrivial generators of gauge symmetries cannot vanish asymptotically and thus cannot belong to $\mathscr{M}^p_{s,\delta}$.

We can now apply the techniques developed for general relativity by Fischer and Marsden[18-20] to show that the constraint subset is a submanifold of the phase space $\mathscr{A} \times \mathscr{E}$. Let

$$\mathscr{C}: \mathscr{M}^p_{s-1,\delta+1} \times \mathscr{M}^p_{s-1,\delta+1} \rightarrow \mathscr{M}^p_{s-2,\delta+2} \tag{4.23}$$

be the constraint map given by

$$\mathscr{C}(\tilde{a},\tilde{e}) = -\nabla_{\tilde{a}}\cdot\tilde{e} = \mathscr{B}^*_{\tilde{a}}\cdot\tilde{e}. \tag{4.24}$$

To show that the constraint subset $\mathscr{C}^{-1}(0)$ is a submanifold of $\mathscr{A} \times \mathscr{E}$ through any $(\tilde{a},\tilde{e}) \in \mathscr{C}^{-1}(0)$ we shall show that the linearized constraint [i.e., the derivative $D\mathscr{C}(\tilde{a},\tilde{e})$ of $\mathscr{C}$ at $(\tilde{a},\tilde{e})$] given by

$$D\mathscr{C}(\tilde{a},\tilde{e}): \mathscr{M}^p_{s-1,\delta+1} \times \mathscr{M}^p_{s-1,\delta+1} \rightarrow \mathscr{M}^p_{s-2,\delta+2},$$

$$(\tilde{a}',\tilde{e}') \rightarrow \mathscr{B}^*_{\tilde{a}}\cdot\tilde{e}' - i[\tilde{e}\cdot,\tilde{a}'] \tag{4.25}$$

$$= \mathscr{B}^*_{\tilde{a}}\cdot\tilde{e}' + C^{\alpha}_{\beta\gamma}e^{(\beta)}_i a'^{(\gamma)}_i\Theta_\alpha$$

is surjective and has splitting kernel (see Refs. 21 and 22). That $D\mathscr{C}(\tilde{a},\tilde{e})$ is surjective follows at once from the fact that $\Delta_{\tilde{a}} = -\mathscr{B}^*_{\tilde{a}}\cdot\mathscr{B}_{\tilde{a}}$ is an isomorphism; $\mathscr{B}^*_{\tilde{a}}$ maps the range of $\mathscr{B}_{\tilde{a}}$ onto $\mathscr{M}^p_{s-2,\delta+2}$.

To show that $D\mathscr{C}(\tilde{a},\tilde{e})$ has splitting kernel, we establish the L_2 orthogonal decomposition of the tangent space

$$T_{(\tilde{a},\tilde{e})}\mathscr{A} \times \mathscr{E} = \ker D\mathscr{C}(\tilde{a},\tilde{e})$$

$$\oplus\, \mathrm{range} D\mathscr{C}(\tilde{a},\tilde{e})^*, \tag{4.26}$$

where $D\mathscr{C}(\tilde{a},\tilde{e})^*$ is the formal L_2 adjoint of $D\mathscr{C}(\tilde{a},\tilde{e})$ given by

$$D\mathscr{C}(\tilde{a},\tilde{e})^*: \mathscr{M}^p_{s,\delta} \rightarrow \mathscr{M}^p_{s-1,\delta+1} \times \mathscr{M}^p_{s-1,\delta+1},$$

$$\tilde{\psi} \rightarrow \{+i[\tilde{\psi},\tilde{e}]; -\mathscr{B}_{\tilde{a}}\tilde{\psi}\}$$

$$= \{-C^{\alpha}_{\beta\gamma}\psi^{(\beta)}e^{(\gamma)}_i\Theta_\alpha dx^i; -[\partial_i\psi^{(\alpha)} - C^{\alpha}_{\beta\gamma}\psi^{(\beta)}a^{(\gamma)}_i]\Theta_\alpha dx^i\}. \tag{4.27}$$

Let $(\tilde{a}',\tilde{e}')$ be any element of $T_{(\tilde{a},\tilde{e})}\mathscr{A} \times \mathscr{E} \approx \mathscr{M}^p_{s-1,\delta+1} \times \mathscr{M}^p_{s-1,\delta+1}$. We shall show that there is a unique splitting

$$(\tilde{a}',\tilde{e}') = (\tilde{a}^*,\tilde{e}^*) + D\mathscr{C}(\tilde{a},\tilde{e})^*\cdot\tilde{\psi}, \tag{4.28}$$

$$(\tilde{a}^*,\tilde{e}^*) \in \ker D\mathscr{C}(\tilde{a},\tilde{e}).$$

Acting on Eq. (4.28) with $D\mathscr{C}(\tilde{a},\tilde{e})$, we get

$$D\mathscr{C}(\tilde{a},\tilde{e})\cdot(\tilde{a}',\tilde{e}') = D\mathscr{C}(\tilde{a},\tilde{e})\cdot D\mathscr{C}(\tilde{a},\tilde{e})^*\cdot\tilde{\psi}. \tag{4.29}$$

Since the left-hand side is contained in $\mathscr{M}^p_{s-2,\delta+2}$, we need to show that the operator $D\mathscr{C}(\tilde{a},\tilde{e})\cdot D\mathscr{C}(\tilde{a},\tilde{e})^*$ is an isomorphism of $\mathscr{M}^p_{s,\delta}$ and $\mathscr{M}^p_{s-2,\delta+2}$. This operator is given explicitly by

$$D\mathscr{C}(\tilde{a},\tilde{e})\cdot D\mathscr{C}(\tilde{a},\tilde{e})^*\cdot\tilde{\psi}$$

$$= +\mathscr{B}^*_{\tilde{a}}\cdot\mathscr{B}_{\tilde{a}}\tilde{\psi} - [\tilde{e}\cdot,[\tilde{\psi},\tilde{e}]]$$

$$= -\Delta_{\tilde{a}}\tilde{\psi} - C^{\alpha}_{\beta\gamma}C^{\nu}_{\alpha\mu}\psi^{(\beta)}e^{(\gamma)}_i e^{(\mu)}_i\Theta_\nu. \tag{4.30}$$

As this differs from $-\Delta_{\tilde{a}}$ only in its zeroth order term which lies in $\mathscr{M}^p_{s-2,2}$ we can argue as before using the curve of operators

$$c(\lambda) \equiv D\mathscr{C}(\tilde{a}(\lambda),\tilde{e}(\lambda))\cdot D\mathscr{C}(\tilde{a}(\lambda),\tilde{e}(\lambda))^*, \tag{4.31}$$

where

$$\lambda\in[0,1], \quad (\tilde a(\lambda),\tilde e(\lambda)) = (\lambda\tilde a,\lambda\tilde e),$$
$$c(0) = -\Delta. \tag{4.32}$$

Since $c(0)$ is an isomorphism, we need to show that $D\mathscr{C}(\tilde a,\tilde e)\cdot D\mathscr{C}(\tilde a,\tilde e)^*$ is an injection for arbitrary $(\tilde a,\tilde e)\in\mathscr{A}\times\mathscr{E}$. Assuming $\widetilde\psi$ to be in its kernel, we get

$$\int_{\mathbb{R}^3} d^3x[-\delta\cdot(\widetilde\psi\cdot\mathscr{B}_{\tilde a}\widetilde\psi) + (\mathscr{B}_{\tilde a}\widetilde\psi)\cdot(\mathscr{B}_{\tilde a}\widetilde\psi) + C^{\alpha}_{\beta\gamma}C^{\alpha}_{\nu\mu}\psi^{(\nu)}\psi^{(\beta)}e_i^{(\gamma)}e_i^{(\mu)}] = 0. \tag{4.33}$$

The first term (which converts to a surface integral) vanishes as before and the remaining terms are at least positive semi-definite. It follows that $\widetilde\psi$ must be in the kernel of $\mathscr{B}_{\tilde a}$ and thus vanish identically as we showed previously.

The L_2 orthogonality of the summands in Eq. (4.24) follows from the identity

$$\int_{\mathbb{R}^3} d^3x\langle(\tilde a',\tilde e');D\mathscr{C}(\tilde a,\tilde e)^*.\widetilde\psi\rangle$$
$$\equiv \int_{\mathbb{R}^3} d^3x\{\tilde a'\cdot i[\widetilde\psi,\tilde e] - \tilde e'\cdot\mathscr{B}_{\tilde a}\widetilde\psi\}$$
$$= \int_{\mathbb{R}^3} d^3x\{-\delta\cdot(\widetilde\psi\cdot\tilde e') + \widetilde\psi\cdot D\mathscr{C}(\tilde a,\tilde e)\cdot(\tilde a',\tilde e')\}. \tag{4.34}$$

Since $(\widetilde\psi\cdot\tilde e') = \psi_{(\alpha)}\tilde e_i'^{(\alpha)}dx^i$ decays faster than $1/|\mathbf{x}|^2$, the boundary integral vanishes giving

$$\int_{\mathbb{R}^3} d^3x\langle(\tilde a',\tilde e');D\mathscr{C}(\tilde a,\tilde e)^*.\widetilde\psi\rangle = \int_{\mathbb{R}^3} d^3x[\widetilde\psi\cdot D\mathscr{C}(\tilde a,\tilde e)\cdot(\tilde a',\tilde e')], \tag{4.35}$$

which vanishes if $(\tilde a',\tilde e')\in\ker D\mathscr{C}(\tilde a,\tilde e)$. This completes the argument that $\mathscr{C}^{-1}(0)$ is a submanifold of $\mathscr{A}\times\mathscr{E}$.

We can refine the decomposition of $T_{(\tilde a,\tilde e)}\mathscr{A}\times\mathscr{E}$ by splitting $\ker D\mathscr{C}(\tilde a,\tilde e)$ into a subspace tangent to the gauge group orbit through $(\tilde a,\tilde e)$ and a complementary (L_2 orthogonal) subspace. We wish to establish the splitting

$$\ker D\mathscr{C}(\tilde a,\tilde e) = (\ker D\mathscr{C}(\tilde a,\tilde e)\cap\ker D\mathscr{C}(\tilde a,\tilde e)\circ J) \oplus \operatorname{range} J\circ D\mathscr{C}(\tilde a,\tilde e)^* \tag{4.36}$$

for all $(\tilde a,\tilde e)\in\mathscr{C}^{-1}(0)$. Here J is the symplectic matrix

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \tag{4.37}$$

and the second summand on the right of Eq. (4.36) represents the tangent space of the gauge group orbit through $(\tilde a,\tilde e)$. This identification follows from the observation that

$$J\circ D\mathscr{C}(\tilde a,\tilde e)^*\cdot\widetilde\psi = (+\mathscr{B}_{\tilde a}\widetilde\psi; i[\widetilde\psi,\tilde e]), \tag{4.38}$$

where the two slots are just the infinitesimal gauge transformations of $\tilde a$ and $\tilde e$ respectively generated by $\widetilde\psi$. It is straightforward to check that

$$\operatorname{range} J\circ D\mathscr{C}(\tilde a,\tilde e)^* \subset \ker D\mathscr{C}(\tilde a,\tilde e) \tag{4.39}$$

for all $(\tilde a,\tilde e)\in\mathscr{C}^{-1}(0)$. This is simply the statement that infinitesimal gauge transformations satisfy the linearized constraints.

Suppose that $(\tilde a^*,\tilde e^*)\in\ker D\mathscr{C}(\tilde a,\tilde e)$. We wish to show that there is a unique $\widetilde\psi\in\mathscr{M}^p_{s,\delta}$ and a unique $(\tilde a^{**},\tilde e^{**})\in(\ker D\mathscr{C}(\tilde a,\tilde e)\cap\ker D\mathscr{C}(\tilde a,\tilde e)\circ J)$ such that

$$\begin{pmatrix}\tilde a^*\\ \tilde e^*\end{pmatrix} = \begin{pmatrix}\tilde a^{**}\\ \tilde e^{**}\end{pmatrix} + J\circ D\mathscr{C}(\tilde a,\tilde e)^*\cdot\widetilde\psi, \tag{4.40}$$

Applying $D\mathscr{C}(\tilde a,\tilde e)\circ J$ to this equation gives

$$[D\mathscr{C}(\tilde a,\tilde e)\circ J]\cdot(\tilde a^*,\tilde e^*) = -D\mathscr{C}(\tilde a,\tilde e)\cdot D\mathscr{C}(\tilde a,\tilde e)^*\cdot\widetilde\psi, \tag{4.41}$$

which we must solve for $\widetilde\psi$. Since we have taken both $\mathscr{A}$ and $\mathscr{E}\approx\mathscr{M}^p_{s-1,\delta+1}$ the source term on the left lies in $\mathscr{M}^p_{s-2,\delta+2}$. Furthermore, we have already shown the $D\mathscr{C}(\tilde a,\tilde e)\cdot D\mathscr{C}(\tilde a,\tilde e)^*$ is an isomorphism of $\mathscr{M}^p_{s-2,\delta+2}$ and $\mathscr{M}^p_{s,\delta}$ so a unique solution $\widetilde\psi$ always exists. The L_2 orthogonality of the summands in Eq. (4.36) follows as before since

$$\int_{\mathbb{R}^3} d^3x\{\langle(\tilde a^*,\tilde e^*);J\circ D\mathscr{C}(\tilde a,\tilde e)^*\cdot\widetilde\psi\rangle\}$$
$$= \int_{\mathbb{R}^3} d^3x\{+\delta\cdot(\widetilde\psi\cdot\tilde a^*) - \widetilde\psi\cdot[(D\mathscr{C}(\tilde a,\tilde e)\circ J)\cdot(\tilde a^*,\tilde e^*)]\}. \tag{4.42}$$

The first term on the right vanishes by Gauss' theorem and the fact that $(\widetilde\psi\cdot\tilde a^*) = \psi_{(\alpha)}a_i^{(\alpha)}dx^i$ vanishes faster than $1/|\mathbf{x}|^2$. The second term on the right vanishes whenever $(\tilde a^*,\tilde e^*)\in\ker D\mathscr{C}(\tilde a,\tilde e)\circ J$ proving the desired result.

We can summarize the two decomposition results for $\mathscr{A}\times\mathscr{E} = \mathscr{M}^p_{s-1,\delta+1}\times\mathscr{M}^p_{s-1,\delta+1}$ with the formula

$$T_{(\tilde a,\tilde e)}\mathscr{A}\times\mathscr{E} = [\ker D\mathscr{C}(\tilde a,\tilde e)\cap\ker D\mathscr{C}(\tilde a,\tilde e)\circ J] \oplus \operatorname{range} J\circ D\mathscr{C}(\tilde a,\tilde e)^* \oplus \operatorname{range} D\mathscr{C}(\tilde a,\tilde e)^* \tag{4.43}$$

for any $(\tilde a,\tilde e)\in\mathscr{C}^{-1}(0)$.

The three summands are L_2 orthogonal; the second summand represents the tangent space to the gauge group orbit while the first represents the orbit's orthogonal complement in $\ker D\mathscr{C}(\tilde a,\tilde e)$. The corresponding decomposition for gravitational perturbations (on compact hypersurfaces) was given in Ref. 9.

The gauge transformations $U(\tilde\omega)$ act on $\mathscr{A}\times\mathscr{E}$ by sending

$$\tilde a\to U(\tilde\omega)\tilde a U^{-1}(\tilde\omega) + iU(\tilde\omega)[dU^{-1}(\tilde\omega)],$$
$$\tilde e\to U(\tilde\omega)\tilde e U^{-1}(\tilde\omega). \tag{4.44}$$

The infinitesimal form of this action is given in Eq. (4.38). This group action restricts to the constraint submanifold $\mathscr{C}^{-1}(0)$ since, as is well known, the gauge transformations (4.44) preserve the constraint equations [see Eq. (4.39) for the infinitesimal form of this result]. It is therefore natural to contemplate making a further reduction of the constraint submanifold by dividing out the gauge group action. Marsden and Weinstein[23] have a general method for reducing symplectic manifolds on which there is a suitable group action and for studying the geometry of the resulting quotient space. Clearly a symplectic manifold on which the constraints are trivial and the gauge transormation have been factored out would provide a natural setting for the Hamiltonian dynamics of Yang–Mills fields.

The symmetries of Yang–Mills fields on spacially compact manifolds have been extensively studied by Arms.[24] She presents a decomposition of the electric field of which ours is the noncompact extension. She also allows for coupling to the gravitational field. It would be interesting to consider the problem of gauge degeneracies in this more general setting since ultimately one wants to quantize this fully coupled system.

ACKNOWLEDGMENTS

I am grateful to S. MacDowell, I. Bars, C. Sommerfield, M. Cantor, J. Marsden, A. Fischer, and J. Arms for valuable conversations.

[1]V.N. Gribov, Lecture at the 12th Winter School of the Leningrad Nuclear Physics Institute, 1977 (unpublished).
[2]V.N. Gribov, "Quantization of non-Abelian gauge theories," Leningrad Nuclear Physics Institute preprint (1977).
[3]L. Nirenberg and H. Walker, J. Math. Ann. Appl. **42**, 271 (1973).
[4]M. Cantor, Ind. Univ. Math. J. **24**, 897 (1975).
[5]M. Cantor, "The existence of non-trivial asymptotically flat initial data for vacuum spacetimes," Univ. of Texas preprint (1977).
[6]M. Cantor, "Some problems of global analysis on asymptotically simple manifolds," Univ. of Texas preprint (1977).
[7]Y. Choquet-Bruhat, A. Fischer, and J. Marsden, "Maximal hypersurfaces and positivity of mass," Univ. of Paris and Univ. of Calif. preprint (1977).
[8]Samuel MacDowell (private communication).
[9]V. Moncrief, J. Math. Phys. **16**, 1556 (1975).
[10]I.M. Singer, "Some remarks on the Gribov ambiguity," Univ. of Calif. preprint (1978).
[11]C. Bender, T. Eguchi, and H. Pagels, Phys. Rev. D **17**, 1086 (1978).
[12]R.D. Peccei, Phys. Rev. D **17**, 1097 (1978).
[13]R. Jackiw, I. Muzinich, and C. Rebbi, Phys. Rev. D **17**, 1576 (1978).
[14]M. Ademollo, E. Napolitano, and S. Sciuto, "A further pathology of the Coulomb gauge in non-Abelian Yang–Mills theory," CERN Preprint Th. 2412 (1977).
[15]J. York, J. Math. Phys. **14**, 456 (1973).
[16]N. O'Murchadha, J. York, Phys. Rev. D **10**, 428 (1974).
[17]V. Moncrief, Ann. Phys. (N.Y.) **108**, 387 (1977).
[18]A. Fischer and J. Marsden, Bull. Am. Math. Soc. **79**, 997 (1973).
[19]A. Fischer, J. Mardsen, and V. Moncrief, "The structure of the space of solutions of Einstein's equations," (in preparation).
[21]S. Lang, *Differential Manifolds* (Addison-Wesley, Reading, Mass., 1972), see esp. Chap. II.
[22]J. Marsden, *Applications of Global Analysis in Mathematical Physics* (Publish or Perish, Boston, 1974), see esp. Chap. I.
[23]J. Marsden and A. Weinstein, Rep. Math. Phys. **5**, 121 (1974).
[24]J. Arms, J. Math. Phys. **20**, 443 (1979).

Quartic trace identity for exceptional Lie algebras[a)]

Susumu Okubo

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627

(Received 5 September 1978)

Let X be a representation matrix of generic element x of a simple Lie algebra in generic irreducible representation $\{\lambda\}$ of the Lie algebra. Then, for all exceptional Lie algebras as well as A_1 and A_2, we can prove the validity of a quartic trace identity

$\mathrm{Tr}(X^4) = K(\lambda)[\mathrm{Tr}(X^2)]^2$,

where the constant $K(\lambda)$ depends only upon the irreducible representation $\{\lambda\}$, and its explicit form is calculated. Some applications of second and fourth order indices have also been discussed.

1. INTRODUCTION AND SUMMARY OF MAIN RESULTS

Let $\mathfrak{g}$ stand for any of the simple Lie algebras $A_1, A_2, G_2, F_4, E_6, E_7$, and E_8 over the complex number field. Let X denote a representation matrix of the generic element x of any of these algebras in any given irreducible representation $\{\lambda\}$. The main purpose of this note is to prove the validity of the quartic trace identity,

$$\mathrm{Tr}X^4 = K(\lambda)(\mathrm{Tr}X^2)^2, \tag{1.1}$$

$$K(\lambda) = \frac{d(\lambda_0)}{2[2+d(\lambda_0)]d(\lambda)}\left(6 - \frac{I_2(\lambda_0)}{I_2(\lambda)}\right), \tag{1.2}$$

where $\{\lambda_0\}$ designates the adjoint representation and where $d(\lambda)$ and $I_2(\lambda)$ are dimension and eigenvalue of the second-order Casimir invariant of $\mathfrak{g}$ in the irreducible representation $\{\lambda\}$, respectively. Note that the adjoint representation $\{\lambda_0\}$ is also irreducible because of the simplicity of algebras under consideration.

As we shall see in the next section, our identity, Eq. (1.1), is intimately connected with a fact[1] that all exceptional Lie algebras G_2, F_4, E_6, E_7, and E_8 as well as A_1 and A_2 have no genuine fourth-order Casimir invariants. To show some special examples of Eq. (1.1), let us adopt a lexicographical numbering of simple roots of these algebras as by Patera and Sankoff.[2] Then, we calculate

$$A_2{:}\Lambda = \Lambda_1,\quad d(\lambda) = 3,\quad K(\lambda) = \tfrac{1}{2}, \tag{1.3}$$

$$G_2{:}\Lambda = \Lambda_2,\quad d(\lambda) = 7,\quad K(\lambda) = \tfrac{1}{4}, \tag{1.4}$$

$$F_4{:}\Lambda = \Lambda_4,\quad d(\lambda) = 26,\quad K(\lambda) = \frac{1}{12}, \tag{1.5}$$

$$E_6{:}\Lambda = \Lambda_1,\quad d(\lambda) = 27,\quad K(\lambda) = \frac{1}{12}, \tag{1.6}$$

$$E_7{:}\Lambda = \Lambda_6,\quad d(\lambda) = 56,\quad K(\lambda) = \frac{1}{24}, \tag{1.7}$$

$$E_8{:}\Lambda = \Lambda_1,\quad d(\lambda) = 248,\quad K(\lambda) = \frac{1}{100}, \tag{1.8}$$

where Λ and Λ_j $(j = 1,2,...,n)$ are respectively the highest weight of the irreducible representation $\{\lambda\}$ and n fundamental weights of the simple Lie algebras with the rank n. We may note that the validity of the quartic trace identity (1.1) for the case of Eq. (1.4) has already been verified elsewhere,[3] while cases for Eqs. (1.4), (1.5), and (1.6) have also been directly derived by Cvitanovic.[4] If we choose X to correspond to a generic element of a Cartan subalgebra of these Lie algebras, then Eq. (1.1) is rewritten as

$$\sum_M M^4 = K(\lambda)\Big(\sum_M M^2\Big)^2, \tag{1.9}$$

where M is the eigenvalue of X in the irreducible representation $\{\lambda\}$ and the summation is over all eigenvalues of X, counting the same M as many times as its multiplicity. Especially, for the algebra A_1, we find an amusing formula,

$$\sum_{m=-j}^{j} m^4 = \tfrac{1}{15}j(j+1)(2j+1)\{3j(j+1)-1\}, \tag{1.10}$$

for any positive integer or half-integer j. Also for the algebras A_2 and G_2, Eq. (1.9) for the cases of (1.3) and (1.4) is found to be equivalent to an identity

$$a^4 + b^4 + c^4 = \tfrac{1}{2}(a^2+b^2+c^2)^2 \tag{1.11a}$$

where a, b, and c are arbitrary constants subject to a constraint

$$a + b + c = 0. \tag{1.11b}$$

This equation has been used to construct[5] a pseudo-octonion algebra which permits a nondegenerate composition without unit element.

As a byproduct in the course of proving Eqs. (1.1) and (1.2), we have also found the following rather amusing results. First, *any* simple Lie algebra (not necessarily limited now to exceptional algebra and A_2) *except* for A_1 possess an irreducible representation $\{\rho\}$ such that we have first

$$I_2(\rho) = 2I_2(\lambda_0), \tag{1.12}$$

and second its dimension $d(\rho)$ is given by

$$d(\rho) = \tfrac{1}{2}d(\lambda_0)[d(\lambda_0) - 3] \tag{1.13a}$$

for all algebras other than A_n $(n \geqslant 2)$, and

$$d(\rho) = \tfrac{1}{4}d(\lambda_0)[d(\lambda_0) - 3] \tag{1.13b}$$

for the algebra A_n $(n \geqslant 2)$. If we decompose the product representation $\{\lambda_0\} \otimes \{\lambda_0\}$ as a direct sum of N irreducible representations $\{\lambda_j\}$ as

$$\{\lambda_0\} \otimes \{\lambda_0\} = \sum_{j=1}^{N} \oplus \{\lambda_j\}, \tag{1.14}$$

[a)]Work supported in part by the U.S. Department of Energy under Contract Number EY-76-C-02-3065.000.

 0022-2488/79/040586-08$01.00

then the representation $\{\rho\}$ always appears as one of $\{\lambda_j\}$ in the right side of Eq. (1.14) for all cases other than A_n ($n \geqslant 2$). On the other hand, we have a pair of such $\{\rho\}$'s for the algebras A_n ($n \geqslant 2$) in Eq. (1.14), which are contragradient to each other.

For the exceptional Lie algebras, we have $N = 5$ and we can completely determine $d(\lambda_j)$ and $I_2(\lambda_j)$ ($j = 1,2,3,4,5$) in this case as follows. The right side of Eq. (1.14) always contains a trivial representation $\{0\}$ as well as the adjoint representation $\{\lambda_0\}$. Also, it contains an irreducible representation whose highest weight is precisely twice the highest weight Λ_0 of the adjoint representation. We label them as $\{\lambda_1\}$, $\{\lambda_2\}$, and $\{\lambda_3\}$, respectively. Finally, we choose $\{\lambda_5\}$ to be the special representation $\{\rho\}$ satisfying Eq. (1.12). Then, when we set

$$\theta = \left(\frac{242 + d(\lambda_0)}{2 + d(\lambda_0)} \right)^{1/2}, \tag{1.15}$$

we find

$$d(\lambda_3) = \frac{3}{\theta(\theta - 1)} \{(19 + \theta)d(\lambda_0) - 2(11 - \theta)\}, \tag{1.16a}$$

$$d(\lambda_4) = \frac{3}{\theta(\theta + 1)} \{(19 - \theta)d(\lambda_0) - 2(11 + \theta)\}, \tag{1.16b}$$

$$I_2(\lambda_3) = \tfrac{1}{6}(11 + \theta)I_2(\lambda_0), \tag{1.17a}$$

$$I_2(\lambda_4) = \tfrac{1}{6}(11 - \theta)I_2(\lambda_0). \tag{1.17b}$$

The case for $\lambda = \{\lambda_5\} = \{\rho\}$ has already been given in Eqs. (1.12) and (1.13a). Since other two cases $\{\lambda_1\} = \{0\}$ and $\{\lambda_2\} = \{\lambda_0\}$ are trivial, these completely determine $d(\lambda_j)$ and $I_2(\lambda_j)$ ($j = 1,2,3,4,5$) for all exceptional Lie algebras. Also, because $d(\lambda_3)$ and $d(\lambda_4)$ must be integer and $I_2(\lambda_j)/I(\lambda_0)$ ($j = 3,4$) are also rational numbers, θ has to be a rational number. Indeed, we find $\theta = 4, \frac{7}{3}, 2, \frac{5}{3}$, and $\frac{7}{5}$ for G_2, F_4, E_6, E_7, and E_8, respectively. We have also $\theta = 7$ and 5 for algebras A_1 and A_2. It is amusing to note that $(n + 2)\theta$ are always integers for all these algebras. Finally, we may remark that Eqs. (1.15)–(1.17) are valid also for the algebra A_2 with $N = 6$, if we identify $\{\lambda_4\} = \{\lambda_0\}$, $\{\lambda_5\} = \{\rho\}$, and $\{\lambda_6\} = \{\rho^*\}$.

2. CASIMIR INVARIANTS AND TENSOR OPERATORS

Let x_μ ($\mu = 1,2,\ldots,d_0$) be a basis of any Lie algebra $\mathfrak{g}$ with structure equation

$$[x_\mu, x_\nu] = C^\lambda_{\mu\nu} x_\lambda, \tag{2.1}$$

where the repeated index λ implies an automatical summation on λ from 1 to $d_0 = d(\lambda_0)$. Note that $d(\lambda_0)$ is the dimension of the adjoint representation $\{\lambda_0\}$. Let X_μ be a representation matrix of x_μ in a representation $\{\lambda\}$ which need not be irreducible, and let $d = d(\lambda)$ be the dimension of the representation. Then, a collection of d elements t_j ($j = 1,2,\ldots,d$) belonging to the universal enveloping algebra of $\mathfrak{g}$ is called a tensor operator of the type $\{\lambda\}$, if they satisfy commutation relation

$$[x_\mu, t_j] = \sum_{k=1}^{d} (X_\mu)_{kj} t_k \tag{2.2}$$

for all $\mu = 1,2,\ldots,d_0$ and $j = 1,2,\ldots,d$, where $(X_\mu)_{jk}$ is the (j,k) matrix element of the matrix X_μ. We may note that Eqs. (2.1) and (2.2) are compatible with the Jacobi identity

$$[x_\mu,[x_\nu,t_j]] + [x_\nu,[t_j,x_\mu]] + [t_j,[x_\mu,x_\nu]] = 0$$

because of the equation of representation

$$[X_\mu, X_\nu] = C^\lambda_{\mu\nu} X_\lambda \tag{2.3}$$

for $d \times d$ matrix X_μ. In this paper, we are especially interested in a case that $\{\lambda\}$ is a product of p copies of the adjoint representation $\{\lambda_0\}$,

$$\{\lambda\} = \underbrace{\{\lambda_0\} \otimes \{\lambda_0\} \otimes \cdots \otimes \{\lambda_0\}}_{p\text{-times}} \tag{2.4}$$

which is not in general irreducible. Then, the corresponding tensor operator which may be written as $t_{\mu_1\mu_2\cdots\mu_p}$ ($\mu_1,\mu_2,\ldots,\mu_p = 1,2,\ldots,d_0$) satisfies

$$[x_\lambda, t_{\mu_1\mu_2\cdots\mu_p}] = \sum_{j=1}^{p} C^\alpha_{\lambda\mu_j} t_{\mu_1\cdots\hat{\alpha}\cdots\mu_p}, \tag{2.5}$$

where the symbol $\hat{\alpha}$ in Eq. (2.5) implies that we replace the jth index μ_j by α and sum over $\alpha = 1,2,\ldots,d_0$. We shall call such a tensor operator a p-vector operator since a product of p vector-operators has the same commutation relation (2.5). For the special case $p = 1$, this reduces of course to the usual vector (or adjoint) operator.

We may define also a tensor operator T_j of the type $\{\lambda\}$ in a representation $\{\rho\}$, if Y_μ and T_j are now $d(\rho) \times d(\rho)$ matrices satisfying

$$[Y_\mu, Y_\nu] = C^\lambda_{\mu\nu} Y_\lambda, \tag{2.6a}$$

$$[Y_\mu, T_j] = \sum_{j=1}^{d} (X_\mu)_{kj} T_k. \tag{2.6b}$$

Especially, any representation matrices Y_μ and T_j of x_μ and t_j in the representation $\{\rho\}$, respectively satisfy Eqs. (2.6). In the present paper we are mostly interested in p-vector operators in the adjoint representation $\{\rho\} = \{\lambda_0\}$ so that

$$Y_\mu = \mathrm{ad}x_\mu, \tag{2.7a}$$

$$(\mathrm{ad}x_\mu)_{\lambda\nu} = C^\lambda_{\mu\nu}. \tag{2.7b}$$

In passing, we remark that a study of tensor operators in a reducible or irreducible representation is important in physics, since it determines all matrix elements of tensor operators of a given type between two arbitrary irreducible representations which need not be equivalent to each other. A general study for such tensor operator algebra has been extensively carried out by Biedenharn and Louck[6] for the algebra A_n.

Hereafter, we restrict ourselves to the case that $\mathfrak{g}$ is semisimple[7] over the complex number field. Then setting

$$g_{\mu\nu} = c\,\mathrm{Tr}(\mathrm{ad}x_\mu \mathrm{ad}x_\nu) \tag{2.8}$$

for a nonzero constant c, its inverse $g^{\mu\nu}$ exists in view of Cartan's criteria of semisimplicity, and they satisfy the relations

$$g_{\mu\lambda}C^{\lambda}_{\nu\tau} = -g_{\nu\lambda}C^{\lambda}_{\mu\tau}, \tag{2.9a}$$

$$g^{\mu\lambda}C^{\nu}_{\lambda\tau} = -g^{\nu\lambda}C^{\mu}_{\lambda\tau}. \tag{2.9b}$$

Next, let I_p $(p\geqslant 2)$ be a pth order Casimir invariant given by

$$I_p = b^{\mu_1\mu_2\cdots\mu_p}x_{\mu_1}x_{\mu_2}\cdots x_{\mu_p}$$

$$= b_{\mu_1\mu_2\cdots\mu_p}x^{\mu_1}x^{\mu_2}\cdots x^{\mu_p}, \tag{2.10}$$

where the constants $b_{\mu_1\cdots\mu_p}$ are completely symmetric in the indices $\mu_1,\mu_2,\ldots,\mu_p$ and where we raise or lower indices by means of the metrics $g_{\mu\nu}$ and $g^{\mu\nu}$. Then, the condition

$$[x_\mu, I_p] = 0 \tag{2.11}$$

implies that $b_{\mu_1\cdots\mu_p}$ must satisfy

$$\sum_{j=1}^{p} C^{\alpha}_{\lambda\mu_j}b_{\mu_1\cdots\hat{\alpha}\cdots\mu_p} = 0 \tag{2.12}$$

because of the Poincare–Birkoff–Witt theorem. Conversely, suppose that the constants $b_{\mu_1\cdots\mu_p}$ which need not now be completely symmetric in these indices satisfy Eq. (2.12). Then, I_p, defined by Eq. (2.10) satisfies (2.11), i.e., it is a Casimir invariant although it may not define a completely symmetric irreducible Casimir invariant. If we define a $d(\lambda_0)\times d(\lambda_0)$ matrix $B_{\mu_1\mu_2\cdots\mu_q}$ with $q=p-2$ by

$$(B_{\mu_1\mu_2\cdots\mu_q})_{\alpha\beta} = g^{\nu\alpha}b_{\nu\beta\mu_1\mu_2\cdots\mu_q}, \tag{2.13a}$$

$$q = p-2, \tag{2.13b}$$

then $B_{\mu_1\cdots\mu_q}$ is completely symmetric in $\mu_1,\mu_2,\ldots,\mu_q$ whenever $b_{\mu_1\mu_2\cdots\mu_p}$ is completely symmetric in all p-indices $\mu_1\mu_2\cdots\mu_p$. Moreover, Eq. (2.12) is rewritten into a matrix equation

$$[\mathrm{ad}x_\lambda, B_{\mu_1\cdots\mu_q}] = \sum_{j=1}^{q} C^{\alpha}_{\lambda\mu_j}B_{\mu_1\cdots\hat{\alpha}\cdots\mu_q}. \tag{2.14}$$

In other words, $B_{\mu_1\cdots\mu_q}$ is a completely symmetric q-vector operator in the adjoint space. Conversely, let $B_{\mu_1\cdots\mu_q}$ be any such q-vector operator satisfying Eq. (2.14). Then, reversing the direction of our reasoning, the coefficient $b_{\mu_1\mu_2\cdots\mu_p}$ $(p=q+2)$ defined by

$$b_{\mu_1\mu_2\cdots\mu_p} = g_{\mu_1\lambda}(B_{\mu_3\mu_4\cdots\mu_p})_{\lambda\mu_2} \tag{2.15}$$

satisfy Eq. (2.12) so that I_p given by Eq. (2.10) is a Casimir invariant. However, $b_{\mu_1\cdots\mu_p}$ $(p=q+2)$ just defined by Eq. (2.15) is in general *not* completely symmetric in all indices $\mu_1\mu_2\mu_3\cdots\mu_p$, although it is symmetric in $\mu_3,\mu_4\ldots,\mu_p$ by definition. At any rate, summarizing our results, all pth order Casimir invariants I_p of any semisimple Lie algebras[7] can be constructed first by finding all linearly independent completely symmetric $(p-2)$-ple vector operators in the adjoint space and secondly by choosing a suitable linear combination of them so that $b_{\mu_1\mu_2\cdots\mu_p}$ given by (2.15) is now completely symmetric in all indices.

The special case $p=3$ has been solved elsewhere[8] to show that there is no third-order Casimir invariant I_3 for any simple Lie algebra except for algebras A_n $(n\geqslant 2)$ and $D_3 = A_3$, in conformity with the result of Ref. 1. The next simple case is $p=4$ corresponding to the fourth order Casimir invariant I_4 which is the main subject of the present note. Then we have to find first all linearly independent $d(\lambda_0)\times d(\lambda_0)$ matrices $T_{\mu\nu}$ satisfying

$$[\mathrm{ad}x_\lambda, T_{\mu\nu}] = C^{\alpha}_{\lambda\nu}T_{\mu\alpha} + C^{\alpha}_{\lambda\mu}T_{\alpha\nu} \tag{2.16}$$

which need not yet be symmetric in μ and ν. We can easily find five bivector operators in the adjoint space as follows. Let E be the unit matrix in the adjoint space. Further, let us define $d(\lambda_0)\times d(\lambda_0)$ matrices $G_{\mu\nu}$ by

$$(G_{\mu\nu})_{\alpha\beta} = \delta^{\alpha}_{\mu}g_{\nu\beta} \tag{2.17}$$

which can be easily shown to be a bivector operator. In this way, we find that

$$g_{\mu\nu}E, \quad G_{\mu\nu}+G_{\nu\mu}, \quad \mathrm{ad}x_\mu\mathrm{ad}x_\nu + \mathrm{ad}x_\nu\mathrm{ad}x_\mu \tag{2.18}$$

are symmetric bivector operators, while

$$G_{\mu\nu}-G_{\nu\mu}, \quad [\mathrm{ad}x_\mu, \mathrm{ad}x_\nu] = C^{\lambda}_{\mu\nu}\mathrm{ad}x_\lambda \tag{2.19}$$

are antisymmetric bivector operators. Moreover, only the combination

$$B_{\mu\nu} = g_{\mu\nu}E + G_{\mu\nu} + G_{\nu\mu} \tag{2.20}$$

among these five turns out to give a completely symmetric coefficient $b_{\mu\nu\alpha\beta}$ of the form

$$b_{\mu\nu\alpha\beta} = g_{\alpha\lambda}(B_{\mu\nu})_{\lambda\beta}$$

$$= g_{\mu\nu}g_{\alpha\beta} + g_{\mu\alpha}g_{\nu\beta} + g_{\mu\beta}g_{\nu\alpha}. \tag{2.21}$$

In the next section, we shall prove that all exceptional Lie algebras and A_2 have precisely three linearly independent symmetric bivector operators in the adjoint space, while the algebra A_1 has only two. Now, three bivector operators given by Eq. (2.18) are easily checked to be linearly independent except for the algebra A_1, where only $g_{\mu\nu}E$ and $G_{\mu\nu}+G_{\nu\mu}$ are linearly independent for A_1. Therefore, for all these cases of algebras G_2, F_4, E_6, E_7, E_8, A_1, and A_2, only the completely symmetric $b_{\mu\nu\alpha\beta}$ satisfying Eq. (2.12) for $p=4$, must have the form of Eq. (2.21), apart from a proportional constant c. Therefore, any fourth order Casimir invariants I_4 of these algebras is proportional to

$$\tilde{I} = (g^{\mu\nu}g^{\alpha\beta} + g^{\mu\alpha}g^{\nu\beta} + g^{\mu\beta}g^{\nu\alpha})x_\mu x_\nu x_\alpha x_\beta$$

$$= 3(I_2)^2 - \tfrac{1}{2}I_2(\lambda_0)I_2, \tag{2.22}$$

where I_2 is the second-order Casimir invariant

$$I_2 = g^{\mu\nu}x_\mu x_\nu \tag{2.23}$$

and $I_2(\lambda_0)$ is the eigenvalue of I_2 in the adjoint representation so that

$$g^{\mu\nu}\mathrm{Tr}(\mathrm{ad}x_\mu\mathrm{ad}x_\nu) = d(\lambda_0)I_2(\lambda_0). \tag{2.24}$$

This proves that any fourth order Casimir invariant I_4 of these algebras is a quadratic polynomial of I_2. This fact is again in conformity with the result of Ref. 1, which states that I_4 should be expressed as a polynomial of n fundamental Casimir invariants of orders not equal to 4. However, it may be stressed that the result of Ref. 1 does not necessarily prove the validity of the specific form of Eq. (2.22) for I_4 by the following reason. For example, consider the case of G_2, where I_4 must be a polynomial of I_2 and I_6 which is a six-order Casimir invariant. This would suggest that I_4 be expressed as a quadratic polynomial of I_2 without any dependence on I_6. But this need not be so in principle, since I_4 could be a complicated polynomial of I_6 and I_2 such as

$$I_4 = I_6 - a_1(I_2)^3 - a_2(I_2)^2 - a_3 I_2$$

for some constants a_1, a_2, and a_3 in such a way that the six-order polynomial of generators would cancel each other to leave I_4. This possibility is certainly realizable, if I_2 and I_4 (as well as I_2 and I_6) could generate the space of all Casimir invariants of G_2. Because of our result, Eq. (2.22), such a possibility is, of course, impossible in reality, but the general theorem of Ref. 1 does not appear to preclude it.

After these preparations, let us now prove our main result, Eqs. (1.1) and (1.2). Let X_μ be a representation matrix of x_μ in the generic representation space $\{\lambda\}$ which need not yet be irreducible. Further, let us set

$$b_{\mu_1\mu_2\cdots\mu_p} = \frac{1}{p!}\sum_P \mathrm{Tr}(X_{\mu_1}X_{\mu_2}\cdots X_{\mu_p}), \tag{2.25}$$

where the summation is over $p!$ permutations among $\mu_1,\mu_2,\ldots,\mu_p$. Then, $b_{\mu_1\mu_2\cdots\mu_p}$ is evidently completely symmetric. When we note a trivial identity

$$\mathrm{Tr}([X_\lambda, X_{\mu_1}X_{\mu_2}\cdots X_{\mu_p}]) = 0$$

and compute the commutator by means of Eq. (2.3), then we immediately see that $b_{\mu_1\mu_2\cdots\mu_p}$ satisfy Eq. (2.12). Especially for the case $p = 4$ of all algebras A_1, A_2, G_2, F_4, E_6, E_7, and E_8, we must have

$$b_{\mu\nu\alpha\beta} = \tfrac{1}{3}B(\lambda)(g_{\mu\nu}g_{\alpha\beta} + g_{\mu\alpha}g_{\nu\beta} + g_{\mu\beta}g_{\nu\alpha}), \tag{2.26}$$

where the proportional constant $B(\lambda)$ depends upon the representation $\{\lambda\}$. In passing, we note that Eqs. (2.25) and (2.26) immediately prove the symmetricity of the fourth order indices for these algebras in agreement with the result of Patera *et al.*[9] Now we can compute $B(\lambda)$ as follows. We multiply $g^{\alpha\beta}$ on both sides of Eq. (2.26) and compute $g^{\alpha\beta}b_{\mu\nu\alpha\beta}$ from Eq. (2.25). Hereafter, we restrict ourselves to the case that $\{\lambda\}$ is irreducible. Then, we may use

$$g^{\mu\nu}X_\mu X_\nu = I_2(\lambda)E, \tag{2.27}$$

$$\mathrm{Tr}(X_\mu X_\nu) = \frac{d(\lambda)}{d(\lambda_0)}I_2(\lambda)g_{\mu\nu}, \tag{2.28}$$

where $d(\lambda)$ and $I_2(\lambda)$ are dimension and eigenvalue of I_2 in the irreducible representation $\{\lambda\}$, respectively, and where E in Eq. (2.27) stands for the unit matrix in $\{\lambda\}$. Using Eqs. (2.9) many times, we then find

$$B(\lambda) = \frac{d(\lambda)I_2(\lambda)}{2[2 + d(\lambda_0)]d(\lambda_0)}[6I_2(\lambda) - I_2(\lambda_0)] \tag{2.29}$$

after some computations. Since any generic element X is expressible as

$$X = \sum_{\mu=1}^{d_0}\xi^\mu X_\mu \tag{2.30}$$

for some complex constant ξ^μ, Eqs. (2.25) and (2.26) lead to

$$\mathrm{Tr}X^4 = B(\lambda)(g^{\mu\nu}\xi_\mu\xi_\nu)^2 = K(\lambda)(\mathrm{Tr}X^2)^2,$$

$$K(\lambda) = B(\lambda)\left(\frac{d(\lambda_0)}{d(\lambda)I_2(\lambda)}\right)^2$$

which reproduce Eqs. (1.1) and (1.2).

In this connection, we should note that the result of Ref. 8 for the case $p = 3$ is

$$\mathrm{Tr}X^3 = 0, \tag{2.31a}$$

$$\mathrm{Tr}(X_\mu X_\nu + X_\nu X_\mu)X_\lambda = 0, \tag{2.31b}$$

for any simple Lie algebras except for the algebra A_n $(n\geqslant 2)$, as well as any self-contragradient representation $\{\lambda\}$ of A_n. Then, if we rewrite

$$\begin{aligned}\mathrm{Tr}(X_\mu X_\nu X_\alpha X_\beta) &= \frac{1}{4!}\sum_P \mathrm{Tr}(X_\mu X_\nu X_\alpha X_\beta)\\ &\quad + \tfrac{1}{4}C^\lambda_{\alpha\beta}\mathrm{Tr}(X_\mu X_\nu + X_\nu X_\mu)X_\lambda\\ &\quad + \tfrac{1}{4}C^\lambda_{\mu\nu}\mathrm{Tr}(X_\alpha X_\beta + X_\beta X_\alpha)X_\lambda\\ &\quad + \tfrac{1}{6}(C^\lambda_{\mu\nu}C^\tau_{\alpha\beta} - C^\lambda_{\mu\beta}C^\tau_{\nu\alpha})\mathrm{Tr}(X_\lambda X_\tau),\end{aligned} \tag{2.32}$$

we can compute $\mathrm{Tr}(X_\mu X_\nu X_\alpha X_\beta)$ for all exceptional algebras and A_1 in terms of the second order Casimir invariant $I_2(\lambda)$, while it can be expressed[8] in terms of $I_2(\lambda)$ and $I_3(\lambda)$ for the case of A_2.

For other algebras B_n $(n\geqslant 2)$, C_n $(n\geqslant 2)$, D_n $(n\geqslant 4)$, and A_n $(n\geqslant 3)$, we require one more symmetric bivector operator in addition to that given by Eq. (2.20). Indeed, the results of Ref. 1 and 3 show the existence of algebraically independent fourth-order Casimir invariants for these algebras so that we do not have any relation of similar type for these cases. However, the results of Ref. 1 and 3 suggest that any fifth-order Casimir invariant I_5 for algebra A_2 and A_3 would be expressible as products of I_2 and I_3. If this were indeed the case, then we can prove the validity of

$$\mathrm{Tr}X^5 = A(\lambda)(\mathrm{Tr}X^2)(\mathrm{Tr}X^3), \tag{2.33a}$$

$$A(\lambda) = \frac{5d(\lambda_0)}{2[6 + d(\lambda_0)]d(\lambda)}\left(4 - \frac{I_2(\lambda_0)}{I_2(\lambda)}\right), \tag{2.33b}$$

in any irreducible representation $\{\lambda\}$ of A_2 and A_3. Similarly, we may conjecture that any seventh-order Casimir invariant I_7 of algebras E_6 and D_5 would be expressed as products of I_5 and I_2, since both do not possess genuine third and seventh Casimir invariants. In that case, we would find

$$\mathrm{Tr}X^7 = D(\lambda)(\mathrm{Tr}X^2)(\mathrm{Tr}X^5), \tag{2.34a}$$

$$D(\lambda) = \frac{35d(\lambda_0)}{4[10 + d(\lambda_0)]d(\lambda)}\left(\frac{12}{5} - \frac{I_2(\lambda_0)}{I_2(\lambda)}\right). \tag{2.34b}$$

As checks of these relations, we note that the case of $\Lambda = \Lambda_1$ or Λ_5 for the algebra E_6 gives

$$\mathrm{Tr}X^7 = \frac{7}{24}\mathrm{Tr}X^2\mathrm{Tr}X^5 \tag{2.35}$$

which agrees with the result found by Cvitanovic.[4] Also, for the algebra A_3 with the choice of $\Lambda = \Lambda_1$, then Eqs. (2.33) give an identity

$$a^5+b^5+c^5+d^5$$
$$=\tfrac{5}{6}(a^2+b^2+c^2+d^2)(a^3+b^3+c^3+d^3) \qquad (2.36)$$

whenever arbitrary constants a,b,c, and d satisfy

$$a+b+c+d=0. \qquad (2.37)$$

For the algebra A_2 with $\Lambda=\Lambda_1$, we find the same equation with $d=0$. Note that Eq. (2.36) is an analog of Eq. (1.11) and can be easily verified to be correct. Therefore, the validity of our conjecture for Eqs. (2.33) and (2.34) appears to have a good chance, although its formal proof for their correctness is so far unsuccessful. Note that it involves a study of triple and quintuple vector operators in the adjoint space.

3. BIVECTOR OPERATORS IN ADJOINT SPACE

Let $\mathfrak{g}$ be any simple Lie algebra and let us decompose the product $\{\lambda_0\}\otimes\{\lambda_0\}$ as in Eq. (1.14)

$$\{\lambda_0\}\otimes\{\lambda_0\}=\sum_{j=1}^{N}\oplus\{\lambda_j\} \qquad (3.1)$$

as a direct sum of irreducible component $\{\lambda_j\}$ ($j=1,2,\dots,N$). If any two $\{\lambda_j\}$ and $\{\lambda_k\}$ appearing in Eq. (3.1) are inequivalent to each other whenever $j\neq k$, then we say that the product is nondegenerate or multiplicity-free. By a direct computation, we find that all simple algebras except for A_n ($n\geqslant2$) and $D_3=A_3$ have multiplicity-free decomposition for (3.1) with $N=3$ for A_1, $N=5$ for all exceptional algebras, and $N=6$ for B_n ($n\geqslant2$), C_n ($n\geqslant2$), and D_n ($n\geqslant5$). On the other hand, for the algebra A_n ($n\geqslant2$), the same adjoint representation $\{\lambda_0\}$ appears twice in the right side of Eq. (3.1) with $N=6$ for A_2, and $N=7$ for A_n ($n\geqslant3$); while $N=7$ for D_4.

Now, corresponding to Eq. (3.1), any bivector operator $T_{\mu\nu}$ which need not be symmetric in μ and ν can also be decomposed into a direct sum of N irreducible tensor operators of the type $\{\lambda_j\}$ ($j=1,2,\dots,N$). For the case of multiplicity-free decomposition for all simple Lie algebras excepting the algebra A_n ($n\geqslant2$), the Wigner–Eckart theorem implies that the number of linearly independent bivector operators in the adjoint space for all these simple Lie algebras must be precisely N. On the other hand, the algebra A_n ($n\geqslant2$) will have $N+2$ linearly independent bivector operators in the adjoint space, because of the double multiplicity in the decomposition. Especially, the number of all linearly independent bivector operators are three for A_1, five for all exceptional Lie algebras, six for B_n ($n\geqslant2$), C_n ($n\geqslant2$), and D_n ($n\geqslant5$), seven for D_4, eight for A_2, and nine for A_n ($n\geqslant3$). Hence, Eqs. (1.1) and (1.2) will be valid for all exceptional algebras as well as for A_1, as we have shown in the previous section. For the algebra A_2, we observe that only three out of these eight bivector operators are symmetric ones. Therefore, the same argument given in the previous section is also applicable to the case of A_2. These prove Eqs. (1.1) and (1.2) for all A_1, A_2, G_2, F_4, E_6, E_7, and E_8.

Since the Wigner–Eckart theorem employs the transcendental method of integration over group manifold, it may be of some interest to give another purely algebraic proof, although it is applicable only for cases of multiplicity-free decomposition, i.e., only for simple Lie algebras except for A_n ($n\geqslant2$). Following (I), we set

$$Q=g^{\mu\nu}\mathrm{ad}x_\mu\otimes\mathrm{ad}x_\nu \qquad (3.2)$$

in the product space $\{\lambda_0\}\otimes\{\lambda_0\}$. When P_j ($j=1,2,\dots,N$) represents the projection operator for the jth irreducible space $\{\lambda_j\}$ in the decomposition Eq. (3.1), then it has been shown in (I) that

$$Q=\sum_{j=1}^{N}\xi_jP_j, \qquad (3.3)$$

$$\xi_j=\tfrac{1}{2}[I_2(\lambda_j)-2I_2(\lambda_0)]. \qquad (3.4)$$

Moreover, we can check that we have $\xi_j\neq\xi_k$ for $j\neq k$ except for the algebra D_4. In other words, these algebras satisfy the multiplicity-free (or nondegeneracy) condition with zero deficiency index in the terminology of (I). Therefore, if we set

$$T=T_{\mu\nu}\otimes G^{\mu\nu} \qquad (3.5)$$

for an arbitrary bivector operator $T_{\mu\nu}$ in $\{\lambda_0\}$, we can express T as a polynomial of Q of order $N-1$,

$$T=\sum_{j=0}^{N-1}a_jQ^j,\qquad Q^0\equiv E\otimes E, \qquad (3.6)$$

for some constant a_j by the method following (I). Now, we note the identities

$$G_{\mu\nu}G_{\alpha\beta}=g_{\nu\alpha}G_{\mu\beta}, \qquad (3.7a)$$

$$\mathrm{Tr}G_{\mu\nu}=g_{\mu\nu}. \qquad (3.7b)$$

Then multiplying $E\otimes G_{\beta\alpha}$ by both sides of Eq. (3.6) and taking a partial trace with respect to the second space (but *not* the first), this gives

$$T_{\alpha\beta}=\sum_{j=0}^{N-1}a_jR^{(j)}_{\alpha\beta}, \qquad (3.8a)$$

$$R^{(j)}_{\alpha\beta}=\mathrm{Tr}^{(2)}(Q^j\cdot E\otimes G_{\beta\alpha}). \qquad (3.8b)$$

Clearly, $R^{(j)}_{\alpha\beta}$ ($j=0,1,2,\dots,N-1$) are independent of $T_{\mu\nu}$ and it is easily shown that they are bivector operators in the first adjoint space $\{\lambda_0\}$. This proves that any arbitrary bivector operator is a linear combination of N bivector operators $R^{(j)}_{\alpha\beta}$ for any simple Lie algebras except for the algebra A_n ($n\geqslant2$), and D_4. In other words, these algebras can have at most N linearly independent bivector operators. From this, we can again prove Eqs. (1.1) and (1.2) for all exceptional algebras and A_1 by the same reasoning given in the previous section.

Next, let us give explicit forms of the decomposition (3.1) for all exceptional Lie algebras as well as A_1 and A_2. To this end, we adopt the lexicographical numbering of simple roots as in Ref. 2. Let $\Lambda^{(j)}$ denote the highest weight of the jth irreducible representation $\{\lambda_j\}$ in (3.1), while we write n fundamental weights as $\Lambda_1,\Lambda_2,\dots,\Lambda_n$. Then $\{\lambda_0\}\otimes\{\lambda_0\}$ always contains the trivial representation $\{0\}$ and at least one adjoint representation $\{\lambda_0\}$, so that we label

$$\{\lambda_1\}=\{0\},\quad \{\lambda_2\}=\{\lambda_0\}. \qquad (3.9)$$

Moreover, it always contains an irreducible representation $\{\lambda_3\}$ whose highest weight $\Lambda^{(3)}$ is equal to $2\Lambda_0$, where Λ_0 is the highest weight of the adjoint representation. Thus, we identify

$$\{\lambda_3\}\equiv\{\Lambda^{(3)}\},\qquad \Lambda^{(3)}=2\Lambda_0. \qquad (3.10)$$

This is enough for the algebra A_1 with $N=3$, i.e.,

$$A_1: \{\lambda_0\}\otimes\{\lambda_0\} = \{0\}\oplus\{\lambda_0\}\oplus\{2\Lambda_0\}, \tag{3.11}$$

while we have to specify further $\{\lambda_4\}$ and $\{\lambda_5\}$ for exceptional Lie algebras. Explicit computations show that we have

(i) $G_2: \Lambda^{(4)} = 2\Lambda_2,\quad \Lambda^{(5)} = 3\Lambda_2,$ (3.12a)

(ii) $F_4: \Lambda^{(4)} = 2\Lambda_4,\quad \Lambda^{(5)} = \Lambda_2,$ (3.12b)

(iii) $E_6: \Lambda^{(4)} = \Lambda_1+\Lambda_5,\quad \Lambda^{(5)} = \Lambda_3,$ (3.12c)

(iv) $E_7: \Lambda^{(4)} = \Lambda_5,\quad \Lambda^{(5)} = \Lambda_2,$ (3.12d)

(v) $E_8: \Lambda^{(4)} = \Lambda_7,\quad \Lambda^{(5)} = \Lambda_2.$ (3.12e)

In deriving these Clebsch–Gordan series, the table of the second and fourth indices by McKay and Patera[10] has been very helpful. We may note that for all exceptional Lie algebras, the irreducible representations $\{\rho\} = \{\lambda_5\}$ given in Eqs. (3.12) are chosen to satisfy Eqs. (1.12) and (1.13a), i.e.,

$$I_2(\lambda_5) = 2I_2(\lambda_0),\quad d(\lambda_5) = \tfrac{1}{2}d(\lambda_0)[d(\lambda_0)-3]. \tag{3.13}$$

The reason for this validity will be explained in the next section. Similarly, all $d(\lambda_j)$ and $I_2(\lambda_j)$ for $j=3$ and 4 can be expressed as in Eqs. (1.16) and (1.17) which can be shown also in the next section. Then, those identities imply

$$\xi_1 = -I_2(\lambda_0),\quad \xi_2 = -\tfrac{1}{2}I_2(\lambda_0),\quad \xi_3 = \alpha I_2(\lambda_0),$$
$$\xi_4 = \beta I_2(\lambda_0),\quad \xi_5 = 0, \tag{3.14}$$

where we have set

$$\alpha = \tfrac{1}{12}(\theta-1),\quad \beta = -\tfrac{1}{12}(\theta+1), \tag{3.15}$$

$$\theta = \left(\frac{242+d(\lambda_0)}{2+d(\lambda_0)}\right)^{1/2}. \tag{3.16}$$

Note that α and β are solutions of the quadratic equation

$$6t^2 + t - \frac{10}{2+d(\lambda_0)} = 0. \tag{3.17}$$

If we define J by

$$J = \frac{1}{I_2(\lambda_0)}Q = \frac{1}{I_2(\lambda_0)}g^{\mu\nu}\mathrm{ad}x_\mu\otimes\mathrm{ad}x_\nu, \tag{3.18}$$

Eqs. (3.3) and (3.14) imply the validity of

$$J(J+1)(2J+1)\left(6J^2+J-\frac{10}{2+d(\lambda_0)}\right) = 0. \tag{3.19}$$

Moreover, the explicit forms of projection operators P_j for exceptional Lie algebras are found to be

$$P_1 = \frac{1}{d(\lambda_0)}G^{\mu\nu}\otimes G_{\mu\nu}, \tag{3.20a}$$

$$P_2 = \frac{1}{I_2(\lambda_0)}(\mathrm{ad}x_\mu\mathrm{ad}x_\nu)\otimes G^{\mu\nu}$$
$$= \frac{1}{I_2(\lambda_0)}G^{\mu\nu}\otimes(\mathrm{ad}x_\mu\mathrm{ad}x_\nu), \tag{3.20b}$$

$$P_3 = \frac{J(J+1)(2J+1)(J-\beta)}{\alpha(\alpha+1)(2\alpha+1)(\alpha-\beta)}, \tag{3.20c}$$

$$P_4 = \frac{J(J+1)(2J+1)(J-\alpha)}{\beta(\beta+1)(2\beta+1)(\beta-\alpha)}, \tag{3.20d}$$

$$P_5 = \tfrac{1}{2}(E\otimes E - G^{\mu\nu}\otimes G_{\nu\mu}) - P_2. \tag{3.20e}$$

Corresponding to Eq. (3.14), we also find

$$JP_1 = P_1J = -P_1, \tag{3.21a}$$
$$JP_2 = P_2J = -\tfrac{1}{2}P_2, \tag{3.21b}$$
$$JP_3 = P_3J = \alpha P_3, \tag{3.21c}$$
$$JP_4 = P_4J = \beta P_4, \tag{3.21d}$$
$$JP_5 = P_5J = 0, \tag{3.21e}$$

as well as identity

$$6J^2 + J = P_2 + \frac{5}{2+d(\lambda_0)}\times\{d(\lambda_0)P_1 + G^{\mu\nu}\otimes G_{\nu\mu} + E\otimes E\}. \tag{3.22}$$

We remark that Eqs. (3.20a), (3.20b), (3.20e), (3.21a), (3.21b), and (3.21e) are valid for all simple Lie algebras, while Eqs. (3.19), (3.20c), (3.20d), (3.21c), (3.21d), and (3.22) are only valid for exceptional Lie algebras. For the latter cases, we used Eqs. (1.1) and (1.2) for $\{\lambda\} = \{\lambda_0\}$. The proof of these identities are quite involved and will not be given here.

For the algebra A_2, we have $N=6$ with

$$\{\lambda_0\}\otimes\{\lambda_0\} = \{0\}\oplus\{\lambda_0\}\oplus\{\lambda_0\}\oplus\{2\Lambda_0\}\oplus\{3\Lambda_1\}\oplus\{3\Lambda_2\}. \tag{3.23}$$

Note that $\{\rho\} = \{3\Lambda_1\}$ and $\{\rho^*\} = \{3\Lambda_2\}$ satisfy Eqs. (1.12) and (1.13b) now. If we identify $\{\lambda_4\} = \{\lambda_0\}$, $\{\lambda_5\} = \{\rho\}$ and $\{\lambda_6\} = \{\rho^*\}$, then Eqs. (1.16) and (1.17) are still valid by the reason which will be explained in the next section.

Finally, with respect to other simple Lie algebras, we mention that the special irreducible representation $\{\rho\}$ satisfying Eqs. (1.12) and (1.13) is found to be

(i) A_n $(n\geqslant 2)$

$$\{\rho\} = \{2\Lambda_1+\Lambda_{n-1}\}, \tag{3.24a}$$
$$\{\rho^*\} = \{\Lambda_2+2\Lambda_n\},$$

(ii) B_3 $(n=3)$

$$\{\rho\} = \{\Lambda_1+2\Lambda_3\}, \tag{3.24b}$$

(iii) B_n $(n\geqslant 4)$ and D_n $(n\geqslant 5)$

$$\{\rho\} = \{\Lambda_1+\Lambda_3\}, \tag{3.24c}$$

(iv) D_4 $(n=4)$

$$\{\rho\} = \{\Lambda_1+\Lambda_3+\Lambda_4\}, \tag{3.24d}$$

(v) C_n $(n\geqslant 2)$

$$\{\rho\} = \{2\Lambda_1+\Lambda_2\}, \tag{3.24e}$$

all of which appear in the decomposition of Eq. (3.1).

4. SECOND AND FOURTH ORDER INDICES

Let $\mathfrak{H}_j$ $(j=1,2,\ldots,n)$ and E_α be standard Cartan–Weyl bases of any simple Lie algebra. Let $\Delta(\lambda)$ be the weight system of the representation $\{\lambda\}$ which need not be irreducible. Then, following Dynkin[11] and Patera *et al.*,[9] let us define even-dimensional indices $l_{2p}(\lambda)$ by

$$l_{2p}(\lambda) = \mathrm{Tr}(\mathfrak{H},\mathfrak{H})^p = \sum_{M\in\Delta(\lambda)} (\mathfrak{M},\mathfrak{M})^p, \tag{4.1}$$

where the summation is over all weights $\mathfrak{M}$ belonging to the representation $\{\lambda\}$. Then, these have the following properties:

(i) $l_0(\lambda) = d(\lambda)$, (4.2)

(ii) $l_2(\lambda) = n\,\dfrac{d(\lambda)}{d(\lambda_0)} I_2(\lambda)$, (4.3)

(iii) $l_2(\{\lambda_A\}\otimes\{\lambda_B\}) = d(\lambda_A)l_2(\lambda_B) + d(\lambda_B)l_2(\lambda_A)$, (4.4)

(iv) $l_{2p}(\{\lambda_A\}\oplus\{\lambda_B\}) = l_{2p}(\lambda_A) + l_{2p}(\lambda_B)$, (4.5)

where Eq. (4.3) presupposes $\{\lambda\}$ to be irreducible. Moreover, let us consider a product of two irreducible representations $\{\lambda_A\}\otimes\{\lambda_B\}$ and decompose it into a direct sum of N irreducible representations as in Eq. (3.1),

$$\{\lambda_A\}\otimes\{\lambda_B\} = \sum_{j=1}^{N} \oplus\{\lambda_j\}. \tag{4.6}$$

Then, from Eq. (4.5), we find

$$l_{2p}(\{\lambda_A\}\otimes\{\lambda_B\}) = \sum_{j=1}^{N} l_{2p}(\lambda_j). \tag{4.7}$$

Again, following (I), we set

$$Q = g^{\mu\nu} X_\mu^{(A)} \otimes X_\nu^{(B)}, \tag{4.8}$$

where $X_\mu^{(A)}$ and $X_\nu^{(B)}$ are representation matrices of x_μ in irreducible representations $\{\lambda_A\}$ and $\{\lambda_B\}$, respectively. Then, as in (I), we find

$$Q^l = \sum_{j=1}^{N} (\xi_j)^l P_j \quad (l = 0,1,2,\cdots), \tag{4.9}$$

$$\xi_j = \tfrac{1}{2}[I_2(\lambda_j) - I_2(\lambda_A) - I_2(\lambda_B)], \tag{4.10}$$

where P_j is the projection operator for the jth irreducible space $\{\lambda_j\}$. These generalize Eqs. (3.2)–(3.4) which are a special case of $\{\lambda_A\} = \{\lambda_B\} = \{\lambda_0\}$. Taking the full trace o both sides of Eq. (4.9), this gives

$$g^{\mu_1\nu_1}g^{\mu_2\nu_2}\cdots g^{\mu_l\nu_l}\mathrm{Tr}(X_{\mu_1}^{(A)}X_{\mu_2}^{(A)}\cdots X_{\mu_l}^{(A)})\ \mathrm{Tr}(X_{\nu_1}^{(B)}X_{\nu_2}^{(B)}\cdots X_{\nu_l}^{(B)}) = \sum_{j=1}^{N} d(\lambda_j)(\xi_j)^l \tag{4.11}$$

for all nonnegative integers $l = 0,1,\cdots$. For the special cases $l = 0$ and $l = 1$, this reproduces the results

$$d(\lambda_A)d(\lambda_B) = \sum_{j=1}^{N} d(\lambda_j), \tag{4.12}$$

$$\sum_{j=1}^{N} \xi_j d(\lambda_j) = 0, \tag{4.13}$$

where we used $\mathrm{Tr}X_\mu = 0$ in deriving (4.13). We can rewrite Eq. (4.13) as a more familiar form of

$$d(\lambda_A)l_2(\lambda_B) + d(\lambda_B)l_2(\lambda_A) = \sum_{j=1}^{N} l_2(\lambda_j). \tag{4.14}$$

For $l = 2$, Eq. (4.11) gives a nontrivial result of

$$\frac{4l_2(\lambda_A)l_2(\lambda_B)}{d(\lambda_0)} + d(\lambda_A)d(\lambda_B)\left(\frac{l_2(\lambda_A)}{d(\lambda_A)} + \frac{l_2(\lambda_B)}{d(\lambda_B)}\right)^2 = \sum_{j=1}^{N} \frac{1}{d(\lambda_j)}[l_2(\lambda_j)]^2. \tag{4.15}$$

For $l = 3$, we use Eq. (2.31b) for all simple algebras other than A_n ($n\geqslant 2$) or for cases of self-contragradient representation of A_n to find

$$\sum_{j=1}^{N} \frac{[l_2(\lambda_j)]^3}{[d(\lambda_j)]^2} = -2\frac{l_2(\lambda_A)l_2(\lambda_B)l_2(\lambda_0)}{[d(\lambda_0)]^2} + 12\left(\frac{l_2(\lambda_A)}{d(\lambda_A)} + \frac{l_2(\lambda_B)}{d(\lambda_B)}\right)\frac{l_2(\lambda_A)l_2(\lambda_B)}{d(\lambda_0)} + \left(\frac{l_2(\lambda_A)}{d(\lambda_A)} + \frac{l_2(\lambda_B)}{d(\lambda_B)}\right)^3 d(\lambda_A)d(\lambda_B). \tag{4.16}$$

Finally, for $l = 4$, we use Eqs. (2.32), (2.25), and (2.26) to find a rather complicated formula

$$\frac{d(\lambda_A)d(\lambda_B)}{12[2+d(\lambda_0)]d(\lambda_0)}\left(6 - \frac{I_2(\lambda_0)}{I_2(\lambda_A)}\right)\left(6 - \frac{I_2(\lambda_0)}{I_2(\lambda_B)}\right) \times [I_2(\lambda_A)I_2(\lambda_B)]^2 = \sum_{j=1}^{N}\left((\xi_j)^4 + \frac{1}{2n} l_2(\lambda_0)(\xi_j)^3\right)d(\lambda_j) + \frac{1}{24}\frac{1}{n^4} l_2(\lambda_A)l_2(\lambda_B)[l_2(\lambda_0)]^2 d(\lambda_0) \tag{4.17}$$

for all exceptional Lie algebras and A_1, as well as the case when at least one of $\{\lambda_A\}$ and $\{\lambda_B\}$ for the case A_2 is self-contragradient.

Finally, for all exceptional Lie algebras and A_1, as well as any self-contragradient representation of A_2, we choose all X_μ, X_ν, X_α, and X_β in Eq. (2.32) to be some members $\mathfrak{H}_j$'s of Cartan subalgebra. Then from the definition Eq. (4.1) for $l_4(\lambda)$, we find

$$l_4(\lambda) = \frac{n+2}{n}\frac{1}{2+d(\lambda_0)}\left(\frac{d(\lambda_0)}{d(\lambda)} - \frac{1}{6}\frac{l_2(\lambda_0)}{l_2(\lambda)}\right)[l_2(\lambda)]^2. \tag{4.18}$$

Moreover, after some calculations, we reproduce the relation

$$l_4(\{\lambda_A\}\otimes\{\lambda_B\}) = \sum_{j=1}^{N} l_4(\lambda_j) = d(\lambda_A)l_4(\lambda_B) + d(\lambda_B)l_4(\lambda_A) + \frac{2(n+2)}{n} l_2(\lambda_A)l_2(\lambda_B) \tag{4.19}$$

also for all exceptional Lie algebras as well as A_1 and A_2 in agreement with the result of Ref. 9. However, Eq. (4.19) except for A_2 is not independent of Eqs. (4.15) and (4.18).

For our present case, we are interested in the special case of $\{\lambda_A\} = \{\lambda_B\} = \{\lambda_0\}$. Then, since the adjoint representation $\{\lambda_0\}$ is self-contragradient even for the algebra A_n ($n\geqslant 2$), we can rewrite all Eqs. (4.11)–(4.17) for all cases of exceptional Lie algebras as well as A_1 and A_2, as follows:

(i) $\displaystyle\sum_{j=1}^{N} d(\lambda_j) = [d(\lambda_0)]^2$, (4.20a)

(ii) $\displaystyle\sum_{j=1}^{N} \xi_j d(\lambda_j) = 0$, (4.20b)

(iii) $\sum_{j=1}^{N} (\xi_j)^2 d(\lambda_j) = d(\lambda_0)[I_2(\lambda_0)]^2,$ (4.20c)

(iv) $\sum_{j=1}^{N} (\xi_j)^3 d(\lambda_j) = -\frac{1}{4} d(\lambda_0)[I_2(\lambda_0)]^3,$ (4.20d)

(v) $\sum_{j=1}^{N} (\xi_j)^4 d(\lambda_j) = \frac{1}{12}\frac{27 + d(\lambda_0)}{2 + d(\lambda_0)} d(\lambda_0)[I_2(\lambda_0)]^4.$ (4.20e)

Now, first consider the case of exceptional Lie algebras with $N = 5$. Then, we know $\{\lambda_1\} = \{0\}$, and $\{\lambda_2\} = \{\lambda_0\}$. Moreover, we shall prove soon that $\{\lambda_5\} = \{\rho\}$ satisfies Eqs. (1.12) and (1.13a). Then when $d(\lambda_0)$ and $I_2(\lambda_0)$ are given, only $d(\lambda_j)$ and $I_2(\lambda_j)$ for $j = 3$ and 4 are unknown for five equations in Eqs. (4.20). This completely determines $d(\lambda_j)$ and $I_2(\lambda_j)$ for $j = 3$ and 4 as in Eqs. (1.16) and (1.17). Actually, we may only assume Eq. (1.12), i.e., $I_2(\lambda_5) = 2I_2(\lambda_0)$, then Eq. (1.13a) will emerge also as a consequence of Eqs. (4.20). For the algebra A_2, we notice that $\{\lambda_5\} = \{\rho\}$ and $\{\lambda_6\} = \{\rho^*\}$ have the same dimension and same eigenvalue for I_2. Therefore, in spite of $N = 6$, we can essentially proceed exactly in the same way.

In ending this section, we shall give a reason for the validity of Eqs. (1.12) and (1.13). The product $\{\lambda_0\} \otimes \{\lambda_0\}$ can be first decomposed into a direct sum of symmetric $\{\lambda_S\}$ and antisymmetric $\{\lambda_A\}$ representations as

$$\{\lambda_0\} \otimes \{\lambda_0\} = \{\lambda_A\} \oplus \{\lambda_S\}. \quad (4.21)$$

The dimensions of these in general reducible representations are obviously

$$d(\lambda_A) = \tfrac{1}{2} d(\lambda_0)[d(\lambda_0) - 1], \quad (4.22)$$

$$d(\lambda_S) = \tfrac{1}{2} d(\lambda_0)[d(\lambda_0) + 1]. \quad (4.23)$$

Also, from the definition (4.1), we can compute their second indices as

$$l_2(\lambda_A) = \sum_{M_1 > M_2} (M_1 + M_2, M_1 + M_2) = [d(\lambda_0) - 2] l_2(\lambda_0), \quad (4.24)$$

$$l_2(\lambda_S) = \sum_{M_1 \geqslant M_2} (M_1 + M_2, M_1 + M_2) = [d(\lambda_0) + 2] l_2(\lambda_0), \quad (4.25)$$

where M_1 and M_2 are weights of two adjoint spaces. Now, the antisymmetric part $\{\lambda_A\}$ turns out to be always reducible as a direct sum of the adjoint representation $\{\lambda_0\}$ and one additional irreducible representation $\{\rho\}$ to be

$$\{\lambda_A\} = \{\lambda_0\} \oplus \{\rho\} \quad (4.26)$$

for all simple Lie algebras except for A_n $(n \geqslant 1)$ and $D_3 = A_3$. Then, since we have

$$d(\lambda_A) = d(\lambda_0) + d(\rho), \quad (4.27)$$

$$l_2(\lambda_A) = l_2(\lambda_0) + l_2(\rho), \quad (4.28)$$

Eqs. (4.22) and (4.24) give the desired relations, Eqs. (1.12) and (1.13a) for these algebras. For the algebra A_1, we have simply

$$\{\lambda_A\} = \{\lambda_0\}, \quad d(\lambda_0) = 3$$

so that it gives nothing new. For the algebra A_n $(n \geqslant 2)$, we find

$$\{\lambda_A\} = \{\lambda_0\} \oplus \{\rho\} \oplus \{\rho^*\}, \quad (4.29)$$

where $\{\rho^*\}$ is the inequivalent representation antigradient to $\{\rho\}$. Since we have $d(\rho^*) = d(\rho)$ and $l_2(\rho^*) = l_2(\rho)$, we find in this case Eqs. (1.12) and (1.13b).

For the exceptional Lie algebras, we can identify

$$\{\lambda_S\} = \{\lambda_1\} \oplus \{\lambda_3\} \oplus \{\lambda_4\}, \quad (4.30a)$$

$$\{\lambda_A\} = \{\lambda_2\} \oplus \{\lambda_5\}. \quad (4.30b)$$

Finally, we simply mention an empirical relation

$$(n + 2)\theta + m = 18 \quad (4.31)$$

for all exceptional algebras as well as A_1 and A_2, where the integer m is defined by

$$m = n \text{ (modulo 4)}, \quad (4.32a)$$

$$1 \leqslant m \leqslant 4 \quad \text{for } G_2, F_4, E_6, E_7, \text{ and } E_8, \quad (4.32b)$$

$$-3 \leqslant m \leqslant -1 \quad \text{for } A_1 \text{ and } A_2. \quad (4.32c)$$

Note added in proof: After this paper was written, Dr. Cvitanovic noted that our main equation, Eq. (1.1) with (1.2), is also valid for the adjoint representation of the algebra D_4 which corresponds to the SO(8) group. However, it is not satisfied in general by other irreducible representations of D_4 since D_4 is known to have two genuine fourth order Casimir invariants. The author would like to express his gratitude to Dr. Cvitanovic for this information as well as many other communications.

[1] A. Borel and C. Chevalley, Am. Math. Soc. Mem. **14**, 1 (1955); G. Warner, *Harmonic Analysis in Semi-Simple Lie Groups (I)* (Springer, New York 1972), p. 144.

[2] J. Patera and D. Sankoff, *Tables of Branching Rules for Representations of Simple Lie Algebras* (Les Presses de L'Université de Montreal, Montreal, Canada, 1973); A.K. Bose and J. Patera, J. Math. Phys. **11**, 2231 (1970).

[3] S. Okubo, J. Math. Phys. **18**, 2382 (1977). This paper will herafter be referred to as (I).

[4] P. Cvitanović, private communication, where he uses a technique developed in P. Cvitanović, Phys. Rev. D **14**, 1536 (1976) and University of Oxford Report 40/77 (1977) (unpublished). The author would like to express his gratitude to Dr. Cvitanović for informing him of these results.

[5] S. Okubo, Hadronic J. **1**, 1250 (1978). Also, the identity Eq. (1.1) has been successfully applied to eliminate the presence of any exceptional Lie algebra in any unified gauge theory of weak and electromagnetic interactions which are compatible with the two charge ansatz. See S. Okubo, Phys. Rev. D **18**, 3792 (1978).

[6] J.D. Louck and L.C. Biedenharn, J. Math. Phys. **11**, 2368 (1970); **14**, 1336 (1973); L.C. Biedenharn and J.D. Louck, J. Math. Phys. **13**, 1985 (1972); L.C. Biedenharn, J.D. Louck, F. Chacón, and M. Ciftan, J. Math. Phys. **13**, 1957 (1972).

[7] Actually, we can generalize this to a slightly more general class of classical Lie algebras as is defined in (I).

[8] S. Okubo, Phys. Rev. D **16**, 3528 (1977).

[9] J. Patera, R.T. Sharp, and P. Winternitz, J. Math. Phys. **17**, 1972 (1976); **18**, 1519 (Erratum) (1977). Note that our notation $l_{2p}(\lambda)$ corresponds to $I^{(2p)}(\lambda)$ of theirs.

[10] W. McKay and J. Patera, "Tables of dimensions, second, and fourth indices of representations of simple Lie algebras," Université de Montreal, Montreal, Canada CRM-742 (1977).

[11] E.B. Dynkin, Mat. Sb., N.S. **30** (72), 349 (1952) [Amer. Math. Soc. Transl. Ser. 2, **6**, 111 (1957)].

Spectral and scattering theory for the adiabatic oscillator and related potentials[a)]

Matania Ben-Artzi[b)] and Allen Devinatz

Department of Mathematics, Northwestern University, Evanston, Illinois 60201
(Received 28 August 1978)

We consider the Schrödinger operator $H = -\Delta + V(r)$ on R^n, where $V(r) = a \sin(br^{\alpha})/r^{\beta} + V_S(r)$, $V_S(r)$ being a short range potential and $\alpha > 0$, $\beta > 0$. Under suitable restrictions on α, β, but always including $\alpha = \beta = 1$, we show that the absolutely continuous spectrum of H is the essential spectrum of H, which is $[0, \infty)$, and the absolutely continuous part of H is unitarily equivalent to $-\Delta$. We use these results to show the existence and completeness of the Møller wave operators. Our results are obtained by establishing the asymptotic behavior of solutions of the equation $Hu = zu$ for complex values of z.

1. INTRODUCTION

We shall consider the spectral and scattering theory associated with a Schrödinger operator on R^n

$$H = -\Delta + V, \tag{1.1}$$

where V is a real spherically symmetric potential of a type related to the adiabatic oscillator studied by von Neumann and Wigner.[1]

It is very well known that the rate of decay of V at infinity plays a significant role in determining the spectral properties of H. When V is a short-range potential [i.e., roughly, $|V(x)| \sim |x|^{-1-\epsilon}$, $\epsilon > 0$, as $|x| \to \infty$] there now exists a complete theory (see Agmon[2]) proving that the essential spectrum of H coincides with its absolutely continuous spectrum (apart possibly from a discrete sequence of eigenvalues if V is not sufficiently regular), and also that the Møllerwave operators exist and are complete. For certain long-range potentials (i.e, roughly, $|V(x)| \sim |x|^{-\epsilon}$, $\epsilon > 0$, as $|x| \to \infty$ and short-range decay of the radial derivative), spectral problems have been studied by Weidmann[3] for spherically symmetric potentials, and by Ikebe and Saito[4] and Lavine[5] for general potentials, with the common results that, again, the absolutely continuous spectrum coincides with the essential spectrum.

In this paper we shall suppose that V is a spherically symmetric potential of the form:

$$V(r) = \frac{a \sin br^{\alpha}}{r^{\beta}} + V_S(r), \quad r = |x|, \tag{1.2}$$

where $V_S(r)$ is a short-range perturbation, a, b, α, β real constants, $\alpha > 0$, $\beta > 0$. We shall always assume that the restriction of H to $C_0^{\infty}(R^n)$ (infinitely differentiable compactly supported functions) has a unique self-adjoint realization in $L^2(R^n)$, thus requiring certain regularity properties of V (to be specified in the body of the paper). However, the case $\alpha = \beta = 1$, corresponding to von Neumann and Wigner's adiabatic oscillator, will always be included.

Potentials of the type (1.2) have also been studied by Wintner,[6] Atkinson,[7] Simon,[8] Harris and Lutz,[9] Dollard and Friedman[10] and others. In these papers various results about the asymptotic behavior of eigenfunctions at infinity, structure of the essential spectrum, boundedness of the set of possible positive eigenvalues, and the existence of the Møller wave operators have been obtained.

Note that, if $\beta \leqslant 1$, this is no longer a short-range potential and, moreover, since the derivatives of V may exhibit no better decay at infinity than V itself, it does not belong to the family of long-range potentials mentioned above.

Because our potential is radially symmetric, we can reduce the study of H to the study of ordinary self-adjoint differential operators on the half-line $R^+ = (0, \infty)$. (This is done in Sec. 2 and the Appendix.) Our methods are based on obtaining asymptotic estimates for the solutions of $Hu = zu$, z complex (Sec. 3). The same scheme was followed by Dollard and Friedman,[11] but their methods do not appear to be suitable for the derivation of the kind of asymptotic information needed for a "limiting absorption principle" (with all its ramifications). Our asymptotic studies are based on the elegantly elementary techniques proposed by Harris and Lutz.[12] In Sec. 4 we study the spectral properties of H. By means of a "limiting absorption principle" (roughly speaking, the continuity of the resolvent "down" to the spectrum in an appropriate setting) we show that the essential spectrum of H is absolutely continuous, except for one possible eigenvalue, if $\alpha = 1$. Using multiplicity arguments for ordinary differential operators, we then show the unitary equivalence of the absolutely continuous part of H and $-\Delta$. In Sec. 5 we show the existence and completeness of the Møller wave operators. The *existence* of these wave operators for potentials of the form (1.2) has already been shown by Dollard and Friedman,[11] and we have essentially nothing new to add here, except for a few more values of α and β. However, what we do have to add is the *completeness* of the wave operators.

2. PRELIMINARIES

As we noted in the Introduction, we shall be considering Schrödinger operators $H = -\Delta + V(r)$ in $L^2(R^n)$. We shall write the potential as

[a)]Research partially supported by NSF Grants MCS78-02030 and MCS76-04976 A01.

[b)]Present address: Department of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455.

$$V(r) = V_L(r) + V_S(r),$$

$$V_L(r) = \frac{a \sin br^\alpha}{r^\beta}, \quad \alpha,\beta > 0.$$

In the remainder of the paper we shall impose the following general assumptions on V:

$$V(r) \in L^2_{\text{loc}}(0,\infty), \quad V(r) \to 0 \text{ as } r \to \infty, \tag{V1}$$

$$V_S(r) \in L^1(r_0,\infty), \quad \text{for } r_0 \text{ sufficiently large}, \tag{V2}$$

$$V(r) = O(r^{-2+\epsilon}), \quad \epsilon > 0, \quad \text{as } r \to 0, \tag{V3}$$

$$\int_0^1 V(r)^2 r^{n-1} dr < \infty. \tag{V3'}$$

We note that $(V3')$ follows from $(V3)$, except for $n = 1,2,3$.

As we shall point out in Sec. 4, these assumptions guarantee that the self-adjoint operator which is multiplication by the real potential V is relatively compact with respect to $-\Delta$. In turn this guarantees that $H | C_0^\infty(R^n)$ has a unique self-adjoint extension which we will again denote by H.

The above assumptions will also suffice to prove the absolute continuity of the essential spectrum. However, in order to prove the unitary equivalence of the absolutely continuous part of H and $-\Delta$, and the existence and completeness of the Møller wave operators, we will have to impose further restrictions on V.

The following theorem, which characterizes H as unitarily equivalent to the direct sum of one-dimensional differential operators, has appeared in various forms in the literature (see, e.g., Weidmann[14], Dollard and Friedman[15]):

Theorem 2.1: Let $\mu_j = j(j+n-2), j = 0,1,2,\ldots$, be the eigenvalues of the Laplace–Beltrami operator on the unit sphere of R^n. Let H_j be the ordinary differential operator on $R^+ = (0,\infty)$ given by

$$H_j = -\frac{d^2}{dr^2} + \frac{1}{r^2}\left[\mu_j + \frac{(n-1)(n-3)}{4}\right] + V(r).$$

Under the hypotheses $(V1)$–$(V3 =)$, if $j \geqslant 1$ or $n > 3$, the operator $H_j | C_0^\infty(R^n)$ is essentially self-adjoint and we denote its self-adjoint extension by H_j again. For $j = 0$ and $n = 1,2,3$, the formal operator H_j has two boundary values at $r = 0$. For $j = 0$, $n = 1,3$ we use H_0 to denote that self-adjoint extension which is given by the boundary condition $\lim_{r\to 0} u(r) = 0$. For $j = 0$, $n = 2$ we use H_0 to denote that self-adjoint extension given by the boundary condition $\lim_{r\to 0} r^{1/2}(\log r)^{-1} u'(r) = 0$.

Under the above specifications[16] for the self-adjoint operators $H_j, j = 0,1,2,\ldots$, the direct sum of the H_j, with H_j, $j \neq 0$, repeated

$$\frac{2j+n-2}{j}\binom{j+n-3}{j-1}$$

times, is unitarily equivalent to H.

Since for our purposes the presentation of this theorem as given in the references mentioned above is incomplete, we shall present a proof of Theorem 2.1 in the Appendix.

The following theorem is easily seen to hold, and we shall leave the details for the reader.

Theorem 2.2: Under the hypotheses and notation of Theorem 2.1 we have the following:

(a) $\lambda \in R$ is an eigenvalue of H if and only if it is an eigenvalue of at least one H_j.

(b) The essential spectrum of H contains the essential spectrum of each H_j.

By this theorem, in order to study the spectral structure of H it suffices to study the spectral structure of each H_j, $j = 0,1,\cdots$. An essential ingredient of studying the spectral structure of H_j are certain asymptotic lemmas which will be presented in the next section.

3. ASYMPTOTIC ESTIMATES

In this section we shall study the asymptotic properties of solutions of the equation

$$-\frac{d^2u}{dr^2} + \left(V(r) + \frac{\nu(j,n)}{r^2}\right)u = zu, \tag{3.1}$$

where z is in some complex neighborhood of a point in $R^+ = (0,\infty)$, $\nu(j,n) = \mu_j + (n-1)(n-3)/4$, and $V(r)$ is a potential of the type considered in Secs. 1 and 2. In this section we need only assume $(V2)$ for $V_S(r)$; i.e., V_S is integrable on (r_0,∞), for some sufficiently large r_0. In case $\alpha = 1$ we shall find it necessary to exclude the point $z_1 = b^2/4$ from the considerations of this section. In the sequel the reader should assume this is always the case, unless we make a statement to the contrary.

Let $z_0 \in R^+$ and let Ω be a sufficiently small complex disc around z_0 so that $0 \notin \Omega$. In case $\alpha = 1$ we shall also suppose that $z_1 \notin \Omega$, so that in particular $z_0 \neq z_1$. Let us set

$$\Omega^\pm = \{z \in \Omega \mid \pm \operatorname{Im} z > 0\},$$

$$\tilde{\Omega}^\pm = \{z \in \Omega \mid \pm \operatorname{Im} z \geqslant 0\},$$

$$\Omega_R = \{z \in \Omega \mid \operatorname{Im} z = 0\}.$$

The main result of this section is the following lemma.

Lemma 3.1: If $\alpha > 0, \beta > 0$, either $2\beta + \alpha > 2$ or $2\beta - \alpha > 0$, and $V_S \in L^1(r_0,\infty)$, then for $z \in \tilde{\Omega}^+$ Eq. (3.1) has a solution $\varphi_+(r,z)$ having the following properties:

(a) $\varphi_+(r,z)$ and $\varphi'_+(r,z)$ are continuous on $R^+ \times \tilde{\Omega}^+$.

(b) $\varphi_+(r,z) \sim e^{i\sqrt{z}r}, \quad \varphi'_+(r,z) \sim i\sqrt{z}e^{i\sqrt{z}r}, \quad r \to \infty.$

In particular, if we take $\operatorname{Im}\sqrt{z} > 0$, $\varphi_+(r,z)$ and $\varphi'_+(r,z)$ decay exponentially as $r \to \infty$.

Proof: Our proof will build on the techniques initiated by Harris and Lutz.[17] In what follows we shall always make the choice $\operatorname{Im}\sqrt{z} > 0$. Also in what follows we shall employ the following convention:

$X(r,z)$ shall denote a generic 2×2 matrix, continuous in z for each r, and such that there exists an $r_0 > 0$ and a function $x(r) \in L^1(r_0,\infty)$ which dominates $X(r,z)$ for all $z \in \tilde{\Omega}^+$, namely,

$$|X(r,z)| \leqslant x(r), \quad z \in \tilde{\Omega}^+, \quad r \geqslant r_0.$$

Equation (3.1) may be written equivalently as a first order 2×2 system in the form

$$v'(r,z) = \begin{pmatrix} 0 & 1 \\ V(r) + v(j,n)/r^2 - z & 0 \end{pmatrix} v(r,z). \quad (3.2)$$

Because of the hypothesis made on $V_S(r)$, namely $V_S \in L^1(r_0, \infty)$, this may be written, using the above convention, as

$$v'(r,z) = \left[\begin{pmatrix} 0 & 1 \\ V_L(r) - z & 0 \end{pmatrix} + X(r,z) \right] v(r,z). \quad (3.2')$$

We shall, at first, consider the case $\alpha + 2\beta > 2$, $\alpha \geqslant 1$. Let us set

$$P = \begin{pmatrix} 1 & i\sqrt{z} \\ i\sqrt{z} & z \end{pmatrix},$$

and make the transformation $v = Pw$. An elementary computation shows that Eq. (3.2′) is transformed into the equation

$$w'(r,z) = [\Lambda_0(z) + W(r,z) + X(r,z)]w(r,z), \quad (3.3)$$

where

$$W(r,z) = \frac{V_L(r)}{2z} \begin{pmatrix} -i\sqrt{z} & z \\ 1 & i\sqrt{z} \end{pmatrix},$$

$$\Lambda_0(z) = \begin{pmatrix} i\sqrt{z} & 0 \\ 0 & -i\sqrt{z} \end{pmatrix}.$$

Now set

$$Q(r,z) = \begin{pmatrix} 0 & q_{12}(r,z) \\ q_{21}(r,z) & 0 \end{pmatrix}.$$

We shall suppose that Q is continuously differentiable in r and that $Q(r,z) \to 0$ as $r \to \infty$. If we make the transformation $w = (I + Q)y$, then Eq. (3.3) is transformed (for all sufficiently large r) into

$$y'(r,z) = [\Lambda_0(z) + Y(r,z) + X(r,z)]y(r,z), \quad (3.4)$$

where

$$(I + Q)Y = \Lambda_0 Q - Q\Lambda_0 + W + WQ - Q'. \quad (3.5)$$

We now choose Q in an appropriate manner so as to simplify this equation as much as possible, namely,

$$Q' = \Lambda_0 Q - Q\Lambda_0 + W - \operatorname{diag} W. \quad (3.6)$$

This leads to the equations

$$q'_{12}(r,z) = 2i\sqrt{z}\, q_{12}(r,z) + \tfrac{1}{2} V_L(r), \quad (3.7a)$$

$$q'_{21}(r,z) = -2i\sqrt{z}\, q_{21}(r,z) + \frac{1}{2z} V_L(r). \quad (3.7b)$$

A solution to (3.7a) is given by

$$q_{12}(r,z) = c(z) e^{2i\sqrt{z}r} + \tfrac{1}{2} e^{2i\sqrt{z}r} \int_{r_0}^{r} e^{-2i\sqrt{z}s} V_L(s)\,ds,$$

where we shall take $c(z)$ to be

$$c(z) = -\tfrac{1}{2} \int_{r_0}^{\infty} e^{-2i\xi s} V_L(s)\,ds, \quad \xi = \operatorname{Re}\sqrt{z} > 0.$$

The conditional convergence of the defining integral for $c(z)$ is easily seen after an integration by parts. It should be noted that, if $\alpha = 1$, we get conditional convergence only if $\xi \neq |b|/2$. The reader will see specific computations like this in the next few paragraphs. Thus we have

$$q_{12}(r,z) = \tfrac{1}{2} e^{2i\sqrt{z}r} \left[\int_{r_0}^{r} e^{-2i\sqrt{z}s} V_L(s)\,ds - \int_{r_0}^{\infty} e^{-2i\xi s} V_L(s)\,ds \right]. \quad (3.8)$$

Next, a solution to (3.7b) is given by

$$q_{21}(r,z) = d(z) \tfrac{1}{2} e^{-2i\sqrt{z}r} + \frac{1}{2z} e^{-2i\sqrt{z}r} \int_{r_0}^{r} e^{2i\sqrt{z}s} V_L(s)\,ds,$$

where we shall take $d(z)$ to be

$$d(z) = -\frac{1}{2z} \int_{r_0}^{\infty} e^{2i\sqrt{z}s} V_L(s)\,ds.$$

Again, it is easily established that the defining integral for $d(z)$ is absolutely convergent if $\operatorname{Im}\sqrt{z} > 0$ and conditionally convergent if $\operatorname{Im}\sqrt{z} = 0$. Thus we have

$$q_{21}(r,z) = -\frac{1}{2z} \int_{r_0}^{\infty} e^{-2i\sqrt{z}(r-s)} V_L(s)\,ds. \quad (3.9)$$

For fixed r, $q_{12}(r,z)$ and $q_{21}(r,z)$ are continuous in $z \in \tilde{\Omega}^+$. This is immediate if we establish that $c(z)$ and $d(z)$ are continuous in $\tilde{\Omega}^+$. In order to handle the integral for $c(z)$, we first write $\sin(bs^{\alpha})$ as a sum of exponentials and then perform an integration by parts using these exponentials. We have

$$\int_{r_0}^{\infty} e^{-2i\xi s} \frac{e^{\pm ibs^{\alpha}}}{s^{\beta}}\,ds = \int_{r_0}^{\infty} e^{-2i\xi s \pm ibs^{\alpha}} (-2i\xi \pm i\alpha b s^{\alpha-1}) g(s,z)\,ds,$$

where $g(s,z) = [s^{\alpha+\beta-1}(-2i\xi s^{1-\alpha} \pm i\alpha b)]^{-1}$. Notice that if $\alpha = 1$, we must exclude $z_1 = b^2/4$ so that we do not divide by zero. An integration by parts of the integral on the right yields

$$-e^{-2i\xi r_0 \pm ibr_0^{\alpha}} g(r_0,z) - \int_{r_0}^{\infty} e^{-2i\xi s \pm ibs^{\alpha}} \frac{dg(s,z)}{ds}\,ds.$$

Since we are taking $\alpha \geqslant 1$, an elementary computation shows that $dg(s,z)/ds$ is $O(1/s^{\alpha+\beta})$, independent of $z \in \tilde{\Omega}^+$. Thus $dg(s,z)/ds$ is dominated by an integrable function, independent of z, so that by Lebesgue's dominated convergence theorem, $c(z)$ is continuous in $\tilde{\Omega}^+$. Exactly the same technique shows that $d(z)$ is continuous in $\tilde{\Omega}^+$.

Our next objective is to get estimates on $q_{12}(r,z)$ and $q_{21}(r,z)$, namely,

$$|q_{12}(r,z)| + |q_{21}(r,z)| = O\left(\max\left\{\frac{1}{r^{\alpha+\beta-1}}, \frac{1}{r}\right\}\right). \quad (3.10)$$

The easier one to deal with is $q_{21}(r,z)$, and we start with it. We want to integrate the integral on the right-hand side of (3.9) by parts. As before, we replace $\sin(bs^\alpha)$ by a sum of exponentials. We then get

$$e^{-2i\sqrt{z}r}\int_r^\infty e^{2i\sqrt{z}s}\frac{e^{\pm ibs^\alpha}}{s^\beta}ds$$
$$= -e^{\pm ibr^\alpha}h(r,z) - e^{-2i\sqrt{z}r}\int_r^\infty e^{2i\sqrt{z}s\pm ibs^\alpha}\frac{dh(s,z)}{ds}ds,$$

where $h(s,z) = [s^{\alpha+\beta-1}(2i\sqrt{z}s^{1-\alpha} \pm i\alpha b)]^{-1}$. From this we immediately get the estimate (3.10) for $q_{21}(r,z)$.

Next, we get the same estimate for $q_{12}(r,z)$, but this is considerably more delicate. We rewrite (3.8) in the form

$$q_{12}(r,z) = -\tfrac{1}{2}e^{2i\sqrt{z}r}\int_r^\infty e^{-2i\xi s}V_L(s)ds$$
$$+\frac{e^{2i\sqrt{z}r}}{2}\int_{r_0}^r(e^{-2i\sqrt{z}s} - e^{-2i\xi s})V_L(s)ds.$$

The first integral can be treated in exactly the same way that we treated the integral for q_{21}. To estimate the second integral, let $\sqrt{z} = \xi + i\mu$, $\mu\geqslant 0$, and again replace $\sin(bs^\alpha)$ by sums of exponentials. Integrating by parts, we get

$$\int_{r_0}^r e^{-2i\sqrt{z}s}\frac{e^{\pm ibs^\alpha}}{s^\beta}ds$$
$$= e^{-2i\sqrt{z}s\pm ibs^\alpha}h(s,z)\Big|_{r_0}^r - \int_{r_0}^r e^{-2i\sqrt{z}s\pm ibs^\alpha}\frac{dh(s,z)}{ds}ds. \tag{3.11}$$

In the same way we have

$$\int_{r_0}^r e^{-2i\xi s}\frac{e^{\pm ibs^\alpha}}{s^\beta}ds$$
$$= e^{-2i\xi s\pm ibs^\alpha}g(s,z)\Big|_{r_0}^r - \int_{r_0}^r e^{-2i\xi s\pm ibs^\alpha}\frac{dg(s,z)}{ds}ds, \tag{3.12}$$

where $g(s,z)$ is the same function that we considered in the proof that $c(z)$ is continuous. We now subtract. The integrated term evaluated at r and then multiplied by $\exp(2i\sqrt{z}r)$ is

$$e^{\pm ibr^\alpha}[h(r,z) - e^{-2\mu r}g(r,z)] = O(1/r^{\alpha+\beta-1}),$$

independent of $z\in\widetilde{\Omega}^+$. The integrated term evaluated at r_0 is easily seen to be $O(\mu)$. so that this multiplied by $\exp(2i\sqrt{z}r)$ is estimated by $O(\mu\exp(-\mu r)) = O(1/r)$.

It remains to estimate the difference of the integrals on the right-hand side of (3.11) and (3.12). This difference may be written as

$$\int_{r_0}^r e^{-2i\xi\pm ibs^\alpha}\left\{\frac{dh(s,z)}{ds}(e^{2\mu s}-1)\right.$$
$$\left.+\left[\frac{dh(s,z)}{ds} - \frac{dg(s,z)}{ds}\right]\right\}ds.$$

The integral of the second term in the integrand is easily estimated and is seen to be $O(\mu)$. Thus when this is multiplied by $\exp(2i\sqrt{z}r)$, this term is estimated by $O(1/r)$. To evaluate the integral of the first term in the integrand we integrate by parts once more. An easy, if somewhat tedious, computation shows that this is estimated by

$$O\left(\frac{e^{2\mu r}}{r^{\alpha+\beta-1}}\right) + O(\mu) + O\left(\int_{r_0}^r\frac{e^{2\mu s}-1}{s^{\alpha+\beta+1}}ds\right). \tag{3.13}$$

We now write

$$\int_{r_0}^r\frac{e^{2\mu s}-1}{s^{\alpha+\beta+1}}ds = \int_{r_0}^{r/2}\frac{e^{2\mu s}-1}{s^{\alpha+\beta+1}}ds + \int_{r/2}^r\frac{e^{2\mu s}-1}{s^{\alpha+\beta+1}}ds.$$

If we apply the mean value theorem inside the first integral on the right, we see that it is estimated by

$$\mu\int_{r_0}^{r/2}\frac{e^{2\mu s}}{s^{\alpha+\beta}}ds.$$

The second integral on the right is estimated by

$$e^{2\mu r}\int_{r/2}^r\frac{ds}{s^{\alpha+\beta+1}} = O\left(\frac{e^{2\mu r}}{r^{\alpha+\beta}}\right).$$

If we use these estimates in (3.13) and then multiply by $\exp(2i\sqrt{z}r)$, we see that we have (3.10) for $q_{12}(r,z)$ also.

Let us now go back to (3.5) and (3.6). Noting that $(I+Q)^{-1} = I - (I+Q)^{-1}Q$, we have

$$Y = \operatorname{diag}W + (I+Q)^{-1}Q\operatorname{diag}W + (I+Q)^{-1}WQ. \tag{3.14}$$

The matrix W is order of magnitude $1/r^\beta$. Since we are supposing that $2\beta+\alpha>2$, it follows from (3.10) that the last two terms on the right in the expression (3.14) for Y are dominated by an integrable function which is independent of $z\in\widetilde{\Omega}^+$. Thus we see that we may write (3.4) as

$$y'(r,z) = [A_0(z) + \operatorname{diag}W(r,z) + X(r,z)]y(r,z). \tag{3.15}$$

It is now possible to use (3.15) and Levinson's method to get asymptotic estimates on $y(r,z)$. Before we do this, we want to transform (3.2′) into a form like (3.15) in the case when $2\beta-\alpha>0$. Notice that when $\alpha<1$ and $2\beta+\alpha>2$, then $2\beta-\alpha>0$, so when the lemma has been established for this latter case all possibilities will have been covered.

In order to treat the case $2\beta-\alpha>0$, we must transform (3.2′) a little differently than we did previously. Let

$$\lambda(r,z) = i\sqrt{z - V_L(r)}$$

be an eigenvalue of the first matrix in (3.2′). If $z\in\widetilde{\Omega}^+$ and r_0 is sufficiently large, then for $r\geqslant r_0$, $\operatorname{sgn}\operatorname{Re}\lambda(r,z) = -\operatorname{sgn}\operatorname{Im}\sqrt{z}\leqslant 0$. Let us set

$$P(r,z) = \begin{pmatrix}1 & 1\\ \lambda(r,z) & -\lambda(r,z)\end{pmatrix}.$$

We find that

$$P^{-1}(r,z)P'(r,z,) = \frac{1}{2}\frac{\lambda'(r,z)}{\lambda(r,z)}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}.$$

If we make the transformation $v(r,z) = P(r,z)w(r,z)$, we get

$$w'(r,z) = [A_0(z) + W(r,z) + X(r,z)]w(r,z),$$

where $\Lambda_0(z)$ is taken as before and

$$W(r,z)=\begin{pmatrix}\lambda(r,z)-\lambda_0(z) & 0\\ 0 & -\lambda(r,z)+\lambda_0(z)\end{pmatrix}$$

$$+\frac{1}{2}\frac{\lambda'(r,z)}{\lambda(r,z)}\begin{pmatrix}-1 & 1\\ 1 & -1\end{pmatrix}. \quad (3.16)$$

We now choose $Q(r,z)$ as before and make the transformation $w=(I+Q)y$. We are again led to formulas (3.5) and (3.6). However, this time we have

$$q_{12}'(r,z)=2i\sqrt{z}q_{12}(r,z)+\frac{1}{2}\frac{\lambda'(r,z)}{\lambda(r,z)}, \quad (3.17a)$$

$$q_{21}'(r,z)=-2i\sqrt{z}q_{21}(r,z)+\frac{1}{2}\frac{\lambda'(r,z)}{\lambda(r,z)}. \quad (3.17b)$$

We can now proceed exactly as we did before, but using λ'/λ in place of V_L. Exactly analogously to (3.8) and (3.9) we get now

$$q_{12}(r,z)=\frac{1}{2}e^{2i\sqrt{z}r}\times\left[\int_{r_0}^{r}e^{-2i\sqrt{z}s}\frac{\lambda'(s,z)}{\lambda(s,z)}ds-\int_{r_0}^{\infty}e^{-2i\xi s}\frac{\lambda'(s,z)}{\lambda(s,z)}\right]ds,$$

$$q_{21}(r,z)=-\frac{1}{2}\int_{r_0}^{\infty}e^{-2i\sqrt{z}(r-s)}\frac{\lambda'(s,z)}{\lambda(s,z)}ds.$$

We now note that

$$\frac{\lambda'(r,z)}{\lambda(r,z)}=-\frac{1}{2}\frac{V_L'(r)}{z-V_L(r)},$$

$$V_L'(r)=\frac{ab\alpha\cos(br^{\alpha})}{r^{\beta-\alpha+1}}-\frac{a\beta\sin(br^{\alpha})}{r^{1+\beta}}.$$

Since the second term of V_L' is integrable, we may relegate this part of the W matrix to the X term, so that we are working only with the term $ab\alpha\cos(br^{\alpha})r^{\alpha-\beta-1}$. In getting estimates on q_{12} and q_{21}, it is enough to suppose that $\alpha<1$, since $\alpha\geqslant1$ and $2\beta-\alpha>0$ implies $\alpha\geqslant1$ and $2\beta+\alpha>2$, which has already been worked out previously.

In this case we get the estimate

$$|q_{12}(r,z)|+|q_{21}(r,z)|=O\left(\max\left\{\frac{1}{r^{\beta-\alpha+1}},\frac{1}{r}\right\}\right),\quad \alpha<1,\ 2\beta-\alpha>0. \quad (3.18)$$

Let us indicate the details of the estimate for q_{21}, which is the easier case. Noting the form of q_{21}, λ'/λ, and V_L', and taking account of the above remarks, we replace $\cos(br^{\alpha})$ by the sum of exponentials and integrate by parts. We have

$$e^{-2i\sqrt{z}r}\int_r^{\infty}\frac{e^{2i\sqrt{z}s}}{(z-V_L(s))}\frac{e^{i\pm bs^{\alpha}}}{s^{\beta-\alpha+1}}ds$$

$$=-e^{\pm ibr^{\alpha}}h(r,z)-e^{-2i\sqrt{z}r}\times\int_r^{\infty}e^{2i\sqrt{z}s\pm ibs^{\alpha}}\frac{dh(s,z)}{ds}ds,$$

where

$h(s,z)=\{s^{\beta-\alpha+1}(2i\sqrt{z}\pm i\alpha bs^{\alpha-1})[z-V_L(s)]^{-1}\}$. The function $h(s,z)$ is order of magnitude $O(s^{\alpha-\beta-1})$ so that the integrated term above is of this order of magnitude. The derivative of $h(s,z)$ has terms of order of magnitude $O(s^{\alpha-\beta-2})$ and a term of the form $O(s^{2(\alpha-\beta-1)})\times[z-V_L(s)]^{-2}$, where $O(s^{2(\alpha-\beta-1)})$ has a derivative of order of magnitude $O(s^{2(\alpha-\beta)-3})$ provided $\alpha\leqslant1$. The terms $O(s^{\alpha-\beta-2})$ integrate out to terms of order of magnitude $O(r^{\alpha-\beta-1})$. For the second mentioned term we integrate by parts one more time. We get terms which integrate out to order of magnitude $O(r^{2(\alpha-\beta-1)})$ and a term under the integral sign of order of magnitude $O(s^{3(\alpha-\beta-1)})$. This latter term integrates out to a term $O(r^{3(\alpha-\beta)-2})$. Thus we have

$$q_{21}(r,z)=O\left(\frac{1}{r^{\beta-\alpha+1}}+\frac{1}{r^{2(\beta-\alpha+1)}}+\frac{1}{r^{3(\beta-\alpha)+2}}\right)$$

$$=O\left(\frac{1}{r^{\beta-\alpha+1}}\left[1+\frac{1}{r^{(\beta-\alpha+1)}}+\frac{1}{r^{2(\beta-\alpha)+1}}\right]\right).$$

Now, if $\alpha<1$, then $\beta-\alpha+1>\beta>0$ and $3(\beta-\alpha)+2=2\beta-\alpha+2(1-\alpha)>0$ if $2\beta-\alpha>0$. Thus we see that we have (3.18) in this case. The proof of the estimate (3.18) for q_{12} in this case proceeds exactly as the proof of the estimate for q_{12} in (3.10), and we shall not burden the reader with the details.

From (3.16) we see that W is of order of magnitude $\max\{1/r^{\beta},1/r^{\beta-\alpha+1}\}$, as $r\to\infty$. Thus from (3.18) it follows that QW is order of magnitude $\max\{1/r^{\beta+1},1/r^{2\beta-\alpha+1},1/r^{\beta-\alpha+2}\},1/r^{2(\beta-\alpha+1)}\}$. Now, for $\alpha<1$, $2\beta-\alpha>0$ implies $\beta-\alpha+2>1$ and $2(\beta-\alpha+1)>1$. Hence QW is bounded by an integrable function which is independent of $z\in\widetilde{\Omega}^{+}$. Going back to (3.14), we see that we have (3.15) in this case as well. *In both cases diag W is conditionally integrable and goes to zero as* $r\to\infty$, *independent of* $z\in\widetilde{\Omega}^{+}$. In the second case we must restrict ourselves to $\alpha<1$ in order for this to happen. However, as we have already pointed out, we may do so without loss of generality.

We now use Levinson's[18] method to find the solution required by Lemma 3.1. Let us set

$$\lambda_1(r,z)=i\sqrt{z}+W_{11}(r,z),$$
$$\lambda_2(r,z)=-i\sqrt{z}+W_{22}(r,z). \quad (3.19)$$

Since diag $W\to0$ as $r\to\infty$, we see that there is an r_0 so tht for $r\geqslant r_0$ and $z\in\widetilde{\Omega}^{+}$,

$$\text{Re}[\lambda_1(r,z)-\lambda_2(r,z)]=-2\text{Re}\sqrt{z}+\text{Re}[W_{11}(r,z)-W_{22}(r,z)]\leqslant0.$$

Let B be the Banach space of two-dimensional complex vector valued functions $f(r,z)$ continuous on $[r_0,\infty)\times\widetilde{\Omega}^+$ and with the norm

$$\|f\| = \sup_{\substack{r_0\leqslant r<\infty \\ z\in\widetilde{\Omega}^+}} \left|\exp\left[-\int_{r_0}^{r}\lambda_1(s,z)ds\right]f(r,z)\right|,$$

where we are using any convenient norm for complex vectors. Let us set

$$\Psi(r,z) = \operatorname{diag}\left[\exp\left(\int_{r_0}^{r}\lambda_1\right), \exp\left(\int_{r_0}^{r}\lambda_2\right)\right].$$

Let $\Gamma: B\to B$ be the transformation given by

$$(\Gamma f)(r,z) = -\int_r^\infty \Psi(r,z)\Psi^{-1}(s,z)X(s,z)f(s,z)ds. \quad (3.20)$$

It is a relatively simple matter [using $\mathrm{Re}(\lambda_1-\lambda_2)\leqslant 0$] to get the estimate

$$\left|\exp(-\int_{r_0}^{r}\lambda_1)(\Gamma f)(r,z)\right| \leqslant \int_r^\infty |X(s,z)|ds\|f\|$$
$$\leqslant \int_r^\infty x(s)ds\|f\|. \quad (3.21)$$

We are choosing the matrix and vector norm so that we have $|X(s,z)f(s,z)| \leqslant |X(s,z)|\ |f(s,z)|$. Otherwise we get the inequality (3.21) up to a multiplicative constant. At any rate if we choose r_0 sufficiently large we get $\|\Gamma\|\leqslant\frac{1}{2}$.

Let $\psi_1(r,z)$ be the vector which is the first column of $\Psi(r,z)$. Clearly $\psi_1\in B$, and thus the function

$$\theta(r,z) = (I-\Gamma)^{-1}\psi_1(r,z) \quad (3.22)$$

also belongs to B. It is a straightforward verification that $\theta(r,z)$ satisfies the differential equation (3.15). Now $\theta - \psi_1 = \Gamma\theta$, so that

$$\exp(-\int_{r_0}^{r}\lambda_1)\theta - e_1 = \exp(-\int_{r_0}^{r}\lambda_1)\Gamma\theta, \quad e_1 = (1,0).$$

From the estimate (3.21), it is immediate that the term on the right goes to zero as $r\to\infty$, independent of $z\in\widetilde{\Omega}^+$.

Let us now transform $\theta(r,z)$ back to get a solution of (3.2). If we normalize this solution by the multiplicative constant $\exp[\int_{r_0}^\infty W_{11}(r,z)]$ and take the first component, we have a solution of (3.1) on $[r_0,\infty)\times\widetilde{\Omega}^+$ which satisfies all of the conditions of Lemma 3.1. Clearly, by standard techniques we can extend this solution back to the origin so as to get a solution on $R^+\times\widetilde{\Omega}^+$ which satisfies all of the conclusions of Lemma 3.1.

Lemma 3.2: Under the hypotheses of Lemma 3.1, for $z\in\Omega_R$ there exist two linearly independent solutions, $\eta_1^+(r,z)$ and $\eta_2^+(r,z)$ so that as $r\to\infty$

$$\eta_1^+(r,z)\sim e^{i\sqrt{z}r}, \quad \frac{d\eta_1^+(r,z)}{dr}\sim i\sqrt{z}e^{i\sqrt{z}r},$$
$$\eta_2^+(r,z)\sim e^{-i\sqrt{z}r}, \quad \frac{d\eta_2^+(r,z)}{dr}\sim -i\sqrt{z}e^{-i\sqrt{z}r}. \quad (3.23)$$

Proof: We have already obtained η_1^+ in Lemma 3.1. The solution η_1^+ is obtained in the same way. More precisely, let B_1 be the Banach space of complex vector valued functions $f(r,z)$ with

$$\|f\| = \sup_{\substack{r_0\leqslant r<\infty \\ z\in\Omega_R}} |f(r,z)|,$$

and let $\Gamma_1: B_1\to B_1$ be the map

$$(\Gamma_1 f)(r,z) = -\int_r^\infty \Psi(r,z)\Psi^{-1}(s,z)X(s,z)f(s,z)ds. \quad (3.24)$$

Noting that, for z real, $\mathrm{Re}\lambda_1(r,z) = \mathrm{Re}\lambda_2(r,z) = 0$, we find that if r_0 is sufficiently large $\|\Gamma_1\|\leqslant\frac{1}{2}$.

Now let ψ_1 be the vector which is the first column of Ψ and ψ_2 be the vector which is the second column of Ψ. Clearly $\psi_1, \psi_2\in B_1$. Now set

$$\theta_j(r,z) = (I-\Gamma_1)^{-1}\psi_j(r,z), \quad j=1,2. \quad (3.25)$$

Transform the functions $\theta_j(r,z)$ back to get solutions of (3.2). If we suitably normalize these latter solutions and take their first components, we have the solutions η_1^+ and η_2^+ of the lemma.

Notice that the solution $\theta(r,z)$ which was obtained in Lemma 3.1, when restricted to Ω_R, is the solution $\theta_1(r,z)$, and that $\varphi_+(r,z)$ when restricted to Ω_R is $\eta_1^+(r,z)$.

We shall now study the asymptotic behavior of solutions of (3.1) near the point $r=0$. The following result is well known (see Kodaira,[19] Dollard and Friedman[20],) but we shall give a proof here for completeness.

Lemma 3.3: Let (V3) hold, $\gamma\geqslant -\frac{1}{2}$ be a solution to $\gamma(\gamma+1) = \nu(j,n)$, and let Ω be a bounded domain in the complex plane. Then for every $z\in\Omega$ Eq. (3.1) has a solution $\eta^0(r,z)$ having the following properties:

(a) $\eta^0(r,z)$ and $d\eta^0(r,z)/dr$ are continuous on $R^+\times\Omega$ and for each r are analytic in z,

(b) $\eta^0(r,z)\sim r^{\gamma+1}$ as $r\to 0$,

(c) $\lim_{r\to 0} r^{1/2}(\log r)^{-1}\dfrac{d\eta^0(r,z)}{dr} = 0$ if $\gamma = -\frac{1}{2}$.

Proof: Equation (3.1) may be written in the form

$$\frac{d}{dr}\left(r^{2(\gamma+1)}\frac{d}{dr}(r^{-(\gamma+1)}u)\right) = r^{\gamma+1}[V(r)-z]u.$$

If we let $w(r,z) = r^{-(\gamma+1)}u$ and $q(r,z) = r[V(r)-z]$, then

$$\frac{d}{dr}\left(r^{2(\gamma+1)}\frac{d}{dr}w(r,z)\right) = r^{2\gamma+1}q(r,z)w(r,z). \quad (3.26)$$

Because of the hypotheses (V3), it is clear that $q(r,z)$ is integrable near $r=0$, and indeed there is an ϵ, $0<\epsilon\leqslant 1$, such that

$$\int_{r_0}^{r}|q(s,z)|ds\leqslant Cr^\epsilon,$$

where C is independent of z in Ω.

Let $r_0>0$ be sufficiently small so that

$$C\int_0^{r_0}s^{-1+\epsilon}ds\leqslant\tfrac{1}{2},$$

and let $B = \mathscr{L}^\infty(0,r_0)$ the space of bounded Borel measurable functions on $(0,r_0)$ with the usual supremum norm. For each $z\in\Omega$ let $\Gamma(z)$ be the operator on $B\to B$ defined by

$$[\Gamma(z)g](r) = \int_0^r s^{-2(\gamma+1)}\int_0^s t^{2\gamma+1}q(t,z)g(t)dtds. \tag{3.27}$$

By the way we have chosen r_0 it is clear that $\|\Gamma\|\leqslant\frac{1}{2}$. Let $\psi(r) = 1$ in $[0,r_0]$ and let

$$w(r,z) = [I-\Gamma(z)]^{-1}\psi(r)$$
$$= 1 + \Gamma(z)[I-\Gamma(z)]^{-1}\psi(r).$$

Since, for every $g\in B$, $\Gamma(z)g(r) = o(1)$ as $r\to 0$, it follows that $w(r,z) = 1 + o(1)$ as $r\to 0$. The operator valued function $\Gamma(z)$ is analytic in Ω, and the analyticity and continuity properties of $w(r,z)$ follow immediately from this. Clearly, $w(r,z)$ is a solution of (3.6) in $(0,r_0]$ and can be extended to $R^+\times\Omega$ with the same continuity and analyticity properties so as to be a solution of (3.6) on R^+ for each $z\in\Omega$. If we now take $\eta^0(r,z) = r^{\gamma+1}w(r,z)$, we have proved (a) and (b).

To prove (c), we note that

$$\frac{d\eta^0(r,z)}{dr} = \tfrac{1}{2}r^{-1/2}w(r,z) + r^{1/2}w'(r,z).$$

Noting that $w' = (\Gamma w)'$ we see that $w' = O(r^{-1+\epsilon})$ as $r\to 0$. Since w is bounded in a neighborhood of zero, we have (c).

4. THE SPECTRUM OF H

Let us recall that if A is a self-adjoint operator, then its essential spectrum $\sigma_e(A)$ is the spectrum of A with the isolated eigenvalues of finite multiplicity removed. Let us also recall that if B is a symmetric operator and B is relatively compact with respect to A, then $A+B$ is self-adjoint and $\sigma_e(A) = \sigma_e(A+B)$. We refer the reader to Kato[21] and Schechter[22] as general references on this material.

In this paper we are taking A to be the self-adjoint realization of $-\Delta$ and B to be the operator which is multiplication by the real potential V. The underlying Hilbert space is, of course, $L^2(R^n)$. There are well-known sufficient conditions that V be relatively compact with respect to $-\Delta$ (see, e.g., Schechter[23]). For $V\in L^2_{\rm loc}(R^n)$ these conditions are

$$\sup_y\int_{|x|<1} V^2(x-y)|x|^{\mu-n}dx<\infty,\quad 0<\mu<4.$$

$$\int_{|x|<1} V^2(y-x)dx\to 0,\quad \text{as } |y|\to\infty.$$

These conditions will be satisfied if $V(r)$ satisfies $(V1)$, $(V3)$, and $(V3')$. Thus the essential spectrum of $H = -\Delta + V$ is the same as the essential spectrum of $-\Delta$, namely $R_0^+ = [0,\infty)$.

If we expand H into the direct sum of the operators H_j as stated in Theorem 2.1, then by Theorem 2.2(b) it follows that the essential spectrum of each H_j is contained in R_0^+. In fact, it is equal to R_0^+ as shown by the following lemma.

Lemma 4.1: Under the hypotheses $(V1)$, $(V3)$, and $(V3')$, the essential spectrum of H_j, $j = 0,1,2,\cdots$, is $R_0^+ = [0,\infty)$.

Proof: By the remarks made above we have only to show that the essential spectrum of H_j contains R_0^+. Let $H_j^{(1)}$ and $H_j^{(2)}$ be any self-adjoint realizations of the restrictions of the formal differential operator H_j to the intervals (0,1) and $[1,\infty)$, respectively. By Theorem XIII.7.4 in Dunford and Schwartz,[24] we have

$$\sigma_e(H_j) = \sigma_e(H_j^{(1)})\cup\sigma_e(H_j^{(2)}).$$

But by Corollary XIII. 7.13 in Dunford and Schwartz[25] we have

$$\sigma_e(H_j^{(2)}) = R_0^+,$$

implying that $R_0^+\subseteq\sigma_e(H_j)$.

In order to give the main theorem of this section, we shall briefly recall the meaning of the absolutely continuous spectrum of a self-adjoint operator. Let A be a self-adjoint operator defined in a Hilbert space $\mathscr{H}$ and $E(d\lambda)$ its canonical spectral measure. The subspace $\mathscr{H}_{\rm ac}(A)$ of absolute continuity is the collection of $u\in\mathscr{H}$ so that the measure $(E(d\lambda)u,u)$ is absolutely continuous with respect to Lebesgue measure. The closed subspace $\mathscr{H}_{\rm ac}(A)$ reduces A to a self-adjoint operator $A_{\rm ac}$. The spectrum of $A_{\rm ac}$ called the absolutely continuous spectrum of A and is denoted by $\sigma_{\rm ac}(A)$. We note that the absolutely continuous spectrum may have eigenvalues of A embedded in it. Another definition of the absolutely continuous spectrum if sometimes given as $\sigma_{\rm ac}(A)$ with all the eigenvalues removed. The definition we have adopted appears to be more popular at present.

Theorem 4.2: Under the hypotheses of Lemmas 3.1 and 4.1, if $\alpha\neq 1$, the operator H has no eigenvalues in $(0,\infty)$, and if $\alpha = 1$, it has at most one eigenvalue in $(0,\infty)$ at $b^2/4$. The space of absolute continuity of H is the orthogonal complement in $L^2(R^n)$ of the subspace generated by the eigenvectors of H. Thus $\sigma_{\rm ac}(H) = [0,\infty)$.

Proof: We have already noted that $\sigma_e(H) = \sigma_e(-\Delta) = [0,\infty)$ and that $H\,|\,C_0^\infty(R^n)$ is essentially self-adjoint. Let $E(d\lambda)$ be the canonical spectral measure of the self-adjoint realization of H. Assuming that the first statement in the theorem is true, in order to prove the remaining statements, it is enough to show that, for every $f\in L^2$, the measure $(E(d\lambda)f,f)$ is absolutely continuous when restricted to any compact interval in $(0,\infty)$ which avoids an eigenvalue of H.

From Theorem 2.1 we know that H is unitarily equivalent to a direct sum of ordinary differential operators H_j defined on $(0,\infty)$. The asymptotic estimates of (3.24) (in the proof of Lemma 3.1) show that if $\alpha\neq 1$ or if $z\neq b^2/4$, then H_j cannot have the eigenvalue z for $j = 0,1,2,\cdots$. Thus the first statement of Theorem 4.2 is true.

The canonical spectral measure $E(d\lambda)$ of H is unitarily equivalent to the direct sum of the canonical spectral measures $E_j(d\lambda)$ of H_j. In order to show that $(E(d\lambda)f,f)$ is absolutely continuous on a compact interval for every $f\in L^2(R^n)$, it is enough to show that this is true for a dense set in $L^2(R^n)$. Thus it is enough to show that each measure $(E_j(d\lambda)f,f)$ is absolutely continuous on every compact interval (which avoids the possible eigenvalue) for all $f\in L^2(R^+)$; and for the latter it is enough to prove it for a dense set in $L^2(R^+)$, say $C_0(R^+)$, the space of compactly supported continuous functions on R^+.

We follow the argument given in Ben-Artzi.[26] For $\mathrm{Im} z \neq 0$ let $R_j(z) = (H_j - zI)^{-1}$ be the resolvent for H_j. For $r,s \in R^+$, let $K_j(r,s;z)$ be the resolvent kernel for H_j, i.e., the kernel which satisfies

$$R_j(z)f(r) = \int_{R^+} K_j(r,s;z)f(s)ds, \quad f \in L^2(R^+). \tag{4.1}$$

There is a well-known formula (the Stieltjes inversion formula) which connects the spectral measure of H_j with its resolvent given by ($[\lambda_0,\lambda]$ avoids $b^2/4$)

$$\begin{aligned}(E_j([\lambda_0,\lambda])u,u) &= \lim_{\epsilon\to 0+} \frac{1}{\pi}\int_{\lambda_0}^{\lambda} \mathrm{Im}(R_j(\mu + i\epsilon)u,u)d\mu \\ &= \lim_{\epsilon\to 0+} \frac{1}{\pi}\int_{\lambda_0}^{\lambda} \mathrm{Im}\int_{R^+\times R^+} K_j(r,s;\mu + i\epsilon)u(s)\bar{u}(r)dsdrd\mu.\end{aligned} \tag{4.2}$$

From well-known results of the theory of ordinary differential equations, $K_j(r,s;z)$ is continuous on $R^+\times R^+\times C^+$ (see Dunford and Schwartz[27]). Suppose $\widetilde{\Omega}^+$ avoids a possible eigenvalue of H_j and suppose that $K_j(r,s;z)$ can be extended continuously from $R^+\times R^+\times \Omega^+$ to $R^+\times R^+\times \widetilde{\Omega}^+$. If $u \in C_0(R^+)$, then in (4.2) we may interchange the limit and the integral. This shows that $(E_j(d\lambda)u,u)$ is absolutely continuous on Ω_R.

In order to show that $K_j(r,s;z)$ may be continuously extended to $R^+\times R^+\times \widetilde{\Omega}^+$ we shall exhibit an explicit formula for K_j. Toward this end we recall the following well known fact (Dunford and Schwartz[28]). If L is a second-order symmetric differential operator defined on an open interval I and if c is an end point of I at which L has no boundary values, then for every z the dimension of the space of solutions to $Lu = zu$ which are square integrable near c is one. This remark applies to the end points 0 and ∞ of R^+ for all H_j except at the end point 0 and $j = 0$, $n = 1,2,3$, as noted in Theorem 2.1. According to Theorem 2.1 boundary values have to be added at zero to the definition of H_0 in these exceptional cases. These boundary values are satisfied by the solution $\eta^0(r,z)$ of Lemma 3.3, and indeed there is no other linearly independent solution of $H_0u = zu$ which satisfies these boundary conditions and is square integrable near zero.

Using these remarks and Corollary XIII. 3.15 from Dunford and Schwartz,[29] we see that the following explicit formula for $K_j(r,s;z)$ holds for $z \in \Omega^+$:

$$K_j(r,s;z) = \begin{cases} \dfrac{\varphi_+(r,z)\eta^0(s,z)}{W_z(\varphi_+,\eta^0)}, & s \leqslant r, \\[2ex] \dfrac{\eta^0(r,z)\varphi_+(s,z)}{W_z(\varphi_+,\eta^0)}, & s \geqslant r, \end{cases} \tag{4.3}$$

where φ_+ and η^0 are the solutions obtained in Lemmas 3.1 and 3.3, respectively, and $W_z(\varphi_+,\eta^0)$ is the Wronskian of $\varphi_+(r,z)$ and $\eta^0(r,z)$, which by the standard theory of differential equations is independent of r.

By Lemmas 3.1 and 3.3 we know that $\varphi^+(r,z),\eta^0(r,z)$ and their first derivatives can be extended continuously. Hence by (4.3) we will be done if we can show that

$$W_z(\varphi_+,\eta^0) = W(\varphi_+(r,z),\eta^0(r,z)) \neq 0, \quad z \in \Omega_R. \tag{4.4}$$

Suppose, to the contrary, that there exists a $\lambda_0 \in \Omega_R$ so that $W_{\lambda_0}(\varphi_+,\eta^0) = 0$. This implies that $\varphi_+(r,\lambda_0)$ and $\eta^0(r,\lambda_0)$ are linearly dependent. It follows that there exists a nonzero ξ so that

$$\eta^0(r,\lambda_0) = \xi\varphi_+(r,\lambda_0) = g(r).$$

We now compute $W(g,\bar{g})$ which we have already noted is independent of r. Since $\eta^0(r,\lambda)$ is a real function for real λ, it follows that $W(g,\bar{g}) = W_{\lambda_0}(\eta^0,\eta^0) = 0$. On the other hand, using Lemma 3.1 and the fact that $W(g,\bar{g})$ is independent of r, we get

$$\begin{aligned} W(g,\bar{g}) &= W(\xi\varphi_+(r,\lambda_0),\overline{\xi\varphi_+(r,\lambda_0)}) \\ &= |\xi|^2 \lim_{r\to\infty} W(\varphi_+(r,\lambda_0),\bar{\varphi}_+(r,\lambda_0)) \\ &= -2|\xi|^2 i\sqrt{\lambda_0} \neq 0. \end{aligned}$$

This is a contradiction, which establishes (4.4) and completes the proof of the theorem.

Corollary 4.3: Under the hypotheses of Theorem 4.2, $\sigma_{ac}(H_j) = [0,\infty)$, $j = 0,1,\cdots$.

Proof: An immediate consequence of Lemma 4.1 and the proof of Theorem 4.2.

We now turn to the study of the spectral multiplicity of the operator H_j. We recall that if A is a self-adjoint operator in a separable Hilbert space $\mathscr{H}$, then there is a countable set of measures $\{\mu_k\}$ defined on the Borel field of R, and a unitary operator U from $\mathscr{H}$ into the direct sum $\Sigma_1^m \oplus L^2(\mu_k)$ so that $(UAf)_k(\lambda) = \lambda(Uf)_k(\lambda)$, $1 \leqslant k \leqslant m$. The measure μ_{k+1} is absolutely continuous with respect to the measure μ_k and m may take on any positive integer value or infinity. The number m is called the spectral multiplicity of the operator A. In case A is a second order ordinary differential operator, its spectral multiplicity is at most 2. We refer the reader to Dunford and Schwartz,[30] Chap. X.5 and Chap. XII.3, as a general reference for spectral multiplicity theory and to the same source, Chap. XIII.5, for the special spectral theory associated with second-order ordinary differential operators.

When A is the ordinary differential operator H_j, there exist kernels $W_k(r,\lambda)$, $k = 1,2$, measurable with respect to the Borel field of $R^+\times R$, for each fixed λ are twice differentiable, a.e. Lebesgue, and so that considering H_j as a formal differential operator

$$H_jW_k(r,\lambda) = \lambda W_k(r,\lambda). \tag{4.5}$$

Moreover, $W_1(r,\lambda)$ and $W_2(r,\lambda)$ are linearly independent almost everywhere in λ with respect to the measure μ_2, and for every compactly supported $f \in L^2(R^+)$

$$(Uf)_k(\lambda) = \int_0^\infty f(r)\overline{W}_k(r,\lambda)dr, \quad \text{a.e.} - \mu_k. \tag{4.6}$$

All of these facts may be found essentially in the reference quoted at the end of the last paragraph.

Let now Ω be a bounded domain in the complex plane. By the discussion in the proof of Lemma 3.3 we can choose $r_0 > 0$ small enough so that the mapping $\Gamma(z)$ defined there will satisfy $\|\Gamma(z)\| \leqslant \frac{1}{2}$ for every $z \in \Omega$. Define $H_j^{(1)}$ to be the restriction of the formal differential operator H_j to $(0,r_0]$. If j and n do not take the exceptional values delineated in Theorem 2.1, and if we impose the boundary condition $u(r_0) = 0$, then $H_j^{(1)}$ becomes self-adjoint. If for the exceptional values of j and n we also impose the boundary conditions given in Theorem 2.1, then $H_j^{(1)}$ becomes self-adjoint for $j = 0,1,2,\cdots$.

Let $\eta^0(r,z)$ be the solution obtained in Lemma 3.3 and let $u(r,z)$ be a nontrivial solution to $H_j u = zu$ with $u(r_0,z) = 0$. Suppose j and n do not take on the exceptional values of Theorem 2.1. If $u \in L^2(0,r_0)$, then since the space of solutions to $H_j u = zu$ which are square integrable near zero is one-dimensional, we must have that u is a scalar multiple of η^0. We can make this scalar multiple equal to 1 by choosing $du(r_0,z)/dr = d\eta^0(r_0,z)/dr$. Thus we see that there can be only a finite number of z in Ω where $u(r,z) \in L^2(0,r_0)$. Otherwise, since $u(r,z)$ and $\eta^0(r,z)$ are analytic in z, they would coincide in all of Ω. This is impossible for nonreal z since $H_j^{(1)}$ is self-adjoint. In this argument we have overlooked the fact that there may be a sequence of z converging to the boundary of Ω where $u(r,z)$ and $\eta^0(r,z)$ coincide. However, this is easily taken care of by initially choosing a slightly larger domain. We have thus proved the following for the nonexceptional values of j and n.

Lemma 4.4: Under the hypotheses of Lemma 4.1, $H_j^{(1)}$ has at most a finite number of eigenvalues in any bounded set in the complex plane.

For the exceptional values of j and n, as we pointed out above, we also impose an additional boundary condition at zero. As is well known, the self-adjoint operator $H_j^{(1)}$ defined in this way has a compact resolvent. Thus we have the lemma in this case also.

Lemma 4.5: Under the hypotheses of Lemma 4.1 but with the more stringent requirement that $V(r) = O(r^{-3/2+\epsilon})$ as $r \to 0$, $\epsilon > 0$, the spectrum of $H_j^{(1)}$ contains only eigenvalues.

Proof: Let $z \in \Omega$ which is not in the point spectrum of $H_j^{(1)}$. We shall show that z is in the resolvent set of $H_j^{(1)}$. Toward this end let $f \in L^2(0,r_0)$. We shall find a function u in the domain of $H_j^{(1)}$ so that

$$(H_j^{(1)} - z)u = f.$$

Referring to the proof of Lemma 3.3, we see that the last equation is equivalent to

$$\frac{d}{dr}\left(r^{2(\gamma+1)}\frac{dw}{dr}\right) = r^{2\gamma+1}qw + r^{\gamma+1}f,$$

where $q(r,z) = r(V(r) - z)$, $w = r^{-(\gamma+1)}u$, and $\gamma(\gamma+1) = \nu(j,n)$, $\gamma \geqslant 0$. In integral equation form, this last equation may be written as

$$w(r,z) = \int_{r_0}^{r} \frac{1}{s^{2(\gamma+1)}} \int_0^s t^{2\gamma+1} q(t,z) w(t,z)\,dt\,ds + \int_{r_0}^{r} \frac{1}{s^{2(\gamma+1)}} \int_0^s t^{\gamma+1} f(t)\,dt\,ds. \tag{4.7}$$

Let L be the Hilbert space of functions on $(0,r_0)$ normed by

$$\|g\|_L^2 = \int_0^{r_0} |g(s)|^2 s^{2(\gamma+1)} ds.$$

Let

$$F(r) = \int_{r_0}^{r} \frac{1}{s^{2(\gamma+1)}} \int_0^s t^{\gamma+1} f(t)\,dt\,ds.$$

This is a continuous map from $L^2(0,r_0) \to L$; i.e.,

$$\|F\|_L \leqslant C \|f\|_{L^2(0,r_0)}. \tag{4.8}$$

Indeed, we have

$$|r^{(\gamma+1)}F(r)| \leqslant \left| \int_{r_0}^{r} \frac{1}{s^{(\gamma+1)}} \int_0^s t^{\gamma+1} |f(t)|\,dt\,ds \right| \leqslant \left| \int_{r_0}^{r} \frac{1}{s^{(\gamma+1)}} \frac{1}{\sqrt{2\gamma+3}} s^{(2\gamma+3)/2} ds \right| \times \|f\|_{L^2(0,r_0)} \leqslant \frac{1}{\sqrt{2\gamma+3}} \int_0^{r_0} s^{1/2} ds \|f\|_{L^2(0,r_0)}.$$

From this, of course, (4.8) is immediate.

We now proceed as in Lemma 3.3. Let $\Gamma(z)$ be the mapping from L to L given by

$$\Gamma(z)g(r) = \int_{r_0}^{r} \frac{1}{s^{2(\gamma+1)}} \int_0^s t^{2\gamma+1} q(t,z) q(t)\,dt\,ds. \tag{4.9}$$

Recall that $q(t,z) = t(V(t) - z)$. By our assumption on $V(t)$ we see that $q(t,z) = O(t^{-1/2+\epsilon})$ as $t \to 0$, independent of z in Ω. If we use this in (4.9), an elementary calculation shows that we get the pointwise estimate

$$|r^{(\gamma+1)}\Gamma(z)g(r)| = O(r_0^{\epsilon})\|g\|_L.$$

Thus it follows that

$$\|\Gamma(z)g\|_L \leqslant C r_0^{1/2+\epsilon} \|g\|_L,$$

so that if r_0 is sufficiently small, $\|\Gamma(z)\| \leqslant \frac{1}{2}$, independent of z in Ω.

Equation (4.7), which can be written as $(I - \Gamma)w = F$, can therefore be solved uniquely for $w \in L$, $w = (I - \Gamma)^{-1}F$. By (4.8) we have

$$\|w\|_L \leqslant C\|F\|_L \leqslant C\|f\|_{L^2(0,r_0)}.$$

Since $u = r^{(\gamma+1)}w$, we have

$$\|u\|_{L^2(0,r_0)} = \|w\|_L,$$

so that the last inequality becomes

$$\|u\|_{L^2(0,r_0)} \leqslant C\|f\|_{L^2(0,r_0)}.$$

We have thus found a solution u of $(H_j^{(1)} - z)u = f$, which is square integrable on $(0,r_0]$. But, in order that u belong to the domain of $H_j^{(1)}$, it must also satisfy $u(r_0) = 0$. This can be accomplished by adding to u the function $c\eta^0(r,z)$, which was obtained in Lemma 3.3 as a solution of $(H_j - z)u = 0$. Notice that $\eta^0(0,z)\neq 0$ since z is not an eigenvalue of $H_j^{(1)}$. By adjusting the constant c we can ensure that $u(r_0) = 0$.

In case $j = 0$ and $n = 1,2,3$, the above argument is not valid. However, as we pointed out before, if we add an additional boundary condition at zero, then $H_0^{(1)}$ has a compact resolvent so that the spectrum consists only of eigenvalues.

Lemma 4.6: Under the hypotheses of Lemma 4.5 the spectral multiplicity of H_j, $j = 0,1,2,\cdots$, is equal to 1.

Proof: Let $W_k(r,\lambda)$ be the kernels of (4.5) and U the representation of $L^2(R^+)$ onto $L^2(\mu_1)\oplus L^2(\mu_2)$ which is given by (4.6). As before, let $H_j^{(1)}$ be the restriction of the formal differential operator H_j to an interval $(0,r_0]$ and let Λ be an interval on the real axis so that $\sigma_e(H_j^{(1)})\cap\Lambda = \emptyset$. According to Theorem XIII.6.13 of Dunford and Schwartz,[31] for μ_k-almost all $\lambda\in\Lambda$, $W_k(r,\lambda)\in L^2(0,r_0)$, $k = 1,2$. Lemma 4.5 implies that the last statement is true for every bounded interval Λ in R. The above-mentioned theorem also says that $W_k(r,\lambda)$ must satisfy, for μ_k-almost all $\lambda\in\Lambda$ (if $\sigma_e(H_j^{(1)})\cap\Lambda = \phi$) the boundary condition at zero defining the self-adjoint operator H_j, if such a boundary condition is needed.

If we combine the above remarks with Lemma 3.3, we conclude that, for μ_k-almost all $\lambda\in\Lambda$, $W_k(r,\lambda)$ is a scalar multiple of $\eta^0(r,\lambda)$. This means precisely that there are Borel measurable sets $\Lambda_1, \Lambda_2\subseteq\Lambda$, so that $\mu_k(\Lambda_k) = \mu_k(\Lambda)$ and $W_k(r,\lambda)$ is a scalar multiple of $\eta^0(r,\lambda)$ for $\lambda\in\Lambda_k$. Since μ_2 is absolutely continuous with respect to μ_1, it follows that $\mu_2(\Lambda_1) = \mu_2(\Lambda)$, so that $\mu_2(\Lambda_1\cap\Lambda_2) = \mu_2(\Lambda)$. Further, there is a set $\Lambda_3\subseteq\Lambda$ so that $\mu_2(\Lambda_3) = \mu_2(\Lambda)$ and $W_1(r,\lambda)$ and $W_2(r,\lambda)$ are linearly independent for $\lambda\in\Lambda_3$. Thus on $\Lambda_1\cap\Lambda_2\cap\Lambda_3$, $W_1(r,\lambda)$ and $W_2(r,\lambda)$ are linearly independent and are scalar multiples of each other. This is impossible unless $\mu_2 = 0$. That $\mu_1\neq 0$ follows from the fact that H_j has a nonvoid spectrum.

Corollary 4.7: Under the hypotheses of Lemma 4.5 the spectral multiplicity of $(H_j)_{ac}$, $j = 0,1,2,\cdots$, is equal to 1.

Proof: Suppose that A is a self-adjoint operator which is reduced by a subspace to a self-adjoint operator A_0. The spectral multiplicity of A_0 is $\leqslant$ the spectral multiplicity of A. The corollary is now an immediate consequence of this fact and Lemma 4.6.

Suppose now that A and B are self-adjoint operators with the same spectrum and both having multiplicity 1. Then A and B are unitarily equivalent to the operator of multiplication by λ in $L^2(\sigma,\mu_A)$ and $L^2(\sigma,\mu_B)$, respectively, where $\sigma = \sigma(A) = \sigma(B)$. Moreover, if μ_A and μ_B are absolutely continuous with respect to each other, it is clear that A and B are unitarily equivalent. In our case this leads to the following conclusion:

Theorem 4.8: Under the hypotheses of Theorem 4.2 and Lemma 4.5 the operator H_{ac} is unitarily equivalent to $-\Delta$.

Proof: Let $H_{j,0} = H_j - V(r)$. By Corollary 4.3, $\sigma_{ac}(H_{j,0}) = \sigma_{ac}(H_j)$ and by Corollary 4.7, both operators have spectral multiplicity 1. Let μ_0 and μ be the measure μ_1 of the proof of Lemma 4.6, corresponding to $H_{j,0}$ and H_j, respectively. We know from Theorem 4.2 that both of these measures are absolutely continuous with respect to Lebesgue measure $d\lambda$. To show that $d\lambda$ is absolutely continuous with respect to μ, we observe that if this is not the case, there exists a set $\Lambda\subseteq R^+$ with positive Lebesgue measure such that the Radon–Nikodym derivative $d(E(\lambda)u,u)/d\lambda = 0$ for every $u\in L^2(R^+,d\lambda)$ and a.e. $d\lambda$ in Λ. This contradicts the formula preceding (5.6), or (5.6). Thus μ_0 and μ are absolutely continuous with respect to each other. If we apply Theorem 2.1 the proof is complete.

Remark: Our analysis, specifically Theorem 4.2, shows that when $\alpha = 1$, and under suitable restrictions on β, $\sigma_e(H)\cap R^+$ can have at most one eigenvalue at $b^2/4$. According to Theorem 2.2 of Harris and Lutz[32] [with condition (iii) appropriately corrected] the equation $H_j u = (b^2/4)u$ has two linearly independent solutions u_1 and u_2 so that as $r\to\infty$

$$u_1(r)\sim\sin\tfrac{1}{2}br,\quad u_2(r)\sim\cos\tfrac{1}{2}br,\quad \text{for } \beta>1,$$

$$u_1(r)\sim r^{-a/2b}\sin\tfrac{1}{2}br,\quad u_2(r)\sim r^{a/2b}\cos\tfrac{1}{2}br,\quad \text{for } \beta=1,$$

$$u_1(r)\sim e^{-(a/b)r^{1-\beta}}\sin\tfrac{1}{2}br,$$

$$u_2(r)\sim e^{(a/b)r^{1-\beta}}\cos\tfrac{1}{2}br,\quad \text{for } \beta<1.$$

Thus when $\beta>1$ we see that $\sigma_e(H)\cap R^+$ can have no eigenvalues, which is expected since this is the case of a short range potential. When $\beta = 1$, there is no eigenvalue when $|a/b|\leqslant 1$. When $|a/b|>1$, there is the possibility of an eigenvalue. However, if H_j has no boundary values at zero, the solution which is square integrable at infinity must also be square integrable zero. If, e.g., $j = 0$ and $n = 3$, then H_0 has two boundary values at zero, so that all solutions are square integrable at zero. In this case any solution which is square integrable at infinity is also square integrable at zero. But it must also satisfy the boundary condition $u(0) = 0$. Thus we have not been able to draw any conclusions as to the actual existence of eigenvalues (but see von Neumann and Wigner[33]).

5. EXISTENCE AND COMPLETENESS OF THE WAVE OPERATORS

The Møller wave opoerators for $H = -\Delta + V(r)$ are defined by

$$W^{\pm} = \text{s-}\lim_{t\to\pm\infty} e^{itH}e^{it\Delta}. \tag{5.1}$$

We recall that if these operators exist, then the scattering operator

$$S = (W^+)^* W^- \tag{5.2}$$

is unitary if and only if the range of W_+ is the same as the range of W_-. Moreover, the scattering operators are called *complete* if they have a common range equal to $L^2(R^n)_{ac}$, the space of absolute continuity for H.

Theorem 5.1: Assume the hypotheses of Lemma 4.5 but with the following more stringent requirements:

(a) $V_S(r) = O(r^{-3/2+\epsilon})$ as $r\to\infty$, $\epsilon>0$, and any one of the following conditions:

$$\text{(b)}\quad \begin{array}{ll} \text{(i)} & \beta+\alpha\geqslant 2 \text{ and } \beta>\tfrac{1}{2}, \\ \text{(ii)} & \beta+\alpha<2 \text{ and } 2\beta+\alpha>\tfrac{5}{2}, \\ \text{(iii)} & \beta-\alpha\geqslant 0 \text{ and } \beta>\tfrac{1}{2}, \\ \text{(iv)} & \beta-\alpha<0 \text{ and } 2\beta-\alpha>-\tfrac{1}{2}. \end{array}$$

Then the wave operators exist and are complete.

Proof: In order to prove the existence and completeness of the wave operators, it is enough to prove that the wave operators

$$W_j^{\pm}=\text{s}-\lim_{t\to\pm\infty} e^{itH_j}e^{-itH_{j,0}},\quad j=0,1,2,\cdots, \tag{5.3}$$

exist and are complete, where $H_{j,0}=H_j-V(r)$. Let $E_j(d\lambda)$ and $E_j^0(d\lambda)$ be the canonical spectral measures of H_j and $H_{j,0}$, respectively. In order to prove that (5.3) exists, it is sufficient to prove that, for every compact interval $\Lambda\subseteq(0,\infty)$ which does not contain the possible eigenvalue $b^2/4$, the wave operators

$$W_j^{\pm}(\Lambda)=\text{s}-\lim_{t\to\pm\infty} e^{itH_j}e^{-itH_{j,0}}E_j^0(\Lambda),\quad j=0,1,2,\cdots, \tag{5.4}$$

exist. Let us make the convention that in what follows when we speak of an interval Λ it is always of the type considered above.

By Theorem 4.2, the subspace of absolute continuity of H_j is the orthogonal complement of the space generated by its eigenvectors. Hence, in order to show the completeness of the wave operators, it is enough to show that for every Λ

$$\text{Range}(W_j(\Lambda))=E_j(\Lambda)L^2(R^+), \tag{5.5}$$

where the latter is the subspace of absolute continuity of H_j.

In Lemma 4.6 we showed that the spectral multiplicity of H_j is 1. This means that there is only one component in the direct sum spectral expansion for H_j and consequently one function $W(r,\lambda)$ which gives the unitary map (4.6) from $L^2(R^n)$ onto $L^2(\mu)$. For any compactly supported $f,g\in L^2(R^+)$ we have

$$(E_j(\Lambda)f,g)=(UE_j(\Lambda)f,Ug)$$

$$=\int_\Lambda\int_{R^+\times R^+} f(s)\overline{W}(s,\lambda)\bar{g}(t)W(t,\lambda)\,ds\,dt\,d\mu.$$

If we use Theorem 4.2 again, we see that $d\mu$, when restricted to the Borel field of Λ, is absolutely continuous with respect to Lebesgue measure; i.e., $d\mu=\omega(\lambda)d\lambda$. Consequently, we see that the measure generated by $(E_j(\Lambda)f,f)$ is absolutely continuous with respect to Lebesgue measure. If we denote its Radon–Nikodym derivative by $d(E(\lambda)f,f)/d\lambda$, we have

$$\frac{d(E(\lambda)f,f)}{d\lambda}=\omega(\lambda)\int_{R^+\times R^+} f(s)\bar{f}(t)W(t,\lambda)\overline{W}(s,\lambda)\,ds\,dt.$$

On the other hand, using formula (4.2), we have

$$\frac{d(E(\lambda)f,f)}{d\lambda}=\frac{1}{\pi}\text{Im}\int_{R^+\times R^+} K_j(r,s;\lambda)f(s)\bar{f}(r)\,ds\,dr.$$

Hence, a.e., $\lambda\in\Lambda$, we have by the continuity of $K_j(r,s;\lambda)$ and $W(r,\lambda)$ in s and r,

$$\omega(\lambda)\,|W(r,\lambda)|^2=\text{Im}K_j(r,r;\lambda). \tag{5.6}$$

For almost all λ, we have $H_j\,W(r,\lambda)=\lambda W(r,\lambda)$. If we set $\widetilde{W}(r,\lambda)=\omega(\lambda)^{1/2}W(r,\lambda)$, the same is true for $\widetilde{W}(r,\lambda)$. If H_j has no boundary values at zero, as we pointed out in the proof of Lemma 4.6, for almost all $\lambda\in\Lambda$, $\widetilde{W}(r,\lambda)$ is square integrable at zero. If H_0 has two boundary values at zero, for almost all $\lambda\in\Lambda$, there exists a constant $c(\lambda)$ so that

$$\widetilde{W}(r,\lambda)=c(\lambda)\eta^0(r,\lambda), \tag{5.7}$$

where $\eta^0(r,\lambda)$ is the solution constructed in Lemma 3.3. For a fixed λ_0, say for which (5.6) and (5.7) hold, there is an r_0 so that $\text{Im}K_j(r_0,r_0;\lambda_0)\neq 0$. By continuity there is a neighborhood of λ_0 for which this is true. If we let $c(\lambda)$ take on any value where it is not defined, it certainly is a measurable function. In particular, if we take $c(\lambda)=1$ where it is not defined, then by our above remarks and (5.6) there exists an $m>0$ so that for every $\lambda\in\Lambda$, $|c(\lambda)|\geqslant m>0$.

The map Uf of (4.6) restricted to Λ is a unitary map from $E_j(\Lambda)L^2(R^+)$ to $L^2(\mu_\Lambda)$, where μ_Λ is the restriction of μ to the Borel field of Λ. The inverse map to U, for $f\in L^2(\mu_\Lambda)$, is given by

$$(U^*f)(r)=\int_\Lambda f(\lambda)W(r,\lambda)\,d\mu$$

$$=\int_\Lambda \omega(\lambda)^{1/2}f(\lambda)\widetilde{W}(r,\lambda)\,d\lambda. \tag{5.8}$$

Since $f\in L^2(\mu_\Lambda)$, $\omega^{1/2}f\in L^2(d\lambda)$. Thus, if we use the kernel $\widetilde{W}(r,\lambda)$ instead of $W(r,\lambda)$, (5.8) provides a unitary map from $L^2(d\lambda)$ to $E_j(\Lambda)L^2(R^+)$.

The function $\eta^0(r,\lambda)$ being a solution of $H_ju=\lambda u$, it may be written as a linear combination of the solutions $\eta_1^+(r,\lambda)$ and $\eta_2^+(r,\lambda)$ of (3.23). All of these functions and their derivatives with respect to r are continuous in (r,λ). We may write

$$\eta^0(r,\lambda)=c_1(\lambda)\eta_1^+(r,\lambda)+c_2(\lambda)\eta_2^+(r,\lambda),$$

$$\frac{d\eta^0(r,\lambda)}{dr}=c_1(\lambda)\frac{d\eta_1^+(r,\lambda)}{dr}+c_2(\lambda)\frac{\eta_2^+(r,\lambda)}{dr}.$$

Since the Wronskian of η_1^+ and η_2^+ is different from zero, we may solve these equations for $c_1(\lambda)$ and $c_2(\lambda)$ and thus see that they are continuous functions of λ.

Since $\eta^0(r,\lambda)$ is real, if we use the asymptotic estimates (3.23), we may write

$$\eta^0(r,\lambda)=d_1(\lambda)\sin\sqrt{\lambda}\,r+d_2(\lambda)\cos\sqrt{\lambda}\,r+\xi(\lambda,r)$$

$$=[d_1^2(\lambda)+d_2^2(\lambda)]^{1/2}\sin[\sqrt{\lambda}\,r+\delta_j(\lambda)]$$

$$+\xi(\lambda,r),$$

where $d_1(\lambda)$ and $d_2(\lambda)$ are real continuous functions given by $d_1=i(c_1-c_2)$, $d_2=c_1+c_2$. From this we see that if $d_1^2+d_2^2=0$, then $c_1=c_2=0$, which would contradict the fact that η^0 is not the zero function. Thus, on any compact set, $d_1^2(\lambda) + d_2^2(\lambda) \geqslant m > 0$. Since $c_1(\lambda)$ and $c_2(\lambda)$ are bounded on a compact set, it follows that $\xi(\lambda,r) \to 0$ as $r \to \infty$ independent of λ in a compact set.

We may thus write

$$\widetilde{W}(r,\lambda) = c(\lambda)\sin[\sqrt{\lambda}\, r + \delta_j(\lambda)] + \xi(\lambda,r), \quad (5.9)$$

where $c(\lambda)$ is some measurable function which is bounded away from zero on compact intervals. We may write the phase shift $\delta_j = \delta_{j,0}(\lambda) + \delta_{j,1}(\lambda)$, where $\delta_{j,0}(\lambda)$ is the phase shift corresponding to the operator $H_{j,0} = H_j - V(r)$. As is well known, when $n = 3$, $\delta_{j,0}(\lambda) = -j\pi/2$.

Using the above remarks, we can apply the Dollard–Friedman[34] proof (see also Green and Lanford[35]) to show that the wave operators $W_j^{\pm}(\Lambda)$ exist, provided we establish a uniform estimate of the form

$$\xi(\lambda,r) = O(r^{-1/2+\epsilon}), \quad \lambda \in \Lambda. \quad (5.10)$$

In order to get an estimate on $\xi(\lambda,r)$, we must get estimates on the rates of approach of the functions $\theta_j(r,\lambda)$ given by (3.25), to their asymptotic values. This rate of approach is given by $(\Gamma_1\theta_j)(r,\lambda)$, where Γ_1 is given (3.24). Since λ is real and θ_j is bounded, we see from (3.24) that this rate of approach is estimated by

$$\int_r^\infty |X(s,\lambda)|\,ds.$$

To get an estimate on $|X(r,\lambda)|$, we need an estimate on $V_S(r)$ and the estimates (3.10) and (3.18). The hypothesis we have assumed on $V_S(r)$ gives the estimate (5.10) for this part of $X(r,\lambda)$. Referring now to (3.10), if $\alpha + \beta \geqslant 2$, then the estimate in (3.10) is $O(1/r)$. Thus this part of $X(r,\lambda)$ has an estimate $O(1/r^{1+\beta})$, which gives (5.10) if $\beta > \frac{1}{2}$. If $\alpha + \beta < 2$, then the estimate in (3.10) is $O(1/r^{\beta+\alpha-1})$, which gives (5.10) if $2\beta + \alpha > \frac{5}{2}$. Making a similar analysis of the estimate (3.18), we see that we get (5.10) if $\beta - \alpha \geqslant 0$ and $\beta > \frac{1}{2}$ and also if $\beta - \alpha < 0$ and $2\beta - \alpha > -\frac{1}{2}$.

The Dollard–Friedman proof shows that the range of $W_j^{\pm}(\Lambda)$ is the closure of functions of the form.

$$\int_\Lambda f(\lambda)\widetilde{W}(s,\lambda)\,d\lambda,$$

where $f \in C^\infty(\Lambda)$. But the inversion formula (5.8) shows that this closure is $E_j(\Lambda)L^2(R^+)$, proving the completeness of the wave operators.

APPENDIX

In this appendix we shall give a proof of Theorem 2.1. Let Δ_S be the Laplace–Beltrami operator on the $(n-1)$-dimensional unit sphere S^{n-1} and let $\{\mu_j\}_{j=0}^\infty$ be its increasing sequence of eigenvalues, $\mu_j = j(j+n-2)$. We denote by

$$N_j = \frac{2j+n-2}{j}\binom{j+n-3}{j-1}, \quad j \neq 0, \; N_0 = 1,$$

the multiplicity of μ_j. Let $\{Y_{lj}\}_{l=1}^{N_j}$ be an orthonormal set of eigenfunctions of Δ_S corresponding to μ_j. Thus $\{Y_{lj}: 1 \leqslant l \leqslant N_j,\ 0 \leqslant j < \infty\}$ is a complete orthonormal set for $L^2(S^{n-1})$. We may write

$$\begin{aligned} L^2(R^n) &= L^2(S^{n-1}) \otimes L^2(R^+, r^{n-1}dr) \\ &= \sum_{j=0}^\infty \oplus \left\{\sum_{l=1}^{N_j} \oplus Y_{lj}\right\} \otimes L^2(R^+, r^{n-1}dr) \\ &= \sum_{j=0}^\infty \oplus \left\{\sum_{l=1}^{N_j} \oplus \mathscr{H}_{lj}\right\} \\ &\cong \sum_{j=0}^\infty \oplus \left\{\sum_{l=1}^{N_j} \oplus L^2(R^+, r^{n-1}dr)\right\}. \end{aligned} \quad (A1)$$

As is well known, in the polar coordinates $r = |x|$, $\xi \in S^{n-1}$ we can write

$$\begin{aligned} H &= -\Delta + V(r) \\ &= r^{1-n}\left[-\frac{\partial}{\partial r}\left(r^{n-1}\frac{\partial}{\partial r}\right) + r^{n-3}\Delta_S + r^{n-1}V(r)\right]. \end{aligned} \quad (A2)$$

We are assuming, as we have done throughout the paper, that the operator of multiplication by V is relatively compact with respect to $-\Delta$. This implies that

$$D(H) = D(-\Delta) = H^2(R^n) = \{u \in L^2(R^n) | \Delta u \in L^2(R^n)$$

Here, Δu is taken in the distributional sense.

Let P_{lj} be the projection of $L^2(R^n)$ onto $\mathscr{H}_{lj}$ and

$$\widetilde{H}_j = r^{1-n}\left[-\frac{d}{dr}\left(r^{n-1}\frac{d}{dr}\right) + \mu_j r^{n-3} + r^{n-1}V(r)\right]. \quad (A3)$$

If $\Delta u \in L^2(R^n)$, it is easily established that $\Delta P_{lj}u \in L^2(R^n)$ and $HP_{lj}u = P_{lj}Hu$. Thus $\mathscr{H}_{lj}$ reduces H and $H|\mathscr{H}_{lj}$ is a self-adjoint operator in $\mathscr{H}_{lj}$ which is unitarily equivalent to a self-adjoint extension of $\widetilde{H}_j|C_0^\infty(R^+)$ in $L^2(R^+)$ in $L^2(R^+, r^{n-1}dr)$. Furthermore, by the unitary transformation $Uf(r) = r^{(n-1)/2}f(r)$ of $L^2(R^+, r^{n-1}dr)$ onto $L^2(R^+, dr)$ we see that $H|\mathscr{H}_{lj}$ is unitarily equivalent under U to a self-adjoint extension of $H_j|C_0^\infty(R^+)$ in $L^2(R^+, dr)$, where

$$H_j = -\frac{d^2}{dr^2} + \frac{1}{r^2}[\mu_j + \tfrac{1}{4}(n-1)(n-3)] + V(r). \quad (A4)$$

In what follows we shall study the exact boundary conditions at zero needed to define this selfadjoint extension. Since V is bounded at infinity, there are no boundary conditions to specify at infinity.

We believe that the following lemma is well known. However, since we have been unable to find a proof in the literature we shall present one.

*Lemma A.*1: Suppose that $V(r) = O(r^{-2+\epsilon})$, $\epsilon > 0$, as $r \to 0$. Let ν be real and $\gamma(\gamma+1) = \nu$, $\gamma < \frac{1}{2}$. Then the equation

$$\left[-\frac{d^2}{dr^2}+\frac{\nu}{r^2}+V(r)\right]u=zu+f, \qquad (A5)$$

where z is complex and $f(r)$ is square integrable in a neighborhood of $r=0$, has a solution u such that

$$u(r)\sim r^{\gamma+1},$$
$$u'(r)\sim(\gamma+1)r^{\gamma}, \quad \text{as } r\to 0.$$

If $\gamma=-\frac{1}{2}$ there is also a solution u which satisfies:

$$\begin{matrix} u(r)\sim r^{1/2}\log r, \\ u'(r)\sim\frac{1}{2}r^{-1/2}\log r, \end{matrix} \quad \text{as } r\to 0.$$

Proof: As in the proof of Lemma 4.5, Eq. (A5) can be written as

$$\frac{d}{dr}\left(r^{2(\gamma+1)}\frac{dw}{dr}\right)=r^{2\gamma+1}qw+r^{\gamma+1}f, \qquad (A6)$$

where $q(r,z)=r(V(r)-z)$, $w=r^{-(\gamma+1)}u$.

Suppose, at first, that $\gamma\geqslant-\frac{1}{2}$, and define

$$(\Gamma w)(r)=\int_{r_0}^{r}\frac{1}{s^{2(\gamma+1)}}\int_0^s t^{2\gamma+1}q(t)w(t)dtds,$$
$$F(r)=\int_{r_0}^{r}\frac{1}{s^{2(\gamma+1)}}\int_0^s t^{\gamma+1}f(t)dtds.$$

If w is a solution of

$$w=\Gamma w+F+c, \qquad (A7)$$

then clearly it is a solution of (A6).

Let $B=\mathscr{L}^{\infty}(0,r_0)$. By hypothesis $q(t,z)=O(t^{-1+\epsilon})$ as $t\to 0$. An easy computation shows that $\Gamma: B\to B$ is bounded, and, moreover, $\|\Gamma\|\leqslant Kr_0^{\epsilon}$. Also,

$$\|F\|_B\leqslant K\int_0^{r_0}\frac{ds}{s^{\gamma+1/2}}\|f\|_{L^2(0,r_0)}\leqslant Kr_0^{1/2-\gamma}\|f\|_{L^2(0,r_0)}.$$

Hence, if r_0 is sufficiently small, (A7) is solvable. Since Γw and F converge as $r\to 0$, so does w, and (for a suitably chosen c) by setting $u=r^{\gamma+1}w$ we get one assertion of the lemma. To get the assertion about the derivative, we differentiate (A7) and note that $rw'(r)\to 0$ as $r\to 0$.

If $\gamma<-\frac{1}{2}$, define B as before, but this time take

$$(\Gamma w)(r)=\int_0^{r}\frac{1}{s^{2(\gamma+1)}}\int_{r_0}^s t^{2\gamma+1}q(t,z)w(t)dtds,$$
$$F(r)=\int_0^{r}\frac{1}{s^{2(\gamma+1)}}\int_{r_0}^s t^{\gamma+1}f(t)dtds.$$

Then we get

$$\|\Gamma\|\leqslant K\int_0^{r_0}\frac{1}{s^{2(\gamma+1)}}|(s^{2\gamma+1+\epsilon}-r_0^{2\gamma+1+\epsilon})|ds\leqslant Kr_0^{\epsilon},$$
$$\|F\|_B\leqslant K\int_0^{r_0}\frac{1}{s^{2(\gamma+1)}}(s^{\gamma+3/2}+r_0^{\gamma+3/2})ds\|f\|_{L^2(0,r_0)}.$$

The same argument as before proves the lemma in this case.

Finally, if $\gamma=-\frac{1}{2}$, we define a space B normed by

$$\|g\|_B=\sup_{0<r\leqslant r_0}\left|\frac{g(r)}{\log r}\right|.$$

We now set

$$(\Gamma w)(r)=\int_0^r\frac{1}{s}\int_0^s q(t,z)w(t)dtds,$$
$$F(r)=\int_0^r\frac{1}{s}\int_0^s t^{1/2}f(t)dtds.$$

Instead of (A7) we now use the equation

$$w=\Gamma w+F+\log r. \qquad (A8)$$

Clearly, a solution of (A8) satisfies (A6) for $\gamma=-\frac{1}{2}$. Since $|(\log r)q(r,z)|=O(t^{-1+\epsilon/2})$ as $t\to 0$, it is again easy to show that $\Gamma: B\to B$ is bounded, with bound which becomes small as r_0 becomes small. Hence (A8) is solvable in B and the convergence of $(\Gamma w)(r)$ and $F(r)$ to zero as $r\to 0$ implies that $w\sim\log r$ as $r\to 0$. Thus $u=r^{1/2}w\sim r^{1/2}\log r$ as $r\to 0$. To obtain the last statement of the lemma, we note that differentiating (A8) leads to $dw/dr\sim 1/r$. The proof is complete.

If $\nu\geqslant\frac{3}{4}$, the equation $\gamma(\gamma+1)=\nu$ has a solution $\gamma\leqslant-\frac{3}{2}$. Therefore the solution of (A5), with $f=0$, $z=i$, given by Lemma (A1) for this γ is not square integrable at zero. Thus we have

Corollary A.2: If $\nu(j,n)=\mu_j+(n-1)(n-3)/4\geqslant 3/4$, $H_j|C_0^{\infty}(R^+)$ is essentially self-adjoint, i.e., no boundary condition is required at 0.

The cases not covered by this corollary are easily seen to be $j=0$ and $n=1,2,3$. Note that for $n=1$ we need discuss boundary values only if we discuss the operator H on R^+. For the cases mentioned above we have $\nu(0,1)=\nu(0,3)=0$ or $\nu(0,2)=-\frac{1}{4}$. If $\nu(j,n)=0$, the solutions of (A5) given by Lemma A.1 behave asymptotically at zero like r or like a constant. The case $\nu=-\frac{1}{4}$ was treated explicitly in Lemma A.1.

Lemma A.3: Let $\hat{H}_j$ be the maximal operator corresponding to H_j in $L^2(R^+,dr)$. Then:

(a) For $j=0$, $n=1,3$ the nonzero linear functional

$$G(u)=\lim_{r\to 0}u(r), \quad u\in D(\hat{H}_0),$$

is a boundary value for $\hat{H}_0$.

(b) For $j=0$, $n=2$, the nonzero linear functional

$$G(u)=\lim_{r\to 0}\frac{r^{1/2}}{\log r}u'(r), \quad u\in D(\hat{H}_0),$$

is a boundary value for $\hat{H}_0$.

Proof: We must verify that the limits defined in (a) and (b) exist, are finite, and not all zero. In case (a) let $u\in D(\hat{H}_0)$ and let v be the solution obtained in Lemma A.1, for $z=0$ and $f=\hat{H}_0u$, which behaves asymptotically like r. Let u_1 and u_2 be the solutions obtained in Lemma A.1 for $f=0$ and $z=0$ which behave asymptotically like r and like a constant, respectively. Then $u-v=c_1u_1+c_2u_2$ and we see immediately that $G(u)$ exists and is a nonzero linear functional on $D(\hat{H}_0)$.

For case (b) we may reason as before and write

$u = v + c_1u_1 + c_2u_2$, where v and u_1 behave asymptotically like $r^{1/2}$ and u_2 behaves asymptotically like $r^{1/2}\log r$. Differentiating, we get

$$u' \sim \tfrac{1}{2}[(c_1+1)r^{-1/2} + c_2 r^{-1/2}\log r].$$

From this we see that $G(u)$ in case (b) is well defined also, and is a nonzero linear functional on $D(\hat{H}_0)$.

Lemma A.4: In the case $j = 0$, $n = 1,2,3$, the operator $H|\mathscr{H}_{lj}$ is unitarily equivalent to the self-adjoint extension of $H_0|C_0^\infty(R^+)$ given by $G(u) = 0$, where G is defined in Lemma A.3.

Proof: Since we are concerned only with $j = 0$, it suffices to consider a spherically symmetric function $f(r)$ such that $f \in D(H) \subseteq L^2(R^n)$. The trace of f on the $(n-1)$-dimensional sphere of radius r is given by $r^{(n-1)/2}|f(r)|$, and it is well known (see, e.g., Lions and Magenes[36]) that this is a continuous function which goes to zero as $r \to 0$. Since $Uf(r) = r^{(n-1)/2}f(r)$, we see that every element in $D(\hat{H}_0)$ which comes from an element of H satisfies the boundary condition $G(u) = 0$ for $n = 1,3$. But since $H|\mathscr{H}_0$ is self-adjoint and is unitarily equivalent to a self-adjoint extension of the minimal operator associated with the formal operator H_0, the boundary condition of the self-adjoint extension must be $G(u) = 0$ in the cases $n = 1,3$.

Suppose now that $f \in H^2(R^2)$; then $f' \in H^1(R^2)$. Hence f' has a continuous trace on the circle of radius r which goes to zero as $r \to 0$. Hence $\lim_{r\to 0} r^{1/2}f'(r) = 0$. Now,

$$(UF)'(r) = \tfrac{1}{2}r^{-1/2}f + r^{1/2}f',$$

so that

$$\frac{r^{-1/2}(Uf)'(r)}{\log r} = \frac{1}{2}\frac{f}{\log r} + \frac{rf'}{\log r}.$$

By Sobolev's inequality f is bounded in a neighborhood of zero. Thus we see $G(Uf) = 0$, which concludes the proof.

Remark: In the definition of G in Lemma A.3(b) we needed the factor $\log r$ to guarantee that G is defined for every function $D(\hat{H}_0)$, the domain of the maximal operator. In the proof of the last lemma we saw that if $f \in D(H)$, then

$$\frac{r^{-1/2}(Uf)'(r)}{g(r)} \to 0 \quad \text{as } r \to 0 \tag{A9}$$

for any g such that $|g(r)| \to \infty$ as $r \to 0$. That this does not constitute any discrepancy easily follows from Lemma A.1. Indeed suppose $u \in D(\hat{H}_0)$ and $G(u) = 0$. Then by Lemma A.1,

$$u'(r) \sim c_1 r^{-1/2} + c_2 r^{-1/2}\log r \quad \text{as } r \to 0.$$

Hence, if $r^{1/2}(\log r)^{-1}u'(r) \to 0$ as $r \to 0$ we must have $c_2 = 0$. From this (A9) is an immediate consequence.

For $j = 0$, $n = 2$, the reason we could not take $G(u) = \lim_{r\to 0} u(r)$ is that $G(u) = 0$ for all $u \in D(\hat{H}_0)$ and hence is not a boundary value.

[1] J. von Neumann and E. Wigner, Z. Physik **30**, 465–67 (1929).
[2] S. Agmon, Ann Scuol. Norm. Sup. Pisa **2**, 151–218 (1975).
[3] J. Weidmann, Math. Zeit. **98**, 268–302 (1967).
[4] T. Ikebe and Y. Saito, J. Math. Kyoto Univ. **7**, 513–42 (1972).
[5] R. Lavine, J. Func. Anal. **12**, 3–54 (1973).
[6] A. Wintner, Am. J. Math. **69**, 251–72 (1947).
[7] F. Atkinson, Ann. Mat. Pura Appl. **37**, 347–78 (1954).
[8] B. Simon, Commun. Pure Appl. Math. **22**, 531–38 (1969).
[9] W. A. Harris, Jr. and D. A. Lutz, J. Math. Anal. Appl. **51**, 76–93 (1975).
[10] J. D. Dollard and C. N. Friedman, Ann. Phys. (N.Y.) **111**, 251–66 (1978).
[11] Ref. 10.
[12] W. A. Harris, Jr. and D. A. Lutz, J. Math. Anal. Appl. **57**, 577–86 (1977).
[13] Ref. 10.
[14] Ref. 3, p. 299.
[15] Ref. 10.
[16] For $n = 1$ we do not need to specify a boundary condition at 0, unless we discuss H only on R^+.
[17] Ref. 12.
[18] E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955), p. 92.
[19] K. Kodaira, Am. J. Math. **71**, 921–45 (1949).
[20] Ref. 10.
[21] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[22] M. Schechter, *Spectra of Partial Differential Operators* (North-Holland, Amsterdam, 1971).
[23] Ref. 22, p. 84.
[24] N. Dunford and J. Schwartz, *Linear Operators, Part II*, (Interscience, New York, 1963). We note that the definition of $\sigma_c(A)$ given in this reference, p. 1393, coincides with the usual definition for self-adjoint operators A only when A does not have an eigenvalue of infinite multiplicity. This is, of course, the case for second order ordinary differential operators.
[25] Ref. 24, p. 1445.
[26] M. Ben-Artzi, "Spectral Properties of Linear Ordinary Differential Operators with Slowly Decreasing Coefficients," J. Math. Anal. Appl. (to be published).
[27] Ref. 24, Sec. XIII.3, p. 1310.
[28] Ref. 24, Sec. XIII.2, p. 1285.
[29] Ref. 24, pp. 1327–28.
[30] Ref. 24.
[31] Ref. 24, p. 1402.
[32] Ref. 9.
[33] Ref. 1.
[34] Ref. 10.
[35] T. A. Green and O. E. Lanford, J. Math. Phys. **1**, 139–45 (1960).
[36] J. L. Lions and E. Magenes, *Problémes aux limites nonhomogénes et applications, Vol.* 1 (Dunod, Paris, 1968).

A note on the iteration of the Chandrasekhar nonlinear H-equation

R. L. Bowden
Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
(Received 30 August 1978)

An iteration scheme to solve the Chandrasekhar H equation in the form
$$H(\mu) = \{1-\mu \int_0^1 [\Psi(s)\ H(s)]/(s+\mu) ds\}^{-1}$$
is shown to converge monotonically and uniformly.

I. INTRODUCTION

It is well known that the nonlinear H equation of radiative and neutron transport theory,

$$H(\mu) = 1 + \mu H(\mu) \int_0^1 \frac{\Psi(s)H(s)}{s+\mu} ds, \qquad (1a)$$

does not have a unique solution. However, the "physical" solution of Eq. (1a) subject to the constraints

$$\nu_j \int_0^1 \frac{\Psi(s)H(s)}{\nu_j - s} ds = 1, \quad j = 0,...,\alpha, \qquad (2)$$

where $\nu_j, j = 0,..., \alpha$, are zeros of the dispersion function

$$\Lambda(z) = 1 + z \int_{-1}^{+1} \frac{\Psi(s)}{s-z} ds, \qquad (3)$$

is unique, and an explicit solution can be written down. Nevertheless, a traditional approach to obtaining values for the H function is to attempt to solve numerically Eq. (1a) by iteration. However, it has been only relatively recently that the iteration of Eq. (1a) has been shown to converge. For example, Bittoni, Casadei, and Lorenzutta[1] showed that when the right-hand side of Eq. (1a) is regarded as a bilinear operator from $L_1(0,1)\times L_1(0,1)$ to $L_1(0,1)$ that the norm of its Fréchet derivative is less than unity and that the bilinear operator is contractive in the ball

$$S = \{\Psi H \in L_1(0,1) | \|\Psi(1-H)\| < \tfrac{1}{2}\} \qquad (4)$$

if $\|\Psi\| < \frac{1}{2}$, where $\|\,\|$ denotes the usual L_1 norm. Thus the unique solution of Eq. (1) in S can be obtained by iteration using the scheme

$$H_{n+1}(\mu) = 1 + \mu H_n(\mu) \int_0^1 \frac{\Psi(s)H_n(s)}{s+\mu} d\mu, \quad H_0 \in S. \qquad (5)$$

Subsequently, Bowden and Zweifel[2] showed that the solution so obtained was indeed the "physical" solution. Furthermore, if $\Psi(s)$ is nonnegative and even for $s\in(-1,1)$, the transformation

$$H(z) = \frac{\nu_0(1+z)}{\nu_0 + z} L(z) \qquad (6)$$

leads to an equation for L which is identical in form to Eq. (1), but with the characteristic function Ψ replaced by the function

$$\Psi'(s) = \frac{\nu_0^2(1-s^2)}{\nu_0^2 - s^2} \Psi(s).$$

(Under the condition stated for Ψ, the dispersion function has only two zeros $\pm\nu_0$.) Bowden and Zweifel pointed out that Ψ' is also a nonnegative even function and that $\|\Psi'\| < \frac{1}{2}$. Thus in all instance in which $\Psi(s)$ is nonnegative and even, numerical values of the H function can be obtained from the iteration scheme given by Eq. (5).

On the other hand, Eq. (1a) can also be rewritten as

$$H(\mu) = \left(1 - \mu \int_0^1 \frac{\Psi(s)H(s)}{s+\mu} ds\right)^{-1}, \qquad (1b)$$

and numerical values for the H function can be obtained by iterating on

$$H_{n+1}(\mu) = \left(1 - \mu \int_0^1 \frac{\Psi(s)H_n(s)}{s+\mu} ds\right)^{-1}, \quad H_0(\mu) = 0, \qquad (7a)$$

or equivalently,

$$H_{n+1}(\mu) = 1 + \mu H_{n+1}(\mu) \int_0^1 \frac{\Psi(s)H_n(s)}{s+\mu} ds. \qquad (7b)$$

Unfortunately, rewriting Eq. (1a) removes the bilinear structure of the equation, and the proof of Bittoni, Casadei, and Lorenzutta need no longer apply. Thus, although this iteration scheme has often been used, dating perhaps from Chandrasekhar and Breen,[3] there does not appear to have been a proof that the sequence $\{H_n\}$ obtained from the iteration scheme given by Eq. (7) converges to the solution of Eq. (1), and that if a solution is obtained, that it represents the physical solution given by the constraining condition (2). The purpose of this note is to given a brief proof that the iteration scheme given by Eq. (7) does indeed converge to the "physical" solution of Eq. (1). The proof is based only on the positivity of the integral operator

$$(\mathscr{L}f)(\mu) = \mu \int_0^1 \frac{\Psi(s)f(s)}{s+\mu} ds, \quad \mu\in(0,1), \qquad (8)$$

where the assumptions

$$\Psi(s) = \Psi(-s), \quad \Psi(s) \geqslant 0, \quad s\in(-1,+1), \qquad (9a)$$

and

$$\int_0^1 \Psi(s) ds < \tfrac{1}{2} \qquad (9b)$$

are used throughout this note. However, as pointed out above, if condition (9b) is not satisfied, then transformation (6) can be used to obtain an equation in which the resulting characteristic function does satisfy the inequality (9b). The assumption of evenness of $\Psi(s)$ on $(-1, +1)$ is not used in

 0022-2488/79/040608-03$01.00

the proof of convergence of the iteration scheme *per se*. It is a condition that the solution obey the constraint given by Eq. (2).[2]

2. PROOF OF CONVERGENCE

In this section we present the proof of convergence of the iteration scheme given by Eq. (7). The first step is to show that

$$H_{n+1}(\mu) \geqslant H_n(\mu), \quad \mu \in (0,1), \quad n \geqslant 1, \tag{10}$$

and

$$\|\Psi H_n\| < 1 - \sqrt{1 - 2\|\Psi\|}. \tag{11}$$

The proof of inequalities (10) and (11) proceeds by induction. First note that

$$H_1(\mu) = 1, \quad \mu \in (0,1). \tag{12}$$

Further note that

$$\int_0^1 |\Psi(s) H_1(s)| ds = \|\Psi\| \leqslant 1 - \sqrt{1 - 2\|\Psi\|}. \tag{13}$$

If the inequalities (10) and (11) are assumed true for $n = N$, then subtraction of

$$H_N = 1 + H_N \mathscr{L}(H_{N-1}) \tag{14}$$

from

$$H_{N+1} = 1 + H_{N+1} \mathscr{L}(H_N) \tag{15}$$

yields, after a little rearrangement,

$$(H_{N+1} - H_N)[1 - \mathscr{L}(H_N)] = H_N \mathscr{L}(H_N - H_{N-1}). \tag{16}$$

By simple manipulation it is apparent that

$$1 - (\mathscr{L} H_N)(\mu) = 1 + \int_0^1 \frac{s\Psi(s) H_N(s)}{s + \mu} ds - \int_0^1 \Psi(s) H_N(s) ds > 0, \quad \mu \in (0,1), \tag{17}$$

since by assumption,

$$\int_0^1 \Psi(s) H_N(s) ds \leqslant 1 - \sqrt{1 - 2\|\Psi\|} < 1.$$

Therefore, it follows that

$$H_{N+1}(\mu) - H_N(\mu) \geqslant 0, \quad \mu \in (0,1). \tag{18}$$

This proves inequality (10). To prove inequality (11) it may be noted from Eq. (7b) and inequality (10) that

$$\begin{aligned} &\|\Psi H_{N+1}\| \\ &= \|\Psi\| + \int_0^1 \left(\mu \Psi(\mu) H_{N+1}(\mu) \int_0^1 \frac{\Psi(s) H_N(s)}{s + \mu} ds\right) d\mu \\ &\leqslant \|\Psi\| + \tfrac{1}{2}\left(\int_0^1 \Psi(\mu) H_{N+1}(\mu) d\mu\right)^2 \\ &\quad + \tfrac{1}{2} \int_0^1 \left(\Psi(\mu) H_{N+1}(\mu) \int_0^1 \frac{\Psi(s) H_{N+1}(s)(\mu - s)}{s + \mu} ds\right) d\mu. \end{aligned} \tag{19}$$

However, the last term on the right-hand side of the last inequality vanishes so that

$$2\|\Psi H_{N+1}\| - 2\|\Psi\| \leqslant \|\Psi H_{n+1}\|^2. \tag{20}$$

It is then obvious that either

$$\|\Psi H_{N+1}\| \geqslant 1 + \sqrt{1 - 2\|\Psi\|}, \tag{21a}$$

or

$$\|\Psi H_{N+1}\| \leqslant 1 - \sqrt{1 - 2\|\Psi\|}. \tag{21b}$$

On the other hand, Eq. (7b) and inequality (10) can be used to write

$$\begin{aligned} &\|\Psi H_{N+1}\| \\ &= \|\Psi\| + \int_0^1 \left[\Psi(\mu) H_{N+1}(\mu) \int_0^1 \Psi(s) H_N(s) ds\right] d\mu \\ &\quad - \int_0^1 \left(\Psi(\mu) H_{N+1}(\mu) \int_0^1 \frac{s\Psi(s) H_N(s)}{s + \mu} ds\right) d\mu \\ &\leqslant \|\Psi\| + \|\Psi H_{N+1}\| \cdot \|\Psi H_N\| \\ &\quad - \int_0^1 \left(\Psi(\mu) H_N(\mu) \int_0^1 \frac{s\Psi(s) H_N(s)}{s + \mu} ds\right) d\mu \\ &= \|\Psi\| + \|\Psi H_{N+1}\| \cdot \|\Psi H_N\| - \tfrac{1}{2}\|\Psi H_N\|^2. \end{aligned} \tag{22}$$

Thus it follows that

$$\|\Psi H_{N+1}\| \leqslant (1 - \|\Psi H_N\|)^{-1}(\|\Psi\| - \tfrac{1}{2}\|\Psi H_N\|^2). \tag{23}$$

This last inequality yields

$$\|\Psi H_{N+1}\| \leqslant 1, \tag{24}$$

since

$$(1 - \|\Psi H_N\|)^{-1}(\|\Psi\| - \tfrac{1}{2}\|\Psi H_N\|^2) > 1 \tag{25}$$

would imply that

$$\|\Psi\| > \tfrac{1}{2}[1 + (1 - \|\Psi H_N\|)^2], \tag{26}$$

which contradicts the assumption $\|\Psi\| < \frac{1}{2}$. Combining this result with the earlier estimate of $\|\Psi H_{N+1}\|$ gives inequality (11).

Therefore, the sequence $\{H_n(\mu)\}$ given by Eq. (7) is positive and monotone with $\|\Psi H_n\|$ bounded by the inequality (11). Thus from the monotone convergence theorem there exists an L_1 (0,1) function, say H^*, such that

$$\lim_{n \to \infty} \Psi H_n \to \Psi H^*, \tag{27}$$

and

$$\|\Psi H^*\| = \int_0^1 |\Psi(s) H^*(s)| ds = \int_0^1 \Psi(s) H^*(s) ds \leqslant 1 - \sqrt{1 - 2\|\Psi\|}. \tag{28}$$

A straightforward calculation gives

$$\|\Psi H^* - \Psi - \Psi H^* \mathscr{L}(H^*)\| = 0, \tag{29}$$

that is, H^* is an L_1 solution to

$$\Psi H^* = \Psi + \Psi H^* \mathscr{L}(H^*). \tag{30}$$

Now define the function H by

$$H(z) = [1 - (\mathscr{L} H^*)(z)]^{-1}. \tag{31}$$

It is obvious that

$$\Psi(\mu)H(\mu) = \Psi(\mu)[1 - (\mathscr{L}H^*)(\mu)]^{-1} = \Psi(\mu)H^*(\mu) \quad (32)$$

almost everywhere for $\mu \in (0,1)$. Thus $H(z)$ satisfies the nonlinear H equation (1) and is analytic in the complex plane of z cut along $(-1,0)$ except at that value of z such that $1 - (\mathscr{L}H)(z)$ vanishes. However, since $\|\Psi H\| < 1$ it follows from the results of Bowden and Zweifel[2] that $1 - (\mathscr{L}H) \times (-\nu_0) = 0$, i.e., $H(z)$ represents the "physical" solution according to Eq. (2). In particular $1 - (\mathscr{L}H)(z)$ is bounded in the right half complex z plane.

Finally an estimate of $|H_{n+1}(\mu) - H(\mu)|$ can be written from Eqs. (1) and (7) as

$$\begin{aligned} |H_{n+1}(\mu) - H(\mu)| &= |H_{n+1}(\mu)(\mathscr{L}H_n)(\mu) - H(\mu)(\mathscr{L}H)(\mu)| \\ &\leqslant |H_{n+1}(\mu)[\mathscr{L}(H_n - H)](\mu) \\ &\quad + [H_{n+1}(\mu) - H(\mu)](\mathscr{L}H)(\mu)|. \end{aligned} \quad (33)$$

This inequality in turn yields

$$|H_{n+1}(\mu) - H(\mu)| \leqslant [1 - (\mathscr{L}H)(\mu)]^{-1} \times |H_{n+1}(\mu)[\mathscr{L}(H_n - H)](\mu)|. \quad (34)$$

Since the right-hand side of this last approaches zero as $n \to \infty$, the sequence $\{H_n\}$ approaches H at least pointwise. However, since H is continuous on $(0,1)$ and the $[H_n(\mu)]$ is a positive monotone sequence, it follows from Dini's theorem that the convergence is also uniform.

In summary, it can then be stated that the sequence of functions (H_n) obtained by the iteration scheme given by Eq. (7) converges monotonically and uniformly to the physical solution of Eq. (1).

ACKNOWLEDGMENTS

The author is grateful to Istituto Matematica Applicata, Universita degli Studi di Firenze,and to Professor Aldo Bellini-Morante for their kind hospitality during the period when much of this work was reviewed.

This work was supported in part by the National Science Foundation Grant Number ENG75-15882 A01.

[1]E. Bittoni, G. Casadei, and S. Lorenzutta, Boll. U.M.I. **4**, 535 (1969).
[2]R.L. Bowden and P.F. Zweifel, Astrophys. J. **210**, 178 (1976).
[3]S. Chandrasekhar and Frances H. Breen, Astrophys. J. **106**, 145 (1947).

On the quantization of spin systems and Fermi systems

Ph. Combe and R. Rodriguez

Centre Universitaire de Luminy et Centre de Physique Théorique, CNRS

M. Sirugue-Collin

Universitaire de Provence et Centre de Physique Théorique, CNRS

M. Sirugue[a)]

Centre de Physique Théorique, CNRS
(Received 13 June 1978)

It is shown that spin operators and Fermi operators can be interpreted as the Weyl quantization of some functions on a "classical phase space" which is a compact group. Moreover the transition from quantum spin to Fermi operators is an isomorphism of the "classical phase space" preserving the Haar measure.

1. INTRODUCTION

Canonical anticommutation relations (CAR) and commutation relations of quantum spin system (QSS) can be seen as projective representations of the group $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$, where $\mathfrak{p}_\Lambda$ is the group of finite subsets of a set Λ equipped with the symmetric difference as group law. This allows us to consider the CAR and QSS in a similar way as usual canonical commutation relations (CCR) of systems with finite number n of degrees of freedom with the replacement of the group R^{2n} by $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$ and the multiplier $((x,p);(x',p')) \to \exp(i\hbar/2)(xp' - xp)$ by a suitable multiplier on $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$.

In a paper, hereafter referred as Ref. 1, we have shown that the multipliers on $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$ which satisfy natural physical requirements are only four. Moreover we have shown that the algebras of "canonical commutation relations" associated with these multipliers are all isomorphic.

In this paper, we develop the analogy between the CCR on one hand and the CAR and QSS on the other hand, mostly from the point of view of the Weyl quantization. In Sec. 2, we recall the essential definitions and results which are needed in what follows and we derive explicitly the isomorphism of $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda,\xi)}$ to the usual UHF algebra, ξ being the multiplier on $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$.

In the third section we study the states of $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda,\xi)}$ and in particular we define an important closed convex subset of states, the "quasiclassical states" [definition (3.2)] and we show that it contains a great number of usual states. These states have a nice characterization in terms of probability on the classical phase space [Corollary (3.4)]. In some sense they are as close as possible to classical states.

In the fourth section we quantify, according to the Weyl procedure, functions on the "classical phase space" [definition (4.2)]. A new situation arises, with respect to the usual CCR, for a finite number of degrees of freedom, since there are in our case lots of inequivalent irreducible representations of $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda,\xi)}$ and the set of functions one can quantize depends on the representation. Finally we derive an important result [Theorem (4.15)] which shows that the transition from quantum spin systems to Fermi systems is described by an automorphism of the "classical phase space" which leaves invariant the Haar measure.

2. THE WEYL ALGEBRA FOR SPIN AND FERMI SYSTEMS

We recall that $\mathfrak{p}_\Lambda$ is the group of finite subsets of an at most countable set Λ equipped with the symmetric difference as group law (denoted by $\triangle$). In Ref. 1 we have shown that there exist for all cardinality of Λ only two bicharacters b^S and b^F on $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$ which are nondegenerate, invariant under the finite permutations of points in Λ, invariant under the automorphism τ_{θ_0} of $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$ defined by $\tau_{\theta_0}(X,Y) = (Y,X)$ and which can be written as

$$b((X,Y);(X',Y')) = \xi((X,Y);(X',Y'))\overline{\xi((X',Y');(X,Y))}$$

for some multiplier ξ on $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$.

Correspondingly, up to trivial multipliers, there are two multipliers ξ^S and ξ^F which are

$$\xi^S((X,Y);(X',Y')) = i^{-|X\cap Y| - |X'\cap Y'| + |X\triangle X'\cap Y\triangle Y'|}(-1)^{|Y\cap X'|}, \tag{2.1}$$

$$\xi^F((X,Y);(X',Y')) = i^{|X\cap\theta(X)| + |Y\cap\theta(Y)| + |X||Y|}\, i^{|X'\cap\theta(X')| + |Y'\cap\theta(Y')| + |X'||Y'| - |X\triangle X'\cap\theta(X\triangle X')| - |Y\triangle Y'\cap\theta(Y\triangle Y')|}$$
$$\times i^{-|X\triangle X'||Y\triangle Y'|}(-1)^{|X'||Y| + |X'\cap\theta(X)| + |Y'\cap\theta(Y)|}, \tag{2.2}$$

[a)]Postal address: Centre de Physique Theorique, CNRS Luminy Case 907, Route Léon Lachamp, F13288 Marseille Cedex 2, France.

 0022-2488/79/040611-06$01.00

where $|X|$ is the cardinality of X and θ is the homomorphism of $\mathfrak{p}_\Lambda$ such that

$$\theta(\{x_i\}) = \{x_1 \cdots x_{i-1}\}, \quad \theta(\{x_1\}) = \phi$$

for a given but arbitrary ordering of points in Λ.

In Ref. 1 we have shown that there exist two C^* algebras $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda, \xi^S)}$ and $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda, \xi^F)}$ which are generated by the elements δ^S_{XY} and δ^F_{XY}, $(X,Y) \in \mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$, such that

$$\delta^S_{X,Y}\delta^S_{X',Y'} = \xi^S((X,Y);(X',Y'))\delta^S_{X\triangle X',Y\triangle Y'}, \tag{2.3}$$

$$\delta^F_{X,Y}\delta^F_{X',Y'} = \xi^F((X,Y);(X',Y'))\delta^F_{X\triangle X',Y\triangle Y'}, \tag{2.4}$$

$$\delta^{S\,*}_{X,Y} = \delta^S_{X,Y}, \quad \delta^{F\,*}_{X,Y} = \delta^F_{X,Y}.$$

An explicit computation shows that the $\delta^S_{X,Y}$ have the commutation relations of a quantum spin system on a lattice and the $\delta^F_{X,Y}$ the canonical anticommutation relations of products of Fermi field operators.

Moreover in Ref. 1, we have shown that there exists a *isomorphism α between $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda, \xi^S)}$ and $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda, \xi^F)}$

$$\alpha(\delta^F_{X,Y}) = F(X,Y)\delta^S_{\tau_\theta(X,Y)}, \tag{2.5}$$

where

$$F(X,Y) = (-1)^{|Y\cap\theta(Y)| + |X\cap Y\cap\theta(X\triangle Y)|}$$

$$\times(-1)^{|X||Y|[(|X|+|Y|)/2] + |X||Y| + |X\cap Y|},$$

$$\forall\, X,Y \in \mathfrak{p}_\Lambda \tag{2.6}$$

and τ_θ is the isomorphism of $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$,

$$\tau_\theta(X,Y) = (X\triangle\theta(X\triangle Y), Y\triangle\theta(X\triangle Y)). \tag{2.7}$$

These algebras for the cardinality of Λ infinite are * isomorphic to the UHF algebra and now we want to make this correspondence more explicit.

We shall specialize to $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda, \xi^S)}$ for the sake of simplicity, but our results are also valid for the other multipliers [see Ref. 1, Theorem (2.40)].

If $\Lambda_1 \subset \Lambda$, then $\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}$ is a subgroup of $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$ and we shall denote by the same symbol ξ^S the restriction of the multiplier ξ^S of $\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda$ to $\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}$. It is nondegenerate. [See Ref. 1 definition (1.11).]

Lemma (2.8): Let $\Lambda_1, \Lambda_2 \subset \Lambda$. $\overline{\Delta(\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}, \xi^S)}$ is * isomorphic to a subalgebra of $\overline{\Delta(\mathfrak{p}_{\Lambda_2} \times \mathfrak{p}_{\Lambda_2}, \xi^S)}$ if $|\Lambda_1| \leqslant |\Lambda_2|$; they are * isomorphic if and only if $|\Lambda_1| = |\Lambda_2|$.

By definition if $|\Lambda_1| \leqslant |\Lambda_2|$ there is an injection j of Λ_1 into Λ_2. j extends to an injective homomorphism $\tilde{j}$ of $\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}$ into $\mathfrak{p}_{\Lambda_2} \times \mathfrak{p}_{\Lambda_2}$; a proof similar to the one of Proposition (1.16) of Ref. 1 shows that there exists an injective * homomorphism $\alpha_{\tilde{j}}$ of $\overline{\Delta(\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}, \xi^S)}$ into $\overline{\Delta(\mathfrak{p}_{\Lambda_2} \times \mathfrak{p}_{\Lambda_2}, \xi^S)}$; moreover an application of Proposition (3.11) in Ref. 2 shows that

$$\begin{matrix}\Lambda_1 \not\subseteq \Lambda_2 \\ \Lambda_1, \Lambda_2 \subset \Lambda\end{matrix} \Rightarrow \overline{\Delta(\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}, \xi^S)} \not\subseteq \overline{\Delta(\mathfrak{p}_{\Lambda_2} \times \mathfrak{p}_{\Lambda_2}, \xi^S)}.$$

On the other hand, j and $\tilde{j}$ and $\alpha_{\tilde{j}}$, are surjective if and only if $|\Lambda_1| = |\Lambda_2|$.

Lemma (2.9): If $\Lambda_1 \in \mathfrak{p}_\Lambda$, then

$$\Delta(\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}, \xi^S) \simeq \bigotimes_{i=1}^{|\Lambda_1|} M_{2\times 2},$$

the tensor product of two-by-two matrices. Consequently,

$$\Delta(\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}, \xi^S) = \overline{\Delta(\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}, \xi^S)}.$$

In order to prove the first statement let us describe the correspondence explicitly.

Let σ_x and σ_y be the usual two-by-two Pauli spin matrices; they generate the two-by-two matrices. Let

$$\sigma^j_y = \underbrace{1 \otimes \cdots \otimes \underbrace{\sigma_y}_{i\text{th place}} \otimes \cdots \otimes 1}_{|\Lambda_1|}.$$

$$\sigma^{+j}_y = \underbrace{1 \otimes \cdots \otimes \underbrace{\sigma_y}_{j\text{th place}} \otimes \cdots \otimes 1}_{|\Lambda_1|}.$$

These matrices satisfy the relations

$$\sigma^i_x\sigma^j_x = \sigma^j_x\sigma^i_x,$$

$$\sigma^i_y\sigma^j_y = \sigma^j_y\sigma^i_y,$$

$$\sigma^i_x\sigma^j_y = (1 - 2\delta_{i,j})\sigma^j_y\sigma^i_x,$$

and generate

$$\bigotimes_{i=1}^{|\Lambda_1|} M_{2\times 2}.$$

Then let us define

$$\sigma_x^X = \prod_{x_i \in X} \sigma^i_x, \quad \sigma_y^Y = \prod_{x_j \in Y} \sigma^j_y, \quad X,Y \in \mathfrak{p}_{\Lambda_1}.$$

The elements $i^{-|X\cap Y|}\sigma_x^X\sigma_y^Y$, $X,Y \in \mathfrak{p}_{\Lambda_1}$, are both unitary and Hermitian: They span linearly,

$$\bigotimes_{i=1}^{|\Lambda_1|} M_{2\times 2},$$

and the correspondence

$$\delta^S_{X,Y} = i^{-|X\cap Y|}\sigma_x^X\sigma_y^Y$$

extends to a * isomorphism of $\overline{\Delta(\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}, \xi^S)}$ to

$$\bigotimes_{i=1}^{|\Lambda_1|} M_{2\times 2};$$

but that last algebra is closed hence $\Delta(\mathfrak{p}_{\Lambda_1} \times \mathfrak{p}_{\Lambda_1}, \xi^S)$ is closed for $|\Lambda_1| < \infty$.

Proposition 2.10: $\overline{\Delta(\mathfrak{p}_\Lambda \times \mathfrak{p}_\Lambda, \xi^S)}$ is isomorphic to the UHF algebra

$$\overline{\bigcup_{N\in Z} \bigotimes_{i=1}^{N} M^{\|\cdot\|}_{2\times 2}}.$$

It is clear that

$$\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)=\bigcup_{\Lambda_1\in\mathfrak{p}_\Lambda}\Delta(\mathfrak{p}_{\Lambda_1}\times\mathfrak{p}_{\Lambda_1},\xi^S)$$

$$=\bigcup_{\Lambda_1\in\mathfrak{p}_\Lambda}\overline{\Delta(\mathfrak{p}_{\Lambda_1}\times\mathfrak{p}_{\Lambda_1},\xi^S)}.$$

This last term is * isomorphic to the union of tensor products and the isomorphism is isometric hence it extends to a * isomorphism of the closures.

3. SOME INTERESTING STATES OF $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$

Let Δ be the diagonal of $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$. Since the ξ^S we have considered is such that

$$\xi^S(X,X;Y,Y)=1,\quad X,Y\in\mathfrak{p}_\Lambda, \tag{3.1}$$

the subalgebra generated by the $\delta^S_{X,X}$, $X\in\mathfrak{p}_\Lambda$, is Abelian.

Definition (3.2) (cf. Ref. 4): $\mathscr{C}$ is the convex and weakly closed set of states ω of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$ satisfying

$$\omega(\delta^S_{X,Y})=0\quad\text{if } X\neq Y.$$

This set has a nice characterization in terms of the probability measures on $\mathscr{P}_\Lambda$, the dual of $\mathfrak{p}_\Lambda$, namely:

Theorem (3.3): Let f be a positive type function on $\mathfrak{p}_\Lambda$ normalized to 1, namely $f(\phi)=1$; the correspondence

$$f\rightarrow\omega_f:\delta^S_{X,Y}\rightarrow\omega_f(\delta^S_{X,Y})=\begin{cases}0 & \text{if } X\neq Y\\ f(X) & \text{if } X=Y\end{cases}$$

is a bijection of the positive type functions on $\mathfrak{p}_\Lambda$ normalized to 1 onto $\mathscr{C}$. It sends characters onto pure states.

Indeed if $\omega\in\mathscr{C}$, the restriction of ω to the sublagebra generated by the $\delta^S_{X,X}$ is still a state on an Abelian algebra. $X\in\mathfrak{p}_\Lambda\rightarrow\omega(\delta^S_{X,X})$ is a positive type function. Vice versa, if f is a positive type function on $\mathfrak{p}_\Lambda$, then ω_f defined in (3.3) is a state of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$ [cf. Ref. 3, Corollary (2.15)].

Corollary (3.4): Let $\omega\in\mathscr{C}$. There exists a probability measure μ on $\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda$ such that

$$X,Y\in\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda\rightarrow\omega(\delta^S_{X,Y})=\mathscr{F}\mu(X,Y).$$

Proof is a direct application of Ref. 3, Theorem (2.14).

As we shall see in Sec. 4, the $\delta^S_{X,Y}$ are the quantization of the functions $\hat{X},\hat{Y}\in\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda\rightarrow(-1)^{|X\cap\hat{Y}|+|Y\cap\hat{X}|}$; hence (cf. Ref. 4) Corollary (3.4) justify the name of quasiclassical state for elements of $\mathscr{C}$. Let us remark that we have defined quasiclassical states with respect to the diagonal of $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$; one can as well choose any maximal subgroup of $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$ on which ξ^S is one.

According to Ref. 3 Theorem (2.16) one can write explicitly the GNS representation associated to a quasiclassical state.

Proposition 3.5: Let ν be a probability measure on $\mathscr{P}_\Lambda$. Let f be its Fourier transform and ω_f the quasiclassical state of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$ associated to f by Theorem (3.3). Let $\mathscr{H},\pi,\Omega$ be the Hilbert space, the representation, and the cyclic vector of the GNS construction associated with ω_f. Then

$$\mathscr{H}=l_2(\mathfrak{p}_\Lambda)\otimes L_2(\mathscr{P}_\Lambda,\nu),$$
$$\Omega=\delta_\phi\otimes 1,$$
$$\{\pi(\delta^S_{X,Y})F\}(Z_1,\hat{Z}_2)$$
$$=i^{-|X\cap Y|}i^{|Y|}(-1)^{|Z_1\triangle\hat{Z}_2\cap Y|}F(Z_1\triangle X\triangle Y,\hat{Z}_2).$$

$\mathscr{H}$ is separable.

Proof is an obvious extension of theorem (2.16) of Ref. 3 with the special choice of application S and T:

$$S(X,Y)=(X\triangle Y,\phi), \tag{3.6}$$

$$T(X,Y)=(Y,Y),\quad X,Y\in\mathfrak{p}_\Lambda. \tag{3.7}$$

Let us remark that the diagonal Δ of $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$ is maximal Abelian with respect to ξ^S hence the pure states on the diagonal correspond to pure states of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$. Among quasiclassical states, we have the product states $\omega_{\{a_i\}}$ which are defined by

$$\omega_{\{a_i\}}(\delta^S_{X,Y})=\prod_{x_i\in\Lambda}[(1-|\{x_i\}\cap X|)(1-|\{x_i\}\cap Y|)$$

$$+|\{x_i\}\cap X|\,|\{x_i\}\cap Y|a_i], \tag{3.8}$$

where a_i is any function from $x_i\in\Lambda\rightarrow[-1,+1]$. The product converges since at most a finite number of terms are different from one. Moreover $\omega_{\{a_i\}}\in\mathscr{C}$ since if $X\neq Y$ there is at least a point $x_i\in X\triangle Y$, hence a factor, which is zero. $\omega_{\{a_i\}}$ is pure if and only if $|a_i|=1$; since if it exists an a_i which is different from ± 1, the corresponding factor

$$(1-|x_i\cap X|)(1-|x_i\cap Y|)+|x_i\cap X|\,|x_i\cap Y|a_i$$

is a convex combination of two functions of this type. On the other hand, if $a_i=\pm 1$ let $\Lambda_1\subset\Lambda$ be the set of $x_i\in\Lambda$ for which $a_i=-1$. Then

$$\omega_{\{a_i\}}(\delta^S_{X,X})=(-1)^{|\Lambda_1\cap X|}=\omega_{\Lambda_1}(\delta^S_{X,X}). \tag{3.9}$$

It is a character of $\mathfrak{p}_\Lambda$, hence $\omega_{\{a_i\}}$ is pure.

All these states are quite familiar indeed: $\omega_{\{-1\}}=\omega_\Lambda$ is the usual Fock representation of the UHF algebra, $\omega_{\{+1\}}=\omega_\phi$ is the anti Fock representation of the UHF algebra, $\omega_{\{0\}}$ is the central state of the UHF algebra. In that context the $\omega_{\{a_i\}}$ for $a_i=\pm 1$ are the analog of the coherent states. (Those states are not the coherent states which are usually considered for spin systems, they are the states with a fixed particle number; but they are much more the analog of the usual coherent staes of the CCR.) Indeed we have the following:

Proposition (3.10): Let ω_{Λ_1} and ω_{Λ_2} be two pure product states. Then

$$\omega_{\Lambda_1}(\delta^S_{X,Y})=\omega_{\Lambda_2}(\delta^S_{X,Y})(-1)^{|\Lambda_1\triangle\Lambda_2\cap X|},\quad X,Y\in\mathfrak{p}_\Lambda.$$

Moreover ω_{Λ_1} and ω_{Λ_2} are unitary equivalent if and only if

$$\Lambda_1\triangle\Lambda_2\in\mathfrak{p}_\Lambda.$$

This proposition for the usual CCR is nothing but the fact that two coherent states differ by a character on the phase space. Moreover they are unitary equivalent if and only if this character is continuous. In our case, if $\Lambda_1\triangle\Lambda_2\in\mathfrak{p}_\Lambda$ then $X\in\mathscr{P}_\Lambda\rightarrow(-1)^{|\Lambda_1\triangle\Lambda_2\cap X|}$ is a continuous character.

At this stage it is worthwhile to give other explicit realizations of these representations of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$ which are in a sense closer to the usual form of the Schrödinger x and p representation. For the sake of symmetry, it is easier to diagonalize the $\pi(\delta^S_{X,\phi})$ [resp. $\pi(\delta^S_{\phi,X})$] than the $\pi(\delta^S_{X,X})$; the results are equivalent since it is nothing but a permutation of $\{\sigma_x,\sigma_y,\sigma_z\}$, the familiar Pauli spin matrices.

$\mathfrak{p}_\Lambda$ is a discrete countable Abelian group. It is locally compact and the Haar measure on $\mathfrak{p}_\Lambda$ is the sum over the points. We consider $l_2(\mathfrak{p}_\Lambda)$ which is the space of functions f from $\mathfrak{p}_\Lambda\to\mathbb{C}$ such that

$$\|f\|_l^2=\sum_{X\in\mathfrak{p}_\Lambda}|f(X)|^2<\infty. \tag{3.11}$$

$\mathscr{P}_\Lambda$ its dual group is Abelian and compact; it is isomorphic to the group of subsets of Λ equipped with the symmetric difference. We denote by dX its normalized Haar measure; moreover $L_2(\mathscr{P}_\Lambda)$ is the set of functions $f:\mathscr{P}_\Lambda\to\mathbb{C}$ such that

$$\|f\|_L^2=\int_{\mathscr{P}_\Lambda}|f(X)|^2\,dX<\infty. \tag{3.12}$$

The duality correspondence is given by

$$(X,Y)\in\mathfrak{p}_\Lambda\times\mathscr{P}_\Lambda\to(X\mid Y)=(-1)^{|X\cap Y|}. \tag{3.13}$$

The Fourier transform exchange l_2 and L_2

$$\{\mathscr{F}f\}(Z)=\sum_{X\in\mathfrak{p}_\Lambda}(-1)^{|X\cap Z|}f(X),\quad f\in l_2, \tag{3.14}$$

$$\{\mathscr{F}f\}(Z)=\int_{\mathscr{P}_\Lambda}\partial X\,(-1)^{|X\cap Z|}f(X),\quad f\in L_2.$$

Definition (3.15): π_1^l, π_1^L, π_2^l, and π_2^L, are the four representations of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$ defined by

$$\{\pi_1^l(\delta^S_{X,Y})f\}(Z)=i^{-|X\cap Y|}(-1)^{|X\cap Z|}f(Z\triangle Y),\quad f\in l_2,$$

$$\{\pi_1^L(\delta^S_{X,Y})f\}(Z)=i^{-|X\cap Y|}(-1)^{|X\cap Z|}f(Z\triangle Y),\quad f\in L_2,$$

$$\{\pi_2^l(\delta^S_{X,Y})f\}(Z)=i^{|X\cap Y|}(-1)^{|Y\cap Z|}f(Z\triangle X),\quad f\in l_2,$$

$$\{\pi_2^L(\delta^S_{X,Y})f\}(Z)=i^{|X\cap Y|}(-1)^{|Y\cap Z|}f(Z\triangle X),\quad f\in L_2.$$

The operators we have defined are evidently unitary, hence the only thing we have to prove is that they are indeed representations of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$. The proof is purely algebraic. The next proposition gives the relation between these representations.

Proposition (3.16): The representations defined in the previous definition are irreducible.

If $|\Lambda|=\infty$, π_1^l and π_2^l (resp. π_1^L and π_2^L) are disjoint.

π_1^L and π_2^l (resp. π_1^L and π_2^l) are unitarily equivalent through a Fourier transformation.

Irreducibility is obvious. To prove the disjunction of π_1^l and π_2^l let us consider in l_2 a function δ_{X_0} whose support is $X_0\in\mathfrak{p}_\Lambda$:

$$\{\pi_1^l(\delta^S_{X,\phi})\delta_{X_0}\}(Z)=(-1)^{|X\cap X_0|}\delta_{X_0}(Z),\quad\forall Z\in\mathfrak{p}_\Lambda, \tag{3.17}$$

hence it is an eigenvector of the $\pi_1^l(\delta^S_{X,\phi})$. On the other hand, in l_2 an eigenvector of $\pi_2^l(\delta^S_{X,\phi})$ would satisfy

$$\{\pi_2^l(\delta^S_{X,\phi})f\}(Z)=f(Z\triangle X)=(-1)^{|X\cap X_0|}f(Z), \tag{3.18}$$

it would imply that $f(Z)=(-1)^{|Z\cap X_0|}$ but this function is not in l_2, if $|\Lambda|=\infty$.

To prove that π_1^L and π_2^L are disjoint it is sufficient to prove that π_1^l and π_2^L (resp. π_1^L and π_2^l) are unitarily equivalent. The unitary equivalence is given by Fourier transformation. It is interesting to remark that the disjunction of π_1^l and π_1^L can be seen in the following way: $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$ is a subgroup of $\mathfrak{p}_\Lambda\times\mathscr{P}_\Lambda$ and ξ^S can be extended by continuity to $\mathfrak{p}_\Lambda\times\mathscr{P}_\Lambda$ using the same expression:

$$\bar{\xi}^S(X,Y;X',Y')=i^{-|X\cap Y|-|X'\cap Y'|+|X\triangle X'\cap Y\triangle Y'|+2|X'\cap Y|}$$

for $(X,Y),(X',Y')\in\mathfrak{p}_\Lambda\times\mathscr{P}_\Lambda$. Consequently for $|\Lambda|=\infty$

$$\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}\not\subset\overline{\Delta(\mathfrak{p}_\Lambda\times\mathscr{P}_\Lambda,\bar{\xi}^S)}.$$

π_1^L can be extended by continuity to a representation of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathscr{P}_\Lambda,\bar{\xi}^S)}$, but not π_1^l. In that context, let us remark that $\overline{\Delta(\mathfrak{p}\times\mathscr{P},\bar{\xi}^S)}$ has only one representation up to quasi-equivalence according to the Mackey–Von Neumann uniqueness theorem.[5]

For Λ such that $|\Lambda|<\infty$, the distinction between $\mathfrak{p}_\Lambda$ and $\mathscr{P}_\Lambda$ has no meaning and all the representations defined in (3.15) coincide. Let $\pi=\pi_1$ and define $|\hat{X}_0\rangle$, $X_0\in\mathfrak{p}_\Lambda$ as the function δ_{X_0} which is an eigenvector of $\pi(\delta^S_{X,\phi})$. In the same way define $|\hat{X}_0\rangle$ as the function $\hat{\delta}_{X_0}(Z)=2^{-|\Lambda|/2}\times(-1)^{|X_0\cap Z|}$, the Fourier transform of δ_{X_0}. They both form an orthonormal set in $l_2(\mathfrak{p}_\Lambda)$, moreover:

$$\langle\hat{X}\mid Y\rangle=2^{-|\Lambda|/2}(-1)^{|X\cap Y|} \tag{3.19}$$

and we have the following decomposition of the identity:

$$I=2^{-|\Lambda|/2}\sum_{X,Y\in\mathfrak{p}_\Lambda}(-1)^{|X\cap Y|}|\hat{X}\rangle\langle Y|. \tag{3.20}$$

Up to now we dealt only with the quantum spin commutation rules. But there is a * isomorphism of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^F)}$ onto $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$ [see e.g., Ref. 1, theorem (2.36)] which allows to translate the state of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^S)}$ onto the states of $\overline{\Delta(\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda,\xi^F)}$. The * automorphism $\hat{\alpha}_\theta$ satisfies:

$$\hat{\alpha}_\theta(\delta^F_{X,Y})=\delta^S_{\tau_\theta(X,Y)}. \tag{3.21}$$

Notice that for the sake of simplicity we have incorporated the function F of theorem (2.36) of Ref. 1 into the definition of δ^F. This changes ξ^F by a trivial multiplier which still satisfies $\xi^F(X,Y;X,Y)=1$, $X,Y\in\mathfrak{p}_\Lambda$, [see Ref. 1 remark before proposition (2.32)]. Even this does not change the generators since $F(\{i\},\phi)=F(\phi,\{i\})=1$, $\forall i$. The simplification is especially clear for the classical states previously defined.

Proposition (3.22): Let ω^S be a state of $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}$ then $\omega^S\circ\widehat{\alpha}_\theta=\omega^F$ is a state of $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^F)}$. If moreover $\omega^S\in\mathscr{C}$, then:

$$\omega^S(\delta^S_{X,Y})=\omega^F(\delta^F_{X,Y}),\quad X,Y,\in\mathfrak{p}_A.$$

The first assertion is trivial. For the second one it is sufficient to remark that $\omega^S(\delta^S_{X,Y})$ is zero if $X\neq Y$ but on the diagonal

$$\tau_\theta(X,X)=(X,X),\quad \forall X\in\mathfrak{p}_A.$$

This proposition also shows that the product states, in the sense of Powers, of the Clifford algebra (see Ref. 6) are associated to the product states on the spin algebra. In the next section we shall come back to this isomorphism and we shall describe it as a transformation of the classical phase space which preserves the Haar measure.

4. QUANTIZATION

The formalism which has been developed in the previous sections is intended to quantify functions on the "phase space" $\mathscr{P}_A\times\mathscr{P}_A$. Indeed let f be a function on $\mathscr{P}_A\times\mathscr{P}_A$ with Fourier transform

$$\{\mathscr{F}f\}(X,Y)=\int_{\mathscr{P}_A\times\mathscr{P}_A}dX'\,dY'\,(-1)^{|X\cap Y'|+|X'\cap Y|}f(X',Y'),\tag{4.1}$$

where $dX'\,dY'$ is the normalized Haar measure on the compact group $\mathscr{P}_A\times\mathscr{P}_A$. One is tempted to define the quantized $\mathscr{Q}(f)$ of f according to the Weyl prescription

$$\mathscr{Q}^S(f)=\sum_{X,Y\in\mathfrak{p}_A}\{\mathscr{F}f\}(X,Y)\delta^S_{X,Y},\tag{4.2}$$

whenever it exists.

There are two main differences between our situation and the usual situation of the CCR over a finite number of degrees of freedom, namely:

(i) The multiplier we consider is not a bicharacter.

(ii) There is a whole set of nonequivalent representations of $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}$, which indicates that the $\mathscr{Q}^S(f)$ can be defined only within some representations, for some f.

Later on, we shall return to the first point; but now we shall give at least a set of functions which can be quantized within any representation.

Proposition (4.3): Let f be a l_1 function on $\mathfrak{p}_A\times\mathfrak{p}_A$,

$$A_f=\sum_{X,Y\in\mathfrak{p}_A}f(X,Y)\delta^S_{X,Y},$$

defines an element of $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}$. Moreover,

$$(A_f)^*=A_{f^*},\quad f^*(X,Y)=\overline{f(X,Y)},$$

$$A_{f_1}A_{f_2}=A_{f_1\times f_2},$$

where

$$f_1\times f_2(X,Y)=\sum_{X',Y'\in\mathfrak{p}_A}\xi^S(X,Y;X',Y')f_1(X',Y')$$

$$f_2(X\triangle Y',Y\triangle Y').\tag{4.4}$$

Equipped with this product $l_1(\mathfrak{p}_A\times\mathfrak{p}_A)$ is a Banach * algebra and

$$\|A_f\|\leqslant\|f\|_1.$$

We shall not give the proof since it is obvious. Notice that the set of A_f is nothing but the algebra $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}^{\|\cdot\|_1}$, [cf. Ref. 2 formula (2.10)]. In our case this algebra is important since we have the following:

Theorem 4.5: Let $\mathscr{B}_A$ be the algebra of functions on $\mathscr{P}_A\times\mathscr{P}_A$ which are Fourier transforms of an l_1 function on $\mathfrak{p}_A\times\mathfrak{p}_A$, then

$$\mathscr{Q}^S(f)=\sum_{X,Y\in\mathfrak{p}_A}\{\mathscr{F}f\}(X,Y)\delta^S_{X,Y},$$

where $\mathscr{F}f$ is the Fourier transform of f, defines an element of $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}$. Moreover,

$$\mathscr{Q}^S(f)^*=\mathscr{Q}^S(f^*),\quad f^*(X,Y)=\overline{f(X,Y)},$$

$$\mathscr{Q}^S(f_1)\mathscr{Q}^S(f_2)=\mathscr{Q}^S(f_1\circ f_2),\quad f,f_1,f_2\in\mathscr{B}_A,$$

where $f_1\circ f_2=\mathscr{F}(\mathscr{F}f_1\times\mathscr{F}f_2)$.

$$\mathscr{Q}^S(1)=1,$$

$$\mathscr{Q}^S(\chi_{X,Y})=\delta^S_{X,Y},\quad X,Y\in\mathfrak{p}_A,$$

where

$$\chi_{X,Y}(X',Y')=(-1)^{|X\cap Y'|+|X'\cap Y|}.$$

Notice that since $\mathscr{P}_A\times\mathscr{P}_A$ is compact any continuous function on $\mathscr{P}_A\times\mathscr{P}_A$ can be approximated uniformly by elements of $\mathscr{B}_A$. Conversely let A be an element of $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}$ and ω^S_c be the central state of $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}$ then

$$X,Y\in\mathfrak{p}_A\times\mathfrak{p}_A\rightarrow\omega^S_c(\delta_{X,Y}A),$$

is a bounded (continuous) function on $\mathfrak{p}_A\times\mathfrak{p}_A$. If, moreover, it is a l_1 function then it is the Fourier transform of an element of $\mathscr{B}_A$. Since $\mathscr{P}_A\times\mathscr{P}_A$ is compact we can use the inversion theorem to state the following theorem:

Theorem 4.6: Let $A\in\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}$ be such that

$$X,Y\in\mathfrak{p}_A\times\mathfrak{p}_A\rightarrow\omega^S_c(\delta^S_{X,Y}A)$$

is an l_1 function then

$$A=\mathscr{Q}^S(f_A),$$

where

$$f_A(X,Y)=\sum_{X',Y'\in\mathfrak{p}_A}(-1)^{|X'\cap Y|+|X\cap Y'|}\omega^S_c(\delta^S_{X',Y'}A).$$

We recall that ω^S_c, the central state of $\overline{\Delta(\mathfrak{p}_A\times\mathfrak{p}_A,\xi^S)}$ is defined by

$$\omega^S_c(\delta^S_{X,Y})=\delta_{X,\phi}\delta_{Y,\phi},\quad X,Y\in\mathfrak{p}_A,$$

where $\delta_{X,Y}$ in the right-hand side is a Kronecker symbol.

The next property has an obviously analog for usual CCR.

Proposition (4.7): Let f and g be functions in $\mathscr{B}_A$ such that

$$f(X,Y)=f(X,Y'),\quad \forall\, X,Y,Y'\in \mathfrak{p}_\Lambda,$$
$$[\text{resp. } f(X,Y)=f(X',Y),\quad \forall\, X,X',Y\in \mathfrak{p}_\Lambda],$$
$$[\text{resp. } f(X,Y)=h(X\triangle Y),\quad \forall\, X,Y\in \mathfrak{p}_\Lambda],$$
then $f\circ g=f\cdot g$, the usual product of functions.

All the previous results are still valid for Fermi systems if ξ^S is changed into ξ^F. Nevertheless, we give in what follows another relation between the two quantizations.

τ_θ acting on $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$, it is possible to induce transformation of $\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda$ according to

Proposition (4.8): Let τ_θ be the automorphism of $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$ defined by

$$\tau_\theta(X,Y)=(X\triangle\theta(X\triangle Y),Y\triangle\theta(X\triangle Y)),\quad X,Y\in\mathfrak{p}_\Lambda,$$

with $\theta(\{x_i\})=\{x_j\in\Lambda; j<i\}$ for some arbitrary order on Λ; then there exists a unique automorphism ${}^t\tau_\theta$ of $\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda$ such that, if $b^S(X,Y;X',Y')$ is the bicharacter associated with ξ^S:

$$b^S((X,Y);\tau_\theta(X',Y'))=b_S({}^t\tau_\theta(X,Y);(X',Y')),\qquad (4.9)$$

$\forall\, X,Y\in\mathscr{P}_\Lambda$, $\forall\, X',Y'\in\mathfrak{p}_\Lambda$. It is defined by

$${}^t\tau_\theta(X,Y)=(X_\theta(X,Y),Y_\theta(X,Y)),\quad X,Y\in\mathscr{P}_\Lambda,\qquad (4.10)$$

where

$$X_\theta(X,Y)=\{x_i\in\Lambda;(-1)^{|\theta(\{x_i\})\cap Y|+|\{x_i\}\triangle\theta(\{x_i\})\cap X|}=-1\},\qquad (4.11)$$

$$Y_\theta(X,Y)=\{x_j\in\Lambda;(-1)^{|\theta(\{x_j\})\cap X|+|\{x_j\}\triangle\theta(\{x_j\})\cap Y|}=-1\}.$$

The existence and the uniqueness of ${}^t\tau_\theta$ is ensured by the fact that b^S is a bijection of $(\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda)^\wedge$ onto $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$ the dual of $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$. Moreover

$$X'',Y''\in\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda\to b^S((X,Y);\tau_\theta(X'',Y''))$$

is a character of $\mathfrak{p}_\Lambda\times\mathfrak{p}_\Lambda$, hence it is of the form

$$b^S((X_\theta,Y_\theta);(X'',Y''))=b^S({}^t\tau_\theta(X,Y);(X'',Y'')).$$

It is easy to see that X_θ, Y_θ are defined by

$$X_\theta(X,Y)=\{x_i\in\Lambda;b^S((X,Y);\tau_\theta(\phi,\{x_i\}))=-1\},$$

and similarly for Y_θ.

Since τ_θ is a bijective automorphism and b^S a nondegenerate bicharacter ${}^t\tau_\theta$ is an injective automorphism; moreover it retains some properties of τ_θ as exemplified by the following proposition.

Proposition (4.12): ${}^t\tau_\theta$ is a bijective automorphism of $\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda$ which satisfies

(i) ${}^t\tau_\theta^2=i$ (the identity automorphism).

(ii) ${}^t\tau_\theta(X,X)=(X,X)$, $\forall X\in\mathscr{P}_\Lambda$.

(iii) ${}^t\tau_\theta\circ\tilde\tau_{\theta_0}=\tilde\tau_{\theta_0}\circ{}^t\tau_\theta$,

where $\tilde\tau_{\theta_0}$ denotes the extension of τ_{θ_0} to $\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda$ [see Ref. 1, (2.13)].

(iv) ${}^t\tau_\theta(\{x_i\},\phi)=(\{x_i\}\triangle\Theta_i,\Theta_i)$,

where $\Theta_i=\{x_j\in\Lambda; j>i\}$.

All these properties are easy to prove. Using this automorphism, we can give another expression for $\mathscr{Q}^F(f)$ where $f\in\mathscr{B}_\Lambda$. Indeed

$$\hat\alpha_\theta(\mathscr{Q}^F(f))=\sum_{X,Y\in\mathfrak{p}_\Lambda}\{\mathscr{F}f\}(X,Y)\hat\alpha_\theta(\delta^F_{XY})$$

$$=\sum_{X,Y\in\mathfrak{p}_\Lambda}\{\mathscr{F}f\}(X,Y)\delta^S_{\tau_\theta(X,Y)}$$

$$=\sum_{X,Y\in\mathfrak{p}_\Lambda}\{\mathscr{F}f\}(\tau_\theta(X,Y))\delta^S_{X,Y}$$

$$=\sum_{X,Y\in\mathfrak{p}_\Lambda}\{\mathscr{F}(f\circ{}^t\tau_\theta\}(X,Y)\delta^S_{X,Y}.$$

Consequently the correspondence

$$f\to f\circ{}^t\tau_\theta=f^F,\qquad (4.13)$$

satisfies

$$\hat\alpha_\theta(\mathscr{Q}^F(f))=\mathscr{Q}^S(f^F).\qquad (4.14)$$

Hence we can state the following:

Proposition (4.15): There exists an automorphism ${}^t\tau_\theta$ of $\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda$ which leaves invariant the Haar measure and such that

$$\hat\alpha_\theta(\mathscr{Q}^F(f))=\mathscr{Q}^S(f\circ{}^t\tau_\theta),$$

$\forall f\in\mathscr{B}_\Lambda$.

Indeed the group $\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda$ is compact, hence any continuous function on $\mathscr{P}_\Lambda\times\mathscr{P}_\Lambda$ can be uniformly approximated by finite linear combinations of characters.

[1] Ph. Combe, R. Rodriguez, M. Sirugue-Collin, and M. Sirugue, Commun. Math. Phys. **63**, 219 (1978).
[2] J. Manuceau, M. Sirugue, D. Testard, and A. Verbeure, Commun. Math. Phys. **32**, 231 (1973).
[3] R. Beaume, J. Manuceau, A. Pellet, and M. Sirugue, Commun. Math. Phys. **38**, 29 (1974).
[4] H.J. Brascamp, thesis, University of Groningen (1970).
[5] G.W. Mackey, Duke Math. J. **16**, 313 (1949).
[6] R.T. Powers, thesis, Princeton University (1967).

An analytical theory of pulse wave propagation in turbulent media

K. Furutsu

Radio Research Laboratories, Koganei-shi, Tokyo 184, Japan
(Received 14 April 1978)

The theory of pulse wave propagation in turbulent media is developed starting from the space–time transport equation with the forward-scattering approximation. The solutions are obtained by a fully analytical method based on the eigenfunction expansion, and the averaged intensity of plane wave pulse is presented by two different expressions for both the Gaussian and Kolmogorov turbulence spectra. These two expressions are given in the series, and the convergence of each series is good when the convergence of the other series is poor; in the case of the Gaussian turbulence spectrum, one of these expressions precisely agrees with the previous one obtained by Sreenivasiah *et al.* (1976). In connection with the pulse wave width, the pulse moments are evaluated in detail. The resolvent function is fully used to find the eigenvalues and eigenfunctions.

I. INTRODUCTION AND PRELIMINARIES

Recently, the pulse wave propagation has been studied by many authors[1-6] in connection with the radio signals from pulsars and from communication satellites, and the media in these cases are the interstellar medium and the ionosphere or, more generally, plasma in turbulence. In cases of the radio signals propagated through unionized media, the broadening of pulses are coursed by discrete scatterers such as rains, clouds, and fogs as well as by the atmospheric turbulence.[7-10]

The theoretical works seem to have been achieved so far by using the two-frequency mutual coherence functions which were obtained, except the work by Sreenivasiah *et al.*,[9] by numerically solving their equations, and were followed by the numerical integration over the frequency difference for the Fourier inversion. The analytical works also have been tried by use of a method similar to the Rytov approximation, although their results are applicable only to the cases of short optical distances.[3,7]

In this paper, the basic equation is chosen to be the space–time transport equation for the angular-frequency distribution function of wave, since the equation has a symmetrical form in space and time, and it is then subjected to the forward-scattering approximation:

In the space-time coordinate system $(\bar{x}) = (x,t)$, where x denotes the three-dimensional space coordinates and t the time, the original wave equation is assumed to be of the form

$$\left[\left(c^{-1}\frac{\partial}{\partial t}\right)^2[1+\epsilon(\bar{x})] - \left(\frac{\partial}{\partial x}\right)^2\right]\psi(\bar{x}) = j(\bar{x}), \tag{1.1}$$

where $\psi(\bar{x})$ is the (scalar) real wavefunction, $\epsilon(\bar{x})$ is the fluctuating part of the square of refractive index, $j(\bar{x})$ is the wave source density, and c is the wave velocity in the space of $\epsilon(\bar{x}) = 0$. Generally, $\epsilon(\bar{x})$ is also a function of the operator $\partial/\partial t$ but its explicit dependence will be suppressed in the following, with the understanding that $\epsilon(\bar{x})$ is a well-ordered function in which $\partial/\partial t$ is ordered, e.g., to the left of $\bar{x}$.

Here, $\epsilon(\bar{x})$ is assumed to be a small random quantity of the magnitude of $|\epsilon(\bar{x})| \ll 1$ with $\langle\epsilon(\bar{x})\rangle = 0$, and its space–time change is assumed to be negligibly small within the range of wavelength and period of time considered. The above conditions also enable us, if we are interested in the electromagnetic wave propagation in turbulent media, to use the scalar wave Eq. (1.1) instead of the Maxwell's equation, since, on these conditions, the waves are scattered mostly in the forward direction and the depolarization effect also becomes very small.[5,11,12]

Concerning the statistics of $\epsilon(\bar{x})$, in Ref. 5, the Markov approximation was used, based on the forward scattering approximation or parabolic wave equation, to derive the moment equations of various order without the assumption of Gaussian statistics. In this connection, it will be noted that the condition $|\epsilon| \ll 1$ in (1.1) is generally enough to enable us, when evaluating average values of the form $\langle\epsilon(\bar{x})\psi(\bar{x}_1)\psi(\bar{x}_2)\cdots\rangle$, to use the Gaussian statistics for $\epsilon(\bar{x})$ even when it really does not follow the Gaussian (see Appendix C).

Under a rather general condition, the equation of the coherence function of wave, $\langle\psi(\bar{x}_1)\psi(\bar{x}_2)\rangle$, can be shown to have the form of the Bethe–Salpeter equation, not only in media following the Gaussian statistics but also, e.g., in media of random discrete scatterers where their refractive indexes could be very large and their sizes be very small as compared to the wavelength.[13,14] From the equation of the coherence function, the conventional space–time transport equation can be derived under the conditions to be described later, and has the form[15]

$$\left[\Omega\cdot\frac{\partial}{\partial\rho} + c^{-1}\frac{\partial}{\partial t} + \gamma(\Omega,\omega)\right] I(\Omega,\omega;\bar{\rho}) = \int_{-\infty}^{\infty} d\omega' \int d\Omega'\sigma(\Omega,\omega|\Omega',\omega')I(\Omega',\omega';\bar{\rho}) + J_c(\Omega,\omega;\bar{\rho}), \tag{1.2}$$

where Ω is the three-dimensional unit vector, ω the angular frequency, and $I(\Omega,\omega;\bar{\rho}) = I(\Omega,-\omega;\bar{\rho})$ is the angular-frequency distribution function at $\bar{\rho} = (\rho,t)$; the total irradiance $I(\bar{\rho})$ is given by

$$I(\bar{\rho}) = \int_{-\infty}^{\infty} d\omega \int d\Omega\, I(\Omega,\omega;\bar{\rho}); \tag{1.3}$$

 0022-2488/79/040617-12$01.00

$\sigma(\Omega,\omega|\Omega',\omega')$ is the scattering cross section per unit volume, per unit solid angle, and per unit frequency and, in terms of the correlation function of medium with the Fourier representation

$$D(\bar{x}_1-\bar{x}_2)=\langle\epsilon(\bar{x}_1)\epsilon(\bar{x}_2)\rangle$$
$$=(2\pi)^{-4}\int d\bar{\lambda}\tilde{D}(\bar{\lambda})\exp[-i\bar{\lambda}\cdot(\bar{x}_1-\bar{x}_2)], \quad (1.4)$$

$\bar{\lambda}\cdot\bar{x}=\lambda\cdot x-\omega t,\quad(\bar{\lambda})=(\lambda,\omega),\quad d\bar{\lambda}=d\lambda\, d\omega,$

is given by

$$\sigma(\Omega,\omega|\Omega',\omega')=\sigma(\Omega,-\omega|\Omega',-\omega')$$
$$\simeq(32\pi^3)^{-1}(\omega/c)^4\tilde{D}[(\omega\Omega-\omega'\Omega')/c,\omega-\omega'], \quad (1.5)$$

$|\omega|\gg|\omega-\omega'|;$

it is connected to the extinction coefficient $\gamma(\Omega,\omega) = \gamma(\Omega,-\omega)$ by the relation

$$\omega\gamma(\Omega,\omega)=\int_{-\infty}^{\infty}d\omega'\int d\Omega'\omega'\sigma(\Omega',\omega'|\Omega,\omega); \quad (1.6)$$

$J_c(\Omega,\omega;\bar{\rho})$ is the angular-frequency distribution of the wave source.

Generally, including the case of random discrete scatterers, the transport Eq. (1.2) is derived from the Bethe–Salpeter equation, under the condition[14] that the changes of $I(\Omega,\omega;\bar{\rho})$ and also of $\sigma(\Omega',\omega'|\Omega,\omega)$ are negligibly small for the variation of $\omega\to\omega+\Delta\omega$ over the range of $\Delta\omega/c$ of the order of magnitude of $\gamma(\Omega,\omega)\ll\omega/c$ and of l^{-1}, where l is the smallest scale of $I(\Omega,\omega,\bar{\rho})$ in the direction of Ω and in the "time" ct; thereby the same condition holds also with respect to $\sigma(\Omega,\omega|\Omega',\omega')$ by virtue of the symmetry.

In the particular case of frozen model where the random medium is spatially homogeneous and moving with a constant velocity v, then, the medium correlation function has the form

$$\langle\epsilon(\bar{x}_1)\epsilon(\bar{x}_2)\rangle=D[x_1-x_2-v(t_1-t_2)], \quad (1.7)$$

which, according to (1.5), gives

$$\sigma(\Omega,\omega|\Omega',\omega')=(4\pi)^{-2}(\omega/c)^4\tilde{D}[(\omega\Omega-\omega'\Omega')/c]$$
$$\times\delta[\omega-\omega'-(\omega\Omega-\omega'\Omega')\cdot v/c]$$
$$\simeq\sigma(\Omega-\Omega')\delta[\omega-\omega'-\omega(\Omega-\Omega')\cdot v/c],$$
$$|v/c|\ll1, \quad (1.8)$$

where

$$\sigma(\Omega-\Omega')=(4\pi)^{-2}(\omega/c)^4\tilde{D}[(\Omega-\Omega')\omega/c], \quad (1.9)$$

$$\gamma\simeq\int d\Omega'\sigma(\Omega'-\Omega).$$

In the following where the forward scattering is assumed to be essential, as in the typical case of light-wave propagation in turbulent media, we shall use the spatial coordinate system $(x)=(\mathbf{x},z)$ with the z axis taken in the main direction of wave propagation, and also put $(\Omega)=(\mathbf{\Omega},\Omega_z)$ where $\Omega_z=(1-\mathbf{\Omega}^2)^{1/2}\sim1-\frac{1}{2}\mathbf{\Omega}^2$. Then, (1.2) yields the equation of the angular-frequency distribution function $I(\Omega,\omega;\bar{\rho})=I(\mathbf{\Omega},\omega;\boldsymbol{\rho},z,t)$ of the form $[J_c=0]$

$$\left[\mathbf{\Omega}\cdot\frac{\partial}{\partial\boldsymbol{\rho}}+(1-\tfrac{1}{2}\mathbf{\Omega}^2)\frac{\partial}{\partial z}+c^{-1}\frac{\partial}{\partial t}+\gamma\right]I(\mathbf{\Omega},\omega;\boldsymbol{\rho},z,t)$$
$$=\int_{-\infty}^{\infty}d\omega'\int_{-\infty}^{\infty}d\mathbf{\Omega}'\sigma(\mathbf{\Omega}-\mathbf{\Omega}',\omega-\omega')I(\mathbf{\Omega}',\omega';\boldsymbol{\rho},z,t). \quad (1.10)$$

Here, $\sigma(\mathbf{\Omega}-\mathbf{\Omega}',\omega-\omega')$ on the right-hand side is to be given by (1.8) with the condition $|v/c|\ll1$.

In order to solve (1.10), it is most convenient[16] to make the Fourier transformation, on its both sides, with respect to the variables $\mathbf{\Omega}$ and ω, and hence, on assuming that the wave has a narrow frequency bandwidth with the center at $\pm\omega_0$, we put $k=\omega_0/c$ and define a new function, according to

$$I(r,\tau;\boldsymbol{\rho},z,t)=\int_{-\infty}^{\infty}d\omega\int_{-\infty}^{\infty}d\mathbf{\Omega}I(\mathbf{\Omega},\omega;\boldsymbol{\rho},z,t)$$
$$\times\exp\{i[(\omega\mp\omega_0)\tau\mp k\,\mathbf{\Omega}\cdot\mathbf{r}]\},\quad k=\omega_0/c, \quad (1.11)$$

where signs $\mp$ are to be taken according as $\omega\gtrless0$, respectively. Hence, since $I(\mathbf{\Omega},\omega;\boldsymbol{\rho},z,t)$ is an even function of ω, the contributions from the positive and negative frequency parts are exactly complex conjugate to each other, and hence we shall consider only the contribution from the positive frequency part in the following.

From (1.10), we find the equation of the new function thus defined, in the form

$$\left[c^{-1}\frac{\partial}{\partial t}+\frac{\partial}{\partial z}+ik^{-1}\frac{\partial}{\partial\mathbf{r}}\cdot\frac{\partial}{\partial\boldsymbol{\rho}}+\frac{1}{2}\left(k^{-1}\frac{\partial}{\partial\mathbf{r}}\right)^2\frac{\partial}{\partial z}\right.$$
$$\left.+kV(\mathbf{r}-\mathbf{v}\tau)\right]I(\mathbf{r},\tau;\boldsymbol{\rho},z,t)=0, \quad (1.12)$$

where, by (1.8) and (1.9),

$$V(\mathbf{r}-\mathbf{v}\tau)=k^{-1}\int_{-\infty}^{\infty}d\omega\int_{-\infty}^{\infty}d\mathbf{\Omega}\sigma(\mathbf{\Omega},1)\delta(\omega-k\,\mathbf{\Omega}\cdot\mathbf{v})$$
$$\times\{1-\exp[i(\omega\tau-k\,\mathbf{\Omega}\cdot\mathbf{r})]\}$$
$$=\tfrac{1}{4}k\int_{-\infty}^{\infty}dz[D(0,z)-D(\mathbf{r}-\mathbf{v}\tau,z)]. \quad (1.13)$$

Equation (1.12) becomes further simple by the introduction of new coordinates, according to

$$\zeta=ct-z,\quad\mathbf{r}'=\mathbf{r}-\mathbf{v}\tau, \quad (1.14)$$

and by the change of variable from $z\to\zeta$ and $\mathbf{r}\to\mathbf{r}'$:

$$\left[c^{-1}\frac{\partial}{\partial t}+ik^{-1}\frac{\partial}{\partial\mathbf{r}'}\cdot\frac{\partial}{\partial\boldsymbol{\rho}}-\frac{1}{2}\left(k^{-1}\frac{\partial}{\partial\mathbf{r}'}\right)^2\frac{\partial}{\partial\zeta}\right.$$
$$\left.+kV(\mathbf{r}')\right]I(\mathbf{r}';\boldsymbol{\rho},\zeta,t)=0. \quad (1.15)$$

Here, when the media have the Kolmogorov spectrum, $V(\mathbf{r})$ has the form

$$V(\mathbf{r}) \sim \begin{cases} \beta' |k\mathbf{r}|^2, & |k_m \mathbf{r}| \ll 1, \\ \beta |k\mathbf{r}|^\alpha, \quad \alpha = \frac{5}{3}, & |k_m \mathbf{r}| \gg 1, \end{cases} \tag{1.16}$$

where β' and β are nondimensional constants and k_m is the minimum length associated with the index of refraction fluctuation; in terms of the structure constant $\overline{C}_n^2$ (meter$^{-2/3}$)

$$\beta' = 1.65k^{-1} l_0^{-1/3} \overline{C}_n^2, \quad k_m l_0 = 5.92, \qquad \beta = 1.46k^{-2/3} \overline{C}_n^2. \tag{1.17}$$

Equation (1.15) is obviously equivalent to the equation derived from the equation of the two-frequency mutual coherence function,[18] which is based on the forward-scattering approximation or on the Markov random process approximation,[5,7] and the latter equation has been exclusively used by previous authors. Here, it will be worthwhile to note that the transport Eq. (1.2) has a symmetrical form in space and time and is applicable also to the backscattered waves as well, and therefore the equation of backscattered waves corresponding to (1.15) can also be obtained in a similar fashion, whereas the situation will be more complicated when using the equation of two-frequency mutual coherence function which is based on the forward-scattering approximation. The above statement is valid for the pulse waves as far as they are subjected to the "diffractive" effect of the media.

On the other hand, when deriving from the equation of two-frequency mutual coherence function, it is rather easy to take account of the additional "refractive" effect[17] (although it is usually very small or, in the case of pulsars, is not actually observed[1,5]) and to find the corresponding term to be added to the left-hand side of (1.15), but it is not the case when deriving from the transport equation and, to find the refractive term, the next order correction to the transport Eq. (1.2) is necessary to contain the second-order terms with respect to $\partial/\partial \bar{p}$.

2. PLANE WAVE PULSE

In case of plane wave pulses, the solution of (1.15) becomes independent of the coordinates $\boldsymbol{\rho}$ and may be represented by an integral of the form

$$I(\mathbf{r}';\zeta,t) = (2\pi)^{-1} \int_{-\infty}^{\infty} du e^{iu\zeta} \tilde{I}(\mathbf{r}';u,t), \tag{2.1}$$

where, by (1.15) and (1.16), $\tilde{I}(\mathbf{r}';u,t)$ is a solution of

$$\left[c^{-1}\frac{\partial}{\partial t} - \frac{1}{2}iu\left(k^{-1}\frac{\partial}{\partial \mathbf{r}'}\right)^2 + k\beta\,|k\mathbf{r}'|^\alpha\right]\tilde{I}(\mathbf{r}';u,t) = 0. \tag{2.2}$$

Equation (2.2) can be expressed in the simple form

$$\left[\frac{\partial}{\partial \eta} + H\right]\tilde{I}(\boldsymbol{\xi},\eta) = 0, \tag{2.3}$$

in terms of the new coordinates $\boldsymbol{\xi}$, η and the operator H defined by

$$\boldsymbol{\xi} = (2k\beta/iu)^{1/(\alpha+2)}k\mathbf{r}', \quad \eta = (iu/2k\beta)^{\alpha/(\alpha+2)}2k\beta ct,$$
$$|\arg(\boldsymbol{\xi})| < \pi 2^{-1}(\alpha+2)^{-1}, \quad |\arg(\eta)| < \pi 2^{-1}\alpha(\alpha+2)^{-1}, \quad 2 \geqslant \alpha > 1, \tag{2.4}$$

$$H = \frac{1}{2}\left[-\left(\frac{\partial}{\partial \boldsymbol{\xi}}\right)^2 + |\boldsymbol{\xi}^2|^{\alpha/2}\right]. \tag{2.5}$$

Although the new coordinates $\boldsymbol{\xi}$ are complex according to (2.4), an effective way to solve (2.3) is to construct the eigenfunctions of H, which are defined by the eigenvalue equation

$$H\psi_A(\boldsymbol{\xi}) = A\psi_A(\boldsymbol{\xi}). \tag{2.6}$$

Here, when $\boldsymbol{\xi}$ are real, the operator H is Hermitian and self-adjoint, giving the real eigenvalues A and the orthogonal eigenfunctions $\psi_A(\boldsymbol{\xi})$, which may be normalized according to

$$\int d\boldsymbol{\xi}\psi_A(\boldsymbol{\xi})\psi_B(\boldsymbol{\xi}) = \delta_{AB}. \tag{2.7}$$

All the above properties of the eigenvalues and eigenfunctions remain to be valid even when $\boldsymbol{\xi}$ are changed from real to complex coordinates, in so far as the eigenfunctions are analytically continued. Here, when $\alpha = 2$, H is exactly the same as the Hamiltonian of two-dimensional linear quantum oscillator and therefore $A = 2n + 1$, $n = 0,1,2,\ldots$; also $\psi_A(\boldsymbol{\xi})$ can be obtained exactly and be shown to tend to zero as $|\boldsymbol{\xi}| \to \infty$ for the complex $\boldsymbol{\xi}$ of (2.4). This is also the case for other cases of $\alpha < 2$, as will be seen later.

In terms of the eigenvalues and eigenfunctions, the solution of (2.3) for the initial values $\tilde{I}(\boldsymbol{\xi},0)$ is exhibited by

$$\tilde{I}(\boldsymbol{\xi},\eta) = \sum_A \exp(-A\eta)\psi_A(\boldsymbol{\xi}) \int d\boldsymbol{\xi}'\, \psi_A(\boldsymbol{\xi}')\tilde{I}(\boldsymbol{\xi}',0). \tag{2.8}$$

Here, from the orthogonality and completeness of the eigenfunctions,

$$\sum_A \psi_A(\boldsymbol{\xi})\psi_A(\boldsymbol{\xi}') = \delta(\boldsymbol{\xi} - \boldsymbol{\xi}'). \tag{2.9}$$

A. Pulse wave for a perfectly coherent initial wave

When the initial wave at $t = 0$ is a perfectly coherent plane wave pulse and has the form of $\delta(z)$, given by

$$I(\boldsymbol{\Omega},\omega;\boldsymbol{\rho},z,t)|_{t=0} = \delta(\boldsymbol{\Omega})\delta(\omega - \omega_0)\delta(z), \tag{2.10a}$$

or, according to (1.11), (2.1), and (2.4), by

$$I(\mathbf{r},\tau;\boldsymbol{\rho},z,t)|_{t=0} = \delta(z), \quad \tilde{I}(\boldsymbol{\xi},0) = 1, \tag{2.10b}$$

then, from (2.1) and (2.8), the total irradiance

$$I(\zeta,t) \equiv I(\mathbf{r}';\zeta,t)|_{\mathbf{r}'=0} = \int_{-\infty}^{\infty} d\omega \int d\boldsymbol{\Omega} I(\boldsymbol{\Omega},\omega;z,t) \tag{2.11a}$$

is exhibited by

$$I(\zeta,t) = \sum_A C_A(2\pi)^{-1}\int_{-\infty}^{\infty} du \exp(-A\eta + iu\zeta),$$
$$C_A = \psi_A(0)\int d\boldsymbol{\xi}\psi_A(\boldsymbol{\xi}), \quad \eta = (iu/2k\beta)^{\alpha/(\alpha+2)}2k\beta ct. \tag{2.11b}$$

Here, η is a function of u given by (2.4) and, in order that the condition in (2.4) is fulfillled over the whole range of integration, the branch cut is taken from the origin to $+i\infty$ along the positive imaginary axis of u and also the path of integration along the infinitesimally lower side of the real axis. Thus, we find that $I(\zeta,t) = 0$ for $\zeta < 0$ by deforming the path of integration into the infinite lower semicircle.

In terms of new variables, defined by

$$u\zeta = \Delta\omega(t - z/c), \quad \Delta\omega = uc, \tag{2.12a}$$

$$\eta = (i\Delta\omega/\omega_c)^{\alpha/(\alpha+2)}, \quad \omega_c = t^{-1}(2\beta\omega_0 t)^{-2/\alpha},$$

(2.11b) is also expressed in the form

$$I(\zeta,t) = \frac{1}{2\pi c}\int_{-\infty}^{\infty} d\Delta\omega \Big(\sum_A C_A \exp[-A \times (i\Delta\omega/\omega_c)^{\alpha/(\alpha+2)}] \exp[i\Delta\omega(t - z/c)]\Big), \tag{2.12b}$$

where A and C_A are to be given by the following (2.22) and (2.24). Thus, we find that the factor ( ) in the integrand of (2.12b) provides the two-frequency mutual coherence function of waves for the frequency difference $\Delta\omega$. In Fig. 1 is shown the normalized two-frequency mutual coherence function, given by the real part of the integrand of (2.12b), as a function of $\Delta\omega/\omega_c$ for $\alpha = 2$ and $\frac{5}{3}$.

In order to evaluate the integral in (2.11b) for $\zeta > 0$, we define a new variable of integration, according to $a = -iu\zeta$ with

$$\eta = (-a)^{\alpha/(\alpha+2)}\phi, \quad \phi = (2k\beta ct)(2k\beta\zeta)^{-\alpha/(\alpha+2)}, \tag{2.13}$$

whence

$$I(\zeta,t) = (2\pi i\zeta)^{-1}\sum_A C_A \int_{-i\infty-\delta}^{i\infty-\delta} da \exp[-a - (-a)^{\alpha/(\alpha+2)}A\phi], \quad \delta = +0, \tag{2.14}$$

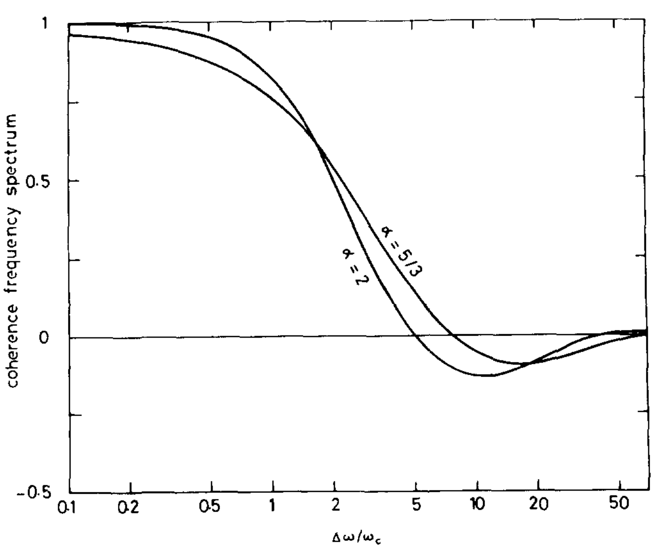


FIG. 1. The normalized frequency spectrum of pulse wave given by (2.12b) for Gaussian ($\alpha = 2$) and Kolmogorov ($\alpha = \frac{5}{3}$) spectra of turbulence.

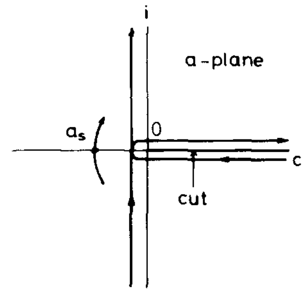


FIG. 2. Path of integration for (2.14).

where $\arg(-a) = 0$ for $a < 0$. The above path of integration may be deformed to the contour path c which starts at $a = +\infty$ and goes to the origin along the infinitesimally lower side of the positive real axis; it then encircles the origin in the closewise and returns to the starting point along the infinitesimally upper side of the real axis (Fig. 2). Thus, after the expansion of the integrand in the power series of $A\phi$, the use of the Hankel contour integral

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i}\int_c da(-a)^{-z}e^{-a} \tag{2.15}$$

brings (2.14) into the form

$$I(\zeta,t) = \sum_A I_A(\zeta,t), \quad \zeta > 0, \tag{2.16}$$

where

$$I_A(\zeta,t) = \zeta^{-1}C_A \sum_{m=0}^{\infty}\{m!\Gamma[-m\alpha/(\alpha+2)]\}^{-1}(-A\phi)^m$$

$$= (\pi\zeta)^{-1}C_A \sum_{m=1}^{\infty}(m!)^{-1}\Gamma[1 + \alpha(\alpha+2)^{-1}m] \times \sin[2\pi m(\alpha+2)^{-1}](A\phi)^m. \tag{2.17}$$

In the special case of $\alpha = 2$, (2.17) gives

$$I_A(\zeta,t) = \zeta^{-1}C_A\pi^{-1/2}(A\phi/2)\exp[-(A\phi/2)^2], \quad \alpha = 2, \tag{2.18}$$

where the coefficients C_A, defined by (2.11b), are found to be $(-)^n 2$ for $A = 2n+1, n = 0,1,2,\ldots$, respectively [see (2.24)].

According to (2.13), $\phi \to \infty$ as $\zeta \to 0$, and the convergence of the series (2.17) becomes poor when $A\phi \gg 1$. The asymptotic expression in this case can be obtained by using the saddle point approximation in the integral in (2.14), where the saddle point a_s exists on the negative real axis at $a_s = -[\alpha(\alpha+2)^{-1}A\phi]^{1+\alpha/2}$ (Fig. 2). Thus,

$$I_A(\zeta,t) \sim C_A(2\zeta)^{-1}[(\alpha+2)/\pi]^{1/2}[\alpha(\alpha+2)^{-1}A\phi]^{(1+\alpha/2)/2} \times \exp\{-(2/\alpha)[\alpha(\alpha+2)^{-1}A\phi]^{1+\alpha/2}\}, \quad A\phi \gg 1. \tag{2.19}$$

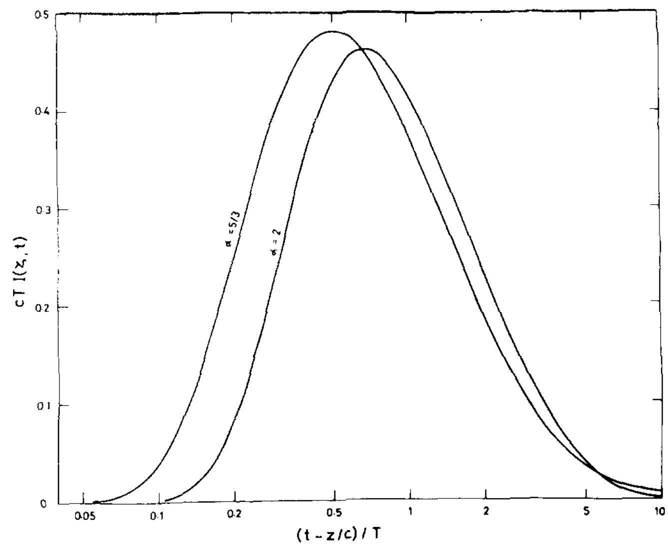


FIG. 3. The total intensity $I(\zeta,t)$ given by (2.16) as a function of $(t-z/c)/T$ for $\alpha=2$ and 5/3.

Note that (2.19) agrees with the exact expression (2.18) when $\alpha=2$.

When $\zeta=ct-z$ is small enough to satisfy the condition $A\phi \gg 1$ for all the A's, then $I(\zeta,t)$ is actually determined by the first term of the series (2.16) with the asymptotic expression (2.19) and, even when $\phi\sim 1$, several terms of the series are good enough for practical computation (Fig. 3).

On the other hand, since the eigenfunctions $\psi_A(\xi)$ are symmetrical around the z axis of wave propagation in the present case, the eigenvalue equation (2.6) with (2.5) is expressed, in the cylindrical coordinate system, by

$$\left(\xi^{-1}\frac{\partial}{\partial\xi}\xi\frac{\partial}{\partial\xi}+2A-\xi^{\alpha}\right)\psi_A(\xi)=0,\quad \xi=|\boldsymbol{\xi}|. \tag{2.20}$$

Here, the asymptotic expression of A for large eigenvalues can be obtained from the quantum condition [refer to (B10) with (B2)], and the result is that the A's are the roots of

$$\int_0^{(2A)^{1/\alpha}} d\xi\,(2A-\xi^{\alpha})^{1/2}\sim\pi(n+1/2),\quad n=0,1,2,\cdots, \tag{2.21}$$

which, on using the new variable of integration $x=\xi^{\alpha}/2A$, gives

$$A\sim(1/2)[\pi(n+1/2)\alpha/B(1/\alpha,3/2)]^{2\alpha/(\alpha+2)},\qquad n=0,1,2,\cdots, \tag{2.22}$$

where $B(\nu,\mu)$ is the Beta function defined by

$$B(\nu,\mu)=\int_0^1 dx\,x^{\nu-1}(1-x)^{\mu-1}=\frac{\Gamma(\nu)\Gamma(\mu)}{\Gamma(\nu+\mu)}. \tag{2.23}$$

Note that, when $\alpha=2$, Eq. (2.22) gives the exact eigenvalues $A=2n+1$, in spite of the fact that (2.22) is valid only in the asymptotic sense.

The corresponding expression of the coefficients C_A, defined by (2.11b), is found to be [see (B15)]

$$C_A\sim(-)^n(2\alpha)^{1/2}\pi/B(1/\alpha,1/2),\quad n=0,1,2,\cdots, \tag{2.24}$$

which again gives the exact expression $C_A=(-)^n 2$ when $\alpha=2$. Thus, we may expect that both (2.22) and (2.24) will be highly reliable even when $\alpha=5/3<2$. In Table I, the first several eigenvalues are shown according to (2.22).

In Fig. 3, the values of $I(\zeta,t)$ are shown as a function of $\zeta/c=t-z/c$ for $\alpha=2$ and $\frac{5}{3}$. Here, use is made of (2.16), (2.17), and (2.19), and the constant T is defined, according to (2.13), by

$$\phi/2=[T/(t-z/c)]^{\alpha/(\alpha+2)},$$
$$cT=\tfrac{1}{2}ct\,(k\beta ct)^{2/\alpha}=\begin{cases}0.79(ct)^{11/5}k^{2/5}\bar{C}_n^{\,12/5}, & \alpha=\frac{5}{3},\\ 0.83(ct)^2 l_0^{-1/3}\bar{C}_n^2, & \alpha=2.\end{cases} \tag{2.25}$$

These curves show a very good agreement with those shown by the previous authors who solved (2.2) by a numerical method[5] for $\alpha=5/3$ and by an analytical method[9] for $\alpha=2$. In the latter case, $I(\zeta,t)$ was expressed by a residual series whose convergence is good when $\zeta\sim\infty$ or $\phi\ll 1$, whereas the convergence is poor when $\zeta\sim 0$ or $\phi\gg 1$ [refer to (3.15)]. Therefore, the situation is just inverse to that of the present series (2.16) with (2.18), whose convergence is good when $\phi\gg 1$, and indeed these two expressions are found to be two different expressions of the same analytical function (Sec. 3).

B. Pulse wave for a perfectly incoherent initial wave

So far we have considered the pulse waves when the initial wave at $t=0$ is the coherent plane wave having the form of $\delta(z)$, as given by (2.10). Another extreme case is where the initial wave is a perfectly incoherent pulse with the same factor $\delta(z)$, i.e.,

$$I(\boldsymbol{\Omega},\omega;\boldsymbol{\rho},z,t)|_{t=0}=(2\pi)^{-2}\delta(\omega-\omega_0)\delta(z), \tag{2.26}$$

or, by (1.11),

$$I(\mathbf{r},\tau;\boldsymbol{\rho},z,t)|_{t=0}=\delta(k\mathbf{r})\delta(z), \tag{2.27}$$

which, in the coordinate system $\boldsymbol{\xi}$ of (2.4), is expressed formally by

$$\tilde{I}(\boldsymbol{\xi},0)=(2k\beta/iu)^{2/(\alpha+2)}\delta(\boldsymbol{\xi})\delta(z),$$
$$\delta(\boldsymbol{\xi})=(2\pi\xi)^{-1}\delta(\xi-0). \tag{2.28}$$

For the initial condition (2.28), the use of (2.8) and the procedure leading to the expressions (2.14) and (2.17) yield [with the superscript i attached to $I(\zeta,t)$]

TABLE I. Asymptotic values of eigenvalues given by (2.22).

α \ η	0	1	2	3	4	5	6
5/3	0.971	2.637	4.195	5.697	7.159	8.591	10.000
2	1	3	5	7	9	11	13

$$I^i(\zeta,t) = (2\pi ct)^{-1}\phi \sum_A \psi_A^2(0)$$

$$\times (2\pi i)^{-1} \int_{-i\infty-\delta}^{i\infty-\delta} da(-a)^{-2/(\alpha+2)}$$

$$\times \exp[-a-(-a)^{\alpha/(\alpha+2)}A\phi] \quad (2.29)$$

$$= \sum_A \psi_A^2(0)(2\pi ct)^{-1} \sum_{m=0}^{\infty} \frac{\phi(-A\phi)^m}{m!\Gamma[(2-m\alpha)/(\alpha+2)]},$$

which may be given in the form

$$I^i(\zeta,t) = \sum_A I_A^i(\zeta,t), \quad (2.30)$$

where, with the aid of (B17),

$$I_A^i(\zeta,t)$$
$$= C_A^i(2\pi ct)^{-1}\pi^{-1} \sum_{m=0}^{\infty} (m!)^{-1} \sin[2\pi(m+1)/(\alpha+2)]$$

$$\times \Gamma[1+(m\alpha-2)/(\alpha+2)](A\phi)^{m+1}, \quad (2.31)$$

$$C_A^i = A^{-1}\psi_A^2(0) \sim 2\alpha(2A)^{-1/2-1/\alpha}\pi/B(1/\alpha,1/2). \quad (2.32)$$

Here, the asymptotic expression of $I_A^i(\zeta,t)$ can be obtained, as in (2.19), by using the saddle-point approximation in the integral in (2.29) and it is found to be

$$I_A^i(\zeta,t) \sim C_A^i(2\pi ct)^{-1}(1/2+1/\alpha)[(\alpha+2)/\pi]^{1/2}$$
$$\times [\alpha(\alpha+2)^{-1}A\phi]^{(1+\alpha/2)/2}$$
$$\times \exp\{-(2/\alpha)[\alpha(\alpha+2)^{-1}A\phi]^{1+\alpha/2}\}, \quad A\phi \gg 1. \quad (2.33)$$

It will be noted that, according to (2.24), C_A has the factor $(-)^n$ whereas C_A^i has not.

When $\alpha = 2$, we obtain $C_A^i = 2/A$ and

$$I_A^i(\zeta,t) \sim C_A^i(2\pi ct)^{-1}\pi^{-1/2}A\phi \exp[-(A\phi/2)^2], \quad \alpha = 2. \quad (2.34)$$

In Fig. 4 are shown the curves of $(2\pi ct)I^i(\zeta,t)$ as a function of the same variables as in Fig. 3. These curves show that $I^i(\zeta,t)$ tends to $(2\pi ct)^{-1}$ as $\zeta \to \infty$ and make a notable difference from the corresponding curves in Fig. 3. Here, since, by (2.13), $\phi \to 0$ for $\beta \to 0$ as will as for $\zeta \to \infty$, the asymptotic expression of $I^i(\zeta,t)$ should agree with the solution of the original equation (1.12) or (1.15) in free space for the same initial wave condition (2.27) and, indeed, the latter solution is found to be $(2\pi ct)^{-1}$ for $\zeta = ct - z > 0$ and 0 for $\zeta < 0$. Thus, the rather peculiar behavior of $I^i(\zeta,t)$ may be understood to be reflecting the physical situation that, at each point on the plane of initial pulse wave at $z = 0$, there is always such a component of wave which can directly reach the point of observation without bending the path of wave propagation, whereas this is not the case when the pulse waves are coherently initiated by (2.10a).

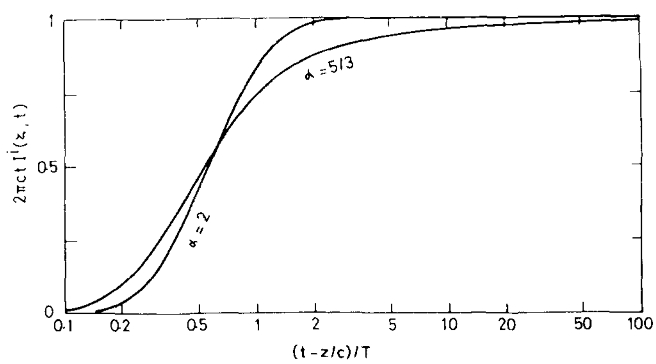


FIG. 4. The total intensity $I^i(\zeta,t)$ given by (2.30) as a function of $(t - z/c)/T$ for $\alpha = 2$ and $\alpha = \frac{5}{3}$.

3. ASYMPTOTIC EXPRESSIONS OF PULSE WAVES FOR LARGE VALUES OF $\zeta = ct - z$

The convergence of the series (2.16) and (2.30) for the total irradiance $I(\zeta,t)$ is very good when $\phi \gg 1$ or $\zeta \sim 0$, but becomes poor when $\phi \ll 1$ or $\zeta \sim \infty$. In order to obtain the various asymptotic expressions in the latter case, we refer to (A24) with (A23) to apply to $I(\xi,\eta)$ in (2.8). Thus, in the case of (2.10) where the initial wave is coherently excited, we obtain

$$\tilde{I}(\xi,\eta) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} d\lambda\, e^{-\lambda\eta} \int_0^{\infty} d\xi'\xi' R_\lambda(\xi\,|\,\xi'). \quad (3.1)$$

Here, $R_\lambda(\xi_1|\xi_2)$ is given by (A15) and, if we are interested only in the asymptotic expressions for $\lambda \gg 1$, $f_\lambda^0(\xi)$ and $f_\lambda^\infty(\xi)$ are given by (B2) and (B7), respectively, and the bracket $[f_\lambda^\infty f_\lambda^0]$ by (B9). Therefore, $f_\lambda^0(0) = 1$, and (B14) with A replaced by λ, may be used to evaluate the last integral in (3.1). Thus, using (A15) in (3.1) gives

$$\tilde{I}(0,\eta) = \pi(2/\alpha)^{1/2}\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} d\lambda\, e^{-\lambda\eta}(2\lambda)^{1/\nu-1}\sec\varphi, \quad (3.2)$$

$$\varphi = \alpha^{-1}B(1/\alpha,3/2)(2\lambda)^{1/\nu},$$

$$\nu = 2\alpha/(\alpha+2) < 1, \quad |\arg(\lambda)| < \pi/2,$$

$$I(\zeta,t) = (2\pi)^{-1}\int_{-\infty}^{\infty} du\, e^{iu\zeta}\tilde{I}(0,\eta). \quad (3.3)$$

In order to evaluate the integral (3.3) with (3.2), we define new variables of integration, according to

$$a = -iu\zeta, \quad b = \lambda\eta, \quad \eta = (-a)^{\nu/2}\phi, \quad (3.4)$$

with a and ϕ defined by (2.13), and hence obtain

$$I(\zeta,t)$$
$$= (2\alpha)^{-1/2}(2/\phi)^{1/\nu}(2\pi i\zeta)^{-1}\int_{-i\infty-\delta}^{i\infty-\delta} da(-a)^{-1/2}e^{-a}$$

$$\times (2i)^{-1}\int_{-i\infty-\delta}^{i\infty-\delta} db\, b^{1/\nu-1}e^{-b}\sec\varphi, \quad \delta = +0, \quad (3.5)$$

with

$$\varphi = (2/\phi)^{1/\nu}\alpha^{-1}B(1/\alpha,3/2)(-a)^{-1/2}b^{1/\nu},$$

$$|\arg(-a)| < \pi/2, \quad 1 \geqslant \nu > \tfrac{2}{3}. \quad (3.6)$$

Here, $\mathrm{Im}[\varphi] \gtrless 0$ depending on $\arg(b) = \pm \pi/2$, respectively, and $\sec(\varphi)$, involved in the last integrand of (3.5), can be expanded in the series

$$\sec(\varphi) = 2 \sum_{n=0}^{\infty} (-)^n e^{\pm i(2n+1)\varphi}, \quad \mathrm{Im}[\varphi] \gtrless 0. \tag{3.7}$$

To evaluate the integral (3.5) with the expansion (3.7), we first integrate with respect to the variable a and thus have the integrals of the form

$$g(b, \pm ic_n) = \frac{1}{2\pi i} \int_{-i\infty-\delta}^{i\infty-\delta} da\,(-a)^{-1/2} \times \exp[-a \pm ic_n(-a)^{-1/2} b^{1/\nu}], \tag{3.8}$$

$$c_n = (2n+1)(2/\phi)^{1/\nu}\alpha^{-1}B(1/\alpha, 3/2), \quad n = 0,1,2,\cdots. \tag{3.9}$$

Here, with the aid of formula (2.15), (3.8) is expanded in the power series of c_n, by

$$g(b, \pm ic_n) = \sum_{m=0}^{\infty} \frac{(\pm ic_n)^m}{m!\Gamma[(m+1)/2]} b^{m/\nu}, \quad \arg(b) = \pm \pi/2. \tag{3.10}$$

Thus, in terms of the function

$$F(c_n) = (2i)^{-1}\left[\int_{i0}^{i\infty} db\, g(b, +ic_n) + \int_{-i\infty}^{-i0} db\, g(b, -ic_n)\right] b^{1/\nu - 1} e^{-b} = \mathrm{Im}\left[\int_0^{\infty} db\, g(b, +ic_n) b^{1/\nu-1} e^{-b}\right], \tag{3.11}$$

(3.5) is exhibited by

$$I(\zeta, t) = (2/\alpha)^{1/2}(2/\phi)^{1/\nu}\zeta^{-1} \sum_{n=0}^{\infty} (-)^n F(c_n). \tag{3.12}$$

Here, on using (3.10), we obtain from (3.11)

$$F(c_n) = \mathrm{Im}\left(\sum_{m=0}^{\infty} \frac{\Gamma[(m+1)/\nu]}{m!\Gamma[(m+1)/2]} e^{im\pi/2} c_n^m\right) = \sum_{l=0}^{\infty} \frac{(-)^l \Gamma[2(l+1)/\nu]}{l!(2l+1)!} c_n^{2l+1}. \tag{3.13}$$

When $\alpha = 2$, (3.13) gives

$$F(c_n) = c_n \exp(-c_n^2), \tag{3.14}$$

and hence, from (3.12) with ϕ given by (2.25), we obtain

$$I(\zeta, t) = 2(\zeta\phi)^{-1} \sum_{n=0}^{\infty} (-)^n c_n \exp(-c_n^2), \quad \alpha = 2,$$
$$c_n = \pi(2n+1)(2\phi)^{-1}, \quad \phi = 2[T/(t - z/c)]^{1/2}, \tag{3.15}$$

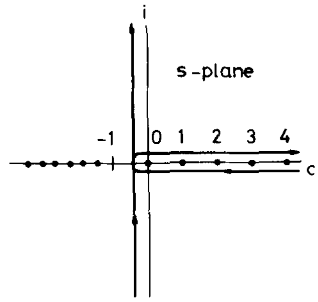


FIG. 5. Path of integration for (3.17).

which gives the exact expression obtained by Sreenivasiah *et al.*[9] using a residual method. Note that the series (3.15) has the same form as the previous series (2.16) with (2.18) where $A = 2n + 1$ and $C_A = (-)^n 2$; the only difference is that ϕ and ϕ^{-1} are interchanged.

Here, as $\zeta \to \infty$, $\phi \to 0$, and hence also $c_n \to \infty$ by (3.9). Therefore, when actually using the expression (3.12), we need the asymptotic expression of $F(x)$ for large values of x and, to this end, first begin with its integral representation suitable for this purpose. In reference to

$$\mathrm{Res}[\Gamma(-s)]_{s=l} = (-)^{l+1}/l!, \quad l = 0,1,2,\cdots, \tag{3.16}$$

it is straightforward to bring the series (3.13) into the integral representation of Barnes-type:

$$F(x) = \frac{1}{2\pi i} \int_{-i\infty-\delta}^{i\infty-\delta} ds \frac{\Gamma(-s)\Gamma[2(s+1)/\nu]}{\Gamma[2(s+1)]} x^{2s+1}, \quad 1 \geqslant \nu > \tfrac{2}{3}, \tag{3.17}$$

which gives the original series in terms of the residual series due to the poles of the factor $\Gamma(-s)$ (Fig. 5).

Here, when $\nu < 1$ or $\alpha < 2$, the integrand of (3.17) has the poles also due to the factor $\Gamma[2(s+1)/\nu]$ at $s = -m\nu/2 - 1$, $m = 1,2,\cdots$, on the negative real axis of s and

$$\mathrm{Res}\{\Gamma[2(s+1)/\nu]\}_{s=-m\nu/2-1} = (\nu/2)(-)^m/m!. \tag{3.18}$$

Therefore, on shifting the path of integration to the left of the imaginary axis and using the residue values (3.18), we obtain the asymptotic expression

$$F(x) \sim (2\pi)^{-1}\nu \sum_{m=1}^{\infty} (m!)^{-1}\Gamma(1 + m\nu/2)\Gamma(1 + m\nu) \times \sin[\pi(1-\nu)m]\, x^{-m\nu-1}, \quad x \gg 1, \quad \nu < 1, \tag{3.19}$$

whereas the exact expression (3.14) is available when $\nu = 1$ or $\alpha = 2$.

Here, when $\phi \ll 1$ and therefore $c_n \gg 1$ by (3.9), the asymptotic expression (3.19) is available for all the terms in the series (3.12). Thus, in terms of the function defined by

$$\beta(x) = \sum_{n=0}^{\infty} (-)^n (2n+1)^{-x}, \tag{3.20}$$

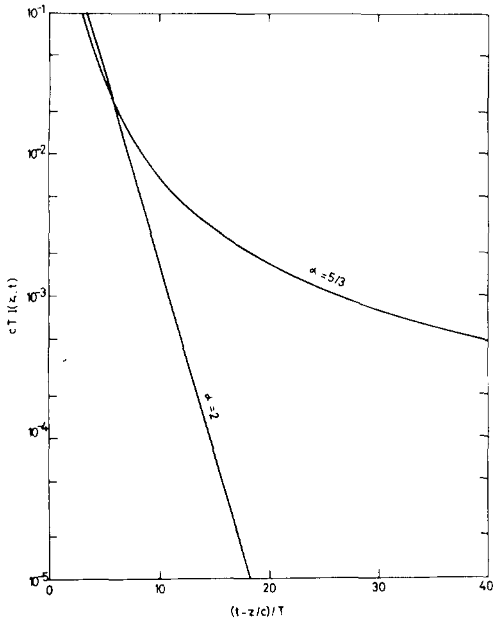


FIG. 6. The tail part of the pulse intensity $I(\zeta,t)$ according to (3.21).

we find that, when $\phi \ll 1$ and $\nu < 1$,

$$I(\zeta,t) \sim \pi^{-2} |C_A| \zeta^{-1} \sum_{m=1}^{\infty} (m!)^{-1} \beta(1+m\nu) \Gamma(1+m\nu/2)$$
$$\times \Gamma(1+m\nu) \sin[\pi(1-\nu)m] \{(\phi/2)$$
$$\times [\alpha/B(1/\alpha,3/2)]^{\nu}\}^{m}, \qquad (3.21)$$

where C_A is given by (2.24) with the relation

$$|C_A| = \pi\nu(\alpha/2)^{1/2}/B(1/\alpha,3/2) = \pi(2\alpha)^{1/2}/B(1/\alpha,1/2). \qquad (3.22)$$

It will be noted that $\phi \propto \zeta^{-\nu/2}$ by (2.13) and hence, as $\zeta \to \infty$, $I(\zeta,t) \sim \zeta^{-\nu/2-1} = \zeta^{-16/11}$ for $\alpha = \frac{5}{3}$, and this makes a notable difference from the asymptotic form (3.15) for $\alpha = 2$, which exponentially decreases with ζ as $\zeta \to \infty$ (Fig. 6).[19]

The asymptotic expression for the pulse wave $I^i(\zeta,t)$ for the incoherent excitation can be obtained in the same way and, on referring to (2.28), Eq. (3.1) is replaced by

$$\tilde{I}(\xi,\eta) = (2k\beta/iu)^{2/(\alpha+2)} [(2\pi)^2 i]^{-1}$$
$$\times \int_{-i\infty}^{i\infty} d\lambda e^{-\lambda\eta} R_\lambda(\xi|0), \qquad (3.23)$$
$$R_\lambda(0|0) \sim \pi \tan\varphi,$$

where the last expression has been derived from (A15) with (B2), (B5), and (B8). The result is

$$I^i(\zeta,t) = (2\pi ct)^{-1}\left[1 + 2\sum_{n=1}^{\infty}(-)^n F^i(c_n^i)\right],$$
$$c_n^i = 2n(2/\phi)^{1/\nu}\alpha^{-1}B(1/\alpha,3/2), \qquad (3.24)$$
$$F^i(x) = \sum_{l=0}^{\infty}(-)^l \frac{\Gamma(1+2l/\nu)}{l!(2l)!}x^{2l},$$

which corresponds to (3.12) and (3.13) with (3.9).

Here, when $\alpha = 2$ or $\nu = 1$, (3.24) gives

$$c_n^i = \pi n/\phi, \quad F^i(x) = \exp(-x^2), \qquad (3.25a)$$

whereas, when $\alpha < 2$, $F(x)$ has the asymptotic expression

$$F^i(x) \sim \pi^{-1}\sum_{m=0}^{\infty}(m!)^{-1}\Gamma[(m+1)\nu/2]\Gamma[(m+1)\nu]$$
$$\times \sin[\pi(1-\nu)(m+1)]\, x^{-(m+1)\nu}, \quad \nu < 1. \qquad (3.25b)$$

Thus, $2\pi ct I^i(\zeta,t) \to 1$ as $\zeta \to \infty$ or $c_n \to \infty$ (Fig. 4).

4. PULSE MOMENTS

Since, by (2.13) and (3.9), $\phi \propto \zeta^{-\nu/2}$ and $c_n \propto \zeta^{1/2}$ and therefore $c_n^{2s+1} \propto \zeta^{s+1/2}$, we find, on using (3.12) with the integral representation (3.17), that, for any $p \geqslant 0$,

$$\int_0^{\zeta} d\zeta' \zeta'^p I(\zeta',t) = (2/\alpha)^{1/2}(2/\phi)^{1/\nu}\zeta^p \sum_{n=0}^{\infty}(-)^n F^p(c_n), \qquad (4.1)$$

where

$$F^p(x) = \frac{1}{2\pi i}\int_{-i\infty-\delta}^{i\infty-\delta} ds \frac{\Gamma(-s)\Gamma[2(s+1)/\nu]}{(s+p+1)\Gamma[2(s+1)]} x^{2s+1},$$
$$p \geqslant 0. \qquad (4.2)$$

Here, as in (3.17), the path of integration in (4.2) can be shifted to the left side to obtain the asymptotic expression of $F^p(x)$ and, then, the contribution from the pole at $s = -p-1$, say $F_0^p(x)$, is found to be

$$F_0^p(x) = \frac{\Gamma(1+p)\,\Gamma(1+2p)\sin(2\pi p)}{\Gamma(1+2p/\nu)\sin(2\pi p/\nu)} x^{-2p-1},$$

which contribution to the nth term on the right-hand side of (4.1) is therefore proportional to

$$(2/\phi)^{1/\nu}\zeta^p c_n^{-2p-1} = (cT)^p[\alpha/B(1/\alpha,3/2)]^{2p+1}$$
$$\times (2n+1)^{-2p-1}, \qquad (4.4)$$

where use has been made of (2.25) and (3.9). Note that (4.4) is independent of ζ. The total contribution from the terms (4.4) to the right-hand side of (4.1), say $\langle \zeta^p \rangle$, is thus given by

$$\langle \zeta^p \rangle/(cT)^p = (2/\alpha)^{1/2}[\alpha/B(1/\alpha,3/2)]^{2p+1}\beta(2p+1)$$
$$\times \frac{\Gamma(1+p)\Gamma(1+2p)\sin(2\pi p)}{\Gamma(1+2p/\nu)\sin(2\pi p/\nu)}, \qquad (4.5)$$

where $\beta(x)$ is defined by (3.20).

When $\nu < 1$, the integrand in (4.2) has the additional poles due to the factor $\Gamma[2(s+1)/\nu]$ at $s = -m\nu/2 - 1$, $m = 1,2,\cdots$, on the negative real axis. Here, in order that the integral (4.1) tends to a finite value as $\zeta \to \infty$, the contributions from all these poles to the right-hand side of (4.1) are to vanish as $\zeta \to \infty$ since, by (4.4), the contribution from the pole at $s = -p-1$ is independent of ζ. Thus, the integral

(4.1) converges as $\zeta\to\infty$ when $p<\nu/2$ (where $p=\nu/2$ is the special case in which the first pole at $s=-\nu/2-1$ coincides with the pole at $s=-p-1$ to make a double pole and yields a term of $\log\zeta$), and we find

$$\int_0^\infty d\zeta\,\zeta^p I(\zeta,t)=\langle\zeta^p\rangle,\quad p<\nu/2, \tag{4.6}$$

where the condition is $p<\frac{5}{11}$ when $\alpha=\frac{5}{3}$.

On the other hand, when $\alpha=2$, the integrand in (4.2) has the pole only at $s=-p-1$ on the negative real axis and therefore (4.5) is valid for all values of $p>-1$. Thus

$$\langle\zeta^p\rangle/(cT)^p=(4/\pi)^{2p+1}\beta(2p+1)\Gamma(1+p),$$
$$\alpha=2,\quad p>-1, \tag{4.7a}$$

which gives

$$\langle\zeta\rangle=2cT,\quad \langle(\zeta-\langle\zeta\rangle)^2\rangle=2.666(cT)^2. \tag{4.7b}$$

In the special case of $p=0$, (4.5) with (3.22) gives

$$\langle 1\rangle=\beta(1)(2\alpha)^{1/2}\nu/B(1/\alpha,3/2)=(2/\pi)\beta(1)|C_A|$$
$$=|C_A/2|,\quad \beta(1)=\pi/4, \tag{4.8}$$

while $\langle 1\rangle$ should always be equal to 1, as may directly be seen from (2.1) with (2.2). Here, according to (4.8), $\langle 1\rangle=1$ for $\alpha=2$, whereas $\langle 1\rangle=1.034$ for $\alpha=5/3$, and therefore the 3.4% discrepancy in the latter case will represent an overall error by the present method based on the asymptotic values of eigenvalue (2.22) and the corresponding eigenfunctions.

5. SUMMARY

An analytical theory of pulse wave propagation in turbulent media, based on the conventional space–time transport equation, was developed and was found, after the forward-scattering approximation was made, to be equivalent to the corresponding theory based on the equation of two-frequency mutual coherence function which has been exclusively used by previous authors. The above conclusion is valid as long as the pulse waves are subjected only to the most important "diffractive" effect of the media, whereas, in order to take the supplemental "refractive" effect also into account, the transport equation needs to be modified to include the higher order terms. The pulse wave intensities were then exhibited in terms of the eigenfunction series for both coherent and incoherent initial waves (Sec. 2), and the associated resolvent function was fully used to find the eigenvalues and eigenfunctions (Appendices A and B). The quantum condition, based on the WKB approximation, was used to obtain the explicit expresson of the eigenvalues as a function of the medium parameter α to apply to the media of both Gaussian ($\alpha=2$) and Kolmogorov ($\alpha=\frac{5}{3}$) turbulence spectra. Here, in case of the Gaussian spectrum, the expression gives the exact eigenvalues and also the exact eigenfunction series for the averaged pulse wave intensities. The convergence of the series is very good in the beginning part of pulse wave, whereas it becomes poor in the tail part. Another expression of the pulse wave intensities was obtained in the series whose convergence is good in the tail part of the pulse wave [Sec. 3]. In case of the Gaussian turbulence medium, the latter series precisely agrees with the previous result obtained by Sreenivasiah *et al.*[9] The pulse wave has a long tail in case of the Kolmogorov spectrum, whereas it has a short and exponentially decreasing tail in case of the Gaussian spectrum. In connection with the pulse width, the pulse moments were evaluated in detail (Sec. 4). When the media are composed of random discrete scatterers such as rains and clouds, the eigenfunction expansion in Sec. 2 needs to be modified to accommodate additional eigenfunctions of continuous spectrum (Appendix A).

APPENDIX A: RESOLVENT FUNCTION OF H AND THE ASSOCIATED EIGENFUNCTIONS

An effective way of constructing the eigenfunctions of (2.6) is to use the resolvent function of H which, according to (2.5), is given in the form

$$H=-\tfrac{1}{2}\xi^{-1}\frac{\partial}{\partial\xi}\xi\frac{\partial}{\partial\xi}+V(\xi) \tag{A1}$$

in the cylindrical coordinate system of $\boldsymbol{\xi}$. We begin with (2.3),

$$\left(\frac{\partial}{\partial\eta}+H\right)\tilde I(\xi,\eta)=0, \tag{A2}$$

which, by the Laplace transformation with respect to η, gives

$$\mathrm{Lap}[\tilde I(\xi,\eta)]=\int_0^\infty d\eta e^{\lambda\eta}\tilde I(\xi,\eta)$$
$$=(H-\lambda)^{-1}\tilde I(\xi,0),\quad \mathrm{Re}[\lambda]<0, \tag{A3}$$

which, by the Laplace inversion, gives

$$\tilde I(\xi,\eta)=\frac{1}{2\pi i}\int_{-i\infty-\epsilon}^{i\infty-\epsilon}d\lambda\,(H-\lambda)^{-1}e^{-\lambda\eta}\tilde I(\xi,0),\quad \epsilon>0. \tag{A4}$$

Here, the operator $R_\lambda=(H-\lambda)^{-1}$ is called the resolvent operator of H.

In (A1), we first suppose $V(\xi)$ to be a real function of ξ so that the operator H becomes real and self-adjoint. The eigenvalue equation is given by

$$H\psi_A=A\psi_A, \tag{A5}$$

where the eigenfunctions ψ_A are to be finite at $\xi=0$ and to tend to zero as $\xi\to\infty$; the eigenvalues A are first assumed to have only discrete values.

For two arbitrary functions ψ' and ψ'', it holds

$$\int_a^b d\xi\,\xi\,(\psi'H\psi''-\psi''H\psi')$$
$$=-\tfrac{1}{2}\xi\left(\psi'\frac{\partial}{\partial\xi}\psi''-\psi''\frac{\partial}{\partial\xi}\psi'\right)\Big|_{\xi=a}^{b}. \tag{A6}$$

Here, by putting $\psi'=\psi_A$, $\psi''=\psi_B$, and $a=0$, $b=\infty$ in (A6), we can directly confirm the reality of eigenvalues and also the orthogonality of eigenfunctions

$$(\psi_A,\psi_B)\equiv\int_0^\infty d\xi\,\xi\,\psi_A(\xi)\psi_B(\xi)=\delta_{AB}, \tag{A7}$$

which, together with the completeness, gives

$$\sum_A \psi_A(\xi_1)\,\psi_A(\xi_2) = \xi_1^{-1}\delta(\xi_1 - \xi_2). \tag{A8}$$

The resolvent operator R_λ can be exhibited in terms of the eigenvalues and eigenfunctions of H as

$$R_\lambda(\xi_1|\xi_2) = \sum_A (A-\lambda)^{-1}\psi_A(\xi_1)\psi_A(\xi_2), \tag{A9}$$

which, on the other hand, is the solution of

$$(H-\lambda)R_\lambda(\xi|\xi') = \xi^{-1}\delta(\xi - \xi'), \tag{A10}$$

as may directly be proven by the substitution of (A9) with (A8). Equation (A9) shows that $R_\lambda(\xi_1|\xi_2)$ has the poles at $\lambda = A,B,\cdots$, with the residue values

$$\mathrm{Res}[R_\lambda(\xi_1|\xi_2)]_{\lambda=A} = -\psi_A(\xi_1)\psi_A(\xi_2). \tag{A11}$$

On the other hand, the solution of (A10) can be expressed in terms of the solutions of

$$(H-\lambda)f_\lambda(\xi) = 0: \tag{A12}$$

Let $f_\lambda^0(\xi)$ and $f_\lambda^\infty(\xi)$ be the solutions of (A12) satisfying the conditions at $\xi = 0$ and $\xi = \infty$, respectively, and substitute $\psi' f_\lambda^0(\xi)$ and $\psi'' f_\lambda^\infty(\xi)$ in (A6). Then, in terms of the notation

$$[f_\lambda^0,f_\lambda^\infty] \equiv \tfrac{1}{2}\xi\left[f_\lambda^0(\xi)\frac{\partial}{\partial\xi}f_\lambda^\infty(\xi) - f_\lambda^\infty(\xi)\frac{\partial}{\partial\xi}f_\lambda^0(\xi)\right] \tag{A13}$$

being equal to the Wronskian of f_λ^0 and f_λ^∞ except for the factor $\frac{1}{2}\xi$, the use of (A12) shows that

$$[f_\lambda^0,f_\lambda^\infty] = -[f_\lambda^\infty,f_\lambda^0] = \text{const}, \tag{A14}$$

independently of ξ. Now, the solution of (A10) is given by

$$R_\lambda(\xi_1|\xi_2) = [f_\lambda^\infty f_\lambda^0]^{-1} f_\lambda^0(\xi_<) f_\lambda^\infty(\xi_>), \tag{A15}$$

where $\xi_>$, $\xi_<$ denote the larger and smaller of ξ_1, ξ_2, respectively. In order to prove (A15), it is only necessary to check, on the left-hand side of (A10), the term that resulted from the discontinuity at $\xi = \xi'$ and, with the aid of (A13) and (A14), this term is easily shown to give the right-hand side of (A10).

In order to evaluate the eigenfunctions $\psi_A(\xi)$ according to (A11) with (A15), we first note that $\psi_A(\xi)$ are the solutions of (A12) satisfying the boundary conditions at $\xi = 0$ and ∞ at the same time and therefore, as $\lambda \to A$, $f_\lambda^0(\xi)$ and $f_\lambda^\infty(\xi)$ tend to the same function except for a constant factor; hence also $[f_\lambda^\infty f_\lambda^0] \to 0$ by the definition (A13). Thus, it follows

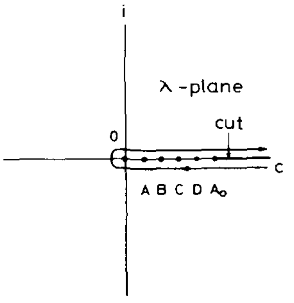


FIG. 7. Path of integration for (A22).

from (A15) that $R_\lambda(\xi_1|\xi_2)$ has the poles at $\lambda = A$ by the factor $[f_\lambda^\infty f_\lambda^0]^{-1}$, and therefore from (A11) also that

$$\psi_A(\xi_1)\psi_A(\xi_2) = -f_A^0(\xi_1)f_A^\infty(\xi_2)\Big/\frac{\partial}{\partial\lambda}[f_\lambda^\infty f_\lambda^0]_{\lambda=A}, \tag{A16}$$

and the eigenvalues A are given by the roots of

$$[f_\lambda^\infty f_\lambda^0]_{\lambda=A} = 0. \tag{A17}$$

Here, the above result is valid even when $V(\xi)$ is a discontinuous function of ξ, and it is then most convenient to evaluate $[f_\lambda^\infty f_\lambda^0]$ at the point of discontinuity since it is to be independent of ξ.

It often happens that, besides the discrete eigenvalues $A,B,\cdots < A_0$, as we have assumed, there exist also the continuous eigenvalues in the range $\infty \geqslant A \geqslant A_0$. However, the extension of the above results to this general case is rather straightforward: (A7)–(A9) are first changed by

$$(\psi_B,\psi_A) \equiv \int_0^\infty d\xi\,\xi\,\psi_B(\xi)\psi_A(\xi) = \delta(B-A), \quad B,A > A_0, \tag{A18}$$

$$\sum_{A<A_0}\psi_A(\xi_1)\psi_A(\xi_2) + \int_{A_0}^\infty dA\,\psi_A(\xi_1)\psi_A(\xi_2) = \xi_1^{-1}\delta(\xi_1-\xi_2),$$

$$R_\lambda(\xi_1|\xi_2) = \sum_{A<A_0}(A-\lambda)^{-1}\psi_A(\xi_1)\psi_A(\xi_2) + \int_{A_0}^\infty dA\,(A-\lambda)^{-1}\psi_A(\xi_1)\psi_A(\xi_2). \tag{A19}$$

The expression (A19) means that, in the complex plane of λ, $R_\lambda(\xi_1|\xi_2)$ has a branch cut starting from $\lambda = A_0$ and ending at $\lambda = +\infty$ along the real axis, besides the poles at $\lambda = A,B,\cdots < A_0$. Here, the discontinuity of the values of R_λ on the upper and lower sides of the branch cut is

$$[R_{\lambda+i0}(\xi_1|\xi_2) - R_{\lambda-i0}(\xi_1|\xi_2)]_{\lambda=A>A_0} = 2\pi i\psi_A(\xi_1)\psi_A(\xi_2). \tag{A20}$$

Therefore, it follows from (A11) and (A20) that, for any function given in the form

$$F(\xi_1|\xi_2) = \sum_{A<A_0} f(A)\psi_A(\xi_1)\psi_A(\xi_2) + \int_{A_0}^\infty dA\,f(A)\psi_A(\xi_1)\psi_A(\xi_2) \tag{A21}$$

with the condition $f(A) \to 0$ for $A \to +\infty$, $F(\xi_1|\xi_2)$ can be exhibited in terms of $R_\lambda(\xi_1|\xi_2)$ as

$$F(\xi_1|\xi_2) = \frac{1}{2\pi i}\int_c d\lambda f(\lambda)R_\lambda(\xi_1|\xi_2), \tag{A22}$$

where, if the smallest eigenvalue is chosen to be zero, the path of integration c starts at $+\infty$ and goes to the origin along the infinitesimally lower side of the real axis; then it encircles the origin in the clockwise and returns to $+\infty$ along the infinitesimally upper side of the real axis (Fig. 7).

Note that the expression (A22) is correct irrespectively of whether the eigenvalues are discrete, continuous or both.

In the case of (2.11), for example, $f(\lambda) = \exp(-\lambda\eta)$ and the expression (2.11b) can be given in terms of $\int_0^\infty d\xi\, \xi F(0|\xi)$ and, in the case of (2.29), in terms of $F(0|0)$. Here, the path of integration in (A22) can be deformed to give

$$F(\xi_1|\xi_2) = \frac{1}{2\pi i}\int_{-i\infty-\epsilon}^{i\infty-\epsilon} d\lambda e^{-\lambda\eta} R_\lambda(\xi_1|\xi_2). \tag{A23}$$

On the other hand, directly from (A4), the use of (A23) gives

$$\tilde{I}(\xi,\eta) = \int_0^\infty d\xi'\xi' F(\xi|\xi')\tilde{I}(\xi',0), \tag{A24}$$

which is the alternative expression of (2.8).

APPENDIX B: ASYMPTOTIC VALUES OF A AND C_A

Although it is difficult to find the exact eigenvalues A and the eigenfunctions ψ_A by solving (2.20), their asymptotic expressions for large values of A are comparatively easy to obtain: We start from (A12) which, in the present case, is given by

$$\left[\xi^{-1}\frac{\partial}{\partial\xi}\xi\frac{\partial}{\partial\xi} - \xi^\alpha + 2\lambda\right] f_\lambda(\xi) = 0. \tag{B1}$$

Here, when $\lambda \gg 1$, the asymptotic expression of $f_\lambda^0(\xi)$ is found to be[21]

$$\begin{aligned} &f_\lambda^0(\xi) = [\varphi(\xi)/\xi\varphi'(\xi)]^{1/2} J_0[\varphi(\xi)], \quad f_\lambda^0(0) = 1, \\ &\varphi(\xi) = \int_0^\xi dx (2\lambda - x^\alpha)^{1/2}, \quad \varphi'(\xi) = (2\lambda - \xi^\alpha)^{1/2}, \end{aligned} \tag{B2}$$

which gives

$$f_\lambda^0(\xi) \sim \begin{cases} J_0[(2\lambda)^{1/2}\xi], & \xi \ll (2\lambda)^{1/\alpha} \\ (2/\pi)^{1/2}[\xi\varphi'(\xi)]^{-1/2}\cos[\varphi(\xi) - \pi/4], \\ \quad \varphi(\xi) \gg 1. \end{cases} \tag{B3}$$

In the neighborhood of $\xi \sim (2\lambda)^{1/\alpha}$ where $\varphi'(\xi) \sim 0$ or $f_\lambda^0(\xi) \sim \infty$, we can make the approximation $2\lambda - \xi^\alpha \sim \alpha(2\lambda)^{1-1/\alpha}[(2\lambda)^{1/\alpha} - \xi]$ and replace (B1) by

$$\begin{aligned} &[(\partial/\partial z)^2 + z/3] f(z) = 0, \quad \xi \sim (2\lambda)^{1/\alpha} \gg 1, \\ &z = h[(2\lambda)^{1/\alpha} - \xi], \quad h = [3\alpha(2\lambda)^{1-1/\alpha}]^{1/3}. \end{aligned} \tag{B4}$$

Here, in the same range of ξ as in (B4), the asymptotic expression in (B3) is expressed by

$$f_\lambda^0(\xi) = f^0(z) \sim (2/\pi)^{1/2}(2\lambda)^{-1/2\alpha}h^{-1/2}(z/3)^{-1/4}$$

$$\times \cos\{\varphi[(2\lambda)^{1/\alpha}] - Z - \pi/4\}, \quad z \gg 1, \tag{B5}$$

$$Z = 2(z/3)^{3/2}. \tag{B6}$$

On the other hand, the solution of (B4) which tends to zero as $z \to -\infty$ (or $\xi \to +\infty$), say $f^\infty(z)$, is given, in terms of Z in (B6), by

$$\begin{aligned} f^\infty(z) &= \frac{1}{2}\int_{-\infty}^\infty dt \exp[i(t^3 - zt)] \\ &= \tfrac{1}{3}|z|^{1/2}K_{1/3}(|Z|), \quad z<0 \\ &= (\pi/3)(z/3)^{1/2}[J_{1/3}(Z) + J_{-1/3}(Z)], \quad z>0 \end{aligned} \tag{B7}$$

and, when $z \gg 1$, has the asymptotic expression

$$f^\infty(z) \sim (\pi/3)^{1/2}(z/3)^{-1/4}\cos[Z - \pi/4], \quad z \gg 1, \tag{B8}$$

which corresponds to (B5) for $f_\lambda^0(\xi)$.

In order to find the bracket $[f_\lambda^\infty, f_\lambda^0]$ defined by (A13), it is most simple to evaluate it in the range of $z \gg 1$ where the asymptotic expressions of f_λ^0 and f_λ^∞ are given by (B5) and (B8), respectively. The result is

$$[f_\lambda^\infty f_\lambda^0] = -6^{-1/2}h^{1/2}(2\lambda)^{1/2\alpha}\sin\{\varphi[(2\lambda)^{1/\alpha}] - \pi/2\}, \tag{B9}$$

which is independent of z, as is required. Thus, according to (A17), the eigenvalues A are the roots of

$$\varphi[(2A)^{1/\alpha}] = \pi(n + \tfrac{1}{2}) > 0, \quad n = 0,1,2,\cdots, \tag{B10}$$

and, correspondingly,

$$\begin{aligned} &\left(\frac{\partial}{\partial\lambda}\right)[f_\lambda^\infty f_\lambda^0]_{\lambda=A} \\ &\quad = (-)^{n+1}6^{-1/2}h^{1/2}(2A)^{1/2\alpha}\left(\frac{\partial}{\partial A}\right)\varphi[(2A)^{1/\alpha}], \\ &\left(\frac{\partial}{\partial A}\right)\varphi[(2A)^{1/\alpha}] = \int_0^{(2A)^{1/\alpha}} dx(2A - x^\alpha)^{-1/2} \\ &\quad = \alpha^{-1}(2A)^{1/\alpha - 1/2}B(1/\alpha, 1/2), \end{aligned} \tag{B11}$$

In terms of the beta function defined by (2.23).

On the other hand, according to (A16), the coefficients C_A in (2.11b) is given by

$$\begin{aligned} C_A &= \psi(0)\int_0^\infty d\xi\xi\psi_A(\xi) \\ &= -f_A^0(0)\int_0^\infty d\xi\xi f_A^\infty(\xi) \Big/ \frac{\partial}{\partial\lambda}[f_\lambda^\infty f_\lambda^0]_{\lambda=A}. \end{aligned} \tag{B12}$$

Here, by (B2), $f_A^0(0) = 1$ and, when $A \gg 1$, the last integral is determined effectively by the integration over the range where the expression (B7) is valid, since the integrand rapidly oscillates in the other range of integration. Here, the integral representation in (B7) immediately gives

$$\int_{-\infty}^\infty dz f^\infty(z) = \pi, \tag{B13}$$

and hence, by the change of variable from ξ to z according to (B4),

$$\begin{aligned} &\int_0^\infty d\xi\xi f_A^\infty(\xi) \sim (2A)^{1/\alpha}h^{-1}\int_{-\infty}^\infty dz f^\infty(z) = \pi h^{-1}(2A)^{1/\alpha}, \\ &\quad \xi \sim (2A)^{1/\alpha} \gg 1. \end{aligned} \tag{B14}$$

Thus, on using (B11), (B12), (B14), and also h in (B4), we finally find

$$C_A \sim (-)^n (2\alpha)^{1/2}\pi/B(1/\alpha,1/2), \quad A \gg 1. \tag{B15}$$

On the other hand, from (B5) and (B8),

$$f_A^\infty(\xi)/f_A^0(\xi) \simeq (-)^n \pi 6^{-1/2}(2A)^{1/2\alpha} h^{1/2}|_{\lambda=A}, \quad f_A^0(0)=1, \tag{B16}$$

which used in (A16) with (B11) yields

$$\psi_A^2(0) = \alpha(2A)^{1/2-1/\alpha}\pi/B(1/\alpha,1/2). \tag{B17}$$

APPENDIX C: STATISTICS OF THE MEDIUM

Let $\mathbf{q}(\bar{x})$ be an operator involving arbitrary $c(\bar{x})$ and $\delta/\delta c(\bar{x})$, defined by

$$\mathbf{q}(\bar{x}) = c(\bar{x}) + \int d\bar{x}' \langle \epsilon(\bar{x})\epsilon(\bar{x}')\rangle \delta/\delta c(\bar{x}') + (1/2)\int d\bar{x}_1'\, d\bar{x}_2' \langle \epsilon(\bar{x})\epsilon(\bar{x}_1')\epsilon(\bar{x}_2')\rangle \times [\delta/\delta c(\bar{x}_1')]\,[\delta/\delta c(\bar{x}_2')] + \cdots; \tag{C1}$$

then, for any function $f(\epsilon)$ of $\epsilon(\bar{x})$, its average value is given, on referring to Secs. VI and VII in Ref. 20, by

$$\langle f[\epsilon + c]\rangle = f[\mathbf{q}], \tag{C2}$$

which tends to $\langle f[\epsilon]\rangle$ as $c \to 0$. Hence, it follows that, for example,

$$\langle \epsilon(\bar{x})\psi[\epsilon]\rangle = \mathbf{q}(\bar{x})\psi[\mathbf{q}]|_{c=0} = \mathbf{q}(\bar{x})\langle \psi[\epsilon + c]\rangle|_{c=0} = \int d\bar{x}' \langle \epsilon(\bar{x})\epsilon(\bar{x}')\rangle\, [\delta/\delta c(\bar{x}')] \times \langle \psi[\epsilon + c]\rangle|_{c=0} + O[\epsilon^3], \tag{C3}$$

in view of the condition that $|\epsilon| \ll 1$ in (1.1). Thus, the second term on the right-hand side of (C3) is certainly negligible, compared to the first term, and the result becomes the same as what is obtained by assuming the Gaussian statistics for ϵ.

[1] V.I. Shishov, Sov. Astron. **17**, 598 (1974) [Astron. Zh. **50**, 941 (1973)].

[2] L.M. Erukhimov, I.G.A. Zarnitsyna, and P.I. Kirsh, Radiophys. Quantum Electron. **16**, 436 (1973).

[3] C.H. Liu, A.W. Wernik, and K.C. Yeh, IEEE Trans. Antennas Propagat. **AP-22**, 624 (1974).

[4] B.J. Uscinski, Proc. R. Soc. London Ser. A **336**, 379 (1974).

[5] L.C. Lee and J.R. Jokipii, Astrophys. J. **196**, 695 (1975); **201**, 532 (1975).

[6] C.H. Liu and K.C. Yeh, Radio Sci. **10**, 1055 (1975).

[7] A. Ishimaru and S.T. Hong, Radio Sci. **10**, 637 (1975).

[8] S.T. Hong and A. Ishimaru, Radio Sci. **11**, 551 (1976).

[9] I. Sreenivasiah, A. Ishimaru, and S.T. Hong, Radio Sci. **11**, 775 (1976).

[10] S.T. Hong, I. Sreenivasiah, and A. Ishimaru, IEEE Trans. Antennas Propagat. **AP-25**, 822 (1977).

[11] J.W. Strohbehn, Proc. IEEE **56**, 1301 (1968), and other articles quoted.

[12] V.I. Tatarski, "The effects of the turbulent atmosphere on wave propagation," English translation available from the U.S. Dept. of Commerce, NTIS, Springfield, VA 22151 (1971).

[13] L.L. Foldy, Phys. Rev. **67**, 107 (1945).

[14] K. Furutsu, Radio Sci. **10**, 29 (1975).

[15] A detailed review of the time-independent transport equation and the related references were given by Yu. N. Barabanenkov, Yu. A. Kravtsov, S.M. Rytov, and V.I. Tatarski, Sov. Phys. Usp. **13**, 551 (1971).

[16] L. Dolin, Radiophys. Quantum Electron. **9**, 40 (1966).

[17] The "refractive" effect means the effect of variation of optical distance due to the fluctuation of index of refraction for each path of wave propagation. In the case of pulsars, the characteristic time of this variation is of the order of a month or a year.[1,5]

[18] Strictly speaking, this equivalence is not perfect since z and ct are exchanged in the two expressions, but this discrepancy is negligible whenever the pulse width ζ is small enough as compared to z.

[19] This difference was emphasized qualitatively in Ref. 5 based on the numerical solution.

[20] K. Furutsu, J. Opt. Soc. Am. **62**, 240 (1972).

[21] G.N. Watson, *Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1952), 2nd ed., §4.31, §6.4.

Lowering and raising operators of IU(n) and IO(n) and their normalization constants

M. K. F. Wong

Fairfield University, Fairfield, Connecticut 06430

Hsin-Yang Yeh

Moorhead State University, Moorhead, Minnesota 56560

(Received 28 June 1978)

Lowering and raising operators for the vector space $U(n) \subset IU(n)$ and $O(n) \subset IO(n)$ have been obtained, and their normalization constants evaluated. For $U(n) \subset IU(n)$, we obtain two forms, one according to Nagel and Moshinsky, and the other according to Bincer. For $O(n) \subset IO(n)$, we obtain the shift operators according to Bincer.

1. INTRODUCTION

In a previous paper,[1] we obtained the invariant operators and their eigenvalues for IU(n) and IO(n). This was made possible by noticing the close similarity between the representations of IU(n) and U(n,1) and between IO(n) and O(n,1). Encouraged by the results, we ask whether there are quantities in IU(n) and IO(n) whose properties might be deducible from similar quantities in U(n,1) and O(n,1). We find that, in fact, the lowering and raising operators of IU(n) and IO(n) and their normalization constants can be obtained from those of U(n,1) and O(n,1). This is the subject of our present paper.

The raising and lowering operators of U(n,1) have been discussed by Patera.[2] Those of O(n,1) have been discussed by Wong.[3] These operators and their normalization constants are practically the same as those in U($n+1$) and O($n+1$).

The basic reason for our statement above is that the vector space of U(n) contained in (1) IU(n), (2) U(n,1), and (3) U($n+1$) is the same vector space. Similarly, the vector space of O(n) contained in (1) IO(n), (2) O(n,1), and (3) O($n+1$) is the same vector space. Therefore, those operators which raise or lower the vector space U(n), or O(n), must have the same defining equations in all three cases, insofar as they are expressed as commutation relations with the generators of U(n), or O(n), acting on the highest weight of U(n), or O(n). Thus the raising and lowering operators of IU(n) can be obtained from those of U(n,1) or U($n+1$), by changing $A^{\,i}_{\,n+1}$ to $I^{\,i}_{\,n+1}$ and $A^{\,n+1}_{\,i}$ to $I^{\,n+1}_{\,i}$. Likewise, the shift operators of IO(n) can be obtained from those of O(n,1) or O($n+1$) recently obtained by Bincer[4–6] by changing $C^{\,0}_{\,a}$ to $I^{(2\nu+1)}_{a}$ for IO(2ν) and $\frac{1}{2}(C^{\,1}_{\,a} - C^{\,-1}_{\,a})$ to $I^{(2\nu)}_{a}$ for IO($2\nu-1$) if $a \neq 0$, and $C^{\,1}_{\,1}$ to $I^{2\nu}_{0}$ if $a=0$, where $I^{(2\nu)}_{0}$ is the diagonal part of $I_{2\nu,2\nu-1}$.

In the case of IU(n), there now exist two simple and elegant expressions for the lowering and raising operators, one due to Nagel and Moshinsky,[7] and the other due to Bincer.[4] We shall discuss both of them in Sec. 2. There we shall also discuss the normalization constants. We find the normalization constants can be most easily obtained by relating them to the matrix elements of the generators.

In the case of IO(n), it is found that the raising and lowering operators of Pang and Hecht,[8] or Wong[9], cannot be immediately extended from O($n+1$) to IO(n), because these operators are not, according to Bincer's definition, "one-tensor" operators. Fortunately, Bincer has obtained recently the shift operators for O($n+1$) in product form which are one-tensor operators. These operators can be easily extended to IO(n). The normalization constants can then be easily obtained too. This is the content of Sec. 3.

In conclusion, we find that the normalization constants of IU(n) differ from those of U($n+1$) by a factor which is just the one by which the matrix elements of $I^{\,n}_{\,n+1}$ differ from those of $A^{\,n}_{\,n+1}$, namely, $\kappa^{-1}[-(h_{1\,n+1} - h_{jn} + j) \times (h_{n+1\,n+1} - h_{jn} - n + j)]^{1/2}$. Similarly, the normalization constants of IO(n) differ from those of O($n+1$) by the same factor as the difference between the matrix elements of $J_{2k+1,2k}$ and $I_{2k+1,2k}$, i.e., $\kappa^{-1}[(l_{2k+1,1} - l_{2kj} - 1)(l_{2k+1,1} + l_{2kj})]^{1/2}$, and between $J_{2k-1,2k}$ and $I_{2k-1,2k}$, i.e., $\kappa^{-1}[l^{\,2}_{2k,1} - l^{\,2}_{2k-1,j}]^{1/2}$, and between the diagonal matrix elements of $J_{2k-1,2k}$ and $I_{2k-1,2k}$, i.e., $\kappa^{-1} i l_{2k,1}$, multiplied by a constant factor which happens to be $(1-k)^k$.

2. RAISING AND LOWERING OPERATORS OF IU(n) AND THEIR NORMALIZATION CONSTANTS

We use the notation of the previous paper.[1] The lowering operator $L^{\,m}_{\,n+1}$ is defined by

$$L^{\,m}_{\,n+1}\left|\begin{matrix} h_2 \cdots h_n \\ q_1 \cdots q_m \cdots q_n \end{matrix}\right\rangle \propto \left|\begin{matrix} h_2 \cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_n \end{matrix}\right\rangle, \quad 1 \leqslant m \leqslant n, \tag{2.1}$$

where h_1 and h_{n+1} are missing from the first row, because we are dealing with the representations of IU(n).

Similarly, the raising operator $R^{\,n+1}_{\,m}$ is defined by

$$R^{\,n+1}_{\,m}\left|\begin{matrix} h_2 \cdots h_n \\ q_1 \cdots q_m \cdots q_n \end{matrix}\right\rangle \propto \left|\begin{matrix} h_2 \cdots h_n \\ q_1 \cdots q_m + 1 \cdots q_n \end{matrix}\right\rangle, \quad 1 \leqslant m \leqslant n. \tag{2.2}$$

We find that (2.1) is equivalent to the following two equations:

$$[A^{\,i}_{\,i}, L^{\,m}_{\,n+1}] = -\delta^m_i L^{\,m}_{\,n+1}, \quad 1 \leqslant i, m \leqslant n, \tag{2.3}$$

 0022-2488/79/040629-05$01.00

$$[A_p^{p+1}, L_{n+1}^m]\left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0, \quad 1 \leqslant p \leqslant n-1. \tag{2.4}$$

Note that these two equations, (2.3) and (2.4), are exactly the same as those for U(n,1) and U($n+1$). Similarly, we have, for R_m^{n+1}

$$[A_p^i, R_m^{n+1}] = \delta_i^m R_m^{n+1}, \quad 1 \leqslant i, m \leqslant n, \tag{2.5}$$

$$[A_p^{p+1}, R_m^n]\left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0, \quad 1 \leqslant p \leqslant n-1. \tag{2.6}$$

Now if we look at the lowering (raising) operator of Nagel and Moshinsky, we find that it is linear in $A_{n+1}^{\mu_p}$ ($A_{\mu_1}^{n+1}$). Therefore its commutation relations with the generators of U(n) must be the same if $A_{n+1}^{\mu_p}$ ($A_{\mu_1}^{n+1}$) is replaced by a similar operator in IU(n) whose commutation relations with the generators of U(n) are the same as $A_{n+1}^{\mu_p}$ ($A_{\mu_1}^{n+1}$). The simplest choice is $I_{n+1}^{\mu_p}$ ($I_{\mu_1}^{n+1}$), i.e., the translation generators. Thus we obtain the lowering and raising operators of IU(n), according to Nagel and Moshinsky:

$$L_{n+1}^m = \Bigg(\sum_{p=0}^{n-m} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = m+1}^{n} A_{\mu_1}^m A_{\mu_2}^{\mu_1} \cdots A_{\mu_p}^{\mu_{p-1}} \times I_{n+1}^{\mu_p} \prod_{i=1}^{p} \epsilon_{m\mu_i}^{-1}\Bigg) \prod_{\mu=m+1}^{n} \epsilon_{m\mu}, \tag{2.7}$$

$$R_m^{n+1} = \Bigg(\sum_{p=0}^{m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = 1}^{m-1} A_m^{\mu_p} A_{\mu_p}^{\mu_{p-1}} \cdots A_{\mu_2}^{\mu_1} \times I_{\mu_1}^{n+1} \prod_{i=1}^{p} \epsilon_{m\mu_i}^{-1}\Bigg) \prod_{\mu=1}^{m-1} \epsilon_{m\mu}, \tag{2.8}$$

where $\epsilon_{m\mu} = A_m^m - A_\mu^\mu + \mu - m$.

The normalization constants are defined in the same way as Nagel and Moshinsky, i.e.,

$$\mathscr{L}_{q_\mu - \delta_{\mu m}}^{q_\mu}\left|\begin{matrix} h_\mu \\ q_\mu \end{matrix}\right\rangle = \left|\begin{matrix} h_\mu \\ q_\mu - \delta_{\mu m} \end{matrix}\right\rangle, \quad 1 \leqslant m \leqslant n, \tag{2.9}$$

$$\mathscr{L}_{q_\mu - \delta_{\mu m}}^{q_\mu} = (N_{q_\mu - \delta_{\mu m}}^{q_\mu})^{-1} L_{n+1}^m. \tag{2.10}$$

It is obvious that the normalization constants for U(n)⊃U($n-1$)⊃⋯⊃U(2)⊃U(1) must be the same as Nagel and Moshinsky, since the states are not changed. The only difference must come from IU(n)⊃U(n), i.e., when $h_{\mu\, n+1}$ are involved. The normalization constants can be most easily obtained if we relate them to the matrix elements of the generators. We use basically Eqs. (7.1) and (7.2) of Nagel and Moshinsky.[10] In the derivation of Eq. (7.1), they also use Eq. (4.2a″). It can be easily checked that these equations are still true in our case. We thus obtain the following equation:

$$\left\langle \begin{matrix} h_2 \cdots h_n \\ q_{1n} \cdots q_{ln}+1 \cdots q_{nn} \\ r_\mu \end{matrix} \middle| I_n^{n+1} \middle| \begin{matrix} h_2 \cdots h_n \\ q_{1n} \cdots q_{ln} \cdots q_{nn} \\ r_\mu \end{matrix} \right\rangle$$

$$= \prod_{\substack{\lambda=1 \\ \lambda \neq l}}^{n} (q_{ln} - q_{\lambda n} + \lambda - l)^{-1} N_{u(n-1)}^{-1}\begin{pmatrix} q_{\mu n} \\ q_{\mu n} \\ q_{\mu\, n-1} \end{pmatrix} \times N_{u(n-1)}\begin{pmatrix} q_{\mu n} + \delta_{\mu l} \\ q_{\mu n} + \delta_{\mu l} \\ q_{\mu\, n-1} \end{pmatrix} N_{u(n)}\begin{pmatrix} q_{\mu n} \\ h_{2n+1} \cdots h_{n\, n+1} \\ q_{\mu n} + \delta_{\mu l} \end{pmatrix}. \tag{2.11}$$

The only difference between our Eq. (2.11) and Nagel and Moshinsky's Eq. (7.2) is that on the left-hand side we have the matrix elements of I_n^{n+1} and on the right hand side we have

$$N_{u(n)}\begin{pmatrix} q_{\mu n} \\ q_{\mu n} + \delta_{\mu l} \end{pmatrix}.$$

If we take the matrix elements of I_n^{n+1} as known, given by Chakrabarti,[11] then $N_{u(n)}$ can be easily calculated. The result is: $N_{u(n)}$ differs from that of Nagel and Moshinsky by a factor $\kappa^{*-1}[-(h_{1\, n+1} - h_{j\, n+1} + j)(h_{n+1\, n+1} - h_{jn} - n + j)]^{1/2}$. Explicitly we have

$$N_{u(n)}^{(R)}\begin{pmatrix} q_{ln} \\ h_{2\, n+1} \cdots h_{n\, n+1} \\ q_{ln} + 1 \end{pmatrix} = \kappa^*(-1)^{l-1}\left[\frac{\prod_{\mu=1}^{l-1}(q_{ln} - q_{\mu n} + \mu - l)}{\prod_{\mu=l+1}^{n}(q_{ln} - q_{\mu n} + \mu - l + 1)} \times \prod_{\mu=2}^{n}(q_{ln} - h_{\mu\, n+1} + \mu - l)\right]^{1/2}, \tag{2.12}$$

$$N_{u(n)}^{(L)}\begin{pmatrix} q_{ln} \\ h_{2\, n+1} \cdots h_{n\, n+1} \\ q_{ln} - 1 \end{pmatrix} = \kappa\left[\frac{\prod_{\mu=l+1}^{n}(q_{ln} - q_{\mu n} + \mu - l)}{\prod_{\mu=1}^{l-1}(q_{ln} - q_{\mu n} + \mu - l - 1)} \times \prod_{\mu=2}^{n}(q_{ln} - h_{\mu\, n+1} + \mu - l - 1)\right]^{1/2}. \tag{2.13}$$

Note that without this method, it will be rather difficult to obtain the normalization constants, since one has to relate them to at least the third order invariants of IU(n) in order to obtain all the quantities in (2.12) or (2.13).

Recently, Bincer[4] has obtained the lowering operators of U(n) in product form. Using his method, we find that the raising operators of U(n) can be expressed as follows:

$$R_m^n = \left\{C \circ \prod_{k=1}^{m-1}(C - c_k \mathbb{1})\right\}_m^n, \tag{2.14}$$

where

$$c_k = q_k - k + 1, \tag{2.15}$$

$$(X \circ Y)_\beta^\alpha = \sum_{\rho=1}^{n-1} X_\rho^\alpha Y_\beta^\rho, \tag{2.16}$$

and, for $m = 1$,

$$\prod_{k=1}^{m-1}(C - c_k \mathbb{1}) = \mathbb{1}. \tag{2.17}$$

The normalization constant for the raising operators of U(n) is found to be

$$N\begin{pmatrix} q_m \\ q_m+1 \end{pmatrix} = \Big\{ -\Big[\prod_{\lambda=1}^{n} (q_m - h_\lambda + \lambda - m) \prod_{\lambda=1}^{m-1} (q_m - q_\lambda + \lambda - m)\Big] \Big[\prod_{\lambda=m+1}^{n-1} (q_m - q_\lambda + \lambda - m + 1)\Big]^{-1}\Big\}^{1/2}. \quad (2.18)$$

Now we can directly take over Bincer's results and write down the lowering and raising operators of IU(n). They are

$$L^m_{n+1} = \Big\{C \prod_{k=m+1}^{n} (C - c_k \mathbb{1})\Big\}^m_{n+1}, \quad (2.19)$$

where

$$c_k = q_k - k + n, \quad (2.20)$$

$$(XY)^\alpha_\beta = \sum_{\rho=1}^{n} X^\rho_\beta Y^\alpha_\rho, \quad (2.21)$$

$$C^i_{n+1} \equiv I^i_{n+1}, \quad i = 1,2,\ldots,n, \quad (2.22)$$

and, for $m = n$,

$$\prod_{k=m+1}^{n} (C - c_k \mathbb{1}) \equiv \mathbb{1}. \quad (2.23)$$

The normalization constant for the lowering operators of IU(n)$\supset$U(n) is obtained as follows. We define

$$V(\mu+1)^m_{n+1} \equiv \{V(\mu)(C - c_\mu \mathbb{1})\}^m_{n+1}. \quad (2.24)$$

Then, taking matrix elements between semimaximal states both in Eqs. (2.25) and (2.26), we have

$$\begin{aligned} L^m_{n+1}(\mu+1) &= \sum_{a=1}^{m} V(\mu)^a_{n+1}(C - c_\mu \mathbb{1})^m_a + \sum_{a=m+1}^{n} V(\mu)^a_{n+1} C^m_a \\ &= V(\mu)^m_{n+1}(C^m_m - c_\mu) + \sum_{a=m+1}^{n} V(\mu)^a_{n+1} C^m_a \\ &= V(\mu)^m_{n+1}(q_m - c_\mu) \\ &\quad + \sum_{a=m+1}^{n} \{C^m_a V(\mu)^a_{n+1}[V(\mu)^a_{n+1}, C^m_a]\} \\ &= V(\mu)^m_{n+1}(q_m - c_\mu) + V(\mu)^m_{n+1}(n-m) \\ &= V(\mu)^m_{n+1}(q_m - m + n - c_\mu) \\ &= V(\mu)^m_{n+1}(c_m - c_\mu). \end{aligned} \quad (2.25)$$

From the recurrence relation in (2.25), we obtain

$$L^m_{n+1} = \prod_{j=m+1}^{n} (c_m - c_j) V(m+1)^m_{n+1} = \prod_{j=m+1}^{n} (c_m - c_j) I^m_{n+1}. \quad (2.26)$$

Therefore the normalization constant is found to be

$$N\begin{pmatrix} q_m \\ q_{m-1} \end{pmatrix} = \Big\langle q_m - 1 \Big| \prod_{j=m+1}^{n} (c_m - c_j) I^m_{n+1} \Big| q_m \Big\rangle. \quad (2.27)$$

If we look at the matrix elements of

$$\Big\langle q_m - 1 \Big| \prod_{j=m+1}^{n} (c_m - c_j) A^m_n \Big| q_m \Big\rangle, \quad (2.28)$$

we find that it agrees with the normalization constant of Bincer for U(n). Thus, from (2.27), we conclude that

$$N\begin{pmatrix} q_m \\ q_{m-1} \end{pmatrix} \quad \text{for} \quad \mathrm{IU}(n)\supset \mathrm{U}(n)$$

differs from the normalization constant of Bincer for U($n+1$)$\supset$U(n) by a factor $\kappa^{-1}[-(h_{1\,n+1} - h_{j\,n+1} + j)(h_{n+1\,n+1} - h_{jn} - n + j)]^{1/2}$. Explicitly we have

$$\begin{aligned} N^L_{u(n)}\begin{pmatrix} q_{ln} \\ h_{2\,n+1} \cdots h_{n\,n+1} \\ q_{ln-1} \end{pmatrix} &= \kappa\Big[(d_l - c_l + 1) \prod_{j=2}^{l-1} (d_j - c_l + 1)/(c_j - c_l + 1) \\ &\quad \times \prod_{k=l+1}^{n} (c_k - c_l)(d_k - c_l + 1)\Big]^{1/2} \\ &= \kappa\Big\{\Big[\prod_{\lambda=2}^{n} (q_l - 1 - h_\lambda + \lambda - 1) \prod_{\lambda=l+1}^{n} (q_l - q_\lambda \\ &\quad + \lambda - l)\Big]\Big[\prod_{\lambda=1}^{l-1} (q_l - q_\lambda + \lambda - l - 1)\Big]^{-1}\Big\}^{1/2}, \end{aligned} \quad (2.29)$$

where $d_j = h_j - k + n$.

The raising operators for IU(n)$\supset$U(n) are the same as (2.14)–(2.17), with n replaced by $n+1$, and

$$C^{n+1}_i = I^{n+1}_i, \quad i = 1,2,\ldots n. \quad (2.30)$$

The normalization constant for the raising operators of IU(n)$\supset$U(n) is found to be

$$\begin{aligned} N^{(R)}_{u(n)}\begin{pmatrix} q_m \\ q_{m+1} \end{pmatrix} &= \kappa^*\Big\{-\Big[\prod_{\lambda=2}^{n} (q_m - h_\lambda + \lambda - m) \\ &\quad \times \prod_{\lambda=1}^{m-1} (q_m - q_\lambda + \lambda - m)\Big] \\ &\quad \times \Big[\prod_{\lambda=m+1}^{n} (q_m - q_\lambda + \lambda - m + 1)\Big]^{-1}\Big\}^{1/2} \end{aligned} \quad (2.31)$$

3. SHIFT OPERATORS OF IO(n) AND THEIR NORMALIZATION CONSTANTS

It turns out that the raising and lowering operators of O(n) obtained by Pang and Hecht[8] and Wong[9] cannot be immediately extended to IO(n), since they are not tensor operators. Fortunately, Bincer[5,6] has recently obtained shift operators for O(n) which are tensor operators. These operators are, moreover, expressed in simple, product form, and can be easily written down. We shall therefore express our results in the formalism obtained by Bincer.

We follow the notation used by Bincer, and note that the generators C_a^b are equivalent to $A\ B\ C\ D\ E\ F$ defined by Wong.[9] In fact, the relations between them are given by Wong and Yeh.[12] Note also that Bincer labels the irreducible representation in the order of

$$h_\nu, h_{\nu-1}, \ldots, h_2, h_1,$$

which is the reverse of the customary way of ordering.

The important point now is to realize that the translation generators in IO(n) behave like one-tensor operators in O(n). Therefore in IO($2\nu+1$) the translation operators are denoted by $I_\mu^{(2\nu+2)}$, $\mu = \nu, \nu-1, \ldots, 0, \ldots, -\nu$. In IO($2\nu$), we have $I_\mu^{(2\nu+1)}$, $\mu = \nu, \ldots, 1, -1, \ldots, -\nu$.

The first thing we do is to map IO(n) to O($n+1$), and discuss shift operators in O($n+1$), eventually changing the "one-tensor" generators of O(n) into translation generators.

Thus for IO(2ν), we look for shift operators of the form $V(\mu)_d$ with respect to O(2ν), where d takes the values $\nu, \nu-1, \ldots, 1, -1, \ldots, -\nu$. According to Bincer's method, we then have

$$V(\bar{\nu})_a = C_a^0, \tag{3.1}$$

where C_a^0 is a generator in O($2\nu+1$) corresponding to $E_{2\nu+1}^{\nu+a+1}$ or $F_{2\nu+1}^{\nu-a+1}$ according to Wong's notation. In the case of IO(2ν), then, we have

$$V(\bar{\nu})_a = I_a^{(2\nu+1)}, \tag{3.2}$$

where $I_a^{(2\nu+1)}$ behaves like $E_{2\nu+1}^{\nu+a+1}$ or $F_{2\nu+1}^{\nu-a+1}$ according to Wong's notation.

In the case of IO($2\nu-1$), we look for shift operators $V(\mu)_d$ with respect to O($2\nu-1$), where d takes the values ν, $\nu-1, \ldots, 2, 0, -2, \ldots, -\nu$. According to Bincer's method, we then have

$$V(\bar{\nu})_a = (C_a^1 - C_a^{-1})/\sqrt{2}, \tag{3.3}$$

where, for $a>0$, C_a^1 corresponds to $C_\nu^{\nu-a+1}$ and C_a^{-1} corresponds to $B_\nu^{\nu-a+1}$, while, for $a<0$, C_a^1 corresponds to $D_\nu^{\nu+a+1}$ and C_a^{-1} corresponds to $A_\nu^{\nu+a+1}$; for $a=0$

$$V(\bar{\nu})_0 = C_1^1, \tag{3.4}$$

where C_1^1 corresponds to the diagonal part of $J_{2\nu-1,2\nu}$. Therefore, for IO($2\nu-1$), we have

$$V(\bar{\nu})_a = I_a^{(2\nu)}, \tag{3.5}$$

where, for $a\neq 0$, $I_a^{(2\nu)}$ behaves like $(C_a^1 - C_a^{-1})/\sqrt{2}$, while, for $a=0$, $I_a^{(2\nu)}$ behaves like C_1^1.

Hence for the shift operators ${}^{2\nu+1}S_\mu$ in IO(2ν), we have

$${}^{2\nu+1}S_\mu = \left\{ V(\bar{\nu}) \prod_{j=\nu}^{\mu-1} (C - c_j^{2\nu}\mathbb{1}) \right\}_\mu^{(2\nu+1)}, \tag{3.6}$$

where

$$(VT)_\mu = \sum_c V_c T_\mu^c, \tag{3.7}$$

$$c_j^{2\nu} = m_j^{2\nu} + j + \nu - 2\theta_j, \tag{3.8}$$

$$\theta_j = \begin{cases} 1 & \text{if } j>0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.9}$$

with $V(\bar{\nu})_a = C_a^0$ replaced by $I_a^{(2\nu+1)}$ according to (3.2).

For the shift operators ${}^{2\nu}S_\mu$ in IO($2\nu-1$), we have

$${}^{2\nu}S_\mu = \left\{ V(\bar{\nu}) \prod_{j=\nu}^{\mu-1} (C - c_j^{2\nu-1}\mathbb{1}) \right\}_\mu^{(2\nu)}, \tag{3.10}$$

where

$$c_j^{2\nu-1} = m_j^{2\nu-1} + j + \ - 3\theta_j - \delta_j^0 \tag{3.11}$$

with $V(\bar{\nu})_a = (C_a^1 - C_a^{-1})/\sqrt{2}$ replaced by $I_a^{(2\nu)}$ for $a\neq 0$, and $V(\bar{\nu})_0 = C_1^1$ replaced by $I_0^{(2\nu)}$.

For the normalization constants of IO(n), we use basically Eq. (2.4) of Bincer.[6] From there we obtain

$$N\left\{ \begin{matrix} m_i^n \\ m_i^n - \delta_{\mu i} + \delta_{\bar{\mu} i} \end{matrix} \right\} = \prod_{j=\nu}^{\mu-1} (c_\mu - c_j) \langle \widehat{m}_i^n | V(\bar{\nu})_\mu | m_i^n \rangle. \tag{3.12}$$

Thus, for IO(2ν), we have

$$V(\bar{\nu})_\mu = C_\mu^0, \tag{3.13}$$

where

$$\begin{aligned} (\mu<0) \quad C_\mu^0 &= -iE_{2\nu+1}^{\nu+\mu+1} = -i(I_{2\nu+1}^{2\nu+2\mu+2} \\ &\quad + iI_{2\nu+1}^{2\nu+2\mu+2})/\sqrt{2}, \\ (\mu>0) \quad C_\mu^0 &= -iF_{2\nu+1}^{\nu-\mu+1} = -i(I_{2\nu+1}^{2\nu+2\mu+2} \\ &\quad - iI_{2\nu+1}^{2\nu-2\mu+1})/\sqrt{2}, \end{aligned} \tag{3.14}$$

Consequently for the normalization constant of the raising operator, $\mu<0$, we look for C_μ^0 which will raise $m_{2\nu,j}$ to $m_{2\nu,j}+1$ according to the Gelfand–Zetlin notation, where $j = \nu + \mu + 1$.

Now

$$\begin{aligned} C_\mu^0 &= -iE_{2\nu+1}^{\nu+\mu+1} = -iE_{2\nu+1}^{j} \\ &= -iD_{j+1}^{j} D_{j+2}^{j+1} \cdots D_\nu^{\nu-1} E_{2\nu+1}^{\nu}, \end{aligned} \tag{3.15}$$

where (3.15) is understood to act on the highest state of O(2ν).

It is clear that the matrix elements of $D_{j+1}^j D_{j+2}^{j+1} \cdots D_\nu^{\nu-1}$ must be the same for IO(2ν) and O($2\nu+1$). Therefore the difference can only come from $E_{2\nu+1}^{\nu} = (I_{2\nu+1}^{2\nu} + iI_{2\nu+1}^{2\nu-1})/\sqrt{2}$. The matrix elements for both these terms have been given explicitly by Wong.[9] There we find that they have a common factor $C_{2\nu+1}^{2\nu}$, defined by Eq. (87), which is the only term containing $l_{2\nu+1,i}$. We now calculate the term which should be missing from the normalization constant IO(2ν), i.e., the term containing $l_{2\nu+1,1}$ in $C_{2\nu+1}^{2\nu}$. This is the term

$$\begin{aligned} &(l_{2\nu+1\,1} - l_{2\nu j} - 1)(l_{2\nu+1\,1} + l_{2\nu j}) \\ &\quad = (m_{2\nu+1,1} - m_{2\nu j} + j - 1) \\ &\qquad \times (m_{2\nu+1,1} + 2\nu + m_{2\nu j} - j). \end{aligned} \tag{3.16}$$

Changing now to Bincer's notation, we have, for (3.16),

$$(m_\nu^{2\nu+1} - m_{\bar{\mu}}^{2\nu} + \nu + \mu)(m_\nu^{2\nu+1} + m_{\bar{\mu}}^{2\nu} + \nu - \mu - 1). \tag{3.17}$$

Now we look at the normalization constant of Bincer, and find that an identical term is found in Eq. (1.7) of Ref. 6, i.e.,

$$-(c_{\bar\mu}^{n-1}-c_{\bar\nu}^{n}+1)(c_{\bar\mu}^{n-1}-c_{\bar\nu}^{n}+1)$$
$$=-(m_{\bar\mu}^{2\nu}+\bar\mu+\nu+2\nu-2-2\nu-m_{\nu}^{2\nu+1}-2\nu+2)(m_{\bar\mu}^{2\nu}+\bar\mu+\nu-2-m_{\bar\nu}^{2\nu+1}+1)$$
$$=(m_{\nu}^{2\nu+1}-m_{\bar\mu}^{2\nu}+\nu+\mu)(m_{\nu}^{2\nu+1}+m_{\bar\mu}^{2\nu}+\nu-\mu-1)(n=2\nu+1) \tag{3.18}$$

which agrees with (3.17).

Thus we see that for IO(2ν), we can write the normalization constant as

$$\kappa\left\{\left[\prod_{j=\bar\nu}^{\mu-1}(c_{\bar\mu}^{2\nu}-c_j^{2\nu}+1)(c_{\mu}^{2\nu}-c_j^{2\nu})\prod_{j=\bar\nu+1}^{\nu-1}(c_{\bar\mu}^{2\nu}-c_j^{2\nu+1}+1)\right]\left[\prod_{s=\bar\nu}^{\nu}(c_{\bar\mu}^{2\nu}-c_s^{2\nu}+1)\right]^{-1}\right\}^{1/2}. \tag{3.19}$$

For IO($2\nu-1$), we have, for $\mu<0$,

$$V(\bar\nu)_\mu=(C_\mu^1-C_\mu^{-1})/\sqrt{2}=A_\nu^{\nu+\mu+1}+D_\nu^{\nu+\mu+1}=-i(I_{2\nu}^{2\nu+2\mu+2}+iI_{2\nu}^{2\nu+2\mu+1})=A_\nu^j+D_\nu^j,\quad j=\nu+\mu+1. \tag{3.20}$$

Now we write

$$A_\nu^j|m_i^{2\nu-1}\rangle=D_{j+1}^jD_{j+2}^{j+1}\cdots D_{\nu-1}^{\nu-2}A_\nu^{\nu-1}|m_i^{2\nu-1}\rangle, \tag{3.21}$$
$$D_\nu^j|m_i^{2\nu-1}\rangle=D_{j+1}^jD_{j+2}^{j+1}\cdots D_{\nu-1}^{\nu-2}D_\nu^{\nu-1}|m_i^{2\nu-1}\rangle. \tag{3.22}$$

Therefore we only need to look at the term $A_\nu^{\nu-1}+D_\nu^{\nu-1}$, or $(I_{2\nu}^{2\nu-2}+iI_{2\nu}^{2\nu-3})$.

Again both terms contribute a common factor $A_{2\nu}^{2\nu-1}$ defined by Eq. (85) of Ref. 9, which is the only term containing $l_{2\nu,i}$. Thus the term missing from the normalization constant of IO($2\nu-1$) is

$$(l_{2\nu 1}^2-l_{2\nu-1j}^2)=(m_1^{2\nu}-1+m_j^{2\nu-1}+2\nu-j)(m_1^{2\nu}-1-m_j^{2\nu-1}+j)$$
$$=(m_\nu^{2\nu}+m_{\bar\mu}^{2\nu-1}+\nu-\mu-2)(m_\nu^{2\nu}-m_{\bar\mu}^{2\nu-1}+\nu+\mu). \tag{3.23}$$

The contribution from Bincer's term is

$$-(c_{\bar\mu}^{n-1}-c_{\bar\nu}^{n}+1)(c_{\bar\mu}^{n-1}-c_{\nu}^{n}+1)=-(m_{\bar\mu}^{2\nu-1}+\bar\mu+\nu-3-m_{\bar\nu}^{2\nu}+1)(m_{\bar\mu}^{2\nu-1}+\bar\mu+\nu-m_\nu^{2\nu}-2\nu+2+1)$$
$$=(m_\nu^{2\nu}+m_{\bar\mu}^{2\nu-1}+\nu-\mu-2)(m_\nu^{2\nu}-m_{\bar\mu}^{2\nu-1}+\nu+\mu) \tag{3.24}$$

which agrees with (3.23).

Thus the normalization constants for both IO(2ν) and IO($2\nu-1$) can be expressed in the form of Bincer as

$$\kappa\left\{\left[\prod_{j=\bar\nu}^{\mu-1}(c_{\bar\mu}^{n-1}-c_j^{n-1}+1)(c_\mu^{n-1}-c_j^{n-1})\prod_{j=\bar\nu+1}^{\nu-1}(c_\mu^{n-1}-c_j^{n-1}+1)\right]\left[\prod_{s=\bar\nu}^{\nu}(c_{\bar\mu}^{n-1}-c_s^{n-1}+1)\right]^{-1}\right\}^{1/2}. \tag{3.25}$$

where

$$c_\mu^n=m_\mu^n+\mu+\nu+(n-2-2\nu)\theta_\mu+\left(\frac{n-1}{2}-\nu\right)\delta_{\mu 0}.$$

Finally, the normalization constant for the weight operator in IO($2\nu-1$) is just the diagonal matrix elements of $I_{2\nu,2\nu-1}$, multiplied by $(1-\nu)^\nu$, i.e.,

$$N_0^{(2\nu)}=\kappa\frac{\prod_{a=2}^{\nu}(l_{2\nu,a})}{\prod_{a=2}^{\nu-1}(l_{2\nu-1,a}-1)}(1-\nu)^\nu. \tag{3.26}$$

ACKNOWLEDGMENTS

We wish to thank Moorhead State University, especially Dr. W. B. Treumann and Dr. B.W. McCashland, for financial support of this work.

[1]M.K.F. Wong and H.Y. Yeh, J. Math. Phys **20**, 247 (1979).
[2]J. Patera, J. Math. Phys. **14**, 279 (1973).
[3]M.K.F. Wong, J. Math. Phys. **15**, 25 (1974).
[4]A.M. Bincer, J. Math. Phys. **18**, 1870 (1977).
[5]A.M. Bincer, J. Math. Phys. **19**, 1173 (1978).
[6]A.M. Bincer, J. Math. Phys. **19**, 1179 (1978).
[7]J.G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).
[8]S.C. Pang and K.T. Hecht, J. Math. Phys. **8**, 1233 (1967).
[9]M.K.F. Wong, J. Math. Phys. **8**, 1899 (1967).
[10]J.G. Nagel and M. Moshinsky, Rev. Mex. Fis. **14**, 29 (1965).
[11]A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968).
[12]M.K.F. Wong and H.Y. Yeh, J. Math. Phys. **16**, 1239 (1975).

Quasisteady primitive equations with associated upper boundary conditions

Paul Gordon

Associate Professor, Department of Mathematical Sciences, Purdue University, Calumet Campus, Hammond, Indiana

(Received 7 June 1977; revised manuscript received 25 January 1978)

This paper presents another approach to the problem of modeling large scale atmospheric flow. The major thrust of the method is to search for quasi-steady-state phenomena. This leads to sets of diagnostic and predictive equations that differ from those presently in use. Another important feature of the analysis is the introduction of a slowly floating upper boundary. In addition to simplifying the question of boundary conditions at the upper boundary, the floating top requires a highly significant change in the set of diagnostic variables. Two possible upper boundary conditions are derived in conjunction with the floating top. The first assumes continuous flow at the upper boundary, while the second assumes a compression-wave type discontinuity. Two specific criteria are formulated for checking the validity of the quasi-steady-state model. One is a scale assumption, between the physical scale and the time scale. The other is the requirement that the solution of the diagnostic equations be the steady-state limit of the original time-dependent equations. Various examples are given in order to attempt to clarify the techniques and philosophy of this approach. In addition, a specific test case is solved numerically with three models: The fixed top quasi-steady-state model, the floating top quasi-steady-state model, and a hydrostatic model. At the same time various upper boundary conditions are tested and compared. The results of the investigation indicate several significant advantages in favor of the floating top quasi-steady-state model.

I. INTRODUCTION

In many atmospheric problems, one encounters two natural difficulties in attempting to solve the hydrodynamic equations. The first is a scale problem (the height scale is much smaller than the horizontal). The second is a problem of boundary conditions (as in many problems of fluid dynamics, the lateral and top boundaries are free and consequently physical considerations do not give much insight regarding proper specification of boundary conditions). In this paper only the problem of upper boundary conditions will be considered. (The lateral boundary condition will be discussed in a later paper).

Several sets of "primitive" equations have been proposed in order to resolve the scale problem (see, for example, Ref. 1). In the present paper, this problem is resolved by analysis of quasisteady relationships. Essentially, one needs to find the dependent variables which are in quasisteady equilibrium with respect to the remaining variables. The quasisteady primitive equations thereby derived contain, as would be expected, two quasisteady diagnostic equations. One of these is in effect the usual hydrostatic equation. However, the derivation simultaneously produces a diagnostic equation for vertical velocity; this equation seems to be significantly different than those presently in use.

Given the correct time dependence at some point in the flow, the diagnostic variable is obtained, presumably to desired accuracy, by integrating the diagnostic equation vertically. The function, then, of the upper boundary condition, is to provide the time dependence. Two upper boundary conditions are derived in this paper; each of these is based on a specific assumption regarding time dependence. The following gives the organization of the paper.

Section II discusses the mathematical basis for a quasisteady system of equations. The scale assumption, upon which the quasisteady equations are based, is formulated. Also, an important criterion is formulated regarding steady-state limits: The solution of the quasi-steady equations should be the steady-state limit of the time-dependent equations. In Section II we also discuss briefly the fundamental differences between our model and other models presently being used in the study of atmospheric problems.

Section III discusses examples which are intended to clarify the use of quasisteady equations. In particular, we attempt to establish the importance of maintaining contact, mathematically, with the original time-dependent problem upon which the quasisteady equations are based. It is also shown, through simple example, that in general steady-state limits do not exist.

The derivations are given in Secs. IV, V, and VI. In Sec. IV we derive the quasisteady equations, both for a fixed coordinate system and for a floating top coordinate system. The floating top is introduced to minimize the number of boundary conditions required at the upper boundary. Section V gives the derivation of the continuous upper boundary condition; the assumption here is that the flow at the upper boundary remains continuous with a hypothetical flow existing above the boundary. Section VI gives the derivation of the discontinuous upper boundary condition; the assumption here is that the flow at the upper boundary produces a compression-wave type discontinuity with respect to the flow above.

Comparisons of the various models and boundary conditions are based essentially on a specific test case. Section VII defines this test problem and describes the numerical approximations that are used to obtain a solution.

The problem of convergence and stability of the numerical calculations is discussed in Sec. VIII. A simple criterion is derived by which a sequence of calculations can be examined for convergence. For a fixed top and constant pressure as the upper boundary condition, a solution is ob-

 0022-2488/79/040634-25$01.00

tained to the test case. The results of this calculation are then described in some detail.

This same test problem, with the same boundary conditions, is solved with a hydrostatic model developed by Kreitzberg. Section IX gives a comparison between the results of this hydrostatic calculation and our quasisteady calculation.

In Sec. X we attempt to study the difference between the quasisteady fixed top model and the quasisteady floating top model. In particular, a somewhat dramatic result is described: For one specific boundary condition (a physically realistic one), the solution to the fixed top model becomes unstable, while with the floating top analog of this boundary condition the solution to the floating top model remains well behaved.

The results obtained with the various upper boundary conditions are compared in Sec. XI. The discussion here is centered around the study of six specific questions. We attempted to choose these questions so as to reflect major points of interest and importance, both from a mathematical point of view and a physical point of view.

In Sec. XII we study the criterion regarding steady-state limits, namely that the solution to the quasisteady equations should be the steady-state limit of the complete time-dependent equations. One major result is that the floating top quasisteady model, in conjunction with the upper boundary conditions derived in Sec. V and VI, satisfies the criterion, while other commonly used upper boundary conditions do not.

In Sec. XIII two simple problems are solved with the quasisteady equations. Because the solutions are clearly physically unrealistic, we have an opportunity to examine several of the assumptions inherent in our derivations. One of these assumptions is that of the scale relationship (the height scale is much less than the length scale). Another of these assumptions, perhaps more subtle, is that an upper boundary condition can provide proper time dependence for the quasisteady variables.

Finally, in Sec. XIV an attempt is made to summarize the purpose, techniques, and results of the investigation.

II. THE MATHEMATICAL BASIS FOR QUASISTEADY PRIMITIVE EQUATIONS

This paper is concerned with the "primitive" equations, and associated boundary conditions, to be used for describing atmospheric flow. The initial assumption is that, for the problem under consideration, the standard hydrodynamic equations are suitable. Ignoring all external forces, except gravitational, one can write the two-dimensional equations in the form,

$$V_t = -AV_z - BV_x - F, \qquad (1)$$

$$A = \begin{bmatrix} w & 0 & \bar\rho & 0 \\ 0 & w & \dfrac{R\bar T}{c_v} & 0 \\ \dfrac{RT_0\bar T}{\bar\rho} & RT_0 & w & 0 \\ 0 & 0 & 0 & w \end{bmatrix},$$

$$B = \begin{bmatrix} u & 0 & 0 & \bar\rho \\ 0 & u & 0 & \dfrac{R\bar T}{c_v} \\ 0 & 0 & u & 0 \\ \dfrac{RT_0\bar T}{\bar\rho} & RT_0 & 0 & u \end{bmatrix},$$

$$F = \begin{bmatrix} 0 \\ 0 \\ g \\ 0 \end{bmatrix}, \quad V = \begin{bmatrix} \bar\rho \\ \bar T \\ w \\ u \end{bmatrix},$$

$$\bar\rho = \rho/\rho_0, \quad \bar T = T/T_0,$$

where ρ = density, T = temperature, w = vertical velocity, u = horizontal velocity, R = gas constant, c_v = specific heat, x = horizontal distance, z = vertical distance, t = time, g = acceleration due to gravity, P = pressure = ρRT, and ρ_0 and T_0 are reference values.

Suppose the problem is confined to a region of the xz plane of height h and length L. Assume further that appropriate boundary conditions are specified at all boundaries so that, together with Eq. (1), one has a well-posed mathematical problem, Let c_s = the speed of sound = $\sqrt{\gamma RT}$, where $\gamma = 1 + R/c_v$. In atmospheric problems for which $h \ll L$ one finds a natural stability criterion for many numerical methods to be essentially $\Delta t < \Delta z/c_s$, where Δz is the vertical mesh size. In many situations the features of the problem which are of interest to the scientist do not change significantly over such time scales. In our view, this is not a numerical problem, but rather constitutes an additional physical assumption, which is phrased as follows:

> Equation (1) is to be solved in a region of height h and length L, with $h \ll L$. It is assumed that boundary conditions and initial conditions to be imposed on Eq. (1) are such that the flow variables will experience significant variations only over time scales which are large compared to h/c_s. (2)

Mathematically, assumption (2) is interpreted as follows: Dependent variables, which can react to perturbation on a time scale of the order of h/c_s, are in quasisteady equilibrium with the slowly reacting dependent variables. The problem is to find these variables. Simultaneously, it is necessary to consider appropriate quasisteady limits of equations being used at the boundary.

Lateral boundary conditions, which are particularly troublesome, are currently being studied, and the results will be announced in a later paper. In the present paper the study of upper boundary conditions will consider the following questions:

(1) What are appropriate physical boundary conditions, and what is their mathematical representation?

(2) Are the boundary conditions consistent with the quasisteady assumption?

In regard to question (1), the upper boundary is considered an interface between two regions of flow. By allowing the upper boundary to "float," one achieves a situation in which one and only one boundary condition is required at the upper boundary. On the basis of two extreme physical assumptions, two upper boundary conditions are derived. The first physical assumption is that the flow remains continuous at the interface; the second physical assumption is that a compression-wave type discontinuity occurs above, but near, the interface. When used with a diagnostic equation, the purpose of the upper boundary condition is to provide the appropriate time dependence for the diagnostic variable. This aspect of the question is discussed further in Sec. XIII.

Question 2 is concerned with the existence of steady-state limits. Inherent in the use of a quasisteady equation is the thought that the solution of the quasisteady equations is in fact achieved as the steady-state limit of the time-dependent equations. This criterion is considered a necessary condition for any combination of upper boundary conditions and quasisteady equations. Since the dependent variables of the quasisteady equations satisfy wave equations, at least locally, it is perhaps surprising that actual steady-state limits can be achieved. Several simple examples, to illustrate the use of a quasisteady assumption and boundary conditions, are discussed in Sec. III.

This emphasis on the quasi-steady-state needs to be contrasted with existing techniques, presently used by atmospheric scientists, to resolve the problem of the sound speed. The hydrostatic approximation[2] is probably the most commonly used approach; it is easily justified on the basis of physical observation. Another widely used model is the anelastic set of equations; Ogura and Phillips[1] derived these equations with very specific reliance on an assumption regarding the time scale. The Boussinesq approximation is also prominent; Dutton and Fichtl[3] have justified this approximation on the basis of a scale analysis. In all of these models, one can demonstrate (for example, by a perturbation analysis) that the higher frequencies (that is, sound waves) have been eliminated or "filtered" from the solutions. There are basic differences, both philosophical and practical, between these models and the quasisteady approach. In the following, we attempt to discuss some of the these.

First, the author believes that Eq. (1) contains inherently only one set of fundamental frequencies, and these are the ones given by the characteristics. Relative to the atmospheric problem, then, the sound speed is the only pertinent frequency. This is not to say that the real atmosphere does not contain many other physically relevant scales; rather, the thought is that these scales are present because of initial conditions and boundary conditions, and not because of inherent properties of the equations. For example, when Eq. (1) is applied to a problem in aerodynamics, gravity waves are generally not encountered (even if the gravitational term is retained); this is due to the fact that in these problems the stratification in the initial data is negligible. The point is that the author does not believe that one can maintain mathematical consistency when purely physical frequencies are used to mathematically manipulate the equations. For example, in Ref. 1, the Brunt–Väisälä frequency is a fundamental parameter in the derivation. In the quasisteady analysis we will see that the entire derivation is based on assumption (2), in which only the sound speed is relevant.

Secondly, to the author's knowledge, none of the presently used sets of primitive equations maintains contact with the original initial-boundary value problem. This has severe implications:

(1) One can now attempt to specify boundary conditions without regard to mathematical considerations of Eq. (1). This undue flexibility leads to a proliferation and even overspecification of boundary conditions.

(2) Furthermore, if one desires to specify precisely the appropriate number of boundary conditions, presumably one would need somehow to make use of characteristic equations. But since the internal derivation has proceeded without regard to characteristics, there would seem to be no consistent rationale for dealing with characteristics at boundaries.

(3) An entirely different definition is given to the mathematical question of establishing a well-posed problem. It has been shown by Oliger and Sundström[4] that the hydrostatic meterological equations are "ill-posed with any specification of local, pointwise boundary conditions." Similar analysis would also show that the quasisteady equations are ill posed: Well posed implies continuous dependence on data, but a small perturbation (containing a high frequency) will destroy any possibility of a quasisteady state. It is our point of view that determination, of whether or not the problem is well posed, cannot be made without considering the original time-dependent problem.

III. EXAMPLES OF THE USE OF QUASISTEADY ASSUMPTIONS

A simple example in which assumption (2) would be applicable is given by the following:

$$\frac{\partial f}{\partial t} = c\frac{\partial f}{\partial z} \quad \text{for } 0 \leqslant z < L, \quad t > 0, \tag{3.1}$$

where c is a positive constant and initial and boundary conditions are $f(0,z) = \phi(z), f(t,L) = \psi(t)$.

In this equation the constant c plays the role of the speed of sound. Assumption (2) states that the boundary condition $\psi(t)$ changes slowly over time increments on the order of L/c, or $|\psi(t \pm L/c) - \psi(t)|$ is negligible for all t of interest. Thus, rather than following the time-dependent development of $f(t,z)$, a good approximation can be obtained by solving the quasisteady form of Eq. (3.1). In this case, the quasisteady form is $\partial f/\partial z = 0$, whose solution becomes (using the given boundary condition),

$$f(t,z) = \psi(t). \tag{3.2}$$

Equation (3.2) is to be contrasted with the analytic solution,

$$f(t,z) = \psi\left(t + \frac{z-L}{c}\right), \quad ct + z > L. \tag{3.3}$$

It is also of interest to solve the above problem through some type of scale analysis (as discussed in Sec. II with regard to primitive equations such as the hydrostatic equations). For this simple equation, one would obtain $\partial f/\partial z = 0$, and consequently the solution given by Eq. (3.2) would still be valid. However, there is one important difference: Assuming that the time-dependent problem is ignored, one could now specify $\psi(t) = f(t,0)$. The solution thereby obtained has no mathematical difficulties. However, this solution would be inconsistent with the time-dependent problem defined by Eq. (3.1), in which it is required that $\psi(t) = f(t,L)$.

For the second example, consider the following system of two equations, which occurs in the study of sound waves (Ref. 5, p. 245),

$$U_t = -AU_z, \quad U = \begin{pmatrix} u \\ p \end{pmatrix}, \quad A = \begin{pmatrix} 0 & c^2 \\ 1 & 0 \end{pmatrix}. \tag{3.4}$$

Here c is a constant, and the region is $0 \leqslant t$ and $0 < z < L$. Since the characteristics are determined by $dz/dt = \pm c$, one boundary condition is required at each boundary, say

$$u(t,L) = h_1(t), \quad p(t,0) = h_2(t). \tag{3.5}$$

If assumption (2) is satisfied, the quasisteady solution gives

$$u(t,z) \approx h_1(t), \quad p(t,z) \approx h_2(t). \tag{3.6}$$

As in the first example, assuming that the constant c is large, one might argue that $p_z \equiv 0$ is a good approximation. This would give

$$p_z \equiv 0, \quad u_z = -p_t. \tag{3.7}$$

Equation (3.7) is analogous to the hydrostatic equations; see Sec. IX, Eq. (27.1) There are important differences between Eq. (3.8) and the solution resulting from Eq. (3.7). First, Eq. (3.7) allows too much freedom in the choice of boundary conditions. Secondly, even if we correctly use Eq. (3.5), the solution to Eq. (3.7) would be

$$p(t,z) \equiv h_2(t), \quad u(t,z) = h_1(t) + (L-z)h_2'(t). \tag{3.8}$$

Equations (3.8) clearly differs from Eq. (3.6). Thirdly, there is no relationship between Eq. (3.8) and the steady-state solution of Eq. (3.4). This is particularly significant because steady-state solutions do not in general exist for Eq. (3.6). On this basis we must reject the solution given by Eq. (3.6), whereas one might still attempt to use Eq. (3.8).

The following shows that Eq. (3.4) does not in general have steady-state solutions. If $u(0,z) \equiv 0$, $p(0,z) \equiv P_0$, $h_1(t) \equiv 0$, and $h_2(t) \equiv P_0$, the solution is $p \equiv P_0$ and $u \equiv 0$. Thus, the steady-state solution is trivially maintained as the limit of the time-dependent process. Suppose the solution is perturbed by modifying h_1,

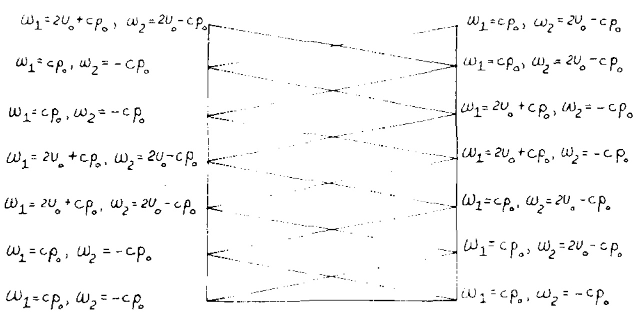
$$h_1(t) = \begin{cases} \dfrac{t}{\tau} U_0 & :0 < t < \tau, \\ U_0 & :t > \tau. \end{cases} \tag{3.9}$$

A steady-state solution is clearly given by $u \equiv U$ and $p \equiv P_0$. However, this steady-state solution is not achieved. This is well known and can be seen by introducing characteristic variables. Letting $W = QU$, where $W = \binom{\omega_1}{\omega_2}$ and $Q = \begin{pmatrix} 1 & c \\ 1 & -c \end{pmatrix}$, the solution can be written in the form

$$\omega_1(t,z) = \begin{cases} \omega_1(0, -ct + z) & :0 < -ct + z < L, \\ \omega_1\left(t - \dfrac{z}{c}, 0\right) & :ct - z > 0, \end{cases}$$

$$\omega_2(t,z) = \begin{cases} \omega_2(0, ct + z) & :0 < ct + z < L, \\ \omega_2\left(t - \dfrac{L-z}{c}, L\right) & :ct + z - L > 0. \end{cases} \tag{3.10}$$

This gives the following schematic:



The lines are intended to have slope $\pm 1/c$. From the data $\omega_1(0,z) = cp_0$ and $\omega_2(0,z) = -cp_0$. If $\tau = L/c$, the solutions are as indicated, and the steady-state solution would be $\omega_1 = U_0 + cp_0$ and $\omega_2 = U_0 - cp_0$. One notes that the time-dependent solution oscillates about the steady-state solution.

The third example is a modification of Eq. (3.4), so as to be somewhat closer in form to the atmospheric problem:

$$U_t = -AU_z - F, \quad U = \begin{pmatrix} u \\ p \end{pmatrix}, \quad F = \begin{pmatrix} g \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} \alpha & c \\ 1 & \alpha \end{pmatrix}, \tag{3.11}$$

where g, c, and α are constant with $|\alpha| \ll c$. Let initial conditions be $u(0,z) \equiv 0$ and $c^2 p_z + g \equiv 0$. Thus, $P(0,z) = P_0 - (g/c^2)(z - L)$, where p_0 is specified.

For $\alpha = 0$, the initial state is a steady-state solution. For α small, but nonzero, a steady-state solution still exists, namely,

$$u_z = \frac{\alpha g}{c^2 - \alpha^2}, \quad \text{or, if } u = 0 \text{ at } z = 0, \quad u = \frac{\alpha g}{c^2 - \alpha^2} z,$$

$$P_z = \frac{-g}{c^2-\alpha^2}, \quad \text{or, if} \quad P=P_0 \text{ at } z=L,$$

$$P = P_0 - \frac{g}{c^2-\alpha^2}(z-L).$$

Yet the steady-state solution is not the limit of the initial value problem. This can be seen in the same manner as in the previous example, although the situation now is considerably more complicated.

It is of course possible that the real atmosphere does contain small, but rapidly oscillating, vertical sound waves. If one believes that the solution, given by Eqs. (1) for a particular problem, does in fact exhibit such waves, then one expects that solutions obtained by diagnostic equations (such as given by the hydrostatic approximation) are in fact averages over long time scales. For ordinary differential equations a rather well-developed theory exists for such averaging. (See, for example Ref. 6 where the "stroboscopic method" is discussed in some detail.) In this paper the averaging approach is not accepted. Instead the criterion, that the quasisteady equations must have a steady-state solution which is the limit of the time-dependent process, will be tested. There are several reasons for this:

(1) There seems to be no mathematical basis in the atmospheric problem for an averaging assumption, as exists for the case of problems in ordinary differential equations; again, see Ref. 6. Equations (1) are a very stable set of equations. This is well-known heuristically and has been verified by the calculation of a local stability parameter.[7] Thus, if these wave motions exist because of perturbations in the initial data, then they should remain in the solution for only a short while. The only other possibility for generating such behavior is through boundary conditions. But assumption (2) precludes this possibility.

(2) There has been much concern in the literature regarding reflective behavior at boundaries. Thus, in many instances small amplitude, vertically propagating wave energy seems to reflect from the upper boundary back into the region of solution. A preliminary safeguard for any set of primitive equations, but particularly for the quasisteady equations, should be the existence of quasisteady-state solutions.

IV. THE DERIVATION OF THE QUASISTEADY EQUATIONS

Ideally, one would identify the quasisteady variables for Eq. (1) by introducing an analytic transformation,

$$W=\begin{pmatrix} f_1(V) \\ f_2(V) \\ f_3(V) \\ f_4(V) \end{pmatrix}=\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \end{pmatrix},$$

such that Eq. (1) transformed to

$$W_t = DW_z - \tilde{B}W_x - \tilde{F}, \tag{4}$$

where

$$D=\operatorname{diag}\{w,w,w+c_s,w-c_s\}.$$

The diagonal elements of D are the characteristics of the matrix A of Eq. (1). In Eq. (4) it is clear that significant variations of ω_3 and ω_4 occur on a time scale of the order of h/c_s. Consequently, in order to accommodate assumption (2) it would be necessary to treat ω_3 and ω_4 as quasisteady variables.

Although it is not possible to achieve Eq. (4), it is possible to find a transformation such that the matrix D of Eq. (4) will have a block diagonal form; i.e.,

$$D=\begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix},$$

where D_1, D_2, 0 are 2×2 matrices, and where D_1 has the characteristics $\{w,w\}$ while D_2 has the characteristic $\{w+c_s,w-c_s\}$. Under these conditions one would again conclude that, in order to incorporate assumption (2), the variables ω_3 and ω_4 are in quasisteady equilibrium with respect to ω_1 and ω_2.

A set of variables which achieves this block diagonal form is given by the following,

$$W=\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \end{pmatrix}=\begin{pmatrix} \Theta \\ u \\ \pi \\ w \end{pmatrix}, \tag{5.1}$$

where

$$\pi=\bar{p}^{[(\gamma-1)/\gamma]}, \quad \Theta=\frac{1}{\rho}\bar{p}^{(1/\gamma)}=\frac{T}{T_0}(\bar{p})^{-R/c_p},$$

$$\bar{p}=\overline{\rho T}, \quad \gamma=c_p/c_v.$$

A similar set of variables was used by Ogura and Phillips,[1] but apparently for different reasons. Equation (1) then takes the form,

$$W_t = -DW_z - \tilde{B}W_x - \tilde{F}, \tag{5.2}$$

where

$$D = \begin{pmatrix} w & 0 & 0 & 0 \\ 0 & w & 0 & 0 \\ 0 & 0 & w & (\gamma-1)\pi \\ 0 & 0 & \dfrac{c_s^2}{(\gamma-1)\pi} & w \end{pmatrix},$$

$$\tilde{B} = \begin{pmatrix} u & 0 & 0 & 0 \\ 0 & u & \dfrac{c_s^2}{(\gamma-1)\pi} & 0 \\ 0 & (\gamma-1)\pi & u & 0 \\ 0 & 0 & 0 & u \end{pmatrix},$$

$$F = \begin{pmatrix} 0 \\ 0 \\ 0 \\ g \end{pmatrix}, \quad c_s^2 = \Theta\pi c_{s_0}^2, \quad c_{s_0}^2 = \gamma R T_0.$$

We now assume that the horizontal scale is so large that the speed of sound in the matrix $\tilde{B}$ does not adversely affect the time scale. The mathematical analog of assumption (2), then, is that π and w are in quasisteady equilibrium with respect to Θ and u. The proposed system of equations consists of the following:

predictive equations:

$$\frac{d\Theta}{dt} = 0, \tag{6.1}$$

$$\frac{du}{dt} = -\frac{c_s^2}{(\gamma-1)\pi}\frac{\partial\pi}{\partial x}, \tag{6.2}$$

diagnostic equations:

$$(\gamma-1)\pi\left(\frac{\partial w}{\partial z} + \frac{\partial u}{\partial x}\right) + \left(w\frac{\partial\pi}{\partial z} + u\frac{\partial\pi}{\partial x}\right) = 0 \tag{6.3}$$

$$\frac{c_s^2}{(\gamma-1)\pi}\frac{\partial\pi}{\partial z} + \left(w\frac{\partial w}{\partial z} + u\frac{\partial w}{\partial x}\right) + g = 0. \tag{6.4}$$

Note that Eq. (6.4) is essentially the usual hydrostatic approximation. Equation (6.3) will be compared in Sec. IX with Kreitzberg's diagnostic equation for vertical velocity.[15] Several comments are perhaps in order regarding the equation set (6.1)–(6.4):

(i) From the point of view by which these equations were derived, one must conclude that it is not correct to use the hydrostatic approximation by itself. In fact, both Eqs. (6.3) and (6.4) must be used simultaneously.

(ii) It would not be correct to use Eq. (6.3) and (6.4) with arbitrary predictive equations, say for example the temperature equation. For in fact the predictive equations must be such that they do not contain information which can propagate vertically at the speed of sound.

(iii) A perturbation analysis, about a constant base flow, has been performed on Eqs. (6). The resulting frequencies are essentially the same as in the case of the hydrostatic equation.[2]

The following coordinate system will also be used:

$$\tau = t, \quad \eta = \frac{x}{L}, \quad \zeta = \frac{z - f_1(t,x)}{f_2(t,x) - f_1(t,x)}. \tag{7}$$

A solution to Eq. (1) is then to be obtained in the region $0 \leqslant \eta \leqslant 1$, $0 \leqslant \zeta \leqslant 1$, $\tau > 0$. (This transformation has been used previously; see Ref. 8.) $z = f_2(t,x)$ represents the top of the region. As discussed in Sec. V, this boundary will "float" so as to always remain as a streamline of the flow. Thus, the vertical region covered by $0 \leqslant \zeta \leqslant 1$ will be the maximum region in which the solution is determined by initial data and one upper boundary condition. Also, f_1 is included to allow nonuniform terrain. After introducing Eq. (7) the characteristics involving the sound speed are no longer uncoupled from the remainder of the system. The following transformation is used to obtain an uncoupled form,

$$\tilde{u}=u-aw, \quad \tilde{w}=w+au, \tag{8.1}$$

where

$$a=\zeta_x/\zeta_z.$$

Equation (5.2) then becomes,

$$U_\tau=-A_1U_\xi-B_1U_\eta-F_1, \tag{8.2}$$

$$A_1=\begin{pmatrix} r & 0 & 0 & 0 \\ 0 & r & 0 & 0 \\ 0 & 0 & r & (\gamma-1)\pi\zeta_z \\ 0 & 0 & \dfrac{c_s^2 b}{(\gamma-1)\pi} & r \end{pmatrix},$$

$$B_1=\begin{pmatrix} u/L & 0 & 0 & 0 \\ 0 & u/L & \dfrac{c_s^2}{(\gamma-1)\pi L} & 0 \\ 0 & \dfrac{(\gamma-1)\pi}{L(1+a^2)} & u/L & \dfrac{(\gamma-1)\pi a}{L(1+a^2)} \\ 0 & 0 & \dfrac{ac_s^2}{(\gamma-1)\pi L} & u/L \end{pmatrix},$$

$$U=\begin{pmatrix} \Theta \\ \tilde{u} \\ \pi \\ \tilde{w} \end{pmatrix}, \quad F_1=\begin{pmatrix} 0 \\ h_1 \\ h_2 \\ h_3 \end{pmatrix}, \quad b=\zeta_z(1+a^2), \quad r=\zeta_t+w\zeta_z+u\zeta_x,$$

$$h_1=-ag+w\left(a_\tau+ra_\xi+\frac{u}{L}a_\eta\right),$$

$$h_2=-(\gamma-1)\pi\zeta_z ua_\xi+\frac{(\gamma-1)\pi}{L}\left[\tilde{w}\left(\frac{a}{1+a^2}\right)_\eta+\tilde{u}\left(\frac{1}{1+a^2}\right)_\eta\right],$$

$$h_3=g-u\left(a_\tau+ra_\xi+\frac{u}{L}a_\eta\right)$$

From the matrix A_1 it can be seen that only the equations for $\tilde{w}$ and π involve the sound speed in the vertical direction. As before, these are taken to be the quasisteady variables. The quasisteady primitive equations are then as follows:

$$\Theta_\tau=-r\Theta_\xi-\frac{u}{L}\Theta_\eta, \tag{9.1}$$

$$\tilde{u}_\tau=-r\tilde{u}_\xi-\frac{u}{L}\tilde{u}_\eta-\frac{c_s^2}{(\gamma-1)\pi L}\pi_\eta-h_1, \tag{9.2}$$

$$0=r\pi_\xi+(\gamma-1)\pi\zeta_z\tilde{w}_\xi+\frac{(\gamma-1)\pi}{L(1+a^2)}\tilde{u}_\eta+\frac{u}{L}\pi_\eta+\frac{(\gamma-1)\pi a}{L(1+a^2)}\tilde{w}_\eta+h_2, \tag{9.3}$$

$$0=\frac{c_s^2 b}{(\gamma-1)\pi}\pi_\xi+r\tilde{w}_\xi+\frac{ac_s^2}{(\gamma-1)\pi L}\pi_\eta+\frac{u}{L}\tilde{w}_\eta+h_3. \tag{9.4}$$

Thus, Eqs. (9.1) and (9.2) are predictive equations for Θ and $\tilde{u}$, while Eqs. (9.3) and (9.4) are diagnostic equations for π and $\tilde{w}$.

V. DERIVATION OF THE CONTINUOUS UPPER BOUNDARY CONDITION

In this section we derive an upper boundary condition which displays the time dependence required if the solution is to remain continuous at the upper boundary. Consider first Eq. (1). It is assumed that at $z=z_{\rm top}$ the number of required boundary conditions is determined by the eigenvalues of the matrix A.[9] The eigenvlaues are $w,w,w\pm c_s$. Thus, at $z=z_{\rm top}$, one boundary condition is specified if $w\geqslant 0$ and three are specified if $w<0$. Mathematically, this can be considerably simplified by forcing $z=z_{\rm top}$ to be a streamline. This is accomplished by the transformation (7), where the following condition is imposed at $\zeta=1$,

$$(f_2)_t=v_n=w+\frac{\zeta_x}{\zeta_z}u=\text{normal component of velocity.} \qquad (10)$$

Boundary conditions at $\zeta=1$ are determined by the eigenvalues of the matrix A_1 of Eq. (8.2). These are $r,r,r\pm[c_s^2(\zeta_x^2+\zeta_z^2)]^{1/2}$. From Eq. (10), $r=0$ at $\zeta=1$. Thus, at $\zeta=1$, one and only one boundary condition needs to be specified for Eq. (8.2).

The goal now is to derive an upper boundary condition under the constraint that the flow above $\zeta=1$ be continuous at $\zeta=1$ with the solution obtained with the assumption of diagnostic equations in $0\leqslant\zeta\leqslant 1$. Return now to Eq. (8.2), before any quasisteady assumptions have been made. Suppose that in a region above $\zeta=1$ the equation are being solved simultaneously with the region below $\zeta=1$. Then, at $\zeta=1$, one characteristic variable will propagate information from below $\zeta=1$ to $\zeta\geqslant 1$, while another will propagate information from above $\zeta=1$ to $\zeta\leqslant 1$. These two characteristic variables can be obtained locally (that is, in a linear form) as follows. First, from Eq. (8.2) put the π and $\tilde{w}$ equations in the from

$$\pi_\tau=-r\pi_\zeta-(\gamma-1)\pi\zeta_z\tilde{w}_\zeta-H_1,$$
$$\tilde{w}_\tau=-\frac{c_s^2 b}{(\gamma-1)\pi}\pi_\zeta-r\tilde{w}_\zeta-H_2, \qquad (11)$$

where H_1 and H_2 represent the remaining terms. Let

$$v_1=[c_s(1+a^2)_0^{1/2}]\pi+(\gamma-1)\pi_0\tilde{w},$$
$$v_2=[-c_s(1+a^2)_0^{1/2}]\pi+(\gamma-1)\pi_0\tilde{w}, \qquad (12)$$

where the subscript indicates that the quantity is evaluated at a base point, which for our purposes lies on the surface $\zeta=1$. Then,

$$(v_1)_\tau=-\lambda_1(v_1)_\zeta-[c_s(1+a^2)_0^{1/2}]H_1-(\gamma-1)\pi_0H_2, \qquad (13.1)$$
$$(v_2)_\tau=-\lambda_2(v_2)_\zeta+[c_s(1+a^2)_0^{1/2}]H_1-(\gamma-1)\pi_0H_2, \qquad (13.2)$$

where

$$\lambda_1=[r+c_sc_\zeta(1+a^2)_0^{1/2}] \text{ and } \lambda_2=[r-c_s\zeta_z(1+a^2)_0^{1/2}].$$

Insofar as the region $0\leqslant\zeta\leqslant 1$ is concerned, the equations being used at $\zeta=1$ are,

(a) the two equations for Θ and $\tilde{u}$ given by Eq. (8.2); that is, Eqs. (9.1) and (9.2);

(b) Equation (13.1) for v_1;

(c) $(v_2)_\tau$ is in effect a boundary condition, which is specified from the solution above $\zeta=1$.

Next, assume the validity of the quasisteady equations, that is, Eq. (9), for $0\leqslant\zeta\leqslant 1$, but continue to use the full equations for $\zeta>1$. Again, insofar as the region $0\leqslant\zeta\leqslant 1$ is concerned, the equations being used at $\zeta=1$ are as follows:

(a) Equations (9.1) and (9.2) to predict Θ and $\tilde{u}$,

(b) $(v_2)_\tau$ is still in effect a boundary condition.

Finally, assume that the solution thereby obtained (for $0\leqslant\zeta\leqslant 1$ and for $\zeta>1$) will be continuously differentiable at $\zeta=1$. Note that the right side of Eq. (13.2) is a linear combination of the right side of Eqs. (9.3) and (9.4). But since Eqs. (9.3) and (9.4) are now valid for all time, then at $\zeta=1$ the right side of Eq. (13.2) will be zero.

The derivation can be summarized as follows:

Assumption: The number of boundary conditions required for Eqs. (8.2) at $\zeta=1$ is determined by the eigenvalues of the matrix A_1. (14.1)

Assumption: Let a streamline determine two regions of a differentiable solution of Eqs. (8.2). Then, the interaction between the two regions is determined by the appropriate characteristic variable. (14.2)

Assumption: Given an initial state, then for some time period of interest, Eqs. (8.2) can be solved in some region $\zeta>1$ simultaneously with Eqs. (9) for $0\leqslant\zeta\leqslant 1$ to produce a continuously differentiable solution. (14.3)

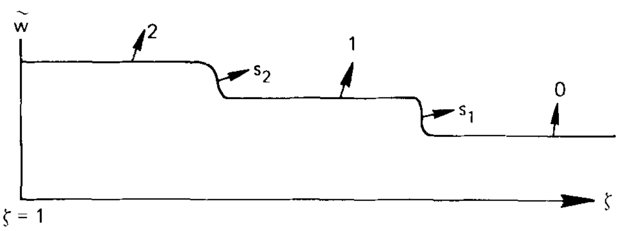
Conclusion: The same solution can be obtained for $0\leqslant\zeta\leqslant 1$, without obtaining the solution for $\zeta>1$, by using at $\zeta=1$ the boundary condition $(v_2)_\tau=0$, or

$$\frac{1}{\pi}\pi_\tau=\frac{(\gamma-1)}{c_s(1+a^2)^{1/2}}\tilde{w}_\tau. \qquad (15)$$

VI. DERIVATION OF THE DISCONTINUOUS UPPER BOUNDARY CONDITION

This section considers the possibility that the flow, when $\tilde{w}>0$ at $\zeta=1$, becomes discontinuous. It is assumed that the flow above $\zeta=1$ is in quasisteady equilibrium with the solution below. Thus, a positive $\tilde{w}$ at $\zeta=1$ represents a weak compression wave moving into the upper region. Within the limits of quasisteady analysis, such one-dimensional waves will form discontinuities which will propagate into the region above $\zeta=1$ very rapidly, namely at the sound speed. At a later time it is assumed that a "new" discontinuity is formed, but again at $\zeta=1$. Thus, without attempting to follow the propagation of this discontinuity, one can calculate the jump conditions at $\zeta=1$.

Assuming an oblique compression wave,[10] the appropriate equations would be the usual Rankine–Hugoniot jump conditions plus the additional requirement that $\tilde{u}$ be continuous at $\zeta=1$. In order to use these conditions, one requires that one variable be known behind the jump and all variables be known in front of the jump. In our case, however, both Θ and $\tilde{w}$ are known behind the jump. This implies that our problem is not that of a "piston" moving into a flow, but rather of a mass injection problem, the point of injection being at $\zeta=1$. For such a problem it is appropriate to specify two conditions at $\zeta=1$, say Θ and $\tilde{w}$, while π would then be calculated.[11] The quasisteady solution to this problem appears qualitatively as follows (for a fixed t):



There are in fact two discontinuities, with Rankine–Hugoniot conditions holding over each and with each moving at different velocities. (This result has been verified by unpublished numerical calculations.)

In the following equations, subscripts will refer to the positions $\{0,1,2\}$ noted in the above diagram. The Rankine–Hugoniot conditions over the jump at s_1 are as follows:

$$v_0^2+\frac{\gamma+1}{2}(\Delta w_1)v_0-c_0^2=0, \tag{16.1}$$

$$\rho_1/\rho_0=v_0/v_1, \tag{16.2}$$

$$\rho_0 v_0^2-\rho_1 v_1^2=P_1-P_0, \tag{16.3}$$

where $v_0=w_0-U_1$, $v_1=w_1-U_1$, $\Delta w_1=w_1-w_0$, and U_1 is the velocity of the discontinuity at s_1. Over the jump at s_2 the equations are as follows:

$$V_1^2+\frac{\gamma+1}{2}(\Delta w_2)V_1-c_1^2=0, \tag{16.4}$$

$$\rho_2/\rho_1=V_1/V_2, \tag{16.5}$$

$$\rho_1 V_1^2-\rho_2 V_2^2=P_2-P_1, \tag{16.6}$$

where $V_1=w_1-U_2$, $V_2=w_2-U_2$, $\Delta w_2=w_2-w_1$, and U_2 is the velocity of the discontinuity at s_2.

There are 11 quantities to be determined, namely $\{\rho_0, P_0, w_0, \rho_1, P_1, w_1, \rho_2, P_2, w_2, U_1, U_2\}$. Of these, five are considered to be known, namely $\{\rho_0, P_0, w_0, w_2, \Theta_2\}$. The six equations of Eqs. (16) are to be used to determine the remaining variables. (Of these variables only P_2 is of actual interest.) A solution to second order accuracy will be obtained by expanding about terms of the form w/c. In the derivation below, equality will imply second order accuracy. The following series expansion will be used:
Let

$$f(\epsilon,\delta)=1+a_1\epsilon+a_2\delta+a_3\epsilon^2+a_4\epsilon\delta+a_5\delta^2,$$

Then,

$$f^\alpha=1+b_1\epsilon+b_2\delta+b_3\epsilon^2+b_4\epsilon\delta+b_5\delta^2,$$

where

$$b_1=\alpha a_1,\quad b_2=\alpha a_2,\quad b_3=\alpha\left(a_3+\frac{\alpha-1}{2}a_1^2\right),$$

$$b_4=\alpha[(\alpha-1)a_1a_2+a_4],\quad b_5=\alpha\left(a_5+\frac{\alpha-1}{2}a_2^2\right). \tag{17}$$

Using this expansion, Eqs. (16) take the following form:

$$\frac{v_0}{c_0}=-\left[1+\left(\frac{\gamma+1}{4}\right)\frac{\Delta w_1}{c_0}+\frac{1}{2}\left(\frac{\gamma+1}{4}\right)^2\left(\frac{\Delta w_1}{c_0}\right)^2\right], \tag{18.1}$$

$$\frac{\rho_1}{\rho_0}=1+\frac{\Delta w_1}{c_0}+\frac{3-\gamma}{4}\left(\frac{\Delta w_1}{c_0}\right)^2, \tag{18.2}$$

$$\frac{P_1}{P_0}=1+\gamma\frac{\Delta w_1}{c_0}+\frac{\gamma(\gamma+1)}{4}\left(\frac{\Delta w_1}{c_0}\right)^2, \tag{18.3}$$

$$\frac{V_1}{c_1}=-\left[1+\left(\frac{\gamma+1}{4}\right)\frac{\Delta w_2}{c_1}+\frac{1}{2}\left(\frac{\gamma+1}{4}\right)^2\left(\frac{\Delta w_2}{c_1}\right)^2\right], \tag{18.4}$$

$$\frac{\rho_2}{\rho_1}=1+\frac{\Delta w_2}{c_1}+\frac{3-\gamma}{4}\left(\frac{\Delta w_2}{c_1}\right)^2, \tag{18.5}$$

$$\frac{P_2}{P_1}=1+\gamma\frac{\Delta w_2}{c_1}+\frac{\gamma(\gamma+1)}{4}\left(\frac{\Delta w_2}{c_1}\right)^2. \tag{18.6}$$

Noting that

$$\frac{c_0^2}{c_1^2}=\frac{P_0}{P_1}\frac{\rho_1}{\rho_0},$$

one obtains

$$\frac{c_0}{c_1}=1-\frac{\gamma-1}{2}\left(\frac{\Delta w_1}{c_0}\right)+\left(\frac{\gamma-1}{2}\right)^2\left(\frac{\Delta w_1}{c_0}\right)^2. \tag{18.7}$$

Then, using Eqs. (18.6), (18.7), and (18.3), and letting $\Delta w=\Delta w_1+\Delta w_2=w_2-w_0$, one obtains

$$\frac{P_2}{P_0}=1+\gamma\frac{\Delta w}{c_0}+\frac{\gamma(\gamma+1)}{4}\left(\frac{\Delta w}{c_0}\right)^2. \tag{18.8}$$

Note that, to second order, P_2/P_0 depends only on Δw and not on the other known upstream condition $\Delta\Theta$. (This agrees with the known result that, in the case of a simple compression wave, $\Delta\Theta$ is third order in Δu.)

Finally, using

$$\frac{\pi_2}{\pi_0}=\left(\frac{P_2}{P_0}\right)^{(\gamma-1)/\gamma}$$

and letting $\Delta\pi=\pi_2-\pi_0$, one obtains to first order,

$$\frac{\Delta\pi}{\pi_0}=(\gamma-1)\frac{\Delta w}{c_0}. \tag{18.9}$$

Noting the quasisteady environment, Eq. (18.9) can be written as

$$\frac{1}{\pi}\pi_\tau=\frac{(\gamma-1)}{c_s}\tilde{w}_\tau. \tag{19}$$

Equation (19) is the proposed upper boundary condition. In form, it is identical to Eq. (15) for the continuous case, except for the factor $(1+a^2)^{1/2}$. However, proper interpretation must be given to the use of Eq. (19).

First, if $\tilde{w}<0$ at $\zeta=1$, the proper analogy would be to an expansion wave. This does not give a discontinuity and it would not seem reasonable to attempt to apply the above analysis to such situations.

Secondly, recall that $\zeta=1$ is a moving interface in the fluid. If pressure at $\zeta=1$, $P(\tau,\eta,1)$, were unaffected by the flow, then $P(\tau,\eta,1)$ would change with time so as to attain the "initial" quasisteady pressure at that point. The appropriate equation at $\zeta=1$ would be the usual hydrostatic equation,

$$\frac{\partial \pi}{\partial z} = \frac{-g(\gamma-1)\pi}{c_s^2}.$$

At $\zeta = 1$, $dz/d\tau = f_2'(\tau) = \tilde{w}$. Thus, the above equation becomes

$$\pi_\tau = -\frac{g(\gamma-1)\pi}{c_s^2}\tilde{w}. \tag{20}$$

Remark: This is essentially the boundary condition used by Pielke and Mahrer.[8]

When using Eq. (15), the continuous boundary condition, the interpretation is that the equation gives the entire pressure change, regardless of the fact that $\zeta = 1$ may be a moving surface. On the other hand, the discontinuous boundary condition, Eq. (19), is intended to represent a perturbation of the undisturbed flow. Thus, the quasisteady pressure change needs to be calculated and the result of Eq. (19) superimposed.

Summarizing, with the assumption of a jump discontinuity at $\zeta = 1$ in the presence of $\tilde{w} > 0$, and with the assumption of continuous flow if $\tilde{w} < 0$, the following upper boundary condition is proposed,

$$\pi_\tau = \begin{cases} \dfrac{\pi(\gamma-1)}{c_s(1+a^2)^{1/2}}\tilde{w}_\tau : \tilde{w} \leqslant 0, \\[2ex] \dfrac{\pi(\gamma-1)}{c_s}\tilde{w}_\tau - \dfrac{g(\gamma-1)}{c_s^2}\tilde{w} : \tilde{w} > 0. \end{cases} \tag{21}$$

VII. SPECIFICATION OF A TEST PROBLEM AND NUMERICAL APPROXIMATIONS

In the previous sections we have derived the quasisteady equations and several possible upper boundary conditions. It is necessary now to evaluate these equations and to compare their solutions with solutions obtained with other techniques presently being used in atmospheric science. As a first step in this direction, it was decided to choose one test problem and to solve it for various choices of the mathematical model.

A severe complication in such a study is the strong effect of lateral boundary conditions. The difficulty stems from the fact that in general all required boundary conditions are not known physically. For purposes of the present paper, it was desired to choose lateral boundary conditions which completely determine the flow. This then removes any difficulties that may arise from imperfectly specified lateral boundary conditions. The following were chosen,

$$u(t,0,z) = 0, \tag{22.1}$$

$$u(t,L_1,z) = 0. \tag{22.2}$$

A further boundary condition imposed for all cases is,

$$w(t,x,0) = 0. \tag{22.3}$$

It was decided to begin with stationary flow and a rectangular region $0 \leqslant x \leqslant L_1, 0 \leqslant z \leqslant L_2$. The following were chosen:

$$w(0,x,z) = 0, \tag{22.4}$$

$$u(0,x,z) = 0, \tag{22.5}$$

$$f_2(0,x) = L_2, \tag{22.6}$$

$$P(0,x,L_2) = P_0 = \text{const.} \tag{22.7}$$

L_1 was taken as 480 km and L_2 as 10 km. Given the initial Θ distribution, $P(0,x,z)$ is obtained from Eq. (9.4), which is the hydrostatic equation for these conditions. Thus, the driving force for this problem is the initial Θ distribution, which was chosen as,

$$\Theta(0,x,z) = \Theta(0,0,z)[1 + f(x,z)], \tag{22.8}$$

where

$$f(x,z) = \alpha\left(3 - \frac{2x}{L_1}\right)\left(\frac{x}{L_1}\right)^2\left(\frac{4z}{L_2}\right)\left(1 - \frac{z}{L_2}\right)^2,$$

$$\Theta(0,0,z) = \alpha_0 + \alpha_1\left(\frac{z}{L_2} - 1\right) + \alpha_2\left(\frac{z}{L_2} - 1\right)^2 + \alpha_3\left(\frac{z}{L_2} - 1\right)^3,$$

$\alpha_0 = 1.267, \quad \alpha_1 = 0.5197, \quad \alpha_2 = 0.3884,$

$\alpha_3 = 0.0857, \quad \alpha = 0.05.$

Remark: From Eq. (22.8),

$$f(0,z) = f(x,(0) = \frac{\partial \Theta}{\partial x}(0,z) = \frac{\partial \Theta}{\partial x}(L_1,z) = \frac{\partial \Theta}{\partial x}(x,0) = \frac{\partial \Theta}{\partial x}(x,L_2) = \frac{\partial^2 \Theta}{\partial x \partial z}(x,L_2) = 0.$$

This produces an initial field as shown in Fig. 1 and 2. Note that Θ is constant at $z = 0$ and $z = 10$ km. Pressure is constant at $z = 10$ km, but for lower values of z there is a pressure gradient (see Fig. 2). This should produce a flow which moves from left to right. Because $u = 0$ at both lateral boundaries, the wave (at some later time) will reflect and move from right to left. Physically this wave should dampen out with time and the flow should return to a stationary flow. However, our model does not contain the viscous mechanism that are usually associated with such behavior. Nevertheless, as we examine the results, the damping effect on the wave will be of particular interest.

Although the physical problem is not of direct interest to atmospheric scientists, because of the reflective behavior at the lateral boundaries, the scales and magnitudes of the variables are comparable to those encountered in real problems.

To complete specification of the mathematical problem, it is necessary to choose upper boundary conditions. This will be discussed later.

The mathematical problem is defined by Eqs. (9), Eqs. (22), and the yet-to-be specified upper boundary conditions. In choosing a numerical scheme, the primary concern was to obtain a solution; questions of efficiency and degree of accuracy were not considered important for this investigation. The author had previously obtained solutions to Eqs. (1)

FIG. 1. Initial flow field.

$\left(\frac{z}{10\text{ km.}}\right)$													
								$\theta(0,x,z)$					
1.	1.2670	1.2670	1.2670	1.2670	1.2670	1.2670	1.2670	1.2670	1.2670	1.2670	1.2670	1.2670	1.2670
0.75	1.1600	1.1602	1.1608	1.1617	1.1628	1.1641	1.1654	1.1668	1.1681	1.1692	1.1701	1.1707	1.1709
0.50	1.0935	1.0941	1.0956	1.0978	1.1006	1.1038	1.1072	1.1106	1.1138	1.1166	1.1189	1.1203	1.1209
0.25	1.0595	1.0601	1.0618	1.0642	1.0673	1.0708	1.0744	1.0781	1.0816	1.0847	1.0871	1.0888	1.0893
0.	1.0500	1.0500	1.0500	1.0500	1.0500	1.0500	1.0500	1.0500	1.0500	1.0500	1.0500	1.0500	1.0500
							$\pi(0,x,z)$						
1.	0.68721	0.68721	0.68721	0.68721	0.68721	0.68721	0.68721	0.68721	0.68721	0.68721	0.68721	0.68721	0.68721
0.75	0.76093	0.76092	0.76090	0.76087	0.76083	0.76079	0.76075	0.76070	0.76066	0.76062	0.76060	0.76058	0.76057
0.50	0.84023	0.84020	0.84010	0.83996	0.83978	0.83958	0.83937	0.83916	0.83896	0.83878	0.83864	0.83855	0.83852
0.25	0.92318	0.92311	0.92289	0.92257	0.92217	0.92171	0.92123	0.92076	0.92031	0.91992	0.91960	0.91940	0.91932
0.	1.00783	1.00773	1.00745	1.00703	1.00651	1.00592	1.00529	1.00467	1.00409	1.00359	1.00318	1.00291	1.00282
							$P(0,x,z)$						
1.	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039
0.75	384.325	384.313	384.278	384.226	384.160	384.086	384.008	383.930	383.856	383.792	383.740	383.706	383.694
0.50	543.744	543.666	543.452	543.129	542.726	542.271	541.791	541.314	540.866	540.473	540.160	539.954	539.879
0.25	755.976	755.753	755.139	754.216	753.064	751.765	750.398	749.040	747.767	746.653	745.768	745.185	744.974
0.	1027.670	1027.310	1026.316	1024.822	1022.959	1020.858	1018.650	1016.457	1014.403	1012.604	1011.177	1010.236	1009.897
							$T(0,x,z)$						
1.	237.831	237.831	237.831	237.831	237.831	237.831	237.831	237.831	237.831	237.831	237.831	237.831	237.831
0.75	241.105	241.147	241.264	241.440	241.661	241.912	242.178	242.444	242.694	242.916	243.092	243.209	243.251
0.50	250.977	251.090	251.403	251.876	252.468	253.141	253.853	254.565	255.237	255.829	256.302	256.614	256.727
0.25	267.183	267.308	267.655	268.178	268.835	269.579	270.367	271.156	271.900	272.555	273.078	273.424	273.549
0.	289.053	289.024	288.944	288.824	288.674	288.504	288.326	288.148	287.982	287.836	287.720	287.643	287.616
$\left(\frac{x}{480\text{ km.}}\right)$	0.	0.08333	0.16667	0.25	0.33333	0.41667	0.50	0.58333	0.66667	0.75	0.83333	0.91667	1.

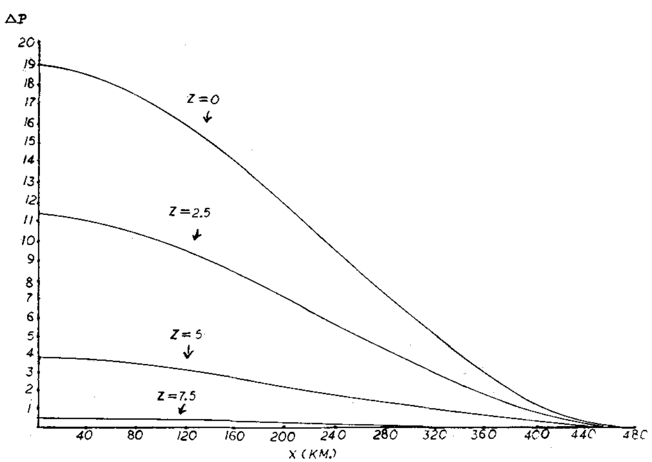


FIG. 2. Over-pressure at initial time, $\Delta P = P(0,x,z) - P(0,480,z)$.

with a first order method,[12,13] and it was decided to use the same difference equation, where applicable, to Eq. (9). The following notation is used:

$$f^n_{i,j}=f(n\Delta t,i\Delta\eta,j\Delta\xi), \tag{23.1}$$

$$f^+_t=\frac{1}{\Delta t}(f^{n+1}_{i,j}-f^n_{i,j}), \tag{23.2}$$

$$f^0_\eta=\frac{1}{2\Delta\eta}(f^n_{i+1,j}-f^n_{i-1,j}), \tag{23.3}$$

$$f_{\eta\eta}=\frac{1}{\Delta\eta^2}(f^n_{i+1,j}-2f^n_{i,j}+f^n_{i-1,j}), \tag{23.4}$$

$$f^0_\xi=\frac{1}{2\Delta\xi}(f^n_{i,j+1}-f^n_{i,j-1}) \tag{23.5}$$

$$f_{\xi\xi}=\frac{1}{\Delta\xi^2}(f^n_{i,j+1}-2f^n_{i,j}+f^n_{i,j-1}). \tag{23.6}$$

For the two predictive equations for Θ and $\tilde{u}$, Eq. (9.1) and (9.2), the difference equations are,

$$\Theta^+_\tau=-r\Theta^0_\xi-\frac{u}{L_1}\Theta^0_\eta+\frac{\Delta\xi}{2}|r|\Theta_{\xi\xi}+\frac{\Delta\eta}{2}\left|\frac{u}{L_1}\right|\Theta_{\eta\eta}, \tag{24.1}$$

$$\begin{aligned}\tilde{u}^+_\tau=&-r\tilde{u}^0_\xi-\frac{u}{L_1}\tilde{u}^0_\eta-\frac{c_s^2}{(\gamma-1)\pi L_1}\pi^0_\eta\\&+\frac{\Delta\xi}{2}|r|\tilde{u}_{\xi\xi}-h_1\\&+\frac{\Delta\eta}{2}\left[\left(\frac{c_s+|u|a^2}{L_1(1+a^2)}\right)\tilde{u}_{\eta\eta}+\left(\frac{uc_s}{(\gamma-1)\pi L_1}\right)\pi_{\eta\eta}\right.\\&\left.+\left(\frac{c_s-|u|}{L_1}\right)\left(\frac{a}{1+a^2}\right)\tilde{w}_{\eta\eta}\right],\end{aligned} \tag{24.2}$$

where all coefficients are evaluated at $\binom{n}{i,j}$. [These equations are obtained from Eqs. (8.2) by locally transforming to diagonal form, differencing according to the sign of the characteristics, and then transforming back.[12]]

The two diagnostic equations, (9.3) and (9.4), were considered as an "independent" system of two equations. Both were differenced with a standard trapezoidal method. Essentially, $\tilde{w}$ is obtained from Eq. (9.3) with Eq. (22.3), and π is obtained from Eq. (9.4) with the aid of an upper boundary condition. When solving Eqs. (9.3) and (9.4) for $\tilde{w}$ and π, it is assumed that $\tilde{u}$ and Θ are known everywhere. These two equations are still coupled and need to be solved iteratively.

Summarizing, the numerical computations proceed as follows. Given an entire distribution of all variables at time n, Eqs. (24.1) and (24.2) are used, in conjucntion with previously given boundary conditions, to obtain Θ and $\tilde{u}$ at time $n+1$. [Note that Eq. (24.1) is used at all boundaries and that Eq. (24.2) is used at $\xi=0$ and $\xi=1$.] Then, Eqs. (9.3) and (9.4) are solved to obtain π and $\tilde{w}$ at time $n+1$.

VIII. CONVERGENCE AND STABILITY OF THE NUMERICAL CALCULATION

No procedures are presently available for analytically establishing convergence and stability of the numerical techniques. Nevertheless, one should at least be able to demonstrate that the solutions behave as though they are a part of a convergent and stable sequence of calculation. A criterion, to be used as a necessary condition for convergence of the difference equations, was obtained by the following argument. Suppose $\Delta t=hc_1$, $\Delta\eta=hc_2$, $\Delta\xi=hc_3$. Let $f(t,\eta,\xi)$ be any of the dependent variables, and let $f^*(t,\eta,\xi,h)$ be its numerical approximation. Assume an expression of the form $f(t,\eta,\xi)=f^*(t,\eta,\xi,h)+A_1h+O(h^2)$, where $A_1=A_1(t,\eta,\xi)$. Ignoring the higher order term,

$$f(t,\eta,\xi)\simeq f^*(t,\eta,\xi,h)+A_1h. \tag{25.1}$$

Using Eq. (25.1) with $h/2$ and $h/4$, one obtains,

$$f^*(t,\eta,\xi,h)-f^*(t,\eta,\xi,h/2)\simeq 2[f^*(t,\eta,\xi,h/2)-f^*(t,\eta,\xi,h/4)]. \tag{25.2}$$

Equation (25.2) is similar to that used in Romberg integration.[14] The validity of Eq. (25.2) is adopted as a reasonable indication that the numerical solutions are converging to $f(t,\eta,\xi)$, and is at the same time a valid test for both stability of the numerical scheme and stability of solutions of the partial differential equations. At any stage of such a sequence, any given quantity is obtained with eight times as many calculations as in the previous stage. Any numerical instabilities will certainly be seen. Also, convergence, in the sense of Eq. (25.2), would hardly be possible if the solution to the differential equation were unstable with respect to perturbations of initial data.

We first check Eqs. (6). In terms of Eq. (7), this requies that

$$f_1(t,x)=0, \tag{26.1}$$

$$f_2(t,x)=\text{const}=L_2. \tag{26.2}$$

The additional boundary conditions chosen for the test case (see Sec. V) were as follows:

$$\frac{\partial u}{\partial z}(t,x,L_2)=\begin{cases}0 & :w<0,\\ \left(\dfrac{\partial u}{\partial z}\right)^- & :w\geqslant 0,\end{cases} \tag{26.3}$$

$$\frac{\partial\Theta}{\partial z}(t,x,L_2)=\begin{cases}\dfrac{(\gamma-1)g}{c_s^2\pi} & :w<0,\\ \left(\dfrac{\partial\Theta}{\partial z}\right)^- & :w\geqslant 0,\end{cases} \tag{26.4}$$

where $(\partial/\partial z)^-$ indicates the backward difference. Note that, for $w\geqslant 0$, the implication is that Eqs. (6.1) and (6.2) are used to calculate u and Θ, the idea being to use characteristic equations as required by the mathematics. (The precise characteristic equations were not employed in the present study, but will be a part of our later study in regard to lateral boundary conditions.) The difficulty actually occurs for $w < 0$, for in this case there seems to be neither mathematical nor physical insight to dictate appropriate boundary conditions. Those given above are clearly artifical.

For the first calculation we also chose to keep pressure constant at the top. That is,

$$\pi(t,x,L_1) \equiv \text{const.} \tag{26.5}$$

Summarizing, a solution (using the numerical technique described in Sec. VII) is to be obtained to Eqs. (6), with initial conditions given by Eqs. (22.4)–(22.8), and boundary conditions given by Eqs. (22.1), (22.2), (22.3), and (26).

To test convergence, a set of four calculations was made. In the first run, the values of the mesh were $\Delta x = 40$ km, $\Delta z = 2.5$ km, and $\Delta t = 80$ seconds. (The stability criterion is essentially $\Delta t < \Delta x/c_s$, which produces $\Delta t \sim 133$.) These values give a 13×5 spatial mesh. In the second run all three increments were halved, producing a 25×9 spatial mesh. In the third run all three increments were halved again, producing a 49×17 spatial mesh, and in the fourth run the spatial mesh was 97×33. This produces an initial field as shown in Figs. 1 and 2.

Some of the results of this calculation are shown in Figs. 3–5. Figures 3 and 4 give the entire field for P, T, u, and w at a later time. Figure 5 displays $P(t,x,0)$. (These results are for the 97×33 mesh.)

One notes that from a physical standpoint the behavior of the solution is generally as expected. The initial Θ distribution produces a wave which then moves to the right and proceeds to reflect from the lateral boundaries. Vertical velocity is monotonic, for fixed x, and achieves maximum absolute values at the top. The period of the oscillation is about 4000 seconds. One such period is shown in Fig. 5. At time zero the maximum pressure at $z=0$ is at $x=0$. In 4000 seconds, the pressure distribution has essentially reversed.

Figure 5 shows that the magnitude of the pressure gradient at time $t=0$ is much larger than that existing at time $t=4000$. This does not indicate any damping of the flow field. In fact, when continued to time $t=8000$, the pressure gradient was much like that existing at $t=0$. (These extended calculations were not thought to be very accurate and so were not shown.)

Further comments regarding this calculation will be made later when comparing results of different upper boundary conditions.

The test for convergence in terms of Eq. (25.2) is shown in Fig. 6. Calculated values of several values of the flow field are shown for all of the mesh sizes used (runs 1 through 4). $\Delta 1 = f^*(t,x,z,h) - f^*(t,x,z,h/2)$ is obtained by simply subtracting the value given by run 2 from that given by run 1. Similarly, one obtains $\Delta 2$ and $\Delta 3$. The criterion is satisfied if $\Delta 3/\Delta 2 = 0.5$. In practice, we look to see if the quantities $\{\Delta 2/\Delta 1, \Delta 3/\Delta 2\}$ appear to be part of a sequence whose limit is less than or equal to $\frac{1}{2}$.

We first note that, as should be expected, better convergence is obtained at $t=2080$ than at $t=4000$. Convergence for all quantities is quite well indicated, with the exception possibly of $T(t,0,L_2)$. This seems to indicate a difficulty with the boundary conditions given by Eqs. (26.3) and (26.4) for $w<0$. As noted earlier, these conditions are artificial ones, and the lack of convergence of $T(t,0,L_2)$ might indicate that the mathematical problem is not well posed. Note that at $x=0$, $w<0$ for most of the range $0 \leqslant t \leqslant 4000$, whereas at $x=L_1$, $w>0$ for most of the calculation (see Fig. 4). Thus, the difficulty with the upper boundary should occur primarily near $x=0$. This is consistent with the results of Fig. 6: $T(t,L_1,L_2)$ converges much better than $T(t,0,L_2)$.

IX. COMPARISON OF THE FIXED TOP QUASISTEADY MODEL WITH A HYDROSTATIC MODEL

There are two goals, somewhat contradictory, to the comparison to be made in this section. The first goal is to show that, at least in some circumstances results obtained with the quasisteady equations are similar to those obtained with standard atmospheric models. It is known that for a large class of problems these standard models give results that are physically realistic. For such problems the quasisteady model should give similar results. The second goal is to ascertain what differences can be attributed to the fact that the different models employ different sets of differential equations. As discussed in Sec. II, we feel that the quasisteady model differs from the standard models in two important ways: (1) the differential equations are different, and (2) with the quasisteady assumption it should be possible to derive physically realistic and mathematically consistent boundary conditions. The second of these two differences will be discussed in Sec. XI. In this section we attempt to examine differences that can be attributed to the differential equations.

The test problem of Sec. VIII was solved with the hydrostatic model developed by Kreitzberg,[15] which is based on the model developed at NCAR.[16] In his model, the primitive equations consist of predictive equations for T and horizontal velocity, and diagnostic equations for P (hydrostatic) and w. For the case of our problem and boundary conditions the equation for w has the form

$$\frac{\partial w}{\partial z} = -\frac{\partial u}{\partial x} - \frac{u}{\gamma P}\frac{\partial P}{\partial x} + \frac{g}{\gamma P} \times \int_z^{L_2} \frac{\partial}{\partial x}(\rho u)d\tilde{z} + \frac{g}{\gamma P}(\rho w)_{z=L_2}, \tag{27.1}$$

where $w_{z=L_2}$ is first obtained by integrating Eq. (27.1) from 0 to L_2 and then solving. All conditions were kept the same for this run except at the upper boundary where temperature is predicted using backward differences in the vertical terms [this replaces Eqs. (26.3)–(26.4) for $w<0$].

Because the convective terms in Eq. (6.4) are negligible for this problem, the two models differ essentially only to the

FIG. 3. Flow field at time 2710 sec.

$\left(\frac{z}{0\text{ km.}}\right)$													
							$P(2710,x,z)$						
1.	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039	269.039
0.9375	293.691	293.705	293.748	293.822	293.925	294.056	294.209	294.376	294.540	294.685	294.799	294.871	294.895
0.8750	320.623	320.653	320.743	320.896	321.108	321.373	321.679	322.005	322.324	322.605	322.823	322.961	323.009
0.8125	350.011	350.056	350.193	350.423	350.740	351.133	351.580	352.054	352.512	352.914	353.226	353.423	353.490
0.75	382.011	382.071	382.252	382.552	382.964	383.470	384.043	384.644	385.223	385.728	386.118	386.364	386.448
0.6875	416.778	416.850	417.068	417.428	417.919	418.519	419.194	419.899	420.573	421.157	421.607	421.891	421.988
0.6250	454.462	454.544	454.790	455.194	455.743	456.412	457.160	457.937	458.676	459.314	459.804	460.111	460.216
0.5625	495.210	495.298	495.560	495.989	496.570	497.276	498.063	498.878	499.649	500.311	500.816	501.133	501.241
0.50	539.164	539.253	539.517	539.948	540.531	541.238	542.028	542.841	543.608	544.263	544.759	545.069	545.175
0.4375	586.458	586.543	586.792	587.200	587.753	588.425	589.176	589.950	590.675	591.291	591.754	592.040	592.138
0.3750	637.216	637.291	637.510	637.869	638.359	638.959	639.634	640.329	640.978	641.524	641.930	642.178	642.263
0.3125	691.553	691.611	691.784	692.071	692.469	692.965	693.528	694.111	694.653	695.103	695.431	695.629	695.695
0.25	749.568	749.605	749.719	749.915	750.199	750.565	750.991	751.437	751.849	752.184	752.420	752.558	752.602
0.1875	811.348	811.363	811.411	811.506	811.663	811.887	812.163	812.459	812.732	812.943	813.080	813.153	813.175
0.1250	876.968	876.959	876.942	876.940	876.974	877.060	877.193	877.347	877.486	877.579	877.623	877.636	877.636
0.0625	946.487	946.459	946.390	946.309	946.247	946.225	946.245	946.289	946.325	946.328	946.299	946.264	946.248
0.	1019.954	1019.917	1019.824	1019.707	1019.602	1019.533	1019.507	1019.505	1019.498	1019.463	1019.403	1019.347	1019.324
							$T(2710,x,z)$						
1.	243.495	243.359	242.945	242.242	241.255	240.005	238.536	236.899	235.276	233.847	232.739	232.039	231.801
0.9375	243.279	243.148	242.750	242.082	241.157	240.014	238.720	237.369	236.073	234.945	234.076	233.530	233.344
0.8750	243.181	243.067	242.726	242.162	241.399	240.478	239.461	238.419	237.434	236.587	235.940	235.535	235.396
0.8125	243.523	243.434	243.168	242.737	242.165	241.486	240.749	240.006	239.318	238.735	238.294	238.019	237.925
0.75	244.356	244.293	244.110	243.821	243.444	243.008	242.544	242.089	241.681	241.344	241.095	240.942	240.889
0.6875	245.664	245.629	245.532	245.387	245.207	245.007	244.807	244.625	244.477	244.368	244.297	244.256	244.242
0.6250	247.429	247.423	247.414	247.412	247.422	247.450	247.497	247.568	247.659	247.758	247.848	247.909	247.930
0.5625	249.637	249.660	249.737	249.872	250.062	250.299	250.572	250.872	251.176	251.459	251.692	251.844	251.897
0.50	252.279	252.330	252.487	252.747	253.098	253.521	253.992	254.489	254.976	255.416	255.769	255.998	256.078
0.4375	255.348	255.424	255.650	256.017	256.504	257.080	257.714	258.370	259.004	259.567	260.016	260.307	260.408
0.3750	258.840	258.936	259.217	259.666	260.254	260.944	261.696	262.466	263.202	263.852	264.366	264.699	264.815
0.3125	262.753	262.862	263.178	263.676	264.323	265.077	265.894	266.724	267.513	268.205	268.751	269.102	269.225
0.25	267.083	267.197	267.522	268.030	268.684	269.444	270.263	271.092	271.874	272.558	273.095	273.440	273.561
0.1875	271.827	271.934	272.237	272.707	273.310	274.007	274.757	275.512	276.221	276.839	277.321	277.630	277.739
0.1250	276.976	277.064	277.308	277.685	278.167	278.724	279.321	279.920	280.480	280.966	281.344	281.585	281.669
0.0625	282.518	282.570	282.713	282.934	283.216	283.542	283.893	284.244	284.571	284.852	285.069	285.206	285.254
0.	288.431	288.428	288.421	288.411	288.403	288.397	288.395	288.395	288.394	288.391	288.387	288.382	288.380
	0.	0.08333	0.16667	0.25	0.33333	0.41667	0.50	0.58333	0.66667	0.75	0.83333	0.91667	1 $\left(\frac{x}{480\text{ km.}}\right)$

FIG. 4. Flow field at time = 2710 sec.

$\left(\frac{z}{10\text{ km.}}\right)$ \ $\left(\frac{x}{480\text{ km.}}\right)$	0.	0.08333	0.16667	0.25	0.33333	0.41667	0.50	0.58333	0.66667	0.75	0.8333	.91667	1.
							$u(2710,x,z)$ (m/sec)						
1.	0.0000	−.0000	−.0000	−.0000	−.0002	−.0014	−.0092	−.0441	−.1195	−.1732	−.1714	−.1069	0.0000
0.9375	0.0000	−.1761	−.3489	−.5134	−.6665	−.8048	−.9185	−.9832	−.9695	−.8675	−.6648	−.3641	0.0000
0.8750	0.0000	−.3919	−.7622	−1.0886	−1.3563	−1.5556	−1.6733	−1.6885	−1.5810	−1.3543	−1.0047	−.5395	0.0000
0.8125	0.0000	−.5601	−1.0774	−1.5133	−1.8455	−2.0645	−2.1617	−2.1232	−1.9396	−1.6269	−1.1884	−.6325	0.0000
0.75	0.0000	−.6595	−1.2579	−1.7449	−2.0952	−2.3024	−2.3643	−2.2757	−2.0397	−1.6835	−1.2159	−.6435	0.0000
0.6875	0.0000	−.6824	−1.2903	−1.7680	−2.0913	−2.2587	−2.2753	−2.1446	−1.8833	−1.5281	−1.0912	−.5745	0.0000
0.6250	0.0000	−.6257	−1.1702	−1.5788	−1.8323	−1.9353	−1.8999	−1.7387	−1.4809	−1.1706	−.8218	−.4297	0.0000
0.5625	0.0000	−.4903	−.9008	−1.1840	−1.3287	−1.3462	−1.2554	−1.0769	−.8512	−.6275	−.4200	−.2154	0.0000
0.50	0.0000	−.2817	−.4936	−.6011	−.6032	−.5182	−.3707	−.1886	−.0219	.0778	.0973	.0594	0.0000
0.4375	0.0000	−.0096	.0319	.1417	.3092	.5095	.7128	.8853	.9700	.9145	.7077	.3830	0.0000
0.3750	0.0000	.3116	.6479	1.0053	1.3612	1.6847	1.9416	2.0930	2.0776	1.8435	1.3831	.7406	0.0000
0.3125	0.0000	0.6633	1.3186	1.9401	2.4936	2.9423	3.2489	3.3704	3.2431	2.8175	2.0893	1.1139	0.0000
0.25	0.0000	1.0223	2.0005	2.8863	3.6348	4.2044	4.5551	4.6408	4.3980	3.7795	2.7854	1.4815	0.0000
0.1875	0.0000	1.3615	2.6423	3.7738	4.7013	5.3798	5.7669	5.8149	5.4618	4.6632	3.4234	1.8179	0.0000
0.1250	0.0000	1.6495	3.1857	4.5224	5.5979	6.3640	6.7774	6.7898	6.3418	5.3919	3.9481	2.0942	0.0000
0.0625	0.0000	1.8510	3.5643	5.0418	6.2164	7.0387	7.4652	7.4486	6.9327	5.8786	4.2970	2.2773	0.0000
0.	0.0000	1.9260	3.7039	5.2301	6.4355	7.2708	7.6939	7.6602	7.1166	6.0257	4.4000	2.3305	0.0000
							$w(2710,x,z)$ (m/sec)						
1.	−.1980	−.1955	−.1874	−.1715	−.1455	−.1089	−.0614	−.0012	.0797	.1696	.2510	.3067	.3263
0.9375	−.1872	−.1849	−.1772	−.1622	−.1378	−.1034	−.0587	−.0019	.0746	.1599	.2374	.2906	.3094
0.8750	−.1802	−.1779	−.1704	−.1558	−.1323	−.0993	−.0564	−.0017	.0717	.1533	.2276	.2789	.2971
0.8125	−.1766	−.1743	−.1665	−.1516	−.1282	−.0956	−.0537	−.0003	.0704	.1488	.2204	.2700	.2877
0.75	−.1754	−.1727	−.1643	−.1486	−.1246	−.0921	−.0506	.0018	.0703	.1457	.2146	.2627	.2799
0.6875	−.1752	−.1722	−.1627	−.1459	−.1211	−.0883	−.0471	.0045	.0706	.1430	.2091	.2556	.2723
0.6250	−.1749	−.1715	−.1609	−.1428	−.1171	−.0839	−.0431	.0074	.0709	.1399	.2030	.2476	.2638
0.5625	−.1732	−.1693	−.1577	−.1385	−.1121	−.0789	−.0387	.0103	.0707	.1357	.1954	.2377	.2532
0.50	−.1691	−.1648	−.1523	−.1323	−.1058	−.0731	−.0341	.0127	.0693	.1299	.1854	.2251	.2397
0.4375	−.1617	−.1571	−.1441	−.1240	−.0979	−.0664	−.0294	.0145	.0666	.1219	.1726	.2090	.2225
0.3750	−.1502	−.1455	−.1326	−.1130	−.0882	−.0587	−.0246	.0154	.0621	.1113	.1565	.1891	.2012
0.3125	−.1343	−.1298	−.1175	−.0994	−.0767	−.0502	−.0199	.0152	.0556	.0981	.1369	.1651	.1757
0.25	−.1139	−.1098	−.0989	−.0830	−.0635	−.0409	−.0154	.0139	.0473	.0821	.1140	.1371	.1459
0.1875	−.0893	−.0859	−.0770	−.0643	−.0488	−.0310	−.0111	.0116	.0371	.0637	.0880	.1056	.1123
0.1250	−.0612	−.0588	−.0525	−.0436	−.0329	−.0207	−.0070	.0083	.0255	.0433	.0596	.0714	.0759
0.0625	−.0309	−.0296	−.0264	−.0218	−.0164	−.0102	−.0033	.0043	.0129	.0217	.0298	.0356	.0379
0.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

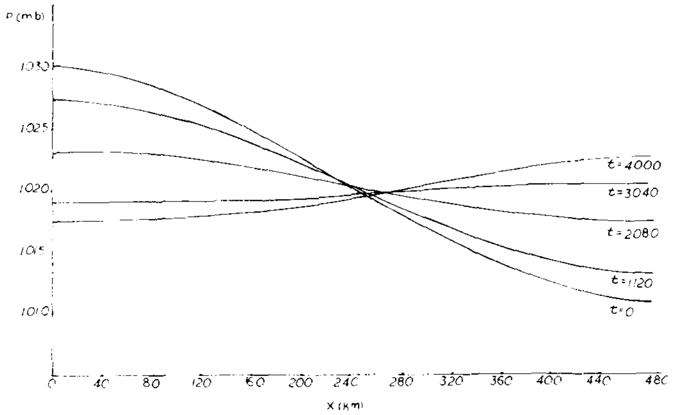


FIG. 5. Pressure distribution at $z = 0$ at selected times.

extent that Eq. (6.3) differs from Eq. (27.1). Ignoring the convective terms in Eq. (6.4), these diagnostic equations for w can be written as,

$$\frac{\partial w}{\partial z}+\frac{\partial u}{\partial x}+\frac{u}{\gamma P}\frac{\partial P}{\partial x}=f_1\frac{g}{c_s^2}, \tag{27.2}$$

$$f_1=\begin{cases}\dfrac{1}{\rho}\displaystyle\int_z^{L_2}\frac{\partial(\rho u)}{\partial x}\,d\tilde{z}+\frac{(\rho w)_{z=L_2}}{\rho} & \text{:from Eq. (27.1),}\\[2ex] w & \text{:from Eq. (6.3).}\end{cases}$$

In Eq. (27.2), the equations agree at $z=L_2$, while they appear to have the greatest discrepancy at $z=0$. It is also interesting to note that if,

$$(\rho u)_x+(\rho w)_z\equiv 0, \tag{27.3}$$

then the two expression are identical. For the boundary conditions to be discussed in Sec. X, Eq. (27.1) would have an additional term involving $\partial P/\partial t$, while Eq. (6.3) remains unchanged. For these cases the discrepancies in the two diagnostic equations for w become more pronounced.

A table of comparisons for the two runs at time 4160 is given below:

	Results using Eq. (6)	Results using Kreitzberg Program
Period of Oscillation	4000 sec.	4160 sec.
$P(4160,0,0)$	1017.013	1016.953
$P(4160,L_1,0)$	1022.578	1023.331
$T(4160,0,L_2)$	244.265	244.636
$T(4160,L_1,L_2)$	230.468	230.283
$u(4160,L_1/2,0)$	6.0837	7.327
$w(4160,0,L_2)$	.1165	.0507

As expected, the basic features of the flow field agree reasonably well. For example, the period of oscillation differs by 4%. The increase in pressure at $z=0$, $z=L_1$ [that is, $P(4160,L_1,0)-P(0,L_1,0)$] differs by approximately 6%. The percentage variation in the velocity field appears to be significantly larger. However, this may be due to a relatively small phase shift (namely, the 4% variation in the period). This can be seen from Fig. 7, where w is displayed at the two top corner points. Figure 7 also shows that differences became more pronounced as the run progressed. This is consistent with our previous discussion: Equation (27.3) is valid at $t=0$, but becomes increasingly less valid as the run continues.

Summarizing, for this test case (in which the effect of boundary conditions was minimized as much as seemed possible) it was not expected that large differences would occur. The results verified this expectation. However, it was also

FIG. 6. Convergence check.

	$t = 2080$ sec. $P(t,0,0)$	$P(t,L_1,0)$	$T(t,0,L_2)$	$T(t,L_1,L_2)$	$w(t,0,L_2)$	$u(t,L_1/2,0)$
run 1	1020.820	1014.139	241.438	234.913	−.2241	5.3806
run 2	1022.287	1015.665	241.755	234.164	−.2470	6.5485
run 3	1022.686	1016.328	241.923	233.760	−.2602	7.1234
run 4	1022.798	1016.640	242.012	233.556	−.2673	7.4005
$\Delta 1$	−1.467	−1.526	−.317	.749	.0229	−1.1679
$\Delta 2$	−.399	−.663	−.168	.404	.0132	−.5749
$\Delta 3$	−.112	−.312	−.089	.204	.0071	−.2771
$\Delta 2/\Delta 1$	.272	.434	.530	.539	.576	.492
$\Delta 3/\Delta 2$	.281	.471	.529	.505	.538	.482

	$t = 4000$ sec. $P(t,0,0)$	$P(t,L_1,0)$	$T(t,0,L_2)$	$T(t,L_1,L_2)$	$w(t,0,L_2)$	$u(t,L_1/2,L_2)$
run 1	1015.777	1018.246	243.478	232.519	.0429	3.2671
run 2	1016.824	1020.690	243.943	231.290	.0624	4.8880
run 3	1017.011	1021.911	244.241	230.641	.0716	5.8344
run 4	1916.009	1022.543	244.417	230.312	.0761	6.3252
$\Delta 1$	−1.047	−2.444	−.465	1.229	−.195	−1.6209
$\Delta 2$	−.187	−1.221	−.298	.649	−.092	−.9464
$\Delta 3$	~0.	−.632	−.176	.329	−.045	−.4908
$\Delta 2/\Delta 1$	.179	.500	.641	.528	.472	.586
$\Delta 3/\Delta 2$	~0.	.518	.591	.507	.489	.518

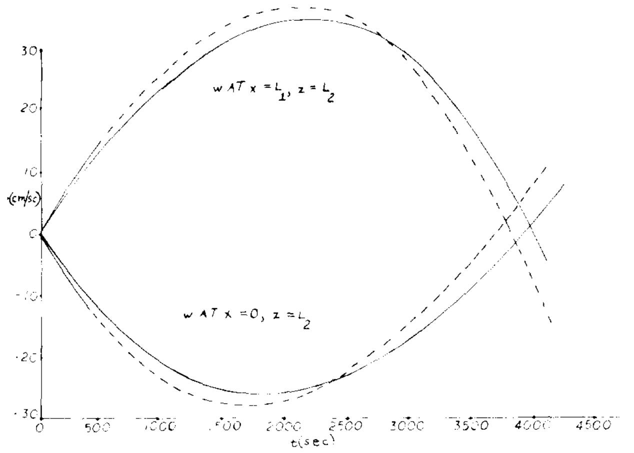


FIG. 7. Comparison for ω. [——: Kreitzberg results, - - - - : results from Eq. (6)]

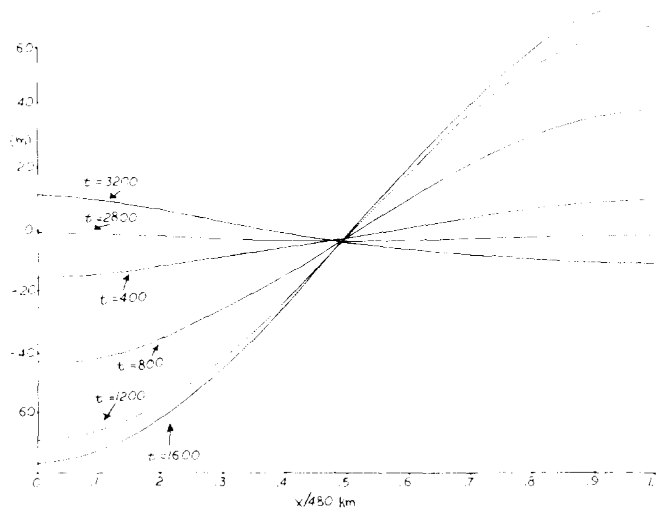


FIG. 8. Position of floating top.

seen that the quantitative differences, although small, were significant.

X. COMPARISON OF THE FIXED TOP QUASISTEADY MODEL WITH THE FLOATING TOP QUASISTEADY MODEL

It was shown in Sec. V that the floating top, in conjunction with Eq. (10), has a significant mathematical effect: We now ask whether there is any significant effect on the solution itself.

All of the upper boundary conditions that will be discussed in this section and in Sec. XI are written in the from

$$\pi_\tau = c_1 \tilde{w}_\tau - c_2 \tilde{w}. \tag{28}$$

For the fixed top, the following three cases will be considered:

$$(1)\ c_1 = 0.0, \quad c_2 = 0, \tag{28.1}$$

$$(2)\ c_1 = 0.0, \quad c_2 = -\frac{g(\gamma-1)\pi}{c_s^2}(0.15), \tag{28.2}$$

$$(3)\ c_1 = 0.0, \quad c_2 = -\frac{g(\gamma-1)\pi}{c_s^2}. \tag{28.3}$$

Cases (1), (2), and (3) are derived from $\partial P/\partial t = \rho g w \alpha$, where $\alpha = 0, 0.15, 1$. $\alpha = 0$ gives the test case of Sec. VIII. The case $\alpha = 1$, which arises from the boundary condition $P_t + wP_z = 0$ and the hydrostatic equation $P_z + \rho g = 0$, has been shown to be unstable; Kreitzberg has found that the damping factor, $\alpha = 0.15$, gives reasonable results.

For the floating top, the following four cases are considered:

$$(4)\ c_1 = \frac{(\gamma-1)\pi}{c_s(1+a^2)^{1/2}}, \quad c_2 = 0, \tag{28.4}$$

$$(5)\ c_1 = \frac{(\gamma-1)\pi}{c_s(1+a^2)^{1/2}}, \quad c_2 = 0 : \tilde{w} \leqslant 0,$$

$$c_1 = \frac{(\gamma-1)\pi}{c_s}, \quad c_2 = \frac{g(\gamma-1)\pi}{c_s^2} : \tilde{w} > 0, \tag{28.5}$$

$$(6)\ c_1 = 0.0, \quad c_2 = \frac{g(\gamma-1)\pi}{c_s^2}, \tag{28.6}$$

$$(7)\ c_1 = 0, \quad c_2 = 0. \tag{28.7}$$

Cases (4), (5), and (6) correspond respectively to Eqs. (15), (21), and (20), while case (7) maintains constant pressure at $\zeta = 1$.

We first discuss the movement of the top surface. Figure 8 shows the configuration of the top for the case of the continuous upper boundary condition, Eq. (28.4). Plots for the range $1600 < t < 2800$ are not shown; the curves essentially retrace the earlier ones. The movement of this surface is actually quite small. For Eq. (28.4), the maximum is about 70 m (to be compared with the initial height of 10,000 m). For Eq. (28.5) the behavior is qualitatively as in Fig. 8, but the extreme values are approximately -102 m and 270 m.

Figure 9 displays the actual change in pressure that occurs at the top surface for cases (2), (3), and (4). For each case pressure is shown at times of maximum deviation from the initial pressure P_0. The pressure increment ΔP is particularly small for the continuous upper boundary condition. Results for Eq. (28.5), which are not given in Fig. 9, show that the maximum $\Delta P = 9.4$.

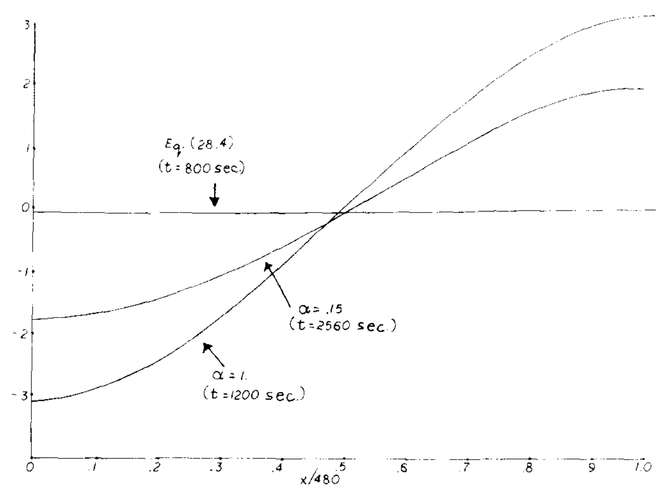


FIG. 9. $P(t,x,z_{top}) - P_0$.

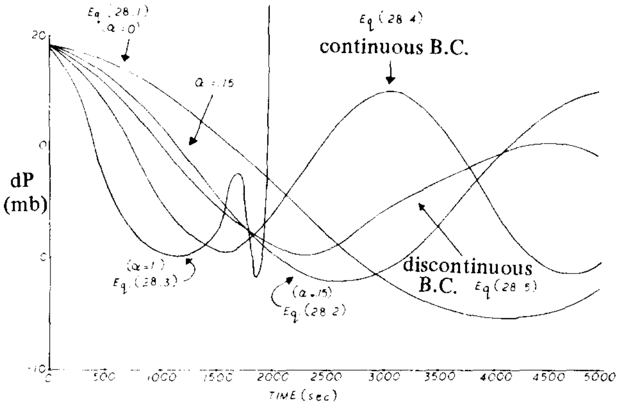


FIG. 10. $dP = P(t,0,0) - P(t,L_1,0)$.

Equation (28.4) can be put in the form of Eq. (28.3) by the following argument. Let "1" denote quantities at $\zeta = 1$ and "0" denote quantities at $z = 10$ km. From the hydrostatic approximation, $P_0 - P_1 \approx -\rho g \Delta z$, where $\Delta z = 10 - z_1$. Differentiating with respect to τ,

$$\frac{\partial P_0}{\partial \tau} - \frac{\partial P_1}{\partial \tau} \approx -\rho g\left(\frac{-z_1}{\partial \tau}\right).$$

Since pressure changes so little at $\zeta = 1$, let $\partial P_1/\partial \tau = 0$; from Eq. (10), $\partial z_1/\partial \tau = \tilde{w}_1$, and since the movement of the top is so small, let $\tilde{w} \approx \tilde{w}_0$. Then $\partial P/\partial \tau = \rho g \tilde{w}$ at $z = 10$ km. Summarizing, with the hydrostatic approximation, Eq. (28.3) arises from $P_t + wP_z = 0$, while Eq. (28.4) is essentially obtained from $P_\tau + \tilde{w}P_\zeta = 0$.

Figure 10 illustrates a remarkable difference in behavior between Eq. (28.3) and Eq. (28.4), namely that the solution with Eq. (28.3) becomes unstable. The equations were examined further in order to attempt to isolate the significant differences. In regard to the predictive equations, two differences are seen. First, the fixed top model, Eqs. (6), requires the additional arbitrary upper boundary conditions given by Eq. (26.3) and (26.4). Secondly, the predictive variables for Eqs. (6) are Θ and u, while for Eqs. (9) they are Θ and $\tilde{u}$. These differences in the predictive equations do have an effect on the solution, but not enough to account for the fact that the fixed top model was divergent while the floating top model was well behaved.

More profound effects can be attributed to the diagnostic equations. Note first that, for this test case at least, both Eqs. (6.4) and (9.4) are essentially the usual hydrostatic equation; in both equations all other terms are negligible. The other diagnostic equation differs, in that Eq. (6.4) is a diagnostic equation for w, while in Eq. (9.4) the diagnostic variable is $\tilde{w}$. As will be seen in Sec. XI, the vertical velocity field $\tilde{w}$ differs greatly, both quantitatively and qualitatively, from the vertical velocity field w produced by the fixed top model. With the boundary conditions being used, Eqs. (28.4) and (28.3) respectively, these changes in the vertical velocity field are directly fed into the pressure calculation.

Summarizing, we have investigated a realistic physical upper boundary condition, Eq. (28.3), with both the fixed top and floating top models. Other investigators have encountered highly unstable results with this boundary condition. Similarly, our fixed top model was unstable. However, the floating top model was well behaved. For this test case, we attributed this primarily to the fact that the floating top requires $\tilde{w}$ in its equations, while for the fixed top w is the required quantity.

XI. COMPARISON OF RESULTS OBTAINED WITH DIFFERENT UPPER BOUNDARY CONDITIONS

In this paper two upper boundary conditions were derived. The goal of this section is to study results obtained with these boundary conditions and to compare with other commonly used upper boundary conditions. The specific questions to be considered are the following:

(1) Which boundary conditions are consistent with the requirement of existence of a quasi-steady-state limit?

(2) Do results with these boundary conditions indicate convergence and stability?

(3) What is the effect of the boundary conditions on the period of the wave, and do all solutions represent gravity waves?

(4) Do the boundary conditions succeed in producing any damping of the wave motion?

(5) Large differences can be seen between the results obtained with the boundary conditions derived in the paper. What accounts for these differences?

(6) What is the effect of doubling the height of the region of computation?

All calculations display the same qualitative features as discussed in Sec. VIII; a wave moves to the right and reflects to the left. These features will not be shown. Rather we will attempt to display those features that explain the effect of the various boundary conditions.

Question (1) will be studied in detail in Sec. XII. There it will be shown that, of the boundary conditions considered, only Eqs. (28.4) and (28.5) satisfied the requirement for existence of a quasi-steady-state solution. The results for Eqs. (28.6) and (28.7) are inconclusive, but all cases for the fixed top, Eqs. (28.1)–(28.3), will be seen to diverge.

We next consider Question (2). The criterion for convergence and stability will be the same as that given in Sec. VIII. In all cases convergence, in the sense of Eq. (25.2), with the floating top was better than with the fixed top. The convergence check for one case with the floating top is shown in Fig. 11. Here the boundary condition is Eq. (28.4), the continuous upper boundary condition. Figure 11 should be compared with the fixed top result given in Fig. 6.

For the continuous boundary condition, temperature at the top changes very little; thus, the convergence test is not meaningful. For the other cases of the floating top temperature does change significantly at the top; in these cases convergence for the temperature field was good (comparable to the velocity field). This is shown in Fig. 11 at $t = 4000$, where the data for $T(t,0,L_2)$ is taken from the case of the discontinu-

FIG. 11. Convergence check for top boundary condition given by Eq. (28.4).

			$t = 2000$ sec.				
$P(t,0,0)$	$P(t,L,0)$	$T(t,0,L_2)$	$\Delta F2(0)$	$\tilde{w}(t,0,L_2)$	$\tilde{u}(t,L_1/2,0)$		
run 1	1020.201	1016.853	237.847	-55.78	0.0535	2.6080	
run 2	1021.596	1017.873	237.849	-56.48	0.0606	3.2979	
run 3	1021.943	1018.198	237.850	-56.67	0.0634	3.6239	
run 4	1022.026	1018.318	237.849	-56.72	0.0647	3.7743	
$\Delta 1$	-0.396	1.020	$\sim 0.$	0.70	-0.0071	-0.6899	
$\Delta 2$	-0.347	-0.325	$\sim 0.$	0.19	-0.0028	-0.3260	
$\Delta 3$	-0.083	-0.120	$\sim 0.$	0.05	-0.0013	-0.1504	
$\Delta 2/\Delta 1$	-0.249	0.319		0.27	0.39	0.4725	
$\Delta 3/\Delta 2$	-0.239	0.369		0.26	0.46	0.4613	
			$t = 400$ sec.				
$P(t,0,0)$	$P(t,L_1,0)$	$T(t,0,L_2)$	$\Delta F2(0)$	$\tilde{w}(t,0,L_1)$	$\tilde{u}(t,L_1/2,0)$		
run 1	1020.990	1015.312	235.717	-21.73	-0.0417	5.4742	
run 2	1022.175	1016.836	235.418	-19.19	-0.0460	7.4129	
run 3	1022.508	1017.447	235.254	-17.43	-0.0480	8.5542	
run 4	1022.609	1017.721	235.167	-16.41	-0.0489	9.1605	
$\Delta 1$	-1.185	-1.524	0.299	-2.54	0.0043	1.9387	
$\Delta 2$	-0.333	-0.611	0.164	-1.76	0.0020	-1.1413	
$\Delta 3$	-0.101	-0.274	0.87	-1.02	0.0009	-0.6063	
$\Delta 2/\Delta 1$	0.281	0.401	0.548	0.693	0.465	0.5887	
$\Delta 3/\Delta 2$	0.303	0.448	0.530	0.570	0.45	0.5312	
			Eq. (28.5)				

ous boundary condition, Eq. (28.5). Also, in Fig. 11 we have added the quantity $\Delta F_2(x)$, which represents the movement of the top surface, in place of $T(t,L_1,L_2)$ of Fig. 6.

Question (3) is concerned with a description of the wave motion itself. The period of oscillation was taken as the time at which a minimum "over-pressure" is achieved, as in Fig. 5. Figure 12 tabulates the period of oscillation for the three fixed top cases and two floating top cases. We were particularly interested in comparing the two boundary conditions derived in this paper, Eqs. (28.4) and (28.5), with the fixed top cases. As perhaps should be expected, the period depends strongly on the boundary condition. In particular, the continuous and discontinuous boundary conditions show large differences. (Further results regarding physical applicability of these boundary conditions will be discussed in a later paper.)

The atmospheric scientist expects a test case such as this to generate a gravity wave, rather than a sound wave. It is of interest to note that, based on the calculations themselves, it is difficult to make such a judgment. To determine whether or not the wave is in fact a gravity wave, the calculations were repeated with the value of g arbitrarily increased by a factor of 4: If the wave were a sound wave, the period should be relatively unaffected, while for a gravity wave the period should be increased by a factor of 2 (Ref. 2, p. 25). The resulting periods were all increased by a factor of 2, indicating that all solutions represent gravity waves.

Question (4) is concerned with whether or not the mathematical model used here (namely), the inviscid equations with an assumed upper boundary condition), contains any such mechanism by which it might be possible to dampen the flow to a steady-state. A primary parameter here is "over pressure", which represents for this problem the driving mechanism for the flow. Figure 10 displays over pressure dP, defined as the difference in pressure at the lower corner points, as a function of time. The case with $\alpha = 0$ has such a

FIG. 12. Comparison of upper boundary conditions.

Boundary Condition	Period of oscillation (sec)	maximum $\tilde{w}$ at $x = 0$ (m/sec)	time max $\tilde{w}$ achieved (sec)	maximum $\tilde{w}$ at $x = L_1$ (m/sec)	time max $\tilde{w}$ achieved (sec)
Eq. (28.1): $\alpha = 0$	4000.	-0.28	1600.	0.37	1920
Eq. (28.2): $\alpha = .15$	2640.	-0.20	1120.	0.22	1280
Eq. (28.3): $\alpha = 1.$	1200.	-0.10	560.	0.10	640.
Eq. (28.4): continuous upper B.C.	1550.	-0.089	1200.	0.092	1200.
Eq. (28.5): discontinuous upper B.C.	2360.	-0.11	200.	0.190	1600.
Eq. (28.6): hydrostatic at top (height doubled)	2280.	-0.14	1000.	0.148	1000.

long period that the calculation really does not reveal the oscillatory behavior of pressure. The case with $\alpha = 1$ becomes unstable at $t = 2000$. Thus, the three cases of interest are the continuous and discontinuous boundary conditions for the floating top and $\alpha = 0.15$ for the fixed top.

Define a complete wave cycle time for this flow as the time required for the wave to travel to the right, reflect from the boundary $x = L_1$, travel to the left, and begin to reflect from the boundary $x = 0$. One measure of the damping for this problem is the extent to which over pressure at the end of the cycle is decreased. These over pressures are tabulated below from Fig. 10.

	dP time = 0	dP cycle time
$\alpha = 0.15$, Eq. (28.2)	19.	15.5
Continuous boundary condition, Eq. (28.4)	19.	15.2
Discontinuous boundary condition, Eq. (28.5)	19.	10.3

It is clear that all produce some damping. The effect produced by Eqs. (28.2) and (28.4) are much alike (although the cycle times are far different). However, Eq. (28.5) produces more than twice as much damping in over pressure as the other cases.

We now consider Question (5). We attempt to pick out those features which are particularly related to the differences in results. First, the qualitative and quantitative differences between w and $\tilde{w}$ are shown in Fig. 13. Curve I shows a typical w curve at $x = 0$ for a fixed top case. Curves II and III show typical $\tilde{w}$ curves for the discontinuous and continuous boundary conditions. For the fixed top cases, the maximum value of w occurs at the top, while for the floating top cases the maximum value of $\tilde{w}$ occurs near the middle. This is important since for the floating top cases $\tilde{w}$ is the pertinent quantity with respect to the boundary conditions: See Eq. (28) and (10). Also, the period of oscillation, as discussed earlier, varies greatly. Figure 12 tabulates the maxima for $\tilde{w}$ and the time at which the maxima are achieved. One sees that for the floating top cases, Eqs. (28.4) and (28.5), there seemed to be some correlation between the time of the period and the time of the maximum w; this was not so for the fixed top cases. One concludes from this that the floating top upper boundary conditions produce a much stronger coupling, between vertical velocity and the total flow, than the fixed top upper boundary conditions.

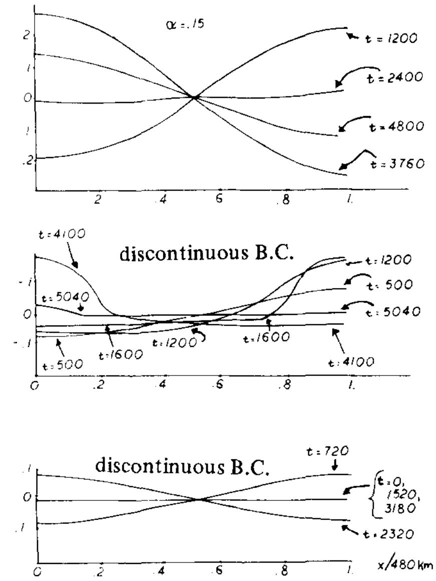


FIG. 14. $\tilde{\omega}$ (m/sec) at $z \sim 10$ km.

Profiles of vertical velocity are shown in Fig. 14. This figure attempts to give an indication of how $\tilde{w}$ changes with time. The figure shows $\tilde{w}$ at $\zeta = 1$ for three of the runs. In the previous discussion we defined the "period" of the oscillation and the "cycle time. Figure 14 shows that Eqs. (28.2) and (28.4) produce $\tilde{w} \approx 0$ at both the period and cycle time, while Eq. (28.5) produces $\tilde{w} \approx 0$ only at the cycle time.

Next, Fig. 15 displays $P(t,0,0)$. From Fig. 10 and 15 one sees that dP and $P(t,0,0)$ are closely in phase for all cases, except for the discontinuous boundary conditions.

For the continuous case Eq. (28.4), the effect of the term $c_1 \tilde{w}_T$ was very small for this test case. [The case of Eq. (28.7)

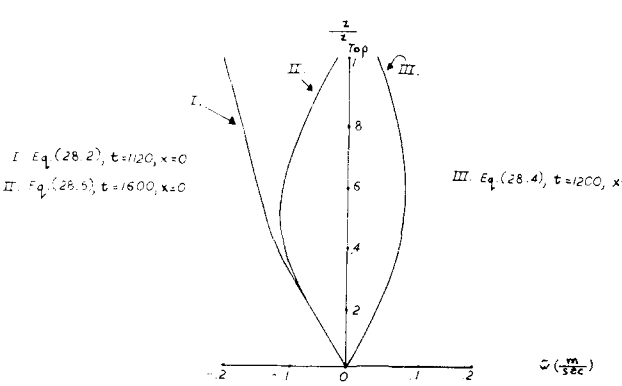


FIG. 13. $\tilde{\omega}$ distribution for fixed x.

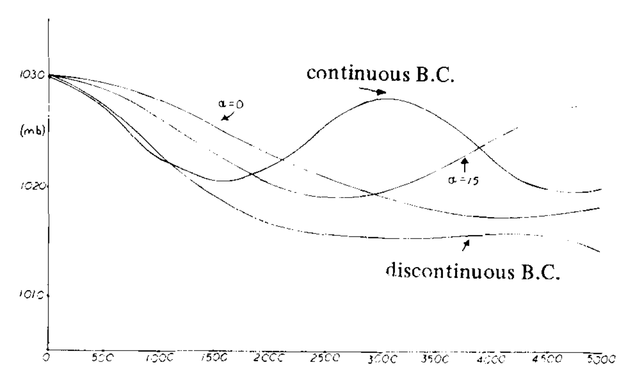


FIG. 15. $P(t,0,0)$.

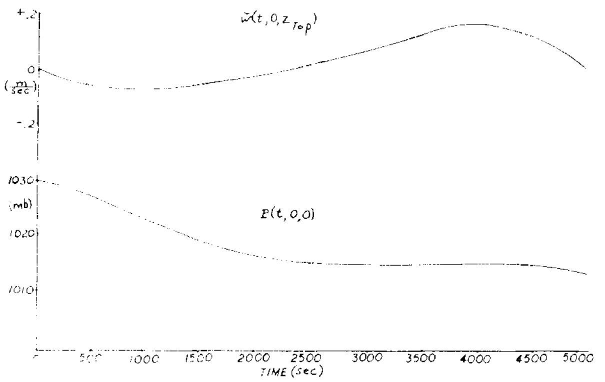


FIG. 16. Discontinuous upper boundary condition [Eq. (28.5)].

which sets $c_1 = 0$ for the moving boundary, produced results that differed from Eq. (28.4) by only a few percent.] Thus, the pressure change at the bottom occurs not so much because of the pressure change at the top, but rather because the top pressure is imposed at the moving surface. In the case of $\tilde{w} \leqslant 0$, Eq. (28.5) is identical to Eq. (28.4). [Note in Fig. 15 that the results for Eq. (28.4) and (28.5) are in good agreement for the earlier part of the run.] But for $\tilde{w} > 0$, which occurs at $x = L_1$ during the early part of the run, Eq. (28.5) adjusts the top pressure in much the same way as Eq. (28.6); that is, essentially pressure at the top is set at the pressure existing at that height at the initial time. Figure 16 further illustrates this point. Note that $P(t,0,0)$ changes significantly when $\tilde{w} < 0$, but much less so when $\tilde{w} > 0$.

Summarizing, the floating top boundary condition behaves roughly in the following manner:

(1) Equation (28.4) tends to maintain constant pressure at the moving surface. Thus, if $\tilde{w} > 0$ at the top, in which case the top moves upward, then through an essentially hydrostatic integration pressure will be increased at the lower levels. Similarly, if $\tilde{w} < 0$, pressure will decrease at lower levels.

(2) Equation (28.6) tends to adjust pressure at the top so that pressures at lower levels are unchanged. Thus, if $\tilde{w} > 0$, pressure decreases at the top so that the following hydrostatic integration downward can approximately leave the pressure unchanged.

(3) Equation (28.5) behaves as Eq. (28.4) for $\tilde{w} < 0$ at the top and as Eq. (28.6) for $\tilde{w} > 0$. Thus, as shown in Fig. 16, for $\tilde{w} < 0$ pressure tends to decrease at lower levels, while for $\tilde{w} > 0$ pressure tends to remain constant at lower levels.

Also note that the results for the continuous and discontinuous boundary conditions differed greatly in all aspects of the flow field. This is disappointing in the following sense. Equations (28.4) and (28.5) were derived on the basis of two "extreme" physical assumptions, one assumption being continuity of the flow and the other being discontinuity of the flow. One would then like to find features of the flow which are similar for the two cases; one could then conclude that these features are relatively independent of the upper boundary condition.

Finally, we consider Question (6), the effect of doubling the height of the region of computation. This test case represents another approach to the problem of the upper boundary condition, namely to place the upper boundary so high that the primary flow occurs far below. For this test case the initial height is doubled from 10 km to 20 km. The input for $0 < z < 10$ km is as in the other cases. In the region 10 km $\leqslant z \leqslant 20$ km, Θ varied linearly with slope $d\Theta/dz = 0.5197 \times 10^{-4}$ m^{-1} and was constant with x; π was integrated upward, hydrostatically, from 10 km, while the velocity was zero. The boundary condition used at the top was Eq. (28.6). This boundary condition allows a moving top, but adjusts pressure so as to remain at the hydrostatic pressure corresponding to its instantaneous position.

Initially there was no pressure gradient above $z = 10$ km. Although this changed during the course of the calculation, relatively small horizontal velocities and gradients were generated in the vicinity of $z = 20$ km. However, because of the diagnostic equation for vertical velocity, large values for this variable will extend through the entire height of the region. The upper boundary condition, Eq. (28.6), is such that the effect on the pressure field, of this presumably unrealistic value of w, is minimized. The question is whether such a procedure can negate the necessity of specifying a boundary condition at the top that is physically correlated with the actual flow field. (This is discussed further in Sec. XIII.) It was thought that the doubled height calculation would agree reasonably well with one of the other cases with an initial height of 10 km. In fact this did not occur. For example, Fig. 17 compares horizontal velocity for the doubled height case. There seems to be good correlation between the doubled height case and two of the other cases ($\alpha = 0.15$ and the discontinuous boundary condition). On the other hand, over pressure at the cycle time 4450 [see question (4)] was $dP = 13.7$; this aggees with none of the other cases.

Also, it did not appear that the test case was successful in achieving a flow in which the solution in the lower region remained independent of the flow generated in the upper region. There seem to be several reasons for this:

(a) For the case of the doubled height, one must still specify an upper boundary condition. Equation (28.6) is intended to minimize the effect of "extraneous" calculations near $z = 20$ km. However, as will be discussed in Sec. XIII, effects occurring at the top still propagate down, and significantly affect the solution in the lower 10 km.

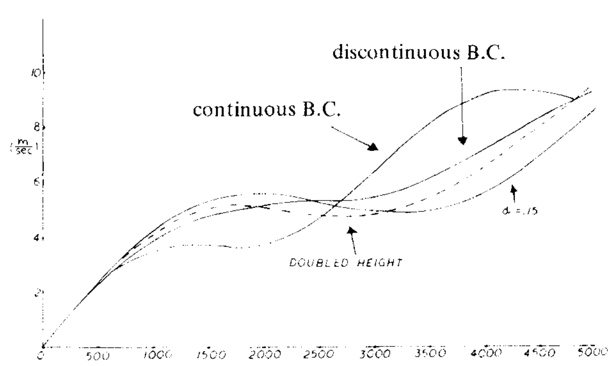


FIG. 17. $\tilde{u}(t, L_1/2, 0)$.

(b) The test cases indicate that it is apparently impossible to specify an internally derived upper boundary condition in a region of relatively undisturbed flow. (In this regard the discussion of Sec. XIII is very relevant.) That is, the boundary condition would need to be independent of the local calculation of the flow variables (say, for example, specified independently as a function of time). More work needs to be done on this aspect of the problem, but at the moment it does not appear to be promising.

XII. THE EXISTENCE OF A STEADY STATE SOLUTION TO THE QUASISTEADY EQUATIONS

In Sec. II we discussed the thought that, for mathematical consistency, the solution to the quasisteady equations should be the steady-state limit of the time-dependent equations. We consider this to be a very important requirement. Thus, the following criterion is adopted as a necessary condition for mathematical justification of the quasisteady assumption:

> The solution of the quasi-steady equations, in conjunction with an imposed upper boundary condition, is the steady-state limit of the time-dependent equations. (29)

The test on criterion (29) was made as follows. For a fixed x and t and for fixed $\tilde{u}$ and Θ fields, the time-dependent π and $\tilde{w}$ equations, given by Eqs. (8.2), were integrated to see if a steady-state solution would be achieved. Since $|\tilde{w}| < c_s$, one boundary condition is needed at each boundary. At $\zeta = 0$ this was $\tilde{w} = 0$ and at $\zeta = 1$ each of the seven cases given by Eq. (28) was tested. The equations were differenced by the same technique as discussed in Sec. VII. One characteristic equation is used, in conjunction with the above boundary conditions, at $\zeta = 0$ and at $\zeta = 1$. (Details of this calculation will not be given. However, it is a somewhat difficult one, particularly because of the occurrence of various arithmetic problems.)

The problem discussed in Sec. VII was run until time $t = 800$ sec., at which point the time-dependent equations for π and $\tilde{w}$ were solved. Figure 18 shows results for the fixed top. Behavior, in terms of convergence to a steady-state, was similar for cases (28.1), (28.2), and (28.3), and only case (28.1) is shown. Also, to isolate the effect of the moving boundary from the boundary condition itself, we ran another test case in which the boundary condition Eq. (28.4) was used with a fixed top (that is, $f_{2_T} \equiv 0$). This case is labeled as Eq. (28.4)*. It seems clear that case (28.1) is not converging to a steady-state. In fact, it appears to be slowly diverging. On the other hand, Eq. (28.4)* does converge. Also, the final steady-state value achieved is the correct value, in that it agrees with the value obtained by the use of the diagnostic equations themselves.

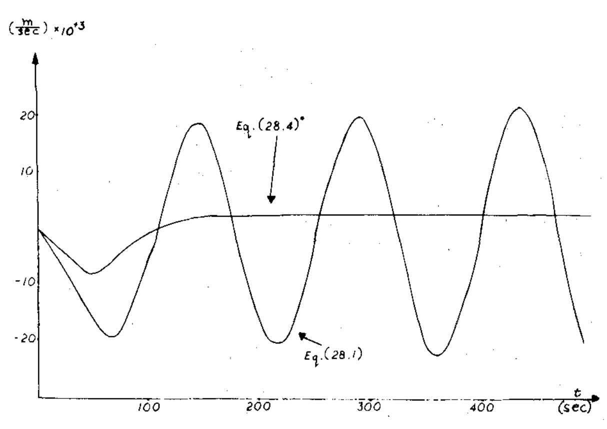


FIG. 18. Quasi-steady-state for fixed top: $\tilde{w}_\epsilon$ at $\zeta = 1$.

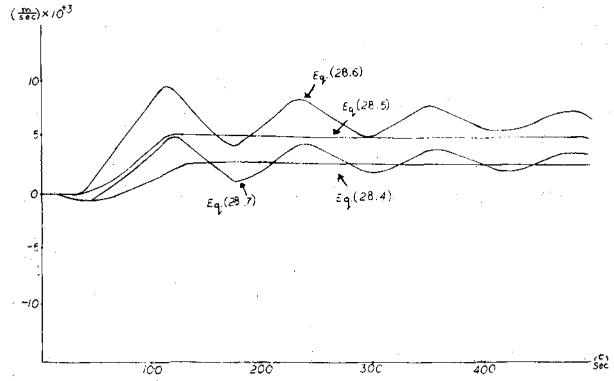


FIG. 19. Quasi-steady-state for the floating top: $\tilde{w}_\epsilon$ at $\zeta = 1$.

Another aspect of this problem, not yet investigated, is the time required to achieve steady state. The time required for Eq. (28.4)* to converge is about 150 sec., while the actual time step used in the numerical calculation was 80 sec. For the moment this "discrepancy" is assumed not to be significant. The equations are in fact perturbation equations, and the calculation clearly begins with incorrect data. That is, data is given only at time t and time $t + \Delta t$, whereas the correct problem would have all intermediate data. If the steady-state solution is "stable," this type of error should not destroy the convergence but obviously would distort the approach to steady state. Note that the solution shown in Fig. 18 for Eq. (28.4)* begins by moving in the wrong direction.

Figure 19 shows results for the four cases with the floating top. The cases corresponding to the continuous and discontinuous boundary conditions, Eq. (28.4) and (28.5), converge smoothly. Conclusions for Eqs. (28.6) and (28.7) are not so clear. At best one might argue that the oscillations appear to be decreasing in magnitude, as opposed to the oscillation shown in Fig. 18, and that the results are oscillating about the correct steady-state value.

It is also interesting to note that the solutions generated by Eqs. (28.4) and (28.7) differed by only a few percent (see Sec. XI). Yet in terms of criterion (29) there is a clear difference in behavior, in favor of Eq. (28.4). As noted at the end of Sec. III, this might well be related to the problem of reflective behavior at the upper boundary.

Summarizing, Eqs. (28.4)*, (28.4), and (28.5) satisfy criterion (29). Equations (28.1), (28.2), and (28.3) do not satisfy criterion (29). Results of Eqs. (28.6) and (28.7) are as yet inconclusive. In particular, the two boundary conditions derived in this paper, Eqs. (28.4) and (28.5), satisfy the criterion.

XIII. THE SCALE ASSUMPTION AND THE TIME DEPENDENCE EXPRESSED BY UPPER BOUNDARY CONDITIONS

Equations (1) are well understood, both physically and mathematically. For example, problems at outflow points are recognized to be physical in nature (one simply does not know the information required by the equations). However, with the introduction of quasisteady equations, or with any of the existing models discussed in Sec. II, this lack of ambiguity is no longer present. In particular, one is no longer certain that problems occurring at the upper boundary are not due to the additional mathematical assumptions inherent in the derivation. In this section, two simple problems will be solved with the quasisteady equations and a given upper boundary condition. It will be clear that the solutions obtained are clearly invalid from a physical point of view. This then provides an opportunity to assess the validity of the assumptions involved.

For the first example take the case of a one-dimensional horizontal flow in which the gravitational force is zero. (Thus, the analytic solution contains no vertical gradients.) Suppose, however, that one chooses to solve the problem in a region $0 \leqslant x \leqslant L$, $0 \leqslant z \leqslant h$. Suppose further that assumption (2) is valid, so that the quasisteady analysis also holds. The quasisteady version of the hydrostatic equation would be $\partial\pi/\partial z = 0$, which is valid equation for this case. The quasisteady equation for w would presumably yield nonzero values at the top. Since the actual solution has $\tilde{w} \equiv 0$, it is clear that the calculated value has no physical significance. Consequently, the value of π at $\zeta = 1$, calculated by any of the cases in Eq. (28), can have no physical significance.

It is the author's opinion that in this problem the difficulty lies with the upper boundary condition, rather than with the quasisteady equations. As already noted, quasisteady, or diagnostic equations, contain no direct time-dependence, the implication being that if correct time-dependence is given at any level, then it is known to desired accuracy at all levels. In the above example, correct time-dependence for π has not been specified. Note that for the case of the continuous upper boundary condition, given by Eq. (15), assumption (14.3) is not valid for this problem. Also, in regard to Eq. (21), any discontinuity produced in a real fluid by the calculated $\tilde{w}$ would form essentially at $z=0$, and thus Eq. (21) cannot be applied at an arbitrary height.

To do this problem, one could choose boundary conditions as follows: At some level, say $z=0$, assume one-dimensional horizontal flow (that is, assume symmetry about $z=0$), and obtain π from the corresponding time-dependent one-dimensional equations. Then let $\tilde{w}=0$ at $z=h$. The diagnostic equations will propagate $\tilde{w}$ and π vertically. A nonzero value of $\tilde{w}$ may be calculated below $z=h$, but a valid approximation to the solution will still be obtained.

In the second example, the lower boundary moves through a stationary fluid, in which $\tilde{w}=0$ is maintained at $\zeta=0$. $f_1(t,\eta)$ varied slowly from a value of zero at time zero to the final distribution, at time 1000, shown in Fig. 20. Boundary and initial conditions were as given by Eqs. (22), except

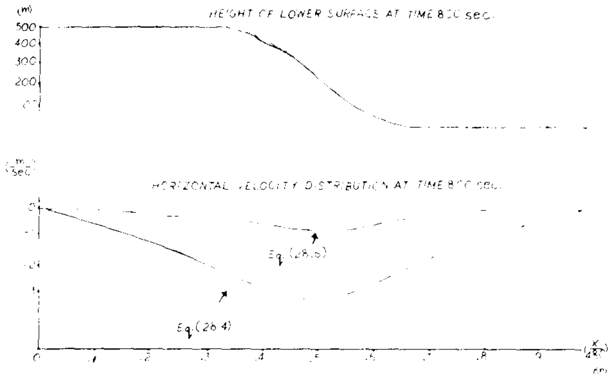


FIG. 20. Moving lower boundary.

that $f(x,z) \equiv 0$. This is simply a movement of the coordinate system and stationary flow should be maintained. For $0 < t < 1000$, the nonzero values of $\partial f_1/\partial t$ produce nonzero values of $\tilde{w}$. This coupled with the continuous boundary condition, Eq. (28.4), produces a distribution for $\tilde{u}$ as shown in Fig. 20. Clearly, the solution has no physical relevance.

One explanation, for the poor results obtained in the calculations, might be that again, as in the first example, correct time-dependence has not been specified for π at any point. This was shown to be false by repeating the calculation with Eq. (28.6); this assumes that π remains at its hydrostatic value at $\zeta = 1$. The results shown in Fig. 20 are much better for Eq. (28.6), but overall the calculation is still not satisfactory.

It is the author's opinion that the difficulty in this problem lies with assumption (2) and not with the upper boundary conditions. More precisely, the assumption that $h \ll L$ is actually not true. For, referring to Fig. 20, the drop in the curve for the lower surface occurs essentially over a distance of approximately 50 km. Thus, in terms of assumption (2) one should consider $h = 10$ km, and $L = 50$ km.

Summarizing, we have considered two examples in this section. The first showed that none of the boundary conditions considered in this paper could provide the correct time-dependence required by the diagnostic equations. The second example showed that violation of assumption (2), as would occur if one used the quasisteady equations over a steep mountain range, can lead to erroneous results.

XIV. SUMMARY

Atmospheric scientists have found that the inviscid, time-dependent equations are extremely difficult to use in regard to large-scale atmospheric flow. Various sets of "primitive" equations have been proposed for the purpose of alleviating the problems. These models have been reasonably successful, but there remain various unresolved inconsistencies, particularly in regard to boundary conditions.

The purpose of this paper was to obtain a model which, (1) would be suitable for large-scale atmospheric flow, and (2) would be consistent with well-defined realistic physical and mathematical assumptions. The basic features of the analysis were, (1) quasisteady state phenomena, and (2) a slowly floating upper boundary. The floating upper boundary reduces the situation at the upper boundary to a more tractible mathematical form. A model, which we have called the quasisteady primitive equations, was derived on the basis of a specific scale assumption. In addition, two upper boundary conditions were derived. The first is based on the assumption of continuity at the upper boundary between the solution obtained using the quasisteady equations and the solution given by the full equations above. The second upper boundary condition assumes a discontinuity, caused by a positive vertical velocity at the upper boundary.

The above derivations thus provide two specific criterion for assessing the validity of the quasisteady approximations. The first is the validity of assumption (2), namely the physical assumption regarding the time and length scales of the problem. The second, criterion (29), is the mathematical restriction that the solution of the quasisteady equations be the steady-state limit of the time-dependent equations.

A major philosophical difference between the quasisteady model and other existing models is the following: With the quasisteady model it is required that direct contact be maintained with the original time-dependent problem upon which the analysis was based; this contact is not required of the other models. Several examples, involving simplified mathematical models, were discussed in Sec. III for the purpose of (1) clarifying the use of quasisteady analysis, and (2) emphasizing the importance of this contact with the original time-dependent problem.

A specific physical problem was defined. This problem is not directly of interest to atmospheric scientists, namely because it used a zero horizontal velocity at the lateral boundaries. However, the problem does involve scales and magnitudes which are comparable to those encountered in large-scale atmospheric flow. Numerical solutions were obtained through use of a first order numerical method. Solutions to this test problem were obtained with several mathematical models and with a variety of upper boundary conditions.

Specific conclusions to our study are as follows:

(1) By comparing the results of a sequence of numerical computations, it was shown that the calculations seemed to be convergent and stable. However, the fixed top case displayed some difficulty at the upper boundary. Overall the floating top cases showed faster convergence than the fixed top cases.

(2) For the case of fixed pressure at the fixed top, the quasisteady model was compared with Kreitzberg's hydrostatic model. As was expected, the results were reasonably close, with differences in the diagnostic equations becoming more significant as the calculations proceeded in time.

(3) The differences between the fixed and floating top models were analyzed. A particularly significant difference is the fact that w is required in the fixed top model, while $\tilde{w}$ is the required variable in the floating top model. It was found that a much stronger coupling, between pressure and vertical velocity, existed for the floating top model than for the fixed top. This may explain, at least in part, the dramatic advantage of the floating top that was demonstrated by the following calculation:

Eq. (28.3) is a physically realistic upper boundary condition, but has been found by other investigators to lead to unstable solutions. When used with the fixed top, our calculations displayed the same phenomenon. However, the floating top analogue of Eq. (28.3), namely Eq. (28.4), produced stable and physically reasonable results.

(4) Over pressure was taken as a reasonable criterion by which one could judge the damping effect produced by the model. This parameter showed that the discontinuous upper boundary condition produced much more damping than any other upper boundary condition that was investigated.

(5) One possible procedure for alleviating the problem of the upper boundary is to place the upper boundary far above the region of disturbed flow. This procedure was investigated. The results of the calculation indicated that this approach is probably not a fruitful one, in that it does not appear possible to remove the necessity of specifying physically meaningful upper boundary conditions.

(6) Criterion (29) was tested for all upper boundary conditions considered. The solution of the quasisteady equations, in conjunction with either of the two upper boundary conditions derived in the paper, was shown to be the steady-state limit of the time-dependent equations. Other commonly used upper boundary conditions did not satisfy the criterion.

(7) The scale assumption, assumption (2) was shown to be of critical importance.

(8) Several simple problems were presented, for which the quasisteady equations did not give adequate solutions. These examples demonstrated the importance of the scale assumption and of the philosophy of maintaining contact with the original time-dependent problem.

ACKNOWLEDGMENTS

This work has been supported by NSF Grant GA-35093. Computational assistance has been provided by the National Center for Atmospheric Research (NCAR), an organization sponsored by the National Science Foundation. Some CDC 7600 computer time was obtained under the joint NCAR/Drexel University Project with the Small Scale Analysis and Prediction Project at NCAR. This work has been supported by the United States Energy Research and Development Adminstration under Contract No. E(11-1)-2360 and by the Atmospheric Sciences Division of Brookhaven National Laboratories.

The author would also like to point out that this investigation was aided greatly by Professor Carl W. Kreitzberg. In addition to actively supporting this type of research, Professor Kreitzberg contributed many hours of discussion and criticism, as well as many useful suggestions.

[1]Yoshimitsu Ogura and N.A. Phillips, "Scale Analysis of Deep and Shallow Convection in the Atmosphere," J. Atmos. Sci. **19**, 173–9 (1962).
[2]G.J. Haltiner, *Numerical Weather Prediction* (Wiley, New York, 1971).
[3]J.A. Dutton and G.H. Fichtl, "Approximate Equations of Motion for Gases and Liquids," J. Atmos. Sci. **26**, 241–54 (1969).
[4]J. Oliger and A. Sundström, "Theoretical and Practical Aspects of Some Initial-Value Problems in Fluid Dynamics," Computer Science Dept., School of Humanities and Sciences, Stanford University, STAN-CS-76-578, November, 1976.
[5]L.D. Landau and E.M. Lifschitz, *Fluid Mechanics* (Pergamon, New York, 1966).
[6]R.P. Fedorenko, "Derivation and Justification of Equations in Slow Time (The Stroboscopic Method)," USSR Comput. Math.&Math. Phys. **14**, 81–118 (1966).
[7]P. Gordon, "Scaling of a System of Differential Equations," SIAM J. Appl. Math. **30**, 391–401 (1976).
[8]Roger A. Pielke and Ytzhaq Mahrer, "Representation of the Heated Planetary Boundary Layer in Mesoscale Models with Course Vertical Resolution," J. Atmos. Sci. **32**, xxx (1975).
[9]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Vol. 2, Partial Differential Equations) (Interscience, New York, 1962).
[10]R. Courant and K.O. Friedrichs, *Supersonic Flow and Shock Waves* (Interscience, New York, 1948).
[11]P. Gordon and S.M. Scala, "Numerical Investigation of the Formation of Discontinuities in Nonisentropic Inviscid Flows," Proceedings of the International Symposium on High Speed Computing in Fluid Dynamics, Monterey, Calif., Aug. 19–24, 1968, Phys. Fluids Suppl. II **12**, Part II, II-41 (1969).
[12]P. Gordon, "The Diagonal Form of Quasi-Linear Hyperbolic Systems as a Basis for Differential Equations," Final Report–Theoretical Investigation of Nuclear Burst Phenomena, Naval Ordinance Laboratory, Contract No. N60921-7164 (General Electric Company Document No. 69SD7056).
[13]P. Gordon and S.M. Scala, "Nonlinear Theory of Pulsatile Blood Flow Through Viscoelastic Blood Vessels," *in Proceedings of the AGARD Specialists' Meeting on Fluid Dynamics of Blood Circulation and Respiratory Flow*, May 4–6 1970, Naples, Italy (Technical Editing and Reproduction Ltd., London, 1970).
[14]F.L. Bauer, H. Rutishauser, and E. Stiefel, "New Aspects in Numerical Quadrature," Proc. Symp. Appl. Math. **15**, 199 (1963).
[15]Carl W. Kreitzberg and Donald J. Perkey, "Release of Potential Instability; Part I. A Sequential Plume Model within a Hydrostatic Primitive Equation Model," J. Atmos. Sci. **33**, 456–75 (1976).
[16]J.E. Oliger, R.E. Wellck, A. Kasahara, and N.M. Washington, "Description of the NCAR Global Circulation Model," Publication of LAS, NCAR, Boulder, Colorado, 1970.

Clebsch–Gordan coefficients: General theory

R. Dirl

Institut für Theoretische Physik, TU Wien, A-1040 Wien, Karlsplatz 13, Austria
(Received 28 December 1977)

A general method is given for obtaining Clebsch–Gordan coefficients for finite groups, by considering the columns of the Clebsch–Gordan matrices as G-adapted vectors and by identifying the multiplicity index as special column indices of the Kronecker product. The matrix representations are assumed to be projective ones, however not necessarily belonging to equivalent factor systems.

INTRODUCTION

An important application of group theory to physics is the problem of decomposing a Kronecker product of irreducible representations into the direct sum of its irreducible constituents. In order to be as general as possible, we shall consider different sets of projective unitary irreducible matrix representations (unirreps) of a given finite group G. These representations, however, do not belong necessarily to equivalent standard factor systems.[1-4] The reason for investigating this rather complicated situation is, that just for space group representations, which we want to consider in the following papers, such cases may occur.

Instead of considering in more detail several methods[5-14] which have been used until now to actually calculate Clebsch–Gordan coefficients CG coefficients) for a given group, we reexamine the defining equation for CC coefficients. In doing so we rewrite this equation in such a way that the columns of the CG matrices can be considered as G-adapted complex vectors, i.e., vectors transforming according to the corresponding projective unirrep. Therefore, these G-adapted vectors can be determined by the usual projection formalism.[15-17] This implies that difficulties which may arise in this approach can only concern the multiplicity problem. This problem appears immediately, if a unirrep is contained more than once into the Kronecker product. In order to resolve this problem, we investigate conditions which allow one to identify the multiplicity index with special column indices of the Kronecker product. For this purpose we consider a set of special vectors. Finally we mention that the present method can be easily transferred to compact continuous groups.

The material is organized as follows: In Sec. I we summarize the properties of projective unirreps and state the problem of determining CG coefficients for a finite group G. In Sec. II we reformulate the theory in such a way that the columns of the CG matrices can be considered as G-adapted vectors. Their explicit construction is carried out by the usual projection formalism. Concerning the multiplicity index, we derive conditions which guarantee whether special column indices of the Kronecker product can be chosen as a multiplicity index. In so doing we obtain by means of simple formulas the CG matrices quite generally.

I. CLEBSCH–GORDAN COEFFICIENTS: STATEMENT OF THE GENERAL PROBLEM

In order to be able to state the problem of how CG matrices can be determined systematically, we have to make several assumptions. First of all we assume for the sake of simplicity that G is a finite group of order $|G|$, since the generalization of the described method to compact continuous groups is obvious. Concerning the representations we want to be as general as possible and consider projective representations. In this connection we assume that for G two different complete sets of projective unirreps are known, which belong to not necessarily equivalent standard factor systems R and S. (A situation which may occur just for nonsymmorphic space groups.)

$$\mathbb{D}^{\alpha}:=\{\mathbb{D}^{\alpha}(x):x\in G\},\quad \alpha\in A_{G(R)},\tag{I.1}$$

$$\mathbb{D}^{\beta}:=\{\mathbb{D}^{\beta}(x):x\in G\},\quad \beta\in A_{G(S)}.\tag{I.2}$$

Thereby $\mathbb{D}^{\alpha}$ ($\mathbb{D}^{\beta}$) denotes n_α (n_β)-dimensional projective matrix unirreps of G belonging to the standard factor system R (S); $A_{G(R)}$ ($A_{G(S)}$) denotes the set of all equivalence classes of G with respect to R (S). The matrix elements of the corresponding projective unirreps must satisfy the following equations:

$$\sum_{r=1}^{n_\alpha}\mathbb{D}^{\alpha}_{pr}(s)\mathbb{D}^{\alpha}_{rq}(y)=R(x,y)\mathbb{D}^{\alpha}_{pq}(xy)\quad\text{for all } x,y\in G,\tag{I.3}$$

$$\frac{1}{|G|}\sum_{x\in G}\mathbb{D}^{\alpha}_{pq}{}^{*}(x)\,\mathbb{D}^{\alpha'}_{p'q'}(x)=n_\alpha^{-1}\,\delta_{\alpha\alpha'}\,\delta_{pp'}\,\delta_{qq'},\tag{I.4}$$

$$\sum_{\alpha pq}n_\alpha\,\mathbb{D}^{\alpha}_{pq}{}^{*}(x)\,\mathbb{D}^{\alpha}_{pq}(y)=|G|\,\delta_{x,y},\tag{I.5}$$

respectively,

$$\sum_{t=1}^{n_\beta}\mathbb{D}^{\beta}_{rt}(x)\,\mathbb{D}^{\beta}_{ts}(y)=S(x,y)\,\mathbb{D}^{\beta}_{rs}(xy),\tag{I.6}$$

$$\frac{1}{|G|}\sum_{x\in G}\mathbb{D}^{\beta}_{rs}{}^{*}(x)\,\mathbb{D}^{\beta'}_{r's'}(x)=n_\beta^{-1}\,\delta_{\beta\beta'}\,\delta_{rr'}\,\delta_{ss'},\tag{I.7}$$

$$\sum_{\beta rs}n_\beta\,\mathbb{D}^{\beta}_{rs}{}^{*}(x)\,\mathbb{D}^{\beta}_{rs}(y)=|G|\,\delta_{x,y}.\tag{I.8}$$

It is well known that CG matrices are special subduction matrices, where the supergroup is the direct product group $G\times G$ and the subgroup the Kronecker product $G[x]G$, which is isomorphic to G. Provided that the complete sets (I.1) and (I.2) of projective matrix unirreps are known, those for $G\times G$ follow immediately by

$$\mathbb{D}^{\alpha\beta}:=\{\mathbb{D}^{\alpha\beta}(x,y):=\mathbb{D}^{\alpha}(x)\otimes\mathbb{D}^{\beta}(y):\ x,y\in G\},\quad \alpha\in A_{G(R)},\ \beta\in A_{G(S)},\tag{I.9}$$

 0022-2488/79/040659-05$01.00

which, however, belong to the standard factor system

$$Q((x,y),(x',y')) = R(x,x')S(y,y'), \tag{I.10}$$

and whose matrix elements take the form

$$\mathbb{D}^{\alpha\beta}_{pr,qs}(x,y) := \mathbb{D}^{\alpha}_{pq}(x)\mathbb{D}^{\beta}_{rs}(y), \quad p,q = 1,2,...,n_\alpha, \quad r,s = 1,2,...,n_\beta. \tag{I.11}$$

Now we are in the position to give a preliminary version of the "subduction problem" in terms of the formula

$$\mathbb{D}^{\alpha\beta}\downarrow G[x]G \sim \sum_{\gamma\in A_{G(RS)}} \oplus\, m_{\alpha\beta;\gamma}\mathbb{D}^{\gamma}, \tag{I.12}$$

where the projective matrix unirreps $\mathbb{D}^{\gamma}$ of G must belong to the factor system

$$T(x,y) := Q((x,x),(y,y)) = R(x,y)S(x,y), \tag{I.13}$$

$$\mathbb{D}^{\gamma}(x)\mathbb{D}^{\gamma}(y) = T(x,y)\mathbb{D}^{\gamma}(xy), \quad \gamma\in A_{G(RS)} = A_{G(T)}. \tag{I.14}$$

Thereby we have to note that the factor system $T = RS$ is, however, in general not equivalent to R, respectively S. The quantities $m_{\alpha\beta;\gamma}$ are called "multiplicities" and give information regarding how many times the projective unirrep $\mathbb{D}^{\gamma}$; $\gamma\in A_{G(RS)}$ is contained in the reducible representation $D^{\alpha\beta}\downarrow G[x]G$. These multiplicies can be calculated by means of the character formula[3]

$$m_{\alpha\beta;\gamma} = \frac{1}{|G|}\sum_{x\in G}\mathbb{X}^{\alpha}(x)\mathbb{X}^{\beta}(x)\mathbb{X}^{\gamma*}(x), \quad \gamma\in A_{G(RS)}, \tag{I.15}$$

where the characters are given by the traces of the corresponding matrices. Thereby it should be noted that in case we change our factor systems R, S, and T by trivial ones the corresponding unimodular factors must appear in the equivalent version of (I.15). The following equation,

$$n_\alpha n_\beta = \sum_{\gamma\in A_{G(RS)}} m_{\alpha\beta;\gamma}\, n_\gamma, \tag{I.16}$$

is the trivial consequence of (I.12), but nevertheless important when checking the multiplicity formula.

As usual we define the CG matrices, for every pair α,β by

$$C^{+}_{\alpha\beta}\,\mathbb{D}^{\alpha\beta}(x)C^{\alpha\beta} = \sum_{\gamma\in A_{G(RS)}} \oplus\, m_{\alpha\beta;\gamma}\mathbb{D}^{\gamma}(x) \quad \text{for all } x\in G, \tag{I.17}$$

where we have introduced the notation

$$\mathbb{D}^{\alpha\beta}(x) := \mathbb{D}^{\alpha\beta}(x,x), \tag{I.18}$$

$$\mathbb{D}^{\alpha\beta}_{pr,qs}(x) := \mathbb{D}^{\alpha}_{pq}(x)\mathbb{D}^{\beta}_{rs}(x). \tag{I.19}$$

Thereby we remark that the CG matrices $C^{\alpha\beta}$ are assumed to be unitary which, however, implies no loss of generality. Equation (I.17) written down in more detail reads

$$\sum_{p,q=1}^{n_\alpha}\sum_{r,s=1}^{n_\beta} C^{\alpha\beta*}_{pr;\gamma wk}\,\mathbb{D}^{\alpha\beta}_{pr,qs}(x)\,C^{\alpha\beta}_{qs;\gamma' w' l} = \delta_{\gamma\gamma'}\,\delta_{ww'}\,\mathbb{D}^{\gamma}_{kl}(x), \tag{I.20}$$

for all $x\in G$ and $w = 1,2,...,m_{\alpha\beta;\gamma}$.

Sometimes a more convenient form of the CG coefficients is used in the following

$$\begin{pmatrix}\alpha & \beta & | & \gamma & w\\ p & r & | & k & \end{pmatrix} := C^{\alpha\beta}_{pr;\gamma wk},$$
$$p = 1,2,...,n_\alpha, \quad r = 1,2,...,n_\beta \quad \gamma\in A_{G(RS)}, \quad w = 1,2,...,m_{\alpha\beta;\gamma}, \quad k = 1,2,...,n_\gamma. \tag{I.21}$$

Hereafter we call the index w a "multiplicity index." Of course, we can imagine that the main difficulties arise from the multiplicities, if we want to determine explicitly the CG coefficients. By utilizing the unitarity of the CG matrices $C^{\alpha\beta}$ Eq. (I.20) can be rewritten as follows:

$$\sum_{q=1}^{n_\alpha}\sum_{s=1}^{n_\beta}\mathbb{D}^{\alpha\beta}_{pr,qs}(x)\,C^{\alpha\beta}_{qs;\gamma wl} = \sum_{k=1}^{n_\gamma}\mathbb{D}^{\gamma}_{kl}(x)C^{\alpha\beta}_{pr;\gamma wk} \quad \text{for all } x\in G \text{ and } w = 1,2,...,m_{\alpha\beta;\gamma}, \tag{I.22}$$

which is, of course, an equivalent form of Eq. (I.20). Equation (I.22) represents for a fixed pair α,β a system of linear equations for the unknown CG coefficients, which is sometimes used as method of their determination. Using once more the unitarity of $C^{\alpha\beta}$ and the orthogonality relation for the matrix elements of the projective unirreps $\mathbb{D}^{\gamma}$ [compare (I.4), (I.7) respectively], we obtain

$$\frac{1}{|G|}\sum_{x\in G}\mathbb{D}^{\alpha}_{pq}(x)\mathbb{D}^{\beta}_{rs}(x)\mathbb{D}^{\gamma}_{kl}{}^{*}(x) = \sum_{w=1}^{m_{\alpha\beta;\gamma}} C^{\alpha\beta}_{pr;\gamma wk}\,C^{*\alpha\beta}_{qs;\gamma wl}, \tag{I.23}$$

which is also used to actually calculate the CG coefficients.[6-12] Equation (I.23) has, however, the unpleasant feature that the multiplicity index w does not occur on the left-hand side. Consequently, the summation about all group elements of G on the left-hand side of (I.23) must imply a loss of information in contrast to (I.20) or (I.22). Finally we mention that all formulas can be easily transferred to ordinary vector representations.

II. REFORMULATION OF THE THEORY

Instead of investigating Eq. (I.23) in more detail we start once again with Eq. (I.22) in order to gain more insight into the problem of how the multiplicity index w could be determined more systematically. To be more concerete, we reinterpret Eq. (I.22) in the following way: We collect the $n_\alpha n_\beta$ matrix elements $C^{\alpha\beta}_{pr;\gamma wk}$ for fixed γ, w, and k to a column vector $\vec{\mathrm{C}}^{\alpha\beta;\gamma w}_{k}$ whose components

$$\{\vec{\mathrm{C}}^{\alpha\beta;\gamma w}_{k}\}_{pr} := C^{\alpha\beta}_{pr;\gamma wk}, \quad p = 1,2,...,n_\alpha, \quad r = 1,2,...,n_\beta \tag{II.1}$$

are just the CG coefficients. Obviously the complex vectors $\vec{\mathrm{C}}^{\alpha\beta;\gamma w}_{k}$ can be seen as elements of a $n_\alpha n_\beta$-dimensional complex Euclidean space $\mathscr{V}^{\alpha\beta}$ with the usual scalar product

$$\langle\vec{\mathrm{A}},\vec{\mathrm{B}}\rangle := \sum_{p=1}^{n_\alpha}\sum_{r=1}^{n_\beta} A^{*}_{pr}B_{pr} \quad \text{for all } \vec{\mathrm{A}},\vec{\mathrm{B}}\in\mathscr{V}^{\alpha\beta}. \tag{II.2}$$

Consequently, Eq. (I.22) turns out to be

$$\mathbb{D}^{\alpha\beta}(x)\,\vec{\mathbb{C}}_k^{\alpha\beta;\gamma w} = \sum_{l=1}^{n_\gamma} \mathbb{D}^{\gamma}_{lk}(x)\vec{\mathbb{C}}_l^{\alpha\beta;\gamma w}$$

for all $x\in G$,

$$\gamma\in A_{G(RS)},\quad w = 1,2,\ldots,m_{\alpha\beta;\gamma},\quad k = 1,2,\ldots,n_\gamma, \tag{II.3}$$

which suggests, together with the unitarity

$$\langle\vec{\mathbb{C}}_k^{\alpha\beta;\gamma w},\vec{\mathbb{C}}_l^{\alpha\beta;\gamma' w'}\rangle = \delta_{\gamma\gamma'}\,\delta_{ww'}\,\delta_{kl}, \tag{II.4}$$

how the multiplicity index w can be determined. First of all let us mention that the vectors $\vec{\mathbb{C}}_k^{\alpha\beta;\gamma w}$ are, due to (II.3), G-adapted vectors, i.e., vectors which transform according to projective unirreps of G belonging to the standard factor system $T = RS$.

In order to be able to determine more systematically the multiplicity index w, we introduce, by means of

$$\mathbb{A}^{\alpha\beta}(G) := \left\{\frac{1}{|G|}\sum_{x\in G} F(x)\,\mathbb{D}^{\alpha\beta}(x) : F(x)\in\mathbb{C}\right\}, \tag{II.5}$$

$n_\alpha n_\beta$-dimensional projective representations of the group algebra $\mathbb{A}(G)$.[18,19] Therefore, the matrices

$$\mathbb{E}_{kl}^{\alpha\beta;\gamma} := \frac{n_\gamma}{|G|}\sum_{x\in G}\mathbb{D}^{\gamma}_{kl}{}^*(x)\mathbb{D}^{\alpha\beta}(x),$$
$$\gamma\in A_{G(RS)},\quad k,l = 1,2,\ldots,n_\gamma, \tag{II.6}$$

are a representation of the units of the group algebra $\mathbb{A}(G)$. Thereby we have to note, if $m_{\alpha\beta;\gamma'} = 0$ for a given $\gamma'\in A_{G(RS)}$ then the corresponding matrices $\mathbb{E}_{kl}^{\alpha\beta;\gamma'}$; $k,l = 1,2,\ldots,n_{\gamma'}$ are identically zero. In the following we summarize the well-known properties of the matrices (II.6) which are different from the zero matrix:

$$\{\mathbb{E}_{kl}^{\alpha\beta;\gamma}\}^+ = \mathbb{E}_{lk}^{\alpha\beta;\gamma}, \tag{II.7}$$

$$\mathbb{E}_{kl}^{\alpha\beta;\gamma}\,\mathbb{E}_{k'l'}^{\alpha\beta;\gamma'} = \delta_{\gamma\gamma'}\,\delta_{lk'}\,\mathbb{E}_{kl'}^{\alpha\beta;\gamma}, \tag{II.8}$$

$$\mathbb{D}^{\alpha\beta}(x)\mathbb{E}_{kl}^{\alpha\beta;\gamma} = \sum_{j=1}^{n_\gamma}\mathbb{D}^{\gamma}_{jk}(x)\mathbb{E}_{jl}^{\alpha\beta;\gamma}, \tag{II.9}$$

$$\mathbb{1} = \sum_{\gamma,k}\mathbb{E}_{kk}^{\alpha\beta;\gamma}. \tag{II.10}$$

$\mathbb{1}$ denotes the unit operator of $\mathscr{V}^{\alpha\beta}$.

As a consequence of the transformation law (II.3) we obtain

$$\mathbb{E}_{jk}^{\alpha\beta;\gamma'}\,\vec{\mathbb{C}}_l^{\alpha\beta;\gamma w} = \delta_{\gamma\gamma'}\,\delta_{kl}\,\vec{\mathbb{C}}_j^{\alpha\beta;\gamma w},$$
$$w = 1,2,\ldots,m_{\alpha\beta;\gamma}, \tag{II.11}$$

which suggests how the multiplicity index w can be determined in principle. The procedure must be as follows. We construct by means of the projection operator $\mathbb{E}_{aa}^{\alpha\beta;\gamma}$ (for a given $\gamma\in A_{G(RS)}$ and an appropriated chosen index a) $m_{\alpha\beta;\gamma}$-dimensional subspaces of $\mathscr{V}^{\alpha\beta}$:

$$\mathscr{V}_a^{\alpha\beta;\gamma} := \{\mathbb{E}_{aa}^{\alpha\beta;\gamma}\vec{\mathbb{A}} : \vec{\mathbb{A}}\in\mathscr{V}^{\alpha\beta}\}, \tag{II.12}$$

$$\dim\mathscr{V}_a^{\alpha\beta;\gamma} = m_{\alpha\beta;\gamma},\quad \gamma\in A_{G(RS)}, \tag{II.13}$$

whose dimension must be independent from the chosen index a. In this connection it should be noted that a clever choice of the index a (depending of course on the explicit structure of the projective unirrep $\mathbb{D}^\gamma$) can greatly simplify the calculation of the CG coefficients, for example the space group CG coefficients. Now it is obvious that any orthonormal basis of $\mathscr{V}_a^{\alpha\beta;\gamma}$,

$$\{\vec{\mathbb{Z}}_a^{\alpha\beta;\gamma w} : w = 1,2,\ldots,m_{\alpha\beta;\gamma}\}, \tag{II.14}$$

$$\vec{\mathbb{Z}}_{pr;\gamma wa}^{\alpha\beta} := \{\vec{\mathbb{Z}}_a^{\alpha\beta;\gamma w}\}_{pr},$$
$$p = 1,2,\ldots,n_\alpha,\quad r = 1,2,\ldots,n_\beta, \tag{II.15}$$

$$\langle\vec{\mathbb{Z}}_a^{\alpha\beta;\gamma w},\vec{\mathbb{Z}}_a^{\alpha\beta;\gamma w'}\rangle = \delta_{ww'}, \tag{II.16}$$

already represents a part of the desired columns of the CG matrices where the other columns must be constructed by means of

$$\vec{\mathbb{Z}}_k^{\alpha\beta;\gamma w} := \mathbb{E}_{ka}^{\alpha\beta;\gamma}\vec{\mathbb{Z}}_a^{\alpha\beta;\gamma w},$$
$$k = 1,2,\ldots,n_\gamma,\quad w = 1,2,\ldots,m_{\alpha\beta;\gamma}, \tag{II.17}$$

so that

$$\mathbb{D}^{\alpha\beta}(x)\vec{\mathbb{Z}}_k^{\alpha\beta;\gamma w} = \sum_{j=1}^{n_\gamma}\mathbb{D}^{\gamma}_{jk}(x)\vec{\mathbb{Z}}_j^{\alpha\beta;\gamma w}, \tag{II.18}$$

$$\langle\vec{\mathbb{Z}}_k^{\alpha\beta;\gamma w},\vec{\mathbb{Z}}_l^{\alpha\beta;\gamma' w'}\rangle = \delta_{\gamma\gamma'}\,\delta_{ww'}\,\delta_{kl}, \tag{II.19}$$

and therefore

$$\sum_{pr,qs} Z^{\alpha\beta *}_{pr;\gamma wk}\mathbb{D}^{\alpha\beta}_{pr,qs}(x)Z^{\alpha\beta}_{qs;\gamma' w' l} = \delta_{\gamma\gamma'}\,\delta_{ww'}\,\mathbb{D}^{\gamma}_{kl}(x)$$

$$\text{for all } x\in G \tag{II.20}$$

is satisfied. The proof of (II.20) is trivial. Obviously the main problem of this procedure consists in determining for each subspace $\mathscr{V}_a^{\alpha\beta;\gamma}$ ($\gamma\in A_{G(RS)}$, a = fixed) of $\mathscr{V}^{\alpha\beta}$ an orthonormal basis. This can be done in any way by Schmidt's procedure. [In this connection we remark that the remaining columns of the CG matrices must be defined in any case by (II.17); otherwise, (II.20) cannot be satisfied.]

Instead of applying Schmidt's procedure we consider in more detail the $m_{\alpha\beta;\gamma}$-dimensional subspaces $\mathscr{V}_a^{\alpha\beta;\gamma}$ of $\mathscr{V}^{\alpha\beta}$, where it is assumed that for every $\gamma\in A_{G(RS)}$ the index a is appropriately chosen. In order to obtain a (not necessarily orthogonal or normalized) basis for $\mathscr{V}_a^{\alpha\beta;\gamma}$, it suffices to apply the projection operator $\mathbb{E}_{aa}^{\alpha\beta;\gamma}$ to each element of the orthonormalized basis

$$\{\vec{\mathbf{B}}_{qs} : q = 1,2,\ldots,n_\alpha;\ s = 1,2,\ldots,n_\beta\}, \tag{II.21}$$

$$\{\vec{\mathbf{B}}_{qs}\}_{pr} = \delta_{pq}\delta_{rs} \tag{II.22}$$

of $\mathscr{V}^{\alpha\beta}$. This yields the following $n_\alpha n_\beta$ vectors:

$$\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)} := \mathbb{E}_{aa}^{\alpha\beta;\gamma}\vec{\mathbf{B}}_{qs},$$
$$q = 1,2,\ldots,n_\alpha,\quad s = 1,2,\ldots,n_\beta \tag{II.23}$$
$$B^{\alpha\beta}_{pr;\gamma(qs)a} = \{\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)}\}_{pr}$$

$$= \frac{n_\gamma}{|G|}\sum_{x\in G}\mathbb{D}^{\alpha}_{pq}(x)\,\mathbb{D}^{\beta}_{rs}(x)\,\mathbb{D}^{\gamma}_{aa}{}^*(x). \tag{II.24}$$

First of all we realize that exactly $m_{\alpha\beta;\gamma}$ linearly independent vectors of the type (II.23) must exist. In order to find out what vectors of the type (II.23) are different from the zero vector and which are linearly independent, we investigate the norm square

$$\|\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)}\|^2 := \langle \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)}, \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)} \rangle$$

$$= \frac{n_\gamma}{|G|} \sum_{x\in G} \mathbb{D}^\alpha_{qq}(x)\, \mathbb{D}^\beta_{ss}(x)\, \mathbb{D}^\gamma_{aa}{}^*(x), \quad \text{(II.25)}$$

and furthermore the scalar product of any two such vectors

$$\langle \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(pr)}, \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)} \rangle = \langle \vec{\mathbf{B}}_{pr}, \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)} \rangle$$

$$= \frac{n_\gamma}{|G|} \sum_{x\in G} \mathbb{D}^\alpha_{pq}(x) \mathbb{D}^\beta_{rs}(x)\, \mathbb{D}^\gamma_{aa}{}^*(x). \quad \text{(II.26)}$$

Equations (II.25) and (II.26) are already [together with (II.17)] the key equations for our approach to fix the multiplicity index w. Namely what we want to do is to identify the multiplicity index w (if possible) with special values of the column indices of the representation $\mathbb{D}^{\alpha\beta}$.

In doing so we start with a vector $\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)}$ whose norm exists. This implies that

$$\vec{\mathbf{C}}_a^{\alpha\beta;\gamma w=(qs)} := \|\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)}\|^{-1}\, \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)}, \quad \text{(II.27)}$$

$$\{\vec{\mathbf{C}}_a^{\alpha\beta;\gamma w=(qs)}\}_{pr} = \|\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)}\|^{-1}\, B^{\alpha\beta}_{pr;\gamma(qs)a} \quad \text{(II.28)}$$

is already one of the desired columns of the CG matrix. Now if we can find, by means of (II.26), a further vector $\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs')} \in \mathscr{V}_a^{\alpha\beta;\gamma}$ with $(qs)\neq(qs')$ whose norm exists and which is orthogonal to the original one, i.e.,

$$\langle \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs)}, \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs')} \rangle = 0, \quad \text{(II.29)}$$

then we will have found a further column of the CG matrix by means of

$$\vec{\mathbf{C}}_a^{\alpha\beta;\gamma w=(qs')} := \|\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs')}\|^{-1} \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(qs')}, \quad \text{(II.30)}$$

since

$$\langle \vec{\mathbf{C}}_a^{\alpha\beta;\gamma w=(qs)}, \vec{\mathbf{C}}_a^{\alpha\beta;\gamma w=(qs')} \rangle = \delta_{(qs),(qs')}. \quad \text{(II.31)}$$

Proceeding in the same way as before, we obtain a set of pairwise orthogonal vectors

$$\{\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(q_v s_v)}:\ v = 1,2,\ldots,\overline{m}_{\alpha\beta;\gamma}\} \quad \text{(II.32)}$$

which are already the desired columns of $C^{\alpha\beta}$, if (II.27) has already been taken into account. In case we find only $\overline{m}_{\alpha\beta;\gamma}$ orthogonal vectors at which

$$\overline{m}_{\alpha\beta;\gamma} < m_{\alpha\beta;\gamma} \quad \text{(II.33)}$$

we have to apply, for the remaining $m_{\alpha\beta;\gamma} - \overline{m}_{\alpha\beta;\gamma}$ vectors, Schmidt's procedure in order to obtain the complete CG matrix. If, however,

$$\overline{m}_{\alpha\beta;\gamma} = m_{\alpha\beta;\gamma}, \quad \text{(II.34)}$$

the multiplicity index can be explained completely by special values of the column indices of the matrix $\mathbb{D}^{\alpha\beta}$, i.e.,

$$w = (q_v s_v), \qquad v = 1,2,\ldots,m_{\alpha\beta;\gamma}. \quad \text{(II.35)}$$

Summarizing the results for $\overline{m}_{\alpha\beta;\gamma} = m_{\alpha\beta;\gamma}$, we write down the matrix elements of $C^{\alpha\beta}$ in a more comprehensive form,

$$C^{\alpha\beta}_{pr;\gamma w=(q_v s_v)a} = \{\vec{\mathbf{C}}_a^{\alpha\beta;\gamma w=(q_v s_v)}\}_{pr}$$

$$= \sqrt{\frac{n_\gamma}{|G|}} \left\{ \sum_{x\in G} \mathbb{D}^\alpha_{q_v q_v}(x)\, \mathbb{D}^\beta_{s_v s_v}(x)\, \mathbb{D}^\gamma_{aa}{}^*(x) \right\}^{-1/2}$$

$$\times \sum_{y\in G} \mathbb{D}^\alpha_{pq_v}(y)\, \mathbb{D}^\beta_{rs_v}(y)\, \mathbb{D}^\gamma_{aa}{}^*(y),$$

$$v = 1,2,\ldots,m_{\alpha\beta;\gamma} \quad p = 1,2,\ldots,n_\alpha, \quad r = 1,2,\ldots,n_\beta. \quad \text{(II.36)}$$

Pairwise orthogonality can be expressed in terms of

$$C^{\alpha\beta}_{q_v s_v;\gamma w=(q_{v'} s_{v'})a} = \delta_{vv'} \left(\frac{n_\gamma}{|G|} \sum_{x\in G} \mathbb{D}^\alpha_{q_v q_v}(x)\, \mathbb{D}^\beta_{s_v s_v}(x)\, \mathbb{D}^\gamma_{aa}{}^*(x) \right)^{1/2},$$

$$v,v' = 1,2,\ldots,m_{\alpha\beta;\gamma}, \quad \text{(II.37)}$$

or in the complete equivalent form

$$\sum_{x\in G} \mathbb{D}^\alpha_{q_v q_{v'}}(x)\, \mathbb{D}^\beta_{s_v s_{v'}}(x)\, \mathbb{D}^\gamma_{aa}{}^*(x) = 0 \Leftrightarrow v \neq v'. \quad \text{(II.38)}$$

Conversely, the problem is now to find $m_{\alpha\beta;\gamma}$ special column indices of $\mathbb{D}^{\alpha\beta}$ such that (II.38) is satisfied. However, we must confess that the problem of finding such a set of indices seems to be an art at this stage of development rather than a science. Whether this task is actually solvable, respectively easy, or not cannot be estimated for the general case, but it is obvious that the explicit form of $\mathbb{D}^{\alpha\beta}\downarrow G\,[x]G$ of the unirreps $\mathbb{D}^\gamma$ $(\gamma\in A_{G(RS)})$ of G and the chosen value for the index a will play an essential role. In anticipation of the following we shall show that in the case of space group CG coefficients the task of identifying the multiplicity index w with special column indices of the Kronecker product can be solved systematically for all cases, because of the special structure of the space group unirreps.

Presupposing that a set of column indices (II.35) is found, such that (II.37) is satisfied, the remaining columns of the CG matrix must be calculated by means of (II.17).

$$C^{\alpha\beta}_{pr;\gamma w=(q_v s_v)j} := \{\vec{\mathbf{C}}_j^{\alpha\beta;\gamma w=(q_v s_v)}\}_{pr}$$

$$= \|\vec{\mathbf{B}}_a^{\alpha\beta;\gamma(q_v s_v)}\|^{-1} \{\mathbb{E}_{ja}^{\alpha\beta;\gamma}\, \vec{\mathbf{B}}_a^{\alpha\beta;\gamma(q_v s_v)}\}_{pr}$$

$$= \sqrt{\frac{n_\gamma}{|G|}} \left\{ \sum_{x\in G} \mathbb{D}^\beta_{q_v q_v}(x)\, \mathbb{D}^\beta_{s_v s_v}(x)\, \mathbb{D}^\gamma_{aa}{}^*(x) \right\}^{-1/2} \quad \text{(II.39)}$$

$$\times \sum_{y\in G} \mathbb{D}^\alpha_{pq_v}(y)\, \mathbb{D}^\beta_{rs_v}(y)\, \mathbb{D}^\gamma_{ja}{}^*(y),$$

$$v = 1,2,\ldots,m_{\alpha\beta;\gamma}, \quad j = 1,2,\ldots,n_\gamma, \quad p = 1,2,\ldots,n_\alpha, \quad r = 1,2,\ldots,n_\beta.$$

These quantities satisfy

$$\mathbb{D}^{\alpha\beta}(x)\vec{\mathbf{C}}_l^{\alpha\beta;\gamma w=(q_v s_v)} = \sum_{k=1}^{n_\gamma} \mathbb{D}^\gamma_{kl}(x)\vec{\mathbf{C}}_k^{\alpha\beta;\gamma w=(q_v s_v)}, \quad \text{(II.40)}$$

$$\langle \vec{\mathbf{C}}_j^{\alpha\beta;\gamma w=(q_v s_v)}, \vec{\mathbf{C}}_l^{\alpha\beta;\gamma' w=(q_{v'} s_{v'})} \rangle = \delta_{\gamma\gamma'}\, \delta_{vv'}\, \delta_{jl}, \quad \text{(II.41)}$$

$$\sum_{pr,qs} C^{\alpha\beta *}_{pr;\gamma w=(q_v s_v)j} \mathbb{D}^{\alpha\beta}_{pr,qs}(x)\, C^{\alpha\beta}_{qs;\gamma' w=(q_{v'},s_{v'})l}$$

$$= \delta_{\gamma\gamma'}\, \delta_{vv'}\, \mathbb{D}^{\gamma}_{jl}(x) \quad \text{for all } x \in G, \tag{II.42}$$

which can be proven directly with the aid of Eqs. (I.3), (I.6), and a corresponding equation for $\mathbb{D}^{\gamma}$. The simplification to ordinary vector representations is obvious.

Finally we have to note that the difficult problem remains in general unsolved, whether we can always trace back the multiplicity index w to special column indices of the corresponding Kronecker product, i.e., whether we can find just $m_{\alpha\beta;\gamma}$ vectors satisfying the orthogonality condition (II.38). For the general case (II.35) we must apply Schmidt's procedure in order to determine the remaining orthogonal vectors. The construction of the other columns of the CG matrix itself has to be carried out in any way by Eq. (II.17).

CONCLUDING REMARKS

In the preceding sections we have demonstrated a general method for obtaining CG coefficients for a finite group and where in particular the multiplicity problem is resolved to the task of finding pairwise orthogonal vectors, so that special column indices of the Kronecker product can be chosen as a multiplicity index. Whether any multiplicity problem can be solved in this manner cannot be answered for the general case. However, when applying the present method to determine space group CG coefficients we shall show that for all cases we can derive simple defining equations.

the multiplicity index without reference to a special space group.

[1]G.W. Mackey, Acta Mathematica **99**, 265 (1958).
[2]Ch. W. Curtis, I. Reiner, *Representation Theory of Finite Groups and Associative Algebras* (Wiley, New York, 1962).
[3]S.L. Altmann, *Induced Representation in Crystals and Molecules* (Academic, London, 1977).
[4]M. Hamermesh, *Group Theory and Its Applications to Physical Problems* (Addison-Wesley, Reading, Mass., 1962).
[5]G.F. Koster, J.O. Dimmock, R.G. Wheeler, and H. Statz, *Properties of the Thirty-two Point Groups* (Technology Press, Cambridge, Mass., 1963).
[6]D.B. Litvin and J. Zak, J. Math. Phys. **9**, 212 (1968).
[7]I. Sakata, J. Math. Phys. **15**, 1702 (1974).
[8]J.L. Birman, "Theory of Crystal Space Groups and Infra-red Raman Processes of Insulated Crystals," in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1974).
[9]R. Berenson and J.L. Birman, J. Math. Phys. **16**, 227 (1975).
[10]R. Berenson, I. Itzkan, and J.L. Birman, J. Math. Phys. **16**, 236 (1975).
[11]P. Gard, J. Phys. **A 6**, 1837 (1973).
[12]S. Schindler and R. Mirman, J. Math. Phys. **18**, 1678 (1977).
[13]S. Schindler and R. Mirman, J. Math. Phys. **18**, 1697 (1977).
[14]L.K. Saulevich, D.T. Sviridov, and Yu.F. Smirnov, Sov. Phys. Crystallogr. **15**, 355 (1970).
[15]L. Jansen and M. Boon, *Theory of Finite Groups, Applications in Physics* (North-Holland, Amsterdam, 1967).
[16]R. Dirl and P. Kasperkovitz, *Gruppentheorie, Anwendungen in der Atom- und Festkörperphysik* (Vieweg, Braunschweig, 1977).
[17]D.J. Klein, "Finite Groups and Semisimple Algebras in Quantum Mechanics," in *Group Theory and Its Applications*, edited by M. Loebl (Academic, New York, 1975), Vol. III.
[18]R. Dirl, J. Math. Phys. **18**, 2065 (1977).
[19]R. Dirl, "Induced projective representations," in *Group Theoretical Methods in Physics*, edited by R.T. Sharp and B. Kolman (Academic, New York, 1977).

Multiplicities for space group representations

R. Dirl

Institut für Theoretische Physik, TU Wien, A-1040 Wien, Karlsplatz 13, Austria
(Received 28 December 1977)

The multiplicity formula for nonsymmorphic space group representations is reinvestigated by using explicitly projective representations for the little cogroups $P^{\vec{q}} \simeq G^{\vec{q}}/T$. Thereby useful identities and relations concerning the wave vector selection rules are derived for various cases which may occur for the elements of the Brillouin zone. These relations allow for nearly all cases a closed expression for the multiplicity without reference to a special space group.

INTRODUCTION

It is well known that a systematic calculation of CG coefficients for a given group presupposes in some way the explicit knowledge of the corresponding multiplicities. Since we want to determine space group CG coefficients by means of the proposed method,[1] we have to consider the multiplicity formula for space group representations. This formula gives first information for the theory of selection rules in crystals.

Instead of considering the well-known multiplicity formula (4.7.28) of Ref. 2, which has been successfully applied by various authors [3-6] in order to calculate the multiplicities for several examples, we discuss an equivalent version of the multiplicity formula. Thereby we start from trivial cases and proceed to the most complicated case, in order to gain more insight into the structure of the wavevector selection rules[3-7] and to derive useful relations which should greatly simplify the calculation of the space group CG coefficients.

The organization of this paper is as follows: In Sec. I we summarize the basic notation and properties of the space group unirreps by using explicitly projective representations for the little cogroups.[8] In Sec. II we derive a general expression for the multiplicity formula for nonsymmorphic space groups where the characters of the projective unirreps essentially enter. This formula is, however, not identical to (4.7.29) of Ref. 2. In Sec. III. A we specialize our multiplicity formula to the special case where $\vec{q}$ and $\vec{q}'$ are assumed to belong to general stars. In Sec. III.B we consider the more general case where $\vec{q}$ is an element of a general star, but $\vec{q}'$ of a star of higher symmetry, i.e., $P^{\vec{q}'}$ contains more than one group element. The most complicated case is discussed in Sec. III.C where both vectors $\vec{q}$ and $\vec{q}'$ are assumed to belong to stars of higher symmetry. Apart from two special cases (where the corresponding multiplicities can be calculated quite generally without reference to a special space group like in the previous cases), we arrive at the most complicated case and obtain a multiplicity formula which is identical to (4.7.29) of Ref. 2, but which probably contains more information.

I. UNIRREPS OF NONSYMMORPHIC SPACE GROUPS

In this section we briefly summarize the basic notation, definitions, and properties of unirreps of nonsymmorphic space groups, which are used throughout this and the following papers. We recall that a space group G consists of all elements whose corresponding symmetry operations leave the crystal structure invariant:

$$G = \{(\alpha|\vec{\tau}(\alpha) + \vec{t}):\alpha\in P, \vec{t}\in T\}, \tag{I.1}$$

$$(\alpha|\vec{\tau}(\alpha) + \vec{t})(\beta|\vec{\tau}(\beta) + \vec{t}') = (\alpha\beta|\vec{\tau}(\alpha\beta) + \vec{t}(\alpha,\beta) + D(\alpha)\vec{t}' + \vec{t}); \tag{I.2}$$

$$\vec{t}(\alpha,\beta) = \vec{\tau}(\alpha) + D(\alpha)\vec{\tau}(\beta) - \vec{\tau}(\alpha\beta). \tag{I.3}$$

The symbol $\vec{\tau}(\alpha)$ denotes nonprimitive lattice translations which are uniquely determined when the multiplication law of the nonsymmorphic space group is established. The vectors $\vec{t}(\alpha,\beta)$ defined by Eq. (I.3) are elements of the translation group T and their appearance is typical for nonsymmorphic space groups. Finally $D := \{D(\alpha): \alpha\in P\}$ is a n-dimensional (n = dimension of the crystal lattice) orthogonal representation of the point group P (being isomorphic to the factor group G/T) of the crystal.

It is known[7,8] that the matrix elements of the (vector) unirreps of a nonsymmorphic space group G can be written in the following form:

$$D^{(\kappa,\vec{q})\downarrow G}_{\sigma a,\sigma' b}(\beta|\vec{\tau}(\beta) + \vec{t}) = \Delta^{\vec{q}}(\sigma,\beta\sigma')\exp[-i\vec{q}(\sigma)\cdot\{\vec{t} + \vec{t}(\beta,\sigma') + \vec{t}(\sigma,\sigma^{-1}\beta\sigma')\}]\mathbb{D}^{\kappa}_{ab}(\sigma^{-1}\beta\sigma')$$

$$\vec{q}\in\Delta BZ;\quad \kappa\in A_{P^{\vec{q}}(R^{\vec{q}})}\quad a,b, = 1,2,\cdots,n_\kappa\quad \sigma,\sigma'\in:P^{\vec{q}}, \tag{I.4}$$

$$P^{\vec{q}} = \{\alpha:D(\alpha)\vec{q} = \vec{q} + \vec{Q}\{\vec{q}(\alpha)\},\ \alpha\in P\}, \tag{I.5}$$

$$\Delta^{\vec{q}}(\gamma,\gamma') = \delta_{\gamma P^{\vec{q}},\gamma' P^{\vec{q}}},\quad \text{for all } \gamma,\gamma'\in P\supseteq P^{\vec{q}}\simeq G^{\vec{q}}/T, \tag{I.6}$$

$$\vec{q}(\gamma) := D(\gamma)\vec{q},\quad \text{for all } \gamma\in P. \tag{I.7}$$

Thereby $G^{\vec{q}}$ denotes the group of the $\vec{q}$ vectors, $\vec{Q}\{\vec{q}(\alpha)\}$ reciprocal lattice vectors, σ, $\sigma'\in P:P^{\vec{q}}$ left coset representatives of $P^{\vec{q}}$ with respect to P, and $\mathbb{D}^{\kappa} := \{\mathbb{D}^{\kappa}(\alpha):\alpha\in P^{\vec{q}}\}$ n_κ-dimensional projective unirreps of $P^{\vec{q}}\simeq G^{\vec{q}}/T$ which belong to the standard factor system

$$R^{\vec{q}}(\alpha,\alpha') = e^{-i\vec{q}\cdot\vec{t}(\alpha,\alpha')},\quad \text{for all } \alpha,\alpha'\in P^{\vec{q}}. \tag{I.8}$$

Furthermore ΔBZ is the fundamental domain of the Brillouin zone and $A_{P^{\vec{q}}(R^{\vec{q}})}$ denotes the set of all equivalence classes of the projective unirreps of $P^{\vec{q}}$ which belong to the factor system $R^{\vec{q}}$.

 0022-2488/79/040664-07$01.00

In order to simplify the following considerations we change every factor system $R^{\vec{q}}$ by means of

$$\mathbb{D}^{\kappa}(\alpha) = e^{-i\vec{q}\cdot\vec{\tau}(\alpha)}\mathbb{R}^{\kappa}(\alpha), \quad \text{for all } \alpha\in P^{\vec{q}}, \tag{I.9}$$

to the new ones

$$S^{\vec{q}}(\alpha,\beta) = \exp[-i\vec{q}\cdot(D(\alpha) - \mathbb{1})\vec{\tau}(\beta)]$$

$$= \exp[i\vec{Q}\{\vec{q}(\alpha)\}\cdot[\vec{\tau}(\alpha\beta) - \vec{\tau}(\alpha)]] \quad \text{for all } \alpha,\beta\in P^{\vec{q}}. \tag{I.10}$$

Therefore, the matrix elements of the $|P{:}P^{\vec{q}}|n_\kappa$-dimensional unirreps of G are rewritten as follows:

$$D^{(\kappa,\vec{q})\uparrow G}_{\sigma a,\sigma' b}(\beta\,|\,\vec{\tau}(\beta) + \vec{t})$$

$$= \Delta^{\vec{q}}(\sigma,\beta\sigma')e^{-i\vec{q}(\sigma)\cdot\vec{t}}B^{\vec{q}}_{\sigma,\sigma'}(\beta)\mathbb{R}^{\kappa}_{ab}(\sigma^{-1}\beta\sigma'),\vec{q}\in\Delta BZ;$$

$$\kappa\in A_{P^{\vec{q}}(S^{\vec{q}})},\ a,b = 1,2,\cdots n_\kappa \qquad \sigma,\sigma'\in P{:}P^{\vec{q}} \tag{I.11}$$

$$B^{\vec{q}}_{\sigma,\sigma'}(\beta) = \exp[-i\vec{q}(\sigma)\cdot\{\vec{\tau}(\beta) + D(\beta)\vec{\tau}(\sigma') - \vec{\tau}(\sigma)\}],$$

$$\text{for all } \beta\in P. \tag{I.12}$$

The main reason for preferring (I.11) instead of (I.9) is that for all $\vec{q}$'s not lying on the surface of ΔBZ the factor systems $S^{\vec{q}}$ reduce to one, which implies that the corresponding unirreps of $P^{\vec{q}}$ must be ordinary vector representations. It should be noted that the case of symmorphic space groups is contained in the formulas in a consistent way.

II. MULTIPLICITIES FOR NONSYMMORPHIC SPACE GROUPS

Equation (I.12) of Ref. 1 has to be rewritten for space group representations,

$$D^{(\kappa,\vec{q})\uparrow G}(\beta\,|\,\vec{\tau}(\beta) + \vec{t})\otimes D^{(\kappa',\vec{q}')\uparrow G}(\beta\,|\,\vec{\tau}(\beta) + \vec{t})$$

$$\sim \sum_{\substack{\vec{q}''\in\Delta BZ\\ \kappa''\in A_{P_{\vec{q}''}(S_{\vec{q}''})}}} \oplus\, m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa'',\vec{q}'')}D^{(\kappa'',\vec{q}'')\uparrow G}(\beta\,|\,\vec{\tau}(\beta) + \vec{t}). \tag{II.1}$$

In order to be able to calculate the multiplicities $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa'',\vec{q}'')}$ we need the characters of the corresponding unirreps. These characters are obtainable from (I.11)

$$X^{(\kappa,\vec{q})\uparrow G}(\beta\,|\,\vec{\tau}(\beta) + \vec{t})$$

$$= \sum_{\sigma\in P:P^{\vec{q}}} \Delta^{\vec{q}}(\sigma,\beta\sigma)B^{\vec{q}}_{\sigma,\sigma}(\beta)\mathbb{X}^{\kappa}(\sigma^{-1}\beta\sigma)e^{-\vec{q}(\sigma)\cdot\vec{t}} \tag{II.2}$$

$$\mathbb{X}^{\kappa}(x) = \operatorname{trace}\mathbb{R}^{\kappa}(x), \quad \text{for all } x\in P^{\vec{q}}, \tag{II.3}$$

where $\mathbb{X}^{\kappa}$ denotes the character of projective unirreps of $P^{\vec{q}}$ which belong to the factor system $S^{\vec{q}}$. Instead of dirrectly applying Eq. (I.15) of Ref. 1 we start from

$$X^{(\kappa,\vec{q})\uparrow G}(\beta\,|\,\vec{\tau}(\beta) + \vec{t})X^{(\kappa',\vec{q}')\uparrow G}(\beta\,|\,\vec{\tau}(\beta) + \vec{t})$$

$$= \sum_{\vec{q}'',\kappa''} m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa'',\vec{q}'')}X^{(\kappa'',\vec{q}'')\uparrow G}(\beta\,|\,\vec{\tau}(\beta) + \vec{t}). \tag{II.4}$$

which is a consequence of Eq. (II.1). Multiplying both sides of Eq. (II.4) by $\exp(i\vec{q}_0\cdot\vec{t})$ with $\vec{q}_0\in\Delta BZ$ and summing about all $\vec{t}\in T$ we have to take the orthogonality relations

$$\frac{1}{|T|}\sum_{t\in T}\exp[-i\{\vec{q}(\sigma) + \vec{q}'(\sigma') - \vec{q}_0\}\cdot\vec{t}\,]$$

$$= \delta_{\vec{q}(\sigma) + \vec{q}'(\sigma'),\vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')]} \tag{II.5}$$

into account in order to obtain

$$\sum_{\kappa''\in A_{P_{\vec{q}_0}(S_{\vec{q}_0})}} m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa'',\vec{q}_0)}B^{\vec{q}_0}_{e,e}(\beta)\mathbb{X}^{\kappa''}(\beta)\Delta^{\vec{q}_0}(e,\beta e)$$

$$= \sum_{\sigma\in P:P_{\vec{q}},\ \sigma'\in P:P_{\vec{q}'}} \delta_{\vec{q}(\sigma) + \vec{q}'(\sigma'),\vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')]}$$

$$\times\Delta^{\vec{q}}(\sigma,\beta\sigma)\Delta^{\vec{q}'}(\sigma',\beta\sigma')B^{\vec{q}}_{\sigma,\sigma}(\beta)B^{\vec{q}'}_{\sigma',\sigma'}(\beta)\mathbb{X}^{\kappa}(\sigma^{-1}\beta\sigma)\mathbb{X}^{\kappa'}(\sigma'^{-1}\beta\sigma'), \tag{II.6}$$

where e denotes the identity element of P. Relation (II.5) is sometimes called the "wave vector selection rules."[3,4,7] In this connection we remark that the reciprocal lattice vectors $\vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')]$ should not be confused with $\vec{Q}\{\vec{q}(\alpha)\}$ which enter into the definitionof $G^{\vec{q}}$ (respectively, $P^{\vec{q}}$). Consequently the general formula for the multiplicities takes the form

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = \sum_{\sigma,\sigma'}\delta_{\vec{q}(\sigma) + \vec{q}'(\sigma'),\vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')]}$$

$$\times\frac{1}{|P^{\vec{q}_0}|}\sum_{\beta\in P^{\vec{q}_0}}\Delta^{\vec{q}_0}(e,\beta e)\Delta^{\vec{q}}(\sigma,\beta\sigma)\Delta^{\vec{q}'}(\sigma',\beta\sigma')$$

$$\times B^{\vec{q}_0 *}_{e,e}(\beta)B^{\vec{q}}_{\sigma,\sigma}(\beta)B^{\vec{q}'}_{\sigma',\sigma'}(\beta)^{\prime}\mathbb{X}^{\kappa}(\sigma^{-1}\beta\sigma)\mathbb{X}^{\kappa'}(\sigma'^{-1}\beta\sigma')\mathbb{X}^{\kappa_0 *}(\beta), \tag{II.7}$$

where we have used the orthogonality relations for the characters of the projective unirreps of $P^{\vec{q}_0}$ which, however, must belong to the original factor system $R^{\vec{q}_0}$ [compare the comments to Eq. (I.15) of Ref. 1]. Now we realize once again that one has to be very careful if using orthogonality relations of characters (later, of matrix elements) of projective unirreps. This fact explains the occurence of a part of the unimodular factors on the right-hand side of Eq. (II.7). Finally we mention that

$$\dim\{D^{(\kappa,\vec{q})\uparrow G}\otimes D^{(\kappa',\vec{q}')\uparrow G}\} = |P{:}P^{\vec{q}}|n_\kappa|P{:}P^{\vec{q}'}|n_{\kappa'}$$

$$= \sum_{\substack{\vec{q}_0\in\Delta BZ\\ \kappa_0\in A_{P_{\vec{q}_0}(S_{\vec{q}_0})}}} m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}|P{:}P^{\vec{q}_0}|n_{\kappa_0} \tag{II.8}$$

must be valid in general in analogy to the general formula (I.16) of Ref. 1. By means of the definitions (intersections of conjugate subgroups)

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} := \sigma P^{\vec{q}}\sigma^{-1}\cap\sigma' P^{\vec{q}'}\sigma'^{-1}\cap P^{\vec{q}_0}, \quad \sigma\in P{:}P^{\vec{q}}, \quad \sigma'\in P{:}P^{\vec{q}'}, \tag{II.9}$$

which are subgroups of P, we can omit the product of the generalized Kronecker delta's in Eq. (II.7),

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = \sum_{\sigma,\sigma}\delta_{\vec{q}(\sigma) + \vec{q}'(\sigma'),\vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')]}$$

$$\times\frac{1}{|P^{\vec{q}_0}|}\sum_{\beta\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}} B^{\vec{q}_0 *}_{e,e}(\beta)B^{\vec{q}}_{\sigma,\sigma}(\beta)B^{\vec{q}'}_{\sigma',\sigma'}(\beta)$$

$$\times\mathbb{X}^{\kappa}(\sigma^{-1}\beta\sigma)\mathbb{X}^{\kappa'}(\sigma'^{-1}\beta\sigma')\mathbb{X}^{\kappa_0}(\beta)^{*}. \tag{II.10}$$

III. DISCUSSION OF VARIOUS CASES

We start from simple cases and proceed to more compli cated ones when investigating the multiplicity formula in order to gain more insight into the structure of the wave vector selection rules. For this purpose we introduce the following useful definition of the product sets,

$$P(\vec{q},\vec{q}') := \{(\sigma,\sigma'):\sigma\in P:P^{\vec{q}},\sigma\in P:P^{\vec{q}'}\};\quad \vec{q},\vec{q}'\in\Delta BZ. \qquad \text{(III.1)}$$

A. $\vec{q},\vec{q}'$ = general stars

In this case Eq. (II.10) reduces, because $P^{\vec{q}} = P^{\vec{q}'} = \{e\}$, to

$$m_{(0,\vec{q})(0,\vec{q}');(\kappa_0,\vec{q}_0)} = \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\sum_{\sigma,\sigma'}\delta_{\vec{q}(\sigma)+\vec{q}'(\sigma'),\vec{q}_0+\vec{Q}[\vec{q}(\sigma)+\vec{q}'(\sigma')]}, \qquad \text{(III.2)}$$

where we have furthermore used the symbol 0 for the identical representation of $P^{\vec{q}}$ (respectively $P^{\vec{q}'}$).

If $\vec{q}_0$ is itself an element of a general star, i.e., $P^{\vec{q}_0} = \{e\}$, we have to look for pairs $(\sigma,\sigma')\in P(\vec{q},\vec{q}')$ such that

$$\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')] \qquad \text{(III.3)}$$

is satisfied for the fixed $\vec{q}_0\in\Delta BZ$. We call this set

$$P(\vec{q},\vec{q}';\vec{q}_0) := \{(\sigma,\sigma'):\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')];(\sigma,\sigma')\in P(\vec{q},\vec{q}')\}, \qquad \text{(III.4)}$$

whose determination requires only geometrical considerations. This implies for the multiplicity

$$m_{(0,\vec{q})(0,\vec{q}');(0,\vec{q}_0)} = |P(\vec{q},\vec{q}';\vec{q}_0)|, \qquad \text{(III.5)}$$

where $|P(\vec{q},\vec{q}';\vec{q}_0)|$ is the order of the set $P(\vec{q},\vec{q}';\vec{q}_0)$.

If $\vec{q}_0$ belongs to a star of higher symmetry, i.e., if $P^{\vec{q}_0}$ is a nontrivial subgroup of P,

$$\{e\}\subset P^{\vec{q}_0}\subseteq P, \qquad \text{(III.6)}$$

then it follows for a fixed pair $(\sigma,\sigma')\in P(\vec{q},\vec{q}')$ with

$$\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')] \qquad \text{(III.7)}$$

that

$$\vec{q}(x\sigma) + \vec{q}'(x\sigma') = \vec{q}_0(x) + D(x)\vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')] = \vec{q}_0 + \vec{Q}\{\vec{q}_0(x)\} + D(x)\vec{Q}[\vec{q}(\sigma + \vec{q}'(\sigma')]$$

for all $x\in P^{\vec{q}_0}$ (III.8)

must be correct. This has (because $P^{\vec{q}} = P^{\vec{q}'} = \{e\}$) as a consequence that

$$\{(x\sigma,x\sigma'):x\in P^{\vec{q}_0}\}\subset P(\vec{q},\vec{q}') \qquad \text{(III.9)}$$

is already a subset of pairs of left coset representatives which satisfy (III.7). If there exists another pair $(\vec{\tau},\vec{\tau}')\in P(\vec{q},\vec{q}')$ with

$$\vec{q}(\tau) + \vec{q}'(\tau') = \vec{q}_0 + \vec{Q}[\vec{q}(\tau) + \vec{q}'(\tau')] \qquad \text{(III.10)}$$

which is not contained in (III.9) then the same argumentation holds as before, i.e.,

$$\{(y\tau,y\tau'):y\in P^{\vec{q}_0}\}\subset P(\vec{q},\vec{q}') \qquad \text{(III.11)}$$

is also a subset of pairs of coset representatives which satisfy (III.10) and has to be taken into account when calculating the multiplicities according to (III.2). In this connection we have to note that such a possibility cannot be excluded in general for space group representations. Now it can be easily proven that the intersection of $\{(x\sigma,x\sigma'):x\in P^{\vec{q}_0}\}$ with $\{y\tau,y\tau'):y\in P^{\vec{q}_0}\}$ either is empty or both sets coincide. In order to show this proposition we have to use

$$\tau = z\sigma,\quad \tau' = z'\sigma',\quad z',z'\in P \qquad \text{(III.12)}$$

for (III.10) which yields

$$\vec{q}(\sigma) + \vec{q}'(z^{-1}z'\sigma') = \vec{q}_0(z^{-1}) + D(z^{-1})\vec{Q}[\vec{q}(\tau) + \vec{q}'(\tau')]. \qquad \text{(III.13).}$$

Now we substract (III.7) from this equation and obtain

$$\vec{q}'(z^{-1}z'\sigma') - \vec{q}'(\sigma') = \vec{q}_0(z^{-1}) - \vec{q}_0 + D(z^{-1})\vec{Q}[\vec{q}(\tau) + \vec{q}'(\tau')] - \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')], \qquad \text{(III.14)}$$

which proves our proposition

$$z\in P^{\vec{q}_0}\Leftrightarrow z = z'\in P^{\vec{q}_0}, \qquad \text{(III.15)}$$

$$z\notin P^{\vec{q}_0}\Leftrightarrow z\neq z'\notin P^{\vec{q}_0}, \qquad \text{(III.16)}$$

since for the first case the left-hand side must be a reciprocal lattice vector, whereas for the second case the left-hand side must be different from a reciprocal lattice vector. Analogously to (III.4) we introduce by means of

$$P(\vec{q},\vec{q}';\vec{q}_0) := \{(\sigma_j,\sigma_j'):\vec{q}(\sigma_j) + \vec{q}'(\sigma_j') = \vec{q}_0 + \vec{Q}\,[\vec{q}(\sigma_j) + \vec{q}'(\sigma_j')]\} \qquad \text{(III.17)}$$

with

$$\{(x\sigma_j,x\sigma_j') : x\in P^{\vec{q}_0}\}\cap\{(y\sigma_k,y\sigma_k'):y\in P^{\vec{q}_0}\} = \emptyset,\quad \text{for all } j\neq k, \qquad \text{(III.18)}$$

a set of pairs of coset representatives which generate pairwise empty intersections. As was already pointed out we cannot expect for space group representations that the order of the set $P(\vec{q},\vec{q}';\vec{q}_0)$ is always one. This implies for the multiplicity

$$m_{(0,\vec{q})(0,\vec{q}');(\kappa_0,\vec{q}_0)} = n_{\kappa_0}|P(\vec{q},\vec{q}';\vec{q}_0)| \qquad \text{(III.19)}$$

since the sum in (III.2) reduces to

$$m_{(0,\vec{q})(0,\vec{q}');(\kappa_0,\vec{q}_0)} = \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\sum_{\substack{(\sigma_j,\sigma_j')\in P(\vec{q},\vec{q}';\vec{q}_0)\\ x\in P^{\vec{q}_0}}}\delta_{\vec{q}(x\sigma_j)+\vec{q}'(x\sigma_j'),\vec{q}_0+\vec{Q}[\vec{q}(x\sigma_j)+\vec{q}'(x\sigma_j')]}. \qquad \text{(III.20)}$$

B. $\vec{q}$ = general star, $\vec{q}'$ = star of higher symmetry

For this case, which is of course more complicated that the first one, Eq. (II.10) takes (because of $P^{\vec{q}} = \{e\}$ and $P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} = \{e\}$) the form

$$m_{(0,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = \frac{n_{\kappa'}n_{\kappa_0}}{|P^{\vec{q}_0}|}\sum_{\substack{\sigma\in P\\ \sigma'\in P:P^{\vec{q}'}}}\delta_{\vec{q}(\sigma)+\vec{q}'(\sigma'),\vec{q}_0+\vec{Q}[\vec{q}(\sigma)+\vec{q}'(\sigma')]}, \qquad \text{(III.21)}$$

which can be simplified in a similar way as the previous one. We distinguish once again two cases, namely $\vec{q}_0$ belongs to a general star or to a star of higher symmetry.

For the first possibility we have to look, since $P^{\vec{q}_0} = \{e\}$, for pairs $(\sigma,\sigma')\in P(\vec{q},\vec{q}')$ such that

$$\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')], \quad \sigma\in P, \ \sigma'\in P{:}P^{\vec{q}'} \tag{III.22}$$

is satisfied for the fixed $\vec{q}_0\in\Delta BZ$. We denote this set by

$$P(\vec{q},\vec{q}';\vec{q}_0) := \{(\sigma,\sigma'){:}\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')], \quad (\sigma,\sigma')\in P(\vec{q},\vec{q}')\}, \tag{III.23}$$

which implies for the multiplicity

$$m_{(0,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)} = n_{\kappa'}|P(\vec{q},\vec{q}';\vec{q}_0)|. \tag{III.24}$$

The second possibility, namely that $\vec{q}_0$ belongs to a star of higher symmetry

$$\{e\}\subset P^{\vec{q}_0}\subseteq P, \tag{III.25}$$

has to be investigated with more care. Starting with a fixed pair $(\sigma,\sigma')\in P(\vec{q},\vec{q}')$ satisfying

$$\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')] \tag{III.26}$$

it follows immediately that

$$\begin{aligned}\vec{q}(x\sigma) + \vec{q}'(x\sigma') &= \vec{q}_0(x) + D(x)\vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')]\\ &= \vec{q}_0 + \vec{Q}\{\vec{q}_0(x)\} + D(x)\vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')]\\ &\text{for all } x\in P^{\vec{q}_0}\end{aligned} \tag{III.27}$$

must be correct. However, in contrast to (III.9) the set

$$\{(x\sigma,x\sigma'){:}x\in P^{\vec{q}_0}\}\not\subset P(\vec{q},\vec{q}') \tag{III.28}$$

is in general not a subset of $P(\vec{q},\vec{q}')$, since $x\sigma'\notin P{:}P^{\vec{q}'}$ is in general (because of $\{e\}\subset P^{\vec{q}'}\subseteq P$) not an element from the set of left coset representatives $P{:}P^{\vec{q}'}$, whereas $x\sigma\in P{:}P^{\vec{q}} = P$ must in any event be an element of the left coset representatives. For this reason we introduce the notation

$$x\sigma' = \sigma'_x x', \quad \text{with } \sigma'_x\in P{:}P^{\vec{q}'} \quad \text{and} \quad x'\in P^{\vec{q}'}, \tag{III.29}$$

which concerns the left coset representatives of $P^{\vec{q}'}$ with respect to P. Now we prove by means of

$$\begin{aligned}\vec{q}(x\sigma) - \vec{q}(y\sigma) &= \vec{q}'(y\sigma') - \vec{q}'(x\sigma') + \vec{Q}[\vec{q}(x\sigma) + \vec{q}'(x\sigma')]\\ &\quad - \vec{Q}[\vec{q}(y\sigma) + \vec{q}'(y\sigma')] \quad \text{for all } x,y\in P^{\vec{q}_0}\end{aligned} \tag{III.30}$$

that for any pair $x,y\in P^{\vec{q}_0}$

$$x\neq y\Leftrightarrow\sigma'_x\neq\sigma'_y \tag{III.31}$$

is satisfied, which implies that the order of the set

$$|\{\sigma'_x = x\sigma'x'^{-1}{:}x\in P^{\vec{q}_0}\}| = |P^{\vec{q}_0}| \tag{III.32}$$

is equal to the order of the group $P^{\vec{q}_0}$. Therefore,

$$|P^{\vec{q}_0}|\leqslant|P{:}P^{\vec{q}'}|, \tag{III.33}$$

which is, however, only correct if $P^{\vec{q}} = \{e\}$. A further important consequence of (III.31) is

$$x\neq y\Leftrightarrow\delta_{\vec{q}(x\sigma) + \vec{q}'(y\sigma' y'^{-1}),\vec{q}_0 + \vec{Q}[\vec{q}(x\sigma) + \vec{q}'(y\sigma' y'^{-1})]} = 0 \quad \text{for all } x,y\in P^{\vec{q}} \tag{III.34}$$

which has to be taken into account when calculating the multiplicities. Now if there exists a further pair $(\tau,\tau')\in P(\vec{q},\vec{q}')$ which is not contained in the set

$$\{(x\sigma,x\sigma'x'^{-1}){:}x\in P^{\vec{q}_0}\}\subset P(\vec{q},\vec{q}'), \tag{III.35}$$

such that

$$\vec{q}(\tau) + \vec{q}'(\tau') = \vec{q}_0 + \vec{Q}[\vec{q}(\tau) + \vec{q}'(\tau')], \tag{III.36}$$

then the same argumentation holds as before, i.e.,

$$\{(y\tau,y\tau'y'^{-1}){:}y\in P^{\vec{q}_0}\}\subset P(\vec{q},\vec{q}') \tag{III.37}$$

is also a subset of pairs of left coset representatives which satisfy (III.36). As in the previous part of this section we show that the intersection of (III.35) with (III.37) is either empty or both sets coincide. For this purpose we set

$$\tau = z\sigma, \quad \tau' = z'\sigma', \quad \text{with } z,z'\in P, \tag{III.38}$$

which yields for (III.36)

$$\vec{q}(\sigma) + \vec{q}'(z^{-1}z'\sigma') = \vec{q}_0(z^{-1}) + D(z^{-1})\vec{Q}[\vec{q}(\tau) + \vec{q}'(\tau')]. \tag{III.39}$$

Substracting (III.26) from this equation we obtain

$$\begin{aligned}&\vec{q}'(z^{-1}z'\tau') - \vec{q}'(\tau')\\ &= \vec{q}_0(z^{-1}) - \vec{q}_0 + D(z^{-1})\vec{Q}[\vec{q}(\tau) + \vec{q}'(\tau)] - \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')],\end{aligned} \tag{III.40}$$

which together with (III.31) proves our proposition

$$z\in P^{\vec{q}_0}\Leftrightarrow z = z'\in P^{\vec{q}_0}, \tag{III.41}$$

$$z\notin P^{\vec{q}_0}\Leftrightarrow z\neq z'\notin P^{\vec{q}_0}. \tag{III.42}$$

Now we introduce by means of

$$P(\vec{q},\vec{q}';\vec{q}_0) := \{(\sigma_j,\sigma'_j){:}\vec{q}(\sigma_j) + \vec{q}'(\sigma'_j) = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma_j) + \vec{q}'(\sigma'_j)]\} \tag{III.43}$$

with the property

$$\{(x\sigma_j,x\sigma'_jx'^{-1}){:}x\in P^{\vec{q}_0}\}\cap\{(y\sigma_k,y\sigma'_ky'^{-1}){:}y\in P^{\vec{q}_0}\} = \emptyset, \quad \text{for all } j\neq k, \tag{III.44}$$

a set of pairs of coset representatives which generate pairwise empty intersections. Hence we obtain for the multiplicity

$$m_{(0,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = n_{\kappa'}n_{\kappa_0}|P(\vec{q},\vec{q}';\vec{q}_0)|, \tag{III.45}$$

since the sum in (III.21) reduces to the same expression as in (III.20).

Now we are in the position to consider an interesting special case. If

$$P^{\vec{q}'} = P, \tag{III.46}$$

then it follows from (III.33) and (II.8) that

$$P^{\vec{q}_0} = \{e\} \quad \text{and} \quad |P(\vec{q},\vec{q}';\vec{q}_0)| = 1, \tag{III.47}$$

and therefore

$$m_{(0,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)} = n_{\kappa'}. \tag{III.48}$$

C. $\vec{q},\vec{q}'$ = stars of higher symmetry

As already pointed out, this case is the most complicated one, and we have to start from the general formula (II.10). Of course the cases A and B must be contained in the general formula. As in the previous cases we distinguish several cases.

For the simplest case, where $\vec{q}_0$ is an element of a general star, we have to look for pairs $(\sigma,\sigma')\in P(\vec{q},\vec{q}')$ such that for the fixed $\vec{q}_0\in\Delta BZ$

$$\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')],$$
$$\sigma\in P{:}P^{\vec{q}}, \quad \sigma'\in P{:}P^{\vec{q}'} \tag{III.49}$$

is satisfied. We call this set, in analogy to the previous cases,

$$P(\vec{q},\vec{q}';\vec{q}_0) := \{(\sigma,\sigma'){:}\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')],(\sigma,\sigma')\in P(\vec{q},\vec{q}')\} \tag{III.50}$$

which immediately yields the multiplicity

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)} = n_\kappa n_{\kappa'}|P(\vec{q},\vec{q}';\vec{q}_0)|, \tag{III.51}$$

since (II.10) reduces to

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)} = n_\kappa n_{\kappa'} \sum_{(\sigma,\sigma')\in P(\vec{q},\vec{q}';\vec{q}_0)} \delta_{\vec{q}(\sigma) + \vec{q}'(\sigma'),\vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')]}. \tag{III.52}$$

If $\vec{q}_0$ belongs to a star of higher symmetry, i.e., if

$$\{e\}\subset P^{\vec{q}_0}\subseteq P, \tag{III.53}$$

then we are confronted with the most complicated case apart from the special case where

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} = \{e\} \tag{III.54}$$

which shall be discussed at the end of this section. Except for this special case we have to take into account in general

$$\{e\}\subset P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}\subseteq P^{\vec{q}_0}\subseteq P. \tag{III.55}$$

We start our investigations once again with a fixed pair $(\sigma,\sigma')\in P(\vec{q},\vec{q}')$ which satisfies

$$\vec{q}(\sigma) + \vec{q}'(\sigma') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')],$$
$$\sigma\in P{:}P^{\vec{q}}, \quad \sigma'\in P{:}P^{\vec{q}'}. \tag{III.56}$$

It is obvious that

$$\vec{q}(x\sigma) + \vec{q}'(x\sigma') = \vec{q}_0 + \vec{Q}\{\vec{q}_0(x)\} + D(x)\vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')] \quad \text{for all } x\in P^{\vec{q}_0} \tag{III.57}$$

must remain valid. For similar reasons as for (III.28), the set

$$\{(x\sigma,x\sigma'){:}x\in P^{\vec{q}_0}\}\not\subset P(\vec{q},\vec{q}') \tag{III.58}$$

cannot be in general a subset of $P(\vec{q},\vec{q}')$. In order to show this proposition we have to realize that

$$x\sigma = \sigma z, \quad \text{with } z\in P^{\vec{q}}, \text{ and } x\sigma' = \sigma' z', \quad \text{with } z'\in P^{\vec{q}'}$$
$$\text{for all } x\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}, \tag{III.59}$$

and that for any $x_1,x_2\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}$ with

$$x_1\neq x_2,$$
$$x_i\sigma = \theta z_i \text{ and } x_i\sigma' = \sigma' z_i' \Longleftrightarrow z_1\neq z_2(\in P^{\vec{q}})$$
$$\text{and } z_1'\neq z_2'(\in P^{\vec{q}'}) \tag{III.60}$$

is true. Since in general we have

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}\subset P^{\vec{q}_0}, \tag{III.61}$$

we introduce

$$v_j\in P^{\vec{q}_0}{:}P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} \tag{III.62}$$

left coset representatives of $P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}$ with respect to the $P^{\vec{q}_0}$ in order to be able to simplify (II.10). Furthermore we define by means of

$$\sigma_j := v_j\sigma z_j^{-1}\in P{:}P^{\vec{q}}, \quad z_j\in P^{\vec{q}}, \tag{III.63}$$
$$\sigma_k' = v_k\sigma' z_k'^{-1}\in P{:}P^{\vec{q}'}, \quad z_k'\in P^{\vec{q}'}, \tag{III.64}$$

the other left coset representatives which are connected (in the above-mentioned sense) by the elements of the group $P^{\vec{q}_0}$. Now it is easy to prove that just

$$\{(v_j\sigma z_j^{-1},v_j\sigma' z_j'^{-1}){:}v_j\in P^{\vec{q}_0}{:}P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}\}\subset P(\vec{q},\vec{q}') \tag{III.65}$$

is a subset of $P(\vec{q},\vec{q}')$ whose elements satisfy (III.57), since we have for any pair $v_j,v_k\in P^{\vec{q}_0}{:}P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}$ (because of

$$v_j^{-1}v_k\notin P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}, \quad \text{for all } v_j\neq v_k) \tag{III.66}$$

the following relations:

$$v_j\neq v_k\Leftrightarrow\sigma_j\neq\sigma_k \text{ and } \sigma_j'\neq\sigma_k', \tag{III.67}$$

$$v_j\neq v_k\Leftrightarrow\delta_{\vec{q}(v_j\sigma z_j^{-1}) + \vec{q}'(v_k\sigma' z_k'^{-1}),\vec{q}_0 + \vec{Q}[\vec{q}(\sigma_j) + \vec{q}'(\sigma_k')]} = 0 \tag{III.68}$$

If there exists a further pair $(\tau,\tau')\in P(\vec{q},\vec{q}')$ which is not contained in (III.65) such that

$$\vec{q}(\tau) + \vec{q}'(\tau') = \vec{q}_0 + \vec{Q}[\vec{q}(\tau) + \vec{q}'(\tau')], \tag{III.69}$$

then the same argumentation holds, i.e.,

$$\{(w_j\tau z_j^{-1},w_j\tau' z_j'^{-1}){:}w_j\in P^{\vec{q}_0}{:}P^{\vec{q},\vec{q}';\vec{q}_0}_{\tau,\tau'}\}\subset P(\vec{q},\vec{q}') \tag{III.70}$$

is a subset of $P(\vec{q},\vec{q}')$. By means of

$$\tau = t\sigma, \quad \tau' = t'\sigma' \quad \text{with } t,t'\in P \tag{III.71}$$

we obtain as a consequence of

$$\vec{q}'(t^{-1}t'\sigma') - \vec{q}'(\sigma') = \vec{q}_0(t^{-1}) - \vec{q}_0 + D(t^{-1})\vec{Q}[\vec{q}(\tau) + \vec{q}'(\tau')] - \vec{Q}[\vec{q}(\sigma) + \vec{q}'(\sigma')], \tag{III.72}$$

$$t\in P^{\vec{q}_0}\Leftrightarrow t^{-1}t'\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}, \tag{III.73}$$

$$t\notin P^{\vec{q}_0}\Leftrightarrow t^{-1}t'\notin P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}, \tag{III.74}$$

which proves, together with (III.67), that either the intersection of (III.65) with (III.70) is empty or both sets coincide. Therefore, we define by

$$P(\vec{q},\vec{q}',\vec{q}_0) = \{(\sigma_l,\sigma_l'){:}\vec{q}(\sigma_l) + \vec{q}(\sigma_l') = \vec{q}_0 + \vec{Q}[\vec{q}(\sigma_l) + \vec{q}'(\sigma_l')], (\sigma_l,\sigma_l')\in P(\vec{q},\vec{q}')\} \tag{III.75}$$

with the property

$$\{(v_j\sigma_l z_j^{-1},v_j\sigma_l' z_j'^{-1}){:}v_j\in P^{\vec{q}_0}{:}P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma_l,\sigma_l'}\}$$
$$\cap\{(w_k\sigma_m z_k^{-1},w_k\sigma_m' z_k'^{-1}){:}w_k\in P^{\vec{q}_0}{:}P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma_m,\sigma_m'}\}$$
$$= \emptyset \quad \text{for all } l\neq m \tag{III.76}$$

a set of pairs of coset representatives which generate pairwise empty intersections where however, the notation for the elements should not be confused with (III.63) and (III.64). Now we are in a position to study in more detail Eq. (II.10) for the most complicated case. Because of (III.67), (III.68), and (III.76), (II.10) reduces to

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = \sum_{(\sigma,\sigma')\in P(\vec{q},\vec{q}';\vec{q}_0)} \sum_{v_j\in P^{\vec{q}_0}{:}P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}} \frac{1}{|P^{\vec{q}_0}|} \sum_{x\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}} B^{\vec{q}_0 *}_{e,e}(x) \times B^{\vec{q}}_{\sigma_j,\sigma_j}(x)B^{\vec{q}'}_{\sigma_j',\sigma_j'}(x)$$

$$\times \mathbb{X}^{\kappa}((v_j\sigma z_j^{-1})^{-1}xv_j\sigma z_j^{-1})$$

$$\times\ \mathbb{X}^{\kappa'}((v_j\sigma' z_j'^{-1})^{-1}xv_j\sigma' z_j'^{-1})\mathbb{X}^{\kappa_0 *}(x), \qquad \text{(III.77)}$$

where we have already used the notation

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{v_j\sigma,v_j\sigma'} := v_j P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} v_j^{-1}, \quad v_j \in P^{\vec{q}_0}:P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}. \qquad \text{(III.78)}$$

The groups (III.78) are conjugate subgroups of $P^{\vec{q}_0}$, which are all isomorphic to $P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}$. For a fixed pair $(\sigma,\sigma')\in P(\vec{q},\vec{q}';\vec{q}_0)$, formula (III.77) contains $|P^{\vec{q}_0}:P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}|$ partial sums of the type

$$\sum_{y\in P^{\vec{q},\vec{q}';\vec{q}_0}_{v_j\sigma,v_j\sigma'}} B^{\vec{q}_0 *}_{e,e}(y) B^{\vec{q}}_{\sigma_j,\sigma_j}(y) B^{\vec{q}'}_{\sigma_j',\sigma_j'}(y)$$

$$\times \mathbb{X}^{\kappa}(\sigma_j^{-1}y\sigma_j)\mathbb{X}^{\kappa'}(\sigma_j'^{-1}y\sigma_j')\mathbb{X}^{\kappa_0 *}(y) =: S^{\vec{q},\vec{q}';\vec{q}_0}_{v_j\sigma,v_j\sigma'}. \qquad \text{(III.79)}$$

Replacing $v_j^{-1}yv_j = x$, which implies

$$y\in P^{\vec{q},\vec{q}';\vec{q}_0}_{v_j\sigma,v_j\sigma'} \Leftrightarrow v_j^{-1}yv_j \in P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}, \qquad \text{(III.80)}$$

and using formulas of the type

$$\mathbb{X}^{\kappa}(yxy^{-1}) = S^{\vec{q}_0 *}(x,y^{-1})S^{\vec{q}_0}(y^{-1},yxy^{-1})\mathbb{X}^{\kappa_0}(x), \quad \text{for all } x,y\in P^{\vec{q}_0}, \qquad \text{(III.81)}$$

which are a consequence of Eq. (3.10) of Ref. 8, we obtain, after a straightforward calculation,

$$S^{\vec{q},\vec{q}';\vec{q}_0}_{v_j\sigma,v_j\sigma'} = S^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} \quad \text{for all } v_j\in P^{\vec{q}_0}:P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}. \qquad \text{(III.82)}$$

Therefore, we arrive at the final formula

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = \sum_{(\sigma,\sigma')\in P(\vec{q},\vec{q}';\vec{q}_0)} \frac{1}{|P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}|} \sum_{x\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}} B^{\vec{q}_0 *}_{e,e}(x)$$

$$\times B^{\vec{q}}_{\sigma,\sigma}(x) B^{\vec{q}'}_{\sigma',\sigma'}(x)\mathbb{X}^{\kappa}(\sigma^{-1}x\sigma)$$

$$\times \mathbb{X}^{\kappa'}(\sigma'^{-1}x\sigma')\mathbb{X}^{\kappa_0}(x)^*, \qquad \text{(III.83)}$$

which has to be used when calculating multiplicities for space group representations. Of course cases A and B are contained in this formula. (For symmorphic space groups which are semidirect products of point groups with translation groups, the unimodular factors become one and the projective unirreps of $P^{\vec{q}}$ reduce to ordinary vector representations.) Finally we mention that Eq. (III.83) must be identical to formula (4.7.29) of Ref. 2. In our case we believe that the nonsymmorphicness of the space group is better expressed by emphasizing the explicity use of projective representations of the factor groups $P^{\vec{q}} \simeq G^{\vec{q}}/T$. In this connection we confess that the proof of (4.7.29) of Ref. 2 seems to be somewhat more elegant by using Mackey's theorem[2] than our approach to Eq. (III.83). However, we believe that our formula (III.83) contains in a more precise way information about how the summation about the coset representatives $(\sigma,\sigma')\in P(\vec{q},\vec{q}')$ has to be carried out in reality, in contrast to (4.7.29) of Ref. 2 where only a verbal hint is given. Finally we remark that for many examples $|P(\vec{q},\vec{q}';\vec{q}_0)| = 1$ is satisfied which makes the summation about several elements of $P(\vec{q},\vec{q}';\vec{q}_0)$ superfluous and simplifies (III.83).

We conclude this section by discussing special cases of (III.83), like (III.54), which has not been considered until now. As a first special case we consider

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} = \{e\}, \quad \text{for all } (\sigma,\sigma')\in P(\vec{q},\vec{q}';\vec{q}_0), \qquad \text{(III.84)}$$

which has as a consequence for the multiplicity

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = |P(\vec{q},\vec{q}';\vec{q}_0)|n_\kappa n_{\kappa'} n_{\kappa_0}. \qquad \text{(III.85)}$$

For the following special cases we assume

$$|P(\vec{q},\vec{q}';\vec{q}_0)| = 1, \qquad \text{(III.86)}$$

which is valid for a large number of cases. Consequently formula (III.83) reduces to

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = \frac{1}{|P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}|} \sum_{x\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}} B^{\vec{q}_0 *}_{e,e}(x)$$

$$\times B^{\vec{q}}_{\sigma,\sigma}(x) B^{\vec{q}'}_{\sigma',\sigma'}(x)\mathbb{X}^{\kappa}(\sigma^{-1}x\sigma)\mathbb{X}^{\kappa'}(\sigma'^{-1}x\sigma')\mathbb{X}^{\kappa_0 *}(x) \qquad \text{(III.87)}$$

If, furthermore,

$$\sigma = \sigma' = e \qquad \text{(III.88)}$$

is satisfied, (III.87) takes the simple form

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = \frac{1}{|P^{\vec{q},\vec{q}';\vec{q}_0}_{e,e}|} \sum_{x\in P^{\vec{q},\vec{q}';\vec{q}_0}_{e,e}} e^{i(\vec{q}_0-\vec{q}-\vec{q}')\cdot\tau(x)}\mathbb{X}^{\kappa}(x)$$

$$\times\ \mathbb{X}^{\kappa'}(x)\mathbb{X}^{\kappa_0 *}(x). \qquad \text{(III.89)}$$

However (III.88) does not represent the general situation, since at least one of the two left coset representatives σ,σ' is different from the identity element e, so that $\vec{q}_0$ belongs to ΔBZ. The appearance of the unimodular factors in (III.89) is due to the fact that we changed our factor system $R^{\vec{q}}$ trivially to $S^{\vec{q}}$. Furthermore (III.89) makes sense, if and only if [in analogy to (I.13) of Ref. 1]

$$R^{\vec{q}}(x,y)R^{\vec{q}'}(x,y) = R^{\vec{q}_0}(x,y), \quad \text{for all } x,y\in P^{\vec{q},\vec{q}';\vec{q}_0}_{e,e}, \qquad \text{(III.90)}$$

is valid. However this relation is, because $\vec{q} + \vec{q}' = \vec{q}_0 + \vec{Q}[\vec{q} + \vec{q}']$, automatically guaranteed. Finally we can imagine that further subductions, like $\mathbb{D}^{\kappa}\downarrow P^{\vec{q},\vec{q}';\vec{q}_0}_{e,e}$, $\mathbb{D}^{\kappa'}\downarrow P^{\vec{q},\vec{q}';\vec{q}_0}_{e,e}$, and $\mathbb{D}^{\kappa_0}\downarrow P^{\vec{q},\vec{q}';\vec{q}_0}_{e,e}$ yield to further simplifications of (III.89). However we do not want to discuss these simplifications in more detail, since the explicit calculation of (III.87) [respectively its special cases, like (III.88)] or, e.g.,

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{e,e} = P^{\vec{q}_0} = P^{\vec{q}} = P^{\vec{q}'} \qquad \text{(III.91)}$$

presupposes the explicit knowledge of the space group with its unirreps.

CONCLUDING REMARKS

The aim of this paper was to derive the multiplicity formula for space group representations by emphasizing projective representations for the little cogroups and to calculate for a series of cases the multiplicities without reference to a special space group. Thereby we have obtained for even the most complicated case relations concerning the wave vector selection rules which should be very useful when calculating space group CG coefficients. In this connection we have shown that just the little cogroup $P^{\vec{q}_0} \simeq G^{\vec{q}_0}/T$ plays an essential role.

[1]R. Dirl, "Clebsch–Gordan coefficients: General theory," J. Math. Phys. **20**, 659 (1979).
[2]C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[3]A.P. Cracknell and B.L. Davies, "Wave Vector Selection Rules for Space Groups" in *Lecture Notes in Physics*, edited by A. Janner, T. Boon, and M. Boon (Springer, Berlin, 1976), Vol. 50.
[4]A.P. Cracknell and B.L. Davies, "Computer Programs for Determining Wave Vector Selection Rules for Space Groups" in *Proceedings of the 6th Int. Colloquium on Group Theoretical Methods in Physics* (1977, in press).
[5]K. Olbrychski, R. Kolodziejski, M. Suffczynski, and H. Kunert, J. Physique **36**, 985 (1975).
[6]K. Olbrychski, R. Kolodziejski, M. Suffczynski, and H. Kunert, J. Physique **37**, 1483 (1976).
[7]J.L. Birman, "Theory of Crystal Space Groups and Infra-red Raman Processes of Insulated Crystals" in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1974).
[8]R. Dirl, J. Math. Phys. **18**, 2065 (1977).

Clebsch–Gordan coefficients for space groups

R. Dirl

Institut für Theoretische Physik, TU Wien, A-1040 Wien, Karlsplatz 13, Austria
(Received 28 December 1977)

A general method for finding Clebsch–Gordan coefficients is used to calculate them for nonsymmorphic space groups. This method is based on the fact that the columns of the Clebsch–Gordan matrices can be seen as G-adapted vectors and that the multiplicity index can be traced back to special column indices of the Kronecker product. Using this method we obtain simple defining equations for the multiplicity index and for nearly all cases without reference to a special space group by a simple calculation the corresponding Clebsch–Gordan matrices.

INTRODUCTION

In this article we shall report on our investigations of how the proposed method[1] works for the determination of space group CG coefficients. The aim of this paper is to show that the proposed method is indeed a useful and practical one in order to determine quite generally CG coefficients for nonsymmorphic space groups. In this connection we remark that the derived formulas contain the special case of symmorphic space groups in a consistent way. The present method utilizes the fact that the columns of the CG matrices can be seen as G adapted vectors and that the multiplicity index can be traced back to special column indices of the Kronecker product. As in Refs. 2 and 3 we emphasize the use of projective representations for the little cogroups $P^{\vec{q}} \simeq G^{\vec{q}}/T$. However it should be noted that our approach is different to those given until now.[2–11]

The paper is organized as follows: In Sec. I we rewrite our general formulas of Ref. 1 into nonsymmorphic space group representations. As in the previous paper[12] we apply our formulas to various cases. In Sec. II. A we discuss the case where $\vec{q}$ and $\vec{q}'$ are assumed to belong to general stars and where $\vec{q}_0$ is either an element of a general star or of a star of higher symmetry. The case where $\vec{q}$ is assumed to be an element of a general star and $\vec{q}'$ of a star of higher symmetry is investigated in Sec. III. B. Thereby we have to consider once again the two possibilities, where $\vec{q}_0$ belongs either to a general star, or to a star of higher symmetry. As in the previous cases we derive very simple formulas for the multiplicity index and obtain immediately the corresponding CG coefficients. The most complicated case, where $\vec{q}$ and $\vec{q}'$, belong to stars of higher symmetry, is considered in Sec. III. C. In analogy to the foregoing sections we have to distinguish the cases where $\vec{q}_0$ is either an element of a general star or a star of higher symmetry. The latter case is divided into further subcases. Even for the most complicated case we derive simple formulas which have to be inspected when determining the multiplicity index. The CG coefficients are obtained by simple calculations for nearly all cases without reference to a special space group.

I. CLEBSCH–GORDAN COEFFICIENTS FOR NONSYMMORPHIC SPACE GROUPS

In this section we specify our general formulas (II. 24)–(II. 26), (II. 36), and (II. 39) of Ref. 1 by means of the formulas (I. 11) and (I. 12) of Ref. 12 to ordinary vector representations of nonsymmorphic space groups. The special case of symmorphic space groups is, of course, included in the general formulas.

In order to be able to transfer the abovementioned formulas of Ref. 1 to space group representations, the equivalence classes, row and column indices of the corresponding unirreps have to be replaced in the following way:

$$\begin{aligned}
&\alpha \to (\kappa,\vec{q}) := (\kappa,\vec{q})\uparrow G;\quad \vec{q}\in\Delta BZ,\ \kappa\in A_{P^{\vec{q}}(S^{\vec{q}})},\\
&\beta \to (\kappa',\vec{q}') := (\kappa',\vec{q}')\uparrow G;\quad \vec{q}'\in\Delta BZ,\ \kappa'\in A_{P^{\vec{q}'}(S^{\vec{q}'})},\\
&\gamma \to (\kappa_0,\vec{q}_0) := (\kappa_0,\vec{q}_0)\uparrow G;\quad \vec{q}_0\in\Delta BZ,\ \kappa_0\in A_{P^{\vec{q}_0}(S^{\vec{q}_0})},\\
&q \to \underline{\sigma}, c;\quad p \to \underline{\tau}, d;\quad \underline{\sigma},\underline{\tau}\in P:P^{\vec{q}},\ d,c=1,2,\ldots,n_\kappa,\\
&s \to \underline{\sigma}', c';\quad r \to \underline{\tau}', d';\quad \underline{\sigma}',\underline{\tau}'\in P:P^{\vec{q}'},\ d',c'=1,2,\ldots,n_{\kappa'},\\
&a \to \underline{e}, a\quad \text{(fixed)}.
\end{aligned} \tag{I. 1}$$

Thereby it should be noted that the index a has already been replaced by an appropriately chosen pair of indices of the corresponding space group unirrep, namely $\underline{e}, a$ with $\underline{e}\in P:P^{\vec{q}_0}$ and a fixed.

Presupposing that the corresponding multiplicity $m_{(\kappa,\vec{q})(\kappa,\vec{q}');(\kappa_0,\vec{q}_0)}$ is different from zero, we construct, according to the proposed method vectors of the type (II. 23) of Ref. 1,

$$\left\{\vec{B}_{\underline{e},a}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma},c;\underline{\sigma}',c')} : (\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}'),\ c=1,2,\ldots,n_\kappa,\ c'=1,2,\ldots,n_{\kappa'}.\right\} \tag{I. 2}$$

Of course, taking the wave vector selection rules (II. 5) of Ref. 12 into account, the order of the set (I. 2) becomes essentially smaller than $|P:P^{\vec{q}}|n_\kappa|P:P^{\vec{q}'}|n_{\kappa'}$, which implies an important simplification for the determination of the multiplicity index and therefore of the corresponding CG coefficients. The components of the vectors (I. 2) turn out to be

$$\begin{aligned}
&B_{\underline{\tau}d,\underline{\tau}'d',(\kappa_0,\vec{q}_0)(\underline{\sigma},c,\underline{\sigma}',c')\underline{e}a}^{(\kappa,\vec{q})(\kappa',\vec{q}')}\\
&\quad=\left\{\vec{B}_{\underline{e}a}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma},c;\underline{\sigma}',c')}\right\}_{\underline{\tau}d,\underline{\tau}'d'}\\
&\quad=\delta_{\vec{q}(\underline{\tau})+\vec{q}'(\underline{\tau}'),\vec{q}_0+\vec{Q}[\vec{q}(\underline{\tau})+\vec{q}'(\underline{\tau}')]}\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\\
&\qquad\times\sum_{\beta\in P^{\vec{q}_0}}\Delta^{\vec{q}_0}(\underline{e},\beta\underline{e})\Delta^{\vec{q}}(\underline{\tau},\beta\underline{\sigma})\Delta^{\vec{q}'}(\underline{\tau}',\beta\underline{\sigma}')
\end{aligned}$$

 0022-2488/79/040671-08$01.00

$$\times B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\beta) B'_{\underline{\tau},\underline{\sigma}}(\beta) B^{\vec{q}}_{\tau',\sigma'}(\beta) \mathbb{R}^{\kappa}_{dc}(\underline{\tau}^{-1}\beta\underline{\sigma})$$

$$\times \mathbb{R}^{\kappa'}_{d'c'}(\underline{\tau}'^{-1}\beta\underline{\sigma}') \mathbb{R}^{\kappa_0 *}_{aa}(\beta),$$

$$\tau \in P:P^{\vec{q}},\ d=1,2,\ldots n_\kappa,\ \underline{\tau}' \in P:P',\ d'=1,2,\ldots n_{\kappa'}, \tag{I.3}$$

where (I.11) and (I.12) of Ref. 12 is used, where the summation about all elements of the translation group T has been carried out, and where

$$n_\gamma = n_{(\kappa_0,\vec{q}_0)\uparrow G} = |P:P^{\vec{q}_0}| n_{\kappa_0} \tag{I.4}$$

was inserted into the general formula (II.24) of Ref. 1. In order to prove the following equations we have to use the orthogonality relations for the projective unirreps $\mathbb{R}^{\kappa_0}$ of $P^{\vec{q}_0}$. This explains why we have generalized our method to projective representations.

As a first step of the procedure we have to look for vectors (I.2) whose norm exists,

$$\|\vec{\mathbb{B}}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma},c,\underline{\sigma}',c')}_{\underline{e}a}\|^2$$

$$\delta_{\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')\vec{q}_0+\mathfrak{Q}[\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')]}\ \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}$$

$$\times \sum_{\beta\in P^{\vec{q}_0}} \Delta^{\vec{q}_0}(\underline{e},\beta\underline{e}) \Delta^{\vec{q}}(\sigma,\beta\underline{\sigma}) \Delta^{\vec{q}'}(\underline{\sigma},\beta\underline{\sigma}')$$

$$\times B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\beta) B^{\vec{q}}_{\sigma,\sigma}(\beta)\, B'_{\underline{\sigma}',\underline{\sigma}'}(\beta) \mathbb{R}^{\kappa}_{cc}(\underline{\sigma}^{-1}\beta\underline{\sigma}) \mathbb{R}^{\kappa'}_{c'c'}(\underline{\sigma}'^{-1}\beta\underline{\sigma}') \mathbb{R}^{\kappa_0 *}_{aa}(\beta). \tag{I.5}$$

As a second step we have to investigate the scalar product of any two such vectors whose norm exists. Thereby we introduce the notation

$$\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\tau},d & \underline{\tau}',d'\end{matrix}\middle|\begin{matrix}(\kappa_0,\vec{q}_0) & (\underline{\sigma},c;\underline{\sigma}',c') \\ \underline{e}a & \end{matrix}\right\}:$$

$$= \langle \vec{\mathbb{B}}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\tau},d;\underline{\tau}',d')}_{\underline{e}a},\ \vec{\mathbb{B}}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma},c;\underline{\sigma}',c')}_{\underline{e}a}\rangle$$

$$= \delta_{\vec{q}(\underline{\tau})+\vec{q}'(\underline{\tau}'),\vec{q}_0+\mathfrak{Q}[\vec{q}(\underline{\tau})+\vec{q}'(\underline{\tau}')]}\ \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}$$

$$\times \sum_{\beta\in P^{\vec{q}_0}} \Delta^{\vec{q}_0}(\underline{e},\beta\underline{e}) \Delta^{\vec{q}}(\underline{\tau},\beta\underline{\sigma}) \Delta^{\vec{q}'}(\underline{\tau}',\beta\underline{\sigma}')$$

$$\times B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\beta) B^{\vec{q}}_{\underline{\tau},\underline{\sigma}}(\beta) B^{\vec{q}'}_{\underline{\tau}',\underline{\sigma}'}(\beta) \mathbb{R}^{\kappa}_{dc}(\underline{\tau}^{-1}\beta\underline{\sigma}) \mathbb{R}^{\kappa'}_{d'c'}(\underline{\tau}'^{-1}\beta\underline{\sigma}') \mathbb{R}^{\kappa_0 *}_{aa}(\beta). \tag{I.6}$$

In case we can find, with the aid of (I.6), just $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ orthogonal vectors (I.2), the corresponding column indices of the space group representation $D^{(\kappa,\vec{q})\uparrow G} \otimes D^{(\kappa',\vec{q}')\uparrow G}$ can be chosen as a multiplicity index w,

$$w = (\underline{\sigma}_v, c_v; \underline{\sigma}'_v, c'_v);\quad v=1,2,\ldots m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}. \tag{I.7}$$

In this connection we remark that for the general case more than one pair of left coset representatives $(\underline{\sigma}_v, \underline{\sigma}'_v) \in P(\vec{q},\vec{q}')$ is necessary in case we try to identify the multiplicity index w with column indices of the considered Kronecker product. Thereby we must expect that at least one of the two coset representatives must be different from the identity element e, otherwise $\vec{q}_0$ would not belong to ΔBZ. Presupposing that we have found $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ orthogonal vectors which are labeled by the indices (I.7), we immediately obtain by means of of (II.36) of Ref. 1 the corresponding CG coefficients,

$$C^{(\kappa,\vec{q})(\kappa',\vec{q}')}_{\underline{\tau}d,\underline{\tau}'d';(\kappa_0,\vec{q}_0)w=(\underline{\sigma}_v,c_v;\underline{\sigma}'_v,c'_v)\underline{e}a}$$

$$=:\begin{pmatrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') & (\kappa_0,\vec{q}_0) & w=(\underline{\sigma}_v,c_v;\underline{\sigma}'_v,c'_v) \\ \underline{\tau},d & \underline{\tau}',d' & \underline{e}a & \end{pmatrix}$$

$$= \|\vec{\mathbb{B}}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma}_v,c_v;\underline{\sigma}'_v,c'_v)}_{\underline{e}a}\|^{-1}$$

$$\times \left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\tau},d & \underline{\tau}',d'\end{matrix}\middle|\begin{matrix}(\kappa_0,\vec{q}_0 & (\underline{\sigma}_v,c_v;\underline{\sigma}'_v,c'_v) \\ \underline{e}a & \end{matrix}\right\},$$

$$\underline{\tau}\in P:P^{\vec{q}},\quad d=1,2,\ldots,n_\kappa,\quad \underline{\tau}'\in P:P',\quad d'=1,2,\ldots,n_{\kappa'}. \tag{I.8}$$

In order to obtain the remaining columns of the CG matrices we have to specialize (II.39) of Ref. 1 to space group representations. The corresponding formula reads

$$C^{(\kappa,\vec{q})(\kappa',\vec{q}')}_{\underline{\tau}d,\underline{\tau}'d';(\kappa_0,\vec{q}_0)w=(\underline{\sigma}_v,c_v;\underline{\sigma}'_v,c'_v)\underline{\sigma}_0 j}$$

$$=:\begin{pmatrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') & (\kappa_0,\vec{q}_0) & w=(\underline{\sigma}_v,c_v;\underline{\sigma}'_v,c'_v) \\ \underline{\tau},d & \underline{\tau}',d' & \underline{\sigma}_0,j & \end{pmatrix}$$

$$= \|\vec{\mathbb{B}}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma}_v,c_v;\underline{\sigma}'_v,c'_v)}_{\underline{e}a}\|^{-1}$$

$$\times \delta_{\vec{q}(\underline{\tau})+\vec{q}'(\underline{\tau}'),\vec{q}_0(\underline{\sigma}_0)+\mathfrak{Q}[\vec{q}(\underline{\tau})+\vec{q}'(\underline{\tau}')]}\ \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}$$

$$\times \sum_{\beta\in P} \Delta^{\vec{q}_0}(\underline{\sigma}_0,\beta\underline{e}) \Delta^{\vec{q}}(\underline{\tau},\beta\underline{\sigma}_v) \Delta^{\vec{q}'}(\underline{\tau}',\beta\underline{\sigma}'_v)$$

$$\times B^{\vec{q}_0 *}_{\underline{\sigma}_0,\underline{e}}(\beta) B^{\vec{q}}_{\underline{\tau},\underline{\sigma}_v}(\beta) B^{\vec{q}'}_{\underline{\tau}',\underline{\sigma}'_v}(\beta) \mathbb{R}^{\kappa}_{dc_v}(\underline{\tau}^{-1}\beta\underline{\sigma}_v)$$

$$\times \mathbb{R}^{\kappa'}_{d'c'_v}(\underline{\tau}'^{-1}\beta\underline{\sigma}'_v) \mathbb{R}^{\kappa_0 *}_{ja}(\underline{\sigma}_0^{-1}\beta),$$

$$v=1,2,\ldots,m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)},$$

$$\underline{\sigma}_0 \in P:P^{\vec{q}_0},\quad j=1,2,\ldots,n_{\kappa_0}, \tag{I.9}$$

where the summation about all elements of the translation group has been already carried out.

If, however [by analogy to (II.33) of Ref. 1], some of the $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ orthonormalized vectors must be determined by Schmidt's procedure, the remaining columns of the CG matrix nevertheless have to be calculated by the same formula [compare (II.17) of Ref. 1], otherwise the corresponding equation to (II.42) of Ref. 1 cannot be satisfied. We shall show in the following sections that for nearly all cases the space group CG coefficients can be determined quite generally without reference to a special space group. Nevertheless we realize now the reason why the method given in Ref. 1 has been generalized to projective representations.

II. DISCUSSION OF VARIOUS CASES

As in the previous paper[12] we start from simple cases and proceed to more complicated ones. In order to obtain a better view we shall subdivide our cases A, B, and C of Sec. III of Ref. 12 into further parts.

A. 1. $\vec{q},\vec{q}'$ = general stars, $\vec{q}_0$ = general star

According to (III.4) and (III.5) of Ref. 12 we have

$$\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')=\vec{q}_0+\vec{Q}[\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')]\quad (\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0). \tag{II.1}$$

Specializing (I.3) to this simple case we omit for the sake of simplicity the row and column index 1 of the projective unirrep of $P^{\vec{q}}=P^{\vec{q}'}=P^{\vec{q}_0}=\{e\}$ and denote the identity representation of these groups by 0,

$$\|\vec{B}_{\underline{e}}^{(0,\vec{q})(0,\vec{q}');(0,\vec{q}_0)(\underline{\sigma},\underline{\sigma}')}\|^2=\delta_{\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}'),\vec{q}_0+\vec{Q}[\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')]}=1$$
$$\text{for all } (\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0). \tag{II.2}$$

The scalar product (I.6) yields for any pair of $P(\vec{q},\vec{q}';\vec{q}_0)$

$$\left\{\begin{matrix}(0,\vec{q}) & (0,\vec{q}') & (0,\vec{q}_0) & (\underline{\sigma}_j,\underline{\sigma}_j')\\ \underline{\sigma}_i & \underline{\sigma}_i' & \underline{e} & \end{matrix}\right\}=\delta_{ij} \tag{II.3}$$

which implies that the multiplicity index w can be chosen as

$$w=(\underline{\sigma}_j,\underline{\sigma}_j'),\quad (\underline{\sigma}_j,\underline{\sigma}_j')\in P(\vec{q},\vec{q}';\vec{q}_0). \tag{II.4}$$

Because of (I.8), we obtain as CG coefficients

$$\left(\begin{matrix}(0,\vec{q}) & (0,\vec{q}') & (0,\vec{q}_0) & w=(\underline{\sigma}_j,\underline{\sigma}_j')\\ \underline{\tau} & \underline{\tau}' & \underline{e} & \end{matrix}\right)=\delta_{\underline{\tau},\underline{\sigma}_j}\delta_{\underline{\tau}',\underline{\sigma}_j'},\quad \underline{\tau},\underline{\tau}'\in P. \tag{II.5}$$

The remaining CG coefficients are immediately calculable by means of (I.9),

$$\left(\begin{matrix}(0,\vec{q}) & (0,\vec{q}') & (0,\vec{q}_0) & w=(\underline{\sigma}_j,\underline{\sigma}_j')\\ \underline{\tau} & \underline{\tau}' & \underline{\sigma}_0 & \end{matrix}\right)=\delta_{\underline{\tau},\underline{\sigma}_0\underline{\sigma}_j}\delta_{\underline{\tau}',\underline{\sigma}_0\underline{\sigma}_j'},\quad \underline{\sigma}_0\in P.$$
$$\times B^{\vec{q}}_{\underline{\sigma}_0\underline{\sigma}_j,\underline{\sigma}_j}(\underline{\sigma}_0)\,B^{\vec{q}'}_{\underline{\sigma}_0\underline{\sigma}_j',\underline{\sigma}_j'}(\underline{\sigma}_0),\quad \underline{\sigma}_0\in P. \tag{II.6}$$

A. 2. $\vec{q},\vec{q}'$ = general stars, $\vec{q}_0$ = star of higher symmetry

In this case we have to take into account (III.17)–(III.19) of Ref. 12 when investigating (I.3). As in the previous case we denote the identity representation of $P^{\vec{q}}=P^{\vec{q}'}=\{e\}$ by 0 and suppress the superfluous row and column index 1 of this representation. According to the multiplicity formula (III.19) of Ref. 12 we have to look for $n_{\kappa_0}|P(\vec{q},\vec{q}';\vec{q}_0)|$ orthogonal vectors (I.3). For this purpose it suffices, because of the wave vector selection rules, to consider the norm of the following vectors:

$$\{\vec{B}_{\underline{e}a}^{(0,\vec{q})(0,\vec{q}');(\kappa_0,\vec{q}_0)(y\underline{\sigma}_j;y\underline{\sigma}_j')}:\ y\in P^{\vec{q}_0};\ (\underline{\sigma}_j,\underline{\sigma}_j')\in P(\vec{q},\vec{q}';\vec{q}_0)\}, \tag{II.7}$$

$$\|\vec{B}_{\underline{e}a}^{(0,\vec{q})(0,\vec{q}');(\kappa_0,\vec{q}_0)(y\underline{\sigma}_j;y\underline{\sigma}_j')}\|^2$$
$$=\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\delta_{\vec{q}(y\underline{\sigma}_j)+\vec{q}'(y\underline{\sigma}_j'),\vec{q}_0+\vec{Q}[\vec{q}(y\underline{\sigma}_j)+\vec{q}'(y\underline{\sigma}_j')]}=\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|},$$
$$\text{for all } y\in P^{\vec{q}_0} \text{ and } (\underline{\sigma}_j,\underline{\sigma}_j')\in P(\vec{q},\vec{q}';\vec{q}_0). \tag{II.8}$$

Consequently we have found $|P^{\vec{q}_0}|\cdot|P(\vec{q},\vec{q}';\vec{q}_0)|$ vectors whose norm exists and is independent of the special index $(y\underline{\sigma}_j,y\underline{\sigma}_j')$. Now we have to solve the problem whether we can find just $n_{\kappa_0}|P(\vec{q},\vec{q}';\vec{q}_0)|$ orthogonal vectors by means of (I.6). Considering the scalar product for two vectors of (II.7), we obtain by using (III.18) of Ref. 12

$$\left\{\begin{matrix}(0,\vec{q}) & (0,\vec{q}') & (\kappa_0,\vec{q}_0) & (y\underline{\sigma}_j;y\underline{\sigma}_j')\\ x\underline{\sigma}_i & x\underline{\sigma}_i' & \underline{e}a & \end{matrix}\right\}\text{for all } x,y\in P^{\vec{q}_0},$$
$$=\delta_{ij}\left\{\begin{matrix}(0,\vec{q}) & (0,\vec{q}') & (\kappa_0,\vec{q}_0) & (y\underline{\sigma}_j;y\underline{\sigma}_j')\\ x\underline{\sigma}_j & x\underline{\sigma}_j' & \underline{e}a & \end{matrix}\right\}, \tag{II.9}$$

$$\left\{\begin{matrix}(0,\vec{q}) & (0,\vec{q}') & (\kappa_0,\vec{q}_0) & (y\underline{\sigma};y\underline{\sigma}')\\ x\underline{\sigma} & x\underline{\sigma}' & \underline{e}a & \end{matrix}\right\}$$
$$=\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(xy^{-1})B^{\vec{q}}_{x\underline{\sigma},y\underline{\sigma}}(xy^{-1})B^{\vec{q}'}_{x\underline{\sigma}',y\underline{\sigma}'}(xy^{-1})\mathbb{R}^{\kappa_0 *}_{aa}(xy^{-1})$$
$$\text{for all } x,y\in P^{\vec{q}_0} \text{ and } (\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0). \tag{II.10}$$

Thereby it should be noted that we have omitted the index j of the elements of $P(\vec{q},\vec{q}';\vec{q}_0)$ and that (II.10) reduces to (II.8) if $y=x$. Now it is obvious how we have to proceed. Namely, in case we can find n_{κ_0} different group elements $x_i\in P^{\vec{q}_0}$, $i=1,2,\dots,n_{\kappa_0}$ such that

$$\mathbb{R}^{\kappa_0 *}_{aa}(x_ix_j^{-1})=\delta_{ij}\quad\text{for all } i,j=1,2,\dots,n_{\kappa_0}, \tag{II.11}$$

we have solved the multiplicity problem for this case. We can choose

$$w=(x_i\underline{\sigma};x_i\underline{\sigma}')\quad\text{for all } (\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0) \text{ and}$$
$$i=1,2,\dots,n_{\kappa_0} \tag{II.12}$$

as a multiplicity index w, because (II.9) and (II.11) guarantee the orthogonality of the corresponding vectors,

$$\left\{\begin{matrix}(0,\vec{q}) & (0,\vec{q}') & (\kappa_0,\vec{q}_0) & (x_i\underline{\sigma}_k;x_i\underline{\sigma}_k')\\ x_l\underline{\sigma}_j & x_l\underline{\sigma}_j' & \underline{e}a & \end{matrix}\right\}=\delta_{jk}\delta_{il}. \tag{II.13}$$

Of course (II.11) is the key equation when fixing the multiplicity index w. It can be easily proven that we can find at most n_{κ_0} group elements $x_i\in P^{\vec{q}_0}$ such that (II.11) is satisfied. For this purpose we rewrite (II.11) as

$$\mathbb{R}^{\kappa_0 *}_{aa}(x_ix_j^{-1})=S^{\vec{q}_0}(x_i,x_j^{-1})S^{\vec{q}_0 *}(x_j,x_j^{-1})\langle\vec{\mathbb{R}}^{\kappa_0}_a(x_i),\vec{\mathbb{R}}^{\kappa_0}_a(x_j)\rangle, \tag{II.14}$$

$$\langle\vec{\mathbb{R}}^{\kappa_0}_a(x_i),\vec{\mathbb{R}}^{\kappa_0}_a(x_j)\rangle:=\sum_{b=1}^{n_{\kappa_0}}\mathbb{R}^{\kappa_0 *}_{ab}(x_i)\mathbb{R}^{\kappa_0}_{ab}(x_j), \tag{II.15}$$

which implies that we consider the ath row of the n_{κ_0}-dimensional projective unirrep $\mathbb{R}^{\kappa_0}$

$$\{\vec{\mathbb{R}}^{\kappa_0}_a(x):\ x\in P^{\vec{q}_0}\} \tag{II.16}$$

as (normalized) elements of an n-dimensional Euclidean space. This proves our assertion that just n_{κ_0} linear independent vectors can be found. Hence

$$\mathbb{R}^{\kappa_0 *}_{aa}(x_ix_j^{-1})=\delta_{ij}\Longleftrightarrow\langle\vec{\mathbb{R}}^{\kappa_0}_a(x_i),\vec{\mathbb{R}}^{\kappa_0}_a(x_j)\rangle=\delta_{ij}. \tag{II.17}$$

In this connection we remark that a suitable choice of the row index a of $\mathbb{R}^{\kappa_0}$ can much simplify the problem of finding n_{κ_0} group elements $x_i\in P^{\vec{q}_0}$ which satisfy (II.11).

Presupposing that we have found n_{κ_0} group elements $x_i \in P^{\vec{q}_0}$ satisfying (II. 11), the CG coefficients can be calculated immediately by means of (I. 8) and (II. 8),

$$\begin{pmatrix} (0,\vec{q}) & (0,\vec{q}') & (\kappa_0,\vec{q}_0) & w=(x_i\underline{\sigma};x_i\underline{\sigma}') \\ \underline{\tau} & \underline{\tau}' & \underline{e}a & \end{pmatrix}$$

$$= \left(\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\right)^{1/2} \Delta^{\vec{q}_0}(\underline{e}, \underline{\tau\sigma^{-1}x_i^{-1}e})\delta_{\underline{\tau}', \underline{\tau\sigma^{-1}\sigma}'}$$

$$\times B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\underline{\tau\sigma^{-1}x_i^{-1}})B^{\vec{q}}_{\underline{\tau}, x_i\underline{\sigma}}(\tau\sigma^{-1}\ x_i^{-1})B^{\vec{q}'}_{\underline{\tau}', x_i\underline{\sigma}'}(\underline{\tau\sigma^{-1}x_i^{-1}})$$

$$\mathbb{R}^{\kappa_0 *}_{aa}(\underline{\tau\sigma^{-1}x_i^{-1}}), (\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0). \quad \text{(II. 18)}$$

A straightforward calculation by directly using (II. 18) yields the orthonormality condition for CG coefficients,

$$\sum_{\underline{\tau},\underline{\tau}'}\begin{pmatrix} (0,\vec{q}) & (0,\vec{q}') & (\kappa_0,\vec{q}_0) & w=(x_i\underline{\sigma};x_i\underline{\sigma}') \\ \underline{\tau} & \underline{\tau}' & \underline{e}a & \end{pmatrix}^*$$

$$\times\begin{pmatrix} (0,\vec{q}) & (0,\vec{q}') & (\kappa_0,\vec{q}_0) & w=(x_j\underline{\sigma};x_j\underline{\sigma}') \\ \underline{\tau} & \underline{\tau}' & \underline{e}a & \end{pmatrix} = \delta_{ij}. \quad \text{(II. 19)}$$

[Orthogonality with respect to different pairs $(\sigma,\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0)$ guarantees Eq. (II. 13).] The remaining CG coefficients must be calculated by means of (I. 9),

$$\begin{pmatrix} (0,\vec{q}) & (0,\vec{q}') & (\kappa_0,\vec{q}_0) & w=(x_i\underline{\sigma};x_i\underline{\sigma}') \\ \underline{\tau} & \underline{\tau}' & \underline{\sigma}_0,j & \end{pmatrix}$$

$$= \left(\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\right)^{1/2} \Delta^{\vec{q}_0}(\underline{\sigma}_0, \underline{\tau\sigma^{-1}x_i^{-1}e})\delta_{\underline{\tau}', \underline{\tau\sigma^{-1}\sigma}'}$$

$$\times B^{\vec{q}_0 *}_{\underline{\sigma}_0,\underline{e}}(\underline{\tau\sigma^{-1}x_i^{-1}})B^{\vec{q}}_{\underline{\tau}, x_i\underline{\sigma}}(\underline{\tau\sigma^{-1}x_i^{-1}})B^{\vec{q}'}_{\underline{\tau}', x_i\underline{\sigma}'}(\underline{\tau\sigma^{-1}x_i^{-1}})$$

$$\times \mathbb{R}^{\kappa_0 *}_{ja}(\underline{\sigma}_0^{-1}\underline{\tau\sigma^{-1}x_i^{-1}}),$$

$$\underline{\sigma}_0 \in P: P^{\vec{q}_0} \quad j=1,2,\ldots,n_{\kappa_0} \quad \text{(II. 20)}$$

B. 1. $\vec{q}$ = general star, $\vec{q}'$ = star of higher symmetry, $\vec{q}_0$ = general star

Due to (III. 24) of Ref. 12 we must find $n_{\kappa'}|P(\vec{q},\vec{q}';\vec{q}_0)|$ orthogonal vectors. For this purpose it suffices to consider the norm of the following vectors:

$$\{\vec{B}_{\underline{e}}^{(0,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)(\underline{\sigma};\underline{\sigma}',c')}; (\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0);$$

$$c'=1,2,\ldots,n_{\kappa'}\}, \quad \text{(II. 21)}$$

$$\|\vec{B}_{\underline{e}}^{(0,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)(\underline{\sigma};\underline{\sigma}',c')}\|^2 = \delta_{\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}'),\vec{q}_0+\mathfrak{Q}[\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')]} = 1,$$

for all $(\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0)$ and $c'=1,2,\ldots n_{\kappa'}$, (II. 22)

where we have already used the symbol 0 for the identity representation of $P^{\vec{q}} = P^{\vec{q}_0} = \{e\}$ and omitted superfluous indices. We obtain for the scalar product (I. 6) of elements of (II. 21) the following expressions:

$$\left\{\begin{matrix} (0,\vec{q}) & (\kappa',\vec{q}') & (0,\vec{q}_0) & (\underline{\sigma}_j;\underline{\sigma}_j',c') \\ \underline{\sigma}_i & \underline{\sigma}_i',d' & \underline{e} & \end{matrix}\right\}$$

$$= \delta_{ij}\left\{\begin{matrix} (0,\vec{q}) & (\kappa',\vec{q}') & (0,\vec{q}_0) & (\underline{\sigma}_j;\underline{\sigma}_j',c') \\ \underline{\sigma}_j & \underline{\sigma}_j',d' & \underline{e} & \end{matrix}\right\}$$

for all $(\underline{\sigma}_j,\underline{\sigma}_j')\in P(\vec{q},\vec{q}';\vec{q}_0)$, $d',c'=1,2,\ldots,n_{\kappa'}$, (II. 23)

$$\left\{\begin{matrix} (0,\vec{q}) & (\kappa',\vec{q}') & (0,\vec{q}_0) & (\underline{\sigma};\underline{\sigma}',c') \\ \underline{\sigma} & \underline{\sigma}',d' & \underline{e} & \end{matrix}\right\} = \mathbb{R}^{\kappa'}_{d'c'}(e) = \delta_{d'c'}. \quad \text{(II. 24)}$$

Hence we have to choose

$w=(\underline{\sigma};\underline{\sigma}',c')$ for all $(\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0)$ and $c'=1,2,\ldots,n_{\kappa'}$ (II. 25)

as a multiplicity index, because (II. 23) and (II. 24) guarantee the orthogonality of the corresponding vectors. Therefore, the corresponding CG coefficients turn out to be

$$\begin{pmatrix} (0,\vec{q}) & (\kappa',\vec{q}') & (0,\vec{q}_0) & w=(\underline{\sigma};\underline{\sigma}',c') \\ \underline{\tau} & \underline{\tau}',d' & \underline{e} & \end{pmatrix} = \delta_{\underline{\tau},\underline{\sigma}}\delta_{\underline{\tau}',\underline{\sigma}'}\delta_{d'c'},$$

for all $(\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0)$ and $c'=1,2,\ldots,n_{\kappa'}$. (II. 26)

The remaining CG coefficients follow immediately with the aid of (I. 9),

$$\begin{pmatrix} (0,\vec{q}) & (\kappa',\vec{q}') & (0,\vec{q}_0) & w=(\underline{\sigma};\underline{\sigma}',c') \\ \underline{\tau} & \underline{\tau}',d' & \underline{\sigma}_0 & \end{pmatrix}$$

$$=\delta_{\underline{\tau},\underline{\sigma}_0\underline{\sigma}}\Delta^{\vec{q}'}(\underline{\tau}',\underline{\sigma}_0\underline{\sigma}')B^{\vec{q}_0 *}_{\underline{\sigma}_0,\underline{e}}(\underline{\sigma}_0)B^{\vec{q}}_{\underline{\tau},\underline{\sigma}}(\underline{\sigma}_0)B^{\vec{q}'}_{\underline{\tau}',\underline{\sigma}'}(\underline{\sigma}_0)$$

$$\times\mathbb{R}^{\kappa'}_{d'c'}(\underline{\tau}'^{-1}\underline{\sigma}_0\underline{\sigma}'), \quad \underline{\sigma}_0\in P. \quad \text{(II. 27)}$$

B. 2. $\vec{q}$ = general star, $\vec{q}'$ = star of higher symmetry, $\vec{q}_0$ = star of higher symmetry

According to (III. 43)—(III. 45) of Ref. 12, we consider, for the same reasons as before, the following set of vectors,

$$\{\vec{B}_{\underline{e}a}^{(0,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(x\underline{\sigma};x\underline{\sigma}'x'^{-1},c')}:(\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0);\ x\in P^{\vec{q}_0};$$

$$c'=1,2,\ldots,n_{\kappa'}\} \quad \text{(II. 28)}$$

where we have already used (III. 29) of Ref. 12, i.e., $x'\in P^{\vec{q}'}$. The norm of all $|P(\vec{q},\vec{q}';\vec{q}_0)|n_{\kappa'}|P^{\vec{q}_0}|$ elements of (II. 28) takes the same value,

$$\|\vec{B}_{\underline{e}a}^{(0,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(x\underline{\sigma};x\underline{\sigma}'x'^{-1},c')}\|^2$$

$$=\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\delta_{\vec{q}(x\underline{\sigma})+\vec{q}'(x\underline{\sigma}'x'^{-1}),\vec{q}_0+\mathfrak{Q}[\vec{q}(x\underline{\sigma})+\vec{q}'(x\underline{\sigma}'x'^{-1})]} = \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}$$

for all $(\underline{\sigma},\underline{\sigma}')\in P(\vec{q},\vec{q}';\vec{q}_0)$ and $c'=1,2,\ldots,n_{\kappa'}$. (II. 29)

Now we are confronted with the problem of finding by means of (I. 6), $|P(\vec{q},\vec{q}';\vec{q}_0)|n_{\kappa'}n_{\kappa_0}$ orthogonal vectors. Inspecting the scalar product of two such vectors we find

$$\left\{\begin{matrix} (0,\vec{q}) & (\kappa',\vec{q}') & (\kappa_0,\vec{q}_0) & (y\underline{\sigma}_j;y\underline{\sigma}_j'y'^{-1},c') \\ x\underline{\sigma}_i & x\underline{\sigma}_i'x'^{-1},d' & \underline{e}a & \end{matrix}\right\}$$

$$=\delta_{ij}\left\{\begin{matrix} (0,\vec{q}) & (\kappa',\vec{q}') & (\kappa_0,\vec{q}_0) & (y\underline{\sigma}_j;y\underline{\sigma}_j'y'^{-1},c') \\ x\underline{\sigma}_j & x\underline{\sigma}_j'x'^{-1},d' & \underline{e}a & \end{matrix}\right\}$$

for all $x, y \in P^{\vec{q}_0}$ $d', c' = 1, 2, \ldots, n_{\kappa'}$, (II. 30)

$$\left\{\begin{matrix}(0,\vec{q}) & (\kappa',\vec{q}') \\ x\underline{\sigma} & x\underline{\sigma}'x'^{-1}, d'\end{matrix}\middle|\begin{matrix}(\kappa_0,\vec{q}_0) & (y\underline{\sigma}; y\underline{\sigma}'y'^{-1}, c') \\ \underline{e}a & \end{matrix}\right\}$$

$$= \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|} B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(xy^{-1}) B^{\vec{q}}_{x\underline{\sigma}, y\underline{\sigma}}(xy^{-1}) B^{\vec{q}'}_{x\underline{\sigma}'x'^{-1}, y\underline{\sigma}'y'^{-1}}(xy^{-1}) \times \mathbb{R}^{\kappa'}_{d'c'}(x'^{-1}y') \mathbb{R}^{\kappa_0 *}_{aa}(xy^{-1})$$

for all $(\underline{\sigma}, \underline{\sigma}') \in P(\vec{q}, \vec{q}'; \vec{q}_0)$ and $x, y \in P^{\vec{q}_0}$. (II. 31)

Thereby we have to note that (II. 31) reduces to (II. 29) if $x = y$ and $d' = c'$, since the first equation implies $x' = y'$. In case we can find n_{κ_0} group elements $x_i \in P^{\vec{q}_0}$, $i = 1, 2, \ldots, n_{\kappa_0}$, such that

$$\mathbb{R}^{\kappa_0 *}_{aa}(x_i x_j^{-1}) = \delta_{ij} \quad \text{for all } i, j = 1, 2, \ldots, n_{\kappa_0}, \qquad \text{(II. 32)}$$

it follows from

$$x_i = x_j \Longleftrightarrow x_i' = x_j' (\in P^{\vec{q}'}), \qquad \text{(II. 33)}$$

$$\mathbb{R}^{\kappa'}_{d'c'}(e) = \delta_{d'c'}, \qquad \text{(II. 34)}$$

that we have solved the multiplicity problem. We can choose

$$w = (x_i\underline{\sigma}; x_i\underline{\sigma}'x_i'^{-1}, c') \quad \text{for all } (\underline{\sigma}, \underline{\sigma}') \in P(\vec{q}, \vec{q}'; \vec{q}_0)$$
$$i = 1, 2, \ldots, n_{\kappa_0}$$
$$c' = 1, 2, \ldots, n_{\kappa'} \qquad \text{(II. 35)}$$

as the multiplicity index, since (II. 30), (II. 32), and (II. 34) guarantee the orthogonality of the corresponding vectors (II. 28) with respect to the multiplicity index (II. 35).

Presupposing that we have found n_{κ_0} group elements $x_i \in P^{\vec{q}_0}$ satisfying (II. 32), the CG coefficients are immediately obtainable with the aid of (I. 8),

$$\begin{pmatrix}(0,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\tau} & \underline{\tau}', d'\end{pmatrix|\begin{matrix}(\kappa_0,\vec{q}_0) & w = (x_i\underline{\sigma}; x_i\underline{\sigma}'x_i'^{-1}, c') \\ \underline{e}a & \end{matrix}\end{pmatrix}$$

$$= \left(\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\right)^{1/2} \Delta^{\vec{q}_0}(\underline{e}, \underline{\tau}\underline{\sigma}^{-1}x_i^{-1}\underline{e}) \Delta^{\vec{q}'}(\underline{\tau}', \underline{\tau}\underline{\sigma}^{-1}x_i^{-1}\underline{\sigma}')$$
$$\times B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\underline{\tau}\underline{\sigma}^{-1}x_i^{-1}) B^{\vec{q}}_{\underline{\tau}, x_i\underline{\sigma}}(\underline{\tau}\underline{\sigma}^{-1}x_i^{-1})$$
$$\times B^{\vec{q}'}_{\underline{\tau}', x_i\underline{\sigma}'x_i'^{-1}}(\underline{\tau}\underline{\sigma}^{-1}x_i^{-1}) \mathbb{R}^{\kappa'}_{d'c'}(\underline{\tau}'^{-1}\underline{\tau}\underline{\sigma}^{-1}x_i^{-1}\sigma') \mathbb{R}^{\kappa_0 *}_{aa}(\underline{\tau}\underline{\sigma}^{-1}x_i^{-1})$$

for all $(\underline{\sigma}, \underline{\sigma}') \in P(\vec{q}, \vec{q}'; \vec{q}_0)$

$i = 1, 2, \ldots, n_{\kappa_0}$ $c' = 1, 2, \ldots, n_{\kappa'}$. (II. 36)

The remaining CG coefficients have to be determined by means of (I. 9),

$$\begin{pmatrix}(0,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\tau} & \underline{\tau}', d'\end{matrix}\middle|\begin{matrix}(\kappa_0,\vec{q}_0) & w = (x_i\underline{\sigma}; x_i\underline{\sigma}'x_i'^{-1}, c') \\ \underline{\sigma}_0, j & \end{matrix}\end{pmatrix}$$

$$= \left(\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\right)^{1/2} \Delta^{\vec{q}_0}(\underline{\sigma}_0, \underline{\tau}\underline{\sigma}^{-1}x_i^{-1}\underline{e}) \Delta^{\vec{q}'}(\underline{\tau}', \underline{\tau}\underline{\sigma}^{-1}x_i^{-1}\underline{\sigma}') B^{\vec{q}_0 *}_{\underline{\sigma}_0,\underline{e}}(\underline{\tau}\underline{\sigma}^{-1}x_i^{-1})$$
$$\times (\underline{\tau}\underline{\sigma}^{-1}x_i^{-1}) B^{\vec{q}}_{\underline{\tau}, x_i\underline{\sigma}}(\underline{\tau}\underline{\sigma}^{-1}x_i^{-1}) B^{\vec{q}'}_{\underline{\tau}', x_i\underline{\sigma}'x_i'^{-1}}$$
$$\times \mathbb{R}^{\kappa'}_{d'c'}(\tau'^{-1}\underline{\tau}\underline{\sigma}^{-1}x_i^{-1}\underline{\sigma}') \quad \mathbb{R}^{\kappa_0 *}_{ja}(\underline{\sigma}_0^{-1}\underline{\tau}\underline{\sigma}^{-1}x_i^{-1}),$$

$\underline{\sigma}_0 \in P : P^{\vec{q}_0}$, $j = 1, 2, \ldots, n_{\kappa_0}$. (II. 37)

C. 1. $\vec{q}, \vec{q}'$ = stars of higher symmetry, $\vec{q}_0$ = general star

Due to (III. 50) and (III. 51) of Ref. 12 it suffices to consider the vectors

$$\{\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)(\underline{\sigma},c;\underline{\sigma}',c')}_{\underline{e}} : (\underline{\sigma}, \underline{\sigma}') \in P(\vec{q}, \vec{q}'; \vec{q}_0);$$
$$c = 1, 2, \ldots, n_\kappa, \; c' = 1, 2, \ldots, n_{\kappa'}\}, \qquad \text{(II. 38)}$$

whose norm is given by

$$\|\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)(\underline{\sigma},c;\underline{\sigma}',c')}_{\underline{e}}\|^2 = \delta_{\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}'),\vec{q}_0+\vec{Q}[\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')]} = 1,$$

for all $(\underline{\sigma}, \underline{\sigma}') \in P(\vec{q}, \vec{q}'; \vec{q}_0)$, $c = 1, 2, \ldots, n_\kappa$,

$c' = 1, 2, \ldots, n_{\kappa'}$. (II. 39)

In order to find $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(0,\vec{q}_0)} = |P(\vec{q}, \vec{q}'; \vec{q}_0)| n_\kappa n_{\kappa'}$ orthogonal vectors we have to investigate the scalar product for vectors of the set (II. 38). The scalar product (I. 6) turns out to be

$$\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\sigma}_i, d & \underline{\sigma}_i', d'\end{matrix}\middle|\begin{matrix}(0,\vec{q}_0) & (\underline{\sigma}_j, c; \underline{\sigma}_j', c') \\ \underline{e} & \end{matrix}\right\} = \delta_{ij}\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\sigma}_j, d & \underline{\sigma}_j', d'\end{matrix}\middle|\begin{matrix}(0,\vec{q}_0) & (\underline{\sigma}_j, c; \underline{\sigma}_j', c') \\ \underline{e} & \end{matrix}\right\},$$

for all $(\underline{\sigma}_j, \underline{\sigma}_j') \in P(\vec{q}, \vec{q}'; \vec{q}_0)$, $d, c = 1, 2, \ldots, n_\kappa$,

$d', c' = 1, 2, \ldots, n_{\kappa'}$, (II. 40)

$$\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\sigma}, d & \underline{\sigma}', d'\end{matrix}\middle|\begin{matrix}(0,\vec{q}_0) & (\underline{\sigma}, c; \underline{\sigma}', c') \\ \underline{e} & \end{matrix}\right\} = \mathbb{R}^{\kappa}_{dc}(e) \mathbb{R}^{\kappa'}_{d'c'}(e) = \delta_{dc}\delta_{d'c'}. \qquad \text{(II. 41)}$$

Hence we can choose

$$w = (\underline{\sigma}, c; \underline{\sigma}', c') \quad \text{for all } (\underline{\sigma}, \underline{\sigma}') \in P(\vec{q}, \vec{q}'; \vec{q}_0),$$
$$c = 1, 2, \ldots, n_\kappa,$$
$$c' = 1, 2, \ldots, n_{\kappa'}, \qquad \text{(II. 42)}$$

as a multiplicity index, since (II. 40) and (II. 41) guarantee the orthogonality of the corresponding vectors. Therefore, the CG coefficients take the form

$$\begin{pmatrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\tau}, d & \underline{\tau}', d'\end{matrix}\middle|\begin{matrix}(0,\vec{q}_0), & w = (\underline{\sigma}, c; \underline{\sigma}', c') \\ \underline{e} & \end{matrix}\end{pmatrix} = \delta_{\underline{\tau},\underline{\sigma}}\delta_{\underline{\tau}',\underline{\sigma}'}\delta_{dc}\delta_{d'c'}$$

for all $(\underline{\sigma}, \underline{\sigma}') \in P(\vec{q}, \vec{q}'; \vec{q}_0)$, $c = 1, 2, \ldots, n_\kappa$,

$c' = 1, 2, \ldots, n_{\kappa'}$, (II. 43)

and the remaining CG coefficients must be calculated by means of (I. 9),

$$\begin{pmatrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\tau}, d & \underline{\tau}', d'\end{matrix}\middle|\begin{matrix}(0,\vec{q}_0) & w = (\underline{\sigma}, c; \underline{\sigma}', c') \\ \underline{\sigma}_0 & \end{matrix}\end{pmatrix}$$

$$= \Delta^{\vec{q}}(\underline{\tau}, \underline{\sigma}_0\underline{\sigma}) \Delta^{\vec{q}'}(\underline{\tau}', \underline{\sigma}_0\underline{\sigma}') B^{\vec{q}_0 *}_{\underline{\sigma}_0,\underline{e}}(\underline{\sigma}_0) B^{\vec{q}}_{\underline{\tau},\underline{\sigma}}(\underline{\sigma}_0) B^{\vec{q}'}_{\underline{\tau}',\underline{\sigma}'}(\underline{\sigma}_0)$$
$$\times \mathbb{R}^{\kappa}_{dc}(\underline{\tau}^{-1}\underline{\sigma}_0\sigma) \mathbb{R}^{\kappa'}_{d'c'}(\underline{\tau}'^{-1}\underline{\sigma}_0\underline{\sigma}'), \quad \underline{\sigma}_0 \in P. \qquad \text{(II. 44)}$$

C. 2. $\vec{q},\vec{q}'$ = stars of higher symmetry, $\vec{q}_0$ = star of higher symmetry

For the sake of simplicity we assume for the following that

$$|P(\vec{q},\vec{q}';\vec{q}_0)|=1 \tag{II. 45}$$

is satisfied, which implies, however, no loss of generality, since orthogonality with respect to different pairs of $P(\vec{q},\vec{q}';\vec{q}_0)$ can be achieved completely in the same way as for all previous cases. First of all let us recall (II. 9), (III. 59), (III. 62)–(III. 64), and (III. 56) of Ref. 12,

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'}:=\underline{\sigma}P^{\vec{q}}\underline{\sigma}^{-1}\cap\underline{\sigma}'P^{\vec{q}'}\underline{\sigma}'^{-1}\cap P^{\vec{q}_0}, \tag{II. 46}$$

$$x\in P^{\vec{q},\vec{q}';\vec{q}_0}\Longleftrightarrow x\underline{\sigma}=\underline{\sigma}z \quad\text{and}\quad x\underline{\sigma}'=\underline{\sigma}'z', \tag{II. 47}$$

$$v_j\in P^{\vec{q}_0}:P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'}, \tag{II. 48}$$

$$\underline{\sigma}_j=v_j\underline{\sigma}z_j^{-1}\in P:P^{\vec{q}}\text{ with } z_j\in P^{\vec{q}}, \tag{II. 49}$$

$$\underline{\sigma}_k'=v_k\underline{\sigma}'z_k'^{-1}\in P:P^{\vec{q}'}\text{ with } z_k'\in P^{\vec{q}'}, \tag{II. 50}$$

where [according to (II. 45)] for the fixed pair $(\underline{\sigma},\underline{\sigma}')$

$$\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')=\vec{q}_0+\vec{Q}[\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')] \tag{II. 51}$$

is satisfied. Because of (III. 68) of Ref. 12, it is suggested that it is sufficient to consider the set

$$\{\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(v_j\underline{\sigma}z_j^{-1},c;v_j\underline{\sigma}'z_j'^{-1},c')}_{\underline{e}a}:v_j\in P^{\vec{q}_0}:P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'}, \quad c=1,2,\ldots,n_\kappa,\ c'=1,2,\ldots,n_{\kappa'}\} \tag{II. 52}$$

in order to be able to find $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ orthogonal vectors. In this connection let us remark that even for this case (being, of course, the most complicated one when space group CG coefficients should be determined), the proposed method gives rise to essential simplifications. One of these simplifications (being a consequence of the wave vector selection rules) is, that we have to only consider $|P^{\vec{q}_0}:P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'}|n_\kappa n_{\kappa'}$ vectors instead of $|P:P^{\vec{q}}|n_\kappa|P:P^{\vec{q}'}|n_{\kappa'}$ in order to find $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ orthogonal vectors. The components of the elements of (II. 52) are given by (I. 3) and their norm by (I. 5),

$$\begin{aligned}\{\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma}_j,c;\underline{\sigma}_j',c')}_{\underline{e}a}\}_{\underline{\tau}d,\underline{\tau}'d'}
&=\delta_{\vec{q}(\underline{\tau})+\vec{q}'(\underline{\tau}'),\vec{q}_0+\vec{Q}[\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')]}\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\\
&\times\sum_{\beta\in P^{\vec{q}_0}}\Delta^{\vec{q}_0}(\underline{e},\beta\underline{e})\Delta^{\vec{q}}(\underline{\tau},\beta\underline{\sigma}_j)\Delta^{\vec{q}'}(\underline{\tau}',\beta\underline{\sigma}_j')\\
&\times B^{\vec{q}_0}_{\underline{e},\underline{e}}{}^*(\beta)B^{\vec{q}}_{\underline{\tau},\underline{\sigma}_j}(\beta)B^{\vec{q}'}_{\underline{\tau}',\underline{\sigma}_j'}(\beta)\mathbf{R}^\kappa_{dc}(\underline{\tau}^{-1}\beta\underline{\sigma}_j)\mathbf{R}^{\kappa'}_{d'c'}(\underline{\tau}'^{-1}\beta\underline{\sigma}_j')\mathbf{R}^{\kappa_0}_{aa}{}^*(\beta),
\end{aligned} \tag{II. 53}$$

$$\underline{\tau}\in P:P^{\vec{q}},\quad d=1,2,\ldots,n_\kappa,\quad \underline{\tau}'\in P:P^{\vec{q}'},\quad d'=1,2,\ldots,n_{\kappa'},$$

$$\begin{aligned}\|\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma}_j,c;\underline{\sigma}_j',c')}_{\underline{e}a}\|^2
&=\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\sum_{x\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'}}B^{\vec{q}_0}_{\underline{e},\underline{e}}{}^*(x)B^{\vec{q}}_{\underline{\sigma}_j,\underline{\sigma}_j}(x)B^{\vec{q}'}_{\underline{\sigma}_j',\underline{\sigma}_j'}(x)\\
&\times\mathbf{R}^\kappa_{cc}(\underline{\sigma}_j^{-1}x\underline{\sigma}_j)\mathbf{R}^{\kappa'}_{c'c'}(\underline{\sigma}_j'^{-1}x\underline{\sigma}_j')\mathbf{R}^{\kappa_0}_{aa}{}^*(x),\\
&\text{for all } v_j\in P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'},\quad c=1,2,\ldots,n_\kappa,\quad c'=1,2,\ldots,n_{\kappa'}.
\end{aligned} \tag{II. 54}$$

In order to find $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ orthogonal vectors we have to inspect once more the scalar product (I. 6) for vectors (II. 52),

$$\begin{aligned}&\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') & (\kappa_0,\vec{q}_0) & (\underline{\sigma}_j,c;\underline{\sigma}_j',c')\\ \underline{\sigma}_\kappa,d & \underline{\sigma}_\kappa',d' & \underline{e}a & \end{matrix}\right\}\\
&=\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\sum_{\beta\in P^{\vec{q}_0}}\Delta^{\vec{q}_0}(\underline{e},\beta\underline{e})\Delta^{\vec{q}}(\underline{\sigma}_\kappa,\beta\underline{\sigma}_j)\Delta^{\vec{q}'}(\underline{\sigma}_\kappa',\beta\underline{\sigma}_j')\\
&\times B^{\vec{q}_0}_{\underline{e},\underline{e}}{}^*(\beta)B^{\vec{q}'}_{\underline{\sigma}_k',\underline{\sigma}_j}(\beta)B^{\vec{q}}_{\underline{\sigma}_k,\underline{\sigma}_j'}(\beta)\mathbf{R}^\kappa_{dc}(\underline{\sigma}_\kappa^{-1}\beta\underline{\sigma}_j)\\
&\times\mathbf{R}^{\kappa'}_{d'c'}(\underline{\sigma}_k'^{-1}\beta\underline{\sigma}_j')\mathbf{R}^{\kappa_0}_{aa}{}^*(\beta)\quad\text{for all } v_j,v_k\in P^{\vec{q}_0}:P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'},\\
&d,c=1,2,\ldots,n_\kappa,\quad d',c'=1,2,\ldots,n_{\kappa'}.
\end{aligned} \tag{II. 55}$$

Therefore, we have to note that for even the most complicated case we are able to derive a simple defining equation for the multiplicity index. In case we can find, with the aid of (II. 55), just $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ pairwise orthogonal vectors the corresponding indices can be chosen as multiplicity index.

$$w=(\underline{\sigma}_v,c_v;\underline{\sigma}_v',c_v'),\quad v=1,2,\ldots,m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}. \tag{II. 56}$$

Consequently the whole set of CG coefficients can be readily calculated by means of (I. 8) and (I. 9).

However we have to realize that for the most complicated case (apart from special cases which shall be discussed in the following) the scalar product (II. 55) (being the key equation for fixing the multiplicity index) yields further information, if and only if the space group and its unirreps are explicitly known. Therefore we cannot expect for the general case to find general formulas (such as for A. 1 and A. 2; B. 1 and B. 2 and C. 2. a) for the multiplicity index without specification of the space group with its unirreps.

C. 2. a. $\vec{q},\vec{q}',\vec{q}_0$ = stars of higher symmetry; $P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'}=\{e\}$

Presupposing that (II. 45) is valid, we have to consider the set

$$\{\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(x_i\underline{\sigma}z_i^{-1},c;x_i\underline{\sigma}'z_i'^{-1},c')}_{\underline{e}a}:x_i\in P^{\vec{q}_0},\quad c=1,2,\ldots,n_\kappa,\quad c'=1,2,\ldots,n_{\kappa'}\}, \tag{II. 57}$$

whose order is equal to $|P^{\vec{q}_0}|n_\kappa n_{\kappa'}$. In order to simplify the following considerations we have to take in some way the corresponding version of (III. 67) and (III. 68) of Ref. 12 into account. The norm of the elements of (II. 57) follows from (II. 54).

$$\begin{aligned}\|\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma}_i,c;\underline{\sigma}_i',c')}_{\underline{e}a}\|^2
&=\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\delta_{\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}'),\vec{q}_0+\vec{Q}[\vec{q}(\underline{\sigma})+\vec{q}'(\underline{\sigma}')]}=\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\\
&\text{for all } x_i\in P^{\vec{q}_0};\quad c=1,2,\ldots,n_\kappa;\quad c'=1,2,\ldots,n_{\kappa'}.
\end{aligned} \tag{II. 58}$$

In order to find $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}=n_\kappa n_{\kappa'}n_{\kappa_0}$ pairwise orthogonal vectors we inspect the scalar product (II. 55) for the elements of (II. 57) by inserting for $\underline{\sigma}_j$, $\underline{\sigma}_j'$ and $\underline{\sigma}_k$, $\underline{\sigma}_k'$ the notation (II. 49, 50),

$$\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\sigma}_k, d & \underline{\sigma}_k', d'\end{matrix}\middle| \begin{matrix}(\kappa_0,\vec{q}_0) & (\underline{\sigma}_j, c; \underline{\sigma}_j', c') \\ \underline{e}a & \end{matrix}\right\}$$

$$= \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|} \sum_{\beta\in P^{\vec{q}_0}} \Delta^{\vec{q}_0}(\underline{e}, \beta\underline{e})\Delta^{\vec{q}}(\underline{\sigma}_k, \beta\underline{\sigma}_j)\Delta^{\vec{q}'}(\underline{\sigma}_k', \beta\underline{\sigma}_j')$$
$$B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\beta)B^{\vec{q}}_{\underline{\sigma}_k,\underline{\sigma}_j}(\beta)B^{\vec{q}'}_{\underline{\sigma}_k',\underline{\sigma}_j'}(\beta)$$
$$\times \mathbb{R}^{\kappa}_{dc}(z_k\underline{\sigma}^{-1}x_k^{-1}\beta x_j\underline{\sigma}z_j^{-1})\mathbb{R}^{\kappa'}_{d'c'}(z_k'\underline{\sigma}'^{-1}x_k^{-1}\beta x_j\underline{\sigma}'z_j'^{-1})\mathbb{R}^{\kappa_0 *}_{aa}(\beta). \quad \text{(II. 59)}$$

Because of $P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'}=\{e\}$ we obtain, as a consequence of the generalized Kronecker deltas, that

$$x_k^{-1}\beta x_j \in P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'} \Longleftrightarrow \beta = x_k x_j^{-1}. \quad \text{(II. 60)}$$

Hence

$$\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\sigma}_k, d & \underline{\sigma}_k', d'\end{matrix}\middle| \begin{matrix}(\kappa_0,\vec{q}_0) & (\underline{\sigma}_j, c; \underline{\sigma}_j', c') \\ \underline{e}a & \end{matrix}\right\}$$

$$= \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|} B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(x_k x_j^{-1})B^{\vec{q}}_{\underline{\sigma}_k,\underline{\sigma}_j}(x_k x_j^{-1})B^{\vec{q}'}_{\underline{\sigma}_k',\underline{\sigma}_j'}(x_k x_j^{-1})$$
$$\times \mathbb{R}^{\kappa}_{dc}(z_k z_j^{-1})\mathbb{R}^{\kappa'}_{d'c'}(z_k' z_j'^{-1})\mathbb{R}^{\kappa_0 *}_{aa}(x_k x_j^{-1}),$$
$$\text{for all } x_j, x_k \in P^{\vec{q}_0}, \quad \text{(II. 61)}$$

thereby we have to note that $x_i = x_j$ implies $z_i = z_j$ and $z_i' = z_j'$. Now in case we can find n_{κ_0} different group elements $x_i \in P^{\vec{q}_0}$ satisfying

$$\mathbb{R}^{\kappa_0 *}_{aa}(x_i x_j^{-1}) = \delta_{ij} \quad \text{for all } i,j = 1,2,\ldots,n_{\kappa_0}, \quad \text{(II. 62)}$$

it follows from

$$x_i = x_j \Longleftrightarrow z_i = z_j \quad \text{and} \quad z_i' = z_j', \quad \text{(II. 63)}$$

$$\mathbb{R}^{\kappa}_{dc}(e) = \delta_{dc} \quad \text{and} \quad \mathbb{R}^{\kappa'}_{d'c'}(e) = \delta_{d'c'}, \quad \text{(II. 64)}$$

that we can choose as multiplicity index w

$$w = (x_i\underline{\sigma}z_i^{-1}, c; x_i\underline{\sigma}'z_i'^{-1}, c')$$
$$\text{for all } i = 1,2,\ldots,n_{\kappa_0}, c = 1,2,\ldots,n_{\kappa},$$
$$c' = 1,2,\ldots,n_{\kappa'}, \quad \text{(II. 65)}$$

since (II. 62)—(II. 64) guarantee the orthogonality of the corresponding vectors. The CG coefficients are obtained immediately with the aid of (I. 8),

$$\left(\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\tau}, d & \underline{\tau}', d'\end{matrix}\middle| \begin{matrix}(\kappa_0,\vec{q}_0) & w = (\underline{\sigma}_i, c; \underline{\sigma}_i', c') \\ \underline{e}a & \end{matrix}\right)$$

$$= \left(\frac{n_{\kappa_0}}{|P^{\vec{q}_0}|}\right)^{1/2} \sum_{\beta\in P^{\vec{q}_0}} \Delta^{\vec{q}_0}(\underline{e}, \beta\underline{e})\Delta^{\vec{q}}(\underline{\tau}, \beta\underline{\sigma}_i)\Delta^{\vec{q}'}(\underline{\tau}', \beta\underline{\sigma}_i')$$
$$\times B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\beta)B^{\vec{q}}_{\underline{\tau},\underline{\sigma}_i}(\beta)B^{\vec{q}'}_{\underline{\tau}',\underline{\sigma}_i'}(\beta)\mathbb{R}^{\kappa}_{dc}(\underline{\tau}^{-1}\beta\underline{\sigma}_i)\mathbb{R}^{\kappa'}_{d'c'}(\underline{\tau}'^{-1}\beta\underline{\sigma}_i')$$
$$\times \mathbb{R}^{\kappa_0 *}_{aa}(\beta). \quad \text{(II. 66)}$$

The sum on the right-hand side of (II. 66) reduces to a single term [compare (II. 61)], if $\underline{\tau} = x\underline{\sigma}z^{-1}$ ($z \in P^{\vec{q}}$) and $\underline{\tau}' = x\underline{\sigma}'z'^{-1}$ ($z' \in P^{\vec{q}'}$) with $x \in P^{\vec{q}_0}$, or is otherwise zero. The remaining CG coefficients can be readily calculated with (I. 9).

C. 2. b. $\vec{q},\vec{q}',\vec{q}_0$ = stars of higher symmetry, $P^{q,q';q_0}_{\sigma,\sigma} = P^{q_0}$

This implies, because of

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'} = \underline{\sigma}P^{\vec{q}}\underline{\sigma}^{-1} \cap \underline{\sigma}'P^{\vec{q}'}\underline{\sigma}'^{-1} \cap P^{\vec{q}_0} = P^{\vec{q}_0}, \quad \text{(II. 67)}$$

that the corresponding set (II. 52) is given by

$$\{B^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma},c;\underline{\sigma}',c')}_{\underline{e}a}: c = 1,2,\ldots,n_{\kappa};$$
$$c' = 1,2,\ldots,n_{\kappa'}\}. \quad \text{(II. 68)}$$

Due to (II. 45) the pair $(\underline{\sigma},\underline{\sigma}')$ is fixed. Now it is obvious that the task to fix the multiplicity index is resolved into a simpler problem, since the order $n_{\kappa}n_{\kappa'}$ of the set (II. 68) is essentially smaller than $|P: P^{\vec{q}}|n_{\kappa}|P: P^{\vec{q}'}|n_{\kappa'}$. Due to (II. 54) we have

$$\|\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{\sigma},c;\underline{\sigma}',c')}_{\underline{e}a}\|^2$$
$$= \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|} \sum_{\beta\in P^{\vec{q}_0}} B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\beta)B^{\vec{q}}_{\underline{\sigma},\underline{\sigma}}(\beta)B^{\vec{q}'}_{\underline{\sigma}',\underline{\sigma}'}(\beta)(\mathbb{R}^{\kappa}_{cc}(\underline{\sigma}^{-1}\beta\underline{\sigma})$$
$$\times \mathbb{R}^{\kappa'}_{c'c'}(\underline{\sigma}'^{-1}\beta\underline{\sigma}')\mathbb{R}^{\kappa_0 *}_{aa}(\beta)$$
$$\text{for all } c = 1,2,\ldots,n_{\kappa}; \quad c' = 1,2,\ldots,n_{\kappa'}, \quad \text{(II. 69)}$$

and the $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ pairwise orthogonal vectors can be found by means of the simple formula

$$\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{\sigma}, d & \underline{\sigma}', d'\end{matrix}\middle| \begin{matrix}(\kappa_0,\vec{q}_0) & (\underline{\sigma}, c; \underline{\sigma}', c') \\ \underline{e}a & \end{matrix}\right\}$$

$$= \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|} \sum_{\beta\in P^{\vec{q}_0}} B^{\vec{q}_0 *}_{\underline{e},\underline{e}}(\beta)B^{\vec{q}}_{\underline{\sigma},\underline{\sigma}}(\beta)B^{\vec{q}'}_{\underline{\sigma}',\underline{\sigma}'}(\beta)\mathbb{R}^{\kappa}_{dc}(\underline{\sigma}^{-1}\beta\underline{\sigma})$$
$$\times \mathbb{R}^{\kappa'}_{d'c'}(\underline{\sigma}'^{-1}\beta\underline{\sigma}')\mathbb{R}^{\kappa_0 *}_{aa}(\beta), \quad \text{(II. 70)}$$
$$d, c = 1,2,\ldots, n_{\kappa} \quad d', c' = 1,2,\ldots, n_{\kappa'}.$$

In case we can find $m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)}$ pairwise orthogonal vectors with the aid of the scalar product (II. 70) the multiplicity problem is solved and the CG coefficients have to be determined by means of (I. 8) and (I. 9).

We conclude this section by considering the special case

$$P^{\vec{q},\vec{q}';\vec{q}_0}_{\underline{\sigma},\underline{\sigma}'} = P^{\vec{q}_0} = P^{\vec{q}} = P^{\vec{q}'}, \quad \text{with } \underline{\sigma} = \underline{\sigma}' = e, \quad \text{(II. 71)}$$

of (II. 67), which of course does not reflect the general situation where at least one of the two left coset representatives must be different from the identity element e, so that $\vec{q}_0$ belongs to ΔBZ. Equations (II. 69) and (II. 70) take the simple form

$$\|\vec{B}^{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)(\underline{e},c;\underline{e},c')}_{\underline{e}a}\|^2$$
$$= \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|} \sum_{\beta\in P^{\vec{q}_0}} e^{i(\vec{q}_0-\vec{q}-\vec{q}')\cdot\vec{\tau}(\beta)}\mathbb{R}^{\kappa}_{cc}(\beta)\mathbb{R}^{\kappa'}_{c'c'}(\beta)\mathbb{R}^{\kappa_0 *}_{aa}(\beta), \quad \text{(II. 72)}$$

$$\left\{\begin{matrix}(\kappa,\vec{q}) & (\kappa',\vec{q}') \\ \underline{e}d & \underline{e}, d'\end{matrix}\middle| \begin{matrix}(\kappa_0,\vec{q}_0) & (\underline{e}, c; \underline{e}, c') \\ \underline{e}a & \end{matrix}\right\}$$

$$= \frac{n_{\kappa_0}}{|P^{\vec{q}_0}|} \sum_{\beta\in P^{\vec{q}_0}} e^{i(\vec{q}_0-\vec{q}-\vec{q}')\cdot\vec{\tau}(\beta)}\mathbb{R}^{\kappa}_{dc}(\beta)\mathbb{R}^{\kappa'}_{d'c'}(\beta)\mathbb{R}^{\kappa_0 *}_{aa}(\beta), \quad \text{(II. 73)}$$

which represents a simple example that the general formulas (II. 36) (II. 38) and (II. 39) of Ref. 1 must be applied to the projective point group unirreps. This application of the above-mentioned formulas of Ref. 1

shows directly why we have generalized our method to projective representations. In this connection we have to realize that we need this generalized procedure in some way, since otherwise (I. 3), (I. 5), (I. 6), (I. 8) and (I. 9) cannot be derived. We conclude from (II. 70) that only for the special case (II. 67) is the problem of determining the multiplicity index completely reduced to the multiplicity problem for (projective) point group representations.

CONCLUDING REMARKS

In the preceding sections we have demonstrated how useful and practical the present method is for determining quite generally space group CG coefficients. Thereby we have shown that without reference to a special space group the multiplicity index and the corresponding CG coefficients can be determined immediately for nearly all cases. Even for the most complicated case, we were able to derive simple formulas which allow us to identify the multiplicity index with special column indices of the Kronecker product. However for the last case we must specify the space group with its unirreps in order to be able to actually carry out the calculation of the CG coefficients. But in comparison with other methods we have to note that apart from special cases there enter different pairs of left coset representatives $(\underline{\sigma}_\nu, \underline{\sigma}'_\nu) \in P(\vec{q}, \vec{q}')$, which are connected by elements of $P^{\vec{q}_0}$, when fixing the multiplicity index. Finally it should be noted that the considerations connecting the multiplicity index conversely prove the correctness of the multiplicity formula.

[1]R. Dirl, "Clebsch–Gordan coefficients: General theory," J. Math. Phys. **20**, 659 (1979).
[2]R. Berenson and J. L. Birman, J. Math. Phys. **16**, 227 (1975).
[3]R. Berenson, I. Itzkan, and J. L. Birman, J. Math. Phys. **16**, 236 (1975).
[4]G. F. Koster, J. O. Dimmock, R. G. Wheeler, and H. Statz, *Properties of the Thirty-two Point Groups* (Technology Press, Cambridge, Mass., 1963).
[5]J. L. Birman, "Theory of Crystal Space Groups and Infrared Raman Processes of Insulated Crystals" in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1974).
[6]D. B. Litvin and J. Zak, J. Math. Phys. **9**, 212 (1968).
[7]I. Sakata, J. Math. Phys. **15**, 1702 (1974).
[8]L. K. Saulevich, D. T. Sviridov, and Yu. F. Smirnov, Sov. Phys. Crystall. **15**, 355 (1970).
[9]P. Gard, J. Phys. A **6**, 1837 (1973).
[10]S. Schindler and R. Mirman, J. Math. Phys. **18**, 1678 (1977).
[11]S. Schindler and R. Mirman, J. Math. Phys. **18**, 1697 (1977).
[12]R. Dirl, "Multiplicities for space group representations," J. Math. Phys. **20**, 664 (1979).

Clebsch–Gordan coefficients for $Pn3n$

R. Dirl

Institut für Theoretische Physik, TU Wien, A-1040 Wien, Karlsplatz 13, Austria
(Received 28 December 1977)

A general method for calculating Clebsch–Gordan coefficients is applied to determine such coefficients for the nonsymmorphic space group $Pn3n$.

INTRODUCTION

In the previous paper[1] we specialized our general formulas of Ref. 2 to space group representations. In particular we showed how the CG coefficients can be determined quite generally by using projection techniques and by a special manner of choosing the multiplicity index. In this paper we shall give the results of calculation of CG coefficients for $Pn3n$. Thereby we consider examples which belong to the various cases which have been discussed in the previous papers.[1,3]

The organization of the material of this article is as follows: In Sec. I we summarize the notations and definitions concerning $Pn3n$ and its representations which are needed for the following. Furthermore we list the examples we want to consider. The corresponding multiplicities are calculated in Sec. II. In Sec. III we actually determine the multiplicity index for the various examples. The corresponding CG coefficients can be readily obtained by means of the general formulas of Ref. 1.

I. DISCUSSION OF VARIOUS CASES: DEFINITIONS AND NOTATIONS

We choose $Pn3n$ as an example in order to show how the present method works when determining the multiplicity index of space group CG coefficients for various cases. For this purpose we recall some definitions and notations which concern the nonsymmorphic space group $Pn3n$. We choose the lattice constant of the crystal as one which can be done without any loss of generality. Furthermore we define by

$$\vec{\tau}(n) = \vec{0}, \quad \text{for all } n \in \mathscr{O}, \tag{I.1}$$

$$\vec{\tau}(In) = (\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}), \quad \text{for all } n \in \mathscr{O}, \tag{I.2}$$

as nonprimitive lattice translations where the group element I of the point group $\mathscr{O}_h$ denotes the inversion. In this connection we mention that we hereafter use the notation of Ref. 4 for the elements of the point group $\mathscr{O}_h$, respectively, elements of the fundamental domain ΔBZ of the corresponding Brillouin zone,

$$\Delta BZ := \{\vec{q} = \pi(x,y,z): x,y,x \in [0,1] \text{ and } 1 \geqslant y \geqslant x \geqslant z \geqslant 0\}. \tag{I.3}$$

The orthogonal 3×3 matrices $D(\alpha)$; $\alpha \in P \simeq G/T$ which essentially enter into the definition of the groups of the $\vec{q}$ vectors (respectively, their homomorphic images $P^{\vec{q}} \simeq G^{\vec{q}}/T$) can be readily obtained from Table 1.4 of Ref. 4. Complete sets of projective matrix unirreps with their corresponding factor systems of the little cogroups $P^{\vec{q}}$ are listed in full detail in Ref. 5 for all points of the surface of ΔBZ and for some $\vec{q}$'s lying inside of ΔBZ.

It is obvious when calculating CG coefficients for any group that the explicit knowledge of the corresponding unirreps is required in some way. This implies for our case that we have to known not only the projective unirreps of the little cogroups $P^{\vec{q}}$, but also to fix a corresponding set of left coset representatives $P : P^{\vec{q}}$. In doing so we are in fact able to calculate, by means of the present method, the space group CG coefficients.

In the following we want to consider several examples which belong to the cases A.2, B.2, C.2, C.2.a, and C.2.b of Ref. 1. The reason for not investigating the cases A.1, B.1, and C.1 is that we have already solved the multiplicity problem for these cases without any reference to a special space group, so that a discussion would be trivial.

Now we list some examples which belong to the cases A.2, B.2, C.2, C.2.a, and C.2.b of Ref. 1. If necessary for the following considerations we shall also write down the corresponding set $P : P^{\vec{q}}$ of left coset representatives whose elements, for the sake of simplicity, will not be underlined in the following.

Case A.2: $\vec{q},\vec{q}'$ = general stars, $\vec{q}_0$ = star of higher symmetry

$$\vec{q} = \pi(x,y,z) \in \Delta BZ \Longleftrightarrow 1 > y > x > z > 0, \tag{I.4}$$

$$\vec{q}' = \pi(1-x,1-z,1-y) \in \Delta BZ, \tag{I.5}$$

$$\vec{q}_0 = :\vec{q}_R = \vec{q} + \vec{q}'(\sigma_{df}) = \pi(1,1,1), \tag{I.6}$$

$$\vec{q}'(\sigma_{df}) = D(\sigma_{df})\vec{q}', \quad D(\sigma_{df}) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}, \tag{I.7}$$

$$P^{\vec{q}} = \sigma_{df} P^{\vec{q}'} \sigma_{df} = \{E\}, \quad P^R = \mathscr{O}_h, \tag{I.8}$$

$$P(\vec{q},\vec{q}';\vec{q}_0) = \{(E,\sigma_{df})\}. \tag{I.9}$$

In order to convince ourselves of the correctness of (I.9) we have to observe that the same arguments hold as for (III.47) of Ref. 1.

Case B.2: $\vec{q}$ = general star, $\vec{q}'$ = star of higher symmetry, $\vec{q}_0$ = star of higher symmetry

$$\vec{q} = \pi(a,1-y,a-x) \in \Delta BZ \Longleftrightarrow 1 > 1-y > a > x > 0, \tag{I.10}$$

$$\vec{q}' = \pi(x,y,0) \in \Delta BZ, \tag{I.11}$$

 0022-2488/79/040679-06$01.00

$$P^{\vec{q}'} = \{E,\sigma_z\},$$
$$\vec{q}_0 = :\vec{q}_s = \vec{q} + \vec{q}'(\sigma_{de}) = \pi(a,1,a), \tag{I.12}$$
$$D(\sigma_{de}) = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}, \tag{I.13}$$
$$P^{\vec{q}_s} = \{E,C_{2c},\sigma_{de},\sigma_y\}, \tag{I.14}$$
$$P(\vec{q},\vec{q}';\vec{q}_0) = \{(E,\sigma_{de})\}. \tag{I.15}$$

The proof of (I.15) is rather complicated and requires somewhat more extensive geometrical considerations concerning the wave vector selection rules.

Case C.2.a: $\vec{q},\vec{q}',\vec{q}_0$ = stars of higher symmetry, $P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} = \{E\}$

$$\vec{q} = \pi(x,1-y,0)\in\Delta BZ \Longleftrightarrow 1 > 1-y > x > 0, \tag{I.16}$$
$$\vec{q}' = \pi(y,y,y)\in\Delta BZ, \tag{I.17}$$
$$P^{\vec{q}'} = P^{\Lambda} = \{E,C^{\pm}_{31},\sigma_{db},\sigma_{de},\sigma_{df}\}, \tag{I.18}$$
$$\vec{q}_0 = \vec{q} + \vec{q}' = \pi(x+y,1,y)\in\Delta BZ, \tag{I.19}$$
$$P^{\vec{q}_0} = \{E,\sigma_y\}, \tag{I.20}$$
$$P(\vec{q},\vec{q}';\vec{q}_0) = \{(E,E)\}. \tag{I.21}$$

Thereby the proof of (I.21) is trivial.

Case C.2: $\vec{q},\vec{q}',\vec{q}_0$ = stars of higher symmetry, $\{E\} \subset P^{\vec{q},\vec{q}';\vec{q}_0}_{\sigma,\sigma'} \subseteq P^{\vec{q}_0}$.

We shall consider this case (containing Case C.2.b) in order to show how the present method works for the most complicated case. Thereby we investigate the whole CG matrix and not only parts like in the previous cases:

$$\vec{q} = :\vec{q}_\Delta = \pi(0,y,0)\in\Delta BZ \Leftarrow 1 > y > \tfrac{1}{2}, \tag{I.22}$$
$$\vec{q}' = :\vec{q}'_\Delta = \pi(0,1-y,0)\in\Delta BZ, \tag{I.23}$$
$$P^{\vec{q}} = P^{\vec{q}'} = P^{\Delta} = \{E,C_{2y},C^{\pm}_{4y},\sigma_x,\sigma_z,\sigma_{dc},\sigma_{de}\}$$
$$= \{E,\sigma_x\} \circledS \{E,C_{2y},C^{\pm}_{4y}\}, \tag{I.24}$$
$$P{:}P^{\vec{q}'} = \{E,C^{-}_{4z},C^{+}_{4x},I,IC^{-}_{4z},IC^{+}_{4x}\}. \tag{I.25}$$

Now we admit all possible cases which must be considered, if the whole CG matrix shall be calculated. Three cases occur,

Case (i):
$$\vec{q}_0 = :\vec{q}_x = \vec{q} + \vec{q}' = \pi(0,1,0)\in\Delta BZ, \tag{I.26}$$
$$P^{\vec{q}_0} = P^{x} = \{E,I\} \times P^{\Delta}, \tag{I.27}$$
$$P^{\vec{q},\vec{q}';\vec{q}_0}_{E,E} = P^{\Delta} \subset P^{x} \Rightarrow P^{\vec{q}_0}{:}P^{\vec{q},\vec{q}';\vec{q}_0}_{E,E} = P^{x}{:}P^{\Delta} = \{E,I\}. \tag{I.28}$$

Of course this case belongs to the most general one.

Case (ii):
$$\vec{q}_0 = :\vec{q}''_\Delta = \vec{q} + \vec{q}'(I) = \pi(0,2y-1,0)\in\Delta BZ, \tag{I.29}$$
$$\vec{q}'(I) = D(I)\vec{q}', \quad D(I) = \begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix}, \tag{I.30}$$
$$P^{\vec{q}_0} = P^{\Delta} = IP^{\Delta}I, \tag{I.31}$$
$$P^{\vec{q},\vec{q}';\vec{q}_0}_{E,I} = P^{\Delta} \cap IP^{\Delta}I \cap P^{\vec{q}_0} = P^{\vec{q}_0} = P^{\Delta}. \tag{I.32}$$

This example belongs, because of (I.31), to the special case (II.71) of Ref. 1.

Case (iii):
$$\vec{q}_0 = \vec{q} + \vec{q}'(C^{-}_{4z}) = \pi(1-y,y,0)\in\Delta BZ, \tag{I.33}$$
$$\vec{q}'(C^{-}_{4z}) = D(C^{-}_{4z})\vec{q}', \quad D(C^{-}_{4z}) = \begin{vmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \tag{I.34}$$
$$P^{\vec{q}_0} = \{E,\sigma_z\}, \tag{I.35}$$
$$P^{\vec{q},\vec{q}';\vec{q}_0}_{E,C^{-}_{4z}} = P^{\Delta} \cap C^{-}_{4z} P^{\Delta} (C^{-}_{4z})^{-1} \cap P^{\vec{q}_0} = P^{\vec{q}_0}. \tag{I.36}$$

This is just an example for Case C.2.b.

It should be noted that we have restricted in (I.22) the variable y. This however implies no loss of generality, but guarantees merely that the corresponding $\vec{q}_0$'s belong to ΔBZ.

II. MULTIPLICITIES FOR VARIOUS CASES

The next step of the method is to determine the multiplicities for the cases which we want to consider. For this purpose we list again our cases. Thereby we use for the equivalence classes κ of the projective unirreps of $P^{\vec{q}}$ the same symbols as in Ref. 5.

Case A.2: $\kappa_0 = (\mu = 4)\uparrow \mathcal{O}_h$ (six-dimensional unirrep).

Because of (III.19) of Ref. 3 and (I.9) we obtain for the corresponding multiplicity,

$$m_{(0,\vec{q})(0,\vec{q}');(\kappa_0,\vec{q}_R)} = 6. \tag{II.1}$$

Therefore, we use once again the symbol 0 for the identity representation of $P^{\vec{q}}$, respectively, $P^{\vec{q}'}$. The reason for choosing just $\kappa_0 = (\mu = 4)\uparrow \mathcal{O}_h$ is that this is the most complicated case.

Case B.2: $\kappa_0 = (0)\uparrow P^{s}$ (two-dimensional unirrep).

According to (III.45) of Ref. 3, (I.11), and (I.15) we obtain for

$$m_{(0,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_s)} = 2 \quad (\kappa' = \text{fixed}), \tag{II.2}$$

since (I.11) has as consequence that the two possible inequivalent vector unirreps of $P^{\vec{q}'}$ must be one-dimensional. That the unirreps of $P^{\vec{q}'}$ are ordinary vector representations is due to the fact[5] that the factor system $S^{\vec{q}}$ reduces to one for every $\vec{q}$ which does not lie on the surface of ΔBZ. This is a fact which shall be used often for the following cases.

Case C.2.a: $\kappa_0 \in A_{P^{\vec{q}_0}}$ (one-dimensional unirrep)

For this case we have to apply (III.85) of Ref. 3 and (I.15),

$$m_{(\kappa,\vec{q})(\kappa',\vec{q}');(\kappa_0,\vec{q}_0)} = n_\kappa n_{\kappa'} n_{\kappa_0} = n_{\kappa'}. \tag{II.3}$$

Thereby we have already used $n_\kappa = n_{\kappa_0} = 1$. Furthermore it is well known that for $P^{\Lambda} = C_{3v}$ there exist two inequivalent one-dimensional and one two-dimensional vector unirreps.

Taking $n_{\kappa_0} = 1$ into account, the determination of the multiplicity index is due to (II.62) and (II.64) of Ref. 1 trivially.

Case C.2: At the beginning we list a complete set of unirreps of P^Δ and P^x which have been calculated in Ref. 5. For the sake of brevity we write down these sets for appropriately chosen sets of generating elements.

The vector unirreps of $P^\Delta = C_{4v}$ for the set $\{C_{4y}^+, \sigma_x\}$ of generating elements are given by [see (5.25)–(5.27) of Ref. 5]

$$R^\mu:\mu = 0,1 \quad C_{4y}^+ \to 1, \quad \sigma_x \to (-1)^\mu, \qquad \text{(II.4)}$$

$$\mu = 2,3 \quad C_{4y}^+ \to -1, \quad \sigma_x \to (-1)^\mu, \qquad \text{(II.5)}$$

$$R^\mu:\mu = 5 \quad C_{4y}^+ \to \begin{vmatrix} -i & 0 \\ 0 & i \end{vmatrix}, \quad \sigma_x \to \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}. \qquad \text{(II.6)}$$

The occurence of vector unirreps for P^Δ is due to the fact that the factor system S^Δ reduces to one, since $\vec{q} = \vec{q}_\Delta$ does not lie on the surface of ΔBZ.

Because of the general theory of space group representations we have to consider for $\vec{q}_x$ projective representations. These projective unirreps written down for the set $\{C_{4y}^+, \sigma_x, I\}$ have the form [see (5.33), (5.35) and (5.36) of Ref. 5]

$\mathbb{R}^{(\kappa,\mu)\uparrow P^x}:\kappa = 0,1, \quad \mu = 5:$

$$C_{4y}^+ \to \begin{vmatrix} -i & 0 \\ 0 & i \end{vmatrix}, \sigma_x \to \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad I \to (-1)^\kappa \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix}, \qquad \text{(II.7)}$$

$\mathbb{R}^{(\mu = 0)\uparrow P^x}$:

$$C_{4y}^+ \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad \sigma_x \to \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad I \to \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}, \qquad \text{(II.8)}$$

$\mathbb{R}^{(\mu = 2)\uparrow P^x}$:

$$C_{4y}^+ \to \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix}, \quad \sigma_x \to \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix},$$

$$I \to \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}. \qquad \text{(II.9)}$$

For the following we consider the case

$$\kappa = \kappa' = \mu = 5 \quad (\in A_{P^\Delta}). \qquad \text{(II.10)}$$

Hence due to (II.8) of Ref. 3,

$$\dim\{D^{(\mu = 5,\vec{q}_\Delta)\uparrow G} \otimes D^{(\mu = 5,\vec{q}'_\Delta)\uparrow G}\}$$
$$= |P:P^\Delta| n_5 |P:P^\Delta| n_5 = 144$$
$$= \sum_{\substack{\vec{q}_0 \in \Delta BZ \\ \kappa_0 \in A_{P^{q_0}(S^{q_0})}}} m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\kappa_0,\vec{q}_0)} |P:P^{\vec{q}_0}| n_{\kappa_0} \qquad \text{(II.11)}$$

must be always satisfied otherwise the multiplicities would be incorrect. Since there occur three cases (i)–(iii) for $\vec{q}_0$, we have to calculate the corresponding multiplicities.

Case (i): In this case we have to inspect the general formula (III.83) of Ref. 3. We obtain with the aid of (I.12) of Ref. 3, (I.1), and (I.28),

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\kappa_0,\vec{q}_x)} = \frac{1}{8} \sum_{x \in P^\Delta} X^5(x) X^5(x) \mathbb{X}^{\kappa_0 *}(x). \qquad \text{(II.12)}$$

The characters of D^5 of P^Δ are easily obtained by using (II.6) and the multiplication law of P^Δ:

$$C_{4y}^+ C_{4y}^+ = C_{2y}, \quad C_{2y}\sigma_x = \sigma_z,$$
$$(C_{4y}^+)^3 = C_{4y}^-, \quad C_{4y}^+ \sigma_x = \sigma_{dc}, \qquad \text{(II.13)}$$
$$C_{4y}^- \sigma_x = \sigma_{de},$$

$$X^5(E) = -X^5(C_{2y}) = 2, \quad X^5(x) = 0,$$
otherwise. (II.14)

Together with the characters of (II.7–9), formula (II.12) yields

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\kappa,\mu = 5)\uparrow P^x,\vec{q}_x)} = 0, \quad \kappa = 0,1, \qquad \text{(II.15)}$$

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu = 0)\uparrow P^x,\vec{q}_x)} = 2, \qquad \text{(II.16)}$$

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu = 2)\uparrow P^x,\vec{q}_x)} = 2. \qquad \text{(II.17)}$$

Case (ii): Due to (III.87) of Ref. 3 we obtain by means of (I.12) of Ref. 3, (I.1), (I.32), and $I \times I = x$ for the multiplicity,

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\kappa_0,\vec{q}''_\Delta)} = \frac{1}{8} \sum_{x \in P^\Delta} X^5(x) X^5(x) X^{\kappa_0 *}(x). \qquad \text{(II.18)}$$

Inserting the characters of (II.4–6) in (II.18) we arrive at the results

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(5,\vec{q}''_\Delta)} = 0, \qquad \text{(II.19)}$$

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\mu,\vec{q}''_\Delta)} = 1 \quad \text{for } \mu = 0,1,2,3. \qquad \text{(II.20)}$$

Case (iii): By using (I.35), (I.36), (I.1), and the characters of the two one-dimensional vector unirreps of $P^{\vec{q}_0}$,

$$R^\kappa:\kappa = 0,1 \quad \sigma_z \to (-1)^\kappa, \qquad \text{(II.21)}$$

formula (III.87) of Ref. 3 turns out to be

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\kappa,\vec{q}_0)} = \frac{1}{2} \sum_{x \in \{E,\sigma_z\}} X^5(x) X^5(x) X^{\kappa *}(x), \qquad \text{(II.22)}$$

where also $C_{4z}^+ \sigma_z C_{4z}^- = \sigma_z$ has been taken into account. Therefore,

$$m_{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\kappa,\vec{q}_0)} = 2 \quad \text{for } \kappa = 0,1. \qquad \text{(II.23)}$$

Now it is easy to prove the correctness of (II.11).

III. MULTIPLICITY INDEX FOR VARIOUS CASES

As already pointed out the key problem is to solve the multiplicity problem when calculating CG coefficients. In case we can fix the multiplicity index w by means of the proposed method, the corresponding CG coefficients are immediately obtainable with the aid of our general formulas (I.8) and (I.9) of Ref. 1. Therefore, we focus our attention to the determination of the multiplicity index.

Case A.2: Due to (II.11) of Ref. 1 and (II.1) we have to solve the problem whether we can find six group elements $x_i \in \mathcal{O}_h$ $(= P^{\vec{q}_0})$ such that

$$\mathbb{R}_{11}^{(\mu = 4)\uparrow \mathcal{O}_h}(x_i x_j^{-1}) = \delta_{ij}, \quad i,j = 1,2,\ldots,6 \qquad \text{(III.1)}$$

is satisfied, where we have chosen $a = 1$. A simple inspection of the six-dimensional projective unirrep of $\mathcal{O}_h$ [see (5.23) of Ref. 5] shows that the following six group elements satisfy (III.1):

$$x_1 = E, \quad x_4 = C_{31}^{+} I,$$
$$x_2 = C_{31}^{+}, \quad x_5 = C_{2d} C_{31}^{+}, \qquad \text{(III.2)}$$
$$x_3 = I, \quad x_6 = C_{2d} C_{31}^{+} I.$$

Therefore, we can choose as multiplicity index

$$w = (x_j; x_j \sigma_{df}), \quad j = 1,2,...,6. \qquad \text{(III.3)}$$

Case B.2: In this case [because of (II.32) of Ref. 1] we are confronted with the problem of whether we can find two group elements $x_i \in P^s$ such that

$$\mathbb{R}_{11}^{(0)\uparrow P^s}(x_i x_j^{-1}) = \delta_{ij}, \quad i,j = 1,2 \qquad \text{(III.4)}$$

is satisfied. In Ref. 5 two equivalent projective unirreps of P^s have been calculated. For (5.57) of Ref. 5 we can choose

$$x_1 = E, \quad x_2 = \sigma_y, \qquad \text{(III.5)}$$

and for (5.62) of Ref. 5

$$x_1 = E, \quad x_2 = \sigma_{de} \qquad \text{(III.6)}$$

as group elements of P^s satisfying (III.4). This example shows as already pointed out, that the determination of the multiplicity index (and therefore of the corresponding CG coefficients) always depends on the explicit form of the space group unirreps. The multiplicity index turns out to be

$$w = (x_j; x_j \sigma_{de}), \quad j = 1,2, \qquad \text{(III.7)}$$

where the elements x_j have to be taken either from (III.5) or from (III.6) depending what projective unirreps for P^s is used.

Case C.2.a: This case is trivial because of (II.62) and (II.64) of Ref. 1. Consequently,

$$w = (E;E,c'), \quad c' = 1,2,...,n_{\kappa'}, \qquad \text{(III.8)}$$

where we have suppressed the superfluous column index 1 of the one-dimensional unirrep R^{κ}.

Case C.2.(i): As already pointed out this case belongs to the most complicated kind which may occur, when calculating CG coefficients. Because of (II.15)–(II.17), we have to consider only vectors of the kind (II.52) of Ref. 1 which belong to the unirreps $(\mu = 0)\uparrow P^x$ and $(\mu = 2)\uparrow P^x$ of P^x. For both cases we must find two linear independent vectors. According to the proposed procedure we have to calculate (II.54) of Ref. 1 for the corresponding vectors as first step. Thereby we have to take (I.28) and (I.12) of Ref. 3, (I.1), and (II.6)–(II.9) into account.

We start with the case $(\mu = 0)\uparrow P^x$, which decomposes into two subproblems according to [compare (II.54) of Ref. 1]

$$P^{\vec{q}_\Delta} : P^{\vec{q},\vec{q}';\vec{q}_x}_{\sigma,\sigma'} = P^x : P^\Delta = \{E, I\}, \qquad \text{(III.9)}$$

$$\|\vec{\mathbf{B}}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=0)\uparrow P^x,\vec{q}_x)(E,c;E,c')}\|^2$$
$$= \frac{1}{4} \sum_{x \in P^\Delta} R^{5}_{cc}(x) R^{5}_{c'c'}(x) \mathbb{R}_{11}^{(\mu=0)\uparrow P^x}(x)^*$$

$$= \begin{cases} 0, & c = c' = 1, \\ 1, & c = 1, c' = 2, \\ 1, & c = 2, c' = 1, \\ 0, & c = c' = 2. \end{cases} \qquad \text{(III.10)}$$

Inspecting the scalar product (II.55) of Ref. 1 for the vectors which belong to $c = 1$, $c' = 2$, and $c = 2$, $c' = 1$, we obtain

$$\left\{\begin{matrix} (5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ E,2 & E,1 \end{matrix} \middle| \begin{matrix} ((\mu=0)\uparrow P^x,\vec{q}_x) & (E,1;E,2) \\ E,1 & \end{matrix}\right\}$$
$$= \frac{1}{4} \sum_{x \in P^\Delta} R^{5}_{21}(x) R^{5}_{12}(x) \mathbb{R}_{11}^{(\mu=0)\uparrow P^x}(x)^* = 1, \qquad \text{(III.11)}$$

which implies

$$\vec{\mathbf{B}}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=0)\uparrow P^x,\vec{q}_x)(E,1;E,2)}$$
$$= \vec{\mathbf{B}}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=0)\uparrow P^x,\vec{q}_x)(E,2;E,1)}; \qquad \text{(III.12)}$$

that both vectors are identical. Analogously [in accordance with the second possibility of (III.9)] the corresponding expressions turn out to be

$$\|\vec{\mathbf{B}}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=0)\uparrow P^x,\vec{q}_x)(I,c;I,c')}\|^2 = \begin{cases} 0, & c = c' = 1, \\ 1, & c = 1, c' = 2, \\ 1, & c = 2, c' = 1, \\ 0, & c = c' = 2, \end{cases} \qquad \text{(III.13)}$$

$$\left\{\begin{matrix} (5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ I,2 & I,1 \end{matrix} \middle| \begin{matrix} ((\mu=0)\uparrow P^x,\vec{q}_x) & (I,1;I,2) \\ E,1 & \end{matrix}\right\} = 1, \qquad \text{(III.14)}$$

$$\vec{\mathbf{B}}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=0)\uparrow P^x,\vec{q}_x)(I,1;I,2)}$$
$$= \vec{\mathbf{B}}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=0)\uparrow P^x,\vec{q}_x)(I,2;I,1)}. \qquad \text{(III.15)}$$

Now there remains the problem whether the vectors (III.12) and (III.15) are orthogonal or not. Inspecting (II.55) of Ref. 1 we obtain for

$$\left\{\begin{matrix} (5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ I,1 & I,2 \end{matrix} \middle| \begin{matrix} ((\mu=0)\uparrow P^x,\vec{q}_x) & (E,1;E,2) \\ E,1 & \end{matrix}\right\}$$
$$= \frac{1}{4} \sum_{x \in P^\Delta} B^{\vec{q}_x *}_{E,E}(Ix) B^{\vec{q}_\Delta}_{I,E}(Ix) B^{\vec{q}'_\Delta}_{I,E}(Ix) R^{5}_{11}(x) R^{5}_{22}(x)$$
$$\mathbb{R}_{11}^{(\mu=0)\uparrow P^x}(Ix)^*$$
$$= 0,$$

since

$$\mathbb{R}_{11}^{(\mu=0)\uparrow P^x}(Ix) = 0, \quad \text{for } x = E, C_{4y}^{\pm}, C_{2y}. \qquad \text{(III.17)}$$

Hence we can choose the special column indices

$$w = (E,1;E,2) \quad \text{and} \quad (I,1;I,2) \qquad \text{(III.18)}$$

as the multiplicity index w. Because of (III.10) and (III.13) the vectors (III.12) and (III.15) are already a part of the columns of the CG matrix. The remaining CG coefficients can be readily calculated by means of (I.9) of Ref. 1.

For the case $(\mu = 2)\uparrow P^x$ we proceed in completely the same way,

$$\|\vec{\mathbf{B}}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=2)\uparrow P^x,\vec{q}_x)(E,c;E,c')}\|^2 = \delta_{cc'}, \quad c,c' = 1,2. \qquad \text{(III.19)}$$

As in the previous case we obtain for the scalar product of the two vectors ($c = c' = 1$ and $c = c' = 2$) one

$$\left\{\begin{matrix}(5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ E,1 & E,1\end{matrix}\middle|\begin{matrix}((\mu=2)\uparrow P^x,\vec{q}_x) & (E,2;E,2) \\ E,1 & \end{matrix}\right\} = 1. \qquad \text{(III.20)}$$

Therefore,

$$\vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=2)\uparrow P^x,\vec{q}_x)(E,1;E,1)} = \vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=2)\uparrow P^x,\vec{q}_x)(E,2;E,2)}. \qquad \text{(III.21)}$$

Considering the second case we obtain, due to (III.9),

$$\|\vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=2)\uparrow P^x,\vec{q}_x)(I,c;I,c')}\|^2 = \delta_{cc'}, \quad c,c'=1,2, \qquad \text{(III.22)}$$

$$\left\{\begin{matrix}(5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ I,1 & I,1\end{matrix}\middle|\begin{matrix}((\mu=2)\uparrow P^x,\vec{q}_x) & (I,2;I,2) \\ E,1 & \end{matrix}\right\} = 1. \qquad \text{(III.23)}$$

$$\vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=2)\uparrow P^x,\vec{q}_x)(I,1;I,1)} = \vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);((\mu=2)\uparrow P^x,\vec{q}_x)(I,2;I,2)}. \qquad \text{(III.24)}$$

Because of the same reason as before [see (III.16) and (III.17)], the scalar product of the vectors (III.21) and (III.24) yields zero,

$$\left\{\begin{matrix}(5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ I,1 & I,1\end{matrix}\middle|\begin{matrix}((\mu=2)\uparrow P^x,\vec{q}_x) & (E,1;E,1) \\ E,1 & \end{matrix}\right\} = 0. \qquad \text{(III.25)}$$

This implies that the column indices

$$w = (E,1;E,1) \quad \text{and} \quad (I,1;I\,1) \qquad \text{(III.26)}$$

can be chosen as a multiplicity index and that because of (III.19) and (III.22) the components of the corresponding vectors (III.21) and (III.24) are already CG coefficients.

Case C.2.(ii): Due to (II.19) and (II.20) we can restrict our considerations to vectors of the type (II.68) of Ref. 1 which belong to the unirreps $\mu = 0,1,2,3$ of P^Δ. Since the corresponding multiplicities are one, we have to look only for vectors whose norm is different from zero. By means of (I.12) of Ref. 3, (I.29), (I.32), (I.1), and (II.4)–(II.6), Eq. (II.72) of Ref. 1 turns out to be (for $\kappa_0 = \mu = 0,1,2,3$ with $n_\mu = 1$)

$$\|\vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\mu,\vec{q}''_\Delta)(E,c,I,c')}\|^2 = \frac{1}{8}\sum_{x\in P^\Delta} R^5_{cc}(x) R^5_{c'c'}(x) R^{\mu *}_{11}(x)$$

$$= \begin{cases} \left.\begin{matrix} 0, & c=c'=1, \\ \frac{1}{2}, & c=1,c'=2, \\ \frac{1}{2}, & c=2,c'=1, \\ 0, & c=c'=2, \end{matrix}\right\} \mu = 0,1, & \text{(III.27)} \\ \left.\begin{matrix} \frac{1}{2}, & c=c'=1, \\ 0, & c=1,c'=2, \\ 0, & c=2,c'=1, \\ \frac{1}{2}, & c=c'=2, \end{matrix}\right\} \mu = 2,3. & \text{(III.28)} \end{cases}$$

(III.28)

Now there remains the task of showing that the corresponding vectors are linearly dependent.

$$\vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\mu,\vec{q}''_\Delta)(E,1;I,2)} = \gamma \vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\mu,\vec{q}''_\Delta)(E,2;I,1)}, \quad \mu = 0,1, \qquad \text{(III.29)}$$

$$\vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\mu,\vec{q}''_\Delta)(E,1;I,1)} = \gamma' \vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\mu,\vec{q}''_\Delta)(E,2;I,2)}, \quad \mu = 2,3. \qquad \text{(III.30)}$$

A simple inspection of the scalar product (II.73) of Ref. 1 yields

$$\left\{\begin{matrix}(5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ E,2 & I,1\end{matrix}\middle|\begin{matrix}(\mu,\vec{q}''_\Delta) & (E,1;I,2) \\ E,1 & \end{matrix}\right\} = \frac{(-1)^\mu}{2}, \quad \mu = 0,1, \qquad \text{(III.31)}$$

$$\left\{\begin{matrix}(5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ E,2 & I,2\end{matrix}\middle|\begin{matrix}(\mu,\vec{q}''_\Delta) & (E,1;I,1) \\ E,1 & \end{matrix}\right\} = \frac{(-1)^{\mu+1}}{2}, \quad \mu = 2,3, \qquad \text{(III.32)}$$

which has as a consequence that the constants γ and γ' must take the values

$$\gamma = (-1)^\mu \ \text{for } \mu = 0,1, \qquad \gamma' = (-1)^{\mu+1} \ \text{for } \mu = 2,3, \qquad \text{(III.33)}$$

which can also be proven by directly comparing the considered vectors. This implies that we can take the column indices

$$w = (E,1;I\,2) \quad \text{for } \mu = 0,1, \qquad \text{(III.34)}$$

$$w = (E,1;I,1) \quad \text{for } \mu = 2,3 \qquad \text{(III.35)}$$

as the multiplicity index. The CG coefficients itself must be calculated with the aid of (I.8) and (I.9) of Ref. 1. In this connection it should be noted that the same column index can be used as the multiplicity index for different cases.

Case C.2.(iii): Because of (II.23) both unirreps $D^{(\kappa,\vec{q}_0)\uparrow G}$ occur twice in the reducible representation $D^{(5,\vec{q}_\Delta)\uparrow G} \otimes D^{(5,\vec{q}'_\Delta)\uparrow G}$. Therefore, we have to consider vectors of the type (II.68) of Ref. 1 which belong to the unirreps R^κ, $\kappa = 0,1$, of $P^{\vec{q}_0} = \{E,\sigma_z\}$. When calculating (II.69) of Ref. 1 we have to take (I.12) of Ref. 3, (I.1), (I.36), (I.33), (II.6), (II.21), and $C^-_{4z}\sigma_z C^+_{4z} = \sigma_z$ into account

$$\|\vec{B}_{E,1}^{(5,\vec{q}_\Delta)(5,\vec{q}'_\Delta);(\kappa,\vec{q}_0)(E,c;C_{4z},c')}\|^2 = \tfrac{1}{2}, \quad \kappa = 0,1 \quad c,c' = 1,2. \qquad \text{(III.36)}$$

As our next step we investigate the scalar product (II.70) of Ref. 1,

$$\left\{\begin{matrix}(5,\vec{q}_\Delta) & (5,\vec{q}'_\Delta) \\ E,d & C^-_{4z},d'\end{matrix}\middle|\begin{matrix}(\kappa,\vec{q}_0) & (E,c;C^-_{4z},c') \\ E,1 & \end{matrix}\right\} = \frac{1}{2}\sum_{x\in\{E,\sigma_z\}} R^5_{dc}(x) R^5_{d'c'}(x) R^{\kappa *}_{11}(x)$$

$$= \tfrac{1}{2}(\delta_{dc}\delta_{d'c'} + (-1)^{\kappa}\delta_{d,c+1}\delta_{d',c'+1}),$$

$$\kappa = 0,1. \qquad \text{(III.37)}$$

This has, as consequences,

$$\vec{\mathrm{B}}_{E,1}^{(5,\vec{q}_{\Delta})(5,\vec{q}'_{\Delta});(\kappa,\vec{q}_0)(E,1;C_{4z}^{-},1)} = \vec{\mathrm{B}}_{E,1}^{(5,\vec{q}_{\Delta})(\kappa,\vec{q}'_{\Delta});(5,\vec{q}_0)(E,2;C_{4z}^{-},2)}, \quad \kappa = 0,1, \qquad \text{(III.38)}$$

$$\vec{\mathrm{B}}_{E,1}^{(5,\vec{q}_{\Delta})(5,\vec{q}'_{\Delta});(\kappa,\vec{q}_0)(E,1;C_{4z}^{-},2)} = (-1)^{\kappa}\vec{\mathrm{B}}_{E,1}^{(5,\vec{q}_{\Delta})(\kappa,\vec{q}'_{\Delta});(5,\vec{q}_0)(E,2;C_{4z}^{-},1)}, \quad \kappa = 0,1, \qquad \text{(III.39)}$$

so that we can choose the column indices

$$w = (E,1;C_{4z}^{-},1) \quad \text{and} \quad (E,1;C_{4z}^{-},2) \quad \text{for } \kappa = 0,1 \qquad \text{(III.40)}$$

for both cases ($\kappa = 0,1$) as multiplicity index, since the scalar product of the corresponding vectors is, due to (III.37), zero. In this connection it should be noted that equations like (III.38) or (III.39) can be seen conversly as a proof of the corresponding multiplicity formula. The corresponding CG coefficients are readily obtained by using (I.8) and (I.9) of Ref. 1.

CONCLUDING REMARKS

It was the aim of this paper to show examples of the usefulness of the present method. We have shown that even for the most complicated case the determination of the multiplicity index requires the inspection of simple equations. But in comparison with other methods[6,7] we have to note that apart from special cases [such as (III.8), (III.34), and (III.35), and (III.40)] there enter different pairs of left coset representatives $(\sigma_v, \sigma'_v) \in P(\vec{q},\vec{q}')$ [such as (III.3), (III.7), (III.18), and (III.26)] when fixing the multiplicity index in terms of special column indices of the corresponding Kronecker product. Thereby we have shown that just the elements of the group $P^{\vec{q}_0}$ play an essential role.

ACKNOWLEDGMENTS

I want to thank Professor Rhoda Berenson and Professor Joseph Birman for discussions and especially for a fruitful correspondence concerning the multiplicity problem.

[1]R. Dirl, "Clebsch–Gordan coefficients for space groups," **20**, 659 (1979).
[2]R. Dirl, "Clebsch–Gordan coefficients: General theory," **20**, 664 (1979).
[3]R. Dirl, "Multiplicities for space group representations," **20**, 671 (1979).
[4]C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[5]R. Dirl, J. Math. Phys. **18**, 2065 (1977).
[6]R. Berenson and J.L. Birman, J. Math. Phys. **16**, 227 (1975).
[7]R. Berenson, I. Itzkan, and J.L. Birman, J. Math. Phys. **16**, 236 (1975).

Resonances in one-dimensional Stark effect and continued fractions

S. Graffi, V. Grecchi,[a] S. Levoni, and M. Maioli

Istituto Matematico, Universitá di Modena, Modena, Italy [b]
(Received 12 June 1978)

The Stieltjes type continued fraction (i.e., any diagonal Padé approximants sequence) of the perturbation series for the resonances of the so-called one-dimensional Stark effect converges to the resonances.

I. INTRODUCTION

The existence of resonances has recently been proved[1,18] for the hydrogen Stark effect, i.e., for the quantum mechanical system described by the Schrödinger operator $-\Delta - Z/r + Fx$ acting on $L^2(R^3)$. Here Z denotes the atomic number, $F>0$ the uniform electric field strength, and, as usual $r=(x^2+y^2+z^2)^{1/2}$.

In addition, it has been proved[1] that the resonances are uniquely determined by (divergent) time-independent (i.e., Rayleigh–Schrödinger) perturbation theory through the Borel summation method. All these results are currently being generalized to any atomic system, as announced in Ref. 19.

In this paper we deal with the one-dimensional Stark effect, described by the Schrödinger equation,

$$\left(\frac{-d^2}{dx^2}+\frac{m^2-1}{4x^2}-\frac{Z}{x}-Fx\right)\psi=E(F)\psi,\qquad m=0,1,\cdots,\qquad (1.1)$$

acting on $L^2(R_+)$, where the existence of the resonances has long been known from the work of Titchmarsh.[2]

Equation (1.1) results from the separation of $-\Delta - Z/r + Fx$ in parabolic coordinates; here m is the magnetic quantum number. The equation could also be obtained by separating $-\Delta - Z/r + Fr$ in polar coordinates: In this case $m=2l+1$.

Equation (1.1) represents, of course, a very simplified version of the system, but will enable us to obtain, for the problem of determining the resonances starting from divergent perturbation theory, a result that is remarkably stronger than that of simple Borel summability: I.e., convergence to the resonances of the Stieltjes-type continued fractions associated with the formal Rayleigh–Schrödinger expansion. This convergence also yields a rigorous approximation statement, since it is well known[3] to be equivalent to the convergence of any diagonal sequence of Padé approximants on the formal perturbation series, this convergence being, in addition, monotonic for negative F.

As in the three-dimensional case, the whole treatment is based upon the close analogy of the problem under examination to the anharmonic oscillator. Actually, in this one-dimensional case, there is more than a close analogy between the two problems. As remarked by Titchmarsh,[2] under the transformation $x=v^2$, $\psi(x)=v^{1/2}\phi(v)$, Eq. (1.1) goes into

$$\left(-\frac{d^2}{dv^2}+\frac{m^2-\frac{1}{4}}{v^2}-4Ev^2-4Fv^4\right)\phi(v)=4Z\phi(v),\qquad (1.2)$$

which, apart from the centrifugal term, is $(m^2-\frac{1}{4})/v^2$, the Schrödinger equation of a quartic anharmonic oscillator.

However, we shall make use of the analogy only to obtain information about the energy spectrum at fixed charge by means of an inversion argument.

Instead, the key point of the present paper is the identification of the resonances as the analytic continuation to negative F of the eigenvalues of

$$\left(-\frac{d^2}{dx^2}+\frac{m^2-1}{4x^2}-\frac{Z}{x}+Fx\right)\psi=E(F)\psi,\qquad (1.3)$$

by which it is possible to extend Simon's results[4] on the anharmonic oscillator, as well as the Loeffel–Martin arguments[5] yielding analyticity of the eigenvalues for $|\arg(F)|<\pi$. This will yield a proof of the existence of the resonances within the framework of the by now standard Balslev–Combes[6] theory of analytic dilatations. The form (1.2) of the problem will then allow direct application of the results of Ref. 1 to obtain the upper bound required on the perturbation coefficients to prove the convergence of the continued fraction. The exposition proceeds as follows: In Sec. II we apply the spectral theory to Eq. (1.3) for complex F, and identify thereby the resonances within the framework of the Balslev–Combes theory. In Sec. III we adapt to the present case the Loeffel–Martin arguments and apply the results of Ref. 1 to obtain the convergence of the continued fraction. Finally, we briefly discuss the approximations to the position and width of the resonances obtained from the approximants of the continued fraction, i.e., from the Padé approximants.

II. SPECTRAL THEORY AND RESONANCES

Consider in $L^2(R_+)$ the formal differential expression,

$$H_m(Z,F)=-\frac{d^2}{dx^2}+\frac{m^2-1}{4x^2}-\frac{Z}{x}+Fx,\qquad m=0,1,\cdots,\qquad (2.1)$$

Z and F real. The one-dimensional Stark Hamiltonian is obtained from (2.1) for F negative. Hence, we rewrite the

[a]Partially supported by I.N.F.N., Sezione di Bologna. Also at Istituto di Fisica, Universitá di Modena, Modena, Italy.

[b]Postal address: Istituto Matematico, Universitá di Modena, Via Campi 181, 41100 Modena, Italy.

Schrödinger equation (1.1) as

$$H_m(Z, -F)\psi = E\psi. \tag{2.2}$$

As is well known,[7] the formal differential expression (2.1) is, for all real F, a singular Sturm–Liouville operator which is in the limit-point case for all m at infinity, and at the origin is in the limit-circle case for $m = 0,1$ and in the limit point case for $m \geqslant 2$. Hence, a boundary condition at $x = 0$ has to be imposed in order to realize (2.1) as a self-adjoint operator family in $L^2(R_+)$ for $m = 0,1$. A possible realization is given by[8] the following.

Definition 2.1: From now on, by $H_m(Z,F)$ we shall mean the operator family defined by the unique self-adjoint extension of the symmetric differential operator in $L^2(R_+)$ specified by (2.1) on the dense domain D

$$D = \left\{u \,|\, u \in L^2(R_+) \cap C^2(R_+) \,|\, u(0) = 0 \,|\left(\frac{-d^2}{dx^2} + \frac{m^2-1}{4x^2} - \frac{Z}{x} + Fx\right)u \in L^2(R_+)\right\}.$$

Then the results of Titchmarsh[2] on the spectral properties of $H_m(Z,F)$ and the existence of resonances can be summarized in the following way.

Theorem 2.1 (Titchmarsh): Let $Z > 0$ and $m = 0,1,\cdots$. Then (a) If $F > 0$, the spectrum of $H_m(Z,F)$, $\sigma(H_m)$, is simple, positive, and discrete, i.e., it consists of simple, isolated, and positive eigenvalues accumulating only at $+\infty$; (b) If $F < 0$, $\sigma(H_m)$ is absolutely continuous over $(-\infty, +\infty)$, and $H_m(Z,F)$ has an infinite set of resonances, in the following meaning. Consider the Green's function of $H_m(Z,F) - E$, i.e., the integral kernel of the resolvent $[H_m(Z,F) - E]^{-1}$. This is an analytic bounded-operator valued function of E for all E in the upper half plane $\mathrm{Im}(E) > 0$. Then it has a meromorphic continuation to the lower half plane $\mathrm{Im}(E) < 0$ across the cut along the real axis, with an infinite set of simple poles.

Let us proceed now to an examination of $\sigma(H_m)$ when F and Z are allowed to take complex values, $H_m(Z,F)$ being defined in a natural way as the closure of (2.1) on D (Definition 2.1) for Z and F complex as well.

Lemma 2.1: For any F belonging to the cut plane $|\arg(F)| < \pi$, $H_m(0,F)^{-1}$, $m = 0,1,\cdots$, is a holomorphic family of Hilbert–Schmidt operators.

Proof: By means of the standard Green's function technique,[8] one easily proves that $H_m(0,F)^{-1}$, $|\arg(F)| < \pi$ is the integral operator on $L^2(R_+)$ whose kernel is given by

$$G_m(x,y) = \begin{cases} f_\nu(x)g_\nu(y), & 0 \leqslant x \leqslant y + \infty, \\ f_\nu(y)g_\nu(x), & 0 \leqslant y \leqslant x + \infty, \end{cases} \tag{2.3}$$

where

$$\begin{aligned} f_\nu(x) &= (\tfrac{2}{3}x)^{1/2} I_\nu(\tfrac{2}{3}F^{1/2}x^{3/2}), \\ g_\nu(x) &= (\tfrac{2}{3}x)^{1/2} K_\nu(\tfrac{2}{3}F^{1/2}x^{3/2}), \end{aligned} \tag{2.4}$$

where $\nu = \frac{1}{3}m$, and I_ν and K_ν are Bessel functions of the third kind.

Taking into account the well-known asymptotic behavior of the Bessel functions,[9] a short computation yields

$$\int_0^\infty \int_0^\infty |G_m(x,y)|^2 dx\, dy < +\infty,$$

the convergence being uniform with respect to F on any compact of the cut plane. This proves Lemma 2.1.

Lemma 2.2: Let V_1 denote the maximal multiplication operator by $1/x$ in $L^2(R_+)$. Then if $m \geqslant 1$ V_1 is $H_m(0,F)$-compact, and if $m = 0$ V_1 is $H_m(0,F)$-form compact. In both cases the bounds are uniformly bounded on compacts of the cut F plane.

Proof: The integral kernel of $V_1 H_m(0,F)^{-1}$ is, of course, given by $(1/x)G_m(x,y)$, with $G_m(x,y)$ as in (2.4). Then, as above, an easy computation shows that

$$\int_0^\infty \int_0^\infty |x^{-1}G_m(x,y)|^2 dx\, dy < \infty, \quad m \geqslant 1,$$

the convergence being uniform on compacts of the cut F plane. Hence $V_1 H_m(0,F)^{-1}$ is Hilbert–Schmidt and therefore, *a fortiori*, compact. For $m = 0$, consider the operator $V^{1/2} \times H_m(0,F)^{-1} V^{1/2}$, whose integral kernel is $x^{-1/2} G_m(x,y) y^{-1/2}$. Again, this kernel is easily seen to be uniformly Hilbert–Schmidt on compacts in the cut F plane, and this means that V_1 is uniformly form compact with respect to $H_m(0,F)$. This proves Lemma 2.2.

Corollary 2.1: For any fixed F belonging to the cut plane $|\arg(F)| < \pi$, $H_m(Z,F)$ is a holomorphic family in Z, of type A if $m \geqslant 1$ and of type B if $m = 0$ (in the sense of Ref. 10, Chap. VII).

Proof: This is a direct consequence of the former Lemma 2.2, in view of standard results in perturbation theory (see Ref. 10, Chap. VII), since the relative compactness implies *a fortiori* the relative boundedness of $1/x$ with respect to $H_m(0,F)$ with $H_m(0,F)$-bound zero when $m \geqslant 1$, and the relative form compactness implies the relative form boundedness for $m = 0$.

Lemma 2.3: Let V_2 be the maximal multiplication operator by x in $L^2(R_+)$, and let F belong to the cut plane. Then, for any finite complex Z, V_2 is $H_m(Z,F)$-bounded when $m \geqslant 1$, uniformly on compacts in F, and is $H_m(Z,F)$-form bounded when $m = 0$, uniformly on compacts in F.

Proof: The integral kernel of $V_2 H_m(0,F)^{-1}$ is, of course, given by $xG_m(x,y)$, $G_m(x,y)$ as in (2.4). Now, an easy computation shows that there are constants M and N (independent of F, as long as it varies on compacts of the cut plane) such that

$$\int_0^\infty x\,|G_m(x,y)|\,dy < M, \quad \int_0^\infty x\,|G_m(x,y)|\,dy < N,$$

$$m = 0,1,\cdots.$$

Hence, by a known result (see, e.g., Ref. 10, Example III.2.4), V_2 is $H_m(0,F)$-bounded. The results thus follows by Problem IV.1.2 of Ref. 10, since by Lemma 2.2 Z/x is $H_m(0,F)$-bounded [$H_m(0,F)$-form bounded for $m = 0$] with relative bound zero.

Therefore, we have the following corollary.

Corollary 2.2: For any fixed complex Z, $H_m(Z,F)$ is a holomorphic family of type A when $m \geqslant 1$, and of type B when $m = 0$, with compact resolvents, as long as F belongs to the cut plane $|\arg(F)| < \pi$.

A further analogy with the anharmonic oscillator which is also basic to the present treatment is represented by the Symanzik scaling.

Theorem 2.2 (Symanzik): Let $F > 0$, $\lambda > 0$, Z real. If $E_i^m(Z,F)$, $m,i = 0,1,\cdots$, stands for the ith eigenvalue of $H_m(Z,F)$, one has

$$E_i^m(Z,F) = \lambda E_i^m(Z\lambda^{-1/2},F\lambda^{-3/2}). \tag{2.5}$$

Proof: Exactly as in Ref. 4, Theorem II.2.1.

By the same remark after Theorem II.2.1 of Ref. 4, (2.5) will hold by analytic continuation for all complex Z and F for which the continuation is possible.

We are now in a position to give a proof of the existence of the resonances within the framework of the Balslev–Combes theory[6] of analytic dilatations. As pointed out by Simon,[11] this allows the synthesis between the two standard concepts of resonance, i.e., those of Lifsic–Grossmann and Friedrichs–Howland.

Let $U(\theta)$, $\theta \in R$, be the operation of unitary dilatation in $L^2(R_+)$, defined as

$$[U(\theta)f](x) = e^{\theta/2} f(e^\theta x), \quad f(x) \in L^2(R_+).$$

Set

$$H_m(Z,F,\theta) = U(\theta) H_m(Z,F) U(\theta)^{-1}. \tag{2.6}$$

It is easily seen that $H_m(Z,F,\theta) \restriction D$ is given by the differential operator

$$\begin{aligned} H_m(Z,F,\theta) \restriction D &= -e^{-2\theta}\frac{d^2}{dx^2} - e^{-\theta}\frac{Z}{x} + e^{-2\theta}\frac{m^2-1}{x^2} + Fe^\theta x \\ &= e^{-2\theta} H_m(Ze^\theta, Fe^{3\theta}), \end{aligned} \tag{2.7}$$

so that the definition of $H_m(F,Z,\theta)$ extends to complex θ.

Remark: The operation of (unitary) dilatation is nothing else than a Symanzik scaling, with λ replaced by $e^{-2\theta}$. In particular, the eigenvalues of $H_m(Z,F,\theta)$ do not also depend on θ for θ complex (if the analytic continuation to complex θ is possible).

Let us also recall that a dilatation analytic vector is any vector $\Phi \in L^2(R_+)$ such that $U(\theta)\Phi$ can be analytically continued to some strip $|\mathrm{Im}(\theta)| < \alpha$, $\alpha > 0$. For any fixed α, the set N_α of all dilatation analytic vectors is dense.

Now let us prove the existence of the resonances along the lines of Ref. 1, Sec. IV, with some minor changes due to the greater simplicity of the problem.

Lemma 2.4: Let F be fixed in the cut plane. Then the operator $H_m(Z,F,\theta)$ has an analytic continuation to the strip $-\arg(F)/3 < \mathrm{Im}(\theta) < \pi/3 - \arg(F)/3$; i.e., for these values of θ $[H_m(Z,F,\theta) - E]^{-1}$ is a holomorphic operator family in θ with compact resolvents.

Proof: Putting $F' = Fe^{3\theta}$, $Z' = Ze^{-\theta}$, one has $|\arg(F')| < \pi$ when θ belongs to the above strip. Hence we can apply Corollary 2.2 and the lemma is proved.

Theorem 2.3: Let $\Psi \in L^2(R_+)$ be a dilatation analytic vector in some class N_α, $\alpha > \pi/2$. Then the function $f_\Psi(E) = (\Psi [H_m(Z,F) - E]^{-1}\Psi)$, F real and negative, $Z > 0$, originally defined as an analytic function of E in the upper half-plane $\mathrm{Im}(E) > 0$, has a meromorphic continuation to the lower half-plane $\mathrm{Im}(E) < 0$ across the cut at E real. The poles of $f_\Psi(E)$ for $\mathrm{Im}(E) < 0$ are the eigenvalues of $H_m(Z,F,\theta)$, $0 < \mathrm{Im}(\theta) < \pi/3$.

Remark: By F negative we mean $e^{-i\pi}F$, $F > 0$, i.e., the lower edge of the cut. This choice is in agreement with the convention of defining the half-plane $\mathrm{Im}(E) > 0$ as the physical sheet of the energy.

If we had chosen $-F = e^{i\pi}F$, the specular statement would be true, i.e., analyticity for $\mathrm{Im}(E) < 0$, and resonances for $\mathrm{Im}(E) > 0$.

Proof: For F in the cut plane, the function $f_\psi(E)$ exists and is analytic in E as long as E is different from all eigenvalues of $H_m(Z,F)$. For $\theta \in R$ and $\Psi(\theta) = U(\theta)\Psi$, we have

$$f_\Psi(E) = (\Psi(\theta),[H_m(Z,F,\theta) - E]^{-1}\Psi(\theta)) \tag{2.8}$$

by the unitary equivalence between $H_m(Z,F)$ and $H_m(Z,F,\theta)$. Since ψ is a dilatation analytic vector, by Lemma 2.4 Eq. (2.8) holds, also by analytic continuation, for all θ in the strip $-\frac{1}{3}\arg(F) < \mathrm{Im}(\theta) < \pi/3 - \frac{1}{3}\arg(F)$.

Now restricting θ to the strip $0 < \mathrm{Im}(\theta) < \pi/3 - \epsilon$, $\epsilon > 0$, for any fixed value of θ we can perform the analytic continuation of the rhs of (2.8), and hence of $f_\Psi(E)$, to F real and negative, for all E different from the eigenvalues of $H_m(Z,F,\theta)$. It can be easily seen, by direct comparison of the Green functions, that this analytic continuation coincides with $(\Psi,(H_m(Z,-F) - E)^{-1}\Psi)$ for $\mathrm{Im}(E) > 0$.

To conclude the proof it remains to check analyticity for $\mathrm{Im}(E) > 0$. Since the only singularities of $f_\Psi(E)$, $|\arg(F)| < \pi$, are the eigenvalues $E_i^m(F)$ of $H_m(Z,F,\theta)$, i.e., of $H_m(Z,F)$, it will be enough to prove $\mathrm{Im}E_i^m(F) < 0$ as $\arg(F) = -\pi + 0$. Now if ψ_i^m is an eigenvector corresponding to $E_i^m(F)$, one has (Herglotz property; see Lemma 3.1)

$$\mathrm{Im}E_i^m(F) = (\psi_i^m, x\psi_i^m)\cdot\mathrm{Im}(F) < 0.$$

Hence $\lim_{\arg F \to -\pi}[\mathrm{Im}E_i^m(F)]$, which under the present conditions exists at least as a measure by a theorem of Herglotz, cannot be positive. Moreover, it cannot be zero because by Theorem 2.1(a) there are no real eigenvalues. Hence the theorem is proved.

Remarks: (a) The second sheet poles of the above theorem are thus resonances of $H_m(Z,-F)$ according to the Friedrichs–Howland notion.[11]

(b) For an alternative proof, based on strong resolvent convergence as $\theta \to 0$, see Ref. 18.

(c) These second-sheet poles are resonances, too, according to the Lifsic–Grossmann notion. (See again Ref. 11.) Indeed, they coincide with the eigenvalues of $H_m(Z,F,\theta)$, which do not depend on θ as long as $0 < \mathrm{Im}(\theta) < \pi/3 - \epsilon$, $\epsilon > 0$.

We have thus found that the resonances in this case are nothing but the object to which, when F gets negative, the eigenvalues of the operator family $H_m(Z,F)$ are turning, eigenvalues which exist as long as $|\arg(F)| < \pi$.

III. STIELTJES SUMMABILITY TO THE RESONANCES

In this section we proceed to verify that the eigenvalues of $H_m(Z,F)$ turn into resonances of $H_m(Z,-F)$ by analytic continuation. (This is exactly Howland's mechanism[12] for the onset of resonances.) Combining this with the Herglotz property and with an estimate on the perturbation coefficients, we will get the Stieltjes summability.

Let us begin by stating the Herglotz property, already used in Theorem 2.3, under the form of a lemma. From now on, let us denote by $E(F)$ any eigenvalues of $H_m(1,F)$, and by $E'(Z)$ any eigenvalue of $H_m(Z,1)$. By Symanzik scaling, they are related by

$$E'(Z) = F^{-2/3}E(F), \quad F = Z^{-3},$$
$$E(F) = Z^{-2}E'(Z), \quad Z = F^{-1/3}. \tag{3.1}$$

Lemma 3.1: When F is complex, $|\arg F| < \pi$, $E(F)$ enjoys the Herglotz property $\mathrm{Im}E(F)/\mathrm{Im}(F) > 0$. The same is true for $E'(-Z)$ when $-Z$ is complex, $|\arg(-Z)| < \pi$.

Proof: If ψ is an eigenvalue for $E(F)$, one has

$$\mathrm{Im}E(F)/\mathrm{Im}(F) = (\psi, x\psi)/(\psi,\psi) > 0.$$

$$\mathrm{Im}E'(-Z)/\mathrm{Im}(-Z) = (\psi, x^{-1}\psi)/(\psi,\psi) > 0, \tag{3.2}$$

and this proves the lemma.

To prove analyticity of $E(F)$ in the whole first sheet $|\arg(F)| < \pi$ we adapt to this case the arguments of Loeffel and Martin[5] and of Loeffel *et al.*,[13] who proved the analogous statement for the anharmonic oscillator $p^2 + x^2 + Fx^4$.

Let us begin by listing those steps of Ref. 5 which can be taken over to the present case without modifications.

(a) It is enough to prove analyticity of $E'(Z)$ for $|\arg Z| < \pi/3$: by (3.1), this is equivalent to analyticity of $E(F)$ for $|\arg F| < \pi$.

(b) As a consequence of the Herglotz property, $E'(Z)$ has no pole nor isolated essential singularities for $|\arg Z| < \pi/3$.

(c) Any eigenfunction $\psi(x)$ corresponding to $E'(Z)$ is an entire function of F and Z, and an analytic function of x in the whole sector $|\arg x| < \pi$. $\psi(x)$ has the asymptotic behavior $\psi(x) \sim \exp(-x^{3/2})$ as $|x| \to +\infty$ within the sector $|\arg x| < \pi/3$. (These properties are proved in Ref. 2.)

Hence the only argument to be adapted concerns the number of zeros of the wavefunction. Therefore, let us prove the following lemma.

Lemma 3.2: Let $\psi(x)$ be an eigenfunction corresponding to an arbitrary eigenvalue $E'(Z)$. Then the number of zeros of $\psi(x)$ is constant as Z ranges within the sector $|\arg Z| < \pi/3$.

Proof: Following Loeffel and Martin[5] we divide the proof in two steps.

Step 1: Let Z be real. Then the zeros of $\psi(x)$ for $|\arg x| < \pi/3$ lie on the real axis. Indeed, along the ray $x = re^{i\phi}$, $0 \leqslant \phi < \pi/3$, the Schrödinger equation $H_m(Z,1)\psi = E'(Z)\psi$ becomes

$$\left(-\frac{d^2}{dr^2} + re^{3i\phi} + \frac{m^2-1}{4r^2} - \frac{Ze^{i\phi}}{r} - E'e^{2i\phi}\right)\psi = 0. \tag{3.3}$$

After multiplication by $\bar\psi$ and integration between 0 and R, $0 < R < +\infty$, we get [by means of a partial integration taking into account that $\psi(0) = 0$]

$$\left(\bar\psi\frac{d\psi}{dr}\right)_{r=R} = \int_0^R \left\{\left|\frac{d\psi}{dr}\right|^2 + \left(re^{3i\phi} + \frac{m^2-1}{4r^2} - \frac{Ze^{i\phi}}{r} - E'e^{2i\phi}\right)|\psi|^2\right\} dr. \tag{3.4}$$

By the Herglotz property $E'(Z)$ is real for Z real. Hence the imaginary part of (3.4) yields

$$\mathrm{Im}\left(\bar\psi\frac{d\psi}{dr}\right)_{r=R} = \int_0^R Q(r,\phi)r^{-1}|\psi|^2 dr, \tag{3.5}$$

where

$$Q(r,\phi) = r^2\sin 3\phi - E' r\sin 2\phi - 2\sin\phi. \tag{3.6}$$

Since $\psi \to 0$ as $|x| \to +\infty$, in an analogous way one has

$$-\mathrm{Im}\left(\bar\psi\frac{d\psi}{dr}\right)_{r=R} = \int_R^{+\infty} Q(r,\phi)r^{-1}|\psi|^2 dr. \tag{3.7}$$

The quadratic form $Q(r,\phi)$ with $0 < \phi < \pi/3$ is positive as $r \to \infty$ and negative at $r = 0$ because $Z > 0$. Hence, it vanishes only once at a certain value r_0. Let R be fixed in $(0,\infty)$. If $R < r_0$ we take (3.5), and if $R > r_0$ we take (3.7); in both cases we can conclude that

$$\mathrm{Im}\left(\bar\psi\frac{d\psi}{dr}\right)_{r=R} \neq 0, \tag{3.8}$$

since in both cases one has the integral of a quantity of constant sign. [If $R = r_0$, we can take indifferently (3.5) or (3.7), or both.]

By (3.8), $\psi(x)$ never vanishes in the sector $0 < r < +\infty$, $0 < \phi < \pi/3$. The same conclusion holds for $-\pi/3 < \phi < 0$, since the quadratic form $Q(r,\phi)$ is now negative for $r \to \infty$ and positive at $r = 0$, and the above argument can be repeated.

Hence, the only zeros of $\psi(x)$ in the sector $|\arg\phi| < \pi/3$ lie on the real axis; i.e., they are its nodes as an eigenfunction.

Step 2: Let Z belong to the sector $|\arg Z| < \pi/3$. Then the number of zeros of $\psi(x)$ in $|\arg x| < \pi/3$ does not depend on Z.

To see this, consider again the imaginary part of (3.4).

In this case (3.5) and (3.7) become

$$\text{Im}\left(\bar{\psi}\frac{d\psi}{dr}\right)_{r=R} = \int_0^R Q(r,\phi,E')r^{-1}|\psi|^2\,dr, \tag{3.9}$$

$$-\text{Im}\left(\bar{\psi}\frac{d\psi}{dr}\right)_{r=R} = \int_R^\infty Q(r,\phi,E')r^{-1}|\psi|^2\,dr, \tag{3.10}$$

where now

$$Q(r,\phi,E') = r^2\sin 3\phi - |Z|\sin(\phi + \arg Z) - |E'|r\sin(2\phi + \arg E'). \tag{3.11}$$

Let $0 < \arg Z < \pi/3$, which yields $\text{Im}Z > 0$. One has

$$\sin(\phi + \arg Z) > 0, \quad \text{for } \pi/3 - \epsilon < \phi < \pi/3,$$

$$\sin(\phi + \arg Z) < 0, \quad \text{for } -\pi/3 < \phi < -\pi/3 + \epsilon.$$

Hence the quadratic form $Q(r,\phi,E')$ is positive for $r \to \infty$ and negative at $r = 0$, when $\pi/3 - \epsilon < \phi < \pi/3$, and conversely for $-\pi/3 < \phi < -\pi/3 + \epsilon$.

In both cases it vanishes only once in $(0,\infty)$ and hence, as in Step 1, ψ does not vanish.

We can thus conclude that for $|\arg Z| < \pi/3$ there are two angular sectors, $-\pi/3 < \phi < -\pi/3 + \epsilon$ and $\pi/3 - \epsilon < \phi < \pi/3$ (the "Loeffel–Martin walls" as called by Simon[14]) in which ψ does not vanish.

Hence, by exactly the same argument of Loeffel and Martin (no zero can penetrate the walls, and no zero can come from infinity because of the asymptotic behavior of the wavefunction; see Ref. 5 for details) the lemma is proved.

As a direct consequence, we have the following theorem.

Theorem 3.1: Any eigenvalue $E(F)$ of $H_m(Z,F)$ $(Z > 0$ fixed) is analytic for $|\arg(F)| < \pi$.

Proof: As in Ref. 5, we have only to check that $E'(Z)$ remains bounded as long as Z ranges in $|\arg Z| < \pi/3$. The only modification in the Loeffel–Martin argument is that the Volterra integral form of the Schrödinger equation is now given by

$$\psi'(x) = \sqrt{x}J_{-\nu}(\sqrt{E'}x) + \int_0^x J_\nu(\sqrt{E'}(x - x')) \times (x' + Z/x')\psi'(x')dx', \tag{3.12}$$

$$\nu = \tfrac{1}{3}m, \quad m = 0,1,\dots .$$

By means of (3.12), one easily shows that for $|x| < R$, $R > 0$ arbitrary, one has

$$[\psi' - \sqrt{x}J_{-\nu}(\sqrt{E'}x)]/e^{\text{Im}\sqrt{E'}x} \to 0 \quad \text{as } E' \to \infty.$$

At this point we can repeat word by word the arguments of Ref. 5, since the results of Sec. III therein showing that $E' \to \infty$ only in the direction $\arg E' = 0$ can be trivially taken over to the present case.

Hence $E'(Z)$ is analytic for $|\arg Z| < \pi/3$, and, by scaling, the theorem is proved.

As a consequence of this theorem, the resonances of the one-dimensional Stark operator are thus identified as the boundary values of the analytic functions $E_i^m(F)$, the eigenvalues of $H_m(Z,F)$ as $|\arg(F)| < \pi$, at the lower edge of the cut for F negative.

This is a partial realization of Howland's[12] notion; a full realization of it can be obtained, with restriction to a small field, through the same implicit function arguments of Ref. 1 applied to the anharmonic oscillator form (1.2) for the one-dimensional Stark problem. Furthermore, the implicit function argument also yields the upper bound needed on the perturbation coefficients to prove the convergence of the continued fraction.

More precisely, we can state the following theorem.

Theorem 3.2: Let m and i be given nonnegative integers. Then there are constants $B' = B'(m,i) > 0$ such that all eigenvalues $E_i^m(F)$ are analytic in the region $\{F\,|\,F$ on a three-sheeted Riemann surface$|\ 0 < |F| < B'|\ \ |\arg(F)| < \eta\}$ for any $\eta < 3\pi/2$.

Proof: As remarked by Titchmarsh,[2] under the transformation $y = x^2$, $\Phi(y) = x^{1/2}\Psi(x)$, the differential equation

$$\left(-\frac{d^2}{dy^2} + \frac{m^2-1}{4y^2} - Z/y + Fy\right)\Phi = E\Phi$$

goes into

$$\left(-\frac{d^2}{dx^2} + \frac{m^2 - \frac{1}{4}}{x^2} - 4Ex^2 + 4Fx^4\right)\Psi = 4Z\Psi. \tag{3.13}$$

This last equation is a quartic anharmonic oscillator with a centrifugal term. Here the charge Z appears as the spectral parameter, and the energy E as the harmonic coupling. Hence, defining the operator $K_m(E,F)$ out of the differential expression (3.13) exactly as in Ref. 1, Appendix, $\sigma(H_m(E,F))$ will, of course, coincide with $\tilde{\sigma}(K_m(E,F))$, the generalized spectrum of $K_m(E,F)$. This generalized spectrum is defined[10] as the set of all complex E for which the resolvent $[K_m(E,F) - \lambda]^{-1}$ at fixed $\lambda = Z$ does not exist as a bounded operator in $L^2(R_+)$. Since all the results of Simon[4] on $-d^2/dx^2 - 4Ex^2 + 4Fx^4$ can be taken over to $K_m(E,F)$ (see Ref. 1, Appendix), for $|\arg(F)| < \pi$ $K_m(E,F)$ has discrete spectrum. Its eigenvalues $\lambda_i^m(E,F)$, $m,i = 0,1,\cdots$, for any fixed E with $|\arg(-E)| < \epsilon$ are analytic functions of F in the region

$\{F\,|\,F$ on a three-sheeted Riemann surface$|$

$$0 < |F| < B(m,i)|\ \ |\arg(F)| < \eta'\}$$

for any $\eta' = < 3\pi/2 - \epsilon$, with

$$\frac{\partial\lambda_i^m(E,F)}{\partial F} \neq 0, \quad \lim_{|F|\to 0}\lambda_i^m(E,F) = \lambda_i^m(E),$$

when E and F belong to the regions specified above. Now, the eigenvalues $E_i^m(F)$ are implicitly defined by the relation

$$\lambda_i^m(E,F) = Z, \tag{3.14}$$

and the result follows by means of exactly the same implicit function argument of Ref. 1, Theorem 3.2.

Remark: As already mentioned, this result allows the realization of the resonances, for small field, according to

Howland's mechanism, because the motion of the eigenvalues to the resonances takes place by analytic continuation.

For any fixed $E<0$ we know that the formal Taylor expansion of $\lambda_i^m(E,F)$ near $F=0$, coinciding with the Rayleigh–Schrödinger perturbation series, is divergent as fast as $i!$ but uniformly asymptotic to $\lambda_i^m(E,F)$ in any sector $|\arg(F)|<\eta$, $\eta<3\pi/2$, and even Borel summable.[15]

Hence we can directly apply the explicitation arguments of Sec. V, Ref. 1, to conclude the following theorem.

Theorem 3.3: Let $E(F)$ stand for an arbitrary eigenvalue $E_i^m(F)$, $m,i=0,1,\cdots$, and let $\Sigma_0^\infty a_n F^n$ be its formal Rayleigh–Schrödinger perturbation expansion. Then $\Sigma_0^\infty a_n F^n$ is divergent but fulfills a strong asymptotic condition to $E(F)$ in any sector $|\arg(F)|<\eta$, $\eta<3\pi/2$.

This means that there are constants $A>0$, $B>0$ independent of F, such that, for all $N=1,2,\cdots$,

$$\left|E(F)-\sum_0^N a_n F^n\right|<AB^{N+1}(N+1)!|F|^{N+1}, \quad (3.15)$$

uniformly for $F\to 0$ in the above sector.

Remarks: (a) (3.15) implies *a fortiori*

$$|a_n|<AB^n n! \quad (3.16)$$

and hence

$$\sum_0^\infty a_n^{-1/n}=\infty. \quad (3.17)$$

(b) The bound (3.15), together with analyticity in the sector $|\arg(F)|<\eta$, $\eta<3\pi/2$, implies the Borel summability of $\Sigma_{n=0}^\infty a_n F^n$ to $E(F)$ in the whole sector $|\arg(F)|<\eta-\pi/2$.[15]

(c) By the Symanzik scaling we have $E(Z,F)=F^{1/3}E(ZF^{-2/3},1)$. This yields the following asymptotic behavior

$$E(F)\sim CF^{1/3} \quad (3.18)$$

as $F\to\infty$ in any direction within $|\arg(F)|<\pi$, C being some constant depending only on m and i.

Now, by piecing together Lemma 3.1 (Herglotz property), Theorem 3.1 (first-sheet analyticity), Theorem 3.3 [bound on the perturbation coefficients (3.16)], and the asymptotic behavior (3.18) we can directly apply Theorem IV.2.2 of Simon[4] to conclude Theorem 3.4.

Theorem 3.4: The Stieltjes moment problem

$$(-1)^{n+1}a_n=\int_0^\infty x^n d\phi(x), \quad n\geqslant 1 \quad (3.19)$$

has a unique solution, i.e., there is one and only one (normalized) positive measure $d\phi$ on R_+ satisfying (3.19).

By the classical Stieltjes theorem (see, e.g., Ref. 16) Theorem 3.4 implies (and is implied) by the following.

Theorem 3.5: The Stieltjes type continued fraction associated with the formal power series $\Sigma_1^\infty a_n F^{n-1}$ exists and converges to $E'(F)=[E(F)-a_0]/F$, uniformly in any compact contained in the cut plane $|\arg(F)|<\pi$.

In other words, any diagonal sequence of Padé approximants on the formal perturbation series converges to $E(F)$, uniformly on compacts as above.

Remarks: (a) The width of the resonances is, of course, given by the boundary value of $\mathrm{Im}[E(F)]$ at the lower edge of the cut. Under the present conditions it is known[16] that this boundary value is given by the measure $F^{-1}\pi d\phi(-F^{-1})$, $F>0$. For small field, by Theorem 3.2 the limit exists as an analytic function, given by $F^{-1}\pi\phi'(-F^{-1})$.

(b) The moment theory[16,17] approximates $d\phi$ through a measure-convergent sequence of "Dirac measures," i.e., of measures $d\phi_n(x)$ defined as follows,

$$d\phi_n(x)=\sum_0^n \nu_i\delta_{\lambda_i}(-x), \quad (3.20)$$

where $\delta_{\lambda_i}(x)$ is the measure concentrated on the ith pole λ_i of the nth approximant of the continued fraction, and the weight ν_i is the corresponding residue. (We recall that any pole λ_i is negative and simple and has positive residue; the nth order approximant is exactly the $[n,n]$ Padé approximant on the formal series for n odd, and the $[n-1,n]$ for n even.) By Helly's first theorem, we can in addition find a subsequence $\{\phi_{n_k}(x)\}_{k=1}^\infty$ which converges to $\phi(x)$ in all points of continuity of ϕ, the distribution function of the measure $d\phi$. To sum up, in the present context the width of the resonances is approximated through convergent sequences of "weighted sums of δ functions" concentrated on the poles of the Padé approximants, the weights being given by the corresponding residues.

(c) Given the approximation to the width, the approximation to the position of the resonance, given, of course, by the boundary value of $\mathrm{Re}[E(F)]$ at the lower edge of the cut, is, of course, specified through the Cauchy principal value,

$$a_0+F\,\mathrm{P.P.}\int_0^\infty (1+Fx)^{-1}x\,d\phi_n(x). \quad (3.21)$$

Note that (3.21) is just the nth approximant of the continued fraction when $F\neq\lambda_i^{-1}$, $i=1,\ldots,n$.

[1] S. Graffi and V. Grecchi, Lett. Math. Phys. **2**, 335 (1978); Commun. Math. Phys. **62**, 83 (1978).
[2] E.C. Titchmarsh, *Eigenfunction Expansions* (Oxford U.P., New York, 1958), Vol. II, Chap. XX.
[3] See, e.g., H.S. Wall, *Continued Fractions* (Van Nostrand, New York, 1947).
[4] B. Simon, Ann. Phys. (N.Y.) **58**, 76 (1970).
[5] J.J. Loeffel and A. Martin, Proc. Programme 25th Conf., Strasburg (May 1970).
[6] E. Balslev and J.M. Combes, Commun. Math. Phys. **22**, 280 (1971).
[7] See, e.g., M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1975), Vol. II, Sec. X.1., Appendix.
[8] See, e.g., G. Hellwig, *Differentialoperatoren der Mathematischen Physik* (Springer–Verlag, Berlin, 1964).
[9] See, e.g., I.S. Gradstein and I.M. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1965).
[10] T. Kato, *Perturbation Theory of Linear Operators* (Springer–Verlag, Berlin, 1966).
[11] B. Simon, Ann. Math. (N.Y.) **97**, 247 (1973).
[12] J.S. Howland, Pac. J. Math. **58**, 157 (1975).
[13] J.J. Loeffel *et al.*, Phys. Lett. B **30**, 656 (1969).
[14] B. Simon, Proc. Cargèse Summer School (1970).
[15] S. Graffi, V. Grecchi, and B. Simon, Phys. Lett. B **32**, 631 (1970).
[16] J. Shohat and J.M. Tamarkin, *The Problem of Moments* (Am. Math. Soc., Math. Surveys, New York, 1943), Vol. I.
[17] N.I. Akhiezer, *The Classical Moment Problem* (Oliver and Boyd, Edinburgh, 1965).
[18] I.W. Herbst, Commun. Math. Phys., to appear.
[19] I.W. Herbst and B. Simon, Phys. Rev. Lett. **41**, 67 (1978).

Linear representations of any dimensional Lorentz group and computation formulas for their matrix elements

Takayoshi Maekawa

Department of Physics, Kumamoto University, Kumamoto, Japan
(Received 7 July 1978)

The representation matrix elements of SO(n,1) are discussed in a space spanned by the representation matrix elements of the maximal compact subgroup SO(n). A multiplier of the representation corresponding to the boost of SO(n,1) is completely determined by requiring the commutation relations of SO(n,1) for the differential operators of the multiplier representation and of the parameter group of SO(n). It is shown that the bases of the space, the representation matrix elements of SO(n), are classified by the group chains of the first and the second parameter groups of SO(n), whose differential operators commute with each other, and the characteristic numbers of SO(n,1) are the same as those of the first parameter group of SO($n-1$) and a complex number appearing in the multiplier. By using the scalar product defined in the space, the matrix elements for the differential operators and the computation formulas for the representation corresponding to the boost of SO(n,1) are given for all unitary representations of SO(n,1) and useful formulas containing the d matrix elements of SO(n) are obtained. By making use of these results, even for the nonunitary representation of SO(n,1) the matrix elements for the differential operators and the computation formula for the representation corresponding to the boost are obtained by defining the matrix elements with respect to the bases of the space. It is also pointed out that the unitary representations (the complementary series) corresponding to some value of the parameter, which appear in the classification using only the matrix elements of the generators, should not be included in our classification table because of divergence of the normalization integral. The continuation to SO($n+1$) and the contraction to ISO(n) from the principal series are discussed.

I. INTRODUCTION

In a previous paper,[1] the formulas for computing the representation matrix elements of the group SO(n,1) are obtained for the complementary series as well as for the principal series of the unitary irreducible representation (UIR). According to these formulas, it will be easy to analyse the D matrix elements of SO(n,1), which are important for studying the orthogonal and the complete set of functions on the group manifold.[2] One of the next problems will be to give the explicit expressions for the representation matrix elements and to study their properties and the other to complete the procedure used in Ref. 1.

In Ref. 1, explicit use is made of the fact that the representations of the group are determined by those of its Lie algebra, and in a UIR of SO(n,1) an irreducible representation of SO(n) is contained either with multiplicity one or not at all.[3] Nevertheless, it seems that a few points to be modified and clarified in the procedure are contained in Ref. 1. The UIR's of SO(n,1) are classified according to the representation of the Lie algebra of SO(n,1).[4,5] The bases of the representations are classified by the group chain SO(n,1) $\supset$SO(n)$\cdots\supset\cdots\supset$ SO(2) and the matrix elements of the generators are given with respect to the bases which will be called the standard bases. In Ref. 1, however, the representations of SO(n,1) are considered in a space consisting of the representation matrix elements of the maximal compact subgroup, i.e., the space spanned by the D matrix elements of SO(n). Though the D matrix elements of SO(n) are defined through the standard bases, the bases of the space, i.e., the D matrix elements, are also classified by the group chains of the first and the second parameter groups of SO(n), that is, by the characteristic numbers of a complete set of the commuting operators formed from the differential operators of the first and the second parameter groups respectively.[6] As the differential operators of the first and the second parameter groups commute with each other, the characteristic numbers in one of the parameter groups can be regarded as invariants relative to the other. Though Ref. 1 uses the fact that the characteristic numbers of one of the parameter groups coincide with those of SO(n,1), it seems not evident from the beginning whether they conicide or not. It is important for us to clarify these points in order to fix completely the bases of the representation. If this is done, we can calculate the matrix elements of the differential operators and construct the computation formulas for the representation matrix elements of SO(n,1) with respect to the bases

Though in Ref. 1 the matrix elements of the generators and the differential operators with respect to the standard bases and the bases in the space are used in order to obtain a relation between the numerical constants, it is evident that the matrix elements calculated by using only the bases in the space must be used in order that our method is consistent. As the scalar product is introduced on the group manifold of SO(n),it is evident that the integral must converge in order to be able to define the representation. This is left untouched in Ref. 1.

The purpose of this article is to study the above points and to construct the computation formula for the representation matrix elements corresponding to the boost together with the matrix elements for the differential operators of the nonunitary irreducible representation of SO(n,1). Though most of the results on the UIR's are the same as those in Ref. 1, a detailed discussion even on the UIR's will be given for completeness of our procedure.

In Sec. 2, Euler parameters for the group SO(n) are introduced and then the differential operators of the first and the second parameter groups are discussed.[6] In Sec. 3, the Gel'fand and Tsetlin bases of the groups SO(n) and SO(n,1) are introduced and the matrix elements for the generators with respect to the bases are given.[4,5,7,8] The tables of the UIR's of SO(n,1) are given with a slight modification of the notations in Ref. 5. In Sec. 4, the representation matrix elements of SO(n) are given together with the orthogonality and the completeness relations.[2,9,10] In Sec. 5, a multiplier of the representation[2] and the differential operator (generator) of the multiplier representation, which is considered in a linear space consisting of functions $f(g), g \in$ SO(n), are determined by the requirement that the differential operator together with the differential operators of the second parameter group of SO(n) satisfy the commutation relations of SO(n,1). In Sec. 6, the UIR's of SO(n,1) are discussed in the space consisting of the representation matrix elements of SO(n), where the scalar products are defined for the principal and the complementary series corresponding to Tables I and II given in Sec. 3. It is shown that the numbers characterizing the irreducible representation of SO(n,1) are intimately connected with the characteristic numbers of the first parameter group and a complex number contained in the multiplier. The bases of the representation are fixed completely and the matrix elements of the differential operators and the computation formulas for the representation matrix elements corresponding to the boost are obtained. In Sec. 7, the matrix elements for the differential operator and the computation formula for the representation corresponding to the boost of the nonunitary representation are obtained. In Sec. 8, a simple discussion on the classification of the UIR's of SO(n,1) is given and it is pointed out that the case of the parameter $\sigma_{n+1} = 0$ ($n+1$ odd), which is contained in the case of the classification due to the representations of the Lie algebra of SO(n,1) used in Sec. 3, should be exluded because of divergence of the normalization integral. In Sec. 9, the irreducible representations of the groups SO($n+1$) and ISO(n) (inhomogeneous rotation group) are obtained from those of the principal series of SO(n,1) by continuation[11] and concentration.[12] In Sec. 10, simple examples for the general results of Sec. 7 are given.

2. EULER PARAMETERS AND INFINITESIMAL OPERATORS

In this section, in order to fix notation, Euler parameters, which are a generalization to any dimension of the three-dimensional Euler angles, are introduced[2,5,6,10] and the differential operators of the first and the second parameter groups are discussed.[6]

Let us consider an n-dimensional Euclidean space, whose orthonormal bases are denoted by two systems of the unit vectors, i.e, $(\mathbf{e}_1,\mathbf{e}_2,...,\mathbf{e}_n)$ and $(\bar{\mathbf{e}}_1,\bar{\mathbf{e}}_2,...,\bar{\mathbf{e}}_n)$ with the inner product $(\mathbf{e}_i,\mathbf{e}_j) = (\bar{\mathbf{e}}_i,\bar{\mathbf{e}}_j) = \delta_{ij}$. These bases are related to each other as follows,

$$\bar{\mathbf{e}}_j = \sum_k g_{jk}^{(n)} \mathbf{e}_k, \tag{2.1}$$

where the matrices $g^{(n)}$ are given in terms of the Euler parameters θ_{jk} ($j = 2,3,...n$, $k = 1,2,...j-1$),

$$\begin{aligned} &g^{(n)} = g^{(n-1)} S_n, \\ &S_n = \prod_{j=n}^{2} t_{jj-1}^{(n)}(\theta_{nn-j+1}), \\ &g^{(1)} = \text{unit matrix}, \\ &0 \leqslant \theta_{jk} \leqslant \pi \quad (k = 1,2,..,j-2), \quad 0 \leqslant \theta_{jj-1} < 2\pi. \end{aligned} \tag{2.2}$$

The $g^{(n)}$ is an orthogonal matrix and an element of the group SO(n). The $t_{jk}^{(p)}(\theta)$ is a $n \times n$ matrix corresponding to the rotation of an angle θ in the jk plane in the p-dimensional space and has the following nonzero elements:

$$\begin{aligned} &(t_{jj-1}^{(p)}(\theta))_{jj-1} = -(t_{jj-1}^{(p)}(\theta))_{j-1j} = \sin\theta, \\ &(t_{jj-1}^{(p)}(\theta))_{jj} = (t_{jj-1}^{(p)}(\theta))_{j-1j-1} = \cos\theta, \\ &(t_{jk}^{(p)}(\theta))_{rs} = \delta_{rs} \quad (r,s \neq j,k). \end{aligned} \tag{2.3}$$

It is noted that the definition of the Euler parameters in (2.2) differs from the usual one for a special case of $n = 3$. It, however, is easy to relate them. It follows from (2.2) and (2.3) that the following relations hold:

$$(g^{(n)'} g^{(n)-1})_{nn} = \cos\theta'_{n1}\cos\theta_{n1} + \sin\theta'_{n1}\sin\theta_{n1}(g^{(n-1)'} g^{(n-1)-1})_{n-1n-1}, \tag{2.4a}$$

$$(g^{(n)})_{nn} = \cos\theta_{n1}. \tag{2.4b}$$

It is noted that the bracketed term on the right hand side is independent of θ'_{n1} and θ_{n1}.

The Haar measure is given as follows:

$$\begin{aligned} &dV_n = dV_{n-1} dS_n, \quad dV_2 = d\theta_{21}, \\ &dS_n = \prod_{j=1}^{n-1} [(\sin\theta_{nj})^{n-j-1} d\theta_{nj}], \end{aligned} \tag{2.5}$$

where dS_n is the surface element of a sphere in a n-dimensional space. The volume V_n of SO(n) is given by $V_n = V_{n-1} 2\pi^{n/2}/\Gamma(n/2)$ and $V_2 = 2\pi$.

The representation D matrix of SO(n) corresponding to the rotation (2.2) is given as follows:

$$D^{(n)}(g^{(n)}) = D^{(n-1)}(g^{(n-1)}) H(S_n), \tag{2.6a}$$

$$H(S_n) = \prod_{j=n}^{2} R_{jj-1}^{(n)}(\theta_{nn-j+1}), \tag{2.6b}$$

where $R_{jj-1}^{(p)}(\theta)$ is the representation corresponding to the rotation $t_{jj-1}^{(p)}(\theta)$ and is given in a form $\exp(i\theta D_{jj-1})$ with the infinitesimal generator D_{jk} ($= -D_{kj}$, Hermitian) of the representation of SO(n), which satisfies the commutation relations

$$[D_{jk}, D_{lm}] = i(\delta_{lj} D_{km} + \delta_{km} D_{jl} - \delta_{kl} D_{jm} - \delta_{jm} D_{kl}). \tag{2.7}$$

In order to obtain the differential operators, we introduce the combined transformations as follows:

$$e^{(i/2)\Sigma_{j,k}u_{jk}D_{jk}}D^{(n)}(g^{(n)}) = D^{(n)}(g^{(n)\prime}), \tag{2.8a}$$

$$D^{(n)}(g^{(n)})e^{(i/2)\Sigma_{j,k}u'_{jk}D_{jk}} = D^{(n)}(g^{(n)\prime\prime}), \tag{2.8b}$$

where the u_{jk}'s are the standard parameters which are assumed to be uniquely determined by the Euler parameters $(\bar{\theta}_{jk})$ and vice versa, and tend to zero as $\bar{\theta}_{jk} \to 0$. $g^{(n)\prime}$ and $g^{(n)\prime\prime}$ depends on θ and $\bar{\theta}$.

Taking a derivative of (2.8a) and (2.8b) with respect to u_{jk} and u'_{jk} at $u = 0$, we obtain

$$D_{jk}D^{(n)}(g^{(n)}) = \bar{J}_{jk}D^{(n)}(g^{(n)}), \tag{2.9a}$$

$$D^{(n)}(g^{(n)})D_{jk} = J_{jk}D^{(n)}(g^{(n)}), \tag{2.9b}$$

where

$$\bar{J}_{jk} = \sum_{l,m} \frac{\partial\theta_{lm}}{\partial u_{jk}} p_{lm}, \tag{2.10a}$$

$$J_{jk} = \sum_{l,m} \frac{\partial\theta_{lm}}{\partial u'_{jk}} p_{lm}, \tag{2.10b}$$

$$p_{jk} = -i\frac{\partial}{\partial\theta_{jk}},$$

and $\bar{J}_{jk}$ and J_{jk} are the differential operators of the first and the second parameter groups, respectively.[13] If the infinitesimal generator D_{jk} are given in the matrix form, the left-hand side of (2.9) is given in a linear combination of the matrix elements, whereas the right-hand side is a linear combination of derivatives of one matrix element. By our assumption on the relations between θ_{jk} and u_{jk} (u'_{jk}), the following relations hold:

$$\frac{1}{2}\sum_{r,s} \frac{\partial\theta_{jk}}{\partial u_{rs}}\frac{\partial u_{rs}}{\partial\theta_{lm}} = \delta_{jl}\,\delta_{km},$$
$$\frac{1}{2}\sum_{r,s} \frac{\partial\theta_{jk}}{\partial u'_{rs}}\frac{\partial u'_{rs}}{\partial\theta_{lm}} = \delta_{jl}\,\delta_{km}, \tag{2.11}$$

Then (2.10) may be rewritten as follows:

$$p_{lm} = \frac{1}{2}\sum_{j,k}\frac{\partial u_{jk}}{\partial\theta_{lm}}\bar{J}_{jk}, \tag{2.12a}$$

$$p_{lm} = \frac{1}{2}\sum_{j,k}\frac{\partial u'_{jk}}{\partial\theta_{lm}}J_{jk}. \tag{2.12b}$$

On the other hand, we obtain from (2.8) by taking derivatives with respect to $\bar{\theta}$ at θ,

$$\frac{i}{2}\sum_{j,k}\frac{\partial u_{jk}}{\partial\theta_{lm}}D_{jk} = \Big(\frac{\partial}{\partial\theta_{lm}}D^{(n)}(g^{(n)})\Big)[D^{(n)}(g^{(n)})]^{-1}, \tag{2.13a}$$

$$\frac{i}{2}\sum_{j,k}\frac{\partial u'_{jk}}{\partial\theta_{lm}}D_{jk} = [D^{(n)}(g^{(n)})]^{-1}\Big(\frac{\partial}{\partial\theta_{lm}}D^{(n)}(g^{(n)})\Big). \tag{2.13b}$$

These relations must hold for any representation and we can take $g^{(n)}$ for $D^{(n)}$. Then we obtain

$$\frac{\partial u_{jk}}{\partial\theta_{lm}} = \left(\frac{\partial g^{(n)}}{\partial\theta_{lm}}g^{(n)^{-1}}\right)_{jk}, \tag{2.14a}$$

$$\frac{\partial u'_{jk}}{\partial\theta_{lm}} = \left(g^{(n)^{-1}}\frac{\partial g^{(n)}}{\partial\theta_{lm}}\right)_{jk}, \tag{2.14b}$$

where the relations $(D_{jk})_{pq} = -i(\delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp})$, which are valid for the infinitesimal generator D_{jk} of $g^{(n)}$, are used in order to derive these relations from (2.13). Thus, we can express the differential operators $\bar{J}_{jk}$ and J_{jk} in terms of the Euler parameters (θ_{jk}) and the differential operators (p_{jk}) by substituting (2.14) into (2.12).[6]

From (2.9), it follows that $\bar{J}_{jk}$ and J_{jk} satisfy the commutation relations:

$$[\bar{J}_{jk},\bar{J}_{lm}] = -i(\delta_{jl}\bar{J}_{km} + \delta_{km}\bar{J}_{jl} - \delta_{jm}\bar{J}_{kl} - \delta_{kl}\bar{J}_{jm}), \tag{2.15a}$$

$$[J_{jk},J_{lm}] = i(\delta_{jl}J_{km} + \delta_{km}J_{jl} - \delta_{jm}J_{kl} - \delta_{kl}J_{jm}), \tag{2.15b}$$

$$[\bar{J}_{jk},J_{lm}] = 0. \tag{2.15c}$$

It is seen that the expressions for $\bar{J}_{jk}$ and J_{jk} are given in the form:

$$\bar{J}_{21} = p_{21},$$
$$\bar{J}_{k+1k} = \cos\theta_{k1}p_{k+11} - \frac{\cos\theta_{k+11}}{\sin\theta_{k+11}}\sin\theta_{k1}p_{k1} + \frac{\sin\theta_{k1}}{\sin\theta_{k+11}}\bar{J}'_{kk-1}, \tag{2.16a}$$

$$J_{21} = p_{nn-1},$$
$$J_{k+1k} = \cos\theta_{nn-k+1}p_{nn-k} - \frac{\cos\theta_{nn-k}}{\sin\theta_{nn-k}}\sin\theta_{nn-k+1}p_{nn-k+1} + \frac{\sin\theta_{nn-k+1}}{\sin\theta_{nn-k}}J'_{kk-1}, \tag{2.16b}$$

$$k = 2,3,\dots,n-1$$

where the primes on $\bar{J}'_{kk-1}$ and J'_{kk-1} mean the substitutions $\theta_{jl}, p_{jl} \to \theta_{j+1l+1}, p_{j+1l+1}$ in $\bar{J}_{kk-1}$ and $\theta_{jl}, p_{jl} \to \theta_{j-1l-1}, p_{j-1l-1}$ in J_{kk-1} respectively. The expressions for any $\bar{J}_{jk}$ and J_{jk} are obtained from (2.15a) and (2.15b) by using (2.16).

A similar parametrization to that of SO(n) for SO(n,1) is possible and given as follows:

$$g^{(n,1)} = g^{(n)}S^{(n,1)},$$
$$S^{(n,1)} = {}^b t^{(n+1)}_{n+1n}(\zeta)\prod_{j=n}^{2} t^{(n+1)}_{jj-1}(\theta_{n+1n-j+2}), \tag{2.17}$$

where $g^{(n,1)} \in$ SO(n,1) and ${}^b t^{(n+1)}_{n+1n}(\zeta)$ is a boost in the nth direction through ζ $(0 \leqslant \zeta < \infty)$ and has nonzero elements as follows:

$$({}^b t^{(n+1)}_{n+1n}(\zeta))_{nn} = ({}^b t^{(n+1)}_{n+1n}(\zeta))_{n+1n+1} = \cosh\zeta,$$
$$({}^b t^{(n+1)}_{n+1n}(\zeta))_{nn+1} = ({}^b t^{(n+1)}_{n+1n}(\zeta))_{n+1n} = -\sinh\zeta,$$
$$({}^b t^{(n+1)}_{n+1n}(\zeta))_{rs} = \delta_{rs} \quad (r,s<n).$$

The infinitesimal generators D_{jk} of the representation of SO$(n,1)$ satisfy the same commutation relations as (2.7) in which Kronecker δ's are replaced by the metric tensor g's, i.e., $\delta_{jk} \to g_{jk}$ with nonzero elements $g_{jj} = -g_{n+1n+1} = 1$ $(j = 1,2,...,n)$. The representation D matrix of SO$(n,1)$ corresponding to (2.17) is also given by the form similar to (2.6) in which we substitute $n+1$ and $R^{\,b}_{n+1n}(\zeta)$ for n and $R^{(n+1)}_{n+1n}(\theta_{n+11})$, where $R^{\,b}_{n+1n}(\zeta)$ is the representation corresponding to the boost ${}^b t^{(n+1)}_{n+1n}(\zeta)$. The Haar measure is obtained from the expression for SO$(n+1)$ by the substitution of $(\sinh\zeta)^{n-1}d\zeta$ for $(\sin\theta_{n+11})^{n-1}d\theta_{n+11}$. The differential operators $(\bar{J}_{jk}, J_{jk}; j,k \leqslant n)$ of the first and the second parameter groups are given by the same form as (2.16), and the expressions for $\bar{J}_{n+1n}$ and J_{n+1n} are given by

$$\bar{J}_{n+1n} = \cos\theta_{n1} p_\zeta - \frac{\cosh\zeta}{\sinh\zeta}\sin\theta_{n1} p_{n1} + \frac{\sin\theta_{n1}}{\sinh\zeta}\bar{J}'_{nn-1}, \tag{2.18a}$$

$$J_{n+1n} = \cos\theta_{n+12} p_\zeta - \frac{\cosh\zeta}{\sinh\zeta}\sin\theta_{n+12} p_{n+12} + \frac{\sin\theta_{n+12}}{\sinh\zeta} J'_{nn-1}, \tag{2.18b}$$

where $p_\zeta = -i\partial/\partial\zeta$ and the primes on $\bar{J}'_{nn-1}$ and J'_{nn-1} have the same meaning as in (2.16). The expressions for any other $\bar{J}_{n+1j}$ and J_{n+1j} $(j<n)$ are calculated from

$$\bar{J}_{n+1j} = -i[\bar{J}_{n+1n}, \bar{J}_{nj}],$$
$$J_{n+1j} = i[J_{n+1n}, J_{nj}],$$

It is noted that the expressions (2.18) are obtained from (2.16) by the substitutions of $n+1$, $-i\zeta$, $i\bar{J}_{n+1n}$, and iJ_{n+1n} for n, θ_{n1}, $\bar{J}_{n+1n}$, and J_{n+1n}.

Another important parametrization of an element of SO$(n,1)$ is that due to the Iwasawa decomposition,[14] by which any element $g^{(n,1)}$ of SO$(n,1)$ can be written uniquely as follows,

$$g^{(n,1)} = n(\xi)\ {}^b t^{(n+1)}_{n+1n}(\eta) g^{(n)}, \tag{2.19}$$

where $(n+1,j)$ and $(j,n+1)$ elements of $g^{(n)}$ are given by $(g^{(n)})_{n+1j} = (g^{(n)})_{jn+1} = \delta_{jn+1}$, $n(\xi) \in N$, the nilpotent subgroup of SO$(n,1)$, and the expression for $n(\xi)$ is given as follows,

$$n(\xi) = \begin{pmatrix} 1 & \xi & \xi \\ -\xi^t & 1-\Xi & -\Xi \\ \xi^t & \Xi & 1+\Xi \end{pmatrix},$$

$$\xi^t = (\xi_1, \xi_2, \cdots, \xi_{n-1}), \quad \Xi = (\sum_i \xi_i^2)/2.$$

Considering the transformation

$$g^{(n,1)} \xrightarrow{\zeta} g^{(n,1)'} = g^{(n,1)}\ {}^b t^{(n+1)}_{n+1n}(\zeta)$$
$$= n(\xi')\ {}^b t^{(n+1)}_{n+1n}(\eta') g^{(n)'}, \tag{2.20}$$

we easily obtain

$$\cos\theta'_{n1} = \frac{\cos\theta_{n1}\cosh\zeta - \sinh\zeta}{\cosh\zeta - \cos\theta_{n1}\sinh\zeta}, \tag{2.21a}$$

$$\theta'_{jk} = \theta_{jk} \quad \text{for} \quad \theta_{jk} \neq \theta_{n1}, \tag{2.21b}$$

$$\exp\eta' = \exp\eta(\cosh\zeta - \cos\theta_{n1}\sinh\zeta). \tag{2.21c}$$

It is noted that the following relation holds from (2.21a) and (2.21b),

$$dV'_n = (\cosh\zeta - \cos\theta_{n1}\sinh\zeta)^{1-n} dV_n \tag{2.22}$$

3. REPRESENTATIONS OF THE LIE ALGEBRA OF SO(*n*) AND SO(*n*, 1)

In this section, the results needed for the following sections on the groups SO(n) and SO$(n,1)$ are summarized.[5,7,8]

The generators D_{jk} of the representation of SO(n) satisfy the commutation relations (2.7). The basis vectors for the unitary irreducible representation (UIR) of SO(n), classified by the group chain SO$(n) \supset$ SO$(n-1) \supset \cdots \supset$ SO(2), are given by the Gel'fand and Tsetlin labels as follows,

$$|m_{jk}\rangle = |\lambda_n, \lambda_{n-1}, ..., \lambda_2\rangle, \tag{3.1}$$

where λ_j stands for $(m_{j1}, m_{j2}, ..., m_{j[j/2]})$, $[j/2]$ means the largest integer smaller or equal to $j/2$, and all λ_j's are written in a row. The numbers m_{jk} are simultaneously integers or half-integers and are subject to the restrictions:

$$\begin{aligned} &m_{2j+1i+1} \leqslant m_{2ji} \leqslant m_{2j+1i} \quad (i = 1,2,...,j-1), \\ &m_{2ji+1} \leqslant m_{2j-1i} \leqslant m_{2ji} \quad (i = 2,3,...,j-1), \\ &|m_{2jj}| \leqslant m_{2j-1j-1} \leqslant m_{2jj-1}, \\ &|m_{2jj}| \leqslant m_{2j+1j}. \end{aligned} \tag{3.2}$$

The UIR of SO(n) is characterized by the $[n/2]$ numbers m_{nj}, while the rows and columns of the matrix of D_{jk} are labeled by $(\lambda_{n-1}, \lambda_{n-2}, ..., \lambda_2)$. It is noted that all matrices of the Lie algebra, i.e., D_{jk}, are completely determined from those of D_{jj-1} $(j = 2,3,...,n)$ because of the commutation relations (2.7).

The dimension of the UIR of SO(n) is determined by the numbers m_{nj} subjected to the restrictions (3.2) and given as follows[9,15]:

$$N(\lambda_n) = \left[(n-2)!(n-4)!! \prod_{j=1}^{(n-5)/2} (n-2j-3)!\right]^{-1} \times \prod_{i=1}^{(n-1)/2} (2l_{ni}-1) \times \prod_{j=1}^{(n-3)/2} \prod_{k=j+1}^{(n-1)/2} [\,(l_{nj}-\tfrac{1}{2})^2 - (l_{nk}-\tfrac{1}{2})^2], \tag{3.3a}$$

for n odd, and

$$N(\lambda_n) = 2^{(n-2)/2}\left[(n-2)!(n-4)!!\prod_{j=1}^{(n-4)/2}(n-2j-3)!\right]^{-1} \times \prod_{j=1}^{(n-2)/2}\prod_{h=j+1}^{n/2}(l_{nj}^2 - l_{nk}^2), \quad (3.3b)$$

for n even. The quantity l_{jk} is defined by

$$l_{jk} = m_{jk} + [(j+1)/2] - k \quad (k = 1,2,...,[j/2]). \quad (3.4)$$

The action of the generators D_{jj+1} $(1 \leqslant j \leqslant n-1)$ on the bases is given by

$$D_{2k\,2k+1}|m_{ij}\rangle = \sum_{j=1}^{k} A(m_{2kj})|m_{2kj}+1\rangle - \sum_{j=1}^{k} A(m_{2kj}-1)|m_{2kj}-1\rangle, \quad (3.5a)$$

$$D_{2k-1\,2k}|m_{ij}\rangle = \sum_{j=1}^{k-1} B(m_{2k-1j})|m_{2k-1j}+1\rangle - \sum_{j=1}^{k-1} B(m_{2k-1j}-1)|m_{2k-1j}-1\rangle + C_{2k}|m_{ij}\rangle. \quad (3.5b)$$

It is noted that in the bases on the right-hand side only the numbers, which change under the action of the generator, are written except for the last $|m_{ij}\rangle$. The matrix elements A, B, and C in these equations are given as follows:

$$A(m_{2kj}) = \frac{i}{2}\left\{\prod_{i=1}^{k-1}[(l_{2k-1i}-\tfrac{1}{2})^2 - (l_{2kj}+\tfrac{1}{2})^2] \times \prod_{i=1}^{k}[(l_{2k+1i}-\tfrac{1}{2})^2 - (l_{2kj}+\tfrac{1}{2})^2]\right\}^{1/2} \times \left\{\prod_{i=1}^{k}{}'(l_{2ki}^2 - l_{2kj}^2)[l_{2ki}^2 - (l_{2kj}+1)^2]\right\}^{-1/2}, \quad (3.6a)$$

$$B(m_{2k-1j}) = i\left[\prod_{i=1}^{k-1}(l_{2k-2i}^2 - l_{2k-1j}^2)\prod_{i=1}^{k}(l_{2ki}^2 - l_{2k-1j}^2)\right]^{1/2} \times \left\{(l_{2k-1j}^2(4l_{2k-1j}^2-1)\prod_{i=1}^{k-1}{}'(l_{2k-1i}^2 - l_{2k-1j}^2) \times [(l_{2k-1i}-1)^2 - l_{2k-1j}^2]\right\}^{-1/2}, \quad (3.6b)$$

$$C_{2k} = \prod_{i=1}^{k-1} l_{2k-2i}\prod_{i=1}^{k} l_{2ki}\left[\prod_{i=1}^{k-1} l_{2k-1i}(l_{2k-1i}-1)\right]^{-1}, \quad (3.6c)$$

where the prime on Π' means the product of factors with $i \neq j$ and l_{jk} is given by (3.4).

The second-order Casimir operator is

$$F^{(n)} = \sum_{j>k}^{n} D_{jk}^2, \quad (3.7)$$

whose eigenvalue with respect to (3.1) is given by[16]

$$F^{(n)} = \sum_{j=1}^{[n/2]} m_{nj}(m_{nj} + n - 1j). \quad (3.8)$$

Similarly, the bases of the irreducible representation of SO(n,1) can be classified by the group chain SO(n,1) $\supset$...SO(n)$\supset\cdots\supset$SO(2) and written in a form

$$|m_{ij}\rangle^{\rho_{n+1}} = |(\rho_{n+1}, m_{n+12},...,m_{n+1[(n+1)/2]}), \lambda_n,...,\lambda_2\rangle, \quad (3.9)$$

where all numbers m_{jk} are subject to (3.2) and ρ_{n+1} is a complex number. The numbers m_{n+1j} $(j \geqslant 2)$ and the complex number ρ_{n+1}, which will sometimes be written as $\rho_{n+1} = (1-n)/2 + \sigma_{n+1} + i\nu_{n+1}$ with real σ_{n+1} and ν_{n+1}, characterize the irreducible representation of SO(n,1).

The action of the generators D_{jj+1} $(j = 1,2,...,n-1)$ on the bases (3.9) and their matrix elements of D_{jj+1} are given by the same formulas as in SO(n), i.e., (3.5) and (3.6). The action of D_{n+1n} on the bases (3.9) is given by the same form as (3.5), i.e.,

(i) for $n+1$ odd,

$$D_{nn+1}|m_{ij}\rangle^{\rho_{n+1}} = \sum_{j=1}^{n/2} {}^bA(m_{nj})|m_{nj}+1\rangle^{\rho_{n+1}} - \sum_{j=1}^{n/2} {}^bA(m_{nj}-1)|m_{nj}-1\rangle^{\rho_{n+1}}, \quad (3.10a)$$

(ii) for $n+1$ even,

$$D_{nn+1}|m_{ij}\rangle^{\rho_{n+1}} = \sum_{j=1}^{(n-1)/2} {}^bB(m_{nj})|m_{nj}+1\rangle^{\rho_{n+1}} - \sum_{j=1}^{(n-1)/2} {}^bB(m_{nj}-1)|m_{nj}-1\rangle^{\rho_{n+1}} + {}^bC_{n+1}|m_{ij}\rangle^{\rho_{n+1}}. \quad (3.10b)$$

The matrix elements ${}^bA$, ${}^bB$, and ${}^bC$ are given as follows[4,5]

$${}^bA(m_{nj}) = \frac{i}{2}\left\{\prod_{i=1}^{(n-2)/2}[(l_{n-1i}-\tfrac{1}{2})^2 - (l_{nj}+\tfrac{1}{2})^2]\right.$$

$$\times \{(l_{nj}+\tfrac{1}{2})^2 - [\rho_{n+1}+(n-1)/2]^2\}$$

$$\times \prod_{i=2}^{n/2} [(l_{n+1i}-\tfrac{1}{2})^2-(l_{nj}+\tfrac{1}{2})^2]\Big]^{1/2}$$

$$\times \left\{ \prod_{i=1}^{n/2}{}' (l_{ni}^2 - l_{nj}^2)[l_{ni}^2-(l_{nj}+1)^2] \right\}^{-1/2}, \tag{3.11a}$$

$${}^bB(m_{nj})$$

$$= i\Bigg(\prod_{i=1}^{(n-1)/2} (l_{n-1i}^2 - l_{nj}^2)\{l_{nj}^2 - [\rho_{n+1}+(n-1)/2]^2\}$$

$$\times \prod_{i=2}^{(n+1)/2} (l_{n+1i}^2 - l_{nj}^2) \Bigg)^{1/2} \Bigg\{ l_{nj}^2(4l_{nj}^2 - 1)$$

$$\times \prod_{i=1}^{(n-1)/2} (l_{ni}^2 - l_{nj}^2)[(l_{ni}-1)^2 - l_{nj}^2] \Bigg\}^{-1/2}, \tag{3.11b}$$

$${}^bC_{n+1} = -i[\rho_{n+1}+(n-1)/2] \prod_{j=1}^{(n-1)/2} l_{n-1j}$$

$$\times \prod_{j=2}^{(n+1)/2} \prod_{j=2}^{(n+1)/2} l_{n+1j} \left[\prod_{j=1}^{(n-1)/2} (l_{nj}-1) \right]^{-1}, \tag{3.11c}$$

where the prime on Π' has the same meaning as in (3.6) and l_{kj} is given by (3.4). It is noted that $\rho_{n+1}+(n-1)/2$ $(=\sigma_{n+1}+iv_{n+1})$ is contained as squared in ${}^bA$ and ${}^bB$ and as linear in ${}^bC$.

The UIR's of SO$(n,1)$ are determined by the requirement of the hermiticity for D_{n+1n} and summarized in Tables I and II with slight modifications of the notation form Ref. 5. It is noted that we may replace σ_{n+1} by $-\sigma_{n+1}$ in Tables I and II because σ_{n+1} is contained as squared in (3.8).

The second-order Casimir operator is

$$F^{(n,1)} = \sum_{j>k}^{n} D_{jk}^2 - \sum_{j}^{n} D_{n+1j}^2, \tag{3.12}$$

whose eigenvalue with respect to the bases (3.9) is given by[17]

$$F^{(n,1)} = \rho_{n+1}(\rho_{n+1}+n-1)$$

$$+ \sum_{j=2}^{[(n+1)/2]} m_{n+1j}(m_{n+1j}+n+1-2j). \tag{3.13}$$

It is noted that (3.13) is the same form as (3.8) in which m_{n+11} is replaced by ρ_{n+1}

TABLE I. The UIR's of SO$(n,1)$ for n even.

Representation and conditions for $m_{n+12}, m_{n+13}, \ldots, m_{n+1n/2}$ and ρ_{n+1}	SO(n) content
(i) $D(m_{n+12},\ldots,m_{n+1n/2};v_{n+1})$ $m_{n+1j}=0,\frac{1}{2},1,\ldots,$ for $2\leqslant j\leqslant n/2$ $\rho_{n+1}=(1-n)/2+iv_{n+1};\quad 0<v_{n+1}$	$\lvert m_{nn/2}\rvert \leqslant m_{n+1n/2}\leqslant\cdots\leqslant m_{n+12}\leqslant m_{n1}$
(ii) $D^0(m_{n+12},\ldots,m_{n+1n/2};\sigma_{n+1})$ $m_{n+1j}=1,2,\ldots,$ for $2\leqslant j\leqslant n/2$ $\rho_{n+1}=(1-n)/2+\sigma_{n+1}; 0\leqslant\sigma_{n+1}<\frac{1}{2}$	$\lvert m_{nn/2}\rvert \leqslant m_{n+1n/2}\leqslant\cdots\leqslant m_{n+12}\leqslant m_{n1}$
(iii) $D^0(m_{n+12},\ldots,m_{n+1n/2};\sigma_{n+1})$ $m_{n+1j}=1,2,\ldots,$ for $2\leqslant j\leqslant n/2$ $\rho_{n+1}=(1-n)/2+\sigma_{n+1};\sigma_{n+1}=\frac{1}{2}$	$m_{nn/2}=0,\quad m_{n+1n/2}\leqslant m_{n(n-2)/2}$ $\leqslant\cdots\leqslant m_{n+12}\leqslant m_{n1}$
(iv) $D^k(m_{n+12},\ldots,m_{n+1(n/2-k)};\sigma_{n+1})$ $1\leqslant k\leqslant(n-2)/2$ $m_{n+1(n/2-j+1)} = \begin{cases} 0 \text{ for } 1\leqslant j\leqslant k \\ 1,2,\ldots, \text{ for} \\ k+1\leqslant j\leqslant(n-2)/2 \end{cases}$ $\rho_{n+1}=(1-n)/2+\sigma_{n+1}; 0\leqslant\sigma_{n+1}<k+\frac{1}{2}$	$m_{n(n/2-j+1)}=0$ for $1\leqslant j\leqslant k$ $0\leqslant m_{nn/2-k}\leqslant m_{n+1n/2-k}\leqslant\cdots$ $\leqslant m_{n+12}\leqslant m_{n1}$
(v) $D^k(m_{n+12},\ldots,m_{n+1(n/2-k)};\sigma_{n+1})$ $1\leqslant k\leqslant(n-4)/2$ $m_{n+1(n/2-j+1)} = \begin{cases} 0 \text{ for } 1\leqslant j\leqslant k \\ 1,2,\ldots, \text{ for} \\ k+1\leqslant j\leqslant(n-2)/2 \end{cases}$ $\rho_{n+1}=(1-n)/2+\sigma_{n+1};\sigma_{n+1}=k+\frac{1}{2}$	$m_{n(n/2-j+1)}=0$ for $1\leqslant j\leqslant k+1$ $m_{n+1(n/2-k)}\leqslant m_{n(n/2-k-1)}\leqslant\cdots$ $\leqslant m_{n+12}\leqslant m_{n1}$
(vi) $D^{\pm}(m_{n+12},\ldots,m_{n+1n/2};\sigma_{n+1})$ $m_{n+1(n/2-j+1)} = \begin{cases} \frac{1}{2},1,\frac{3}{2},\cdots \text{ for} \\ 1\leqslant j\leqslant(n-2)/2 \end{cases}$ $\rho_{n+1}=(1-n)/2+\sigma_{n+1};\sigma_{n+1}=0,\frac{1}{2},1,\cdots$	$\sigma_{n+1}+\frac{1}{2}\leqslant\pm m_{nn/2}\leqslant m_{n+1n/2}\leqslant\cdots$ $\leqslant m_{n+12}\leqslant m_{n1}$

TABLE II. The UIR's of SO(n,1) for n odd.

Representation and conditions for $m_{n+12}, m_{n+13}, \dots, m_{n+1})/2$ and ${}_{1}\rho_{n+1}$	SO(n) content
(i) $D(m_{n+12},\dots,m_{n+1(n+1)/2};\nu_{n+1})$ $m_{n+1[(n+1)/2-j+1]} = 0,1/2,1,\cdots,$ for $1 \leqslant j \leqslant (n-1)/2$ $\rho_{n+1} = (1-n)/2 + i\nu_{n+1}; 0 \leqslant \nu_{n+1}$	$\lvert m_{n+1(n+1)/2} \rvert \leqslant m_{n(n-1)/2} \leqslant \cdots$ $\leqslant m_{n+12} \leqslant m_{n1}$
(ii) $D^{k}(m_{n+12},\dots,m_{n+1[(n+1)/2-k]};\sigma_{n+1})$ $1 \leqslant k \leqslant (n-1)/2$; $m_{n+1[(n+1)/2-j+1]} = \begin{cases} 0 \text{ for } 1 \leqslant j \leqslant k \\ 1,2,\dots, \text{ for} \\ k+1 \leqslant j \leqslant (n-1)/2 \end{cases}$ $\rho_{n+1} = (1-n)/2 + \sigma_{n+1}; 0 < \sigma_{n+1} < k$	$m_{n[(n+1)/2-j]} = 0$ for $1 \leqslant j \leqslant k-1$ $0 \leqslant m_{n[(n+1)/2-k]} \leqslant m_{n+1[(n+1)/2-k]} \leqslant \cdots$ $\leqslant m_{n+12} \leqslant m_{n1}$
(iii) $D^{k}(m_{n+12},\dots,m_{n+1[(n+1)/2-k]};\sigma_{n+1})$ $1 \leqslant k \leqslant (n-3)/2$ $m_{n+1[(n+1)/2-j+1]} = \begin{cases} 0 \text{ for } 1 \leqslant j \leqslant k \\ 1,2,\dots, \text{ for} \\ k+1 \leqslant j \leqslant (n-1)/2 \end{cases}$ $\rho_{n+1} = (1-n)/2 + \sigma_{n+1}; \sigma_{n+1} = k$	$m_{n[(n+1)/2-j]} = 0$ for $1 \leqslant j \leqslant k$ $0 \leqslant m_{n+1[(n+1)/2-k]} \leqslant m_{n+1[(n-1)/2-k]} \leqslant \cdots$ $\leqslant m_{n+12} \leqslant m_{n1}$

4. REPRESENTATION MATRIX ELEMENTS OF SO(n)

In this section, the properties of the representation matrix elements of SO(n) are summarized.[2,10]

The representation D matrix elements of SO(n) are calculated by sandwiching (2.6a) between the bases (3.1), i.e.,

$$D^{(\lambda_n)}_{\{\lambda'_{n-1}\}\{\lambda_{n-1}\}}(\{\theta_n\}) = \langle \lambda_n \{\lambda'_{n-1}\} | D(g^{(n)}) | \lambda_n \{\lambda_{n-1}\} \rangle, \tag{4.1}$$

where the notation $\{\lambda_j\}$ stands for $(\lambda_j, \lambda_{j-1}, \dots, \lambda_2)$ and $\{\theta_n\} = (\theta_{21}, \theta_{31}, \theta_{32}, \dots, \theta_{n1}, \dots, \theta_{nn-1})$ are sometimes used instead of $g^{(n)}$ in order to show the arguments explicitly. The corresponding D matrix elements of SO(n,1) are given in a similar form to (4.1) by using the representation matrix of SO(n,1) and the bases (3.9). It therefore follows that in order to calculate the D matrix elements it is sufficient for us to know the d matrix elements (the boost matrix elements) provided that those of SO($n-1$) [SO(n)] are known. The d matrix elements of SO(n) and the boost matrix elements of SO($n-1$,1) are defined as follows:

$$d^{(\lambda_n)}_{\lambda'_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta) = \langle \lambda_n \lambda'_{n-1} \{\lambda_{n-2}\} | R_{nn-1}(\theta) | \lambda_n \{\lambda_{n-1}\} \rangle, \tag{4.2}$$

$$^{b}d^{(\Lambda_{n-2},\rho_n)}_{\lambda'_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\zeta) = \langle (\Lambda_{n-2},\rho_n); \lambda'_{n-1} \{\lambda_{n-2}\} | R^{b}_{nn-1}(\zeta) | \times | (\Lambda_{n-2},\rho_n); \{\lambda_{n-1}\} \rangle, \tag{4.3}$$

where $\Lambda_{n-2} \equiv (m_{n2}, m_{n3}, \dots, m_{n[n/2]})$. It will be difficult to calculate (4.2) and (4.3) directly by using the matrix elements (3.6) and (3.10).

The orthogonality and the completeness relations for the D matrix elements of SO(n) are given as follows[2,10]:

$$\int_{SO(n)} dV_n \overline{D^{(\lambda_n)}_{\{\lambda_{n-1}\}\{\lambda'''_{n-1}\}}(g^{(n)})} D^{(\lambda'_n)}_{\{\lambda'_{n-1}\}\{\lambda''_{n-1}\}}(g^{(n)}) = \delta_{\{\lambda_n\}\{\lambda'_n\}} \delta_{\{\lambda''_{n-1}\}\{\lambda'''_{n-1}\}} \frac{V_n}{N(\lambda_n)}, \tag{4.4a}$$

$$\sum_{\lambda_n} \frac{N(\lambda_n)}{V_n} \sum_{\{\lambda_{n-1}\}\{\lambda'_{n-1}\}} \overline{D^{(\lambda_n)}_{\{\lambda_{n-1}\}\{\lambda'_{n-1}\}}(g^{(n)})} D^{(\lambda_n)}_{\{\lambda_{n-1}\}\{\lambda'_{n-1}\}}(g^{(n)\prime}) = \delta_{(n)}(\{\theta_n\},\{\theta'_n\}), \tag{4.4b}$$

where $\delta_{\{\lambda\}\{\lambda'\}}$ stands for a product of Kronecker δ's in the individual indices. The expression for $\delta_{(n)}(\cdot,\cdot)$ is given by

$$\delta_{(n)}(\{\theta_n\},\{\theta'_n\}) = \delta_{(n-1)}(\{\theta_{n-1}\},\{\theta'_{n-1}\}) \times \prod_{j=1}^{n-1} (\sin\theta_{nj})^{j-n+1} \delta(\theta_{nj} - \theta'_{nj}), \tag{4.5}$$

$$\delta_{(2)}(\{\theta_2\},\{\theta'_2\}) = \delta(\theta_{21} - \theta'_{21}).$$

The following relations are obtained from (4.4) by taking into account (2.6a), (4.1), and (4.2):

$$\sum_{\{\lambda_{n-2}\}} \int_{S_n} dS_n \overline{H^{(\lambda_n)}_{\{\lambda_{n-1}\}\{\lambda'_{n-1}\}}(S_n)} H^{(\lambda'_n)}_{\{\lambda_{n-1}\}\{\lambda''_{n-1}\}}(S_n) = \delta_{\{\lambda'_{n-1}\}\{\lambda''_{n-1}\}} \delta_{\lambda_n \lambda'_n} \frac{V_n N(\lambda_{n-1})}{V_{n-1} N(\lambda_n)}, \tag{4.6a}$$

$$\sum_{\lambda_{n-2}} N(\lambda_{n-2}) \int_0^{\pi} d\theta \sin^{n-2}\theta \overline{d^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta)} \times d^{(\lambda'_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta)$$

$$= \delta_{\lambda_n \lambda_n'} \frac{\sqrt{\pi}\,\Gamma((n-1)/2) N(\lambda_{n-1}) N(\lambda'_{n-1})}{\Gamma(n/2) N(\lambda_n)}, \quad (4.6b)$$

$$\sum_{\lambda_n} N(\lambda_n) \overline{d^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta)}\, d^{(\lambda_n)}_{\lambda_{n-1}(\lambda'_{n-2})\lambda'_{n-1}}(\theta')$$

$$= \delta_{\lambda_{n-2}\lambda'_{n-2}} [\delta(\theta-\theta')\sqrt{\pi}\,\Gamma((n-1)/2) N(\lambda_{n-1}) \times N(\lambda'_{n-1})] [(\sin\theta)^{n-2} \Gamma(n/2) N(\lambda_{n-2})]^{-1}. \quad (4.6c)$$

(4.6a) and (4.6b) represent the orthogonality relations of the $H(S_n)$ and the d matrix elements, and (4.6c) represents the completeness relation of the d matrix elements.

The D matrix elements (4.1) can be taken as our bases in the representation space for SO(n). It follows from (4.4a) that the orthonormal bases relative to the scalar product defined as in (4.4a) are given by

$$\Psi^{(\lambda_n)}_{\{\lambda_{n-1}\}}(g^{(n)}) = \left(\frac{N(\lambda_n)}{V_n}\right)^{1/2} D^{(\lambda_n)}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \quad (4.7)$$

where $\{\lambda^0_{n-1}\}$ is fixed to some possible value. In fact, if we take the right shifts $g^{(n)} \to g^{(n)} g_1^{(n)}$ and the action of $D^{(n)}(g_1^{(n)})$ to the bases (4.7) as follows,

$$D^{(n)}(g_1^{(n)}) \Psi^{(\lambda_n)}_{\{\lambda_{n-1}\}}(g^{(n)}) = \Psi^{(\lambda_n)}_{\{\lambda_{n-1}\}}(g^{(n)} g_1^{(n)}), \quad (4.8)$$

then it follows that the representation becomes unitary and the following relation holds,

$$D^{(\lambda_n)}_{\{\lambda'_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) = \langle \Psi^{(\lambda_n)}_{\{\lambda'_{n-1}\}}, D^{(n)}(g^{(n)}) \Psi^{(\lambda_n)}_{\{\lambda_{n-1}\}} \rangle, \quad (4.9)$$

where the right-hand side means the scalar product such as (4.4a).

It follows from (2.9b) that the bases (4.7) are characterized and classified by the numbers λ_n and $\{\lambda_{n-1}\}$, i.e., the group chain of the second parameter group SO(n) $\supset$SO($n-1$)$\supset\cdots\supset$SO(2). From (2.9a), $\{\lambda^0_{n-1}\}$ are characterized by the eigenvalues of the invariants formed from $\bar{J}_{jk}$, i.e., the group chain of the first parameter group SO(n) $\supset$SO($n-1$)$\supset\cdots\supset$SO(2). Because of (2.9c), $\{\lambda^0_{n-1}\}$ can be regarded as the numbers characterizing the bases (4.7). It is clear from (2.9b) that the action of J_{jj+1} on (4.7) and their matrix elements are given by (3.5) and (3.6).

5. DIFFERENTIAL OPERATOR OF THE MULTIPLIER REPRESENTATION

In this section, a differential operator of the representation $R^b_{n+1n}(\zeta)$ corresponding to ${}^b t^{(n+1)}_{n+1n}(\zeta)$ is introduced and a multiplier is determined by requiring the commutation relations of SO(n,1) for the differential operators.[1]

Let E_n be a linear space consisting of the functions $f(g^{(n)}), g^{(n)} \in$ SO(n). Since we know the transformation of the Euler parameters of SO(n) under the boost ${}^b t^{(n+1)}_{n+1n}(\zeta)$ as in (2.21), the action of the representation operator $R^b_{n+1n}(\zeta)$ on the function $f(g^{(n)})$ may be written in the following form with a multiplier,[2,18]

$$R^b_{n+1n}(\zeta) f(g^{(n)}) = \frac{\alpha(\theta_{n1})}{\alpha(\theta'_{n1})} f(\{\theta_{n-1}\}, (\theta'_{n1}, \theta_{n2}, \ldots, \theta_{nn-1})), \quad (5.1)$$

where θ'_{n1} is given by (2.21a) and $\alpha(\theta_{n1})$ is assumed to depend only on θ_{n1} because θ's except for θ_{n1} do not change under the boost ${}^b t^{(n+1)}_{n+1n}(\zeta)$. Then it is evident that the condition of the representation is satisfied;

$$R^b_{n+1n}(\zeta_1+\zeta_2) f(g^{(n)}) = R^b_{n+1n}(\zeta_1) R^b_{n+1n}(\zeta_2) f(g^{(n)}). \quad (5.2)$$

In order to determine $\alpha(\theta)$, we consider the infinitesimal transformation on ζ in (5.1). The differential operator (generator) J_{n+1n} of the representation $R^b_{n+1n}(\zeta)$ ($=\exp i\zeta J_{n+1n}$) is given from (5.1) as follows,

$$J_{n+1n} = -i\sin\theta_{n1}\left(\frac{\partial}{\partial\theta_{n1}} - \frac{1}{\alpha(\theta_{n1})}\frac{d\alpha(\theta_{n1})}{d\theta_{n1}}\right). \quad (5.3)$$

The differential operators of the second parameter group of SO(n) are given in (2.16b). We now require the commutation relations of SO(n,1) for J_{n+1n} in (5.3) and J_{jk} ($j,k \leqslant n$) in (2.16b). Putting

$$[J_{n+1n}, J_{nn-1}] = -iJ_{n+1n-+},$$
$$[J_{n+1n}, J_{n+1n-1}] = -iJ_{nn-1}, \quad (5.4)$$

we obtain from the second relation by writting θ instead of θ_{n1};

$$\cos\theta \frac{d}{d\theta}\left(\frac{\sin\theta}{\alpha(\theta)}\frac{d\alpha(\theta)}{d\theta}\right) = \sin\theta \frac{d^2}{d\theta^2}\left(\frac{\sin\theta}{\alpha(\theta)}\frac{d\alpha(\theta)}{d\theta}\right) \quad (5.5)$$

This gives for $\alpha(\theta)$,

$$\alpha(\theta) = c_2(\sin\theta)^{\rho_{n+1}}\left(\tan\frac{\theta}{2}\right)^{c_1}, \quad (5.6)$$

where c_1, c_2, and ρ_{n+1} are some constants. Then the expression for J_{n+1n} becomes

$$J_{n+1n} = -i\left(\sin\theta_{n1}\frac{\partial}{\partial\theta_{n1}} - \rho_{n+1}\cos\theta_{n1} - c_1\right). \quad (5.7)$$

We may put $c_1 = 0$ without loss of generality, because the c_1 term gives a multiplicative factor to the representation matrix of $R^b_{n+1n}(\zeta)$ and we are dealing with the representation of the special group SO(n,1). Therefore, we obtain

$$\alpha(\theta) = c_2(\sin\theta)^{\rho_{n+1}}, \quad (5.8a)$$

$$J_{n+1n} = -i\left(\sin\theta_{n1}\frac{\partial}{\partial\theta_{n1}} - \rho_{n+1}\cos\theta_{n1}\right), \quad (5.8b)$$

where ρ_{n+1} may be a complex number.

We calculate J_{n+1j} ($j = 1,2,\ldots n-1$) from the commutation relation

$$J_{n+1j} = i[J_{n+1n}, J_{nj}]. \quad (5.9)$$

Then it is straightforward to show that J_{jk} ($j,k = 1,2,\ldots,n+1$) satisfies the commutation relations of

$SO(n,1)$. It is noted that J_{n+1n} does not commute with $\overline{J}_{nj}$ but commutes with $\overline{J}_{jk}$ $(j,k \leqslant n-1)$.

It is known that in a UIR of $SO(n,1)$ an irreducible representation of $SO(n,1)$ is contained either with multiplicity one or not at all[3] and the D matrix elements of $SO(n)$ constitute a complete set in L^2 space.[2] As in the case for the representation of the Lie algebra of $SO(n,1)$, the representation of $SO(n,1)$ may be treated in the space (E_n) consisting of the D matrix elements of $SO(n)$.

6. UNITARY REPRESENTATIONS OF SO(*n*, *1*)

In this section, the unitary representations of $SO(n,1)$ are discussed in a space consisting of the representation D matrix elements of $SO(n)$.[1]

A. Principal series

Let us consider a linear space (E_n) consisting of the D matrix elements of the UIR of $SO(n)$ and define a scalar product as follows,

$$\langle \Phi_1,\Phi_2\rangle = \int_{SO(n)} dV_n \, \overline{\Phi_1(g^{(n)})}\, \Phi_2(g^{(n)}), \tag{6.1}$$

where Φ_1, $\Phi_2 \in E_n$ and $\langle \Phi,\Phi\rangle < \infty$ is assumed for $\Phi \in E_n$, i.e., E_n becomes a Hilbert space. The space E_n may be regarded as the same one as that in Sec. 5 because the D matrix elements of $SO(n)$ constitute the complete orthogonal set in E^2_n.

It follows from (5.1) and (5.8a) that the action of the representation operator $R^b_{n+1n}(\zeta)$ on $\Phi \in E_n$ becomes

$$R^b_{n+1n}(\zeta)\Phi(g^{(n)}) = (\cosh\zeta - \cos\theta_{n1}\sinh\zeta)^{\rho_{n+1}} \times \Phi(\{\theta_{n-1}\},(\theta'_{n1},\theta_{n2},\dots,\theta_{nn-1})), \tag{6.2}$$

where θ'_{n1} is given by (2.21a). Then it is easily shown by using (2.22) that the unitarity condition relative to the scalar product (6.1) is satisfied only for

$$\rho_{n+1} = (1-n)/2 + i\nu_{n+1}, \ \nu_{n+1} \text{ real, i.e.,}$$

$$\langle R^b_{n+1n}(\zeta)\Phi_1, R^b_{n+1n}(\zeta)\Phi_2\rangle = \langle \Phi_1,\Phi_2\rangle. \tag{6.3}$$

It therefore follows that the unitary representation of $SO(n,1)$ relative to the scalar product (6.1) is realized for $\rho_{n+1} = (1-n)/2 + i\nu_{n+1}$. It is expected that this representation corresponds to the principal series of the UIR's given in Tables I and II.

As is seen from (4.7) and the remarks given at the last part of Sec. 5, the orthonormal bases of the UIR's of $SO(n,1)$ will be given by

$$\Phi^{(\Lambda_{n-1},\rho_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = N(\nu_{n+1};\lambda_n)\left(\frac{N(\lambda_n)}{V_n}\right)^{1/2} D^{(\lambda_n)}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \tag{6.4}$$

where $\Lambda_{n-1} = (m_{n+12}, m_{n+13},\dots,m_{n+1[(n+1)/2]})$ and a number ν_{n+1}, which characterize the irreducible representation of $SO(n,1)$, are introduced. At first sight, it seems that Λ_{n-1} is not contained in the right-hand side, but Λ_{n-1} is connected with λ^0_{n-1} as shown below. $N(\nu_{n+1};\lambda_n)$ is some constant with unit magnitude which depends on ν_{n+1} and λ_n, and the normalization of (6.4) gives, due to (4.4a),

$$|N(\nu_{n+1};\lambda_n)|^2 = 1. \tag{6.5}$$

It follows from (2.9b) that the action of J_{jj-1} $(j \leqslant n)$ on the base (6.4) is given by the same formulas as (3.5) and their matrix elements by (3.6). As is noted in the preceding section, the commutation relations of J_{n+1n} with J_{jk} are the same as those of D_{n+1n} with D_{jk} in Sec. 3. As the action of D_{n+1n} on the Gel'fand–Tsetlin bases defined in Sec. 3 and their matrix elements are determined by the commutation relations alone,[4,5] the action of J_{n+1n} on (6.4) and their matrix elements must be given by the same form as (3.10) and (3.11), i.e.,

$$J_{nn+1}\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = \sum_{j=1}^{n/2} {}^bA(m_{nj})\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{(\lambda^+_{nj}\{\lambda_{n-1}\})}(g^{(n)}) - \sum_{j=1}^{n/2} {}^bA(m_{nj}-1)\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{(\lambda^-_{nj}\{\lambda_{n-1}\})}(g^{(n)}), \tag{6.6a}$$

for $n+1$ odd, and

$$J_{nn+1}\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = \sum_{j=1}^{(n-1)/2} {}^bB(m_{nj})\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{(\lambda^+_{nj}\{\lambda_{n-1}\})}(g^{(n)}) - \sum_{j=1}^{(n-1)/2} {}^bB(m_{nj}-1)\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{(\lambda^-_{nj}\{\lambda_{n-1}\})}(g^{(n)}) + {}^bC_{n+1}\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{\{\lambda_n\}}(g^{(n)}), \tag{6.6b}$$

for $n+1$ even. $\lambda^{\pm}_{nj}$ stand for $(m_{n1},m_{n2},\dots,m_{nj-1},m_{nj}\pm 1,m_{nj+1},\dots,m_{n[n/2]})$ and the matrix elements ${}^bA$, ${}^bB$, and ${}^bC$ are given by (3.11) with $\sigma_{n+1} = 0$. Of course, these matrix elements are defined by the scalar product (6.1).

In order to obtain the relation between $(\Lambda_{n-1},\nu_{n+1})$ and $(\{\lambda^0_{n-1}\},\nu'_{n+1})$, which is characterized by the group chain of the first parameter group $SO(n-1) \supset SO(n-2) \supset \cdots \supset SO(2)$ and a number ν'_{n+1} contained in J_{n+1n}, we calculate the second-order Casimir operator by taking into account (5.8b), (5.9), (2.15), and (2.16b). By a straightforward procedure, we get

$$[\rho'_{n+1} = (1-n)/2 + i\nu'_{n+1}],$$

$$F^{(n,1)} = \sum_{j>k}^{n} J^2_{jk} - \sum_{j}^{n} J^2_{n+1j} = \rho'_{n+1}(\rho'_{n+1} + n - 1) + \sum_{j>k}^{n-1} \overline{J}^2_{jk}. \tag{6.7}$$

Taking into account (2.9a), (3.7), and (3.8), we obtain the eigenvalue with respect to (6.4),

$$F^{(n,1)} = \rho'_{n+1}(\rho'_{n+1} + n - 1) + \sum_{j=1}^{[(n-1)/2]} m^0_{n-1j}(m^0_{n-1j} + n - 1 - 2j)$$

$$= \rho'_{n+1}(\rho'_{n+1} + n - 1) + \sum_{j=2}^{[(n+1)/2]} m^0_{n-1j-1} \times (m^0_{n-1j-1} + n + 1 - 2j). \qquad (6.8)$$

On the other hand, as in (4.11) and (4.12), the eigenvalue of $F^{(n,1)}$ can be calculated from the first expression of (6.7). The matrix elements in (6.6) together with the other matrix elements are given by the numbers Λ_{n-1}, i.e., m_{n+1j} ($j \geqslant 2$) and a number ρ_{n+1} [$= (1-n)/2 + i\nu_{n+1}$]. As is seen from (3.12) and (3.3), we must have

$$F^{(n,1)} = \rho_{n+1}(\rho_{n+1} + n - 1) + \sum_{j=2}^{[(n+1)/2]} m_{n+1j}(m_{n+1j} + n + 1 - 2j). \qquad (6.9)$$

On comparing (6.8) and (6.9) and considering the general validity of these relations, we may get

$$\rho'_{n+1} = \rho_{n+1}, \quad m^0_{n-1j-1} = m_{n+1j} \quad (j = 2,3,\ldots,[(n+1)/2]). \qquad (6.10)$$

It is necessary to use a complete set of invariants of SO(n,1)[19] in order to complete the above discussion but it is hard to express directly the invariants given by the differential operators of the second parameter group in terms of those of the first parameter group such as (6.7). However, the validity of the above results is confirmed in the following.

It therefore follows that the irreducible representations of SO(n,1) are characterized by the numbers m_{n+1j} ($= m^0_{n-1j-1}$) and a complex number ρ_{n+1} as expected, and the bases (6.4) may be written in a form

$$\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = N(\nu_{n+1};\lambda_n)\left(\frac{N(\lambda_n)}{V_n}\right)^{1/2} D^{(\lambda_n)}_{\{\Lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \qquad (6.11)$$

where $\{\Lambda_{n-1}\} = (\Lambda_{n-1},\Lambda_{n-2},\ldots,\Lambda_2)$ and $\Lambda_j = (m^0_{j1},m^0_{j2},\ldots,m^0_{j[j/2]})$ for $j < n-2$ with some fixed m^0_{jk}.

Next in order to determine the constant $N(\nu_{n+1};\lambda_n)$, we consider the matrix elements of J_{n+1n}, whose action on the bases (6.11) is given by (6.6) with $\sigma_{n+1} = 0$. we first consider the case of $n+1$ odd. From (6.6a) and (6.6b) with its normalization, we obtain

$${}^bA(m_{nj}) = \langle \Phi^{(\Lambda_{n-1},\nu_{n+1})}_{(\lambda^+_{nj}\{\lambda_{n-1}\})}, J_{nn+1}\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{\{\lambda_n\}}\rangle. \qquad (6.12)$$

The right-hand side is rewritten by using (5.8b), (4.4a), and (4.6a) as follows,

$${}^bA(m_{nj}) = -i\frac{N(\nu_{n+1};\lambda_n)}{N(\nu_{n+1};\lambda^+_{nj})} \times \left[\left(\frac{1-n}{2} + i\nu_{n+1}\right) I^{(\lambda^+_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}} + T^{(\lambda^+_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}\right], \qquad (6.13)$$

where

$$I^{(\lambda^+_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}} = \frac{(V_{n-1})^2}{V_nV_{n-2}}\frac{[N(\lambda^+_{nj})N(\lambda_n)]^{1/2}}{N(\Lambda_{n-1})N(\lambda_{n-1})} \times \sum_{\lambda_{n-2}} N(\lambda_{n-2})\int_0^\pi d\theta \sin^{n-2}\theta\cos\theta \times \overline{d^{(\lambda^+_{nj})}_{\Lambda_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta)}\, d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta),$$

$$T^{(\lambda^+_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}} = -\frac{(V_{n-1})^2}{V_nV_{n-2}}\frac{[N(\lambda^\pm_{nj})N(\lambda_n)]^{1/2}}{N(\Lambda_{n-1})N(\lambda_{n-1})} \times \sum_{\lambda_{n-2}} N(\lambda_{n-2})\int_0^\pi d\theta \sin^{n-2}\theta\cos\theta \times \overline{d^{(\lambda^+_{nj})}_{\Lambda_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta)}\,\frac{d}{d\theta} d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta).$$

It is noted that these I and T do not depend on ν_{n+1}.

Taking into account the matrix element of J_{nn+1}, i.e., (3.11a), we can rewrite (6.13) in the form

$$\frac{N(\nu_{n+1};\lambda^+_{nj})}{N(\nu_{n+1};\lambda_n)} = [(1-n)/2 + i\nu_{n+1} + I_1]I_2 \times \{[m_{nj} - i\nu_{n+1} + (n+1)/2 - j] \times [m_{nj} + i\nu_{n+1} + (n+1)/2 - j]\}^{-1/2} \qquad (6.14)$$

where the quantities I_1 and I_2 are independent of ν_{n+1}. As the magnitude on the left-hand side of (6.14) is unity due to (6.5) and the two quantities in the denominator on the right-hand side are complex conjugate to each other, we thus obtain two possibilities for the solution for $m'_{nj} = m_{nj} + 1$ and $m'_{nl} = m_{nl}$ ($l \neq j$).

Case 1:

$$\frac{N(\nu_{n+1};\lambda^+_{nj})}{N(\nu_{n+1};\lambda_n)} = \epsilon\left(\frac{m_{nj} - i\nu_{n+1} + (n+1)/2 - j}{m_{nj} + i\nu_{n+1} + (n+1)/2 - j}\right)^{1/2}. \qquad (6.15a)$$

$$(l_{nj}^2 + \nu_{n+1}^2)^{1/2} B'(m_{nj}) = {}^bB(m_{nj}), \qquad (6.24b)$$

$$\nu_{n+1} C'_{n+1} = {}^bC_{n+1}. \qquad (6.24c)$$

In this way, we can give many integral formulas containing the d matrix elements. The formulas (6.21)–(6.23) will be used in the discussion of the complementary series.

B. Complementary series

Here we consider the case of $\nu_{n+1} = 0$ $[\rho_{n+1} = (1-n)/2 + \sigma_{n+1}]$ but the unitary representation.[1]

The bases and the action of the representation on the bases are the same as in (A), i.e.,

$$\Phi^{(\Lambda_{n+1},\sigma_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = N(\sigma_{n+1};\lambda_n)\left(\frac{N(\lambda_n)}{V_n}\right)^{1/2} D^{(\lambda_n)}_{\{\Lambda_{n+1}\}\{\lambda_{n-1}\}}(g^{(n)}), \qquad (6.25a)$$

$$R^b_{n+1n}(\zeta)\Phi^{(\Lambda_{n+1},\sigma_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = [(\cosh\zeta - \cos\theta_{n1}\sinh\zeta)^{\rho_{n+1}}] \times \Phi^{(\Lambda_{n+1},\sigma_{n+1})}_{\{\lambda_n\}}(\{\theta_{n-1}\},(\theta'_{n1},\theta_{n2},\ldots,\theta_{nn-1})), \qquad (6.25b)$$

where $N(\sigma_{n+1};\lambda_n)$ is a constant to be determined. It is clear that the representation condition is satisfied for ρ_{n+1} but the unitarity condition is not satisfied for $\sigma_{n+1} \neq 0$ under the scalar product (6.1). We must, therefore, find a scalar product relative to which the unitarity condition holds for $\sigma_{n+1} \neq 0$.

According to Ref. 1, we define the following scalar product,

$$\langle\Phi_1,\Phi_2\rangle_c = \int_{SO(n)}\int_{SO(n)} dV'_n dV_n [1 - (g^{(n)'}g^{(n)-1})_{nn}]^{(1-n)/2-\sigma_{n+1}} \times K(g^{(n)'}g^{(n)-1})\overline{\Phi_1(g^{(n)'})}\Phi_2(g^{(n)}), \qquad (6.26)$$

where $\Phi_1, \Phi_2 \in E'_n$ the space consisting of the bases (6.25a) and having the scalar product (6.26). The quantity $[1 - (g^{(n)'}g^{(n)-1})_{nn}]$ is considered to be a generalization of the kernel in the case of the Lorenz group[20] and depends on the Euler parameters as given in (2.4a). The function K is introduced according to Ref. 1 and we assume that the function K exists and does not depend on the parameters θ_{n1} and θ'_{n1}. It will always be possible to construct such functions because we have n^2 elements of $g^{(n)'}g^{(n)-1}$ and only $n(n-1)$ parameters. However, in what follows it is sufficient for us to assume the existence of such a function.

The normalization of the bases (6.25a) relative to the scalar product (6.26) is given by

$$\langle\Phi^{(\Lambda_{n+1},\sigma_{n+1})}_{\{\lambda'_n\}},\Phi^{(\Lambda_{n+1},\sigma_{n+1})}_{\{\lambda_n\}}\rangle_c = \delta_{\{\lambda'_n\}\{\lambda_n\}}, \qquad (6.27)$$

which gives the condition

$$\int_{SO(n)} dV_n(1-\cos\theta_{n1})^{(1-n)/2-\sigma_{n+1}} K(g^{(n)}) \times \overline{D^{(\lambda_n)}_{\{\Lambda_{n+1}\}\{\Lambda_{n+1}\}}}(g^{(n)}) = |N(\sigma_{n+1};\lambda_n)|^{-2}, \qquad (6.28)$$

where (2.4b) is used and it is taken into account that K does not depend on the parameter θ_{n1} due to our assumption on K.

As the function K in (6.26) does not depend on the parameters θ_{n1} and θ'_{n1}, it is easy to see that the unitarity condition relative to the scalar product (6.26) is satisfied for $\nu_{n+1} = 0$, i.e.,

$$\langle R^b_{n+1n}(\zeta)\Phi_1, R^b_{n+1n}(\zeta)\Phi_2\rangle_c = \langle\Phi_1,\Phi_2\rangle_c, \qquad (6.29)$$

where $\Phi_1,\Phi_2 \in E'_n$. This can be shown as follows.

From (6.25) and (6.26), we have

$$\langle R^b_{n+1n}(\zeta)\Phi_1, R^b_{n+1n}(\zeta)\Phi_2\rangle_c = \int_{SO(n)}\int_{SO(n)} dV'_n dV_n [1 - (g^{(n)'}g^{(n)-1})_{nn}]^{(1-n)/2-\sigma_{n+1}} \times K(g^{(n)'}g^{(n)-1})(\cosh\zeta - \cos\theta'_{n1}\sinh\zeta)^{\bar\rho_{n+1}} \times \overline{\Phi_1(\{\theta'_{n-1}\},(\bar\theta'_{n1},\theta'_{n2},\ldots,\theta'_{nn-1}))}(\cosh\zeta - \cos\theta_{n1}\sinh\zeta)^{\rho_{n+1}}\Phi_2(\{\theta_{n-1}\},(\bar\theta_{n1},\theta_{n2},\ldots,\theta_{nn-1})), \qquad (6.30)$$

where

$$\cos\bar\theta'_{n1} = \frac{\cos\theta'_{n1}\cosh\zeta - \sinh\zeta}{\cosh\zeta - \cos\theta'_{n1}\sinh\zeta},$$

$$\cos\bar\theta_{n1} = \frac{\cos\theta_{n1}\cosh\zeta - \sinh\zeta}{\cosh\zeta - \cos\theta_{n1}\sinh\zeta}.$$

Changing the integration variables $\theta_{n1},\theta'_{n1} \to \bar\theta_{n1},\bar\theta'_{n1}$ and using (2.4a), (2.22), and the relation

$$1 - \cos\theta'_{n1}\cos\theta_{n1} - \sin\theta'_{n1}\sin\theta_{n1}(g^{(n-1)'}g^{(n-1)-1})_{n-1n-1} = [1 - \cos\bar\theta'_{n1}\cos\bar\theta_{n1} - \sin\bar\theta'_{n1}\sin\bar\theta_{n1}(g^{(n-1)'} \times g^{(n-1)-1})_{n-1n-1}](\cosh\zeta + \cos\bar\theta'_{n1}\sinh\zeta)^{-1} \times (\cosh\zeta + \cos\bar\theta_{n1}\sinh\zeta)^{-1},$$

we obtain

$$\langle R^b_{n+1n}(\zeta)\Phi_1, R^b_{n+1n}(\zeta)\Phi_2\rangle_c = \int_{SO(n)}\int_{SO(n)} dV'_n dV_n [1 - (g^{(n)'}g^{(n)-1})_{nn}]^{(1-n)/2-\sigma_{n+1}} \times (\cosh\zeta + \cos\theta'_{n1}\sinh\zeta)^{-i\nu_{n+1}}(\cosh\zeta + \cos\theta_{n1}\sinh\zeta)^{i\nu_{n+1}}, \times K(g^{(n)'}g^{(n)-1})\overline{\Phi_1(g^{(n)'})}\Phi_2(g^{(n)}). \qquad (6.31)$$

This gives (6.29) directly in the case of $\nu_{n+1} = 0$.

Thus we obtain the representation for $\rho_{n+1} = (1-n)/2 + \sigma_{n+1}$ which is unitary relative to the scalar product (6.26). It is evident that these representations given by the various choices of K correspond to the complementary series in Tables I and II because the matrix elements for J_{n+1n} are given by (3.11) with $\rho_{n+1} = (1-n)/2 + \sigma_{n+1}$

In order to determine the factor $N(\sigma_{n+1};\lambda_n)$ in (6.25a), we proceed in the same way as in Sec. 6 A. The matrix elements of J_{n+1n} will be given by (3.11) with $\nu_{n+1} = 0$ (cf. the next section),

$$ {}^bA(m_{nj}) = \langle \Phi^{(\Lambda_{n-1},\sigma_{n+1})}_{(\lambda'_{nj}\{\lambda_{n-1}\})}, J_{nn+1}\Phi^{(\Lambda_{n-1},\sigma_{n+1})}_{\{\lambda_n\}}\rangle_c, \tag{6.31a} $$

for $n+1$ odd and

$$ {}^bB(m_{nj}) = \langle \Phi^{(\Lambda_{n-1},\sigma_{n+1})}_{(\lambda'_{nj}\{\lambda_{n-1}\})}, J_{nn+1}\Phi^{(\Lambda_{n-1},\sigma_{n+1})}_{\{\lambda_n\}}\rangle_c, \tag{6.31b} $$

$$ {}^bC_{n+1} = \langle \Phi^{(\Lambda_{n-1},\sigma_{n+1})}_{\{\lambda_n\}}, J_{nn+1}\Phi^{(\Lambda_{n-1},\sigma_{n+1})}_{\{\lambda_n\}}\rangle_c, \tag{6.31c} $$

for $n+1$ even. Of course, the ρ_{n+1} in J_{nn+1} is given with $\nu_{n+1} = 0$.

We first consider the case of $n+1$ odd. Then (6.31a) gives

$$ {}^bA(m_{nj}) = -i\frac{N(\sigma_{n+1};\lambda_n)}{N(\sigma_{n+1};\lambda^{+}_{nj})}\Big[\Big(\frac{1-n}{2}+\sigma_{n+1}\Big) \times I^{(\lambda'_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}} + T^{(\lambda'_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}\Big], \tag{6.32} $$

where use has been made of the relations (4.4a), (4.6a), and (6.28), and I and T are defined below (6.13). We obtain, using (6.21) and (6.24),

$$ \frac{N(\sigma_{n+1};\lambda^{+}_{nj})}{N(\sigma_{n+1};\lambda_n)} = \Big(\frac{m_{nj}-\sigma_{n+1}+(n+1)/2-j}{m_{nj}+\sigma_{n+1}+(n+1)/2-j}\Big)^{1/2}. \tag{6.33} $$

For $n+1$ even, we obtain a similar result to (6.33) but we must have $\sigma_{n+1} = 0$ or $m_{n+1(n+1)/2} = 0$ due to the ${}^bC_{n+1}$ term in (6.6b). The first case of $\sigma_{n+1} = 0$ is contained in the principal series and then we adopt the second case of $m_{n+1}(n+1)/2 = 0$ in what follows. We, therefore, obtain the following expression for any λ_n and λ'_n as in 6 A,

$$ \frac{N(\sigma_{n+1};\lambda_n)}{N(\sigma_{n+1};\lambda'_n)} = \Big(\prod_{j=1}^{[n/2]}\Gamma(m_{nj}-\sigma_{n+1}+(n+1)/2-j) \times \Gamma(m'_{nj}+\sigma_{n+1}+(n+1)/2-j) \times [\Gamma(m_{nj}+\sigma_{n+1}+(n+1)/2-j) \times \Gamma(m'_{nj}-\sigma_{n+1}+(n+1)/2-j)]^{-1}\Big)^{1/2}. \tag{6.34} $$

It is noted that the expression (6.34) is intimately connected with (6.16). For the complementary series contrary to the principal series, we can not determine $N(\sigma_{n+1};\lambda_n)$ uniquely from (6.28) and (6.34), because only the product of $N(\sigma_{n+1};\lambda_n)$ and the function K is determined from (6.28). However we may choose, for instance, as follows

$$ N(\sigma_{n+1};\lambda_n) = \Big(\prod_{j=1}^{[n/2]}\frac{\Gamma(m_{nj}-\sigma_{n+1}+(n+1)/2-j)}{\Gamma(m_{nj}+\sigma_{n+1}+(n+1)/2-j)}\Big)^{1/2}. \tag{6.35} $$

Then it follows that the function K is fixed by (6.28). It is noted that we must determine $N(\sigma_{n+1};\lambda_n)$ from (6.34) in such a way that the singular terms from the gamma functions do not appear in their expression, for example, as (6.35).

The boost matrix elements of the representation are given by

$$ {}^bd^{(\Lambda_{n-1},\sigma_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta) = \langle \Phi^{(\Lambda_{n-1},\sigma_{n+1})}_{(\lambda'_n\{\lambda_{n-1}\})}, R^{\,b}_{n+1n}(\zeta)\Phi^{(\Lambda_{n-1},\sigma_{n+1})}_{\{\lambda_n\}}\rangle_c. \tag{6.36} $$

As in the case of (6 A), this becomes

$$ {}^bd^{(\Lambda_{n-1},\sigma_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta) = \frac{\Gamma(n/2)}{\sqrt{\pi}\,\Gamma((n-1)/2)}\frac{[N(\lambda'_n)N(\lambda_n)]^{1/2}}{N(\Lambda_{n-1})N(\lambda_{n-1})}\frac{N(\sigma_{n+1};\lambda_n)}{N(\sigma_{n+1};\lambda'_n)} \times \sum_{\lambda_{n-2}} N(\lambda_{n-2})\int_0^{\pi} d\theta\,\sin^{n-2}\theta\,\overline{d^{(\lambda'_n)}_{\Lambda_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta)} \times (\cosh\zeta - \cos\theta\sinh\zeta)^{(1-n)/2+\sigma_{n+1}} \times d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta'). \tag{6.37} $$

As is expected, (6.37) is intimately connected with (6.19) for the principal series. For $n+1$ even, $m_{n+1(n+1)/2}$ contained in Λ_{n-1} is zero.

Thus we could have given the matrix elements for the infinitesimal operator J_{nn+1} and the computation formulas for the representation matrix of all the UIR's of SO$(n,1)$.

7. MATRIX ELEMENTS IN ANY IRREDUCIBLE REPRESENTATIONS OF SO$(n,1)$

In this section, it is shown that the relations (6.10), which hold for the principal series and are used in the complementary series, are valid in general and the irreducible representations which may not necessarily be unitary are discussed.

We consider a linear space (E''_n) consisting of the D matrix elements of the UIR's of SO(n). The bases are taken as follows,

$$ \Phi^{(\Lambda_{n-1},\rho_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = N(\rho_{n+1};\lambda_n)\Big(\frac{N(\lambda_n)}{V_n}\Big)^{1/2} D^{(\lambda_n)}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \tag{7.1} $$

where ρ_{n+1} is some complex constant and

$$\frac{N(\rho_{n+1};\lambda_n)}{N(\rho_{n+1};\lambda'_n)} = [\sigma_{n+1} \rightarrow (n-1)/2 + \rho_{n+1} \text{ in } (6.34)]. \qquad (7.2)$$

If necessary, the factor $N(\rho_{n+1};\lambda_n)$ must be chosen in such a way that the singular terms from the gamma functions do not appear.

The action of the representation operator $R^{\,b}_{n+1n}(\zeta)$ on the bases is given by (5.1) with (5.8a) and the differential operator (generator) of the representation is given by (5.8b), i.e.,

$$J_{n+1n} = -i\left(\sin\theta_{n1}\frac{\partial}{\partial\theta_{n1}} - \rho_{n+1}\cos\theta_{n1}\right). \qquad (7.3)$$

(6.21), (6.22), and (6.23) give the following relations:

$$\cos\theta_{n1} D^{(\lambda_n)}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) = \sum_{j=1}^{n/2}\left(\frac{N(\lambda^+_{nj})}{N(\lambda_n)}\right)^{1/2} I^{(\lambda^+_{nj},\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda^+_{nj})}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) + \sum_{j=1}^{n/2}\left(\frac{N(\lambda^-_{nj})}{N(\lambda_n)}\right)^{1/2} I^{(\lambda^-_{nj},\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda^-_{nj})}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \qquad (7.4a)$$

$$-\sin\theta_{n1}\frac{\partial}{\partial\theta_{n1}} D^{(\lambda_n)}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) = \sum_{j=1}^{n/2}\left(\frac{N(\lambda^+_{nj})}{N(\lambda_n)}\right)^{1/2} T^{(\lambda^+_{nj},\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda^+_{nj})}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) + \sum_{j=1}^{n/2}\left(\frac{N(\lambda^-_{nj})}{N(\lambda_n)}\right)^{1/2} T^{(\lambda^-_{nj},\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda^-_{nj})}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \qquad (7.4b)$$

for n even, and

$$\cos\theta_{n1} D^{(\lambda_n)}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) = \sum_{j=1}^{(n-1)/2}\left(\frac{N(\lambda^+_{nj})}{N(\lambda_n)}\right)^{1/2} I^{(\lambda^+_{nj},\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda^+_{nj})}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) + \sum_{j=1}^{(n-1)/2}\left(\frac{N(\lambda^-_{nj})}{N(\lambda_n)}\right)^{1/2} I^{(\lambda^-_{nj},\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda^-_{nj})}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}} \times (g^{(n)}) + I^{(\lambda_n,\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda_n)}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \qquad (7.5a)$$

$$-\sin\theta_{n1}\frac{\partial}{\partial\theta_{n1}} D^{(\lambda_n)}_{\{\lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) = \sum_{j=1}^{(n-1)/2}\left(\frac{N(\lambda^+_{nj})}{N(\lambda_n)}\right)^{1/2} T^{(\lambda^+_{nj},\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda^+_{nj})}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) + \sum_{j=1}^{(n-1)/2}\left(\frac{N(\lambda^-_{nj})}{N(\lambda_n)}\right)^{1/2} T^{(\lambda^-_{nj},\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda^-_{nj})}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}} \times (g^{(n)}) + T^{(\lambda_n,\lambda_n)}_{\lambda^0_{n-1},\lambda_{n-1}} D^{(\lambda_n)}_{\{\lambda^0_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \qquad (7.5b)$$

for n odd. It is noted that these relations hold independent of the existence of scalar product for the bases (7.1).

Therefore, the action of (7.3) on (7.1) is easily obatined as follows,

$$J_{nn+1}\Phi^{(\Lambda_{n+1},\rho_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = \sum_{j=1}^{n/2} {}^bA(m_{nj})\Phi^{(\Lambda_{n+1},\rho_{n+1})}_{(\lambda^+_{nj}\{\lambda_{n-1}\})}(g^{(n)}) - \sum_{j=1}^{n/2} {}^bA(m_{nj}-1)\Phi^{(\Lambda_{n+1},\rho_{n+1})}_{(\lambda^-_{nj}\{\lambda_{n-1}\})}(g^{(n)}), \qquad (7.6a)$$

for $n+1$ odd and

$$J_{nn+1}\Phi^{(\Lambda_{n+1},\rho_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = \sum_{j=1}^{(n-1)/2} {}^bB(m_{nj})\Phi^{(\Lambda_{n+1},\rho_{n+1})}_{\lambda^+_{nj}\{\lambda_{n-1}\}}(g^{(n)}) - \sum_{j=1}^{(n-1)/2} {}^bB(m_{nj}-1)\Phi^{(\Lambda_{n+1},\rho_{n+1})}_{\lambda^-_{nj}\{\lambda_{n-1}\}}(g^{(n)}) + {}^bC_{n+1}\Phi^{(\Lambda_{n+1},\rho_{n+1})}_{\{\lambda_n\}}(g^{(n)}), \qquad (7.6b)$$

for $n+1$ even. The matrix elements ${}^bA$, ${}^bB$, and ${}^bC$ are given by (3.11) with the substitutions $m_{n+1j} = m^0_{n-1j-1}$ ($j = 2,3,\ldots,[(n+1)/2]$) and complex ρ_{n+1}. On the other hand, (6.7) holds for any ρ_{n+1} and the action of (6.7) on (7.1) gives (6.8), i.e.,

$$F^{(n,1)} = \rho_{n+1}(\rho_{n+1} + n - 1) + \sum_{j=2}^{[(n+1)/2]} m^0_{n-1j-1}(m^0_{n-1j-1} + n + 1 - 2j) \qquad (7.7)$$

From the above results, it follows even in the general cases that we can take $m^0_{n-1j-1} = m_{n+1j}$ ($j = 2,3,\ldots,[(n+1)/2]$) and ρ_{n+1} which characterize the irreducible representations of SO(n,1). The bases (7.1) may thus be written in the form

$$\Phi^{(\Lambda_{n+1},\rho_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = N(\rho_{n+1};\lambda_n)\left(\frac{N(\lambda_n)}{V_n}\right)^{1/2} D^{(\lambda_n)}_{\{\Lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}). \qquad (7.8)$$

The action of $R^{\,b}_{n+1n}(\zeta)$ on (7.8) is given by (5.1) with (5.8a), and the representation condition is satisfied for any ρ_{n+1}. Therefore, we may define the general matrix elements of $R^{\,b}_{n+1n}(\zeta)$ by taking (7.6) into account as follows, Case 2:

$$\frac{N(\nu_{n+1};\lambda_{nj}^{+})}{N(\nu_{n+1};\lambda_n)}=\epsilon'\left(\frac{m_{nj}+i\nu_{n+1}+(n+1)/2-j}{m_{nj}-i\nu_{n+1}+(n+1)/2-j}\right)^{1/2}. \tag{6.15b}$$

Here ϵ and ϵ' are some constants of unit magnitude which do not depend on ν_{n+1}. It follows that (6.15a) and (6.15b) are complex conjugate to each other up to a phase factor. We may, therefore, choose one of them as our solution without loss of generality, because we can start with (6.4) or its complex conjugate and the representations with $(\Lambda_{n-1},\nu_{n+1})$ and $(\Lambda'_{n-1},-\nu_{n+1})$ are unitarily equivalent, where $\Lambda'_{n-1}=(m_{n+12},m_{n+13},\ldots,m_{n+1n/2})$ for $n+1$ odd and $\Lambda'_{n-1}=(m_{n+12},\ldots,m_{n+1(n-1)/2},-m_{n+1(n+1)/2})$ for $n+1$ even. In what follows we adopt (6.15a).

Taking into account the restrictions (3.2), we obtain from (6.15a) for an λ_n and λ'_n,

$$\frac{N(\nu_{n+1};\lambda_n)}{N(\nu_{n+1};\lambda'_n)}=\epsilon\Big(\prod_{j=1}^{n/2}\{\Gamma(m_{nj}-i\nu_{n+1}+(n+1)/2-j)$$
$$\times\Gamma(m'_{nj}+i\nu_{n+1}+(n+1)/2-j)\}$$
$$\times\{\Gamma(m_{nj}+i\nu_{n+1}+(n+1)/2-j)$$
$$\times\Gamma(m'_{nj}-i\nu_{n+1}+(n+1)/2-j)\}^{-1}\Big)^{1/2}, \tag{6.16}$$

where ϵ is a phase factor independent of ν_{n+1}. Thus we could determine the ratio of the constants $N(\nu_{n+1};\lambda_n)$ up to a phase factor which we take equal to one. Then we may put, for example,

$$N(\nu_{n+1};\lambda_n)=\left(\prod_{j=1}^{[n/2]}\frac{\Gamma(m_{nj}-i\nu_{n+1}+(n+1)/2-j)}{\Gamma(m_{nj}+i\nu_{n+1}+(n+1)/2-j)}\right)^{1/2}. \tag{6.17}$$

Similarly, for $n+1$ even we can obtain the same formula as (6.17), in which $n/2$ on the Π is to be replaced with $[n/2]$, and thus (6.17) holds for both $n+1$ odd and even. In this way, we could determine the constant $N(\nu_{n+1};\lambda_n)$ as (6.17) and then our bases (6.4) are fixed completely.

The boost matrix elements of the representation $R^{b}_{n+1n}(\zeta)$, which will, of course, coincide with (4.3), are calculated through

$${}^{b}d^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta)=\langle\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{(\lambda'_n\{\lambda_{n-1}\})},R^{b}_{n+1n}(\zeta)\Phi^{(\Lambda_{n-1},\nu_{n+1})}_{\{\lambda_n\}}\rangle. \tag{6.18}$$

Making use of the relations (2.6a), (4.1), (4.4a), and (4.6a), we rewritte (6.18) into the form

$${}^{b}d^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta)=\frac{\Gamma(n/2)}{\sqrt{\pi}\,\Gamma((n-1)/2)}\frac{[N(\lambda'_n)N(\lambda_n)]^{1/2}}{N(\lambda_{n-1})N(\Lambda_{n-1})}\frac{N(\nu_{n+1};\lambda_n)}{N(\nu_{n+1};\lambda'_n)}$$
$$\times\sum_{\lambda_{n-2}}N(\lambda_{n-2})\int_0^{\pi}d\theta\,\sin^{n-2}\theta\,\overline{d^{(\lambda'_n)}_{\Lambda_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta)} \tag{6.19}$$
$$\times(\cosh\zeta-\cos\theta\sinh\zeta)^{(1-n)/2+i\nu_{n+1}}d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta'),$$

where θ' is given by (2.21a), i.e.,

$$\cos\theta'=\frac{\cos\theta\cosh\zeta-\sinh\zeta}{\cosh\zeta-\cos\theta\sinh\zeta}$$

and $N(\nu_{n+1};\lambda_n)/N(\nu_{n+1};\lambda'_n)$ is given by (6.16) with $\epsilon=1$.

From (6.19), we easily obtain the relations

$${}^{b}d^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(0)=\delta_{\lambda'_n\lambda_n}, \tag{6.20a}$$

$$\overline{{}^{b}d^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta)}={}^{b}d^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda_n(\lambda_{n-1})\lambda'_n}(-\zeta), \tag{6.20b}$$

$$\sum_{\lambda''_{n-1}}{}^{b}d^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda''_n}(\zeta_1)\,{}^{b}d^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda''_n(\lambda_{n-1})\lambda_n}(\zeta_2)$$
$$={}^{b}d^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta_1+\zeta_2). \tag{6.20c}$$

(6.20b) and (6.20c) show the unitarity and the representation conditions explicitly.

As a byproduct of the above discussions in (6.12)–(6.17), we obtain useful formulas for the integral containing the d matrix elements. We obtain from (6.14)

$$-I_1=m_{nj}+1-j,\quad -I_2=1.$$

Expressing I_1 and I_2 in terms of I and T and considering the corresponding expressions for $n+1$ even, we obtain

(i) $n+1$ odd, $m'_{nj}=m_{nj}\pm1$ and $m'_{nl}=m_{nl}$ for $l\neq j$,

$$I^{(\lambda^{+}_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}=-iA'(m_{nj}),$$
$$I^{(\lambda^{-}_{nj},\lambda_n)}_{\Lambda_{n-1},\lambda_{n-1}}=-iA'(m_{nj}-1), \tag{6.21a}$$
$$T^{(\lambda^{+}_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}=i(m_{nj}+1-j)A'(m_{nj}),$$
$$T^{(\lambda^{-}_{nj},\lambda_n)}_{\Lambda_{n-1},\lambda_{n-1}}=-i(m_{nj}+n-1-j)A'(m_{nj}-1). \tag{6.21b}$$

(ii) $n+1$ even,

(a) $m'_{nj}=m_{nj}$ for all j,

$$I^{(\lambda_n,\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}=C'_{n+1}, \tag{6.22a}$$
$$T^{(\lambda_n,\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}=\frac{n-1}{2}C'_{n+1}. \tag{6.22b}$$

(b) $m'_{nj}=m_{nj}\pm1$ and $m'_{nl}=m_{nl}$ for $l\neq j$,

$$I^{(\lambda^{+}_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}=-iB'(m_{nj}),\quad I^{(\lambda^{-}_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}=-iB'(m_{nj}-1), \tag{6.23a}$$

$$T^{(\lambda^{+}_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}=i(m_{nj}+1-j)B'Q(m_{nj}),$$
$$T^{(\lambda^{-}_{nj},\lambda_n)}_{\Lambda_{n-1}\lambda_{n-1}}=-i(m_{nj}+n-1-j)B'(m_{nj}-1). \tag{6.23b}$$

The primes on A', B' and C' mean the omission of a factor containing ν_{n+1}, i.e.,

$$[(l_{nj}+\tfrac{1}{2})^2+\nu^2_{n+1}]^{1/2}A'(m_{nj})={}^{b}A(m_{nj}), \tag{6.24a}$$

$$R^{b}_{n+1n}(\zeta)\Phi^{(\Lambda_{n-1},\rho_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = \sum_{\lambda'_n} {}^{b}d^{(\Lambda_{n-1},\rho_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta)$$

$$\times \Phi^{(\Lambda_{n-1},\rho_{n+1})}_{\lambda'_n\{\lambda_{n-1}\}}(g^{(n)}). \quad (7.9)$$

Making use of (4.4a) and (4.6a), we obtain

$${}^{b}d^{(\Lambda_{n-1},\rho_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta) = [i\nu_{n+1} \to (n-1)/2 + \rho_{n+1} \text{ in } (6.19)]. \quad (7.10)$$

As is expected, this expression agrees with that of (6.19) or (6.37) except that ρ_{n+1} appears instead of ν_{n+1} or σ_{n+1}. It is noted that (7.6) and (7.10) hold in general regardless of the existence of a scalar product for the bases. It is evident that (7.10) satisfies the same relations as (6.20a) and (6.20c), i.e., the representation conditions.

It is noted that the following relations hold,

$${}^{b}d^{(\Lambda_{n-1},\rho_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta) = \exp\left[i\pi \sum_{j}^{[n/2]} (m'_{nj} - m_{nj})\right] {}^{b}d^{(\Lambda_{n-1},\rho_{n+1})}_{\lambda_n(\lambda_{n-1})\lambda'_n}(\zeta),$$

$${}^{b}d^{(\Lambda_{n-1},\rho_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta) = {}^{b}d^{(\lambda_{n-1},\rho_{n+1})}_{\lambda'_n(\Lambda_{n-1})\lambda_n}(\zeta). \quad (7.11)$$

The first one is easily seen from the property for the matrix elements of J_{n+1n} and the second from (7.10) with the first one for the d matrix elements of SO(n).

8. CLASSIFICATIONS OF THE UIR'S OF SO(n, 1)

In this section, a simple discussion on Tables I and II is given.

The classification of the UIR's of SO(n,1) is considered to be given by the same Tables I and II, because the matrix elements of J_{n+1n} are the same as those for D_{n+1n} by which the classification is made. However, it is necessary in our case that the integral must converge in each class. Therefore, simple discussions are given on the classes in Tables I and II.

For the principal series, the gamma functions contained in the factor $N(\nu_{n+1};\lambda_n)$ have no singularities for any possible numbers m_{nj} and $\nu_{n+1} \neq 0$ and thus we obtain the class (i) in Tables I and II. For $\nu_{n+1} = 0$, we obtain $N(\nu_{n+1};\lambda_n) = 1$ from (6.17) and thus have the class (i) in Tables I and II. Though the case of $\nu_{n+1} = 0$ is excluded in Table I, our case contains this class because the complementary series cannot include the class corresponding to $\nu_{n+1} = 0$. The orthonormal bases and the boost matrix elements of the representation $R^{b}_{n+1n}(\zeta)$ are given by (6.11) and (6.19) with (6.16) ($\epsilon = 1$).

For the complementary series, there are many classes in Tables I and II. It is seen from (6.28) that at least the following integral must converge,

$$\int_0^{\pi} d\theta \sin^{n-2}\theta (1-\cos\theta)^{(1-2)/2 - \sigma_{n+1}} d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\Lambda_{n-1}}(\theta). \quad (8.1)$$

The d functions are polynomial in trigonometric functions and the same relation as (6.20a) holds. This means that the d function in (8.1) must contain at least a term without $\sin\theta$. Thus, we get the condition $-\sigma_{n+1} > 0$ for the convergence of the integral. Therefore, we must get rid of the cases of $\sigma_{n+1} = 0$ in Table I. For $\sigma_{n+1} \neq 0$, it is easy to confirm the results given in Tables I and II. The orthonormal bases and the boost matrix elements are given by (6.25) and (6.37). The factor $N(\sigma_{n+1};\lambda_n)$ is given by (6.35) except for the D^- case in Table I. For the D^- case, the factor $N(\sigma_{n+1};\lambda_n$ is given by

$$N(\sigma_{n+1};\lambda_n) = \left(\frac{\Gamma(-m_{nn/2} - \sigma_{n+1} + \frac{1}{2})}{\Gamma(-m_{nn/2} + \sigma_{n+1} + \frac{1}{2})} \times \prod_{j=1}^{(n-2)/2} \frac{\Gamma(m_{nj} - \sigma_{n+1} + (n+1)/2 - j)}{\Gamma(m_{nj} + \sigma_{n+1} + (n+1)/2 - j)}\right)^{1/2}$$

9. CONTINUATION TO THE MATRIX ELEMENTS OF SO($n+1$) AND ISO(n)

In this section, it is shown that by continuation with respect to the parameters of SO(n,1) (the principal series) the representations of SO($n+1$) and ISO(n) are obtained.[5,9,11]

It follows from the fact stated below (2.18) and the relation between (3.8) and (3.13) that the representation of SO($n+1$) will be obtained from that of the principal series of SO(n,1) by the analytic continuations of the parameters, i.e., by the substitutions of $\zeta, \rho_{n+1} = (1-n)/2 + i\nu_{n+1}$ and J_{nn+1} into $i\theta_{n+11}, m_{n+11}$, and iJ_{nn+1}. Then the bases and the differential operator are given by

$$\Phi^{(\lambda_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = N(m_{n+11};\lambda_n)\left(\frac{N(\lambda_n)}{V_n}\right)^{1/2} D^{(\lambda_n)}_{\{\Lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \quad (9.1a)$$

$$J_{nn+1} = m_{n+11}\cos\theta_{n1} - \sin\theta_{n1}\frac{\partial}{\partial\theta_{n1}}, \quad (9.1b)$$

where $\lambda_{n+1} = (m_{n+11}, \Lambda_{n-1})$ and

$$N(m_{n+11};\lambda_n) = \exp\left(-\frac{i\pi}{2}\sum_{j}^{[n/2]} m_{nj}\right) \times \left[\prod_{j=1}^{[n/2]} \Gamma(m_{n+11} + m_{nj} + n - j)\Gamma(m_{n+11} - m_{nj} + j)\right]^{-1/2}.$$

The factor $N(m_{n+11};\lambda_n)$ is easily found from (6.16) by taking into account the property of the gamma function, i.e.,

$$\frac{\Gamma(-m)}{\Gamma(-m')} = (-1)^{m-m'}\frac{\Gamma(m'+1)}{\Gamma(m+1)},$$

for m and m' positive integers and a phase factor is fixed as above. The differential operators J_{n+1j} ($j = 1,2,\ldots,n-1$) are obtained from (5.9) by the above substitutions or directly from the commutation relations

$$J_{n+1j} = i[J_{n+1n}, J_{nj}] \quad (9.2)$$

and it is straightforward to check that they satisfy the commutation relations of SO($n+1$) together with (2.16b).

It is easily seen from (6.21), (6.22), and (6.23) that the action of (9.1b) on (9.1a) agrees with (3.5), i.e.,

$$J_{nn+1}\Phi^{(\lambda_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = \sum_{j=1}^{n/2} A(m_{nj})\Phi^{(\lambda_{n+1})}_{\lambda^+_{nj}\{\lambda_{n-1}\}}(g^{(n)}) - \sum_{j=1}^{n/2} A(m_{nj}-1)\Phi^{(\lambda_{n+1})}_{\lambda^-_{nj}\{\lambda_{n-1}\}}(g^{(n)}), \tag{9.3a}$$

for $n+1$ odd, and

$$J_{nn+1}\Phi^{(\lambda_{n+1})}_{\{\lambda_n\}}(g^{(n)}) = \sum_{j=1}^{(n-1)/2} B(m_{nj})\Phi^{(\lambda_{n+1})}_{\lambda^+_{nj}\{\lambda_{n-1}\}}(g^{(n)}) - \sum_{j=1}^{(n-1)/2} B(m_{nj}-1)\Phi^{(\lambda_{n+1})}_{\lambda^-_{nj}\{\lambda_{n-1}\}}(g^{(n)}) + C_{n+1}\Phi^{(\lambda_{n+1})}_{\{\lambda_n\}}(g^{(n)}), \tag{9.3b}$$

for $n+1$ even. The formula for the D matrix elements is obtained from (6.19) by the replacements

$$\begin{aligned} &{}^bd^{(\Lambda_{n-1},\lambda_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta) \to d^{(\lambda_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\theta_{n+11}),\\ &\zeta \to i\theta_{n+11},\\ &\rho_{n+1} = (1-n)/2 + i\nu_{n+1} \to m_{n+11},\\ &N(\nu_{n+1};\lambda_n) \to N(m_{n+11};\lambda_n). \end{aligned} \tag{9.4}$$

Then it follows from (6.20) that the d matrix elements satisfy the representation and the unitarity conditions.

Similarly, we consider the inhomogeneous rotation group ISO(n) (Euclidean motion). The contraction of SO(n,1) (the principal series) to ISO(n) is given by $i\rho_{n+1} \to \infty$ under the condition of $i\rho_{n+1}\zeta = \gamma\xi$, where γ is some constant and $0 \leqslant \xi < \infty$.[9,10] The infinitesimal operator J_n of the representation corresponding to a translation to the nth direction is given by

$$J_n = \lim_{i\rho_{n+1}\to\infty} \frac{\gamma}{i\rho_{n+1}} J_{n+11} = \gamma\cos\theta_{n1}. \tag{9.5}$$

The operators J_j ($j = 1,2,\dots,n-1$) of the representation are obtained from

$$J_j = -i[J_{nj},J_n], \tag{9.6}$$

where J_{nj}'s are, of course, given by (2.16). Then it is easy to check that the commutation relations of ISO(n) are satisfied. The second-order Casimir operator becomes as follows,

$${}^IF^{(n)} = \sum_{j=1}^{n} J_j^2 = \gamma^2. \tag{9.7}$$

The orthonormal bases are given as follows,

$$\Phi^{(\Lambda_{n-1},\gamma)}_{\{\lambda_n\}}(g^{(n)}) = \exp\left(-i\frac{\pi}{2}\sum_j^{[n/2]} m_{nj}\right)\left(\frac{N(\lambda_n)}{V_n}\right)^{1/2} D^{(\lambda_n)}_{\{\Lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}). \tag{9.8}$$

The action of J_n on (9.8) is given by the same form as (6.6), i.e.,

$$J_n\Phi^{(\Lambda_{n-1},\gamma)}_{\{\lambda_n\}}(g^{(n)}) = +i\gamma\sum_{j=1}^{n/2} I^{(\lambda^+_{nj},\lambda_n)}_{\Lambda_{n-1},\lambda_{n-1}}\,\Phi^{(\Lambda_{n-1},\gamma)}_{\lambda^+_{nj}\{\lambda_{n-1}\}}(g^{(n)}) - i\gamma\sum_{j=1}^{n/2} I^{(\lambda^-_{nj},\lambda_n)}_{\Lambda_{n-1},\lambda_{n-1}}\,\Phi^{(\Lambda_{n-1},\gamma)}_{\lambda^-_{nj}\{\lambda_{n-1}\}}(g^{(n)}), \tag{9.9a}$$

for n even, and

$$J_n\Phi^{(\Lambda_{n-1},\gamma)}_{\{\lambda_n\}}(g^{(n)}) = +i\gamma\sum_{j=1}^{(n-1)/2} I^{(\lambda^+_{nj},\lambda_n)}_{\Lambda_{n-1},\lambda_{n-1}}\,\Phi^{(\Lambda_{n-1},\gamma)}_{\lambda^+_{nj}\{\lambda_{n-1}\}}(g^{(n)}) - i\gamma\sum_{j=1}^{(n-1)/2} I^{(\lambda^-_{nj},\lambda_n)}_{\Lambda_{n-1},\lambda_{n-1}}\,\Phi^{(\Lambda_{n-1},\gamma)}_{\lambda^-_{nj}\{\lambda_{n-1}\}}(g^{(n)}) + \gamma I^{(\lambda_n,\lambda_n)}_{\Lambda_{n-1},\lambda_{n-1}}\,\Phi^{(\Lambda_{n-1},\gamma)}_{\{\lambda_n\}}(g^{(n)}), \tag{9.9b}$$

for n odd. The matrix elements I are given by (6.21) or (6.22) and (6.23).

The formula for the ${}^Id$ matrix elements, which are the matrix elements of the representation ($\exp i\xi J_n$) of the nth direction, is obtained directly by sandwiching $\exp i\xi J_n$ between (9.8) or from (6.19) by the replacements

$$\begin{aligned} &{}^bd^{(\Lambda_{n-1},\nu_{n+1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\zeta) \to {}^Id^{(\Lambda_{n-1},\gamma)}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\xi),\\ &(\cosh\zeta - \cos\theta\sinh\zeta)^{(1-n)/2+i\nu_{n+1}} \to \exp(i\gamma\xi\cos\theta), \end{aligned} \tag{9.10}$$

$$\theta' \to \theta,$$

$$N(\nu_{n+1};\lambda_n) \to \exp\left(-i\frac{\pi}{2}\sum_{j=1}^{[n/2]} m_{nj}\right).$$

These are easily seen from $i\rho_{n+1} \to \infty$ under $i\rho_{n+1}\zeta = \gamma\xi$. Then it is seen that the representation and the unitarity conditions are satisfied,

$$\begin{aligned} &{}^Id^{(\Lambda_{n-1},\gamma)}_{\lambda'_n(\lambda_{n-1})\lambda_n}(O) = \delta_{\lambda'_n\lambda_n},\\ &\sum_{\lambda''_n} {}^Id^{(\Lambda_{n-1},\gamma)}_{\lambda'_n(\lambda_{n-1})\lambda''_n}(\xi_1)\,{}^Id^{(\Lambda_{n-1},\gamma)}_{\lambda''_n(\lambda_{n-1})\lambda_n}(\xi_2) = {}^Id^{(\Lambda_{n-1},\gamma)}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\xi_1+\xi_2),\\ &\overline{{}^Id^{(\Lambda_{n-1},\gamma)}_{\lambda'_n(\lambda_{n-1})\lambda_n}(\xi)} = {}^Id^{(\Lambda_{n-1},\gamma)}_{\lambda_n(\lambda_{n-1})\lambda'_n}(-\xi). \end{aligned} \tag{9.11}$$

Finally, it is noted that a finite dimensional nonunitary representation of SO(n,1) is obtained from the irreducible representation in Sec. 7 by replacing ρ_{n+1} by m_{n+11}. The dimension of the representation, of course, agrees with that for SO($n+1$), and the bases of the representation are given by (7.8). The formula for the d matrix elements is obtained from (7.10) by the replacements $\rho_{n+1} \rightarrow m_{n+11}$ and $N(\rho_{n+1};\lambda_n) \rightarrow N(m_{n+11};\lambda_n)$.

10. SIMPLE EXAMPLES

The results of Sec. 7 are easily checked for the cases of low n by using the known relations for the d matrix elements. Simple examples are given to show the validity of our phase convention.

A. SO(2,1)[18]

The bases and the operators are given by

$$\Phi_m^{(\rho)}(\theta) = \frac{N(\rho;m)}{\sqrt{2\pi}} e^{im\theta}, \tag{10.1a}$$

$$\frac{N(\rho;m)}{N(\rho;m')} = \left(\frac{\Gamma(m-\rho)\Gamma(m'+\rho+1)}{\Gamma(m+\rho+1)\Gamma(m'-\rho)}\right)^{1/2},$$

$$J_{21} = p_{21} = -i\frac{\partial}{\partial\theta},$$

$$J_{32} = -i\left(\sin\theta\frac{\partial}{\partial\theta} - \rho\cos\theta\right), \tag{10.1b}$$

where $0 \leqslant \theta < 2\pi$ and ρ is a complex number. The second-order Casimir operator becomes

$$F^{(2,1)} = J_{21}^2 - J_{31}^2 - J_{32}^2 = \rho(\rho+1). \tag{10.2}$$

The action of J_{23} on (10.1a) is really obtained as follows

$$J_{23}\Phi_m^{(\rho)}(\theta) = \frac{i}{2}\sqrt{(m-\rho)(m+\rho+1)}\,\Phi_{m+1}^{(\rho)}(\theta)$$

$$-\frac{i}{2}\sqrt{(m-\rho-1)(m+\rho)}\,\Phi_{m-1}^{(\rho)}(\theta). \tag{10.3}$$

The computation formula for the representation matrix elements corresponding to the boost becomes as follows

$${}^b d_{m'm}^{(\rho)}(\zeta) = \frac{1}{2\pi}\frac{N(\rho;m)}{N(\rho;m')}\int_0^{2\pi} d\theta\, e^{-im'\theta}(\cosh\zeta - \cos\theta\sinh\zeta)^\rho e^{im\theta'}. \tag{10.4}$$

It is noted that the factor $N(\rho;m)$ for class D^- in Table I must be chosen as in the form shown in Sec. 8. (10.4) can easily be integrated and expressed in terms of a hypergeometric function.

B. SO(3,1)[20]

The bases and the operators are given by

$$\Phi_{(jm)}^{(k,\rho)}(\{\theta_3\}) = N(\rho;j)\left(\frac{2j+1}{8\pi^2}\right)^{1/2} D_{km}^{(j)}(\{\theta_3\}), \tag{10.5a}$$

$$D_{km}^{(j)}(\theta_3\}) = \exp i(k\theta_{21} + m\theta_{32}) d_{km}^{(j)}(\theta_{31}),$$

$$\frac{N(\rho;j)}{N(\rho;j')} = \left(\frac{\Gamma(j-\rho)\Gamma(j'+\rho+2)}{\Gamma(j+\rho+2)\Gamma(j'-\rho)}\right)^{1/2},$$

$$J_{21} = p_{32},\quad J_{32} = \cos\theta_{32}p_{31} - \frac{\cos\theta_{31}}{\sin\theta_{31}}\sin\theta_{32}p_{32} + \frac{\sin\theta_{32}}{\sin\theta_{31}}p_{21}, \tag{10.5b}$$

$$\bar{J}_{21} = p_{21},\quad \bar{J}_{32} = \cos\theta_{21}p_{31} - \frac{\cos\theta_{31}}{\sin\theta_{31}}\sin\theta_{21}p_{21} + \frac{\sin\theta_{21}}{\sin\theta_{31}}p_{32},$$

$$J_{43} = -i\left(\sin\theta_{31}\frac{\partial}{\partial\theta_{31}} - \rho\cos\theta_{31}\right),$$

where ρ is a complex number and k, j, and m are used instead of m_{42}, m_{31}, and m_{21}. The second-order Casimir operators become

$$F^{(3,1)} = \sum_{j>k}^{3} J_{jk}^2 - \sum_{j}^{3} J_{4j}^2 = \rho(\rho+2) + \overline{J}_{21}^2,$$

$$G = J_{12}J_{43} + J_{31}J_{42} + J_{23}J_{41} = -i(\rho+1)\overline{J}_{21}. \tag{10.6}$$

The action of J_{43} on (10.5a) is really given by

$$J_{34}\Phi_{(jm)}^{(k,\rho)}(\{\theta_3\}) = B(j)\Phi_{(j+1m)}^{(k,\rho)}(\{\theta_3\}) - B(j-1)\Phi_{(j-1m)}^{(k,\rho)}(\{\theta_3\}) + C_4\Phi_{(jm)}^{(k,\rho)}(\{\theta_3\}), \tag{10.7}$$

with the matrix elements

$$B(j) = i\left(\frac{[(j+1)^2-k^2][(j+1)^2-m^2][(j+1)^2-(\rho+1)^2]}{(2j+3)(j+1)(2j+1)(j+1)}\right)^{1/2},$$

$$C_4 = -i\frac{mk(\rho+1)}{j(j+1)}.$$

The formula for the representation matrix elements corresponding to the boost is

$${}^b d_{j'(m)j}^{(k,\rho)}(\zeta) = \tfrac{1}{2}[(2j+1)(2j'+1)]^{1/2}\frac{N(\rho;j)}{N(\rho;j')}\int_0^{\pi} d\theta \sin\theta\, \overline{d_{km}^{(j')}(\theta)}\,(\cosh\zeta - \cos\theta\sinh\zeta)^{\rho} d_{km}^{(j)}(\theta'). \tag{10.8}$$

This is integrated and expressed in terms of the hypergeometric functions.

C. SO(4,1)

The bases and the operators are given by

$$\Phi_{(kk_0jm)}^{(K_0,\rho)}(\{\theta_4\}) = N(\rho;\lambda_4)\left(\frac{N(\lambda_4)}{V_4}\right)^{1/2} D_{(K_0m_0)(jm)}^{(\lambda_4)}(\{\theta_4\}),$$

$$D_{(K_0m_0)(jm)}^{(\lambda_4)}(\{\theta_4\}) = \sum_{m'} D_{m_0m'}^{(K_0)}(\{\theta_3\}) d_{K_0(m')j}^{(kk_0)}(\theta_{41}) D_{m'm}^{(j)}(0,\theta_{42},\theta_{43}), \tag{10.9a}$$

$$N(\lambda_4) = (k+k_0+1)(k-k_0+1),$$

$$\frac{N(\rho;\lambda_4)}{N(\rho;\lambda_4')} = \left(\frac{\Gamma(k-\rho)\Gamma(k_0-\rho-1)\Gamma(k'+\rho+3)\Gamma(k_0'+\rho+2)}{\Gamma(k+\rho+3)\Gamma(k_0+\rho+2)\Gamma(k'-\rho)\Gamma(k_0'-\rho-1)}\right)^{1/2},$$

$$J_{21} = p_{43}, \quad J_{32} = \cos\theta_{43}p_{42} - \frac{\cos\theta_{42}}{\sin\theta_{42}}\sin\theta_{43}p_{43} + \frac{\sin\theta_{43}}{\sin\theta_{42}}p_{32},$$

$$J_{43} = \cos\theta_{42}p_{41} - \frac{\cos\theta_{41}}{\sin\theta_{41}}\sin\theta_{42}p_{42} + \frac{\sin\theta_{42}}{\sin\theta_{41}}\left(\cos\theta_{32}p_{31} - \frac{\cos\theta_{31}}{\sin\theta_{31}}\sin\theta_{32}p_{32} + \frac{\sin\theta_{32}}{\sin\theta_{31}}p_{21}\right), \tag{10.9b}$$

$$\overline{J}_{21} = p_{21}, \quad \overline{J}_{32} = \cos\theta_{21}p_{31} - \frac{\cos\theta_{31}}{\sin\theta_{31}}\sin\theta_{21}p_{21} + \frac{\sin\theta_{21}}{\sin\theta_{31}}p_{32},$$

$$\overline{J}_{43} = \cos\theta_{31}p_{41} - \frac{\cos\theta_{41}}{\sin\theta_{41}}\sin\theta_{31}p_{31} + \frac{\sin\theta_{31}}{\sin\theta_{41}}\left(\cos\theta_{32}p_{42} - \frac{\cos\theta_{42}}{\sin\theta_{42}}\sin\theta_{32}p_{32} + \frac{\sin\theta_{32}}{\sin\theta_{42}}p_{43}\right),$$

$$J_{54} = -i\left(\sin\theta_{41}\frac{\partial}{\partial\theta_{41}} - \rho\cos\theta_{41}\right),$$

$\lambda_4 = (k,k_0)$, $\lambda_3 = j$, $\lambda_2 = m$, $\Lambda_3 = K_0$, and $\lambda_2^0 = m_0$ are used instead of $\lambda_4 = (m_{41},m_{42})$, $\lambda_3 = m_{31}$, $\lambda_2 = m_{21}$, $\Lambda_3 = m_{52}$, and $\lambda_2^0 = m_{21}^0$.

The invariant operators become

$$F^{(4,1)} = \sum_{j>k}^{4} J_{jk}^2 - \sum_{j}^{4} J_{5j}^2 = \rho(\rho+3) + \sum_{j>k}^{3} \overline{J}_{jk}^2,$$

$$G = (J_{23}J_{14} + J_{31}J_{24} + J_{12}J_{34})^2 - (J_{23}J_{45} + J_{42}J_{35} + J_{25}J_{34})^2 - (J_{31}J_{45} + J_{43}J_{15} + J_{35}J_{14})^2$$
$$- (J_{12}J_{45} + J_{41}J_{25} + J_{24}J_{15})^2 - (J_{21}J_{35} + J_{13}J_{25} + J_{32}J_{15})^2 = (\rho+1)(\rho+2)\sum_{j>k}^{3}\overline{J}_{jk}^2. \tag{10.10}$$

The action of J_{54} on (10.9a) is given by

$$J_{45}\Phi_{(kk_0jm)}^{(K_0,\rho)}(\{\theta_4\}) = {}^bA(k)\Phi_{(k+1k_0jm)}^{(K_0,\rho)}(\{\theta_4\}) + {}^bA(k_0)\Phi_{(kk_0+1jm)}^{(K_0,\rho)}(\{\theta_4\})$$
$$- {}^bA(k-1)\Phi_{(k-1k_0jm)}^{(K_0,\rho)}(\{\theta_4\}) - {}^bA(k_0-1)\Phi_{(kk_0-1jm)}^{(K_0,\rho)}(\{\theta_4\}), \tag{10.11}$$

with the matrix elements

$$ {}^bA(k) = \frac{i}{2}\left(\frac{(k+j+2)(k-j+1)(k+K_0+2)(k-K_0+1)(k+\rho+3)(k-\rho)}{(k+k_0+1)(k-k_0+1)(k+k_0+2)(k-k_0+2)}\right)^{1/2}, $$

$$ {}^bA(k_0) = \frac{i}{2}\left(\frac{(j+k_0+1)(j-k_0)(K_0+k_0+1)(K_0-k_0)(k_0+\rho+2)(k_0-\rho-1)}{(k+k_0+1)(k-k_0+1)(k+k_0+2)(k-k_0)}\right)^{1/2}. $$

The formula for the representation matrix elements corresponding to the boost is

$$ {}^bd^{(K_0,\rho)}_{(k'k'_0)(j)(kk_0)}(\zeta) = \frac{2}{\pi}\frac{\sqrt{N(\lambda_4)N(\lambda'_4)}}{(2K_0+1)(2j+1)}\frac{N(\rho;kk_0)}{N(\rho;k'k'_0)} $$

$$ \times\sum_m\int_0^\pi d\theta\,\sin^2\theta\,\overline{d^{(k'k'_0)}_{K_0(m)j}(\theta)}(\cosh\zeta-\cos\theta\sinh\zeta)^\rho d^{(kk_0)}_{K_0(m)j}(\theta'). \qquad (10.12) $$

In order to derive the above results, it will be convenient to use the expressions for the d matrix elements and the Clebsch–Gordan coefficients of SO(4).[21] In an Appendix, a discussion on the explicit derivation for SO(4,1) is made.

APPENDIX

In this Appendix, it is shown that the results for SO(4,1) in Sec. 10 are obtained explicitly.

The expression for the d matrix elements of SO(4) is given as follows,[9,10]

$$ d^{(k,k_0)}_{j'(m)j}(\theta) = (-1)^{(j'-j)/2}\sum_{m'}(k_+m_+k_-m_-;j'm)(k_+m_+k_-m_-;jm)e^{im'\theta}, \qquad (A1) $$

where $k_\pm=(k\pm k_0)/2$, $m_\pm=(m\pm m')/2$, and $(j_1m_1j_2m_2;jm)$ denotes the Clebsch–Gordan (C–G) coefficient. The C–G coefficients are related to the Racah coefficients[13] as follows,

$$ (j_1m_1j_2m_2;jm)(jmj_3m_3;j_4m_4) = \sum_s\sqrt{(2j+1)(2s+1)}(j_2m_2j_3m_3;sm_2+m_3)(j_1m_1sm_2+m_3;j_4m_4)W(j_1j_2j_4j_3;js), \qquad (A2) $$

where $W(j_1j_2j_3j_4;js)$ is the Racah coefficient with the relations

$$ \sum_s(2s+1)W(j_1j_2j_3j_4;sj)W(j_1j_2j_3j_4;sj') = \frac{1}{2j+1}\delta_{jj'}, \qquad (A3) $$

$$ \begin{aligned} W(j_1j_2j_3j_4;js) &= W(j_2j_1j_4j_3;js) = W(j_3j_4j_1j_2;js) = W(j_1j_3j_2j_4;sj) \\ &= (-1)^{j+s-j_1-j_4}W(jj_2j_3s;j_1j_4) = (-1)^{j+s-j_2-j_3}W(j_1jsj_4;j_2j_3). \end{aligned} \qquad (A4) $$

The orthogonality relation (4.6b) for the d matrix elements is obtained explicitly. Using the orthogonality relation for the C–G coefficients and the following exression for the d matrix elements, which is easily found from (A1) with (A2),

$$ d^{(k,k_0)}_{j'(m)j}(\theta) = (-1)^{(j'-j)/2+k_--m}\sqrt{(2j'+1)(2j+1)}\sum_s(j'mj-m;s0) $$

$$ \times W(j'jk_-k_-;sk_+)e^{im\theta}\sum_{m'}(-1)^{2k_-+m'}(k_-m'k-m';s0)e^{2im'\theta}, $$

we obtain

$$ \sum_m\int_0^\pi d\theta\,\sin^2\theta\,\overline{d^{(k'k'_0)}_{j'(m)j}(\theta)}d^{(kk_0)}_{j'(m)j}(\theta) = \frac{\pi}{2}\sum_s(-1)^{k_--k'_-+2(k_-k'_-)}(2j'+1)(2j+1)W(j'jk_-k_-;sk_+) $$

$$ \times W(j'jk'_-k'_-;sk'_+)\sum_{m',m''}(-1)^{2(m'-m'')}(k_-m'k_--m';s0)(k'_-m''k'_--m'';s0) $$

$$ \times\frac{1}{\Gamma(2+m'-m'')\Gamma(2-m'+m'')}, \qquad (A5) $$

where the formula

$$ \int_0^\pi d\theta\,\sin^\alpha\theta\,e^{i\beta\theta} = \frac{\pi\Gamma(1+\alpha)e^{i\pi\beta/2}}{2^\alpha\Gamma(1+(\alpha+\beta)/2)\Gamma(1+(\alpha-\beta)/2)}, \quad \mathrm{Re}\,\alpha>-1, $$

is used. The following relation is easily obtained by using (A2),

$$\sum_{m',m''} (-1)^{2(m'-m'')}(k_-m'k_- - m';s0)(k'_-m''k'_- - m'';s0)\frac{1}{\Gamma(2+m'-m'')\Gamma(2-m'+m'')} = \frac{2s+1}{2k_-+1}\delta_{k\ k'}. \quad \text{(A6)}$$

Substuting (A6) into (A5) and taking into account (A3), we obtain the orthogonality relation (4.6b) for the d matrix elements of SO(4).

The product of the d matrix elements can be expressed in terms of a linear combination of the d matrix elements with the help of the C–G coefficients of SO(4)[21]

$$d^{(kk_0)}_{j_1'(m_1)j_1}(\theta)d^{(k'k_0')}_{j_2'(m_2)j_2}(\theta)$$
$$= \sum_{\substack{K,K_0\\J_1,J_2}} (j_1'm_1j_2'm_2;J_1M)(j_1m_1j_2m_2;J_2M)(kk_0j_1',k'k_0'j_2';KK_0J_1)(kk_0j_1,k'k_0'j_2;KK_0J_2)d^{(KK_0)}_{J_1(M)J_2}(\theta), \quad \text{(A7)}$$

where the factor $(kk_0j_1,k'k_0'j_2;KK_0J)$ is related with the Racah coefficients as follows,

$$(kk_0j_1,k'k_0'j_2;KK_0J)$$
$$= \sqrt{(2K_++1)(2K_-+1)(2j_1+1)(2j_2+1)}(-1)^{k_++k\ \ +k'_+ +k'\ \ +j_1+j_2}$$
$$\times \sum_s (-1)^{2s}(2s+1)W(j_1k_-Js;k_+j_2)W(j_2k'_-sK_-;k'_+k_-)W(JK_-k_+k'_+;K_+s).$$

From (A7) we obtain the relation

$$\cos\theta d^{(kk_0)}_{j'(m)j}(\theta) = \tfrac{1}{2}(-1)^{j-j'}\sum_{K,K_0}(2K_++1)(2K_-+1)W(K_+k_+K_-k_-;\tfrac{1}{2}j')W(K_+k_+K_-k_-;\tfrac{1}{2}j)d^{(KK_0)}_{j'(m)j}(\theta), \quad \text{(A8)}$$

because $d^{(10)}_{0(0)0}(\theta) = \cos\theta$. It follows that the right-hand side on (A8) can be expressed in terms of the d matrix elements with $K = k \pm 1$ and $K_0 = k_0 \pm 1$.

In order to obtain an action of J_{54} on the bases (10.9a), we must further find the action of $\sin\theta(\partial/\partial\theta)$ on (A1). As we cannot obtain the result in a form needed for us from (A7), we consider the action on (A1) directly. From (A1), we get

$$\sin\theta\frac{\partial}{\partial\theta}d^{(kk_0)}_{j'(m)j}(\theta) = i\sin\theta(-1)^{(j'-j)/2}\sum_{m'} m'(k_+m_+k_-m_-;j'm)(k_+m_+k_-m_-;jm)e^{im'\theta}. \quad \text{(A9)}$$

The right-hand side can be rewritten by using the C–G coefficients as follows

$$\sin\theta\frac{\partial}{\partial\theta}d^{(kk_0)}_{j'(m)j}(\theta) = \tfrac{1}{2}(-1)^{(j'-j)/2}\sqrt{3k_+(k_++1)}\sum_{m'} f(k_+k_-m_+m_-)e^{im'\theta}$$
$$+\tfrac{1}{2}(-1)^{(j-j')/2}\sqrt{3k_-(k_-+1)}\sum_{m'} f(k_-k_+m_-m_+)e^{im'\theta}, \quad \text{(A10)}$$

where

$$f(k_+k_-m_+m_-) = \sum_s (\tfrac{1}{2}s10;\tfrac{1}{2}s)(k_+m_+-s10;k_+m_+-s)(k_+m_+-sk_-m_-+s;j'm)(k_+m_+-sk_-m_-+s;jm).$$

Using (A2) three times in f, we can rewrite f as follows,

$$f(k_+k_-m_+m_-) = (-1)^{j-j'-k_++3/2}\sqrt{2(2k_++1)}\sum_{p,q}(-1)^p(2p+1)(2q+1)$$
$$\times W(\tfrac{1}{2}\tfrac{1}{2}k_+k_+;1p)W(pk_+qk_-;\tfrac{1}{2}j')W(pk_+qk_-;\tfrac{1}{2}j)(pm_+qm_-;j'm)(pm_+qm_-;jm). \quad \text{(A11)}$$

Substitution of (A11) into (A10) together with (A1) gives

$$\sin\theta\frac{\partial}{\partial\theta}d^{(kk_0)}_{j'(m)j}(\theta) = \tfrac{1}{2}(-1)^{j-j'}\sum_{p,q}\{(p-k_+)(p+k_++1)+(q-k_-)(q+k_-+1)-\tfrac{3}{2}\}$$
$$\times(2p+1)(2q+1)W(pk_+qk_-;\tfrac{1}{2}j')W(pk_+qk_-;\tfrac{1}{2}j)d^{(p+q,p-q)}_{j'(m)j}(\theta), \quad \text{(A12)}$$

$$\times (2p+1)(2q+1)W(pk_+qk_-;\tfrac{1}{2}j')W(pk_+qk_-;\tfrac{1}{2}j)d^{(p+q,p-q)}_{j'(m)j}(\theta), \quad \text{(A12)}$$

where the following expression for the Racah coefficient is used,

$$W(aabb;1c) = (-1)^{a+b-c-1}\frac{a(a+1)+b(b+1)-c(c+1)}{\sqrt{4a(a+1)(2a+1)b(b+1)(2b+1)}}.$$

If (A8) and (A12) are used, it is straightforward to confirm the result (10.11) of the action of J_{54} on (10.9a).

[1]T. Maekawa, J. Math. Phys. **19**, 2028 (1978).
[2]N.Ja. Vilenkin, *Special Functions and the Theory of Group Representations*, Math. Monographs Vol. 22, translated by V.N. Singh (Am. Math. Soc. Providence, Rhode Island, 1968).
[3]J. Dixmier, C. R. Acad. Sci. **250**, 3263 (1960).
[4]T. Hirai, Proc. Jpn. Acad. **38**, (1962); U. Ottoson, Commun. Math. Phys. **8**, 228 (1968).
[5]F. Schwartz, J. Math. Phys. **12**, 131 (1971).
[6]T. Maekawa, J. Math. Phys. **16**, 334 (1975).
[7]I.M. Gel'fand and M.L. Tsetlin, Dokl. Akad. Nauk SSSR **71**, 840 (1950).
[8]A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968).
[9]T. Maekawa, J. Math. Phys. **16**, 2372 (1975).
[10]K.B. Wolf, J. Math. Phys. **12**, 197 (1971).
[11]N.Ja. Vilenkin, Dokl. Akad. Nauk SSSR **113**, 16 (1957).
[12]E. Inönü and E.P. Wigner, Proc. Natl. Acad. Sci. (U.S.) **39**, 510 (1953).
[13]G. Racah, Group Theory and Spectroscopy, lecture notes (Princeton, 1951).
[14]K. Iwasawa, Ann. Math. **50**, 507 (1949); R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).
[15]W. Miller, Jr., *Symmetry Groups and Their Applications* (Academic, New York, 1972).
[16]S.C. Pang and K.T. Hecht, J. Math. Phys. **8**, 1233 (1967); M.K.F. Wong, *ibid.* 1899 (1967).
[17]M. Ikeda and T. Maekawa, Math. Jpn. **16**, 173 (1971).
[18]V. Bargmann, Ann. Math. **48**, 586 (1947).
[19]T. Maekawa, J. Math. Phys. **13**, 655 (1972); T. Maekawa, Math. Jpn. **18**, 129 (1973).
[20]M.A. Naimark, *Linear Repesentations of the Lorentz Group* (Pergamon, New York, 1964).
[21]L.C. Biedenharn, J. Math. Phys. **2**, 433 (1961).

Generalized Lie algebras

M. Scheunert[a)]

School of Theoretical Physics, Dublin Institute for Advanced Studies, Dublin 4, Ireland
(Received 19 September 1978)

The generalized Lie algebras, which have recently been introduced under the name of color (super) algebras, are investigated. The generalized Poincaré–Birkhoff–Witt and Ado theorems hold true. We discuss the so-called commutation factors which enter into the defining identities of these algebras. Moreover, we establish a close relationship between the generalized Lie algebras and ordinary Lie (super) algebras.

1. INTRODUCTION

Supersymmetry (Fermi–Bose symmetry) has turned out to be a powerful tool in particle physics.[1] Its mathematical foundation is the theory of Lie superalgebras[2,3,4] as well as of supermanifolds and of Lie supergroups.[5,6] Recently, a wide class of generalized Lie algebras has been described[7,8] which includes Lie algebras and Lie superalgebras as special cases. These algebras have been called color (super) algebras, for there is some hope that they might have a bearing on color and parastatistics.[7,9]

In the present work we shall investigate the generalized Lie algebras from a purely mathematical point of view. Having obtained a better insight into the structure of these algebras one might hope to be able to estimate their physical importance as well.

Our paper is organized as follows. In Sec. 2 we collect our conventions and make a few remarks on graded algebraic structures.

Section 3 contains the definitions of a commutation factor ϵ and of an ϵ Lie algebra; these two concepts are basic for the rest of this article.

It is well-known that the Poincaré–Birkhoff–Witt theorem is an important tool in the theory of Lie (super) algebras. We shall see in Sec. 4 that this theorem holds for the ϵ Lie algebras as well. Consequently, the technique of induced representations is at our disposal. However, we shall not enter into the discussion of this topic and only mention one simple result which will be needed later on.

In Sec. 5 we investigate the commutation factors on an arbitrary finitely generated Abelian group.

The results of the next section are somewhat unexpected. Looking at the general definition of an ϵ Lie algebra one might have the impression that the ϵ Lie algebras are much more general structures than Lie (super) algebras. However, we shall see in Sec. 6 that the theory of ϵ Lie algebras can be reduced to the theory of graded Lie (super) algebras; moreover, the same holds true for the graded representations.

In Sec. 7 we shall use the information obtained thus far to give a proof of the generalized Ado theorem.

Finally, Sec. 8 contains a few concluding remarks.

2. CONVENTIONS AND A FEW GENERALITIES ON GRADED ALGEBRAIC STRUCTURES

Throughout this work K will denote a commutative field of any characteristic and Γ will stand for an Abelian group. The multiplicative group of nonzero elements of K will be denoted by K_*. All vector spaces and algebras are assumed to have K as their field of scalars.

Of course, the physicist will mainly be interested in the case where K is the field of complex numbers. However, since by restricting our attention to the complex case we would save but a few lines we prefer to present the theory in its proper generality. The sporadic adjustments which are necessary in the case where K is a field of prime characteristic will be typified as "Remarks for the case of prime characteristic," and therefore may be skipped by the reader who is not interested in these generalizations.

In the following we shall collect some definitions concerning graded algebraic structures. For a detailed discussion of the subject we refer the reader to the literature.[10,11]

A vector space V is said to be *Γ-graded* if we are given a family $(V_\gamma)_{\gamma\in\Gamma}$ of subspaces of V such that V is their direct sum,

$$V = \bigoplus_{\gamma\in\Gamma} V_\gamma . \tag{2.1}$$

An element of V is said to be *homogeneous of degree* $\gamma\in\Gamma$ if it is an element of V_γ. A subspace V' of V is said to be *graded* if

$$V' = \bigoplus_{\gamma\in\Gamma} (V'\cap V_\gamma). \tag{2.2}$$

Let $V = \oplus_{\gamma\in\Gamma} V_\gamma$ and $W = \oplus_{\gamma\in\Gamma} W_\gamma$ be two Γ-graded vector spaces. A linear mapping,

$$g : V \to W, \tag{2.3}$$

is said to be *homogeneous of degree* $\gamma\in\Gamma$ if

$$g(V_\alpha) \subset W_{\alpha+\gamma}, \quad \text{for all } \alpha\in\Gamma. \tag{2.4}$$

Let Hom(V,W) denote the vector space of all linear mappings of V into W and let Hom$(V,W)_\gamma$ denote the subspace of those linear mappings of V into W which are homogeneous of degree γ. We define Homgr(V,W) to be the sum of these subspaces; obviously, this sum is direct:

$$\mathrm{Homgr}(V,W) = \bigoplus_{\gamma\in\Gamma} \mathrm{Hom}(V,W)_\gamma . \tag{2.5}$$

Thus Homgr(V,W) is a Γ-graded vector space. Note

[a)] Supported by the Deutsche Forschungsgemeinschaft.

0022-2488/79/040712-09$01.00

that Homgr(V,W) is equal to Hom(V,W) if (for example) $V_\gamma = \{0\}$ and $W_\gamma = \{0\}$ for all but a finite number of degrees. In the case where $V = W$ and $V_\gamma = W_\gamma$ for all $\gamma \in \Gamma$ we shall simplify the notation and write Hom(V) and Homgr(V) instead of Hom(V,V) and Homgr(V,V), respectively.

Let U, V, W be three Γ-graded vector spaces and let $h: U \rightarrow V$ and $g: V \rightarrow W$ be two linear mappings. If h is homogeneous of degree δ and g is homogeneous of degree γ, then $g \circ h$ is homogeneous of degree $\gamma + \delta$.

An algebra S is called *Γ-graded* if its underlying vector space is Γ-graded,

$$S = \bigoplus_{\gamma \in \Gamma} S_\gamma, \tag{2.6}$$

and if, furthermore,

$$S_\alpha S_\beta \subset S_{\alpha+\beta}, \quad \text{for all } \alpha,\beta \in \Gamma. \tag{2.7}$$

If S has a unit element e it follows that $e \in S_0$. A subalgebra of S is said to be *graded* if it is graded as a subspace of S.

Let T be a second Γ-graded algebra. A *homomorphism $S \rightarrow T$ of Γ-graded algebras* is by definition a homomorphism of the algebra S into the algebra T which is, in addition, homogeneous of degree zero.

Example: Let V be a Γ-graded vector space. Then the Γ-graded vector space Homgr(V), equipped in addition with the usual multiplication (i.e., composition) of linear mappings, is a Γ-graded algebra.

3. DEFINITION OF COMMUTATION FACTORS AND OF ϵ LIE ALGEBRAS

The following definition is well-known from the theory of graded algebras.[11]

Definition 1: Let Γ be an Abelian group. A *commutation factor* on Γ (with values in K_*) is a mapping,

$$\epsilon: \Gamma \times \Gamma \rightarrow K_*, \tag{3.1a}$$

such that

$$\epsilon(\alpha,\beta)\epsilon(\beta,\alpha) = 1, \tag{3.1b}$$

$$\epsilon(\alpha,\beta+\gamma) = \epsilon(\alpha,\beta)\epsilon(\alpha,\gamma), \tag{3.1c}$$

$$\epsilon(\alpha+\beta,\gamma) = \epsilon(\alpha,\gamma)\epsilon(\beta,\gamma), \tag{3.1d}$$

for all α, β, $\gamma \in \Gamma$.
The commutation factors on finitely generated Abelian groups are discussed in Sec. 5.

Definition 2: Let Γ be an Abelian group and let ϵ be a commutation factor on Γ. A Γ-graded algebra,

$$L = \bigoplus_{\gamma \in \Gamma} L_\gamma, \tag{3.2}$$

whose product mapping is denoted by a pointed bracket $\langle\ ,\ \rangle$, is called a (Γ-graded) *ϵ Lie algebra* if the following identities are satisfied:

$$\langle A,B \rangle = -\epsilon(\alpha,\beta)\langle B,A \rangle \quad (\epsilon \text{ skew symmetry}), \tag{3.3}$$

$$\epsilon(\gamma,\alpha)\langle A,\langle B,C\rangle\rangle + \text{cyclic} = 0 \quad (\epsilon \text{ Jacobi identity}), \tag{3.4}$$

for all $A \in L_\alpha$, $B \in L_\beta$, $C \in L_\gamma$, α, β, $\gamma \in \Gamma$.

Remark for the case of prime characteristic: If the characteristic of K is equal to 2 or 3 we shall require that one more identity is satisfied. In fact, we shall assume that in the case char$K = 2$,

$$\langle A,A \rangle = 0 \quad \text{for all } A \in L_\alpha,\ \alpha \in \Gamma, \tag{3.5}$$

whereas in the case char$K = 3$,

$$\langle A,\langle A,A \rangle\rangle = 0 \quad \text{for all } A \in L_\alpha,\ \alpha \in \Gamma. \tag{3.6}$$

Note that (3.6) is a consequence of (3.4) if char$K \neq 3$, note also that our additional requirements are necessary if the generalized Poincaré–Birkhoff–Witt theorem and the generalized Ado theorem are expected to hold.

Examples: (1) Choose ϵ to be the trivial commutation factor, which is defined by

$$\epsilon(\alpha,\beta) = 1 \quad \text{for all } \alpha,\beta \in \Gamma. \tag{3.7}$$

Then a Γ-graded ϵ Lie algebra is nothing but a Γ-graded Lie algebra.

(2) Let $\Gamma = Z_2$ be the additive group of integers modulo 2 and let ϵ be defined by

$$\epsilon(\alpha,\beta) = (-1)^{\alpha\beta} \quad \text{for all } \alpha,\beta \in Z_2 \tag{3.8}$$

(the right-hand side has an obvious meaning). Then the Z_2-graded ϵ Lie algebras are just the Lie superalgebras.

(3) The following example illustrates why the axioms for an ϵ Lie algebra have been chosen as described above.

Let ϵ be any commutation factor on Γ and let S be any associative Γ-graded algebra. On the Γ-graded vector space S we define a new multiplication $\langle\ ,\ \rangle$ by the requirement that

$$\langle a,b \rangle = ab - \epsilon(\alpha,\beta)ba \quad \text{for all } a \in S_\alpha,\ b \in S_\beta,\ \alpha,\beta \in \Gamma. \tag{3.9}$$

It is easy to see that the bracket $\langle\ ,\ \rangle$ turns S into an ϵ Lie algebra which is said to be associated with S and which will be denoted by $S(\epsilon)$.

Quite generally, the bracket (3.9) yields a notion of ϵ-permutability in S. For example, S is said to be *ϵ-commutative* if $\langle a,b \rangle = 0$ for all $a,b \in S$.

(4) The following special case of example (3) is particularly important. Let ϵ be a commutation factor on Γ and let V be a Γ-graded vector space. We know that Homgr(V) (see Sec. 2) is an associative Γ-graded algebra. The ϵ Lie algebra which is associated with Homgr(V) will be called the *general linear ϵ Lie algebra* of the Γ-graded vector space V and will be denoted by gl(V,ϵ), it is analogous to the general linear Lie (super) algebra of a (Z_2-graded) vector space. In particular, we define

Definition 3: A *graded representation* of an ϵ Lie algebra L in a Γ-graded vector space V is a homomorphism $\rho: L \rightarrow$ gl(V,ϵ) of Γ-graded ϵ Lie algebras. (According to our conventions this implies that ρ is homogeneous of degree zero.)

It follows from the generalized Poincaré–Birkhoff–Witt theorem (see Sec. 4) that every ϵ Lie algebra L has a faithful graded representation in some Γ-graded vector space V, i.e., that L is isomorphic to a graded subalgebra of

$\mathrm{gl}(V,\epsilon)$. Moreover, the generalized Ado theorem (see Sec. 7) says that if L is finite-dimensional, then V can be chosen to be finite-dimensional as well.

Obviously, our discussion suggests that the standard methods for constructing subalgebras of a general linear Lie (super) algebra should have their direct counterparts in the present more general situation. We shall not go into detail here but refer the reader to Ref. 12.

To conclude this section we remark that every ϵ Lie algebra L has a natural Z_2-gradation. In fact, it follows from Eq. (3.1b) that the mapping

$$\Gamma \to K_*, \quad \alpha \to \epsilon(\alpha,\alpha) \tag{3.10}$$

is a homomorphism of groups and that

$$\epsilon(\alpha,\alpha) = \pm 1 \quad \text{for all } \alpha \in \Gamma. \tag{3.11}$$

Let us define

$$\Gamma_0 = \{\alpha \in \Gamma | \epsilon(\alpha,\alpha) = 1\}, \tag{3.12a}$$

$$\Gamma_1 = \{\alpha \in \Gamma | \epsilon(\alpha,\alpha) \neq 1\}. \tag{3.12b}$$

Then Γ_0 is a subgroup of Γ, and either we have $\Gamma_0 = \Gamma$ (and hence $\Gamma_1 = \emptyset$) or else Γ_0 is a subgroup of index 2 in Γ, and Γ_0, Γ_1 are the two residue classes modulo Γ_0.

We now set

$$L^{(i)} = \bigoplus_{\alpha \in \Gamma_i} L_\alpha \quad \text{for } i = 0,1, \tag{3.13}$$

and consider the indices 0,1 as integers modulo 2; then it follows from the above that the decompositon

$$L = L^{(0)} \oplus L^{(1)} \tag{3.14}$$

is a Z_2 gradation of the algebra L. In Ref. 7 the ϵ Lie algebras with $\Gamma_0 = \Gamma$ (resp. $\Gamma_0 \neq \Gamma$) are called color algebras (resp. color superalgebras).

4. THE ENVELOPING ALGEBRA OF AN ϵ LIE ALGEBRA

In the present section we shall introduce the so-called (universal) enveloping algebra of an ϵ Lie algebra and discuss some of its fundamental properties. The constructions and proofs turn out to be completely analogous to those in the case of Lie (super) algebras.[13-15] Thus it should be sufficient to quote the main results. Throughout this section ϵ denotes a fixed commutation factor on Γ.

A. Definition of the enveloping algebra

Let L be an ϵ Lie algebra and let $T(L)$ be the tensor algebra of the Γ-graded vector space L. As is well-known, $T(L)$ has a natural $Z \times \Gamma$-gradation which is fixed by the condition that the degree of a tensor $A_1 \otimes \cdots \otimes A_n$, with $A_i \in L_{\alpha_i}$, $\alpha_i \in \Gamma$ for $1 \leqslant i \leqslant n$, is equal to $(n, \alpha_1 + \cdots + \alpha_n)$. The subspace of $T(L)$ consisting of the homogeneous tensors of order $n \in Z$ will be denoted by $T_n(L)$; of course, $T_n(L) = \{0\}$ if $n \leqslant -1$.

Let $J(L)$ be the two-sided ideal of $T(L)$ which is generated by the tensors of the form

$$A \otimes B - \epsilon(\alpha,\beta) B \otimes A - \langle A,B \rangle,$$
$$\text{with } A \in L_\alpha, \quad B \in L_\beta \quad \alpha,\beta \in \Gamma. \tag{4.1}$$

The quotient algebra

$$U(L) = T(L)/J(L) \tag{4.2}$$

is called the (universal) *enveloping algebra* of the ϵ Lie algebra L. Obviously, $U(L)$ is associative and has a unit element.

By composing the canonical injection $L \to T(L)$ with the canonical mapping $T(L) \to T(L)/J(L) = U(L)$ we obtain the canonical mapping

$$\sigma : L \to U(L), \tag{4.3}$$

which (by definition) satisfies

$$\sigma(\langle A,B \rangle) = \sigma(A)\sigma(B) - \epsilon(\alpha,\beta)\sigma(B)\sigma(A),$$
$$\text{for all } A \in L_\alpha, B \in L_\beta, \ \alpha,\beta \in \Gamma. \tag{4.4}$$

With respect to the Γ-gradation of $T(L)$ the element (4.1) is homogeneous of degree $\alpha + \beta$. It follows that $U(L)$ inherits a natural Γ-gradation from $T(L)$ which turns $U(L)$ into a Γ-graded algebra. On the other hand, the Z-gradation of $T(L)$ in general only leads to a filtration of $U(L)$, as follows. For any element $n \in Z$, let $T^n(L) = \oplus_{m \leqslant n} T_m(L)$ and let $U^n(L)$ be the canonical image of $T^n(L)$ in $U(L)$. The family $(U^n(L))_{n \in Z}$ is called the *canonical filtration* of $U(L)$. Evidently, $(U^n(L))_{n\in Z}$ is an increasing family of Γ-graded subspaces of $U(L)$, the union of the subspaces $U^n(L)$, $n \in Z$, is equal to $U(L)$. Finally,

$$U^n(L) U^m(L) \subset U^{n+m}(L), \quad \text{for all } n,m \in Z. \tag{4.5}$$

The pair $(U(L),\sigma)$ has the following standard universal property.

Proposition 1: Let L be an ϵ Lie algebra, let $U(L)$ be its enveloping algebra and let $\sigma : L \to U(L)$ be the canonical mapping. Suppose we are given an associative algebra S with unit element and a linear mapping

$$g : L \to S, \tag{4.6}$$

such that

$$g(\langle A,B \rangle) = g(A)g(B) - \epsilon(\alpha,\beta)g(B)g(A),$$
$$\text{for all } A \in L_\alpha, B \in L_\beta, \alpha,\beta \in \Gamma. \tag{4.7}$$

Then there exists a unique homomorphism $\bar{g}$ of the algebra $U(L)$ into the algebra S which maps 1 onto 1 such that

$$g = \bar{g} \circ \sigma. \tag{4.8}$$

If, in addition, S is Γ-graded and if g is homogeneous of degree zero, then $\bar{g}$ is homogeneous of degree zero, too.

B. The ϵ-symmetric algebra of a Γ-graded vector space

Let V be a Γ-graded vector space. We endow V with the trivial "Abelian" structure of an ϵ Lie algebra by defining

$$\langle A,B \rangle = 0, \quad \text{for all } A,B \in V. \tag{4.9}$$

The enveloping algebra of this ϵ Lie algebra will be called the *ϵ-symmetric algebra* of the Γ-graded vector space V, and will be denoted by $\widetilde{U}(V)$. Thus, by definition, we have

$$\widetilde{U}(V) = T(V)/\widetilde{J}(V), \quad (4.10)$$

where $\widetilde{J}(V)$ is the two-sided ideal of $T(V)$ which is generated by the tensors of the form

$$A \otimes B - \epsilon(\alpha,\beta)B \otimes A, \quad \text{with } A \in V_\alpha, B \in V_\beta, \alpha,\beta \in \Gamma. \quad (4.11)$$

Obviously, these tensors are homogeneous with respect to the $Z \times \Gamma$-gradation of $T(V)$; thus $\widetilde{U}(V)$ has a natural $Z \times \Gamma$-gradation as well. The algebra $\widetilde{U}(V)$ is associative, ϵ-commutative (see Sec. 3, example 3) and has a unit element.

C. The generalized Poincaré–Birkhoff–Witt theorem

As before, let L be an ϵ Lie algebra, let $U(L)$ be the enveloping algebra of L and let $(U^n(L))_{n\in Z}$ be the canonical filtration of $U(L)$. We define

$$G_n(L) = U^n(L)/U^{n-1}(L) \quad \text{for all } n \in Z, \quad (4.12)$$

and set

$$G(L) = \bigoplus_{n\in Z} G_n(L). \quad (4.13)$$

Evidently, $G_n(L)$ is a Γ-graded vector space, for all $n \in Z$; consequently, $G(L)$ is a $Z \times \Gamma$-graded vector space.

We next introduce on $G(L)$ an algebra structure, as follows. The relation (4.5) shows that, for all $n,m \in Z$, the product mapping in $U(L)$ defines a bilinear mapping $U^n(L) \times U^m(L) \to U^{n+m}(L)$ which, by going to the quotients, induces a bilinear mapping $G_n(L) \times G_m(L) \to G_{n+m}(L)$. The family of these latter mappings yields a bilinear mapping $G(L) \times G(L) \to G(L)$ which is the product mapping we are looking for. It is easy to see that $G(L)$, endowed with the $Z \times \Gamma$-gradation and the multiplication introduced above, is an associative $Z \times \Gamma$-graded algebra with a unit element. The algebra $G(L)$ is called the graded algebra associated with the filtration $(U^n(L))_{n\in Z}$.

On the other hand there exists, for every $n \in Z$, a canonical mapping

$$\varphi_n : T_n(L) \to G_n(L), \quad (4.14)$$

which is defined to be the composition of the canonical mapping $T_n(L) \to U^n(L)$ with the canonical mapping $U^n(L) \to G_n(L)$. Let

$$\varphi : T(L) \to G(L) \quad (4.15)$$

be the linear mapping which is defined by the family $(\varphi_n)_{n\in Z}$. It is easy to see that φ is a surjective homomorphism of $Z \times \Gamma$-graded algebras which vanishes on the ideal $\widetilde{J}(L)$ (see subsec. B). Thus, by going to the quotient, φ defines a canonical surjective homomorphism,

$$\omega : \widetilde{U}(L) \to G(L), \quad (4.16)$$

of $Z \times \Gamma$-graded algebras.

The generalized Poincaré–Birkhoff–Witt theorem now reads as follows.

Theorem 1: Let L be an ϵ Lie algebra. The canonical homomorphism ω: $\widetilde{U}(L) \to G(L)$ is an isomorphism of $Z \times \Gamma$-graded algebras.

Of the various corollaries to this theorem we only mention the following ones. Actually, in a sense, the first corollary is equivalent to the theorem itself.

Corollary 1: Let L be an ϵ Lie algebra and let $(E_i)_{i\in I}$ be a basis of the vector space L which is indexed by a totally ordered set I. We suppose that the elements E_i, $i \in I$, are homogeneous, let $\eta_i \in \Gamma$ be the degree of E_i.

If $(i_1,\cdots,i_r)$ runs through all finite sequences in I such that, for $1 \leqslant s \leqslant r-1$, we have $i_s \leqslant i_{s+1}$ and even $i_s < i_{s+1}$ if $\epsilon(\eta_{i_s},\eta_{i_s}) \neq 1$, then the products,

$$\sigma(E_{i_1})\sigma(E_{i_2})\cdots\sigma(E_{i_r}), \quad (4.17)$$

form a basis of the vector space $U(L)$. [For $r = 0$ we define the product (4.17) to be equal to 1.]

Corollary 2: For any ϵ Lie algebra L the canonical mapping $\sigma: L \to U(L)$ is injective.

This corollary implies that the ϵ Lie algebra L has a faithful graded representation in the Γ-graded vector space $U(L)$.

Remark: Evidently, the definitions of a commutation factor and of an ϵ Lie algebra make sense even if Γ is only supposed to be an Abelian semigroup (i.e., if Γ is a set endowed with a composition which is associative and commutative). Moreover, the constructions described in this section still can be carried through and the generalized Poincaré–Birkhoff–Witt theorem as well as its corollaries remain valid.

Once the generalized Poincaré–Birkhoff–Witt theorem is established, we have the powerful technique of induced representations at our disposal.[2,16] We shall not go into any details, but only mention one result which is easily obtained on these grounds and which will be needed in the proof of the generalized Ado theorem (see Sec. 7).

Lemma 1: Let L be an ϵ Lie algebra. Define Γ_0 and $L^{(0)}$ by Eqs. (3.12) and (3.13) of Sec. 3. Let $\epsilon^{(0)}$ be the restriction of ϵ to $\Gamma_0 \times \Gamma_0$. Evidently, $L^{(0)}$ may be considered as a Γ_0-graded $\epsilon^{(0)}$ Lie algebra.

Suppose that L is finite-dimensional. If the $\epsilon^{(0)}$ Lie algebra $L^{(0)}$ has a faithful finite-dimensional Γ_0-graded representation, then the ϵ Lie algebra L has a faithful finite-dimensional Γ-graded representation.

5. COMMUTATION FACTORS ON FINITELY GENERATED ABELIAN GROUPS

Let Γ be a finitely generated Abelian group (i.e., an Abelian group that is generated by a finite number of its elements). According to a well-known theorem[17] this means that Γ is the direct sum of finitely many cyclic groups Γ_r, $1 \leqslant r \leqslant n$. Thus we may assume that

$$\Gamma = \Gamma_1 \oplus \cdots \oplus \Gamma_n \quad (5.1)$$

and

$$\Gamma_r = Z/m_r Z, \tag{5.2}$$

with some integers $m_r \geqslant 0$, $1 \leqslant r \leqslant n$. For each r, $1 \leqslant r \leqslant n$, we choose a generator ξ_r of the group Γ_r (a natural choice for ξ_r is the residue class of 1 modulo m_r).

Now let ϵ be any commutation factor on Γ and let r,s be any integers such that $1 \leqslant r,s \leqslant n$. We define

$$\epsilon(\xi_r,\xi_s) = A_{rs}. \tag{5.3}$$

Evidently, we have

$$A_{rs}A_{sr} = 1. \tag{5.4}$$

Furthermore, the obvious equation,

$$\epsilon(m_r\xi_r,\xi_s) = \epsilon(\xi_r,m_s\xi_s) = 1, \tag{5.5}$$

means that

$$A_{rs}^{m_r} = A_{rs}^{m_s} = 1. \tag{5.6}$$

Let m_{rs} be the greatest common divisor of m_r and m_s (recall that the greatest common divisor of 0 and any integer $m \geqslant 0$ is equal to m). Then Eq. (5.6) is equivalent to

$$A_{rs}^{m_{rs}} = 1. \tag{5.7}$$

In the special case $r = s$ Eqs. (5.4) and (5.7) combine to yield

$$A_{rr} = \begin{cases} 1 & \text{if } m_r \text{ is odd,} \\ \pm 1 & \text{if } m_r \text{ is even.} \end{cases} \tag{5.8}$$

Finally, we have

$$\epsilon\Big(\sum_r p_r\xi_r,\sum_s q_s\xi_s\Big) = \prod_t A_{tt}^{p_tq_t} \prod_{r<s} A_{rs}^{(p_rq_s - p_sq_r)}$$

$$\text{for all } p_r,q_r \in Z, \quad 1 \leqslant r \leqslant n. \tag{5.9}$$

Conversely, let A_{rs}, $1 \leqslant r \leqslant s \leqslant n$, be any system of elements from K_* such that Eqs. (5.7) and (5.8) are satisfied. Then Eq. (5.9) defines a commutation factor ϵ on Γ. Thus Eqs. (5.7), (5.8), and (5.9) completely characterize the commutation factors on Γ.

The above result implies the following simple lemma, which will be needed in the subsequent section.

Lemma 2: Let Γ be a finitely generated Abelian group and let ϵ be a commutation factor on Γ such that

$$\epsilon(\alpha,\alpha) = 1 \text{ for all } \alpha \in \Gamma. \tag{5.10}$$

Then there exists a mapping,

$$\sigma: \Gamma \times \Gamma \to K_*, \tag{5.11}$$

such that

$$\epsilon(\alpha,\beta) = \sigma(\alpha,\beta)\sigma(\beta,\alpha)^{-1}, \tag{5.12}$$

$$\sigma(\alpha,\beta+\gamma) = \sigma(\alpha,\beta)\sigma(\alpha,\gamma), \tag{5.13}$$

$$\sigma(\alpha+\beta,\gamma) = \sigma(\alpha,\gamma)\sigma(\beta,\gamma), \tag{5.14}$$

for all $\alpha,\beta,\gamma \in \Gamma$.

Proof: We use the notation introduced above. Assumption (5.10) implies that

$$A_{tt} = 1, \quad \text{for } 1 \leqslant t \leqslant n. \tag{5.15}$$

Define

$$\sigma\Big(\sum_r p_r\xi_r, \sum_s q_s\xi_s\Big) = \prod_{r<s} A_{rs}^{p_rq_s},$$

$$\text{for all } p_r,q_r \in Z, \quad 1 \leqslant r \leqslant n. \tag{5.16}$$

Then σ meets our requirements.

As a next step one would like to classify the commutation factors up to equivalence, in the sense of the following definition.

Definition 4: Let Γ be an Abelian group. Two commutation factors ϵ and ϵ' on Γ are said to be equivalent if there exists an automorphism h of Γ such that

$$\epsilon'(\alpha,\beta) = \epsilon(h(\alpha),h(\beta)) \quad \text{for all } \alpha,\beta \in \Gamma. \tag{5.17}$$

In general, this classification will be a difficult problem since the decomposition (5.1) of Γ into a direct sum of cyclic subgroups is far from unique. Thus the automorphisms of Γ will not, in general, preserve this decomposition.

Therefore, we shall restrict our attention to the special case where

$$\Gamma = Z_p^n, \tag{5.18}$$

with $Z_p = Z/pZ$ and p a prime number. In this case Z_p is a field and Z_p^n is a vector space over Z_p. Consequently, the theory of finite-dimensional vector spaces is applicable here.

First of all, we conclude from Eq. (5.7) that if a nontrivial commutation factor on Γ exists at all, then K must contain a pth root of unity which is different from 1. This will be assumed in the following.

Remark for the case of prime characteristic: If $\operatorname{char}K = p$ then 1 is the sole pth root of unity. Thus our assumption implies that $\operatorname{char}K \neq p$.

Let us choose a pth root of unity $E_p \in K$ which is different from 1. [In the complex case a natural choice would be $E_p = \exp(2\pi i/p)$.] Then any other pth root of unity is a power of E_p.

We next remark that for any integer m the power E_p^m only depends on the residue class of m modulo p. Thus for any element $\lambda \in Z_p$ the expression E_p^λ has a well-defined meaning.

Using these remarks as well as our earlier results it is easy to see that the commutation factors on Z_p^n are just the mappings of the form

$$\epsilon(\alpha,\beta) = E_p^{\psi(\alpha,\beta)}, \quad \text{for all } \alpha,\beta \in \Gamma, \tag{5.19}$$

where ψ is a bilinear form on the vector space Z_p^n over Z_p which is symmetric if $p = 2$, but alternating if $p \geqslant 3$. Note that ψ is uniquely determined by ϵ.

Remarks: (1) We recall that a bilinear form ψ on a vector space V over a field F is called alternating if $\psi(x,x) = 0$ for all $x \in V$. If $\operatorname{char}F \neq 2$, a bilinear form is alternating if and only if it is skew-symmetric. However, this is no longer true if $\operatorname{char}F = 2$. Note that in the latter case the symmetry and skew-symmetry of a bilinear form mean just the same thing.

(2) The difference between the cases $p = 2$ and $p \geqslant 3$ is easily read off from Eq. (5.8). If $p \geqslant 3$ this equation yields $A_{rr} = 1$ for $1 \leqslant r \leqslant n$ and hence there are "no diagonal terms"

in Eq. (5.9). On the other hand, for $p = 2$ we may have diagonal terms as well.

Standard results on the normal forms of alternating or symmetric bilinear forms on a finite-dimensional vector space[18] now yield the following proposition.

Proposition 2: Let p be a prime number, let $E_p \in K$ be a nontrivial pth root of unity, and let $n \geqslant 1$ be an integer. For any element $\alpha \in Z_p^n$ the coordinates of α with respect to the canonical basis of Z_p^n will be denoted by $\alpha_1,\ldots,\alpha_n$.

Let r be an integer such that $0 \leqslant 2r \leqslant n$. We define an alternating bilinear form ψ_r on Z_p^n by

$$\psi_r(\alpha,\beta) = \sum_{i=1}^{r} (\alpha_{2i-1}\beta_{2i} - \alpha_{2i}\beta_{2i-1}) \quad \text{for all } \alpha,\beta \in Z_p^n. \tag{5.20}$$

Let s be an integer such that $1 \leqslant s \leqslant n$. We define a symmetric bilinear form ϕ_s on Z_p^n by

$$\phi_s(\alpha,\beta) = \sum_{i=1}^{s} \alpha_i \beta_i, \quad \text{for all } \alpha,\beta \in Z_p^n. \tag{5.21}$$

(a) If $p = 2$, then any commutation factor on Z_2^n is equivalent to $(-1)^{\psi}$ with ψ equal to one of the forms ψ_r or ϕ_s.

(b) If $p \geqslant 3$, then any commutation factor on Z_p^n is equivalent to E_p^{ψ} with ψ equal to one of the forms ψ_r.

Remark: For $p = 2$ the result has already been proved.[7]

6. CHANGE OF THE COMMUTATION FACTOR

In this section we shall establish a close relationship between ϵ Lie algebras corresponding to different commutation factors.

Let Γ be an Abelian group, let ϵ be a commutation factor on Γ, and let L be an ϵ Lie algebra. Given any mapping

$$\sigma : \Gamma \times \Gamma \to K_*, \tag{6.1}$$

we define on the Γ-graded vector space L a new multiplication $\langle\ ,\ \rangle^\sigma$ by the requirement that

$$\langle A,B \rangle^\sigma = \sigma(\alpha,\beta)\langle A,B \rangle \quad \text{for all } A \in L_\alpha,\ B \in L_\beta,\ \alpha,\beta \in \Gamma. \tag{6.2}$$

The Γ-graded vector space L, endowed with the multiplication $\langle\ ,\ \rangle^\sigma$, is a Γ-graded algebra which will be denoted by L^σ.

Let us define a mapping

$$\delta : \Gamma \times \Gamma \to K_* \tag{6.3a}$$

by

$$\delta(\alpha,\beta) = \sigma(\alpha,\beta)\sigma(\beta,\alpha)^{-1}, \quad \text{for all } \alpha,\beta \in \Gamma. \tag{6.3b}$$

Obviously, we have

$$\langle A,B \rangle^\sigma = -\epsilon(\alpha,\beta)\delta(\alpha,\beta)\langle B,A \rangle^\sigma \quad \text{for all } A \in L_\alpha,\ B \in L_\beta,\ \alpha,\beta \in \Gamma. \tag{6.4}$$

We are looking for conditions on σ which ensure that $\epsilon\delta$ (i.e., that δ) is a commutation factor on Γ and that L^σ is an $\epsilon\delta$ Lie algebra.

It is easy to see that δ is a commutation factor if and only if

(I) $$\sigma(\alpha,\beta+\gamma)\sigma(\alpha,\beta)^{-1}\sigma(\alpha,\gamma)^{-1} = \sigma(\beta+\gamma,\alpha)\sigma(\beta,\alpha)^{-1}\,\sigma(\gamma,\alpha)^{-1}, \quad \text{for all } \alpha,\beta,\gamma \in \Gamma. \tag{6.5}$$

Furthermore, the product $\langle\ ,\ \rangle^\sigma$ satisfies the $\epsilon\delta$ Jacobi identity if (and, generically speaking, only if)

(II) $\sigma(\alpha,\beta+\gamma)\sigma(\alpha,\beta)^{-1}\sigma(\alpha,\gamma)^{-1}$ is invariant under cyclic permutations of α,β,γ, for all $\alpha,\beta,\gamma \in \Gamma$.

Finally, it is not difficult to check that Conditions (I) and (II) are satisfied simultaneously if and only if

$$\sigma(\alpha,\beta+\gamma)\sigma(\beta,\gamma) = \sigma(\alpha,\beta)\sigma(\alpha+\beta,\gamma), \quad \text{for all } \alpha,\beta,\gamma \in \Gamma. \tag{6.6}$$

Definition 5: Let Γ be an Abelian group. A mapping $\sigma : \Gamma \times \Gamma \to K_*$ satisfying Condition (6.6) is called a *multiplier* on Γ.

For any multiplier σ, the mapping δ defined by Eq. (6.3) is a commutation factor on Γ which is said to be associated with σ.

Note that

$$\delta(\alpha,\alpha) = 1 \quad \text{for all } \alpha \in \Gamma. \tag{6.7}$$

Our discussion above now shows:

Proposition 3: Let Γ be an Abelian group, let ϵ be a commutation factor on Γ, and let L be an ϵ Lie algebra. Suppose we are given a multiplier σ on Γ; let δ be the commutation factor associated with it. Then the Γ-graded algebra L^σ defined above is an $\epsilon\delta$ Lie algebra.

Remark: Choosing $\beta = 0$ in Eq. (6.6) we conclude that a multiplier σ satisfies

$$\sigma(0,\alpha) = \sigma(\alpha,0) = \sigma(0,0) \quad \text{for all } \alpha \in \Gamma. \tag{6.8}$$

Occasionally, it is convenient to impose the normalization condition $\sigma(0,0) = 1$, and some authors include this condition in the definition of a multiplier.

The defining equation (6.6) of a multiplier has an obvious generalization to arbitrary groups. These general multipliers (or factor systems, as they are also called) are well-known objects in mathematics. For example, they play a crucial role in the theory of projective (i.e., ray) representations of a group.[19]

Resuming our general discussion we remark that the relationship described in Proposition 3 extends to the graded representations of the algebras. In fact, it is easy to prove the following proposition.

Proposition 4: Using the notation introduced in Proposition 3 let ρ be a graded representation of the ϵ Lie algebra L in some Γ-graded vector space V. Define a linear mapping,

$$\rho^\sigma : L^\sigma \to \text{Homgr}(V), \tag{6.9a}$$

by the requirement that

$$\rho^\sigma(A)x = \sigma(\alpha,\xi)\rho(A)x \tag{6.9b}$$

for all $A \in L_\alpha$, $x \in V_\xi; \alpha,\xi \in \Gamma$.

Then ρ^σ is a graded representation of the $\epsilon\delta$ Lie algebra L^σ in V.

The above correspondence between ϵ Lie algebras and $\epsilon\delta$ Lie algebras and between their representations has nice functorial properties. Let us mention a few of them.

(1) Obviously, the correspondence is bijective, the inverse transformation being given by σ^{-1} in place of σ.

(2) A graded subspace of L is a graded subalgebra (resp. a graded ideal) of L if and only if it is a graded subalgebra (resp. a graded ideal) of L^σ.

(3) A graded subspace of V is invariant under the representation ρ if and only if it is invariant under the representation ρ^σ.

(4) Let L' be a second ϵ Lie algebra. A homomorphism of the ϵ Lie algebra L into the ϵ Lie algebra L' is also a homomorphism of the $\epsilon\delta$ Lie algebra L^σ into the $\epsilon\delta$ Lie algebra L'^σ.

It is quite instructive to look at the above constructions from a different point of view. As before, let ϵ be a commutation factor on Γ and let L be an ϵ Lie algebra. Suppose we are given a Γ-graded associative ϵ'-commutative algebra S, where ϵ' is some commutation factor on Γ [see Sec. 3, example 3]. On the Γ-graded vector space,

$$\hat{L} = \bigoplus_{\alpha\in\Gamma} \hat{L}_\alpha, \quad \text{with} \quad \hat{L}_\alpha = S_\alpha \otimes L_\alpha \text{ for all } \alpha\in\Gamma, \tag{6.10}$$

we define a multiplication by the requirement that

$$\langle a \otimes A, b \otimes B\rangle = ab \otimes \langle A,B\rangle$$
$$\text{for all } a\in S_\alpha,\ b\in S_\beta,\ A\in L_\alpha,\ B\in L_\beta,\ \alpha,\beta\in\Gamma. \tag{6.11}$$

Then it is easy to see that $\hat{L}$ is a Γ-graded $\epsilon\epsilon'$ Lie algebra.

Now let σ be any multiplier on Γ. We are going to construct an algebra S of the type described above (the so-called crossed product of K and Γ with respect to the multiplier σ). As a vector space, S is the space of all functions $g:\Gamma\to K$ such that $g(\gamma)=0$ for all but a finite number of elements $\gamma\in\Gamma$. The multiplication in S is defined by

$$(fg)(\gamma) = \sum_{\alpha+\beta=\gamma} \sigma(\alpha,\beta) f(\alpha) g(\beta) \text{ for all } f,g\in S. \tag{6.12}$$

Let $(e_\alpha)_{\alpha\in\Gamma}$ be the canonical basis of S; by definition, $e_\alpha:\Gamma\to K$ is the function which satisfies

$$e_\alpha(\gamma) = \begin{cases} 1 & \text{if } \gamma=\alpha, \\ 0 & \text{if } \gamma\neq\alpha. \end{cases} \tag{6.13}$$

According to Eq. (6.12) we have

$$e_\alpha e_\beta = \sigma(\alpha,\beta) e_{\alpha+\beta} \quad \text{for all } \alpha,\beta\in\Gamma. \tag{6.14}$$

Consequently, Eq. (6.6) is a necessary and sufficient condition for S to be associative, and Eq. (6.8) shows that $\sigma(0,0)^{-1} e_0$ is the unit element of S.

The algebra S has a natural Γ-gradation; we simply define

$$S_\alpha = K e_\alpha \quad \text{for all } \alpha\in\Gamma. \tag{6.15}$$

Let δ be the commutation factor associated with σ. Since, obviously,

$$e_\alpha e_\beta = \delta(\alpha,\beta) e_\beta e_\alpha \quad \text{for all } \alpha,\beta\in\Gamma, \tag{6.16}$$

we have shown that S is a Γ-graded associative δ-commutative algebra.

Therefore, the above construction applies and yields an $\epsilon\delta$ Lie algebra $\hat{L}$. This algebra is isomorphic to L^σ. In fact, the linear mapping,

$$\lambda : L^\sigma \to \hat{L}, \tag{6.17a}$$

defined by

$$\lambda(A) = e_\alpha \otimes A \quad \text{for all } A\in L_\alpha,\ \alpha\in\Gamma, \tag{6.17b}$$

is an isomorphism of Γ-graded algebras.

A similar construction can be carried out for the graded representations.

The foregoing discussion reveals that our construction is completely analogous to the one which, by tensoring with a Grassmann algebra, converts a Lie superalgebra into a Lie algebra.

We shall now give the most important application of the results obtained in this section. Let ϵ be a commutation factor on Γ. As in Sec. 3 we define

$$\Gamma_0 = \{\alpha\in\Gamma | \epsilon(\alpha,\alpha)=1\}, \tag{6.18a}$$

$$\Gamma_1 = \{\alpha\in\Gamma | \epsilon(\alpha,\alpha)\neq 1\}. \tag{6.18b}$$

Recall that Γ_0 is a subgroup of Γ and that either $\Gamma_0=\Gamma$ or else Γ_0 is a subgroup of index 2 in Γ and Γ_0, Γ_1 are the corresponding residue classes. We define a mapping,

$$\epsilon_0 : \Gamma\times\Gamma \to K_*, \tag{6.19a}$$

as follows: If $\alpha,\beta\in\Gamma$, we set

$$\epsilon_0(\alpha,\beta) = \begin{cases} -1 & \text{if } \alpha,\beta\in\Gamma_1, \\ 1 & \text{otherwise.} \end{cases} \tag{6.19b}$$

Obviously, ϵ_0 is a commutation factor on Γ, hence

$$\delta = \epsilon_0\epsilon \tag{6.20}$$

is a commutation factor as well. Note that

$$\delta(\alpha,\alpha) = 1 \quad \text{for all } \alpha\in\Gamma. \tag{6.21}$$

Suppose now that there exists a multiplier σ on Γ such that δ is the commutation factor associated with σ. Then the foregoing constructions apply. But according to Lemma 2 of Sec. 5 the existence of σ is guaranteed if the group Γ is finitely generated. Expressed in a somewhat oversimplified language, we thus have shown,

Theorem 2: Let Γ be a finitely generated Abelian group and let ϵ be a commutation factor on Γ. Define the commutation factor ϵ_0 on Γ by Eq. (6.19), then choose a multiplier σ on Γ such that the commutation factor associated with σ is equal to $\epsilon_0\epsilon$ (this is possible).

Then $L\to L^\sigma$ is a bijection between the ϵ_0 Lie algebras and the ϵ Lie algebras; furthermore, for any ϵ_0 Lie algebra L, the transformation $\rho\to\rho^\sigma$ is a bijection between the graded representations of L and the graded representations of L^σ.

To explain the content of this theorem let us stress that an ϵ_0 Lie algebra is nothing but a Γ-graded Lie superalgebra (with Γ_0 being the subgroup of even degrees). In particular, if ϵ is a commutation factor such that $\epsilon(\alpha,\alpha)=1$ for all $\alpha\in\Gamma$, then ϵ_0 is trivial and the ϵ_0 Lie algebras are just the Γ-graded Lie algebras.

Remark: Theorem 2 generalizes and sharpens a result which has been obtained in Ref. 7. Therein the authors treat the special case where $\Gamma = Z_2^n$ and use a Clifford algebra to establish a correspondence between ϵ Lie algebras and Z_2^n-graded Lie superalgebras.

Incidentally, it is well-known that a Clifford algebra may be considered as a crossed product of K and Z_2^n with respect to a suitable multiplier (provided that char$K \neq 2$).

We conclude this section by a comment on the question of how the construction summarized in Theorem 2 will depend on the special choice of σ.

Let σ_1 and σ_2 be two multipliers on Γ. Then σ_1 and σ_2 yield the same commutation factor if and only if $\tau = \sigma_1\sigma_2^{-1}$ is symmetric, i.e., if and only if

$$\tau(\alpha,\beta) = \tau(\beta,\alpha) \quad \text{for all } \alpha,\beta \in \Gamma. \tag{6.22}$$

It is easy to construct a large class of symmetric multipliers on Γ as follows. Let ω be an arbitrary mapping of Γ into K_*. Then the mapping

$$\tau : \Gamma \times \Gamma \rightarrow K_*, \tag{6.23a}$$

defined by

$$\tau(\alpha,\beta) = \omega(\alpha+\beta)\omega(\alpha)^{-1}\omega(\beta)^{-1} \quad \text{for all } \alpha,\beta \in \Gamma, \tag{6.23b}$$

is a symmetric multiplier on Γ. These multipliers are called trivial.

Conversely, one can prove;

Proposition 5: Let Γ be a finitely generated Abelian group. If the field K is algebraically closed, then every symmetric multiplier on Γ is trivial.

Suppose now that Γ is a finitely generated Abelian group and that ϵ is a commutation factor on Γ. If the field K is algebraically closed we can use Proposition 5 to show that, up to isomorphism, the construction summarized in Theorem 2 does not depend on the special choice of σ. We shall not go into the details.

7. THE GENERALIZED ADO THEOREM

We now are ready to prove the generalized Ado theorem, which reads as follows.

Theorem 3: Let Γ be an Abelian group, let ϵ be a commutation factor on Γ, and let L be a finite-dimensional ϵ Lie algebra. Then L has a faithful finite-dimensional Γ-graded representation.

Proof: Set

$$\Delta = \{\alpha \in \Gamma | L_\alpha \neq \{0\}\}. \tag{7.1}$$

By assumption, L is finite-dimensional, hence the set Δ is finite. Let Γ' be the subgroup of Γ which is generated by Δ and let ϵ' be the restriction of ϵ to $\Gamma' \times \Gamma'$. Of course, L is a Γ'-graded ϵ' Lie algebra in an obvious way.

Suppose we can find a faithful graded representation ρ of the ϵ' Lie algebra L in a finite-dimensional Γ'-graded vector space V. We extend the Γ'-gradation of V into a Γ-gradation by setting $V_\gamma = \{0\}$ if $\gamma \in \Gamma$, $\gamma \notin \Gamma'$. Then ρ is a faithful finite-dimensional graded representation of the ϵ Lie algebra L in the Γ-graded vector space V.

Thus we now may assume that the group Γ is finitely generated. According to Lemma 1 of Sec. 4 we may assume in addition that

$$\epsilon(\alpha,\alpha) = 1 \quad \text{for all } \alpha \in \Gamma. \tag{7.2}$$

But then it follows from Theorem 2 of Sec. 6 that it is sufficient to prove the theorem for Γ-graded Lie algebras. The following theorem by Ross[14] now yields the required result: Let Γ be an Abelian group and let L be a finite-dimensional Γ-graded Lie algebra. Then L has a faithful finite-dimensional Γ-graded representation.

Remark: Ross[14] has made the general assumptions that Γ contains a subgroup of index 2 (the subgroup of even degrees), and that char$K \neq 2$. However, these assumptions do not enter into his proof of the theorem cited above. Incidentally, Ross has used this latter theorem to obtain the generalized Ado theorem for finite-dimensional Γ-graded Lie superalgebras.

8. CONCLUSION

In the present work we have investigated a class of generalized graded Lie algebras, the so-called ϵ Lie algebras. The defining identities (a certain skew-symmetry condition and a generalized Jacobi identity) depend on the degrees of the elements under consideration through the commutation factor ϵ. We have seen (at least on the rather general level of this article) that the ϵ Lie algebras have properties quite similar to those of Lie (super) algebras. For example, the generalized Poincaré–Birkhoff–Witt and Ado theorems hold true.

A deeper reason for this similarity has been discussed in Sec. 6, where we have reduced the theory of ϵ Lie algebras and their graded representations to the theory of graded Lie (super) algebras and their graded representations (provided that the Abelian group of degrees is finitely generated). Any result of this latter theory will therefore lead to a corresponding one for ϵ Lie algebras. This point is clearly visible in our proof of the generalized Ado theorem, where the theorem by Ross has been an important ingredient.

Note added in proof: Part of this work has been presented at the Integrative Conference on Group Theory and Mathematical Physics, Austin, Texas, 1978.

[1]A survey of the subject has been given by P. Fayet and S. Ferrara, Phys. Rep. C **32**, 249 (1977).
[2]V.G. Kac, Adv. Math. **26**, 8 (1977); see also V.G. Kac, Commun. Math. Phys. **53**, 31 (1977).
[3]P.G.O. Freund and I. Kaplansky, J. Math. Phys. **17**, 228 (1976).
[4]M. Scheunert, W. Nahm, and V. Rittenberg, J. Math. Phys. **17**, 1626, 1640 (1976).
[5]B. Kostant, in *Differential Geometrical Methods in Mathematical Physics*, Bonn 1975, Lecture Notes in Mathematics 570 (Springer, Berlin, 1977).
[6]F.A. Berezin, Inst. of Theor. and Exp. Physics, ITEP-66, 75–78 (5 parts), Moscow (1977).
[7]V. Rittenberg and D. Wyler, Nucl. Phys. B **139**, 189 (1978).

[8]V. Rittenberg, "Beyond Supersymmetry," *The Johns Hopkins University Workshop on Current Problems in High Energy Particle Theory* (Baltimore, 1978), p. 254.
[9]J. Lukierski and V. Rittenberg, Phys. Rev. D **18**, 385 (1978).
[10]N. Bourbaki, *Algèbre* (Hermann, Paris, 1962), chapter II, 3rd ed.
[11]N. Bourbaki, *Algèbre* (Hermann, Paris, 1971), chapter III, new ed.
[12]V. Rittenberg and D. Wyler, J. Math. Phys. **19**, 2193 (1978).
[13]N. Bourbaki, *Groupes et algèbres de Lie* (Hermann, Paris, 1960), chapter I.
[14]L.E. Ross, Trans. Am. Math. Soc. **120**, 17 (1965).
[15]L. Corwin, Y. Ne'eman, and S. Sternberg, Rev. Mod. Phys. **47**, 573 (1975).
[16]R.J. Blattner, Trans. Am. Math. Soc. **144**, 457 (1969).
[17]N. Bourbaki, *Algèbre* (Hermann, Paris, 1952), chapters VI–VII.
[18]N. Bourbaki, *Algèbre* (Hermann, Paris, 1959), chapter IX.
[19]K.R. Parthasarathy, *Multipliers on locally compact groups*, Lecture Notes in Mathematics 93 (Springer, Berlin, 1969).

Relations among generalized Korteweg–deVries systems[a)]

W. Symes

Mathematics Research Center, University of Wisconsin-Madison, Madison, Wisconsin
(Received 15 May 1978)

This report presents certain relations among the completely integrable Hamiltonian systems introduced by Gel'fand and Dikii. These relations generalize a formula of A. Lenard linking the higher-order Korteweg–deVries equations, of which the Gel'fand–Dikii systems are a generalization. The general form of the relations, which connect the various isospectral deformations of linear differential operators, is described, and two examples are given explicitly.

INTRODUCTION

In this short note we prove the existence of certain relations amongst the infinite-dimensional Hamiltonian systems constructed by Gel'fand and Dikii.[1] These relations generalize a formula of Lenard linking the higher-order Korteweg–deVries equations, of which the systems in Ref. 1 are a generalization. We refer the reader to Refs. 2 and 3 for a description of Lenard's result, which is also reproduced as an example at the end of this paper, and for an explanation of its importance in the theory of the Korteweg–deVries equation and its higher-order relatives. We refer the reader also to the papers of Lax,[4] Gel'fand and Dikii[5] and McKean and van Moerbeke[6] for further comments on, and uses of, Lenard's formula.

In the preceding references, Lenard's formula is derived by various *ad hoc* arguments; for instance, it is a consequence of a differential equation satisfied by *squares* of eigenfunctions of a Schrödinger operator.[4,6] In contrast, we show that Lenard's result may be regarded as a straightforward consequence of the law of exponents for (fractional) powers of ordinary differential operator. In this form, the Lenard formula generalizes immediately to the systems discussed in Ref. 1.

Mark Adler has derived the generalized Lenard relations of this paper in somewhat different ways, first for hierachies of systems related to the Boussinesq equation and to a certain fourth-order eigenvalue problem,[2] then in general for the Gel'fand–Dikii systems.[7]

1. FORMAL ISOSPECTRAL DEFORMATIONS

The central objects of this paper are the systems of partial differential equations discussed by Gel'fand and Dikii in Ref. 1. These systems have many remarkable properties. For instance, they are infinite-dimensional Hamiltonian systems, having infinitely many constants of motion in involution, which moreover are integrals of local polynomial densities. We refer the reader to Ref. 1 for details; see also Refs. 7 and 8.

These systems are also *formal (infinitesimal) isospectral deformations* of (formal) linear ordinary differential operators. A formal linear ordinary differential operator, or *linear differential expression*, is a polynomial in d/dx [or in $D = -\sqrt{-1}(d/dx)$, which will be more convenient], with coefficients which are functions of x, supposed infinitely differentiable on some open interval, finite or not, on the real line. (We use the term *differential expressions* to emphasize that no appropriate function space domain is assumed, so that we cannot speak properly of operators). An *infinitesimal deformation* of such expressions L is a correspondence

$$L \to L + \dot{L}dt,$$

where $\dot{L}$ is another differential expression, whose coefficients depend on those of L in some fashion. An infinitesimal deformation may thus be thought of as a vector field on the space of differential expressions. We restrict ourselves to deformations satisfying

$$\text{order } \dot{L} \leqslant \text{order } L.$$

In this case, the relation between the coefficients of $\dot{L}$ and those of L may be viewed as a system of differential equations,

$$\dot{L} = \frac{\partial L}{\partial t},$$

where coefficients of corresponding powers of $D = -\sqrt{-1}(d/dx)$ are equated, and coefficients of L are regarded as functions of another variable t. A solution of such a system is thus a one-parameter family of differential expressions, depending on the parameter t.

Let L be a linear differential expression of order n, in normal form:

$$L = D^n + \sum_{k=0}^{n-2} q_k D^k, \quad D = -\sqrt{-1}d/dx. \tag{1}$$

A *formal isospectral deformation* of L is a specification

$$\dot{L} = \sum_{k=0}^{n-2} \dot{q}_k D^k = [P,L] = P \circ L - L \circ P, \tag{2}$$

where P is a linear differential expression whose coefficients depend polynomially on $q_0,\cdots,q_{n-2}$ and their derivatives, having the property that the commutator appearing on the right-hand side of (2) is of order $n-2$ or less. Here the small circle $\circ$ is the symbol for composition of differential expressions according to Leibnitz' rule. Equation (2) may be regarded as a collection of $n-1$ partial differential equations for the coefficients q_k, where the dot is interpreted to mean

a)Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and the National science foundation under Grant No. MCS75-17385 A01.

differentiation with respect to a ("time") parameter t, as explained above.

Since the right-hand side of (2) is a commutator, the infinitesimal deformation defined by (2) is some sort of infinitesimal similarity transformation on the expression L. In fact, Peter Lax has shown that if L is provided with a suitable function space domain, becoming an operator and attaining a spectrum, then a solution of such a system of partial differential equations, determined by (2), represents a deformation of L preserving its spectrum—that is, an isospectral one-parameter family of operators.[9]

For example, let $n = 2$. Choose (with $q \equiv q_0$)

$$L = D^2 + q, \quad p = D^3 = \tfrac{3}{2}qD + \tfrac{3}{4}Dq.$$

Then Eq. (2) is equivalent to the Korteweg–deVries equation for the coefficient q,

$$-i\dot{q} = \tfrac{1}{4}D^3 q + \tfrac{3}{4}qDq.$$

As is well known, suitable solutions $q(x,t)$ of the Korteweg–deVries equation determine one parameter families $L(t) = -d^2/dx^2 + q(\cdot,t)$ of Schrödinger operators having the same spectrum, under a wide variety of boundary conditions.

Of course, the existence of a suitable expression P has nothing whatever to do with the selection of a function-space domain for L. In the Russian literature, a pair consisting of a linear differential expression L of order n, in normal form, and another linear differential expression P so that order $[P,L] \leqslant n-2$, is called a *Lax pair*. We shall adopt this terminology. A Lax pair thus gives rise to a deformation of the type depicted in Eq. (2), i.e., a certain system of partial differential equations. In view of the discovery of Lax regarding the Korteweg–deVries equation, we shall call these systems generalized Korteweg–deVries systems.

Gel'fand and Dikii[1] give a construction of all Lax pairs, based on the construction of fractional powers of the expression L. The original argument of Lax[9] recieves an elegant algebraic setting in their work. The price of elegance, naturally, is the introduction of more powerful machinery.

The linear ordinary differential expressions appearing in Ref. 1, like the Schrödinger operator, having leading coefficient $\equiv 1$, and hence are elliptic. The proper setting for a discussion of complex (or fractional) powers of elliptic operators is the calculus of pseudodifferential operators. (See Ref. 10 for a very readable general exposition of this subject.) As is implicitly recognized in the work of Seeley,[11] and explicitly utilized in Ref. 1, the construction of complex powers of elliptic operators has a formal counterpart on the level of the algebra of symbols. The algebra of symbols is an extension of the algebra of differential expressions, large enough to include inverses and powers of elliptic ("invertible") expressions. We give a very brief review of this subject in Secs. 2 and 3. All of the statements in Secs. 2 and 3 are formal counterparts of results of Seeley,[11] and proofs are also sketched in Ref. 1. We therefore omit the proofs. A reward for the introduction of all this machinery is the very simple construction of Lax pairs, due to Gel'fand and Dikii, which we present in Sec. 4. A further reward is a straightforward derivation of a relation amongst the various Lax pairs, which in the case of the Schrödinger operator ($n = 2$) boils down to Lenard's formula. We detail this result in Secs. 5–7.

The symbol algebra also provides the key to the rather subtle recipe for the Hamiltonian nature of the generalized Korteweg–deVries systems and the involution of the various constants of motion. This collection of results, presented somewhat computationally in Ref. 1, is explained in Refs. 7 and 8.

2. SYMBOL ALGEBRA

By intentional abuse of notation, denote also by L the *symbol of L*,

$$L[q,\xi] = \xi^n + \sum_{k=0}^{n-2} q_k \xi^k, \tag{3}$$

where we have emphasized the dependence of the symbol L on the coefficient vector $q = (q_0,\ldots,q_{n-2})$; the q's are complex valued functions on some open interval U.

More generally, a symbol is a formal sum

$$A(x,\xi) = \sum_{l=0}^{\infty} A_l(x,\xi), \tag{4}$$

where A_l is smooth and complex valued in $U \times \mathbb{R} \backslash \{0\}$, and homogeneous of degree d_l in ξ, with

$$d_0 > d_1 > \cdots > d_l > \cdots \to -\infty$$

a sequence of real numbers tending to $-\infty$.

Symbols are considered the same if they agree, term by term, with each other.

The symbols form a module over $C^\infty(U)$, with the module operations defined pointwise. They form an algebra over $\mathbb{C}$, with the multiplication operation $\circ$ defined by

$$(A,B) \to A \circ B = \sum_{\nu \geqslant 0} \frac{1}{\nu!} \partial^\nu A D^\nu B, \tag{5}$$

where $\partial = \partial/\partial\xi$ and $D = -i\partial/\partial x$ are applied term by term. Note that the sum on the right of (5) is not in the canonical form (4); however, only finitely many products of the same homogeneous degree appear in (5), and rearrangement into the form (4) is easy.

Note that, under the replacement $D \to \xi$, linear differential expressions pass over into polynomial (in ξ) symbols, and that in the case of such polynomial symbols Eq. (5) corresponds exactly to the rule for composition of differential expressions. Indeed, this observation is the motivation for the definition (5) of the product of symbols, and shows that the algebra of differential expressions may be viewed as a subalgebra of the algebra of symbols.

The class of symbols all of whose homogeneous pieces have integral degree, and the class of symbols which can be written

$$A = \sum_{l \geqslant 0} A_l[q,\xi]$$

with $A_l[q,\xi] = a_l \xi^{d_l}\sigma_l(\xi)$, where

$$\sigma_l(\xi) \equiv \begin{Bmatrix} 1, & \xi > 0, \\ \pm 1 & \xi < 0, \end{Bmatrix}$$

and $a_l[q]$ a polynomial in q and its x derivatives (that is, a polynomial in $q_0,\ldots,q_{n-2}$ and their x derivatives) both form subalgebras. The latter subalgebra will be called the class of q *symbols.*

The order $\operatorname{ord} A$ of symbol A is the homogeneous degree of the term of highest homogeneous degree appearing in an expansion (4)—that is, $\operatorname{ord} A = d_0$ in the notation of (4).

Note that

$$\operatorname{ord}(A \circ B) \leqslant \operatorname{ord} A + \operatorname{ord} B$$

and

$$\operatorname{ord}[A,B] \leqslant \operatorname{ord} A + \operatorname{ord} B - 1,$$

where

$$[A,B] = A \circ B - B \circ A.$$

Note that differential expressions such as (1), and more generally differential expressions whose coefficients depend polynomially on q and its x derivatives, correspond under the replacement $D \to \xi$ to polynomial (in ξ) q symbols. Therefore, Lax equations (2) are in 1–1 correspondence with polynomial q symbols P for which $\operatorname{ord}[L,P] \leqslant n-2$.

3. COMPLEX POWERS

Define the *resolvent symbol* $R(\lambda)$ for L by the equation

$$R(\lambda) \circ (L - \lambda) = 1.$$

According to the definition of the product (5), we can rewrite this equation in the form

$$(\xi^n - \lambda) R(\lambda) = 1 - R(\lambda) \circ \left(\sum_{k=0}^{n-2} q_k \xi^k \right).$$

From this formula it is easy to see that $R(\lambda)$ admits an expansion in the form

$$R(\lambda) = \sum_{l=n}^{\infty} R_l(\lambda,\xi), \tag{6}$$

where R_l is homogeneous of degree $-l$ in ξ, $\lambda^{1/n}$, and is defined recursively by the formulas

$$R_n(\lambda) = (\xi^n - \lambda)^{-1}, \quad R_{n+1}(\lambda) \equiv 0,$$

$$R_l(\lambda) = -\sum_{j=0}^{l-n-2} \sum_{k=\max(0,2n+j-1)}^{n-2} \frac{\xi^k (\xi^n - \lambda)^{-1}}{(-2n - j + l + k)!} \times \partial^{-2n-j+l+k} R_{n+j} D^{-2n-j+l+k} q_k.$$

In particular,

$$R_l(\lambda) = \sum_{m=2}^{l-n} B_{lm}(\xi^n + \lambda)^{-m}, \quad l \geqslant n+2, \tag{7}$$

where B_{lm} is a homogeneous polynomial in ξ of degree $nm - l$ whose coefficients are polynomials in q and its x derivatives.

Observe that, by virtue of (7), the expansion (6) can be rearranged for large $|\xi|$ into the form (4), showing that $R(\lambda)$ is a q symbol. $R(\lambda)$ should be regarded as a local version of the resolvent of L. Indeed, R obeys the *resolvent equation*

$$R(\lambda) - R(\mu) = (\lambda - \mu) R(\lambda) \circ R(\mu).$$

In further analogy to the usual theory of the resolvent, we use $R(\lambda)$ to define complex powers of the symbol L.

For the remainder of this paper, define

$$\lambda^s = \exp(s \log \lambda), \quad \lambda, s \in \mathbb{C},$$

where the principal branch of log is selected, with the branch cut down the negative imaginary axis.

Define the symbol L^s, for complex s, by

$$L^s = \frac{1}{2\pi i} \oint d\lambda \lambda^s R(\lambda). \tag{8}$$

The contour goes up the ray $\operatorname{Re} \lambda = \frac{1}{2}$, around the semicircle $|\lambda| = \frac{1}{2}$ counterclockwise, and back down the ray $\operatorname{Re} \lambda = -\frac{1}{2}$.

The integral (8) is evaluated by inserting the expansion (6) and (7) for $R(\lambda)$ in (8) and integrating term by term, using the residue theorem. Each integral converges and admits evaluation by residues for $\operatorname{Re} s < -1$ and $|\xi| > 1$. We obtain

$$L^s = \sum_{p=0}^{\infty} A_P(s), \tag{9}$$

where $\operatorname{ord} A_p = n \operatorname{Re} s - p$,

$$A_p(s) \equiv (\xi^n)^s (\xi^n)^{1-m} \times \sum_{m=2}^{p} B_{p+n,m} (-1)^{m-1} \frac{1}{(m-1)!} \sum_{j=0}^{m-2} (s-j). \tag{10}$$

This result is then extended by homogeneity to $|\xi| > 0$.

Thus L^s is a q symbol, of order $n \operatorname{Re} s$, for $\operatorname{Re} s < -1$. Formulas (9) and (10) allow continuation of L^s as an entire function of s. Mimicking arguments in Ref. 6, one easily shows that the resolvent formula implies the semigroup properties

$$L^0 = 1, \quad L^1 = L, \quad L^s L^t = L^{s+t}.$$

In particular, for any complex s, t,

$$[L^s, L^t] = 0.$$

4. LAX PAIRS

Let m be any positive integer. Then $L^{m/n}$ is a q symbol of order m, whose homogeneous pieces have integral degree. Set

$$L^{m/n} = P_m + N_m,$$

where

$$P_m = \sum_{p=0}^{m} A_p\left(\frac{m}{n}\right),$$

$$N_m = \sum_{p=m+1}^{\infty} A_p\left(\frac{m}{n}\right).$$

Thus P_m is a polynomial q symbol, and ord$N_m \leqslant -1$. If m is not divisible by n, it is easy to check that $A_p(m/n) \not\equiv 0$ for $p \geqslant m+1$.

Because $[L, L^{m+n}] = 0$,

$$[P_m, L] = [L, N_m].$$

Since the left-hand side of this equation is a polynomial q symbol, so is the right. On the other hand, since the right-hand side has order $\leqslant n-2$, so does the left-hand side. Thus

$$\dot{L} = [P_m, L]$$

is a Lax equation for each positive integer m. This is Gel'fand and Dikii's construction of Lax pairs.

A simple degree-counting argument shows that P_m is essentially the unique polynomial q symbol of order m, whose commutator with L is of order $n-2$.

5. GENERALIZED LENARD RELATIONS

Denote by X_m the polynomial q symbol

$$X_m \equiv [P_m, L] = [L, N_m].$$

In view of Eq. (2), it is natural to call X_m the mth Lax vector field for L. The object of this section is to relate X_m to X_{m+n}. Since $L^{m/n+1} = L \circ L^{m/n}$, we have

$$P_{m+n} = (L \circ L^{m/n})_+$$

$$= L \circ P_m + (L \circ N_m)_+,$$

where the subscript " + " signifies the sum of parts of nonnegative homogeneous degree, i.e., the polynomial part. Thus

$$X_{m+n} = L \circ X_m - [L, (L \circ N_m)_+]. \tag{11}$$

The equation

$$X_{m,l} = \sum_{k=l+2}^{n} \sum_{p=m+1}^{m+k-l-1} \left[A_p\left(\frac{m}{n}\right), q_k \xi^k\right]_l \tag{12}$$

(where the subscript l denotes the part of homogeneous degree l) shows that X_m depends only on $A_p(m/n)$, $m+1 \leqslant p \leqslant m+n-1$. We shall show that Eq. (11) can also be rewritten to express X_{m+n} in terms of $A_p(m/n)$, $m+1 \leqslant p \leqslant m+n-1$. This will be our generalized Lenard relation.

In fact,

$$(L \circ N_m)_+ = \xi^n A_{m+n}(m/n) + \cdots,$$

where the terms denoted by dots involve only $A_p(m/n)$, $m+1 \leqslant p \leqslant m+n-1$. On the other hand, according to the result of Sec. 4,

$$0 = [L, N_m]_{-1}$$

$$= \sum_{p=1}^{n} \begin{bmatrix} n \\ n+1-p \end{bmatrix} D^{n+1-p} a_{m+p}\left(\frac{m}{n}\right)$$

$$+ \sum_{k=0}^{n-2} \sum_{p=1}^{k} \begin{bmatrix} k \\ k+1-p \end{bmatrix} \left[q_k D^{k+1-p} a_{m+p}\left(\frac{m}{n}\right) - a_{m+p}\left(\frac{m}{n}\right)(-D)^{k+1-p} q_k \right],$$

where we have written $A_p(m/n) = a_p(m/n)\xi^{m-p}$. This can be rewritten

$$-nDa_{m+n}\left(\frac{m}{n}\right) = \sum_{p=1}^{n-1} \begin{bmatrix} n \\ n+1-p \end{bmatrix} D^{n+1-p} a_{m+p}\left(\frac{m}{n}\right) + \sum_{k=0}^{n-2} \sum_{p=1}^{k} \begin{bmatrix} k \\ k+1-p \end{bmatrix} \left[q_k D^{k+1-p} a_{m+p}\left(\frac{m}{n}\right) - a_{m+p}\left(\frac{m}{n}\right)(-D)^{k+1-p} q_k \right]. \tag{13}$$

The right-hand side of (13) is an exact derivative, so we can express $a_{m+p}(m/n)$ in terms of $a_{m+p}(m/p)$, $1 \leqslant p \leqslant n-1$. On the other hand, we don't even need to do that, since only $DA_{m+n}(m/n)$ appears in $[L, (L \circ N_m)_+]$. In any case, we have expressed X_{m+n} in terms of $A_{m+p}(m/n)$, $1 \leqslant p \leqslant n-1$, as we desired.

6. THE HAMILTONIAN FORMALISM

In order to write the relations derived in the last section in the form in which the Lenard relations are usually presented, we introduce the Hamiltonian formalism of Ref. 1.

Suppose that F is a polynomial in q and its derivatives. Define the *formal variational derivative* of F by

$$\frac{\delta F}{\delta q_k} = \sum_{j=0}^{\infty} (-D)^j \frac{\delta F}{\partial (D^j a_k)},$$

$$\frac{\delta F}{\delta q} = \left(\frac{\delta F}{\delta q_0}, \ldots, \frac{\delta F}{\delta q_{n-2}}\right)^t,$$

where the superscript t denotes transpose (to a column matrix). In Ref. 1 it is proven [Eq. (21)] that

$$a_p\left(\frac{m}{n}\right) = \frac{n}{m+n} \sum_{\nu=0}^{p-m-1} \binom{p-m-1}{\nu} (-D)^\nu \times \frac{\delta}{\delta q_{p-m-\nu-1}} a_{m+n+1}\left(\frac{m+n}{n}\right) \tag{14}$$

for $p = m+1, \ldots, m+n-1$.

Using (12) and (14), it is easy to express the Lax vector field X^m in the form

$$(X_{m,0}, \ldots, X_{m,n-2})^t = \frac{n}{m+n} J \frac{\delta}{\delta q} H_m, \tag{15}$$

where $H_m \equiv a_{m+n+1}(m+n/n)$, and J is a certain matrix of differential expressions, whose coefficients depend polynomially on q and its derivatives.

In Ref. 1 it is shown that J defines a symplectic structure on the space of coefficient vectors q, in a certain local sense, and that the Lax vector field is a Hamiltonian vector field, by virtue of (15). (See also Refs. 7 and 8.)

Now Eqs. (11)–(15) allow us to write the relation derived in Sec. 5 in the succinct form

$$\frac{n}{m+2n} J \frac{\delta}{\delta q} H_{m+n} = \frac{n}{n+m} K \frac{\delta}{\delta q} H_m, \qquad (16)$$

where K is another $(n-1)\times(n-1)$ matrix of differential expressions whose coefficients depend polynomially on q and its derivatives. Equation (16) is, for $n=2$, the form in which the Lenard relations are usually stated. The general form of J is very complicated, and that of K even more so. Rather than exhibit these general forms, we compute in Sec. 7 J and K explicitly for $n=2,3$.

7. EXAMPLES

(1) $n=2$: $L=\xi^2+q_0$, and

$$X_0^m = 2Da_{m+1}(\tfrac{1}{2}m) = \frac{4}{m+2} D \frac{\delta}{\delta q_0} a_{m+3}(\tfrac{1}{2}m+1),$$

so in this case J is the 1×1 matrix $2D$.

Equation (11) yields

$$X_0^{m+2} = D^2 X_0^m + qX_0^m - 2D^3 a_{m+1}(\tfrac{1}{2}m) - D^2 a_{m+2}(\tfrac{1}{2}m) + a_{m+1}(\tfrac{1}{2}m)Dq.$$

Equation (13) yields

$$Da_{m+2} = -1/2D^2 a_{m+1}.$$

Hence

$$X_0^{m+2} = (\tfrac{1}{2}D^3 + 2qD + Dq)a_{m+1}(\tfrac{1}{2}m).$$

So K is the 1×1 matrix

$$K = \tfrac{1}{2}D^3 + qD + Dq.$$

That is the expression discovered by Lenard.

(2) $n=3$. Set $H_m = [3/(m+3)]a_{m+4}(\tfrac{1}{3}m+1)$.

Then

$$a_{m+1} = \frac{\delta H_m}{\delta q_0},$$

$$a_{m+2} = \frac{\delta H_m}{\delta q_1} - D\frac{\delta H_m}{\delta q_0},$$

according to formula (14).

The Lax vector field is

$$X^m = 3\xi Da_{m+1}(\tfrac{1}{3}m) + 3[D^2 a_{m+1}(\tfrac{1}{3}m) + Da_2(\tfrac{1}{3}m)].$$

Thus

$$J = \begin{pmatrix} 0 & 3D \\ 3D & 0 \end{pmatrix}.$$

From (13) we obtain

$$Da_{m+3}(\tfrac{1}{3}m) = -\tfrac{1}{3}[3D^2 a_{m+2}(\tfrac{1}{3}m) + D^3 a_{m+1}(\tfrac{1}{3}m) + (D\circ q_0)a_{m+1}(\tfrac{1}{3}m)]$$

and, after tedious computation,

$$K = \begin{pmatrix} K_{00} & K_{01} \\ K_{10} & K_{11} \end{pmatrix},$$

where

$$K_{00} = -\tfrac{2}{3}D^5 + \tfrac{1}{3}D^3\circ q_1 - \tfrac{2}{3}q_1 D^3 - \tfrac{1}{3}q_1(D\circ q_1) + D\circ(Dq_0) + Dq_0 D,$$

$$K_{01} = D^4 - 2q_1 D^2 + 3q_0 D - Dq_0,$$

$$K_{10} = -D^4 + 2D^2\circ q_1 - 2Dq_0 + 3q_0 D,$$

$$K_{11} = 2D^3 + q_1 D + D\circ q_1.$$

ACKNOWLEDGMENTS

I would like to thank Mark Adler, Charles Conley, and Neil Fenichel for a number of illuminating conversations on this subject.

[1] I.M. Gel'fand and L.A. Dikii, "Fractional Powers of Operators and Hamiltonian Systems," Funct. Anal. Appl. **10**, 13–29 (1976).
[2] M. Adler, "Some Algebraic Relations Common to a Set of Integrable Partial and Ordinary Differential Equations," Tech. Summary Rep. No. 1801 (Math. Res. Cent., Univ. Wisconsin–Madison, 1977).
[3] C.S. Gardner, J.M. Greene, M.D. Kruskal, and R.M. Miura, "Korteweg–deVries Equation and Generalizations, VI: Methods for Exact Solution," Commun. Pure. Appl. Math. **27**, 97–133 (1974).
[4] P.D. Lax, "Periodic Solutions of the *KdV* Equation," Commun. Pure Appl. Math. **28**, 141–188 (1975).
[5] I.M. Gel'fand and L.A. Dikii, "Asymptotics of the Resolvent of the Sturm–Liouville Equation and the Algebra of the Korteweg–deVries Equation," Usp. Mat. Nauk **30**, 67–100 (1975).
[6] H. Mckean and P. van Moerbeke, "The Spectrum of Hill's Equation," Invent. Math. **30**, 217–274 (1975).
[7] M. Adler, "On the Relationship of Some Results of Gel'fand–Dikii and P. Moerbeke, and a Natural Trace Functional for Formal Asymptotic Pseudodifferential Operators," Invent. Math. (to appear).
[8] W. Symes, "Systems in Involution of Toda Type" (to appear).
[9] P.D. Lax, "Integrals of Nonlinear Equations of Evolution and Solitary Waves," Commun. Pure Appl. Math. **21**, 467–490 (1968).
[10] M. Taylor, *Notes on Pseudodifferential Operators*, Springer Lecture Notes in Mathmatics No. 416 (Springer, New York, 1974).
[11] R.T. Seeley, "Complex Powers of an Elliptic Operator," Proc. Symp. Pure Math. **10**, 288–307 (1966).

Null electromagnetic fields in the generalized Einstein–Maxwell field theory

Gregory Walter Horndeski

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada
(Received 21 September 1978)

In this paper solutions to the source-free generalized Einstein–Maxwell field equations with a null electromagnetic field are investigated. It is argued that the principal null congruence of the null electromagnetic field need not be geodesic, shear-free or a repeated principal null congruence of the gravitational field. However, if the principal null congruence of the null electromagnetic field is geodesic and shear-free, then it is shown that it must be hypersurface orthogonal, expansion-free, and a repeated principal null congruence of the gravitational field. The local form of all such solutions of Petrov type III or N is presented.

1. INTRODUCTION

The source-free Einstein-Maxwell field equations are[1]

$$G^{ab} = 8\pi T^{ab}, \quad F^{hj}{}_{;j} = 0, \quad {}^*F^{hj}{}_{;j} = 0, \tag{1.1}$$

where

$$T^{ab} \equiv \frac{1}{4\pi}(F^{ar}F^{b}{}_{r} - \tfrac{1}{4}g^{ab}F^{rs}F_{rs}). \tag{1.2}$$

A null vector field l^a is said to define a *principal null direction* (abbreviated, pnd) *of the electromagnetic field* if it is an eigenvector of F_{ab}. There always exists[2] at least one pnd for F_{ab} and at most two linearly independent pnd's (provided $F_{ab}\neq 0$). If F_{ab} is null; i.e.,

$$F_{ab}F^{ab} = 0 \quad \text{and} \quad F_{ab}{}^*F^{ab} = 0,$$

then the pnd's of F_{ab} coalesce and we may write

$$F_{ab} = l_{[a}p_{b]} \quad \text{and} \quad {}^*F_{ab} = l_{[a}q_{b]}, \tag{1.3}$$

where p_b and q_b are spacelike vectors which are perpendicular to l_a, and to each other.

For a definition of the pnd's of the gravitational field see, e.g., Ref. 2 (pp. 319–22).

A congruence of null curves in spacetime which has its tangent vector field parallel to a pnd of the electromagnetic field (gravitational field, resp.) is called a *principal null congruence* (abbreviated, pnc) *of the electromagnetic field* (gravitational field, resp.).

We have the following theorems concerning solutions of Eq. (1.1).

Theorem 1 (*Mariot*[3]*-Robinson*[4]): If the source-free Einstein–Maxwell field equations are satisfied and F_{ab} is null, then the pnc of the electromagnetic field is geodesic and shear-free.

Theorem 2 (*Kundt and Trumper*[5]): Suppose that the source-free Einstein–Maxwell field equations are satisfied and F_{ab} admits a geodesic and shear-free pnc with tangent vector field l^a. Then l^a defines a repeated pnd of the gravitational field and hence the gravitational field is algebraically special.

Theorems 1 and 2 imply

Thereom 3: If the source-free Einstein-Maxwell field equations are satisfied and the electromagnetic field is null, with pnd determined by l^a, then the gravitational field is algebraically special and l^a defines a repeated pnd for the gravitational field.

The primary purpose of this paper is to explore the degree to which Theorems 1 and 3 can be extended to the generalized Einstein–Maxwell field theory.[6,7] The source-free field equations of this theory are given by

$$G^{ab} = 8\pi(T^{ab} - kA^{ab}), \quad F^{hj}{}_{;j} - \frac{k}{2}F^*{}_{ab;j}{}^*R^{*abhj} = 0, \quad {}^*F^{hj}{}_{;j} = 0, \tag{1.4}$$

where

$$A^{ab} \equiv \frac{1}{8\pi}(F_{rh}{}^*R^{*rbs} + {}^*F^{ar,s}{}^*F^{b}{}_{s;r}) \tag{1.5}$$

and k is a coupling constant with units of (length)[2]. When $k=0$ the equations presented in (1.4) reduce to the source-free Einstein–Maxwell field equations, and thus *we shall assume* that $k\neq 0$.

The field equations of the generalized Einstein–Maxwell field theory are *uniquely* characterized by the following four conditions:

(i) The vacuum field equations are derivable from a Lagrangian of the form $L = L(g_{ij};g_{ij,i_1};\ldots;g_{ij,i_1\cdots i_\alpha};\psi_a;\psi_{a,i_1};\ldots;\psi_{a,i_1\cdots i_\beta})$ $(\alpha\geqslant 2,\beta\geqslant 1)$, and are at most of second order in the derivatives of g_{ij} and ψ_a where $F_{ab} := \psi_{b,a} - \psi_{a,b}$.

(ii) In the presence of sources the field equations are compatible with the notion of charge conservation.

(iii) The field equations governing the vector potential ψ_a reduce to Maxwell's equations in a flat space.

(iv) In the absence of electromagnetic fields the equations governing g_{ij} are Einstein's equations.

For a more detailed discussion of the generalized Einstein–Maxwell field equations see Refs. 8 and 9.

My major tool for studying the field equations presented in Eq. (1.4) will be the Newman–Penrose formalism of Ref. 1. Using this approach I show that if the source-free generalized Einstein–Maxwell field equations are satisfied by a null electromagnetic field, then (in general) the corre-

 0022-2488/79/040726-07$01.00

sponding pnc will probably *not* be geodesic or shear-free. However if it is geodesic and shear-free then it must be hypersurface orthogonal, expansion-free, and a repeated pnc of the gravitational field. After establishing this result I devote the remainder of the paper to a study of various types of solutions to the source-free generalized Einstein–Maxwell field equations which are such that F_{ab} is null and the corresponding pnc is geodesic, shear-free, hypersurface orthogonal, and expansion-free.

2. PROPERTIES OF NULL ELECTROMAGNETIC FIELDS

Throughout this paper it is assumed that the metric and electromagnetic field satisfy the source-free generalized Einstein–Maxwell field equations (1.4). In addition F_{ab} is required to be null and we assume that l^a defines the corresponding pnd. Under these assumptions F_{ab} and $^*F_{ab}$ can be expressed as in Eq. (1.3). We now adjoin to l^a three other null vector fields n^a, m^a, and $\bar{m}^a$ which are such that the set $\{l^a, n^a, m^a, \bar{m}^a\}$ defines a null tetrad. In this tetrad n^a is real while m^a is complex with $\bar{m}^a$ being its complex conjugate. The nonzero inner products of the tetrad vectors are

$$l^a n_a = 1 \text{ and } m^a \bar{m}_a = -1.$$

It is assumed that the null tetrad is oriented so that

$$\epsilon_{abcd} = 4!\, i l_{[a} n_b m_c \bar{m}_{d]}.$$

There exists a complex valued function C which is such that the vectors p_b and q_b appearing in Eq. (1.3) can be written as[10]

$$p_b = i(\bar{C}\bar{m}_b - C m_b), \quad q_b = C m_b + \bar{C}\bar{m}_b. \tag{2.1}$$

In the notation of Ref. 1, $C = 2i\Phi_2$.

As usual we define

$$^+F_{ab} = \tfrac{1}{2}(F_{ab} - i^*F_{ab}).$$

Using Eqs. (1.3) and (2.1) we find

$$^+F_{ab} = \frac{iC}{2}(m_a l_b - m_b l_a). \tag{2.2}$$

In terms of $^+F_{ab}$ the "electromagnetic part" of the generalized Einstein–Maxwell field equations (1.4) becomes

$$2\,^+F^{hj}{}_{;j} - (k/2) F_{ab\,;j}\,^*R^{*abhj} = 0. \tag{2.3}$$

When (2.3) is multiplied by $l_h, m_h, \bar{m}_h$, and n_h, noting Eq. (2.2), we obtain[11]

$$C\kappa = \frac{ik}{2} l_h F_{ab;j}\,^*R^{*abhj}, \tag{2.4}$$

$$C\sigma = \frac{ik}{2} m_h F_{ab;j}\,^*R^{*abhj}, \tag{2.5}$$

$$DC + C(2\epsilon - \rho) = \frac{ik}{2}\bar{m}_h F_{ab;j}\,^*R^{*abhj}, \tag{2.6}$$

and

$$\delta C + C(2\beta - \tau) = \frac{ik}{2} n_h F_{ab;j}\,^*R^{*abhj}, \tag{2.7}$$

where

$$DC \equiv C_{;a} l^a \quad \text{and} \quad \delta C \equiv C_{;a} m^a.$$

Now recall that the pnc associated with l^a is geodesic and shear-free if and only if $\kappa = \sigma = 0$. Thus in view of Eqs. (2.4) and (2.5) it seems quite unlikely that this will be the case (in general) for the pnc associated with a null electromagnetic field which satisfies Eq. (1.4). Consequently the Mariot–Robinson theorem probably cannot be extended to the source-free generalized Einstein–Maxwell field equations.

In passing it should be noted the Eq. (2.4) implies that

$$C\kappa + \bar{C}\bar{\kappa} = 0. \tag{2.8}$$

This fact can be used along with Eq. (2.1) to show that

$$l_{a;b} l^b = (\epsilon + j\bar{\epsilon}) l_a + \frac{i\kappa}{\bar{C}} p_a.$$

Using Eqs. (1.3) and (2.1) we find that the "gravitational part" of Eq. (1.4) can be written as follows:

$$\begin{aligned} G_{ab} = {} & -C\bar{C} l_a l_b + \frac{k}{2} C\bar{C}\,^*R^*{}_{arbs} l^r l^s \\ & - \frac{k}{4}\{ l_{a;r} q^r l_{b;s} q^s + 2 l_{(a} l_{b);r} q^{r;s} q_s - 2 q_{(a} l_{b);r} l^{r;s} q_s \\ & - 2 l_{(a}{}^{;r} q_{b);s} l_r q^s + l_a l_b q_{r;s} q^{s;r} - 2 q_{(a} l_{b)} l_{r;s} q^{s;r} \\ & - 2 l_{(a} q_{b);r} q^{r;s} l_s \\ & + q_a q_b l_{r;s} l^{s;r} + 2 q_{(a} q_{b);r} l^r{}_{;s} l^s + q_{a;r} l^r q_{b;s} l^s \}. \end{aligned} \tag{2.9}$$

Combining this equation with (2.8) shows us that

$$\Phi_{00} \equiv -\tfrac{1}{2} R_{ab} l^a l^b = 0, \tag{2.10}$$

while in general the other tetrad components of the Ricci tensor are nonzero; e.g.,[12]

$$\Phi_{01} \equiv -\tfrac{1}{2} R_{ab} l^a m^b = \frac{k\kappa C}{4}(C\sigma + \bar{C}\bar{\rho}). \tag{2.11}$$

This situation should be contrasted with the corresponding situation for null electromagnetic fields satisfying the source-free Einstein–Maxwell field equations. In this case the only nonzero tetrad component of the Ricci tensor is $R_{ab} n^a n^b = -C\bar{C}$.

We shall now turn our attention to an investigation of null electromagnetic fields which satisfy Eq. (1.4) and are such that their pnc is geodesic and shear-free.

3. NULL ELECTROMAGNETIC FIELDS WITH GEODESIC AND SHEAR-FREE PNC'S

In this section it is assumed that the pnc of the null electromagnetic field is geodesic and shear-free. Consequently we may choose our null tetrad $\{l^a, n^a, m^a, \bar{m}^a\}$ so that $\kappa = \sigma = \epsilon = 0$. Thus Eqs. (2.5), (2.6), and (2.11) imply that

$$DC = C\rho \tag{3.1}$$

and

$$\Phi_{01} = 0. \tag{3.2}$$

Upon multiplying Eq. (2.9) by g^{ab} we discover that due to Eqs. (2.10) and (3.1)

$$\Lambda \equiv \frac{R}{24} = 0. \tag{3.3}$$

When (2.9) is multiplied by $m^a m^b$ we find, through the use of Eq. (3.1), that

$$\Phi_{02} \equiv -\tfrac{1}{2} R_{ab} m^a m^b = \frac{k}{4} \bar{C}^2 \bar{\rho}^2 - \frac{k}{4} C\bar{C} {}^*R^*{}_{arbs} m^a l^r m^b l^s. \tag{3.4}$$

Since $\kappa = \sigma = 0$ the Ricci identities[13] imply that l^a must define a pnd of the gravitational field and thus

$$\Psi_0 \equiv -C_{abcd} l^a m^b l^c m^d = 0. \tag{3.5}$$

However, owing to Eq. (3.3),

$$-C_{abcd} = {}^*R^*{}_{abcd} + \tfrac{1}{2}(g_{ac}R_{bd} - g_{ad}R_{bc} + g_{bd}R_{ac} - g_{bc}R_{ad}), \tag{3.6}$$

and hence Eq. (3.5) implies that

$${}^*R^*{}_{arbs} m^a l^r m^b l^s = 0. \tag{3.7}$$

Thus Eqs. (3.4) and (3.7) imply that

$$\Phi_{02} = \frac{k}{4} \bar{C}^2 \bar{\rho}^2. \tag{3.8}$$

I now want to examine whether or not our present assumptions on l^a imply that l^a is a repeated pnd of the gravitational field, as it would be if we were dealing with the source-free Einstein–Maxwell field equations. For this to be the case Ψ_1 must vanish where

$$\Psi_1 \equiv -C_{abcd} l^a n^b l^c m^d. \tag{3.9}$$

Let us suppose that $\Psi_1 \neq 0$. Due to Eqs. (2.10) and (3.2) the Bianchi identities[13] tell us that

$$D\ln\Psi_1 = 4\rho \quad \text{and} \quad \delta\ln\Psi_1 = 2(2\tau + \beta) + \frac{D\Phi_{02} - \bar{\rho}\Phi_{02}}{\Psi_1}, \tag{3.10}$$

while if ϕ is an arbitrary function, then

$$(\delta D - D\delta)\phi = (\bar{\alpha} + \beta - \bar{\pi})D\phi - \bar{\rho}\delta\phi. \tag{3.11}$$

Upon combining Eqs. (3.10) and (3.11) in conjunction with the Ricci identities

$$\delta\rho = \rho(\bar{\alpha} + \beta) + \tau(\rho - \bar{\rho}) - \Psi_1,$$
$$D\tau = \rho(\tau + \bar{\pi}) + \Psi_1,$$
$$D\beta = \beta\bar{\rho} + \Psi_1,$$

we obtain

$$-10\Psi_1 = D\left(\frac{D\Phi_{02} - \bar{\rho}\Phi_{02}}{\Psi_1}\right) - \bar{\rho}\left(\frac{D\Phi_{02} - \bar{\rho}\Phi_{02}}{\Psi_1}\right). \tag{3.12}$$

Using Eqs. (3.1), (3.8), and the Ricci identity

$$D\rho = \rho^2, \tag{3.13}$$

we discover that Eq. (3.12) reduces to

$$-10\Psi_1^2 = 3k\bar{C}^2\bar{\rho}^3(\bar{\rho} - \rho). \tag{3.14}$$

A straightforward calculation employing Eqs. (2.10), (3.1), (3.2), (3.6), and (3.7) shows that Eq.(2.4) can be rewritten as follows,[14]

$$C\rho\Psi_1 = \bar{C}\bar{\rho}\bar{\Psi}_1. \tag{3.15}$$

Upon acting on this equation with the operator D we easily find, through the use of Eqs. (3.1), (3.10), (3.13), and (3.15), that

$$C\rho\Psi_1(\bar{\rho} - \rho) = 0. \tag{3.16}$$

Evidently Eqs. (3.14) and (3.16) imply that $\Psi_1 = 0$, contrary to our original assumption that $\Psi_1 \neq 0$. Thus when $\kappa = \sigma = 0$ we must have $\Psi_1 = 0$. In this case the Bianchi identities tell us that [cf. Eq. (3.10)]

$$D\Phi_{02} = \bar{\rho}\Phi_{02}. \tag{3.17}$$

When Eqs. (3.8) and (3.17) are combined, noting Eqs. (3.1) and (3.13), we get

$$\tfrac{3}{4}k\bar{C}^2\bar{\rho}^3 = 0$$

and hence

$$\rho = 0.$$

This in turn implies that the pnc of the electromagnetic field must be hypersurface orthogonal and expansion-free (see Ref. 2, pp. 340–1).

Upon drawing the above results together we obtain

Theorem 4: When the source-free generalized Einstein–Maxwell field equations are satisfied by a null electromagnetic field whose corresponding pnc is geodesic and shear-free, then this congruence must also be hypersurface orthogonal, expansion-free, and a repeated pnc of the (necessarily) algebraically special gravitational field.

Theorem 4 represents an extension of Theorem 3 to the generalized Einstein–Maxwell field theory. The basic difference between these two theorems is that when proving Theorem 3 you do not have to assume that the pnc of the null electromagnetic field is geodesic and shear-free, since that follows from Theorem 1. However, in the case of the Einstein–Maxwell field theory there exist solutions with a null electromagnetic field which are such that the corresponding pnc has[15] $\rho \neq 0$. Due to Theorem 4 we see that this is not the case in the generalized Einstein–Maxwell field theory (provided the congruence in question is geodesic and shear-free). This is perhaps the most remarkable aspect of Theorem 4.

4. ALGEBRAICALLY SPECIAL GRAVITATIONAL AND ELECTROMAGNETIC FIELDS

The basic assumption in this section is that the pnc of the null electromagnetic field is geodesic and shear-free. Due to Theorem 4 this congruence must be hypersurface orthogonal, expansion-free, and a repeated pnc of the gravitational field. Thus the gravitational field is algebraically special and in addition we can select a null tetrad $\{l^a, n^a, m^a, \bar{m}^a\}$ so that $\kappa = \sigma = \rho = \epsilon = 0$. In this case Eq.(3.1) becomes

$$DC=0 \tag{4.1}$$

and it can be shown tht Eq. (2.9) may be rewritten as follows,

$$R_{ab}=\frac{k}{2}C\bar{C}{}^*R^*{}_{arbs}l^r l^s-Fl_a l_b, \tag{4.2}$$

where

$$F=C\bar{C}+k\,(\beta^2C^2+\bar{\beta}^2\bar{C}^2+2\alpha\bar{\alpha}C\bar{C}+\beta C\delta C+\bar{\beta}\bar{C}\bar{\delta}\bar{C}$$

$$+\alpha C\delta\bar{C}+\bar{\alpha}\bar{C}\bar{\delta}C)+\frac{k}{4}[(\delta C)^2+2(\bar{\delta}C)(\delta\bar{C})$$

$$+(\bar{\delta}\bar{C})^2+(C\tau+\bar{C}\bar{\tau})^2]. \tag{4.3}$$

We shall now demonstrate that R_{ab} must be proportional to $l_a l_b$.

Due to Eqs. (2.10), (3.2), (3.3), and (3.8) we have

$$\Phi_{00}=\Phi_{01}=\Phi_{02}=\Lambda=0. \tag{4.4}$$

Thus in order to prove that R_{ab} is of the required form all we need do is show that $\Phi_{11}=\Phi_{12}=0$, where

$$\Phi_{11}\equiv-\tfrac{1}{4}R_{ab}(l^an^b+m^a\bar{m}^b)$$

and

$$\Phi_{12}\equiv-\tfrac{1}{2}R_{ab}n^am^b.$$

Upon multiplying Eq. (4.2) by n^am^b we get

$$\Phi_{12}=\frac{-k}{4}C\bar{C}{}^*R^*{}_{arbs}n^al^rm^bl^s.$$

Since $\Psi_1=0$ we may employ Eqs. (3.6), (3.9), and (4.4) to show that ${}^*R^*{}_{arbs}n^al^rm^bl^s=0$, and hence the above equation implies that

$$\Phi_{12}=0.$$

A similar argument can be used to prove that[14]

$$\Phi_{11}=0$$

and hence

$$R_{ab}=-2\Phi_{22}l_al_b. \tag{4.5}$$

In passing one should note that Eqs. (3.6) and (4.5) imply that since the gravitational field is of Petrov type II

$$-C_{arbs}l^rl^s={}^*R^*{}_{arbs}l^rl^s=\Omega l_al_b \tag{4.6}$$

for some function Ω, and thus Eq. (4.2) embodies only one constraint upon the field variables.

A straightforward calculation shows that under our present assumptions the electromagnetic field Eqs. (2.4)–(2.7) reduce to (4.1) along with

$$\delta C+(2\beta-\tau)C=\frac{k}{2}[\delta C+2(\beta+\tau)C]\Psi_2$$

$$-\frac{k}{2}[\bar{\delta}\bar{C}+2(\bar{\beta}+\bar{\tau})\bar{C}]\bar{\Psi}_2, \tag{4.7}$$

where

$$\Psi_2\equiv-\tfrac{1}{2}C_{abcd}(l^an^bl^cn^d-l^an^bm^c\bar{m}^d).$$

If we now assume that $\Psi_2=0$, and hence our gravitational field is of Petrov type III, then we find that the function Ω appearing in Eq. (4.6) vanishes. As a result of this additional assumption Eqs. (4.1), (4.2), (4.5), and (4.7) imply that the source-free generalized Einstein–Maxwell field equations reduce to

$$2\Phi_{22}=F,\quad DC=0,\ \text{and}\quad \delta C=(\tau-2\beta)C, \tag{4.8}$$

where F is given by Eq. (4.3)

To recapitulate the previous work we have

Theorem 5: Let V_4 be a spacetime of Petrov type III containing a null electromagnetic field F_{ab}. This spacetime and electromagnetic field satisfies the source-free generalized Einstein–Maxwell field equations with the pnc of F_{ab} being geodesic and shear-free (and thus the repeated pnc of the gravitational field) if and only if locally there exists a null tetrad $\{l^a,n^a,m^a,\bar{m}^a\}$ which is such that:

(a) l^a defines the repeated pnd of the gravitational field;

(b) $\kappa=\sigma=\rho=\epsilon=0$ for this tetrad;

(c) $F_{ab}=l_{[a}p_{b]}$ where $p_b=i(\bar{C}\bar{m}_b-Cm_b)$;

(d) R_{ab} is proportional to l_al_b; and

(e) Eq. (4.8) is satisfied.

Kundt[16] has determined the local form of the metric for spacetimes of Petrov type III which satisfy conditions (a), (b), and (d) in the above theorem. Kundt's metrics naturally divide into two classes depending upon whether or not it is possible to introduce null tetrads which satisfy (b) along with $\tau=0$. We shall now employ Kundt's results to determine the local form of solutions to the source-free generalized Einstein–Maxwell field equations which have the properties described in Theorem 5.

Case (*i*) $\tau=0$: If a spacetime of Petrov type III admits a null tetrad which satisfies conditions (a), (b), and (d) in the above theorem along with $\tau=0$, then locally there exists a chart (x,y,u,v) which is such that the metric has the form

$$ds^2=-dx^2-dy^2-2fdxdu+2dudv$$

$$+(h-vf_x-f^2)du^2, \tag{4.9}$$

where $h=h(x,y,u)$, $f=f(x,y,u)$, and

$$f_{xx}+f_{yy}=0. \tag{4.10}$$

In terms of this chart the repeated pnd of the gravitational field is defined by $\partial/\partial v$. I choose the vectors of the null tetrad to be

$$l^a\frac{\partial}{\partial x^a}=\frac{\partial}{\partial v},\quad n^a\frac{\partial}{\partial x^a}=\frac{\partial}{\partial u}+\tfrac{1}{2}(f^2+vf_x-h)\frac{\partial}{\partial v},$$

and (4.11)

$$m^a\frac{\partial}{\partial x^a}=\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial x}+\frac{i\partial}{\partial y}+\frac{f\partial}{\partial v}\right).$$

All of the spin coefficients of this null tetrad vanish except for γ,μ and ν and hence this tetrad does indeed satisfy conditions (a), (b), and (d) of Theorem 5 with $\tau=0$.

The "electromagnetic part" of Eq. (4.8) implies that C can be expressed as a differentiable function of $z\equiv x+iy$ and

u. The remaining field equation in (4.8) is

$$\tfrac{1}{2}(h_{xx}+h_{yy})-ff_{xx}+f_{xu}-\tfrac{3}{2}f_x^2-\tfrac{1}{2}f_y^2=C\bar{C}+kC'\bar{C}', \tag{4.12}$$

where

$$C'\equiv\frac{\partial C}{\partial z}$$

and f must satisfy Eq. (4.10).

The electromagnetic field corresponding to this solution to the source-free generalized Einstein–Maxwell field equations is

$$F_{ab}dx^a\wedge dx^b=\sqrt{2}\,[\mathrm{Im}(C)dx\wedge du+\mathrm{Re}(C)dy\wedge du] \tag{4.13}$$

where $\mathrm{Im}(C)$ and $\mathrm{Re}(C)$ denote the imaginary and real parts of C, resp.

At this time I would like to point out that if we desire the gravitational field to be of Petrov type N and such that (a), (b), and (d) remain valid with $\tau=0$, then[16] we can assume (without loss of generality) that $f=0$ in Eqs. (4.9), (4.11), and (4.12). In this case the gravitational field represents a plane-fronted gravitational wave with parallel rays[17] (see Ref. 16 or Ref. 2, p. 355).

Case (ii) $\tau\neq 0$: If a spacetime of Petrov type III admits a null tetrad which satisfies conditions (a), (b), and (d) in Theorem 5 along with $\tau\neq 0$, then locally there exists a chart (x,y,u,v) which is such that the metric has the form

$$ds^2=-dx^2-dy^2-2\left(f+\frac{2v}{x}\right)dx\,du+2du\,dv+\left[h-vf_x-\left(f+\frac{v}{x}\right)^2\right]du^2 \tag{4.14}$$

where $h=h(x,y,u)$, $f=f(x,y,u)$, and

$$xf_{xx}+xf_{yy}+2f_x=0. \tag{4.15}$$

The pnd of the gravitational field is defined by $\partial/\partial v$, and I choose the null tetrad to be

$$l^a\frac{\partial}{\partial x^a}=\frac{\partial}{\partial v},$$

$$n^a\frac{\partial}{\partial x^a}=\frac{\partial}{\partial u}+\frac{1}{2}\left[\left(f+\frac{v}{x}\right)^2+vf_x-h\right]\frac{\partial}{\partial v},$$

and

$$m^a\frac{\partial}{\partial x^a}=\frac{1}{\sqrt{2}}\left[\frac{\partial}{\partial x}+\frac{i\partial}{\partial y}+\left(f+\frac{2v}{x}\right)\frac{\partial}{\partial v}\right]. \tag{4.16}$$

The nonzero spin coefficients of this null tetrad are $\alpha,\beta,\gamma,\mu,\nu,\pi$, and τ. In particular

$$\alpha=\beta=\tfrac{1}{2}\tau=\frac{-1}{2\sqrt{2}x}. \tag{4.17}$$

Consequently this tetrad satisfies conditions (a), (b), and (d) of Theorem 5 with $\tau\neq 0$.

Due to Eq. (4.17) the "electromagnetic part" of Eq. (4.8) once again implies that C can be expressed as a differentiable function of $z\equiv x+iy$ and u. In addition the null electromagnetic field is still given by Eq. (4.13).

The remaining "gravitational" field equation in (4.8) is

$$\tfrac{1}{2}(h_{xx}+h_{yy})-\frac{1}{x}h_x+\frac{h}{x^2}-ff_{xx}+\frac{1}{x}ff_x+f_{xu}-\tfrac{3}{2}f_x^2-\tfrac{1}{2}f_y^2=C\bar{C}+\frac{k}{4}\left(\frac{(C+\bar{C})^2}{x^2}-\frac{2}{x}(C\bar{C}'+\bar{C}C')+4C'\bar{C}'\right), \tag{4.18}$$

where f must satisfy Eq. (4.15). In general the solutions to this equation lead to a spacetime which *cannot* be extended beyond $x=0$.

In concluding this case I would like to point out that if we desire the gravitational field to be of Petrov type N and such that (a), (b), and (d) are satisfied with $\tau\neq 0$, then[16] we can assume (without loss of generality) that $f=0$ in Eqs. (4.14), (4.16), and (4.18). The resulting gravitational field represents a plane-fronted gravitational wave (with nonparallel rays).

As an immediate consequence of the above work we obtain

Theorem 6: Let V_4 be a spacetime of Petrov type III (or N) which contains a null electromagnetic field F_{ab} whose corresponding pnc is geodesic and shear-free. If the source-free generalized Einstein–Maxwell field equations are satisfied, then locally the electromagnetic field is given by Eq. (4.13) and the metric is given by either Eq. (4.9) or (4.14) depending upon whether or not an adapted null tetrad can be found with $\tau=0$ along with $\kappa=\sigma=\rho=\epsilon=0$. The two field equations governing C, f, and h in these metrics are (4.10), (4.12) and (4.15), (4.18), resp. In addition the repeated pnd's of the gravitational and electromagnetic fields are aligned.

5. A RESTRICTED CLASS OF NULL ELECTROMAGNETIC FIELDS

In order to obtain the exact solutions to the source-free generalized Einstein–Maxwell field equations presented in the previous section we had to explicitly assume that the pnc of the null electromagnetic field is geodesic and shear-free. In addition we required the gravitational field to be of Petrov type III. I shall now demonstrate that it is possible to place a "single" algebraic restriction upon the pnd of F_{ab} which ensures that all of the above conditions are satisfied along with the gravitational field being of Petrov type N.

Suppose that l^a is tangent to the pnc of F_{ab} and such that

$$^*R^*{}_{abcd}l^c=0. \tag{5.1}$$

This in turn implies that $R_{ab}l^b=\tfrac{1}{2}Rl_a$ and hence we may employ Eq. (3.6) to conclude that[18]

$$C_{abcd}l^c=\frac{R}{12}(l_ag_{bd}-l_bg_{ad})+\tfrac{1}{2}(l_bR_{ad}-l_aR_{bd}). \tag{5.2}$$

Thus the gravitational field must be algebraically special.

Due to Eqs. (1.3) and (5.1) it is clear that

$$F_{ab}{}^*R^{*abhj} = 0$$

and hence

$$F_{ab;j}{}^*R^{*abhj} = 0$$

since ${}^*R^{*abhj}{}_{;j} = 0$. Consequently the "electromagnetic part" of the source-free generalized Einstein–Maxwell field equations reduces to Maxwell's equations in which case the pnc of F_{ab} is geodesic and shear-free [cf Eqs. (2.4) and (2.5)]. We can now appeal to Theorem 4 to deduce that this pnc must be hypersurface orthogonal and expansion-free. As a result of these facts there exists a null tetrad $\{l^a, n^a, m^a, \bar{m}^a\}$ which is such that $\kappa = \sigma = \rho = \epsilon = 0$. Under these conditions Eq. (4.5) is valid and thus Eq. (5.2) implies that the gravitational field is of Petrov type N.

Due to our work in Sec. 4 (cf. Theorem 6) we have just proved

Theorem 7: Suppose that V_4 is a spacetime containing a null electromagnetic field whose corresponding pnc is such that ${}^*R^*{}_{abcd}l^c = 0$, where l^c is tangent to this congruence. If the source-free generalized Einstein–Maxwell field equations are satisfied, then

(i) the gravitational field is of Petrov type N with repeated pnd defined by l^a,

(ii) the pnc of the electromagnetic field is geodesic, shear-free, hypersurface orthogonal and expansion-free,

(iii) locally the electromagnetic field is given by Eq. (4.13), and

(iv) locally the metric is given by either (4.9) or (4.14) (with $f = 0$) depending upon whether or not an adapted null tetrad can be found with $\tau = 0$ along with $\kappa = \sigma = \rho = \epsilon = 0$. The single field equation governing h in these metrics is given by Eq. (4.12) and (4.18), resp.

Remark: It can be shown that if the pnc of a null electromagnetic field satisfies assumption (5.1) and the source-free Einstein–Maxwell field equations (1.1) hold, then conclusions (i)–(iv) of the above theorem are valid[19] with $k = 0$ in Eqs. (4.12) and (4.18).

6. SUMMARY AND CONCLUSIONS

The previous work has shown us that if the source-free generalized Einstein–Maxwell field equations are satisfied by a null electromagnetic field F_{ab}, then the pnc of F_{ab} *most likely* will not be geodesic, shear-free or a repeated pnc of the gravitational field. However, if it is geodesic and shear-free, then not only must it be a repeated pnc of the gravitational field but it must also be hypersurface orthogonal and expansion free! As of yet no one has produced a solution to Eq. (1.4) corresponding to a null electromagnetic field whose pnc is *not* geodesic and shear-free. I believe that such a solution could prove to be quite interesting.

In Ref. 8 it is argued that a possible alternative to the usual energy–momentum tensor $-T_{ab}$[cf. Eq. (1.2)] of the electromagnetic field used in general relativity is provided by $\mathscr{T}_{ab} \equiv -T_{ab} + kA_{ab}$, where A_{ab} is defined by Eq. (1.5). For the two classes of solutions to the source-free generalized Einstein-Maxwell field equations presented in Theorem 6 we have the following expressions for $-T_{ab}$ and $\mathscr{T}_{ab}$.

Case (*i*) $\tau = 0$:

$$-T_{ab} = \frac{1}{8\pi} C\bar{C} l_a l_b, \quad \mathscr{T}_{ab} = \frac{1}{8\pi}(C\bar{C} + kC'\bar{C}')l_a l_b. \quad (6.1)$$

Case (*ii*) $\tau \neq 0$:

$$-T_{ab} = \frac{1}{8\pi} C\bar{C} l_a l_b,$$

$$\mathscr{T}_{ab} = \frac{1}{8\pi}\left[C\bar{C} + \frac{k}{4}\left(\frac{(C+\bar{C})^2}{x^2} - \frac{2(C\bar{C}' + \bar{C}C')}{x} + 4C'\bar{C}'\right)\right] l_a l_b. \quad (6.2)$$

Suppose that $\mathscr{O}$ is an observer in our spacetime whose world line has unit tangent vector u. Then in either of the above cases $-T(u,u) \equiv -T_{ab}u^a u^b > 0$, as was to be expected. In case (i) if $k > 0$, then $\mathscr{T}(u,u) \equiv \mathscr{T}_{ab}u^a u^b > 0$, while if $k < 0$, then $\mathscr{T}(u,u)$ will be indefinite. Furthermore if we select $k < 0$ in case (i), then we can obtain solutions to the source-free generalized Einstein Maxwell field equations which are such that[20] $\mathscr{T}_{ab} = 0$.

In case (ii) $\mathscr{T}(u,u)$ will be indefinite (in general) no matter what our choice of sign for k, although I suspect that if $k > 0$, then the chances of $\mathscr{T}(u,u)$ being positive for observers located at large values of x are pretty good.

The above observations, along with the result presented in Refs. 6 and 7, lead me to believe that k must be chosen to be positive if the generalized Einstein–Maxwell field theory is to be regarded as a viable alternative to the Einstein–Maxwell field theory.

ACKNOWLEDGMENTS

I would like to thank J. Wainwright for his invaluable assistance in preparing the present paper, and in particular for showing me how the computer program CAMAL can be used to compute the Newman–Penrose spin coefficients. I also wish to thank the Natural Sciences and Engineering Research Council of Canada for financial support.

[1]The notational conventions used in this paper are the same as those employed in E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962), with the following exceptions: (i) tensor indices will be denoted by lower case Latin letters, and (ii) geometrized units will be used in terms of which the speed of light c and the gravitational constant G are equal to 1.
[2]F. A. E. Pirani, "Introduction to Gravitational Radiation," article in *Lectures on General Relativity* (Prentice-Hall, Englewood Cliffs, New Jersey, 1964), pp. 317–20.
[3]L. Mariot, C. R. Acad. Sci. **238**, 2055 (1954).
[4]I. Robinson, J. Math. Phys. **2**, 290 (1961).
[5]W. Kundt and M. Trumper, Akad. Wiss. Mainz **12**, 985 (1962).
[6]G. W. Horndeski, Phys. Rev. D **17**, 391 (1978). Note that our present notational conventions differ from those used in this reference.
[7]G. W. Horndeski, J. Math. Phys. **19**, 668 (1978).
[8]G. W. Horndeski, J. Math. Phys. **17**, 1980 (1976).
[9]G. W. Horndeski and J. Wainwright, Phys. Rev. D **16**, 1691 (1977).

[10]It will be assumed that C is nonzero, except, perhaps on a set of measure zero.

[11]For the definition of the "spin coefficients" appearing on the left-hand side of Eqs. (2.4)–(2.7), see Ref. 1.

[12]If $\kappa = 0$, then it can be shown that l^a is an eigenvector of the Ricci tensor.

[13]The Ricci and Bianchi identities which we shall require can be found on pages 350–1 of Ref. 2.

[14]In carrying out the omitted details you should note that since the Weyl tensor is trace-free $C_{abcd} l^a m^b l^c \bar{m}^d = 0$.

[15]D. W. Trim, "Neutrino and Electromagnetic Fields in Curved Space–Time," Ph. D. thesis (unpublished), University of Waterloo, page 118 (1972). Such solutions can also be obtained using the results in D. W. Trim and J. Wainwright, J. Math. Phys. **15**, 535 (1974).

[16]W. Kundt, Z. Phys. **163**, 77 (1961).

[17]M. Gurses and M. Halil have also produced solutions to the source-free generalized Einstein–Maxwell field equations of this form, cf., Phys. Lett. A **68**, 182 (1978).

[18]Since $R \neq 0$ (yet) Eq. (3.6) must be modified by the addition of $R(g_{ad}g_{bc} - g_{ac}g_{bd})/3$ to the right-hand side.

[19]One can prove that $\rho = 0$ using the Bianchi identities.

[20]When $k < 0$ in case (i) then the general solution for C to $\mathscr{T}_{ab} = 0$ is easily found to be

$$C(z,u) = B \exp[l^{-1} \exp(i\psi)z],$$

where B and ψ are arbitrary complex and real valued functions of u resp., and $k = l^{-2}$, $l \in \mathbb{R}^+$. Each of these choices of C can be combined with solutions to Eqs. (4.10) and (4.12) to produce a solution to the source-free generalized Einstein-Maxwell field equations with $\mathscr{T}_{ab} = 0$.

The classical limit of quantum dissipative generators

J. V. Pulè[a)] and A. Verbeure

Instituut voor Theoretische Fysica, K.U. Leuven, B-3030 Leuven, Belgium
(Received 26 May 1978; revised manuscript received September 1978)

For finite quantum systems the classical limit of the general dissipative generator of a semigroup of completely positive maps is obtained, yielding a generalized Fokker–Planck generator.

1. INTRODUCTION

The connection between quantum-mechanical dynamical semigroups and Markov semigroups has been studied in a number of papers (see, e.g., Refs. 1–4).

A new aspect of the connection was studied in Ref. 5 by taking the classical limit ($\hbar\to 0$) of evolution of the type

$$T_\hbar(t) = \exp[(\hbar^{-1}Z + K)t]$$

acting on a Banach space of states on the bounded operators of $L^2(R)$. The operator Z is the generator of the free automorphic evolution and K is a bounded operator describing a perturbation due to the influence of some kind of reservoir. By the particular choice of K, independent of $\hbar$, a Markov semigroup on phase space, the dissipative part of its generator is given essentially by a convolution operator.

We consider here the classical limit of a general dissipative operator using the point of view of Hepp.[6] For quantum systems with a finite number of degrees of freedom we study the classical limit of the expectation values in coherent states of the quantum generator acting on arbitrary bounded quantum observables.

In Sec. 2 we treat a simple model in which the classical limit of the generator is given by the well known Fokker–Planck differential operator acting on the classical phase space functions. For classical limit considerations this model suggests also the introduction of an explicit $\hbar$ dependence in the dissipative part of the quantum generator, which can be considered as the quantum effect of the reservoir.

Section 3 is devoted to the classical limit of the general dissipative quantum generator. In the limit a generalized Fokker–Planck generator is obtained in the sense that it contains terms with a single Poisson bracket and other terms with a double Poisson bracket.

2. MODEL

Consider the Hilbert space $H = \mathscr{L}^2(R)$ and the self-adjoint extensions Q and P of the operators denoted by the same symbol

$$(Q\psi)(x) = x\psi(x),$$
$$\psi\in H.$$
$$(P\psi)(x) = -i\frac{d}{dx}\psi(x),$$

For any $\alpha = (\alpha_1,\alpha_2)\in R^2$ consider the Weyl operators

$$W(\alpha) = \exp[i(\alpha_1 Q + \alpha_2 P)]$$

satisfying the Weyl relations

$$W(\alpha)W(\beta) = W(\alpha+\beta)\exp[-i\sigma(\alpha,\beta)], \quad \alpha,\beta\in R^2,$$

where σ is the symplectic form:

$$\sigma(\alpha,\beta) = \tfrac{1}{2}(\alpha_1\beta_2 - \alpha_2\beta_1).$$

Note that this is the ordinary set up of quantum mechanics but $\hbar$-independent.

Now we introduce a model for irreversible time evolution on the CCR algebra generated by the Weyl operators: for $t\geqslant 0$ we define

$$\tau_t W(\alpha) = W(e^{Gt}\alpha)\frac{\exp[-\frac{1}{2}\sigma(J\alpha,\alpha)]}{\exp[-\frac{1}{2}\sigma(Je^{Gt}\alpha,e^{Gt}\alpha)]}, \qquad (1)$$

where

$$G = \begin{pmatrix} -\eta & 1 \\ -1 & -\eta \end{pmatrix}, \quad \eta > 0,$$

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

It is easily checked that $(\tau_t)_{t\geqslant 0}$ is a semigroup of completely positive maps.[7,8] The infinitesimal generator of this semigroup in the sense of Ref. 9 is given by

$$i[H,\cdot] + L(v),$$

where

$$H = -\tfrac{1}{2}(Q^2 + P^2),$$
$$L(v)x = v^*xv - \tfrac{1}{2}(v^*vx + xv^*v), \quad x\in\mathscr{A}, \qquad (2)$$
$$v = \eta^{1/2}(Q + iP),$$

as should be expected from Refs. 10 and 11.

This type of evolution (1) can be obtained by coupling the system weakly to a reservoir of harmonic oscillators at zero temperature.[12,13] Now we are interested in the classical limit of the evolution given in (1), by introducing explicitly $\hbar$. As one expects, the classical evolution equation should be obtained in the limit $\hbar\to 0$ in some sense. It is well known that the coherent states are the minimal uncertainty states for the basic observables Q and P of quantum mechanics. In Ref. 6 Hepp gave a rigorous proof that the classical limit ($\hbar\to 0$) of the quantal observables under a conservative time evolution in these states gives their classical evolution.

[a)]On leave of absence from the University of Malta.

 0022-2488/79/040733-03$01.00

Here we take the same point of view for a nonconservative time evolution. Consider now

$$\tau_t^{\hbar} W(\alpha) = W(e^{Gt}\alpha)\frac{\exp[-(\sigma/2\hbar)(J\alpha,\alpha)]}{\exp[-(\sigma/2\hbar)(Je^{Gt}\alpha,e^{Gt}\alpha)]}. \quad (3)$$

Note that, for all $\hbar \leqslant 1$, $(\tau_t^{\hbar})_{t>0}$ remains a semigroup of completely positive maps. We have the following proposition:

Proposition 2.1: With the above notation

$$\text{s-}\lim_{\hbar\to 0} W(\hbar^{-1/2}J\beta)^*\tau_t^{\hbar}(W(\hbar^{1/2}\alpha))W(\hbar^{-1/2}J\beta) = f_{\alpha,t}(\beta), \quad (4)$$

where $f_{\alpha,t}(\beta)$ is the solution of the Fokker–Planck equation

$$\frac{d}{dt}f_{\alpha,t}(\beta) = \left\{\frac{q^2+p^2}{2}, f_{\alpha,t}(\beta)\right\} - \eta p\frac{\partial}{\partial p}f_{\alpha,t}(\beta) - \eta q\frac{\partial}{\partial q}f_{\alpha,t}(\beta) + \frac{1}{2}\eta\frac{\partial^2}{\partial p^2}f_{\alpha,t}(\beta) + \frac{1}{2}\eta\frac{\partial^2}{\partial q^2}f_{\alpha,t}(\beta), \quad (5)$$

where $\beta = (q,p)$ with initial value $f_{\alpha,0}(\beta) = \exp[2i\sigma(J\beta,\alpha)]$; $\{\cdot,\cdot\}$ stands for the Poisson bracket.

Proof: It is easily checked that the left-hand side of (4) is equal to

$$\text{s-}\lim_{\hbar\to 0} W(e^{Gt}\hbar^{1/2}\alpha)\exp[2i\sigma(J\beta,e^{Gt}\alpha)] \times\frac{\exp[-\frac{1}{2}\sigma(J\alpha,\alpha)]}{\exp[-\frac{1}{2}\sigma(J\,e^{Gt}\alpha,e^{Gt}\alpha)]}$$

$$= \exp[2i\sigma(J\beta,e^{Gt}\alpha)]\frac{\exp[-\frac{1}{2}\sigma(J\alpha,\alpha)]}{\exp[-\frac{1}{2}\sigma(Je^{Gt}\alpha,e^{Gt}\alpha)]}$$

$$= f_{\alpha,t}(\beta),$$

which satisfies (5). □

Motivated by the result of Proposition 2.1 which gives the expected results, we use the same procedure in the next section to obtain the classical limit of dissipative generators of the type $\Sigma_i L(v_i)$ as defined in (2).

Furthermore, we remark that the explicit $\hbar$ dependence of the dissipative part of the generator of the semigroup (3) is given by

$$L(v_1^{\hbar}) + L(v_2^{\hbar}),$$

where

$$v_1 = (\eta/\hbar)^{1/2}Q + i(\eta\hbar)^{1/2}P$$

and $\quad (6)$

$$v_2 = i(\eta/\hbar - \eta\hbar)^{1/2}P.$$

This particular $\hbar$ dependence of the generator can be thought of as coming from the quantum effects of the reservoir.

3. THE CLASSICAL LIMIT OF THE QUANTUM SEMIGROUP GENERATORS

Consider the generator $L(v)$ as in (2) with v a bounded operator on $L^2(R)$. The generalization to $L^2(R^n)$, systems with n degrees of freedom is straightforward. If $v = A + iB$, where A and B are self-adjoint, then for $x\in\mathscr{A}$:

$$L(v)x = -\tfrac{1}{2}[A,[A,x]] - \tfrac{1}{2}[B,[B,x]] + \tfrac{1}{2}i\{B[A,x] - A[B,x] + [A,x]B - [B,x]A\}. \quad (7)$$

Let $a_{\hbar}:(q,p)\to a_{\hbar}(q,p)$ and $b_{\hbar}:(q,p)\to b_{\hbar}(q,p)$ be arbitrary, real-valued L^1 functions on the classical phase space R^2 whose first and second derivatives are L_1; the corresponding Weyl quantized observables are given by

$$A_{\hbar}(Q,P) = \frac{1}{2\pi}\int_{R^2} du\,\tilde{a}_{\hbar}(u)W(u), \quad (8a)$$

$$B_{\hbar}(Q,P) = \frac{1}{2\pi}\int_{R^2} du\,\tilde{b}_{\hbar}(u)W(u), \quad (8b)$$

where $\tilde{a}_{\hbar}$ and $\tilde{b}_{\hbar}$ are the Fourier transforms of $a_{\hbar}$ and $b_{\hbar}$ and the integrals are Bochner integrals in the strong operator topology. Motivated by formula (6), we assume the following $\hbar$ dependence for $a_{\hbar}$ and $b_{\hbar}$:

$$a_{\hbar} = \frac{1}{\hbar^{\mu}}(a_0 + \hbar^{\epsilon}a_{1,\hbar}),$$

$$b_{\hbar} = \frac{1}{\hbar^{\nu}}(b_0 + \hbar^{\epsilon'}b_{1,\hbar}), \quad (9)$$

where $\mu,\nu\in R$, $\epsilon,\epsilon' > 0$, and the L^1 norms of a_1 and b_1 are uniformly bounded in $\hbar$ on some compact neighborhood of zero. Denote $L_{\hbar}(Q,P) = L(v_{\hbar}(Q,P))$, where $v_{\hbar}(Q,P) = A_{\hbar} + iB_{\hbar}$. The dissipative time evolution $\tau_t^{\hbar}$ is then given by

$$\tau_t^{\hbar} = \exp(t/\hbar)L_{\hbar}(\hbar^{1/2}Q,\hbar^{1/2}P).$$

Theorem 3.1: With the above notations, for all $\mu,\nu\in R$ Satisfying $\mu + \nu\leqslant$, $\nu\leqslant\frac{1}{2}$, $\mu\leqslant\frac{1}{2}$ the following limit:

$$s\text{-}\lim_{\hbar\to 0} W(\hbar^{-1/2}J\beta)^*[(L_{\hbar}/\hbar)(\hbar^{1/2}Q,\hbar^{1/2}P)W(\hbar^{1/2}\alpha)]W(\hbar^{-1/2}J\beta)$$

exists for all α, $\beta\in R^2$ and is given by

$$\delta_{\mu+\nu,0}[a_0(\beta)\{b_0(\beta),\exp[2i\sigma(J\beta,\alpha)]\} - b_0(\beta)\{a_0(\beta),\exp[2i\sigma(J\beta,\alpha)]\}] + \tfrac{1}{2}\delta_{\mu,1/2}\{a_0(\beta),\{a_0(\beta),\exp[2i\sigma(J\beta,\alpha)]\}\} + \tfrac{1}{2}\delta_{\nu,1/2}\{b_0(\beta),\{b_0(\beta),\exp[2i\sigma(J\beta,\alpha)]\}\}, \quad (10)$$

where $\{\cdot,\cdot\}$ is the Poisson bracket with respect to the variable $\beta = (q,p)$.

Proof: Consider first a term of the type $iB[A,x]$ (single commutator) in (7). Using (8) and (9), we can perform the following calculation:

$$W(\hbar^{-1/2}J\beta)^*\frac{i}{2\hbar}B_{\hbar}(\hbar^{1/2}Q,\hbar^{1/2}P) \times[A_{\hbar}(\hbar^{1/2}Q,\hbar^{1/2}P),W(\hbar^{1/2}\alpha)]W(\hbar^{-1/2}J\beta)$$

$$= -\frac{1}{\hbar^{\nu+\mu}}\left(\frac{1}{2\pi}\right)^2 \int\int du dv\, [\tilde{a}_0(u) + \hbar^\epsilon \tilde{a}_{1,\hbar}(u)]$$

$$\times [\tilde{b}_0(v) + \hbar^{\epsilon'} \tilde{b}_{1\hbar}(v)]\frac{\sin\hbar\sigma(\alpha,u)}{\hbar}$$

$$\times \exp\{i[\hbar\sigma(u+\alpha,v) + 2\sigma(J\beta,u+v+\alpha)]\}$$
$$\times W(\hbar^{1/2}(u+v+\alpha)).$$

It is now clear that the $\lim \hbar \to 0$ exists if $\mu + \nu \leqslant 0$. If $\mu + \nu < 0$, then the limit is clearly zero; if $\mu + \nu = 0$, then the limit is given by

$$-\left(\frac{1}{2\pi}\right)^2 \int du dv\, \tilde{a}_0(u)\tilde{b}_0(v)\sigma(\alpha,u)$$
$$\times \exp[2i\sigma(J\beta,u+v+\alpha)]$$
$$= -\frac{b_0(\beta)}{2\pi}\int du\, \tilde{a}_0(u)\sigma(\alpha,u)\exp[2i\sigma(J\beta,u+\alpha)]$$
$$= \frac{b_0(\beta)}{2}\{e^{2i\sigma(J\beta,\alpha)},a_0(\beta)\}.$$

Now we consider the term $-\frac{1}{2}[A,[A,x]]$ (double commutator term) in (7). Using again (8) and (9), we calculate

$$W(\hbar^{-1/2}J\beta)^*(-\tfrac{1}{2})$$
$$\times [A_\hbar(\hbar^{1/2}Q,\hbar^{1/2}P),[A_\hbar(\hbar^{1/2}Q,\hbar^{1/2}P),W(\hbar^{1/2}\alpha)]]\,W(\hbar^{-1/2}J\beta)$$
$$= \frac{2}{\hbar^{(2\mu-1)}}\left(\frac{1}{2\pi}\right)^2\int\int du dv\, [\tilde{a}_0(u) + \hbar^\epsilon \tilde{a}_{1,\hbar}(u)]$$

$$\times [\tilde{a}_0(v) + \hbar^\epsilon \tilde{a}_{1,\hbar}(v)]\frac{\sin\hbar\sigma(\alpha,v)\sin\hbar\sigma(v+\alpha,u)}{\hbar^2}$$

$$\times \exp[2i\sigma(J\beta,u+v+\alpha)]\,W(\hbar^{1/2}(u+v+\alpha)).$$

The strong limit $\hbar \to 0$ of this expression exists if $\mu \leqslant \frac{1}{2}$. If $\mu < \frac{1}{2}$, the limit equals zero, and if $\mu = \frac{1}{2}$, the limit equals

$$2\left(\frac{1}{2\pi}\right)^2\int\int du dv\, \tilde{a}_0(u)\tilde{a}_0(v)\sigma(\alpha,v)\sigma(\alpha+v,u)$$
$$\times \exp[2i\sigma(J_\beta,u+v+\alpha)]$$
$$= \tfrac{1}{2}\{a_0(\beta),\{a_0(\beta),\exp[2i\sigma(J\beta,\alpha)]\}\}.$$

By treating the other terms similarly, we get the result (10). □

Remark that the result of Theorem 3.1 is trivially generalized to arbitrary functions.

Corollary 3.2: With the above notations, let f furthermore be an L^1-classical observable, whose first and second derivative is L^1, with $F(Q,P) = (1/2\pi)\int_{R^2} d\alpha\, \tilde{f}(\alpha)W(\alpha)$ the corresponding quantum observable; then under the same conditions of the theorem:

$$\text{s-}\lim_{\hbar\to 0} W(\hbar^{-1/2}J\beta)^*$$
$$\times [(L_\hbar/\hbar)(\hbar^{1/2}Q,\hbar^{1/2}P)F(\hbar^{1/2}Q,\hbar^{1/2}P)]\,W(\hbar^{-1/2}J\beta)$$
$$= \delta_{\mu+\nu,0}[a_0(\beta)\{b_0(\beta),f(\beta)\} - b_0(\beta)\{a_0(\beta),f(\beta)\}]$$
$$+ \tfrac{1}{2}\delta_{\mu,1/2}\{a_0(\beta),\{a_0(\beta),f(\beta)\}\}$$
$$+ \tfrac{1}{2}\delta_{\nu,1/2}\{a_0(\beta),\{a_0(\beta),f(\beta)\}\}. \tag{11}$$
□

As a result of this, the quantal generator (7) of a semigroup of completely positive unity preserving maps has as classical limit a differential operator containing terms with a single Poisson bracket as well as terms with a double Poisson bracket. This differential operator on the functions on classical phase space is a generalization of the Fokker–Planck differential operator.

The generator of the dissipative perturbation used in Ref. 14 to prove the KMS relation from a subvariational principle is a special case of (11) with $\mu + \nu = 0, \mu < \frac{1}{2}, \nu < \frac{1}{2}$. This means that it can be obtained as the classical limit of generators of the form (7), where a and b are independent of $\hbar$.

Finally, we mention that a related problem of finding differential equations, for extensive observables only, by thermodynamic limits of master equations, instead of by the classical limit, can be found in Ref. 15.

ACKNOWLEDGMENT

The authors than Dr. G. Lindblad for pointing out an error in the original version and for bringing to their attention Ref. 15.

[1]E.B. Davies, Commun. Math. Phys. **15**, 277 (1969).
[2]E.B. Davies, Commun. Math. Phys. **19**, 83 (1970).
[3]E.B. Davies, Zeit. Wahrschein. **23**, 261 (1972).
[4]A. Kossakowski, Rep. Math. Phys. **3**, 247 (1972).
[5]E.B. Davies, Commun. Math. Phys. **49**, 113 (1976).
[6]K. Hepp, Commun. Math. Phys. **35**, 265 (1974).
[7]D.E. Evans and J.T. Lewis, J. Funct. Anal. **26**, 369–77 (1977).
[8]B. Demoen, P. Vanheuverzwijn, and A. Verbeure, Lett. Math. Phys. **2**, 161 (1977).
[9]P. Vanheuverzwijn, Ann. Inst. Henri Poincaré **29**, 123 (1978).
[10]V. Gorini, A. Kossakowski, and E.C.G. Sudarshan, J. Math. Phys. **17**, 821 (1976).
[11]G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
[12]E.B. Davies, Commun. Math. Phys. **33**, 171 (1973).
[13]J.T. Lewis and L.C. Thomas, in *Functional Integration*, edited by A.M. Arthurs, Proc. Int. Conf. Cumberland Lodge, London, 1974 (Clarendon, Oxford, 974).
[14]B. Demoen and A. Verbeure, Commun. Math. Phys. **62**, 191 (1978).
[15]H.I. to and T. Nakagomi, Prog. Theor. Phys. **57**, 54 (1977).

First order estimates of the error in approximate calculations of scattering phase shifts

J. W. Darewych and J. Sokoloff

Department of Physics, York University, Toronto, Ontario M3J 1P3
(Received 1 August 1978)

We discuss the question of error estimation in approximate calculations of scattering phase shifts. The Kato integral identity between the exact and approximate solutions is used as the starting point to determine an upper bound to the absolute value of the difference between the exact and approximate result. This bound involves the maximum value of the modulus of the exact wavefunction as a factor. For the potential scattering case it is shown that this maximum occurs in the asymptotic region, if the potential is monotonically decreasing (with decreasing r). For more general potentials simple calculable bounds to the maximum of the wavefunction are derived, for energies which are everywhere higher than the potential. The results are illustrated for scattering by an (attractive) exponential potential and are compared to bounds obtained previously by Bardsley, Gerjuoy, and Sukumar, who have used the Lippman–Schwinger equation to determine bounds to the maximum modulus of the wavefunction.

1. INTRODUCTION

In any approximate calculation it is very important to be able to estimate the error, so that some unambiguous statement can be made about the accuracy of the approximate result. For approximate scattering calculations Kato,[1] Bardsley, Gerjuoy, and Sukumar,[2] and Darewych and Pooran[3] have given expressions for the error in the approximate phase shift. These expressions can be applied with ease in the case of potential scattering problems and yield very good results. However, the primary interest in them lies in their possible applicability to scattering by compound targets.

All three methods mentioned above generalize formally to the case of scattering by many-body targets in a straightforward way. The difficulty is that their rigorous application to even the simplest targets, such as hydrogen atoms, becomes computationally intractable. To our knowledge, no such rigorous calculation has yet been made, though Shimamura's[4] application of Kato's method to elastic s-wave scattering of electrons by atomic hydrogen comes closest to the mark. Thus it is important to determine computationally simpler error bounds.

For clarity of discussion, let us consider the l-wave scattering by a short range central potential $V(r)$. We are thus seeking a solution of the equation

$$Lu(r) = 0, \tag{1}$$

where

$$L = \frac{d^2}{dr^2} + k^2 - \frac{l(l+1)}{r^2} - \frac{2m}{\hbar^2} V(r), \tag{2}$$

subject to the condition that

$$u(r=0) = 0 \tag{3}$$

and having the asymptotic form

$$u(r\rightarrow\infty) = A(\eta)\sin(kr - l\frac{\pi}{2} + \eta), \tag{4}$$

where $A(\eta)$ is an arbitrary normalization factor, most commonly chosen to be sec η or 1.

Variational approximations to the exact phase shift η are based on the stationary principle introduced first by Hulthén,[5] which can be written in the form

$$I[u_T] = kA(\eta)B(\eta_T)\sin(\eta_T - \eta) + I[u_T - u], \tag{5}$$

where

$$I[u_T] = \int_0^\infty u_T L u_T dr, \tag{6}$$

and u_T is an approximation to u, satisfying the same boundary conditions (3) and (4) as u, except that the unknown, exact η is replaced by a trial value η_T. The normalization factor $B(\eta)$ need not be the same as $A(\eta)$. [Hulthén used the normalization $A = B = 1$ in his original work. Kohn[6] popularized the form $A = B = \sec\eta$, while Kato[1] introduced the one-parameter form $A(\eta) = \pm(1+\lambda^2)^{1/2}$where $\lambda = \cot(\eta - \theta)$, θ being an adjustable constant. Equation (5), with this last form of asymptotic normalization, has come to be called the Kato identity.] Clearly (5) can be used to get an approximation to η, in terms of u_T and all the adjustable parameters in it, up to the "second order" (in $u_T - u$) error term $I[u_T - u]$. Kato[1] and Darewych and Pooran[3] have given calculable expressions for upper bounds to the absolute value of this second order error.

Bardsley, Gerjuoy, and Sukumar[2] base their error bounds on the nonstationary version of the Kato identity (5), viz

$$kA(\eta)B(\eta_T)\sin(\eta_T - \eta) = I[u, u_T - u], \tag{7}$$

where

$$I[u, u_T - u] = \int_0^\infty uL(u_T - u)dr = \int_0^\infty uLu_T dr. \tag{8}$$

It is evident that $I[u, u_T - u]$ is first order in the "small" quantity $u_T - u$. Hence (7) yields only a first order approximation to η, namely η_T, if I is neglected. Clearly,

$$|I[u, u_T - u]| = \left|\int_0^\infty uLu_T dr\right| \leqslant |u|_m \int_0^\infty |Lu_T| dr, \tag{9}$$

where the subscript m refers to the maximum of the quantity

 0022-2488/79/040736-04$01.00

over the range $0 \leqslant r \leqslant \infty$. Equation (9) thus gives a first order estimate of the error, $|\eta - \eta_T|$, if $|u|_m$, or an upper bound to it, can be determined.

If $w(r)$ is an arbitrary weight function, then it follows from Schwartz's inequality that

$$\mathscr{L}_1 = \int_0^\infty |Lu_T| dr = \int_0^\infty |w^{-1} w L u_T| dr$$

$$\leqslant \left[\int_0^\infty |w|^{-2} dr \int_0^\infty |wLu_T|^2 dr \right]^{1/2} = \mathscr{L}_2, \qquad (10)$$

for any weight function $w(r)$. Bardsley, Gerjuoy, and Sukumar[2] have shown that

$$|u|_m \leqslant (1 - g_m)^{-1} \quad \text{if } g_m < 1, \qquad (11)$$

where

$$g(r) = \int_0^\infty |G(r,r')V(r')| dr', \qquad (12)$$

$G(r,r')$ being the appropriate Green's function. $\mathscr{L}_2/(1-g_m)$ is in fact the first order error bound derived by Bardsley *et al*[2]. It is somewhat poorer than the bound $\mathscr{L}_1/(1-g_m)$ because of the inequality (10). Both of these error bounds are in turn poorer by an order in a "small quantity" than those introduced by Kato,[1] viz.,

$$|I[u_T - u]| < \mathscr{L}_2^2/\lambda \qquad (13)$$

and Darewych and Pooran[3]:

$$|I[u_T - u]| < |f|_m \mathscr{L}_1/(1-g_m), \quad \text{if } g_m < 1, \qquad (14)$$

where

$$f(r) = \int_0^\infty G(r,r') L u_T(r') dr'.$$

In (13) λ is a bound to a particular eigenvalue of an auxiliary eigenvalue problem.

The complicating features, when any of these bounds are applied in practice, particularly for compound targets, are the evaluation of the various integrals $[\mathscr{L}_1, \mathscr{L}_2, f(r)]$ and the determination of the of the auxiliary quantities such as λ and g_m. Assuming that the evaluation of $\mathscr{L}_1$, $\mathscr{L}_2$ can be tackled by one method or other (see Ref. 3, 7), it is then necessary to devise tractable methods for determining the remaining factors that enter into the error bound expression.

In the present paper we shall restrict our attention to the nonstationary approximation based on Eq. (7). In particular, we discuss an alternate method for deriving bounds on $|u|_m$ to that given by Bardsley *et al.*[2] [Eq.(11)], which involves the Green's function and thus becomes intractable, without introducing nonrigorous simplifications, when applied to scattering by a compound target.

We emphasize that the identities (5) and (7) and the inequalities (9), (13), and (14) hold for any approximate solution u_T, not only for variationally derived ones. Thus (7), coupled with a tractable bound on $|u|_m$ [Eq.(9)], could be used to determine bounds on the exact phase shifts, by using nonvariational methods, such as the polarized orbital approximation in low energy electron–atom scattering.

2. Bounds on the Amplitude of the Wavefunction

Bardsley *et al.*[2] have derived their expression for the bound to $|u|_m$ by using the Lippman–Schwinger equation for u. This bound, (11), is limited to incident energies and/or potentials such that the condition $g_m < 1$ is satisfied. It also has the undesirable feature of involving the Green's function, which, for compound targets, has a very complex structure involving a sum over all the states of the target. Thus it is important to investigate other avenues of approach that may circumvent this difficulty.

Let us approximate the effective potential

$$U(r) = \frac{2m}{\hbar^2} V(r) + \frac{l(l+1)}{r^2} \qquad (15)$$

by the function

$$\bar{U}(r) = U(r_{n-1}) \quad \text{for } r_{n-1} \leqslant r < r_n, \qquad (16)$$

where $0 = r_0 < r_1 < r_2 < \cdots < r_N = R$, are a set of N points and R is taken to be large enough so that $U(R)$ is negligible. With this potential, the solution is

$$\bar{u}(r) = A_{n-1} \sin(k_{n-1} r + \eta_{n-1}), \quad r_{n-1} \leqslant r \leqslant r_n, \qquad (17)$$

where

$$k_{n-1}^2 = k^2 - U(r_{n-1}). \qquad (18)$$

Continuity of $\bar{u}(r)$ and its derivative at r_n yields

$$\frac{A_{n-1}^2}{A_n^2} = 1 + \left(\frac{k_n^2}{k_{n-1}^2} - 1 \right) \cos^2(k_n r_n + \eta_n) = q_n \qquad (19)$$

and

$$\frac{k_n}{k_{n-1}} \tan(k_{n-1} r_n + \eta_{n-1}) = \tan(k_n r_n + \eta_n). \qquad (20)$$

We see immediately that $A_{n-1}^2 < A_n^2$ whenever $k_n^2 < k_{n-1}^2$, i.e., whenever $U(r_n) > U(r_{n-1})$ and vice versa.

Since $\bar{u}(r)$ approaches $u(r)$ arbitrarily closely as $N \to \infty$, we conclude that the amplitude of the solution decreases with decreasing r whenever the potential decreases with decreasing r and vice versa. In particular for any monotonically decreasing (with decreasing r) potential (usually this would mean $l = 0$), the amplitude takes on its maximum value in the asymptotic region, $r \geqslant R$.

If the potential increases from the asymptotic value, zero, as r decreases up to a maximum value $U(r_m) < k^2$, and decreases thereafter, then the amplitude takes on its maximum value at r_m. We can obtain an upper bound to this maximum value if we note that (19) implies

$$A_{N-i}^2 = \left(\prod_{j=0}^{i-1} q_{N-j} \right) A_N^2. \qquad (21)$$

Now for $r > r_m$ where $k_n^2 > k_{n-1}^2$, $q_n < Q_n = k_n^2/k_{n-1}^2$, i.e.,

$$A^2_{N-i} < \left(\prod_{j=0}^{i-1} Q_{N-j}\right) A^2_N = \frac{k^2}{k^2 - U(r_i)} A^2_N.$$

Thus the maximum amplitude squared, which occurs at r_m, is smaller than $k^2/[k^2 - U(r_m)]$ times the amplitude in the asymptotic region, i.e.,

$$|u(r)|^2_m < \frac{k^2 A^2(\eta)}{k^2 - U(r_m)} \tag{22}$$

provided $U(r_m) < k^2$. Extending this argumentation to any short range potential that has relative maxima at $R_1, R_3, R_5, \cdots$ and relative minima at $R_2, R_4, R_6, \cdots$ $(R_{n+1} < R_n)$, we obtain

$$|u(r)|^2_m < \frac{k^2}{k^2 - U(R_1)} \frac{k^2 - U(R_2)}{k^2 - U(R_3)} \cdots \frac{k^2 - U(R_{2n})}{k^2 - U(R_{2n+1})} A^2 \tag{23}$$

provided $k^2 > U(r)$.

We have tested the inequality (23) on a number of (numerically obtained) solutions corresponding to various potentials. The right-hand side of Eq. (23) does not overestimate $|u(r)|_m$ much, except if k^2 is close to $U(r_{2i+1})$. Some of these results are given in Table I. The inequality (22) can now be combined with $\mathscr{L}_n$ [Eq. (10)] to give a first order error bound on the approximate phase shift for any potential of the form indicated, at any energy $k^2 > U$.

It has the advantage over the result (11) of Bardsley et al[2]. that no computation of the auxiliary value g_m, is necessary, and there is no restriction, $g_m < 1$, on the strength of the potential.

3. APPLICATION TO SCATTERING BY AN EXPONENTIAL POTENTIAL

We consider the example of s-wave scattering by an expontential potential $V(r) = -2e^{-2r}$ (units: $m = \hbar = 1$), using the trial function

$$u_T(r) = \sin kr + \tan\eta_T \cos kr\,(1 - e^{-\alpha r})$$

$$+ \sum_{i=1}^{N} a_i r^i e^{-pr}, \tag{24}$$

TABLE I. Comparison of actual (numerically obtained) maxima of $|u(r)|$ with the upper bounds (22) in a.u.

$U = re^{-0.5r}$		$\|u(r)\|^2_{\max}$	$k^2A^2/(k^2 - U_{\max})$
$k = 2$		0.275	0.308
$k = 1.5$		0.543	0.669
$k = 1$		2.673	6.023
$U = -\sin(\pi r/2)$,	$r \leqslant 4$		
$= 0$,	$r > 4$		
$k = 2$		0.247	0.329
$k = 1.5$		0.626	0.933
$k = 1.1$		1.85	10.6

TABLE II. First order bounds on the s-wave phase shift for scattering by the potential $V = -2e^{-2r}$(a.u.), N: number of short range functions in u_T Eq. (24).

	$k = 0.02$	$k = 0.5$
η(exact)	0.014740314	0.25537311
$N = 2$		
γ	0.43	0.62
α	1.66	1.93
p	2.2	2.15
η_T	0.014 7497	0.255 365
$\eta_T \pm \sin^{-1}(\mathscr{L}_2/k)$	0.017 2117	0.262 527
	0.012 2878	0.249 337
$\eta_T \pm \sin^{-1}(\mathscr{L}_2/k(1-g_m))$	0.019 6688	0.264 417
	0.009 8308	0.247 403
$N = 6$ (nonlinear parameters optimized only very approximately)		
γ		0.75
α		1.5
p		2.4
η_T		0.255 372 73
$\eta_T \pm \sin^{-1}(\mathscr{L}_2/k)$		0.255 903
		0.254 842
$\eta_T \pm \sin^{-1}[\mathscr{L}_2/k(1-g_m)]$		0.256 366
		0.254 379

which is the same as that considered by Bardsley, Gerjuoy, and Sukumar[2].

The "optimal" values of the parameters η_T, α, a_i, p are determined by minimizing the weaker error bound $\propto \mathscr{L}_2$, with $w(r) = e^{\gamma r}$ and $\gamma > 0$. Though the results could be improved somewhat by using the stronger error bound (proportional to $\mathscr{L}_1$), we shall use the weaker one as it involves the more common variance intergral.[1,2,7,8] Since the effective potential here is monontonically decreasing (with decreasing r), the analysis of Sec. 2 implies that $|u|_m = A(\eta)$; hence it follows from Eq. (7), (9), amd (10) that

$$k\,|B(\eta_T)\sin(\eta_T - \eta)| \leqslant \mathscr{L}_n \tag{25}$$

in contrast to

$$k\,|A(\eta)B(\eta_T)\sin(\eta_T - \eta)| \leqslant \mathscr{L}_n/(1 - g_m), \tag{26}$$

which, if $n = 2$ and $A(\eta) = B(\eta) = \sec\eta$, corresponds to the bound derived by Bardsley, Gerjuoy, and Sukumar.[2]

Table II is a list of the present results obtained by using the inequality (25) (with $B = 1$) together with the corresponding results derived by the method of Bardsley *et al.*[2] [Eq. (26) with $A = B = \sec\eta$]. Note that the numbers quoted by Bardsley *et al.*[2] in their paper contain errors; hence we give here the values obtained by Darewych and Pooran.[3] The present results are basically an improvement by the factor $1/(1 - g_m)$, which is about 2 for the present example. Clearly this will be much more significant for stronger potentials, where g_m approaches 1. Of course, if $g_m > 1$, then (26) does not even give a bound, whereas the present formalism (25) suffers from no such resrictions.

CONCLUDING REMARKS

We have derived a calculable upper bound to the absolute value of the partial wavefunction for the case of potential scattering at incident energies higher than the effective potential. This can be used to derive simple upper bounds to the absolute value of the difference between the (unknown) exact and an approximate (trial) phase shift. The present results are an improvement on those derived previously by Bardsley, Gerjuoy, and Sukumar,[2] and have the advantage that they do not require the calculation of any auxiliary quantities [the g_m, Eq. (12); y_m in Bardsley's notation], and place no restrictions on the strength of the potential.

The generalization of the present results to the case of scattering by compound targets (which is the ultimate aim of such investigations as the present) is by no means obvious and requires further study.

Lastly we should point out that for the present illustrative examples we have used particular forms of asympotic normalization [i.e., the factors $A(\eta)$ and $B(\eta_T)$ were chosen to be either 1 or secη]. Though the choice of the asymptotic normalization is immaterial if an exact computation is done, this is not the case for approximate calculation for which the results can vary with the choice of A and B even if all else is held the same. The question of an optimal choice of asymptotic normalization in the case of approximate calculations also needs further investigation.

[1]T. Kato, Prog. Theor. Phys. (Japan) **6**, 394 (1951).
[2]J.N. Bardsley, E. Gerjuoy, and C.V. Sukumar, Phys. Rev. A **6**, 1813 (1972).
[3]J.W. Darewych and R. Pooran, Can. J. Phys. **56**, 1358 (1978).
[4]I. Shimamura, J. Phys. Soc. Japan **25**, 971 (1968).
[5]L. Hulthén, Kgl. Fysiograf. Sälskap. (Lund) **14**, 257 (1944).
[6]W. Kohn, Phys. Rev. **74**, 1763 (1948).
[7]F.H. Read and J.R. Soto-Montiel, J. Phys. B **6**, L15 (1973).
[8]K.J. Miller, Phys. Rev. A **3**, 607 (1971).

On the static Einstein–Maxwell field equations

A. Das

Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada
(Received 25 July 1978)

The static Einstein–Maxwell field equations are investigated in the presence of *both* electric and magnetic fields. The sources or bodies are assumed to be of finite size and to not affect the connectivity of the associated space. Furthermore, electromagnetic and metric fields are assumed to have reasonable differentiabilities. It is then proved that the electric and magnetic field vectors are constant multiples of one another. Moreover, the static Einstein–Maxwell equations reduce to the static magnetovac case. If, furthermore, the variational derivation of the Einstein–Maxwell equations is assumed, then both the total electric and magnetic charge of each body must vanish. As a physical consequence it is pointed out that if a suspended magnet be electrically charged then it must experience a purely general relativistic torque.

1. INTRODUCTION

The static Einstein–Maxwell field equations have been investigated by many authors.[1] Most of these papers, however, deal with either the static electrovac case or else with the static magnetovac case. Since the second case can always be transformed[2] into the first (the converse is not true), most of these investigations are on the static electrovac universes.

In this paper we consider the static Einstein–Maxwell equations when *both* electric and magnetic fields are present. We can treat the equations either in terms of field intensities or in terms of potentials (the second choice is a must if we assume a variational derivation of the original field equations). The two approaches are not completely equivalent. For example, in the first formulation magnetic monopoles can exist,[3] but such monopoles are not possible for a formulation in terms of potentials (unless the monopole is associated with a wire singularity). In this paper the first approach is pursued, supplemented by conclusions from the second approach. In a static situation the modified magnetic vector $B^{\alpha} = \frac{1}{2}(\epsilon^{\alpha\beta\gamma}/\sqrt{-\gamma})F_{\beta\gamma}$ is easier to use[4] than the magnetic field tensor $F_{\alpha\beta}$. If we assume that the modified magnetic field B^{α} and electric field $F_{\alpha 4}$ have continuous first partial derivatives in a simply connected domain of an associated space, then by six of the static Maxwell equations it follows[5] that two single-valued potentials A, B must exist (unique up to additive constants) such that $F_{\alpha 4} = A_{,\alpha}$ and $B_{\alpha} = B_{,\alpha}$. The question of multiply connected domains (caused, for example, by a toroidal singularity) and consequently multiple-valued potentials A and B will not be discussed here. The remaining Maxwell equations reduce to two "potential equations" each for A and B. Einstein's equation ${}^{(4)}\mathscr{E}^{\alpha}{}_{4} = 0$ implies that "the Poynting vector is zero" so that "the electric and magnetic fields are collinear." This conclusion implies that A and B are functionally related, and, comparing two potential equations, it is proved that this functional relationship is a linear one. Working out the invariant components of the Maxwell tensor in space–time, it is found that $F_{(4\alpha)} = cF_{(\beta\gamma)}$ $(\alpha,\beta,\gamma \neq)$ or in other words "the electric and magnetic fields are constant multiples of each other." The remaining static Einstein–Maxwell equations reduce to a purely magnetovac case.

If we make the further assumption that the original Einstein–Maxwell equations are variationally derived, then $F_{ab} \equiv A_{a,b} - A_{b,a} A_4 \equiv A$. This condition implies by Stokes' theorem that both the total normal electric and magnetic fluxes across any exterior closed surface enclosing a body or source must vanish. Therefore, both the total electric and magnetic charges of each of the sources which generates static electric, magnetic, and gravitational fields (none of them being trivial), according to variationally derived Einstein–Maxwell equations, must vanish. Consequently, the sources like electric and magnetic dipoles are allowed.

These conclusions give rise to various experimental possibilities. For example, if a suspended magnet be electrically charged, it must experience a gravitational torque. The magnitude of this effect is not computed yet. Although it should be quite small, this effect may not be outside the scope of experimental verification.

2. NOTATIONS AND PRELIMINARIES

The metric form of the semi-Riemannian space-time manifold M_4 is denoted by

$$\Phi \equiv \gamma_{ab} dx^a dx^b,$$

where the Latin indices take 1,2,3,4 and the summation convention is followed. The signature of M_4 is assumed to be -2. Physical units are so chosen that $c = G = 1$. Denoting the electromagnetic field tensor by F_{ab} $(= -F_{ba})$, the Einstein–Maxwell field equations can be written as the following[6]:

$$M^a \equiv {}^{(4)}F^{ab}{}_{\|b} = 0, \tag{2.1a}$$

$$M_{abc} \equiv F_{[ab\|c]} = 0, \tag{2.1b}$$

$$\mathscr{E}_{ab} \equiv {}^{(4)}G_{ab} + \kappa\{-{}^{(4)}F_a{}^k F_{bk} + \tfrac{1}{4}\gamma_{ab}\,{}^{(4)}F^{kl}F_{kl}\} = 0. \tag{2.1c}$$

Here the double stroke indicates a covariant derivative, the square bracket denotes cyclic permutation, ${}^{(4)}G_{ab}$ is the Einstein tensor, ${}^{(4)}F^{ab} \equiv \gamma^{ac}\gamma^{bd}F_{cd}$, and $\kappa \equiv 8\pi$.

 0022-2488/79/040740-04$01.00

If, furthermore, the variational derivation of (2.1a)–(2.1c) is postulated, then we must have

$$F_{ab} \equiv A_{a,b} - A_{b,a}, \quad (2.2)$$

where the comma denotes a partial derivative and A_a is the 4-potential.

A static M_4 admits a timelike Killing motion as well as a discrete isometry of time reflection. In the coordinates adapted to the motion, the metric form can be written as

$$\Phi = \gamma_{ab}dx^a dx^b = -e^{-\omega(x)}g_{\alpha\beta}(x)dx^\alpha dx^\beta + e^{\omega(x)}(dx^4)^2, \quad (2.3)$$

where the Greek indices take the values 1,2,3, $(x) \equiv (x^1,x^2,x^3)$, and the metric $g_{\alpha\beta}(x)$ defines an associated (positive definite) Riemannian space M_3. It is further assumed[7] that the electromagnetic field tensor is invariant under the Killing motion, that is,

$$F_{ab,4} = 0. \quad (2.4)$$

Subsequently, the Greek indices will be lowered and raised by $g_{\alpha\beta}$ and $g^{\alpha\beta}$ respectively. The single stroke will indicate a covariant derivative in M_3. In a regular domain D in M_3 where the static Einstein-Maxwell equations hold, the differentiability assumptions are (i) $g_{\alpha\beta}(x) \in C^3(D)$, (ii) $\omega(x) \in C^2(D)$, and (iii) $F_{ab}(x) \in C^2(D)$.

3. REDUCTION OF MAXWELL's EQUATIONS

In the static case the usual solution of the equation $M_{\alpha\beta\gamma} = 0$ in (2.1b) is taken to be

$$F_{\alpha\beta}(x) = A_{\alpha,\beta} - A_{\beta,\alpha}. \quad (3.1)$$

It may be emphasized here that this solution strictly holds only in case $F_{\alpha\beta}(x)$ has continuous first partial derivatives (this condition is implied by our assumption) and the domain of consideration is deformable to a point[8] (that is, the domain cannot enclose sources!). On the other hand, if we assume the variational derivation of (2.1a), (2.1b), (2.1c), then (3.1) holds identically by (2.2).

After explaining the status of (3.1), we shall proceed to make an invertible linear transformation on the three independent components of $F_{\alpha\beta}$ by the equations[4]

$$B^\alpha = \tfrac{1}{2}e^\omega \eta^{\alpha\beta\gamma}F_{\beta\gamma}. \quad (3.2)$$

Here $\eta^{\alpha\beta\gamma} = \epsilon^{\alpha\beta\gamma}/\sqrt{g}$ is the antisymmetric Levi-Civita tensor. It should be noted that $B^\alpha(x)$ in (3.2) is a contravariant vector field in M_3.

If Eq. (3.1) is accepted, then (3.2) implies by application of Stokes' theorem the following integral condition:

$$\oint_{S_2} e^{-\omega}B^\alpha n_\alpha d_2 s = \tfrac{1}{2}\oint_{S_2} \eta^{\alpha\beta\gamma}F_{\beta\gamma}n_\alpha d_2 s = 0. \quad (3.3)$$

Here the surface S_2 is assumed to be piecewise smooth, simply connected, orientable, closed, compact and n^α is the unit outer normal. The surface S_2 lies wholly in vacuum, although it might enclose one or more sources or bodies. (A body is a region where at least one of the field equations does not hold.) The surface integral in (3.3) is the total normal magnetic flux[1] across a closed surface, and the vanishing of it implies the absence of the total magnetic charge of the finite bodies it encloses. Since the choice of exterior closed surface S_2 is arbitrary, the total charge on each finite body must separately vanish. However, for a finite body with at least one semi-infinite tail (or wire or string singularity), the enclosing surface is perforated and the integral condition (3.3) need not hold. Therefore, Dirac-type magnetic monopoles[9] (with wire singularities) are permissible within the framework of Einstein–Maxwell equations even if these are variationally derived.

The vacuum equations $M^\alpha = 0$, $M_{\alpha\beta\gamma} = 0$ in (2.1a), (2.1b) imply respectively that

$$B_\alpha = B_{,\alpha}, \quad (3.4)$$

$$(e^{-\omega|\alpha})_{|\alpha} = 0. \quad (3.5)$$

Equation (3.5) can be interpreted as the magnetic potential equation. In case the twice differentiable potential B exists inside each body the expression $(e^{-\omega}B^{|\alpha})_{|\alpha}$ can be interpreted as the magnetic-charge density. If this magnetic charge density vanishes everywhere (and B attains "zero boundary values"), then, by Hopf's theorem[10] on elliptic partial differential equation and Eq. (3.2), there can be no magnetic field anywhere.

The field equations $M_{\alpha\beta 4} = 0$ in (2.1b) for the static case yield

$$F_{4\alpha,\beta} - F_{4\beta,\alpha} = 0.$$

If the above equations hold in a simply connected domain of M_3 and $F_{4\alpha}$ have continuous partial derivatives, then[5]

$$F_{4\alpha}(x) = A_{,\alpha}, \quad (3.6)$$

where $A(x)$ can be identified with the electrostatic potential. If we assume the domain to be multiply connected (as in the case of a domain enclosing a ring singularity), the function $A(x)$ would be multiple valued. We exclude such cases in this paper. If the Einstein–Maxwell equations are variationally derived, Eq. (3.6) holds identically by (2.2).

The equation $M^4 = 0$ in (2.1a) yields $(e^{-\omega}A^{|\alpha})_{|\alpha} = 0$. So under the assumptions stated previously the static Maxwell equations reduce to

$$(e^{-\omega}A^{|\alpha})_{|\alpha} = 0, \quad (3.7a)$$

$$(e^{-\omega}B^{|\alpha})_{|\alpha} = 0. \quad (3.7b)$$

If, furthermore, the variational derivation of the original equations is assumed, then a necessary integral condition from (3.3) is

$$\oint_{S_2} e^{-\omega}B^\alpha n_\alpha d_2 s = 0. \quad (3.7c)$$

4. REDUCTION OF EINSTEIN'S EQUATIONS

The equations ${}^{(4)}\mathscr{E}^\alpha{}_4 = 0$ in (2.1c) yield ${}^{(4)}F^{\alpha\beta}F_{\beta 4} = 0$, or by (3.1), (3.5), (3.6) we have

$$\eta^{\alpha\beta\gamma}A_{,\beta}B_{,\gamma} = 0, \quad (4.1)$$

The above equation can be recognized as "the Poynting-vector of the electromagnetic field is zero." Four cases arise out of (4.1): case (i) $A_{,\beta} \equiv 0$, $B_{,\gamma} \equiv 0$; case (ii) $A_{,\beta} \not\equiv 0$, $B_{,\gamma} \equiv 0$;

case (iii) $A_{,\beta} \equiv 0$, $B_{,\gamma} \not\equiv 0$; case (iv) $A_{,\beta} \not\equiv 0$, $B_{,\gamma} \not\equiv 0$. Case (i) is that of pure gravity, (ii) is the static electrovac case, (iii) is the static magnetovac case (which can be transformed[2] into the electrovac case). We shall ignore these cases here and concentrate on case (iv), for which (4.1) yields

$$B_{,\gamma} = \lambda(x) A_{,\gamma}, \qquad (4.2)$$

where $\lambda(x) \not\equiv 0$, $\lambda(x) \in C^2(D)$ and is otherwise an arbitrary function. Equation (4.2) means that "the electric and magnetic fields are collinear." Moreover, this equation (4.2) holds iff[5]

$$B = \mathscr{B}(A(x)), \qquad (4.3a)$$

$$B' \equiv \frac{d\mathscr{B}(A(x))}{dA} = \lambda(x). \qquad (4.3b)$$

Substituting (4.3a), (4.3b) into (3.7b) and remembering (3.7a) together with $e^{-\omega} A_{|\alpha} A^{|\alpha} > 0$, we have

$$B'' = 0, \qquad (4.4a)$$

$$B' = c^{-1} = \lambda(x), \qquad (4.4b)$$

$$cB(x) = A(x) + d, \qquad (4.4c)$$

where $c (\neq 0)$ and d are arbitrary constants. The last equation (4.4c) implies a physically significant result. To derive this, we adopt the following positively oriented tetrad $\lambda_{(a)}{}^{b}$ [consistent with the metric (2.2)]:

$$\lambda_{(4)}{}^{b} = e^{-\omega/2} \delta_4{}^{b}, \quad \lambda_{(\alpha)}{}^{\beta} = e^{\omega/2} \Lambda_{(\alpha)}{}^{\beta}, \quad \lambda_{(\alpha)}{}^{4} = 0,$$
$$g_{\beta\gamma} \Lambda_{(\mu)}{}^{\beta} \Lambda_{(\nu)}{}^{\gamma} = \delta_{(\mu\nu)}, \quad \Lambda_{(\alpha)}{}^{\beta} \Lambda_{(\alpha)}{}^{\gamma} = g^{\beta\gamma},$$
$$\sqrt{g}\, \det[\Lambda_{(\alpha)}{}^{\beta}] = 1. \qquad (4.5)$$

The corresponding invariant (or nonholonomic) components of the electromagnetic field tensor become [using (3.1), (3.6), (4.4), (4.5)]

$$F_{(4\alpha)} = F_{ab} \lambda_{(4)}{}^{a} \lambda_{(\alpha)}{}^{b} = F_{4\beta} \Lambda_{(\alpha)}{}^{\beta} = cB_{,\beta} \Lambda_{(\alpha)}{}^{\beta} \equiv cB_{,(\alpha)}, \qquad (4.6a)$$

$$F_{(\mu\nu)} = F_{ab} \lambda_{(\mu)}{}^{a} \lambda_{(\nu)}{}^{b} = \sqrt{g}\, \epsilon_{\alpha\beta\gamma} \Lambda_{(\mu)}{}^{\alpha} \Lambda_{(\nu)}{}^{\beta} B^{|\gamma}$$

$$= \sqrt{g}\, [\epsilon_{\alpha\beta\gamma} \Lambda_{(\mu)}{}^{\alpha} \Lambda_{(\nu)}{}^{\beta} \Lambda_{(\gamma)}{}^{\gamma}] B_{,(\sigma)}$$

$$= \sqrt{g}\, \det[\Lambda_{(\mu)}{}^{\alpha}] B_{,(\sigma)} = B_{,(\sigma)} \quad (\mu,\nu,\sigma \neq). \qquad (4.6b)$$

Equations (4.6a), (4.6b) can be interpreted in terms of the measurable components of the electric and magnetic fields[1] as $E_{(\sigma)} = cH_{(\sigma)}$ or "the electric and magnetic fields are constant multiples of one another." If we demand that original equations are variationally derived, then Eq. (4.4c) has another consequence. The integral condition (3.7c) yields by (4.4c) another integral condition:

$$\oint_{S_2} e^{-\omega} A^{|\alpha} n_\alpha d_2 s = 0, \qquad (4.7)$$

or "the total electric charge on each finite body must vanish."

After this short digression let us go back to Einstein's equations (2.1c) again. We define the algebraically equivalent system $\widetilde{\mathscr{E}}_{ab} \equiv \mathscr{E}_{ab} - \frac{1}{2}\gamma_{ab}{}^{(4)}\mathscr{E}^{k}{}_{k} = 0$. The remaining Einstein's equations are equivalent to

$$\widetilde{\mathscr{E}}_{\alpha\beta} - e^{-2\omega} g_{\alpha\beta} \widetilde{\mathscr{E}}_{44} = R_{\alpha\beta} + \tfrac{1}{2}\omega_{,\alpha}\omega_{,\beta}$$
$$- \kappa e^{-\omega}(A_{,\alpha}A_{,\beta} + B_{,\alpha}B_{,\alpha}) = 0, \qquad (4.8a)$$

$$-2e^{-2\omega} \widetilde{\mathscr{E}}_{44} = \Delta_2 \omega - \kappa e^{-\omega}(\Delta_1 A + \Delta_1 B) = 0, \qquad (4.8b)$$

where $\Delta_2 \omega \equiv (\omega^{|\alpha})_{|\alpha}$, $\Delta_1 A \equiv g^{\alpha\beta} A_{,\alpha} A_{,\beta}$, etc. In the above equations contributions from the electric and magnetic fields are *completely decoupled*.

From Eq. (4.8b) it can be concluded that ω is subharmonic. By Hopf's theorem[10] the regularity of ω throughout M_3 and attainment of zero boundary values implies that ω is a constant which in turn by other field equations imply that the space–time M_4 is flat and $F_{ab} \equiv 0$. Using (4.4c) and making a transformation[4]

$$\widehat{B} = \sqrt{c^2 + 1}\ \mathrm{B}, \qquad (4.9)$$

Eqs. (4.8a), (4.8b), and (3.7b) become

$$R_{\alpha\beta} + \tfrac{1}{2}\omega_{,\alpha}\omega_{,\beta} - \kappa e^{-\omega} \widehat{B}_{,\alpha} \widehat{B}_{,\beta} = 0,$$
$$\Delta_2 \omega - \kappa e^{-\omega} \Delta_1 \widehat{B} = 0, \qquad (4.10)$$
$$\Delta_2 \widehat{B} - \omega_{,\alpha} \widehat{B}^{|\alpha} = 0.$$

The above can be recognized as the static magnetovac equations.

Thus we have proved the following propositon:

Proposition 1: Let D be a bounded, simply connected domain M_3, the associated space of the static universe (2.2). Suppose the boundary of D consists of a finite number of piecewise smooth, orientable, simply connected, closed, compact surfaces. Let the static electromagnetic field tensor ($F_{ab,4} = 0$) beong to the class $C^1(D)$, $\omega \in C^2(D)$ and $g_{\alpha\beta} \in C^3(D)$. Let $(F_{14})^2 + (F_{24})^2 + (F_{34})^2 \not\equiv 0$, $(F_{12})^2 + (F_{23})^2 + (F_{31})^2 \not\equiv 0$ and suppose Einstein–Maxwell equations (2.1a), (2.1b), (2.1c) hold in D. The in D (i) there exist functions $A(x)$, $B(x)$ (unique up to additive constants) such that $F_{4\alpha} = A_{,\alpha}$, $F_{\alpha\beta} = e^{-\omega}\eta_{\alpha\beta\gamma} B^{|\gamma}$; (ii) the space–time invariant components are related by $cF_{(\mu\nu)} = F_{(4\sigma)}$ $(\mu,\nu,\sigma \neq)$, and $c(\neq 0)$ is otherwise an arbitrary constant; (iii) the static Einstein–Maxwell equations reduce to the static magnetovac case. If, furthermore, the original Einstein–Maxwell equations (2.1a), (2.1b), (2.1c) are variationally derived, then over any regular exterior surface S_2 wholly lying in D, the integrals

$$\oint_{S_2} e^{-\omega} A^{|\alpha} n_\alpha d_2 s = \oint_{S_2} e^{-\omega} B^{|\alpha} n_\alpha d_2 s = 0.$$

The physical consequences of the above proposition can now be pursued. If, for example, a suspended bar magnet is charged with electricity, then its electric and magnetic fields will not be a constant multiple of one another. In fact, the Poynting vector field will be generated in coaxial circles, the axis being the magnet itself. This would cause a purely general relativistic torque to act on the magnet. The magnitude of this effect would be very small, and design would need obvious modification due to the earth's magnetism. Nevertheless, this effect may not be completely outside the scope of experimentalists today.

ACKNOWLEDGMENTS

The author acknowledges the support of the NRC Grant No. 3993. He also would like to thank Dr. M.J. Hamilton and Dr. S. Kloster for informal discussions.

[1]J.L. Synge, *Relativity*, The General Theory (North-Holland, Amsterdam, 1964), pp. 354, 363.
[2]W.B. Bonnor, Proc. Phys. Soc. A **67**, 225 (1954).
[3]A.Das, Prog. Theor. Phys. **24**, 915 (1960).
[4]D. Kramer, G. Neugebauer, and H. Stephani, Forschr. Physik **20**, Heft 1, (1972); W. Israel and G.A. Wilson, J. Math. Phys. **13**, 865 (1972); A. Das and S. Kloster, J. Math. Phys. **18**, 586 (1977).
[5]R.C. Buck, *Advanced Calculus* (McGraw-Hill, New York, 1965), p. 431, 292.
[6]Each of Refs. 1–4 has an equivalent formulation.
[7]This condition is physically plausible, but it is not required mathematically. See J.R. Ray and E.L. Thompson, J. Math. Phys. **16**, 345 (1975).
[8]The equivalent topological criterion is all homotopic groups are trivial. With the single condition of simply connectedness of the domain, one may encounter a contradiction. See B. Gelbaum and J. Olmstead, *Counter-Examples in Analysis* (Holden-Day, San Francisco, 1964), p. 126.
[9]P.A.M. Dirac, Phys. Rev. **74**, 817 (1948); A.O. Barut, Acta Phys. Austriaca. Suppl. **11**, 565 (1973).
[10]K. Yano and S. Bochner, *Curvature and Betti Numbers* (Princeton U.P., Princeton, N.J., 1953), pp. 30, 181. The continuous function B attains "boundary values zero" in the following sense: For every $\epsilon > 0$ there exists a compact subset D_ϵ of M_3 such that $|B| < \epsilon$ in $M_3 - D_\epsilon$.

Space–time homogeneous models with torsion[a)]

Marcos Rosenbaum and Michael P. Ryan, Jr.

Centro de Estudios Nucleares, Universidad Nacional Autónoma de México, Circuito Exterior C.U., México 20, D.F. México

L. C. Shepley

Center for Relativity, The University of Texas at Austin, Austin, Texas 78712

(Received 9 September 1978)

A space-time homogeneous model is one in which the metric and the connection are invariant under a four-dimensional transitive Lie group. The Einstein-Cartan theory describes the effect of spin on geometry. The field equations require a stress-energy tensor and a spin tensor for a perfect fluid, and we give the required forms here. The main part of this paper is a presentation of several classes of space-time homogeneous solutions of the Einstein-Cartan equations for perfect fluid cosmological models.

I. INTRODUCTION

Several papers have been written on spatially homogeneous cosmological models in the Einstein–Cartan theory,[1,2] which incorporates the effect of spin on geometry. Here we treat the class of space–time homogeneous models, since they are important as explicit exact solutions in this theory. This article derives two important classes of space–time homogeneous models with torsion, which we call the "fluid compatible" models and the "observer-stationary" models. We also include a discussion of the proper constitutive equations to be used in Einstein–Cartan cosmology. The models cannot represent the observed cosmic expansion. Nonetheless certain space–time homogeneous models, in particular the Gödel model,[3] are quite useful physically in general relativity. They provide arenas for the investigation of rotation, Mach's principle, the breakdown of causality, and other global cosmological properties.

The Einstein–Cartan theory uses the metric of space–time

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu.$$

By space–time homogeneous[4] we mean that our models are each invariant under a four-dimensional transitive group of isometries. The metric can be expressed in an orthonormal basis of invariant forms (for convenience we use greek indices to denote components in this basis, also):

$$ds^2 = \eta_{\mu\nu}\omega^\mu\omega^\nu, \quad (\eta_{\mu\nu}) = \mathrm{diag}(-1,1,1,1), \tag{1}$$

$$d\omega^\mu = \tfrac{1}{2} C^\mu{}_{\alpha\beta}\,\omega^\alpha \wedge \omega^\beta, \tag{2}$$

where the $C^\mu{}_{\alpha\beta}$ are the structure constants of the group algebra. They obey the following algebraic properties:

$$C^\mu{}_{\alpha\beta} = -C^\mu{}_{\beta\alpha}, \tag{3}$$

$$C^\sigma{}_{\alpha\beta} C^\mu{}_{\gamma\sigma} + C^\sigma{}_{\beta\gamma} C^\mu{}_{\alpha\sigma} + C^\sigma{}_{\gamma\alpha} C^\mu{}_{\beta\sigma} = 0. \tag{4}$$

This last equation is the Jacobi identities.

Cartan's modification of Einstein's equations allows a nonzero torsion tensor, which in the $\{\omega^\mu\}$ basis has components $T^\mu{}_{\alpha\beta}$ which obey

$$T^\mu{}_{\alpha\beta} = -T^\mu{}_{\beta\alpha}.$$

The connection coefficients $\Gamma^\mu{}_{\alpha\beta}$ are uniquely determined by the first Cartan equation[5]

$$d\omega^\mu = -\Gamma^\mu{}_{\sigma\alpha}\omega^\alpha \wedge \omega^\sigma + \tfrac{1}{2} T^\mu{}_{\alpha\beta}\omega^\alpha \wedge \omega^\beta$$

and the requirement that the metric tensor be covariant constant.

The torsion is generated by the spin density of matter in the Einstein–Cartan theory. The field equation is algebraic, and consequently the torsion vanishes outside of matter. In our models we will be assuming that the matter has the form of a perfect fluid, invariant, as is the torsion, under the group of isometries. Therefore, the torsion tensor components and the matter variables are constant in the $\{\omega^\mu\}$ basis.

The Einstein–Cartan field equations also include the equation

$$R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu} = T_{\mu\nu},$$

where $R_{\mu\nu}$ is the Ricci tensor, not necessarily symmetric. These equations also are algebraic equations due to our assumption of space–time homogeneity.

Our paper is organized as follows: Section II is a discussion of the field equations and the forms needed for the stress energy tensor and spin tensor of a fluid. Our forms are based on the work of Hojman[6] on equations of motion of spinning particles and differ from the forms used by some others.[7] Section III contains two interesting simple models: One has a flat metric, the second has vanishing connection and therefore vanishing Riemann tensor. Section IV is a catalog of all fluid compatible models, where by fluid compatible we mean that the antisymmetric part of the Ricci tensor vanishes. This compatibility condition guarantees that the effect of torsion mimics that of a perfect fluid with energy density equal to pressure. Section V is a discussion of observer-stationary models, in which the Lie derivatives of metric and torsion with respect to the fluid velocity are zero. A summary and conclusion are given in Sec. VI.

[a)]Supported in part by International Scientific Exchange program grants, National Science Foundation OIP75-09783 A01 and Consejo Nacional de Ciencia y Tecnologia No. 955.

II. FLUIDS WITH SPIN

First of all, the stress-energy tensor of a perfect fluid without spin density in general relativity has a well-known form.[8] We let ρ be the energy density and p be the pressure, assumed isotropic in the rest frame of the fluid. The fluid 4-velocity is denoted u^μ and is normalized: $u^\mu u_\mu = -1$. The stress-energy tensor (as used in special or general relativity) for a spinless fluid is denoted $\bar{T}^{\mu\nu}$ to avoid confusion with the stress-energy tensor used with fluids with spin,

$$\bar{T}^{\mu\nu} = \rho u^\mu u^\nu + p(u^\mu u^\nu + g^{\mu\nu}).$$

This tensor is symmetric.

When the fluid has spin density, the proper form to be used for the stress-energy tensor is not so obvious. At the outset we may postulate either that the spin is randomized or that it is aligned.[9] We choose to work with the latter case; thus each fluid element is viewed as having a spin which varies continuously from one fluid element to another. In addition to quantities defined above, we define π^μ to be the density of linear momentum; the spin tensor will be defined below. One effect of spin is that π^μ and u^μ need not be parallel. Thus we have

$$\pi^\mu = \rho u^\mu + f^\mu, \tag{5}$$

where

$$\rho = -\pi^\mu u_\mu \quad \text{so that } f^\mu u_\mu = 0. \tag{6}$$

The Weysenhoff model[10] for the stress-energy tensor provides a general format for this case and is needed in both special and general relativity. The stress-energy tensor, denoted by $T^{\mu\nu}$, need not be symmetric; it has the form

$$T^{\mu\nu} = u^\mu \pi^\nu + p(u^\mu u^\nu + g^{\mu\nu}). \tag{7}$$

In this paper we exclude as unphysical models with negative ρ or P. We do not, however, exclude models with $p \geqslant \rho$.

We call the spin flux tensor $\tau^\alpha{}_{\beta\gamma}$, antisymmetric in its lower indices: $\tau^\alpha{}_{\beta\gamma} = -\tau^\alpha{}_{\gamma\beta}$. In the Weysenhoff model and in our models the spin tnesor in the rest frame of the fluid is denoted by $S_{\beta\gamma}$, and the spin flux tensor has the form

$$\tau^\alpha{}_{\beta\gamma} = u^\alpha S_{\beta\gamma} \quad \text{with } S_{\beta\gamma} = -S_{\gamma\beta}. \tag{8}$$

The tensor $S_{\beta\gamma}$ is itself restricted by the condition[11]

$$\pi^\alpha S_{\alpha\gamma} = 0 \tag{9}$$

(the original Weysenhoff model used the condition $u^\alpha S_{\alpha\gamma} = 0$).

In justifying Eqs. (8) and (9) we should remember that in a fluid with spin there are two preferred frames. The first, which we call the velocity-rest frame, is one in which u^μ has no spatial components. The second, which we call the momentum-rest frame, is one in which π^μ has no spatial components. (Clearly these frames coincide if f^μ vanishes.) In the expression for $T_{\mu\nu}$, the pressure (that is the stress part), is taken as isotropic in the velocity-rest frame. Equation (8) means the velocity-rest frame is a preferred frame for the spin flux tensor. This choice is based on the equation of motion for spinning particles, which has been derived in various ways using both Lagrangian and non-Lagrangian techniques.[6,12] This equation may be compared to the integrability equation of the Einstein–Cartan field theory which corresponds to the twice-contracted Bianchi identity used in general relativity,

$$(R^\sigma{}_\alpha - \tfrac{1}{2}R\delta^\sigma{}_\alpha)_{;\sigma} = -R^\sigma{}_\tau T^\tau{}_{\sigma\alpha} + R^{\sigma\tau}{}_{\rho\alpha} T^\rho{}_{\sigma\tau} \tag{10}$$

(the tensors in this equation will be defined below). The two equations of motion are the same if Eq. (8) holds.

The choice of Eq. (9) is not so straightforward; indeed others have used the original Weysenhoff condition.[7] Our choice is strongly suggested by reference to the equations for the time development of the spin of a particle.[6,12] In this case a Lagrangian technique is used to give the apropriate equations, but these may be integrated only if some functional restriction among $S_{\beta\gamma}, \pi^\mu, u^\mu$ is assumed. The relation $u^\mu S_{\mu\alpha} = 0$ gives unphysical results, namely motion in a circle in the particle momentum-rest frame. The choice of Eq. (9) allows the equations of motion to be integrated without these strange motions, and we therefore adopt that restriction.

We do not mean the above remarks to be a definitive discussion. In particular we are adopting forms for fluid tensors based on the equations of motion for particles. The derivation of fluid properties from the kinetics of particles deserves a fuller discussion, and a future paper will deal with this subject.

The field equations in the Einstein–Cartan theory use connection coefficients $\Gamma^\alpha{}_{\beta\gamma}$ which are derived from the torsion tensor $T^\alpha{}_{\beta\gamma}$ and from the condition that the metric $g_{\mu\nu}$ has vanishing covariant derivative. The Ricci tensor $R_{\mu\nu}$ and Einstein tensor $G_{\alpha\beta}$ are derived from the Riemann tensor $R^\mu{}_{\nu\alpha\beta}$,

$$G_{\alpha\beta} = R_{\alpha\beta} - \tfrac{1}{2}Rg_{\alpha\beta} = R^\sigma{}_{\alpha\sigma\beta} - \tfrac{1}{2}g^{\mu\nu}R^\sigma{}_{\mu\sigma\nu}g_{\alpha\beta}.$$

The Einstein tensor is not necessarily symmetric but its antisymmetric part appears in an integrability condition. The various forms for the field equations are based on this fact.

One convenient form for the field equations[13] is

$$G_{\mu\nu} = T_{\mu\nu}, \tag{11}$$

$$T^\alpha{}_{\beta\gamma} = \tau^\alpha{}_{\beta\gamma} + \delta^\alpha{}_{[\beta}\tau^\sigma{}_{\gamma]\sigma} = u^\alpha S_{\beta\gamma} + \delta^\alpha{}_{[\beta}s_{\gamma]}, \quad \text{with } s_\gamma = S_{\gamma\sigma}u^\sigma,$$

where the coupling constant is set equal to one. The second equation is the (algebraic) equation relating torsion and spin.

In an invariant orthonormal basis, the metric is $g_{\mu\nu} = \eta_{\mu\nu}$ as given in Eq. (1). The structure coefficients $C^\mu{}_{\sigma\tau}$ are constant because of space–time homogeneity. They are given by Eq. (2) and obey Eqs. (3). The Jacobi identities (4) which come from the requirement $dd\omega^\mu = 0$ are the integrability conditions which allow ω^μ to be expressed in terms of coordinate functions and the associated coordinate basis of forms (we will derive some but not all the coordinate expressions).

The connection coefficients are determined from the $C^\mu{}_{\sigma\tau}$ and the constant torsion components $T^\mu{}_{\alpha\beta}$, by use of the first Cartan equation. The expression for $\Gamma^\alpha{}_{\beta\gamma}$ is unique if the connection is metric compatible; here metric compatibility means

$$\Gamma_{\alpha\beta\sigma}=-\Gamma_{\beta\alpha\sigma} \qquad (12)$$

where indices are lowered (and raised) using $\eta_{\alpha\beta}$. The connection coefficients are

$$\Gamma^{\mu}{}_{\alpha\beta}={}^{c}\Gamma^{\mu}{}_{\alpha\beta}-{}^{t}\Gamma^{\mu}{}_{\alpha\beta}, \qquad (13)$$

where

$${}^{c}\Gamma^{\mu}{}_{\alpha\beta}=\tfrac{1}{2}(C^{\mu}{}_{\alpha\beta}-C_{\alpha}{}^{\mu}{}_{\beta}-C_{\beta}{}^{\mu}{}_{\alpha}), \qquad (14)$$

$${}^{t}\Gamma^{\mu}{}_{\alpha\beta}=\tfrac{1}{2}(T^{\mu}{}_{\alpha\beta}-T_{\alpha}{}^{\mu}{}_{\beta}-T_{\beta}{}^{\mu}{}_{\alpha}), \qquad (15)$$

This latter expression, in the case of a fluid torsion tensor, becomes

$${}^{t}\Gamma^{\mu}{}_{\alpha\beta}=\tfrac{1}{2}(u^{\mu}S_{\alpha\beta}-u_{\alpha}S^{\mu}{}_{\beta}-u_{\beta}S^{\mu}{}_{\alpha}+s^{\mu}\eta_{\alpha\beta}-s_{\alpha}\delta^{\mu}{}_{\beta}). \qquad (16)$$

In a space–time homogeneous model u^{μ} and $S_{\alpha\beta}$ are constant tensors.

The Riemann tensor components $R^{\mu}{}_{\nu\alpha\beta}$ are given by the second Cartan equation,

$$\tfrac{1}{2}R^{\mu}{}_{\nu\alpha\beta}\omega^{\alpha}\wedge\omega^{\beta}=d(\Gamma^{\mu}{}_{\nu\sigma}\omega^{\sigma})+\Gamma^{\mu}{}_{\sigma\alpha}\Gamma^{\sigma}{}_{\nu\beta}\omega^{\alpha}\wedge\omega^{\beta}.$$

The Ricci tensor is given by

$$R_{\alpha\beta}=R^{\sigma}{}_{\alpha\sigma\beta}={}^{c}R_{\alpha\beta}+{}^{t}R_{\alpha\beta}+A_{\alpha\beta}, \qquad (17)$$

where, in the case of a fluid torsion tensor, we have

$${}^{c}R_{\alpha\beta}=-\tfrac{1}{2}C^{\tau}{}_{\sigma\beta}(C_{\tau}{}^{\sigma}{}_{\alpha}+C^{\sigma}{}_{\tau\alpha})$$

$$-\tfrac{1}{2}C^{\tau}{}_{\sigma\tau}(C_{\alpha}{}^{\sigma}{}_{\beta}+C_{\beta}{}^{\sigma}{}_{\alpha})+\tfrac{1}{4}C_{\alpha\sigma\tau}C_{\beta}{}^{\sigma\tau}, \qquad (18)$$

$${}^{t}R_{\alpha\beta}=\tfrac{1}{4}u_{\alpha}u_{\beta}S_{\sigma\tau}S^{\sigma\tau}, \qquad (19)$$

$$A_{\alpha\beta}=-\tfrac{1}{2}C^{\tau}{}_{\sigma\tau}s^{\sigma}\eta_{\alpha\beta}+C^{\tau}{}_{\sigma\tau}S^{\sigma}{}_{(\alpha}u_{\beta)}-\tfrac{1}{2}u_{(\alpha}C_{\beta)\sigma\tau}S^{\sigma\tau}$$

$$-\tfrac{1}{2}C^{\tau}{}_{\sigma\tau}u^{\sigma}S_{\alpha\beta}-\tfrac{1}{2}u^{\tau}S^{\sigma}{}_{\beta}C_{\alpha\sigma\tau}$$

$$+\tfrac{1}{2}(u^{\tau}S^{\sigma}{}_{\alpha}+u^{\sigma}S^{\tau}{}_{\alpha})C_{\sigma\tau\beta}. \qquad (20)$$

We have already made use of part of the Einstein–Cartan field equations, namely the algebraic relation between torsion and spin. The other field equation is $G_{\mu\nu}=T_{\mu\nu}$, or in a somewhat more useful form, with the fluid stress-energy tensor,

$$R_{\mu\nu}=(\rho+p)u_{\mu}u_{\nu}+\tfrac{1}{2}(\rho-p)\eta_{\mu\nu}+u_{\mu}f_{\nu}. \qquad (21)$$

Of course neither $R_{\mu\nu}$ nor $T_{\mu\nu}$ is necessarily symmetric. The antisymmetric part of the above field equation is actually an integrability condition. The curl of the first Cartan equation yields $R^{\mu}{}_{[\alpha\beta\gamma]}=T^{\mu}{}_{[\alpha\beta;\gamma]}+T^{\mu}{}_{\tau[\alpha}T^{\tau}{}_{\beta\gamma]}$. On contraction this equation gives the antisymmetric part of $R_{\beta\gamma}$. We will treat the $R_{[\beta\gamma]}$ equation as a field equation since it permits us to find f_{μ}. Notice that only $A_{\mu\nu}$ has an antisymmetric part in the expression for $R_{\mu\nu}$: $R_{[\alpha\beta]}=A_{[\alpha\beta]}$. Therefore, we have

$$f_{\mu}=2u^{\beta}R_{[\mu\beta]}$$

$$=-C^{\tau}{}_{\sigma\tau}u^{\sigma}s_{\mu}-\tfrac{1}{2}u^{\tau}s^{\sigma}C_{\mu\sigma\tau}+u^{\tau}S^{\sigma}{}_{\mu}u^{\beta}C_{\beta\sigma\tau}$$

$$-\tfrac{1}{2}(u_{\tau}s_{\sigma}+u_{\sigma}s_{\tau})C^{\sigma\tau}{}_{\mu}. \qquad (22)$$

We also have the condition

$$R_{[\mu\nu]}(u^{\mu}u_{\alpha}+\delta^{\mu}{}_{\alpha})(u^{\nu}u_{\beta}+\delta^{\nu}{}_{\beta})=0. \qquad (23)$$

III. SIMPLE MODELS

As an illustration of the above general formalism we here consider two simple cases, a model with a flat metric and one with zero connection forms.

The flat metric case is where the structure constants all vanish,

$$C^{\mu}{}_{\alpha\beta}=0. \qquad (24)$$

The metric may then be written

$$ds^2=\eta_{\mu\nu}dx^{\mu}dx^{\nu}. \qquad (25)$$

The Riemann curvature tensor derived from the Christoffel symbols vanishes, but because of torsion the connection coefficients need not be zero. Thus this model has a nonzero curvature.

First, we notice that $R_{\mu\nu}$ is symmetric because $A_{\mu\nu}$ vanishes. Therefore, we have $f_{\mu}=0$. In this case $T^{\mu}{}_{\alpha\beta}=u^{\mu}S_{\alpha\beta}$ with $s_{\alpha}=u^{\beta}S_{\alpha\beta}=0$. The Ricci tensor then becomes

$$R_{\mu\nu}={}^{t}R_{\mu\nu}=\tfrac{1}{4}u_{\mu}u_{\nu}S_{\sigma\tau}S^{\sigma\tau}. \qquad (26)$$

Here we interpret the result as a "cosmological" model filled with a perfect fluid having symmetric stress-energy tensor

$$T_{\mu\nu}=(p+\rho)u_{\mu}u_{\nu}+pg_{\mu\nu}.$$

The pressure p and density ρ are equal

$$\rho=p=\tfrac{1}{8}S_{\sigma\tau}S^{\sigma\tau}. \qquad (27)$$

This model may be extended to a more general situation with flat metric but nonconstant torsion and perfect fluid content.[14]

A canonical form for this model is found by making use of the freedom to make (constant) Lorentz coordinate transformations without altering the metric. We put u^{μ} into the form $u^{\mu}=(1,0,0,0)$ by such a transformation. In a similar fashion we can take the canonical form $S_{12}=-S_{21}=S$, rest of $S_{\mu\nu}=0$. The density and pressure then have the value

$$\rho=p=\tfrac{1}{4}S^2. \qquad (28)$$

Our other simple case is when the connection coefficients vanish (such a geometry is called a teleparallelism geometry) but $C^{\alpha}{}_{\beta\gamma}$ and $T^{\alpha}{}_{\beta\gamma}$ need not vanish. In this model the Riemann tensor is zero and therefore $\rho=p=0$, and also $f_{\mu}=0$. We accept, conventionally, that $u_{\mu}\neq 0$ and $S_{\alpha\beta}\neq 0$, but now we have $s_{\alpha}=u^{\beta}S_{\alpha\beta}=0$. We presume that $T^{\mu}{}_{\alpha\beta}\neq 0$ for the sake of developing a model. Again we can take a canonical form

$$u^{\mu}=(1,0,0,0),$$

$$S_{12}=-S_{21}=S, \quad \text{rest of } S_{\mu\nu}=0. \qquad (29)$$

However, this canonical form is in a basis of 1-forms $\{\omega^\mu\}$, not in a basis directly derived from coordinates.

When $\Gamma^\alpha{}_{\beta\gamma}=0$, we have

$$C^\alpha{}_{\beta\gamma}=T^\alpha{}_{\beta\gamma}=u^\alpha S_{\beta\gamma}. \tag{30}$$

It is easily verified that these structure constants do satisfy the Jacobi identities $C^\mu{}_{\sigma[\alpha}C^\sigma{}_{\beta\gamma]}=0$. Coordinates x^μ may be found to express the basis forms ω^μ. Using the canonical forms of u^μ and $S_{\alpha\beta}$, we find that the ω^μ are:

$$\omega^0=dx^0+x^1dx^2, \qquad \omega^i=dx^i \quad (i=1,2,3). \tag{31}$$

IV. FLUID COMPATIBLE MODELS

Previously we defined various parts of the Ricci tensor, namely ${}^cR_{\alpha\beta}$, ${}^tR_{\alpha\beta}$, $A_{\alpha\beta}$ where ${}^tR_{\alpha\beta}$ has the form associated with a perfect fluid (with $\rho=p$). Here we treat those models in which $A_{\alpha\beta}$ vanishes. We call such a model "fluid compatible" since the effect of torsion is to add the additional fluid-like term ${}^tR_{\alpha\beta}$ to the stress-energy tensor associated with the Ricci tensor ${}^cR_{\alpha\beta}$ of a space–time homogeneous general relativity model. In the classification that follows we exclude as being unphysical models with negative ρ or p.

The vanishing of $A_{\alpha\beta}$ implies that $R_{[\alpha\beta]}$ vanishes. Hence we have

$$f_\mu=0, \quad \text{or } T^\mu{}_{\alpha\beta}=u^\mu S_{\alpha\beta} \quad \text{with } s_\alpha=S_{\alpha\beta}u^\beta=0. \tag{32}$$

Since u^μ and $S_{\alpha\beta}$ are constant tensors, we may use the freedom of Lorentz transformations of the basis $\{\omega^\mu\}$ to put the torsion into a canonical form as we did in the previous section. We thus set

$$u^\mu=(1,0,0,0), \qquad S_{12}=-S_{21}=S, \quad \text{rest of } S_{\alpha\beta}=0. \tag{33}$$

Then ${}^tR_{\alpha\beta}$ has the form ${}^tR_{\alpha\beta}=\frac{1}{2}u_\alpha u_\beta S^2$ which is the form associated with a perfect fluid having energy density = pressure $=\frac{1}{4}S^2$. We will assume $S\neq0$. We have not yet used up all freedom of basis transformations since a rotation in the 1-2 "plane" (that is, a rotation in the ω^1-ω^2 basis forms) preserves the forms of the metric, velocity, and spin. We will use this freedom to simplify some of the equations below.

The equation $A_{\alpha\beta}=0$ is a set of linear equations in the structure constants. After solving these equations we use the freedom of rotation among ω^1-ω^2 to set one of the remaining structure constants to zero. The solutions are conveniently represented by Table I. The six columns are labeled by the lower, antisymmetric indices on $C^\mu{}_{\alpha\beta}$. An entry indicates an equality. Thus the entry "$C^2{}_{03}$" in the upper right corner signifies the equality $C^0_{23}=-C^0_{32}=C^2_{03}$. The entry "0" indicates that a structure constant vanishes. No entry indicates that the structure constant has not been determined—it is a free parameter to be determined, possibly, by additional considerations. Thus there is no entry in the $C^2{}_{03}$ position of the table. If later we find that $C^2{}_{03}$ vanishes, a "0" will have to be put there and also in the $C^0{}_{23}$ position.

TABLE I.

$C^\mu{}_{\alpha\beta}$	01	02	03	12	13	23
0	0	0		0	$C^1{}_{03}$	$C^2{}_{03}$
1	$-\frac{1}{2}C^3{}_{03}$	0				
2	$-C^1{}_{02}$	$-\frac{1}{2}C^3{}_{03}$				
3	0	0		0	0	0

The eleven blank spots in the table correspond to parameters from which all structure constants are determined. These eleven parameters must obey the Jacobi identities and also the field equations

$${}^cR_{\mu\nu}=(\rho+p-\tfrac{1}{2}S^2)u_\mu u_\nu+\tfrac{1}{2}(\rho-p)\eta_{\mu\nu},$$

and we can now list explicit models.

We call the first class of models Type A. It is characterized by $C^3{}_{03}\neq0$. The Jacobi equations imply that several structure constants must vanish. See Table II. There are thus two free parameters left after solving the Jacobi relations, $C^2{}_{23}$ and $C^3{}_{03}$, the latter being assumed nonzero. We now turn to ${}^cR_{\alpha\beta}$ and the field equations. The equation ${}^cR_{03}=0$ implies $C^2{}_{23}=0$. The other field equations allow $C^3{}_{03}$ to serve as a parameter, with the other nonzero structure constants $C^1{}_{01}$ and $C^2{}_{02}$ being determined by it. The perfect fluid content has

$$\rho=p=-\tfrac{3}{4}(C^3{}_{03})^2+\tfrac{1}{4}S^2. \tag{34}$$

Note that a positive value for ρ requires a nonzero spin density. A coordinate system $\{x^\mu\}$ for this model is readily found. The orthonormal basis forms are:

$$\begin{aligned}\omega^0&=dx,\\ w^1&=e^{-Ax^0}dx^1,\\ w^2&=e^{-Ax^0}dx^2,\\ w^3&=e^{2Ax^0}dx^3, \quad \text{where } A=\tfrac{1}{2}C^3{}_{03}.\end{aligned} \tag{35}$$

TABLE II.

Type A $C^\mu{}_{\alpha\beta}$	01	02	03	12	13	23
0	0	0	$-2C^2{}_{23}$	0	0	0
1	$-\frac{1}{2}C^3{}_{03}$	0	0	0	$C^2{}_{23}$	0
2	0	$-\frac{1}{2}C^3{}_{03}$	0	0	0	
3	0	0		0	0	0

TABLE III.

Type B $C^\mu{}_{\alpha\beta}$	01	02	03	12	13	23
0	0	0		0	$C^1{}_{03}$	0
1	0	0		0		
2	0	0	0	0		
3	0	0	0	0	0	0

TABLE IV.

Type B1 $C^{\mu}{}_{\alpha\beta}$	01	02	03	12	13	23
0	0	0	$C^{2}{}_{23}$	0	$C^{1}{}_{03}$	0
1	0	0		0	$C^{2}{}_{23}$	$-C^{2}{}_{13}$
2	0	0	0	0		
3	0	0	0	0	0	0

TABLE VI.

Type C $C^{\mu}{}_{\alpha\beta}$	01	02	03	12	13	23
0	0	0		0	0	0
1	0	0	0			
2	0	0	0	0		
3	0	0	0	0	0	0

We now assume $C^{3}{}_{03}=0$. Our second category, Type B, assumes that $C^{1}{}_{03}$ or $C^{2}{}_{03}\neq 0$. The Jacobi identities show that $C^{1}{}_{12}=C^{2}{}_{12}=0$ and that then all Jacobi equations are satisfied. Further, enough structure constants now vanish that the freedom of rotation in the ω^1-ω^2 "plane" can again be used to give the canonical form $C^{2}{}_{03}=0$, so that we will take $C^{1}{}_{03}\neq 0$. We now have Table III. The field equations ${}^cR_{01}={}^cR_{02}=0$ imply $C^{0}{}_{03}=C^{1}{}_{13}$ and $C^{1}{}_{23}=-C^{2}{}_{13}$. The equations ${}^cR_{11}={}^cR_{22}$ and ${}^cR_{12}=0$ then allow us to define two subclasses: Type B1 with $C^{1}{}_{13}=C^{2}{}_{23}$ and type B2 with $C^{1}{}_{23}=0$ and $C^{2}{}_{23}=-2C^{0}{}_{03}$. We examine each in turn.

First Type B1 (see Table IV). There are two field equations left to satisfy: The ${}^cR_{00}$ and ${}^cR_{33}$ equations yield values for ρ and p:

$$\rho=\tfrac{1}{2}({}^cR_{00}+3{}^cR_{33})+\tfrac{1}{4}S^2=-3(C^{2}{}_{23})^2+\tfrac{1}{4}S^2,$$
$$p=\rho-2{}^cR_{33}=3(C^{2}{}_{23})^2+\tfrac{1}{4}S^2. \tag{36}$$

The structure constants $C^{1}{}_{03}(=C^{0}{}_{13})$ and $C^{2}{}_{13}(=-C^{1}{}_{23})$ are arbitrary.

The second possibility under B is Type B2, with $C^{1}{}_{23}=0$, $C^{2}{}_{23}=-2C^{0}{}_{03}$: the ${}^cR_{11}={}^cR_{33}$ field equation implies $C^{0}{}_{03}=0$, and we have Table V. The other two equations yield the flat metric result $\rho=p=\tfrac{1}{4}S^2$.

Finally we return to the other possibility under the assumption that $C^{3}{}_{03}=0$. This is Type C: Assume $C^{1}{}_{03}=C^{2}{}_{03}=0$. Once again we reacquire the freedom of making a rotation in the ω^1-ω^2 "plane." Such a rotation can be used to set $C^{2}{}_{12}=0$ as a canonical form, and we have Table VI. Two Jacobi relations are at this point still to be satisfied. They require either $C^{1}{}_{12}=0$ (Type C1) or $C^{2}{}_{13}=C^{2}{}_{23}=0$ (Type C2), which we examine in turn.

First, Type C1: In this case the ω^1-ω^2 rotation freedom can be used again to set $C^{1}{}_{23}=-C^{2}{}_{13}$. See Table VII. If $C^{1}{}_{13}\neq C^{2}{}_{23}$, the field equations ${}^cR_{11}={}^cR_{22}={}^cR_{33}$ imply $C^{1}{}_{13}=C^{2}{}_{23}=C^{0}{}_{03}=0$ and we have the flat metric result $\rho=p=\tfrac{1}{4}S^2$, with $C^{2}{}_{13}$ arbitrary. If $C^{1}{}_{13}=C^{2}{}_{23}$, we have $C^{2}{}_{13}=0$ from the ${}^cR_{12}$ equation, and the ${}^cR_{11}={}^cR_{22}={}^cR_{33}$ equations yield that $C^{0}{}_{03}=0$ or that $C^{0}{}_{03}=C^{1}{}_{13}$. In the latter case we have

$$\rho=-\tfrac{3}{2}(C^{0}{}_{03})^2+\tfrac{1}{4}S^2$$
$$p=\tfrac{9}{2}(C^{0}{}_{03})^2+\tfrac{1}{4}S^2. \tag{37}$$

The last of the classes in this section is Type C2, in which $C^{2}{}_{13}=C^{2}{}_{23}=0$. See Table VIII. The field equation ${}^cR_{11}={}^cR_{22}$ implies $C^{1}{}_{13}=0$ and this result causes ${}^cR_{12}={}^cR_{23}=0$. The field equation ${}^cR_{11}={}^cR_{33}$ implies $C^{0}{}_{03}=\pm C^{1}{}_{12}$. The other field equations give results for ρ and p:

$$\rho=-(C^{0}{}_{03})^2-\tfrac{3}{4}(C^{1}{}_{23})^2+\tfrac{1}{4}S^2,$$
$$p=(C^{0}{}_{03})^2+\tfrac{1}{4}(C^{1}{}_{23})^2+\tfrac{1}{4}S^2. \tag{38}$$

To summarize, we have found all models which are fluid compatible. The Type A models have $\rho=p$. The Type B models all have $p\geqslant\rho$. The Type C models also all have $p\geqslant\rho$. In all cases a positive value of the density ρ requires a strictly nonzero value for S.

V. OBSERVER-STATIONARY MODELS

Here we will discuss the effect of the requirements that the fluid velocity u^μ be a Killing vector and that the Lie derivative of the torsion with respect to u^μ also vanish. The corresponding space–time homogeneous general relativity models (zero torsion) are the Einstein universe and the Gödel model. The requirements mean that the direction followed by an observer moving with the cosmic fluid is a direction along which both metric and torsion are unchanging. We therefore call these models "observer-stationary." As in Sec. IV those models with ρ or p negative are excluded as being unphysical.

TABLE V.

Type B2 $C^{\mu}{}_{\alpha\beta}$	01	02	03	12	13	23
0	0	0	0	0	$C^{1}{}_{03}$	0
1	0	0		0	$C^{0}{}_{03}$	0
2	0	0	0	0	0	$-2C^{0}{}_{03}$
3	0	0	0	0	0	0

TABLE VII.

Type C1 $C^{\mu}{}_{\alpha\beta}$	01	02	03	12	13	23
0	0	0		0	0	0
1	0	0	0	0		$-C^{2}{}_{13}$
2	0	0	0	0		
3	0	0	0	0	0	0

TABLE VIII.

Type C2 $C^{\mu}{}_{\alpha\beta}$	01	02	03	12	13	23
0	0	0		0	0	0
1	0	0	0			
2	0	0	0	0	0	0
3	0	0	0	0	0	0

TABLE X.

Type OSA	$C^{\alpha}{}_{\beta\gamma}$	01	02	03	12	13	23
	0	0	0	0		0	0
	1	0		0	0	0	
	2	$-C^{1}{}_{02}$	0	0	0	$-C^{1}{}_{23}$	0
	3	0	0	0	0	0	0

In a coordinate basis, the requirement that u^{μ} be a Killing vector is

$$(\mathscr{L}_u g)_{\alpha\beta} = g_{\alpha\beta,\mu}u^{\mu} + g_{\alpha\beta}u^{\sigma}{}_{,\alpha} + g_{\alpha\sigma}u^{\sigma}{}_{,\beta} = 0.$$

Here the commas mean partial derivative, but the equation retains its form if instead the derivatives are covariant using the symmetric (metric) connection ${}^{c}\Gamma^{\alpha}{}_{\beta\gamma}$. We use a short vertical bar to denote this type of derivative, and the Killing equation becomes

$$u_{\alpha|\beta}+u_{\beta|\alpha}=0. \tag{39}$$

This equation is the same in the basis $\{\omega^{\mu}\}$, where we raise and lower indices using $g_{\mu\nu}=\eta_{\mu\nu}$.

As in the previous section, we use the freedom of Lorentz transformation to put u^{μ} into the canonical form

$$u^{\mu}=(1,0,0,0).$$

Killing's equation is a restriction on the structure constants, and it requires

$$C^{0}{}_{0i}=0, \quad C^{i}{}_{0j}=-C^{j}{}_{0i} \quad (i,j=1,2,3).$$

In a coordinate basis the equation for the vanishing of the Lie derivative of the torsion in the u^{μ} direction is

$$(\mathscr{L}_u T)^{\alpha}{}_{\beta\gamma} = T^{\alpha}{}_{\beta\gamma,\sigma}u^{\sigma} - T^{\sigma}{}_{\beta\gamma}u^{\alpha}{}_{,\sigma} \tag{40}$$

$$+T^{\alpha}{}_{\sigma\gamma}u^{\sigma}{}_{,\beta}+T^{\alpha}{}_{\beta\sigma}\mu^{\sigma}{}_{,\gamma}=0.$$

Again this equation is the same if comma is replaced by bar (covariant differentiation using ${}^{c}\Gamma^{\alpha}{}_{\beta\gamma}$). With the explicit form of ${}^{c}\Gamma^{\alpha}{}_{\beta\gamma}$, and with the expression $s_{\tau}=S_{\tau\beta}u^{\beta}=S_{\tau 0}$ (note: $s_0=0$), this equation results in the further restrictions

$$s_t C^{t}{}_{i0}=0, \quad S_{it}C^{t}{}_{j0}-S_{jt}C^{t}{}_{i0}=0. \tag{41}$$

These conditions, when combined with Eq. (22) for f_{μ} result in

$$f_0=0, \quad f_i=\tfrac{1}{2}s^{t}C^{0}{}_{ti}. \tag{42}$$

With the use of these conditions, Eq. (23) yields

$$0=S^{t}{}_{i}C^{0}{}_{jt}-S^{t}{}_{j}C^{0}{}_{it}.$$

We are now ready to specify a canonical form for models of this category. We use the freedom of spatial Lorentz transformations (rotations) to set

$$S_{12}=-S_{21}=S, \quad \text{rest of } S_{ij}=0 \tag{44}$$

(s_i remains undetermined). The equations for f_{μ} (with $f_0=0$) now read

$$\rho s_i=(-f_2 S, f_1 S, 0). \tag{45}$$

$$f_i=(-\tfrac{1}{2}s_2C^{0}{}_{12}, \tfrac{1}{2}s_1C^{0}{}_{12}, 0).$$

A rotation among ω^1, ω^2 may be used to set $f_2=0$.

With use of the various conditions so far found for $C^{\alpha}{}_{\beta\gamma}$ the Jacobi identities show there are two possible cases, $C^{1}{}_{02}\neq 0$ and $C^{1}{}_{02}=0$.

If $C^{1}{}_{02}\neq 0$, then $f_i=s_i=0$, and we obtain the following structure constant table (Table IX). At this point the antisymmetric part of the field equations has already been satisfied, and of the rest all of the nondiagonal $R_{(\mu\nu)}$ are zero except for $R_{(03)}$. There are two Jacobi identities remaining.

The Jacobi identities give use two cases: ($C^{0}{}_{12} = C^{3}{}_{12} = 0$), and OSA ($C^{1}{}_{12} = 0$). Calculation of $A_{(\mu\nu)}$ for the first case shows that it is zero, so these models are compatible models, and comparison with Section III shows that they are included in the type C models with $C^{0}{}_{03} = 0$. Type OSA is not a compatible model. The $R_{(03)}$ equation has two solutions, $C^{3}{}_{12} = 0$ or $S = C^{0}{}_{12}$. The second of these is unphysical as it gives negative ρ. The other possibility gives us the structure constant Table X. The field equations result in

$$C^{1}{}_{02}=-\tfrac{1}{2}(S-C^{0}{}_{12}), \tag{46}$$

$$\rho=p=\tfrac{1}{4}(S-C^{0}{}_{12})^2.$$

TABLE IX.

$C^{\alpha}{}_{\beta\gamma}$	01	02	03	12	13	23
0	0	0	0		0	0
1	0		0	0		
2	$-C^{1}{}_{02}$	0	0	0	$-C^{1}{}_{23}$	$C^{1}{}_{13}$
3	0	0	0		0	0

TABLE XI.

Type OSB1	$C^{\alpha}{}_{\beta\gamma}$	01	02	03	12	13	23
	0	0	0	0		0	0
	1	0	0	0		0	0
	2	0	0	0		0	0
	3	0	0	0	0	0	0

The second class of models has $C^1{}_{02}=0$. In this case four Jacobi identities remain. They give us two possibilities: $C^0{}_{12}=0$, or $C^1{}_{13}=-C^2{}_{23}$. In the first of these cases $R_{(0i)}=0$ gives us conditions which show that $A_{\mu\nu}=0$, and these models are of Type C with $C^0{}_{03}=0$.

The models with $C^1{}_{13} = -C^2{}_{23}$ can have either $s_2=f_1=0$, or $s_2=(1/\rho)f_1S=-(1/2\rho)s_2C^0{}_{12}$ which implies $C^0{}_{12}=-2\rho/S$. We must apply this relation to any ρ that depends on $C^0{}_{12}$. The Einstein equations $R_{(02)}=0=R_{(03)}$, $R_{(01)}=\frac{1}{2}u_0f_1$, give us conditions which cause several structure constants to vanish except possibly in unphysical situations. The off-diagonal terms in the remaining field equations again yield subclasses. The first (OSB1) has the structure constant Table XI. The Einstein equations give us

$$\begin{aligned}&\tfrac{1}{2}C^0{}_{12}(S-C^0{}_{12})=-(C^1{}_{12})^2-(C^2{}_{12})^2\\&p=\rho-C^1{}_{12}s^2\\&\rho=C^1{}_{12}s^2+\tfrac{1}{4}(S-C^0{}_{12})^2.\end{aligned}\tag{47}$$

If $s^2=0$, this case is similar to OSA while if $s^2\neq0$ we have to impose $C^0{}_{12}=-2\rho/S$, which gives us

$$\rho=S\,[-(C^1{}_{12}s^2+\tfrac{1}{4}S^2)]^{1/2}.\tag{48}$$

This expression is real if $C^1{}_{12}s^2$ is negative, but in that case $p>\rho$. Notice that when $C^0{}_{12}=-2\rho/S$ and the value of ρ from Eq. (48) is put into the first of Eq. (47), we obtain an additional restriction on ρ, s^2, S, $C^1{}_{12}$, and $C^2{}_{12}$.

The other possibility can be shown to require either $C^1{}_{23}=C^2{}_{13}$, or $C^1{}_{23}=-C^2{}_{13}$ and $C^2{}_{23}C^2{}_{13}=0$. The case $C^1{}_{23}=C^2{}_{13}$ (OSB2) gives us (s^2 arbitrary) the structure constant Table XII. The field equations give us

$$\begin{aligned}&\rho=\tfrac{1}{4}S^2-\tfrac{5}{4}SC^0{}_{12}+(C^0{}_{12})^2\\&p=\tfrac{1}{4}S^2-\tfrac{1}{4}SC^0{}_{12}\\&\tfrac{1}{2}C^0{}_{12}(S-C^0{}_{12})=2(C^1{}_{13})^2+2(C^2{}_{13})^2.\end{aligned}\tag{49}$$

This case will always have $p\geqslant\rho$ and if $s^2\neq0$, the relation $C^0{}_{12}=-2\rho/S$ says that $\rho=-\frac{1}{6}S^2-\frac{2}{3}(C^0{}_{12})^2$ which is unphysical.

The last possibility has an unphysical case and a last subclass (OSB3) with the structure constant Table XIII. The remaining Einstein equations yield

$$\begin{aligned}&\rho=\tfrac{1}{4}S^2-\tfrac{5}{4}SC^0{}_{12}+(C^0{}_{12})^2,\\&p=\tfrac{1}{4}S^2-\tfrac{1}{4}SC^0{}_{12},\\&\tfrac{1}{2}C^0{}_{12}(S-C^0{}_{12})=-2(C^2{}_{23})^2.\end{aligned}\tag{50}$$

TABLE XII.

Type OSB2	$C^\alpha{}_{\beta\gamma}$	01	02	03	12	13	23
	0	0	0	0		0	0
	1	0	0	0	0		
	2	0	0	0	0	$C^1{}_{23}$	$-C^1{}_{13}$
	3	0	0	0	0	0	0

TABLE XIII.

Type OSB3	$C^\alpha{}_{\beta\gamma}$	01	02	03	12	13	23
	0	0	0	0		0	0
	1	0	0	0	0		0
	2	0	0	0	0	0	$-C^2{}_{23}$
	3	0	0	0	0	0	0

We must also have $s^2=0$ [for if $s^2\neq0$, $\rho=-\frac{1}{6}S^2-\frac{2}{3}(C^0{}_{12})^2$].

In summary we have only two models with positive ρ and $p\leqslant\rho$, that is OSA and OSB1 with $s^2=0$. Also, all Type C compatible models with $C^0{}_{03}=0$ are observer-stationary, including the flat metric model. Model OSB2 and OSB3 have $p\geqslant\rho$ and $s^2=0$. The only model that is physically permissible and has $s_i\neq0$ is OSB1. For this case, however, $p>\rho$.

VI. SUMMARY AND CONCLUSIONS

We have discussed some space–time homogeneous fluid-filled models possible in the Einstein–Cartan theory, making use of a form for the stress-energy tensor and spin tensor of a fluid based on the equations of motion of spinning particles. The spin tensor had to be restricted so that the evolution equations could be integrated. Our discussion of these points deserves a fuller development, in a future paper, based on a kinetic theory approach.

The models we derived were selected in three broad categories: First, are the simplest, a flat-metric model and a model with vanishing connection and thus vanishing curvature. Second, are the fluid-compatible models, in which the only effect of torsion is to add a fluidlike term to the Ricci curvature. Finally, we have given the observer-stationary models, in which the Lie derivatives of metric and torsion with respect to the fluid velocity both vanish.

The classes of fluid compatible models are Type A models, with $\rho=p$; and Types B1, B2, C1, and C2, with $p\geqslant\rho$. These models are usually disallowed on the grounds that the speed of sound should be less than the speed of light. However, we have included them for completeness. Type C models are also observer-stationary if the structure constant $C^0{}_{03}$ vanishes. The classes of noncompatible observer-stationary models are Types OSA, OSB1, OSB2, and OSB3. The models for which the linear momentum density is not aligned with the fluid velocity all have $p>\rho$.

These models can be used explicitly for the investigation of the equations of motion of spinning particles. It is generally accepted that a nonspinning particle will follow an extremal line, that is, a curve which is of extremal length. Therefore, its velocity v^μ obeys

$$v^\mu{}_{|\nu}v^\nu=(v^\mu{}_{,\nu}+v^{\sigma c}\Gamma^\mu{}_{\sigma\nu})v^\nu\propto v^\mu.$$

In contrast, an autoparallel curve or "straightest line" would have a velocity vector w^μ which obeys

$$w^\mu{}_{;\nu}w^\nu=(w^\mu{}_{,\nu}+w^\sigma\Gamma^\mu{}_{\sigma\nu}w^\nu\propto w^\mu)$$

and thus in general would feel torsion.[1]

It is clear that a spinning particle will follow neither an extremal nor an autoparallel curve.[6,12] Consequently, a comparison of the paths of spinning and nonspinning particles within the models given here should be of interest. The relative simplicity of our models allows these calculations to be performed in a straightforward fashion. However, similar work done on the Gödel model[15] shows that these calculations deserve separate consideration in a future paper.

[1]J.M. Nester, Ph.D. dissertation, University of Maryland (1977). F.W. Hehl, P. von der Heyde, G.D. Kerlick, and J.M. Nester, Rev. Mod. Phys. **45**, 393 (1976), and the references cited therein.

[2]B. Kuchowicz, Astrophys. Space Sci. **39**, 175 (1976). G.D. Kerlick. Ann. Phys. N.Y. **99**, 127 (1976), and the references cited therein.

[3]K. Gödel, Rev. Mod. Phys. **21**, 447 (1949).

[4]M.P. Ryan and L.C. Shepley, *Homogeneous Relativistic Cosmologies* (Princeton U. P., Princeton, 1975).

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

[6]S. Hojman, Ph.D. thesis, Princeton University (1975).

[7]See A. Papapetrou, Proc. R. Soc. A **209**, 248 (1951).

[8]C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[9]B. Kuchowicz, Acta Cosmol. Z. **9**, 109 (1975).

[10]J. Weysenhoff and A. Raabe, Acta Phys. Pol. **9**, 7 (1947).

[11]W. Tulczyjew, Acta Phys. Pol. **18**, 393 (1959).

[12]F.W. Hehl, Phys. Lett. A **36**, 225 (1971).

[13]A. Trautman, Symp. Math. **12** 139 (1973).

[14]H. Minn and L.C. Shepley, in preparation.

[15]S. Chandrasekhar and J.P. Wright, Proc. Nat. Acad. Sci. **47**, 341 (1961).

The spectral properties of many-electron atomic Hamiltonians and the method of configuration interaction. I. Proof of convergence of the configuration interaction method

M. H. Choudhury and D. B. Pearson

Department of Applied Mathematics, The University, Hull, England
(Received 21 June 1977; revised manuscript received 28 December 1977)

In the method of configuration interaction for atomic structures, the Schrödinger equation for a many electron atom is reduced to an infinite system of linear equations. The eigenvalue problem associated with the finite system of equations obtained by truncating the infinite system is solved to obtain energy eigenvalues and eigenvectors. It is generally assumed that the truncation procedure is a convergent one in the sense that as one increases the size of the truncated equations, the number of eigenvalues and eigenvectors will increase and tend to those of the original infinite set. It is shown that the method of configuration interaction is a convergent procedure in the sense that given any point E in the spectrum of the Schrödinger Hamiltonian of the many-electron system, there exists an eigenvalue of the truncated matrix which is arbitrarily close to E for a sufficiently large size of the truncated matrix. Further, it is shown that the convergence referred to cannot be uniform.

1. INTRODUCTION

In the method of configuration interaction for atomic structures,[1] one reduces the Schrödinger equation for an n-electron atom

$$H\,|\,\Psi\rangle=\Big(\sum_{i=1}^{n}H_{0i}+\sum_{\substack{i,j=1\\ i<j}}^{n}\frac{1}{r_{ij}}\Big)|\,\Psi\rangle=E\,|\,\Psi\rangle \quad (1.1)$$

to an infinite system of linear equations. Here $r_{ij}=|\,\mathbf{r}_i-\mathbf{r}_j\,|$, $\mathbf{r}_i$ and $\mathbf{r}_j$ being the position operators of the ith and jth electron respectively, and

$$H_{0i}=-\tfrac{1}{2}\nabla_i^2-\frac{Z}{r_i},\quad r_i=|\mathbf{r}_i|,\quad i=1,2,\dots,n. \quad (1.2)$$

The above stated reduction to an infinite system of linear equations can be accomplished either directly or by means of a variational principle. The eigenvalue problem associated with the finite system of linear equations obtained by truncating the infinite system is then solved to obtain energy eigenvalues and eigenvectors. In this connection it is argued[1,2] that truncation will very likely lead to a "convergent procedure" in the sense that as we increase the size of the truncated equations, the number of eigenvalues and eigenvectors obtained by solving the eigenvalue problem associated with the truncated system of equations will increase and tend to those of the original infinite system of equations.

In this and subsequent papers the following objectives will be achieved:

(i) It will be shown that the usual method of configuration interaction for atomic structures is a convergent procedure in the sense that given any point E in the spectrum of H (discrete or continuous), there exists an eigenvalue of the truncated $N\times N$ matrix which is arbitrarily close to E provided N is sufficiently large. This result does not guarantee that if one obtains an energy eigenvalue of the truncated $N\times N$ matrix which is very close to the ground state energy eigenvalue of H, then the remaining $N-1$ eigenvalues will be similarly close to the first $N-1$ excited states. In fact, as N is increased, sooner or later one is certain to obtain positive energy eigenvalues.

(ii) A formulation of the configuration interaction method will be given which not only leads to a convergent procedure on truncation of the infinite system of linear equations obtained in this formulation, but also ensures that as the size of the truncated equations is increased, the number of eigenvalues will increase and uniformly converge to those of the original infinite system. This is achieved by deriving from the Schrödinger equation of an n-electron atomic system an infinite set of linear equations which defines a compact linear operator in a suitable region of the complex energy plane. The compactness of this operator then ensures[3]

(a) that the procedure of truncation is a convergent one,

(b) that as the size of the truncated matrix is increased, the number of eigenvalues will increase and uniformly converge to the eigenvalues of the original infinite matrix.

The proof of compactness exhibits in a clear and detailed fashion, the spectral properties of the Schrödinger Hamiltonian operator for many-electron atomic systems. An appreciation of the spectral properties is crucially important and reveals, for instance, which of the various mutiparticle and bound state scattering cuts which constitute the continuum has the most significant contributions to make in the computation of the energy eigenvalues and eigenvectors of the atomic system in question.

In this paper our aim shall be to accomplish objective (i). In two subsequent papers, objective (ii) shall be achieved.

In Sec. 2, we prove the result referred to in (i), namely, that if E is any point in the spectrum of H (discrete or continuous), then there exists an eigenvalue of the truncated $N\times N$ matrix which is arbitrarily close to E for sufficiently large N. We then discuss the implications of this result some of which have already been mentioned earlier in (i). In Sec. 3

 0022-2488/79/040752-05$01.00

it will be shown that the convergence referred to cannot be uniform in the sense that as the size of the truncated matrix is increased, the number of eigenvalues will also increase, but will not converge uniformly to those of the original infinite matrix. As mentioned earlier, one method of obtaining a uniformly convergent truncation procedure satisfying properties (a) and (b) is to derive from Eq. (1.1) an infinite system of linear equations which defines a compact linear operator. A seemingly obvious candidate which may define a compact operator, is the system of Eq. (3.9). A compactness test, which consists in verifying the square summability of the matrix elements of some positive integral power of the linear operator in question,[4] yields an infinite result when applied to the linear operator defined by the system of equations (3.9) for $n \geqslant 3$. We show that the reason for this infinite result lies in the occurrence of kronecker delta's and delta functions in the matrix elements precisely in the same manner as delta functions occur in the corresponding situation in scattering theory,[5] namely the scattering of electrons from atomic systems. It is interesting to note that the usual equations of the configuration interaction method, as given by Eq. (2.1), also suffer from this defect.

The essential results of this paper are that:

(i) the truncation procedure is a convergent one and depends in a significant way on the spectral structure of the Hamiltonian H,

(ii) the convergence is nonuniform and is sensitive in its details on the choice of the complete orthonormal basis set $\{|\phi_m\rangle\}$.

While these facts relating to the sensitivity of the convergence process on the choice of an orthonormal basis set are well known to physicists and chemists undertaking practical configuration interaction calculations, the detailed mathematical reasons as to why they occur and their relation to the spectral structure of the Hamiltonian is not sufficiently understood or appreciated. Once the problem of convergence has been settled, the very important problems relating to the choice of basis sets, ordering and labeling of configuration states can then be dealt with. It is at this juncture, that MacDonald's Theorem[6] on the separation of eigenvalues of the truncated matrix takes on a very significant and important role. These and allied problems have been dealt with in a detailed and original fashion by Luken and Sinanoglu[7] and the references contained therein. Here we stress the importance of an appropriate use of MacDonald's theorem, especially in its relation to the results derived in this paper. It is important to note that MacDonald's theorem is a general result and applies to any system of truncated equations associated with an eigenvalue problem, whose matrix defines a self-adjoint operator. However, it does not tell us (and it is not meant to tell us) whether the eigenvalues of the truncated matrix increase in number and converge in the limit to those of the original infinite configuration interaction matrix (that is, if the truncation procedure is convergent). Even if the truncation procedure does not yield convergence to the eigenvalues of the original infinite configuration interaction matrix, MacDonald's theorem can still be applied, but then its use would be quite inappropriate and meaningless. In fact, it is possible to give examples of Hamiltonians for which the truncation procedure does not yield convergence to the eigenvalues of the original infinite configuration interaction matrix. Here, to put it simply, the usual configuration interaction method breaks down.

2. PROOF OF CONVERGENCE

Before we state and prove the main result of this section, we require a few definitions and concepts. In what follows we shall denote the domain of H by $D(H)$ and the spectrum of H by $\sigma(H)$. Let $\{|\phi_n\rangle\} \in D(H)$ be an orthonormal basis. Taking the inner product on both sides of (1.1) with respect to $\langle\phi_n|$ and using the resolution of the identity

$$\sum_{n=1}^{\infty} |\phi_n\rangle\langle\phi_n| = I_H,$$

one obtains the usual equations of the configuration interaction method,[1]

$$\sum_{m=1}^{\infty} \langle\phi_n | H | \phi_m\rangle\langle\phi_m | \Psi\rangle = E\langle\phi_n | \Psi\rangle. \tag{2.1}$$

Definition: A set S of vectors is said to form a *core* for H if

(i) $|\psi\rangle \in S \Rightarrow |\psi\rangle \in D(H)$

(ii) Given any $|\psi\rangle \in D(H)$, there exists a sequence $\{|\Psi_n\rangle\} \in S$ with the property that $|\psi_n\rangle \to |\psi\rangle$ and $H|\psi_n\rangle \to H|\psi\rangle$ in norm as $n \to \infty$.

In other words S forms a core if the closure of H restricted to S is equal to H.

Suppose that finite linear combinations of the $|\phi_n\rangle$ form a core for H. We shall denote this core by S. We define the matrix $A^{(N)}$ to be the $N \times N$ matrix with elements

$$[A^{(N)}]_{ij} = \langle\phi_i | H | \phi_j\rangle, \quad 1 \leqslant i,j \leqslant N. \tag{2.2}$$

We are now ready to state the main result of this section.

Theorem: Let $E \in \sigma(H)$. Then given any $\epsilon > 0$, there exists $N_0(\epsilon)$ such that for $N > N_0$, the matrix $A^{(N)}$ has an eigenvalue $E^{(N)}$ such that $|E^{(N)} - E| < \epsilon$.

It is convenient to demonstrate the proof in four steps. We prove in

Step 1: There exists a sequence $\{|f_n\rangle\} \in D(H)$ with $\| |f_n\rangle \| = 1$ and

$\|(H - E)|f_n\rangle\| \to 0$ as $n \to \infty$;

Step 2: This sequence $\{|f_n\rangle\}$ can be replaced by a sequence $\{|\psi_n\rangle\} \in S$ with $\| |\psi_n\rangle \| = 1$ and having the same property, that is

$\|(H - E)|\psi_n\rangle\| \to 0$ as $n \to \infty$;

Step 3: If $|\psi_n\rangle = \Sigma_{k=1}^{N} |\phi_k\rangle\langle\phi_k|\psi_n\rangle = \Sigma_{k=1}^{N} C_k|\phi_k\rangle$ (with $C_k = 0$ for $k > N$), then

$\|(A^{(N)} - E)\mathbf{C}\| \leqslant \|(H - E)|\psi_n\rangle\|$,

$\mathbf{C} = (C_1, C_2, \ldots C_N)$;

Step 4: $\epsilon^{-1} < \|(A^{(N)} - E)^{-1}\| \leqslant |E^{(N)} - E|^{-1}$

which of course implies that $|E^{(N)} - E| < \epsilon$.

Proof: Step 1: $(H-E)^{-1}$ is unbounded (Recall the $E\in\sigma(H)$). Hence there exists a sequence $\{|g_n\rangle\}$ with $\||g_n\rangle\| = 1$ and $\|(H-E)^{-1}|g_n\rangle\| \to \infty$ as $n\to\infty$. Now take

$$|f_n\rangle = \frac{(H-E)^{-1}|g_n\rangle}{\|(H-E)^{-1}|g_n\rangle\|},$$

and the result follows.

Step 2: For given $|f_n\rangle$ and $\epsilon > 0$, we can find $|\psi_n\rangle \in S$ such that

$$\|(|\psi_n\rangle - |f_n\rangle)\| < \epsilon, \text{ and } \|(H|\psi_n\rangle - H|f_n\rangle)\| < \epsilon,$$

hence

$$\begin{aligned}\|(H-E)|\psi_n\rangle\| &= \|(H-E)|f_n\rangle + H(|\psi_n\rangle - |f_n\rangle) - E(|\psi_n\rangle - |f_n\rangle)\| \\ &\leqslant \|(H-E)|f_n\rangle\| + \epsilon + |E|\epsilon.\end{aligned}$$

Now take $\epsilon = 1/n$ and the result follows. Note that the $\{|\psi_n\rangle\}$ are not necessarily normalized, but the sequence may be replaced by $\{|\psi_n\rangle/\||\psi_n\rangle\|\}$ since

$$\|(|\psi_n\rangle - |f_n\rangle)\| < \epsilon \Rightarrow \||\psi_n\rangle\| \geqslant 1-\epsilon$$

so that $\||\psi_n\rangle\|$ cannot become arbitrarily small.

Step 3: Given $\epsilon > 0$, chooses $|\psi_n\rangle \in S$ such that $\||\psi_n\rangle\| = 1$ and $\|(H-E)|\psi_n\rangle\| < \epsilon$ (by Step 2). We have

$$\begin{aligned}\|(A^{(N)} - E)\mathbf{C}\|^2 &= \sum_{i=1}^{N} \left| \sum_{j=1}^{N} \langle\phi_i(H-E)|\phi_j\rangle\langle\phi_j|\psi_n\rangle\right|^2 \\ &= \sum_{i=1}^{N} |\langle\phi_i|(H-E)|\psi_n\rangle|^2 \\ &\leqslant \langle\psi_n|(H-E)^2|\psi_n\rangle = \|(H-E)|\psi_n\rangle\|^2.\end{aligned}$$

Hence

$$\|(A^{(N)} - E)\mathbf{C}\| \leqslant \|(H-E)|\psi_n\rangle\| < \epsilon.$$

Step 4: Since $\|\mathbf{C}\| = 1$ and $\|A^{(N)} - E)\mathbf{C}\| < \epsilon$, we have

$$\|(A^{(N)} - E)^{-1}\| \geqslant \frac{\|(A^{(N)} - E)^{-1}\{(A^{(N)} - E)\mathbf{C}\}\|}{\|(A^{(N)} - E)\mathbf{C}\|} > \frac{1}{\epsilon}.$$

Hence, if $E^{(N)}$ is the eigenvalue of $A^{(N)}$ closest to E, we have

$$\frac{1}{\epsilon} < \|(A^{(N)} - E)^{-1}\| \leqslant |E^{(N)} - E|^{-1}.$$

We now discuss the implications of this theorem for the method of configuration interaction in calculating the energy eigenvalues and eigenfunctions of atomic structures. It shows, for instance, that for the ground state energy level of (2.1), there exists an eigenvalue of $A^{(N)}$ which is arbitrarily close to it for sufficiently large N. From a practical point of view, for a given accuracy, it would be desirable to have as small a value for N as possible, thereby minimizing the size of the matrix $A^{(N)}$ whose eigenvalues and eigenvectors are to be computed. This in turn would depend on how cleverly one can choose the complete orthonormal sequence $\{|\phi_n\rangle\}$ for the task. This is not a particularly difficult task if one is primarily interested in the ground state energy levels of atoms containing a few electrons. However, there is no guarantee that the remaining $N-1$ energy eigenvalues obtained will be similarly close to the lowest $N-1$ excited levels. Indeed, one may very well obtain eigenvalues, some of which are positive. It is well to remind ourselves that one implication of this theorem is that for sufficiently large N, there will exist eigenvalues which will be arbitrarily close to positive points of the spectrum of H. Indeed, as $N\to\infty$, the largest of the eigenvalues of $A^{(N)}$ will also tend to infinity.

Finally let us say that the task would be prohibitively difficult if one wishes to find an orthonormal set for a not too large value of N which will be such that the N eigenvalues of $A^{(N)}$ will be very close to the lowest N energy levels of (2.1). To the best of our knowledge there exists no publication which has claimed to do this. One way this could be accomplished is if it were possible for the N energy eigenvalues and eigenvectors of $A^{(N)}$ to converge uniformly to the lowest N energy eigenvalues and eigenvectors of (2.1). In the next section we show this to be impossible.

3. PROOF THAT CONVERGENCE CANNOT BE UNIFORM: ASPECTS OF CONFIGURATION INTERACTION METHOD HAVING COUNTERPARTS IN MULTIPARTICLE SCATTERING THEORY

The truncation of Eq. (2.1) yields a system of $n\times n$ linear equations which define a linear operator, call it $H^{(n)}$, of finite rank and hence compact. We shall show that, as $n\to\infty$, the eigenvalues and eigenvectors of $H^{(n)}$ cannot tend uniformly to those of H.

Let $\mathscr{H}$ be the Hilbert space in which H acts. Denote by $P_k^{(n)}$ the projection operator onto the k th (normalized) eigenvector $|\phi_k^{(n)}\rangle$ of $H^{(n)}$, corresponding to eigenvalue $E_k^{(n)}$ (with $P_k^{(n)} = 0$ for $k > n$.) The $E_k^{(n)}$ need not be distinct if eigenvalues are degenerate. We have therefore

$$P_k^{(n)}|\phi\rangle = |\phi_k^{(n)}\rangle, \quad |\phi\rangle \in \mathscr{H}. \tag{3.1}$$

The spectral resolution of $H^{(n)}$ is

$$H^{(n)} = \sum_{k=1}^{n} E_k^{(n)}P_k^{(n)} = \sum_{k=1}^{\infty} E_k^{(n)}P_k^{(n)}.$$

Using this formula we have

$$\begin{aligned}\|H^{(n)} - H^{(m)}\| &\leqslant \sum_{k=1}^{\infty} \|(E_k^{(n)} - E_k^{(m)})P_k^{(n)} + E_k^{(m)}(P_k^{(n)} - P_k^{(m)})\| \\ &\leqslant \sum_{k=1}^{\infty} |E_k^{(n)} - E_k^{(m)}|\,\|P_k^{(n)}\| + \sum_{k=1}^{\infty} |E_k^{(m)}|\,\|P_k^{(n)} - P_k^{(m)}\|. \end{aligned} \tag{3.2}$$

Also, using (3.1)

$$\begin{aligned}\|P_k^{(n)} - P_k^{(m)}\| &= \sup \frac{\|(P_k^{(n)} - P_k^{(m)})|\phi\rangle\|}{\||\phi\rangle\|} \\ &= \sup \frac{\|(|\phi_k^{(n)}\rangle - |\phi_k^{(m)}\rangle)\|}{\||\phi\rangle\|} < \frac{\epsilon}{2|E_k^{(m)}|},\end{aligned}$$

since, by hypothesis, the eigenvectors $|\phi_k^{(n)}\rangle$ are uniformly convergent with respect to the index n and therefore form a Cauchy sequence. Since the same is true of the eigenvalues $E_k^{(n)}$, the first term on the right-hand side of (3.2) can be made less than $\epsilon/2$. (Note that $\|P_k^{(n)}\|\leqslant 1$.) Hence

$$\|H^{(n)} - H^{(m)}\| < \epsilon.$$

Since the bounded operators on $\mathscr{H}$ (of which the compact operators form a subset) constitute a Banach space with respect to the operator norm, the sequence of operators $\{H^{(n)}\}$ must converge to a bounded operator. Denote this bounded operator by H_B. Since H is unbounded we must have $H\neq H_B$.

Let $\mathscr{H}_i$, $i = 1,2,...,n$ be the space of states associated with electrons $1,2,...,n$ respectively, and $\{|\alpha_i\rangle\}$ denote a (complete) discrete basis in the space $\mathscr{H}_i$ (with $\{|\alpha_i\rangle\}\subset D(H)$.) The resolution of the identity in $\mathscr{H}_i$ is

$$\sum_{\alpha_i} |\alpha_i\rangle\langle\alpha_i| = I_{Hi}, \quad i = 1,2,...,n. \tag{3.3}$$

A discrete basis in the product space $\mathscr{H} = \mathscr{H}_1\otimes\mathscr{H}_2\otimes\cdots\otimes\mathscr{H}_n$ is of the form $|\alpha_1,\alpha_2,...,\alpha_n\rangle = |\alpha_1\rangle|\alpha_2\rangle\cdots|\alpha_n\rangle$ and the resolution of the identity in $\mathscr{H}$ is

$$\sum_{\alpha_i} |\alpha_1,...,\alpha_n\rangle\langle\alpha_1,...,\alpha_n| = I_H. \tag{3.4}$$

Also, denoting the bound states of an hydrogenic atom by $|\mathbf{n}\rangle = |nlm\rangle$ and its continuum states by $|\mathbf{k}_H\rangle = |k_H lm\rangle$, the resolution of identity in this case can be written

$$\sum_{\mathbf{n}_i} |\mathbf{n}_i\rangle\langle\mathbf{n}_i| + \Big(\int\sum_{\mathbf{k}_{Hi}}\Big)|\mathbf{k}_{Hi}\rangle\langle\mathbf{k}_{Hi}| = I_{Hi}, \quad i = 1,2,...,n, \tag{3.5}$$

where the sign $(\int\Sigma)$ will be used to indicate a summation over the discrete set of quantum numbers and an integration over the continuous set. It will prove convenient to use the abridged form

$$\Big(\int\sum_{\mathbf{v}_i}\Big)|\mathbf{v}_i\rangle\langle\mathbf{v}_i| = I_{Hi}, \quad i = 1,2,...,n, \tag{3.6}$$

where $\mathbf{v}_i$ assumes discrete and/or continuous values appropriately. The resolution of the identity in the product space $\mathscr{H}$ is then

$$\Big(\int\sum_{\mathbf{v}_i}\Big)|\mathbf{v}_1,\mathbf{v}_2,...,\mathbf{v}_n\rangle\langle\mathbf{v}_1,\mathbf{v}_2,...,\mathbf{v}_n| = I_H. \tag{3.7}$$

For two-electron atoms one can invert Eq. (1.1) (with $n = 2$) to obtain

$$|\Psi\rangle = (E - H_{01} - H_{02})^{-1}\frac{1}{r_{12}}|\Psi\rangle. \tag{3.8}$$

Taking the inner product with respect to $\langle\alpha_1,\alpha_2|$ and using (3.4) with $n = 2$ leads to the infinite system of equations

$$\langle\alpha_1,\alpha_2|\Psi\rangle = \sum_{\alpha_1',\alpha_2'}\Big\langle\alpha_1,\alpha_2\Big|(E - H_{01} - H_{02})^{-1}\frac{1}{r_{12}}\Big|\alpha_1',\alpha_2'\Big\rangle \times\langle|\alpha_1',\alpha_2'|\Psi\rangle. \tag{3.9a}$$

In the next paper we shall show that the operator $(E - H_{01} - H_{02})^{-1}(1/r_{12})$ is compact in a suitable region of the complex energy plane which excludes the multiparticle or continuum cuts. Hence, in the case of Eq. (3.9a), the truncation procedure will be a uniformly convergent one.

A naive generalization of the procedure to the n-electron case (for $n\geqslant 3$) fails. To see this, we apply a compactness test to the linear operator defined by the system of equations

$$\langle|\alpha_1,...,\alpha_n|\Psi\rangle = \sum_{\alpha_i'}\langle\alpha_1,...,\alpha_n|\Big(E - \sum_{p=1}^{n} H_{0p}\Big)^{-1} \sum_{\substack{i,j=1\\ i<j}}^{n}\frac{1}{r_{ij}}|\alpha_1',...,\alpha_n'\rangle\langle\alpha_1',...,\alpha_n'|\Psi\rangle. \tag{3.9}$$

The compactness test consists in showing that any positive integral power of the linear operator is of the Hilbert–Schmidt type.[4] We have, since the Hilbert–Schmidt norm is independent of the basis[8]

$$\sum_{\alpha,\alpha'}\Big|\langle\alpha_1,...,\alpha_n|\Big(E - \sum_{p=1}^{n} H_0\Big)^{-m}\Big(\sum_{i,j=1}^{n}\frac{1}{r_{ij}}\Big)^m|\alpha_1',...,\alpha_n'\rangle\Big|^2$$

$$= \Big(\int\sum_{\mathbf{v}_i}\Big)\Big(\int\sum_{\mathbf{v}_i'}\Big)\Big|\Big(E - \sum_{p=1}^{n} E_{\mathbf{v}p}\Big)^{-m}\Big|^2 \times\Big|\langle\mathbf{v}_1,...,\mathbf{v}_n|\Big(\sum_{i,j=1}^{n}\frac{1}{r_{ij}}\Big)^m|\mathbf{v}_1',...,\mathbf{v}_n'\rangle\Big|^2$$

$$= \Big(\int\sum_{\mathbf{v}_i}\Big)\Big(\int\sum_{\mathbf{v}_i'}\Big)\Big|\Big(E - \sum_{p=1}^{n} E_{\mathbf{v}p}\Big)^{-m}\Big|^2 \times\Big|\sum_{ij}\langle\mathbf{v}_i,\mathbf{v}_j|\frac{1}{r_{ij}^m}|\mathbf{v}_i',\mathbf{v}_j'\rangle\prod_{\substack{k=1\\ k\neq i,j}}^{n}\delta(\mathbf{v}_k,\mathbf{v}_k') + \langle\mathbf{v}_1,...,\mathbf{v}_n|(\text{cross terms})|\mathbf{v}_1',...,\mathbf{v}_n'\rangle\Big|^2 = \infty$$

for $n\geqslant 3$,

where $\delta(\mathbf{v}_k,\mathbf{v}_k') = \delta(v_{k\,1},v_{k\,1}')\,\delta(v_{k\,2},v_{k\,2}')\delta(v_{k\,3},v_{k\,3}')$ is the product of three kronecker delta's or a delta function and two kronecker delta's depending on whether $\mathbf{v}_k$, $\mathbf{v}_k'$ refer to the discrete or continuous part of the spectrum. The infinite sums over the kronecker delta's and delta functions lead to an infinite result. Another way of looking at this adverse result is to observe that the two-particle interactions couples up only two of the particles, thus allowing the remaining $(n - 2)$ particles to go "straight through" as it were and produce either a kronecker delta or a delta function. This is precisely what happens in multiparticle scattering, where it is found that this infinite result is due to delta functions occurring in the kernel of the integral equation for the scattering process.[5] In the scattering situation this led to the formulation of Faddeev[9] and Weinberg[10], where this difficulty is removed. In the third of our present series of articles, we shall show how to deal with this problem for the general case of the n-electron atom.

[1]P.O. Lowdin, Adv. Chem. Phys. **2**, 259–97 (1959).
[2]E. Holoien, Physica Horvegica **1**, 53–111 (1961).
[3]N. Dunford and J. Schwartz, *Linear Operators* (Interscience, New York, 1963) Part II, p. 1091.
[4]R. Schatten, *Norm Ideals of Completely Continuous Operators* (Springer-Verlag, Berlin, 1960), p.18.
[5]C. Lovelace, Three particle systems and unstable particle in *Strong Interactions and High Energy Physics* (Oliver and Boyd, Edinburgh and London, 1964), R.G. Moorhouse, Ed.
[6]J.K.L.MacDonald, Phys. Rev. **43**, 830 (1933). See also E.A. Hylleraas and B. Undheim, Z. Phys. **65**, 759 (1930).
[7]W.L. Luken and O. Sinanogulu, Phys. Rev A. **13**, 1293 (1976).
[8]Ref. 3, p. 1010.
[9]L.D. Faddeev, Soc. Phys. JETP **12**, 1014 (1960); Dokl. **6**, 384 (1961); Dokl. **7**, 600 (1963). (All translations).
[10]S. Weinberg, Phys. Rev. **133**, 232–56 (1964).

CUMULATIVE AUTHOR INDEX

All authors published so far in the current volume are listed alphabetically with the issue and page numbers following the dash. A cumulative author and subject index covering each volume is published annually. An (E) *after the page number indicates an Erratum.*

Abellanas, L.—(3) 437
Adler, M.—(1) 60
Akyeampong, D. A.—(3) 505
Aldersley, S. J.—(3) 522
Anandan, J.—(2) 260
Anile, A. M.—(1) 177
Arms, Judith M.—(3) 443
Arnal, Didier—(4) 556
Arrowsmith, D. K.—(1) 101

Barnsley, M.—(4) 535
Basu, Debabrata—(3) 492
Bednář, Miroslav—(3) 367
Ben-Artzi, Matania—(4) 594
Bertero, M.—(3) 509
Bessis, D.—(4) 535
Biedenharn, L. C.—(2) 269
Bosanac, S.—(3) 396
Bowden, R. L.—(4) 608

Campanino, M.—(1) 210
Case, K. M.—(2) 299
Castaños, O.—(1) 35
Cetto, A. M.—(3) 469
Chacón, E.—(1) 35
Chen, M.—(2) 254
Choudhury, M. H.—(4) 752
Christensen, Mark J.—(3) 385
Combe, Ph.—(4) 611
Cornille, Henri—(1) 199
Cornwell, J. F.—(4) 547
Cortet, Jean-Claude—(4) 556
Cruz, A.—(1) 126

Darewych, J. W.—(4) 736
Das, A.—(4) 740
Das, U.—(3) 372, 378
Davies, E. B.—(3) 345
de la Peña, L.—(3) 469
Del Grosso, G.—(1) 210
De Mol, C.—(3) 509
Desmier, P. E.—(1) 74
Devinatz, Allen—(4) 594
Dirl, R.—(4) 659, 664, 671, 679
Dudley, Alan L.—(2) 311
Dyer, C. C.—(1) 1; (3) 409

Epifanio, G.—(1) 148

Finley, III, J. D.—(2) 311
Floyd, Edward R.—(1) 83
Frank, A.—(1) 35
Furutsu, K.—(4) 617

Garibotti, C. R.—(1) 141
Geronimo, J. S.—(2) 299; (3) 484
Glassey, R. T.—(3) 454
Gordon, Paul—(4) 634
Graffi, S.—(4) 685
Grecchi, V.—(4) 685
Grinstein, F. F.—(1) 141

Halil, Mustafa—(1) 120
Harvey, Alex—(2) 251
Herrera, L.—(1) 195
Hirondel, D.—(1) 104
Hoenders, B. J.—(2) 329
Honig, E.—(1) 1; (3) 409
Horndeski, Gregory Walter—(4) 726
Horváthy, P.—(1) 49
Horwitz, L. P.—(2) 269
Hudson, R. L.—(1) 114

Inomata, Akira—(3) 499

Jagannathan, R.—(3) 390
Jakubowicz, Oleg—(3) 441
Jiménez, J.—(1) 195

Klein, Abraham—(4) 572

Langguth, Wolfgang—(3) 499
Leach, P. G. L.—(1) 86, 96
Levoni, S.—(4) 685
Lewis, H. Ralph—(3) 413; (4) 564
Li, Ching-teh—(4) 572
Lloyd-Evans, D. J. R.—(1) 68

Maekawa, Takayoshi—(4) 691
Maioli, M.—(4) 685
Maiti, S. R.—(2) 245
Majumdar, S. Datta—(3) 492
Mallick, N.—(3) 378
Martin, J.—(1) 25
Martínez Alonso, L.—(2) 219; (3) 437
McCormick, N. J.—(1) 216
Mitchell, Thomas P.—(3) 382
Miyazaki, Eizo—(1) 184
Moncrief, Vincent—(4) 579
Morgan III, John D.—(2) 344(E)
Moshinsky, M.—(1) 35
Moussa, P.—(4) 535
Mukhopadhyaya, S.—(3) 372

Nouri-Moghadam, M.—(1) 20

Okubo, Susumu—(4) 586

Pantano, P.—(1) 177
Pearson, D. B.—(4) 752
Peck, S. N.—(1) 114; (2) 336
Perez, R. B.—(2) 241
Petry, Herbert Rainer—(2) 231
Poénaru, V.—(1) 13
Pulè, J. V.—(4) 733

Raychaudhuri, A. K.—(2) 245
Reddy, D. R. K.—(1) 23
Reichl, L. E.—(3) 352, 361
Rodriguez, R.—(4) 611
Rosen, Gerald—(3) 441
Rosenbaum, Marcos—(4) 744
Rosensteel, G.—(3) 465
Rowe, D. J.—(3) 465
Ryan,Jr., Michael P.—(4) 744

Šachl, Vladimír F.—(3) 367
Sage, Martin L.—(3) 393
Sakmar, I. A.—(1) 11
Sanz, J. L.—(1) 25
Scheunert, M.—(4) 712
Sen Gupta, M.—(3) 378
Sesma, J.—(1) 126
Sharp, R. T.—(1) 74
Shepley, L. C.—(4) 744
Sirugue, M.—(4) 611
Sirugue-Collin, M.—(4) 611
Smallwood, J.—(3) 459
Snellman, Håkan—(1) 132
So, S. I.—(1) 153
Sokoloff, J.—(4) 736
Soo Jang, Pong—(1) 217(E)
Strauss, W. A.—(3) 454
Strinati, G.—(1) 188
Strottman, D.—(1) 153
Symes, W.—(4) 721
Symon, Keith R.—(3) 413

Talukdar, B.—(3) 372, 378
Tatoian, Vahe—(3) 441
Tengstrand, Göran—(1) 132
Tikochinsky, Yoel—(3) 406
Toulouse, G.—(1) 13
Trapani, C.—(1) 148
Turner, Leaf—(4) 564

Vasudevan, R.—(3) 390
Veeder, J. A. R.—(1) 216
Verbeure, A.—(4) 733
Viano, G. A.—(3) 509

Wali, Kameshwar C.—(3) 393
Weston, V. H.—(1) 53
Wódkiewicz, K.—(1) 45
Wong, M. K. F.—(2) 247; (4) 629

Yeh, Hsin-Yang—(2) 247; (4) 629
Yoshimura, Tetz—(1) 20

Zorbas, John—(1) 6