Shift operator techniques for the classification of multipole-phonon states: IV. Properties of shift operators in the G_2 group

G. Vanden Berghe and H. E. De Meyer[a)]
Seminarie voor Wiskundige Natuurkunde, R. U. G., Krijgslaan 271-S9, B-9000 Gent, Belgium

(Received 4 March 1980; accepted for publication 4 April 1980)

A previously developed shift operator method is applied to the G_2 group, which plays an important role in the classification of nuclear octupole-phonon states. Expressions which connect quadratic products of the considered shift operators with G_2 invariants are derived.

1. INTRODUCTION

The shift operator method, previously introduced by Hughes *et al.*[1,2] and recently applied by the present authors[3-5] to solve the state labeling problem of the nuclear quadrupole-phonon states, will be developed here with a view to obtaining an orthogonal specification of the octupole-phonon states. The symmetry group of the octupole Hamiltonian is the U(7) group, whose symmetric representations play an important role in the classification of the considered phonon states. Obviously, seven labels are needed to classify these states. Four of them are related to the Casimir operators of groups appearing in the chain

$$\mathrm{U}(7)\supset \mathrm{SU}(7)\supset \mathrm{R}(7)\supset G_2\supset \mathrm{R}(3)\supset \mathrm{R}(2), \tag{1.1}$$

i.e., the boson number N for $U(7)$, the seniority v for R(7), the angular momentum l for R(3), and its projection m for R(2). The SU(7) and G_2 labels are redundant for the symmetric representations. For the other three internal labels Rohozinsky[6] has proposed the number of quartets and sextets of phonons coupled to spin zero and a nonphysical label defining a residual factor. These three labels however are not related with the eigenvalue of an operator.

Hughes's technique[1,2] provides us with R(3) scalar operators which can play a fundamental role in the orthogonal specification of the states considered. The apparatus for obtaining their eigenvalues in an analytic way consists of the l-shift operators. The mentioned method can be applied to any group possessing a R(3) subgroup. It is evident that one starts with the group with the smallest number of generators, i.e., G_2. The Lie algebra consists of the generators l_i ($i=0,\pm$) of R(3) and the components p_μ ($\mu=-5,\ldots,5$) of an 11-dimensional irreducible representation of R(3). These group generators are defined in terms of the creation ($b^+_{3\mu}$) and annihilation $[(-1)^\mu b_{3-\mu}]$ operators of the octupole-phonon states by Weber *et al.*[7] Since Hughes and Yadegar[2] assume that the p_μ and the generators l_0, $l_\pm$ of R(3) satisfy the standard commutation relations

$$[l_\pm, p_\mu] = [(5\mp\mu)(6\pm\mu)]^{1/2} p_{\mu\pm1}, \tag{1.2}$$

$$[l_0, p_\mu] = \mu p_\mu, \tag{1.3}$$

a little different form for these generators has to be introduced, namely

$$l_0 = -2\sqrt{7}(b^+_3 b_3)^1_0, \quad l_\pm = \pm 2\sqrt{14}(b^+_3 b_3)^1_{\pm1} \tag{1.4}$$

$$p_\mu = (b^+_3 b_3)^5_\mu, \quad \text{with} \quad \mu=-5,\ldots,5. \tag{1.5}$$

2. EXPLICIT FORMS FOR THE G_2 SHIFT OPERATORS

For the construction of the R(5) shift operators (Ref. 3, to be referred to as I), use has been made of the general expressions derived for any arbitrary $(2j+1)$ dimensional tensor representation of R(3). This way of working is only interesting if analytic expressions for the occurring 3-j symbols are available in the literature. As far as we know this is not the case for the 3-j symbols appearing in the present investigation. Therefore we have preferred to use an alternative method of determining the shift operators, denoted as P^k_l, by requiring that

$$[L^2, P^k_l] = k(k+2l+1)P^k_l \quad \text{for} \quad k=0,1,2,3,4,5. \tag{2.1}$$

This method has firstly been used by Hughes[1] in the SU(3) case. Here $l(l+1)$ is the eigenvalue of the R(3) Casimir operator L^2. It is assumed that the operators P^k_l shift the l value by k, without altering the eigenvalue of l_0. We also require P^k_l to depend linearly on p_μ, and therefore we choose its form as

$$\begin{aligned} P^k_l = {} & ap_0 + b(p_{+1}l_- \pm p_{-1}l_+) \\ & + c(p_{+2}l^2_- \pm p_{-2}l^2_+) + d(p_{+3}l^3_- \pm p_{-3}l^3_+) \\ & + e(p_{+4}l^4_- \pm p_{-4}l^4_+) + f(p_{+5}l^5_- \pm p_{-5}l^5_+), \end{aligned} \tag{2.2}$$

where the upper sign is valid for odd k and the lower sign for even k ("a" becoming zero in the latter case). this sign convention follows directly from the general formula (I.1.4) and form the symmetry relation for 3-j symbols. Introducing (2.2) into (2.1), one may solve for each k value a, b, c, d, e, and f in terms of l and m. On the other hand the expression for $P^{-k}_l (k>0)$ follows immediately due to the relation[2]

$$P^{-k}_l = P^k_{-(l+1)}. \tag{2.3}$$

The explicit expressions of these operators involve the l and m values of the states upon which they act. In order, however, to considerably simplify their calculation we have restricted our considerations to the case when they act on states of zero m value. Since the study of the considered shift operators is closely related to the derivation of the P^0_l eigenvalues, which are independent of m^2, this seemingly drastic condition will not seriously detract from the generality of subsequent calculations. The following results are finally obtained:

[a)]Qualified researcher N.F.W.O. (Belgium).

 0022-2488/80/081967-06$1.00

$$P_l^0 = -\frac{\sqrt{2}\sqrt{5}}{\sqrt{3}\sqrt{7}}$$
$$\times(l-1)(l-2)(l+2)(l+3)(p_{+1}l_- - p_{-1}l_+)$$
$$+\frac{4\sqrt{5}}{\sqrt{6}}(l-2)(l+3)(p_{+2}l_-^2 - p_{-2}l_+^2)$$
$$+\frac{\sqrt{5}}{3}(l^2+l-24)(p_{+3}l_-^3 - p_{-3}l_+^3)$$
$$-2\sqrt{2}\sqrt{5}(p_{+4}l_-^4 - p_{-4}l_+^4) - (p_{+5}l_-^5 - p_{-5}l_+^5), \tag{2.4}$$

$$P_l^{+1}/(l+1)$$
$$= \frac{2}{\sqrt{7}} l(l-1)(l+1)(l+2)(l+3)p_0 + \frac{\sqrt{2}}{\sqrt{3}\sqrt{15}\sqrt{7}}$$
$$\times(l-1)(l+2)(l+3)(l-14)(p_{+1}l_- + p_{-1}l_+)$$
$$-2\frac{\sqrt{2}}{\sqrt{3}\sqrt{5}}(l+3)(l^2+3l-13)(p_{+2}l_-^2 + p_{-2}l_+^2)$$
$$-\frac{1}{\sqrt{5}}(l^2-7l-48)(p_{+3}l_-^3 + p_{-3}l_+^3)$$
$$+\frac{\sqrt{2}}{\sqrt{5}}(l+11)(p_{+4}l_-^4 + p_{-4}l_+^4)$$
$$+(p_{+5}l_-^5 + p_{-5}l_+^5), \tag{2.5}$$

$$P_l^{+2}/(l+1)(l+2) = -\frac{\sqrt{2}\sqrt{7}}{\sqrt{3}\sqrt{5}}$$
$$\times(l-1)(l+2)(l+3)(l+4)(p_{+1}l_- - p_{-1}l_+)$$
$$\frac{\sqrt{2}}{\sqrt{3}\sqrt{5}}(l+3)(l+4)(2l-11)(p_{+2}l_-^2 - p_{-2}l_+^2)$$
$$+\frac{1}{\sqrt{5}}(l+4)(l+17)(p_{+3}l_-^3 - p_{-3}l_+^3)$$
$$+\frac{\sqrt{2}}{\sqrt{5}}(2l+13)(p_{+4}l_-^4 - p_{-4}l_+^4)$$
$$+(p_{+5}l_-^5 - p_{-5}l_+^5), \tag{2.6}$$

$$P_l^{+3}/(l+1)(l+2)(l+3) = -\frac{2\sqrt{7}}{3}$$
$$\times l(l+1)(l+2)(l+3)(l+4)p_0$$
$$-\frac{\sqrt{2}\sqrt{7}}{\sqrt{3}\sqrt{5}}(l+2)(l-3)(l+3)(l+4)(p_{+1}l_- + p_{-1}l_+)$$
$$+\frac{2\sqrt{2}}{\sqrt{3}\sqrt{5}}(l+3)(l+4)(l+12)(p_{+2}l_-^2 + p_{-2}l_+^2)$$
$$+\frac{1}{3\sqrt{5}}(l+4)(13l+81)(p_{+3}l_-^3 + p_{-3}l_+^3)$$
$$+\frac{\sqrt{2}}{\sqrt{5}}(3l+16)(p_{+4}l_-^4 + p_{-4}l_+^4)$$
$$+(p_{+5}l_-^5 + p_{-5}l_+^5), \tag{2.7}$$

$$P_l^{+4}/(l+1)(l+2)(l+3)(l+4) = \frac{\sqrt{2}\sqrt{3}\sqrt{7}}{\sqrt{5}}$$
$$\times(l+2)(l+3)(l+4)(l+5)(p_{+1}l_- - p_{-1}l_+)$$
$$+\frac{4\sqrt{2}\sqrt{3}}{\sqrt{5}}(l+3)(l+4)(l+5)(p_{+2}l_-^2 - p_{-2}l_+^2)$$
$$+\frac{9}{\sqrt{5}}(l+4)(l+5)(p_{+3}l_-^3 - p_{-3}l_+^3)$$
$$+\frac{4\sqrt{2}}{\sqrt{5}}(l+5)(p_{+4}l_-^4 - p_{-4}l_+^4)$$
$$+(p_{+5}l_-^5 - p_{-5}l_-^5), \tag{2.8}$$

$$P_l^{+5}/(l+1)(l+2)(l+3)(l+4)(l+5)$$
$$= 6\sqrt{7}(l+1)(l+2)(l+3)(l+4)(l+5)p_0$$
$$+\sqrt{2}\sqrt{3}\sqrt{5}\sqrt{7}$$
$$\times(l+2)(l+3)(l+4)(l+5)(p_{+1}l_- + p_{-1}l_+)$$
$$+2\sqrt{2}\sqrt{3}\sqrt{5}(l+3)(l+4)(l+5)(p_{+2}l_-^2 + p_{-2}l_+^2)$$
$$+3\sqrt{5}(l+4)(l+5)(p_{+3}l_-^3 + p_{-3}l_+^3)$$
$$+\sqrt{2}\sqrt{5}(l+5)(p_{+4}l_-^4 + p_{-4}l_+^4)$$
$$+(p_{+5}l_-^5 + p_{-5}l_+^5). \tag{2.9}$$

3. THE PRODUCT OPERATORS AND THEIR MUTUAL RELATIONS

With the aid of the introduced P_l^k, *various scalar* R(3) operators which obviously commute with L^2 and l_0 can be constructed. As in the quadrupole case[3] we shall be concerned with the ones of the type $P_{l+k}^{-k}P_l^{+k}$. Because of their scalar character these quadratic operators must be expressible in terms of the other available scalar operators, i.e., L^2, l_0, P_l^0, and the G_2 second-order Casimir operator defined by[8]

$$V^* = -\tfrac{1}{28}L^2 - \sum_{\mu=-5}^{5}(-1)^\mu p_\mu p_{-\mu}, \tag{3.1}$$

having an eigenvalue of $-\frac{1}{3}v(v+5)$. (In the G_2 group a sixth-order Casimir operator exists. For symmetric representations however it is not independent of V^*.) This result can be derived from the general expression given by Judd[8] for an arbitrary irreducible representation. In Appendix A we present a derivation for that eigenvalue by making use of pure Racah algebraic techniques.

The quadratic product operators $P_{l+k}^{-k}P_l^{+k}$ consist of terms composed of two p_μ and ten or less l_i operators. In order to reach a one-to-one relation between all before-mentioned scalar operators it is clear that all operators should be brought into a so-called standard form. The procedure to reach that standard form has been discussed in I. To perform that operation explicit use has been made of the commutation relations between the several components of the p-operator. These relations are summarized in Appendix B. It was now rather straightforward to observe that in order to acheive proper relations between the various operators one needs six of the eleven product operators. By this it also follows that among the various relations which can be constructed only six independent ones exist. Since we are interested in the P_l^0 eigenvalues, we have retained in each relation the $(P_l^0)^2$ term. For the other five product operators we have chosen in a first relation $k = 1, 2, 3, 4$, and 5 in a second relation $k = 1, 2, 3, 4, -1$, and in a third relation $k = 1, 2, 3, -1, -2$. By using afterwards the fact that every P_l^{+k} goes over in a P_l^{-k} on replacing l by $-(l+1)$, one easily deduces from the first three constructed relations three other useful equations. The following relations could be finally withdrawn:

$$3(2l+7)^2(P_l^0)^2 + \frac{8}{7\sqrt{3}}(l+1)(l+2)(l+3)(2l+3)(2l+5)(2l+7)^2\ P_l^0 + \frac{5(2l+3)(2l+17)}{(l+1)^2} P_{l+1}^{-1} P_l^{+1}$$
$$+ \frac{20}{7}\frac{(l+9)(2l+5)(2l+17)}{(l+1)(l+2)^2(l+4)^2} P_{l+2}^{-2} P_l^{+2} + \frac{15}{7}\frac{(l+6)(2l+3)(2l+7)(2l^2+27l+151)}{(l+1)(l+2)^2(l+3)^2(l+4)^2(2l+9)^2} P_{l+3}^{-3} P_l^{+3}$$
$$+ \frac{5}{21}\frac{(2l+3)(2l+17)(2l^3+27l^2+166l+435)}{(l+1)(l+2)(l+3)^2(l+4)^4(l+5)^2(2l+9)} P_{l+4}^{-4} P_l^{+4}$$
$$+ \frac{1}{21}\frac{(2l+3)(2l+5)(2l^4+39l^3+298l^2+1101l+1890)}{(l+1)(l+2)(l+3)^2(l+4)^4(l+5)^4(2l+9)^2} P_{l+5}^{-5} P_l^{+5}$$
$$+ \frac{16}{21}(l+1)^2(l+2)^2(l+3)^2(2l+3)^2(2l+5)^2(2l+7)^2 V^*$$
$$+ \frac{4}{147} l(l+1)^2(l+2)^2(l+3)^2(2l+3)^2(2l+5)^2(2l+7)^2(l+16) = 0, \tag{3.2}$$

$$9(l+3)(2l+1)(2l+7)(4l^3+56l^2+283l+630)(P_l^0)^2$$
$$- \frac{8}{35\sqrt{3}}\frac{l(l+1)(l+2)(l+3)(2l+1)(2l+3)(2l+5)(2l+7)}{(l+5)}(4l^5+88l^4+791l^3+3692l^2+8655l+9450)\ P_l^0$$
$$- \frac{l(2l+3)(4l^5+60l^4-105l^3-5135l^2-29544l-60480)}{(l+1)^2} P_{l+1}^{-1} P_l^{+1}$$
$$- \frac{4l(2l+1)(2l+5)(2l+9)}{(l+1)^2(l+2)^2(l+4)(l+5)(2l+7)}(4l^5+64l^4+251l^3-874l^2-8715l-18900)\ P_{l+2}^{-2} P_l^{+2}$$
$$- \frac{18l(2l+1)}{(l+1)(l+2)^2(l+3)^2(l+4)^2}(4l^4+46l^3+116l^2-271l-1470)\ P_{l+3}^{-3} P_l^{+3}$$
$$+ \frac{6l(2l+1)(2l+3)}{(l+1)(l+2)^2(l+3)^2(l+4)^4(l+5)(2l+7)}(-8l^4-86l^3-277l^2-49l+1050)\ P_{l+4}^{-4} P_l^{+4}$$
$$+ \frac{(l+3)^2(2l+5)(2l+7)}{l^2(l+5)}(2l^4+39l^3+298l^2+1101l+1890)\ P_{l-1}^{+1} P_l^{-1}$$
$$+96\frac{l^2(l+1)^2(l+2)^2(l+3)^2(2l+1)(2l+3)^2(2l+5)^2(2l+7)}{(l+5)} V^*$$
$$- \frac{4}{735}\frac{l^2(l+1)^2(l+2)^2(l+3)^2(2l+1)(2l+3)^2(2l+5)^2(2l+7)}{(l+5)}$$
$$\times(4l^5+88l^4+791l^3+3062l^2-795l+9450) = 0, \tag{3.3}$$

$$(l+3)(2l+5)(24l^6+260l^5+326l^4-5085l^3-17417l^2+2362l+47880)(P_l^0)^2$$
$$- \frac{8}{315\sqrt{3}} l(l+1)(l+2)(l+3)(2l-1)(2l+3)(2l+5)$$
$$\times(24l^7+292l^6+1326l^5+2947l^4+4965l^3+3196l^2-22200l-56700)\ P_l^0$$
$$- \frac{(l+3)(2l-1)(2l+3)}{9(l+1)^2(2l+1)}(12l^7+100l^6-979l^5-8721l^4+2434l^3+111404l^2+85800l-315000)\ P_{l+1}^{-1} P_l^{+1}$$
$$+ \frac{2l(2l-1)(2l+5)}{9(l+1)^2(l+2)^2(2l+1)}(-32l^6-164l^5+1040l^4+3905l^3-10308l^2-26541l+40500)\ P_{l+2}^{-2} P_l^{+2}$$
$$- \frac{2l(l-1)(2l-1)}{(l+1)^2(l+2)^2(l+3)^2}(4l^4+8l^3-49l^2-53l+300)\ P_{l+3}^{-3} P_l^{+3}$$
$$- \frac{(l+3)^2(2l-1)(2l+5)}{9l^2(l-1)(2l+1)}(-12l^7-160l^6-75l^5+7133l^4+36063l^3+43517l^2-75126l-143640)\ P_{l-1}^{+1} P_l^{-1}$$
$$- \frac{2(l+2)(l+3)^2(2l+3)(2l+5)^2}{9l^2(l-1)^3(2l+1)}(8l^4+86l^3+277l^2+49l-1050)\ P_{l-2}^{+2} P_l^{-2}$$
$$- \frac{32}{3} l^2(l+1)^2(l+2)^2(l+3)^2(l-1)(2l-1)^2(2l+3)^2(2l+5)^2 V^*$$
$$- \frac{4}{2205} l^2(l+1)^2(l+2)^2(l+3)^2(l-1)(2l-1)^2(2l+3)^2(2l+5)^2(4l^5+36l^4+7l^3-711l^2-386l-4200) = 0, \tag{3.4}$$

$$
\begin{aligned}
&3(2l-5)^2(P_l^0)^2-\frac{8}{7\sqrt{3}}\,l(l-1)(l-2)(2l-1)(2l-3)(2l-5)^2\,P_l^0+\frac{5(2l-1)(2l-15)}{l^2}\,P_{l-1}^{+1}P_l^{-1}\\
&\quad+\frac{20}{7}\,\frac{(l-8)(2l-3)(2l-15)}{l(l-1)^2(l-3)^2}\,P_{l-2}^{+2}P_l^{-2}+\frac{15}{7}\,\frac{(l-5)(2l-1)(2l-5)(2l^2-23l+126)}{l(l-1)^2(l-2)^2(l-3)^2(2l-7)^2}\,P_{l-3}^{+3}P_l^{-3}\\
&\quad+\frac{5}{21}\,\frac{(2l-1)(2l-15)(2l^3-21l^2+118l-294)}{l(l-1)(l-2)^2(l-3)^4(l-4)^2(2l-7)}\,P_{l-4}^{+4}P_l^{-4}\\
&\quad+\frac{1}{21}\,\frac{(2l-1)(2l-3)(2l^4-31l^3+193l^2-614l+1050)}{l(l-1)(l-2)^2(l-3)^4(l-4)^4(2l-7)^2}\,P_{l-5}^{+5}P_l^{-5}\\
&\quad+\frac{16}{21}\,l^2(l-1)^2(l-2)^2(2l-1)^2(2l-3)^2(2l-5)^2V^*\\
&\quad+\frac{4}{147}\,(l+1)l^2(l-1)^2(l-2)^2(2l-1)^2(2l-3)^2(2l-5)^2(l-15)=0,
\end{aligned}
\tag{3.5}
$$

$$
\begin{aligned}
&9(l-2)(2l+1)(2l-5)(4l^3-44l^2+183l-399)(P_l^0)^2\\
&\quad-\frac{8}{35\sqrt{3}}\,\frac{(l+1)l(l-1)(l-2)(2l+1)(2l-1)(2l-3)(2l-5)}{(l-4)}\,(4l^5-68l^4+479l^3-1807l^2+3312l-3780)\,P_l^0\\
&\quad+\frac{(l+1)(2l-1)(4l^5-40l^4-305l^3+4500l^2-19809l+35910)}{l^2}\,P_{l-1}^{+1}P_l^{-1}\\
&\quad-\frac{4(l+1)(2l+1)(2l-3)(2l-7)}{l^2(l-1)^2(l-3)(l-4)(2l-5)}\,(4l^5-44l^4+35l^3+1283l^2-6450l+11250)\,P_{l-2}^{+2}P_l^{-2}\\
&\quad+\frac{18(l+1)(2l+1)}{l(l-1)^2(l-2)^2(l-3)^2}\,(4l^4-30l^3+2l^2+381l-1125)\,P_{l-3}^{+3}P_l^{-3}\\
&\quad+\frac{6(l+1)(2l+1)(2l-1)}{l(l-1)^2(l-2)^2(l-3)^4(l-4)(2l-5)}\,(-8l^4+54l^3-67l^2-279l+900)\,P_{l-4}^{+4}P_l^{-4}\\
&\quad-\frac{(l-2)^2(2l-3)(2l-5)}{(l+1)^2(l-4)}\,(2l^4+31l^3+193l^2-614l+1050)\,P_{l+1}^{-1}P_l^{+1}\\
&\quad-\frac{96(l+1)^2l^2(l-1)^2(l-2)^2(2l+1)(2l-1)^2(2l-3)^2(2l-5)}{(l-4)}\,V^*\\
&\quad-\frac{4}{735}\,\frac{(l+1)^2l^2(l-1)^2(l-2)^2(2l+1)(2l-1)^2(2l-3)^2(2l-5)}{(l-4)}\\
&\quad\times(4l^5-68l^4+479l^3-1177l^2-4878l-12600)=0,
\end{aligned}
\tag{3.6}
$$

$$
\begin{aligned}
&(l-2)(2l-3)(24l^6-116l^5-614l^4+4269l^3-2446l^2-21793l+33276)(P_l^0)^2\\
&\quad-\frac{8}{315\sqrt{3}}\,(l+1)l(l-1)(l-2)(2l+3)(2l-1)(2l-3)\\
&\quad\times(24l^7-124l^6+78l^5+143l^4+1437l^3+3401l^2-20439l+34380)\,P_l^0\\
&\quad+\frac{(l-2)(2l+3)(2l-1)}{9l^2(2l+1)}\,(12l^7-16l^6-1327l^5+2746l^4+25948l^3-62814l^2-100233l+299484)\,P_{l-1}^{+1}P_l^{-1}\\
&\quad+\frac{2(l+1)(2l+3)(2l-3)}{9l^2(l-1)^2(2l+1)}\,(-32l^6-28l^5+1380l^4+1255l^3-14623l^2-1002l+54000)\,P_{l-2}^{+2}P_l^{-2}\\
&\quad+\frac{2(l+1)(l+2)(2l+3)}{l^2(l-1)^2(l-2)^2}\,(4l^4+8l^3-49l^2-53l+300)\,P_{l-3}^{+3}P_l^{-3}\\
&\quad-\frac{(l-2)^2(2l+3)(2l-3)}{9(l+1)^2(l+2)(2l+1)}\,(12l^7-76l^6-633l^5+5528l^4-9561l^3-23272l^2+82002l-54000)\,P_{l+1}^{-1}P_l^{+1}\\
&\quad-\frac{2(l-1)(l-2)^2(2l-1)(2l-3)^2}{9(l+1)^2(l+2)^3(2l+1)}\,(8l^4-54l^3+67l^2+279l-900)\,P_{l+2}^{-2}P_l^{+2}\\
&\quad+\frac{32}{3}\,(l+1)^2l^2(l-1)^2(l-2)^2(l+2)(2l+3)^2(2l-1)^2(2l-3)^2V^*\\
&\quad-\frac{4}{2205}\,(l+1)^2l^2(l-1)^2(l-2)^2(l+2)(2l+3)^2(2l-1)^2(2l-3)^2\\
&\quad(4l^5-16l^4-97l^3+556l^2+933l+4500)=0.
\end{aligned}
\tag{3.7}
$$

Once more we like to insist on the fact that these six relations are only valid when they are acting to the left upon $m = 0$ states.

4. DISCUSSION

The detailed properties of the shift operators considered in this paper, although rather technical, will prove to be extremely useful in a following paper in obtaining the eigenvalues of P_l^0. However, due to the specific structure of the P_l^0 operator (2.4), it is clear that for all $l = 0$, $l = 1$, and $l = 2$ octupole-phonon states, its eigenvalue will become zero. Since degenerated states with these total angular momenta already occur for $v = 6$,[6] it is evident that it is not possible to solve the octupole state labeling problem with the proposed P_l^0 operator alone. Therefore in the next paper a second R(3) scalar operator built up with the help of a 7-dimensional tensor representation belonging to the R(7) group, will be introduced.

APPENDIX A: THE EIGENVALUE OF V* (3.1)

The Casimir operator V^*, as given by (3.1), can, due to (1.4) and (1.5), be denoted as

$$V^* = \sum_{k=0}^{1} \sqrt{(8k+1)}\,[(b_3^+ b_3)^{4k+1}(b_3^+ b_3)^{4k+1}]^0,$$

which, with the help of Racah algebra, can be brought into the following form:

$$V^* = -\sum_{\Gamma=0}^{3} \cdot \sum_{k=0}^{1} (8k+3)\sqrt{4\Gamma+1} \times \begin{Bmatrix} 3 & 3 & 4k+1 \\ 3 & 3 & 2\Gamma \end{Bmatrix} [(b_3^+ b_3^+)^{2\Gamma}(b_3 b_3)^{2\Gamma}]^0 - \sum_{k=0}^{1} (8k+3)^{3/2} \begin{Bmatrix} 4k+1 & 4k+1 & 0 \\ 3 & 3 & 3 \end{Bmatrix} (b_3^+ b_3)^0.$$

From De-Shalit and Talmi[9] one deduces

$$\sum_{k=0}^{1} (8k+3) \begin{Bmatrix} 3 & 3 & 4k+1 \\ 3 & 3 & 2\Gamma \end{Bmatrix} = \frac{1}{2} - \frac{7}{2}\delta_{2\Gamma,0} - 7\begin{Bmatrix} 3 & 3 & 3 \\ 3 & 3 & 2\Gamma \end{Bmatrix}$$

and

$$\sum_{k=0}^{1} (8k+3)^{3/2} \begin{Bmatrix} 4k+1 & 4k+1 & 0 \\ 3 & 3 & 3 \end{Bmatrix} = -2\sqrt{7}.$$

Using expression (37) of Weber *et al.*,[7] i.e.,

$$N^2 = N + \sum_{\Gamma} \sqrt{4\Gamma+1}\,[(b_3^+ b_3^+)^{2\Gamma}(b_3 b_3)^{2\Gamma}]^0,$$

and the previously mentioned relations, V^* can be transformed to

$$V^* = -\tfrac{1}{2}N(N+3) + \tfrac{1}{2}V + 7\sum_{\Gamma} \sqrt{4\Gamma+1} \begin{Bmatrix} 3 & 3 & 3 \\ 3 & 3 & 2\Gamma \end{Bmatrix} [(b_3^+ b_3^+)^{2\Gamma}(b_3 b_3)^{2\Gamma}]^0.$$

Here V is the seniority operator introduced by Weber *et al.*[7] as $7[(b_3^+ b_3^+)^0(b_3 b_3)^0]^0$, having an eigenvalue of $(N-v)(N+v+5)$. Due to the fact that the 6-*j* symbols

$$\begin{Bmatrix} 3 & 3 & 3 \\ 3 & 3 & 2 \end{Bmatrix} = \begin{Bmatrix} 3 & 3 & 3 \\ 3 & 3 & 4 \end{Bmatrix} = \begin{Bmatrix} 3 & 3 & 3 \\ 3 & 3 & 6 \end{Bmatrix} = \frac{1}{2\times 3\times 7}$$

are equal,[10] the Casimir operator V^* can be totally expressed in terms of N and V:

$$V^* = -\tfrac{1}{3}(N^2 + 5N - V),$$

whose eigenvalue due to the known eigenvalue of V becomes $-\tfrac{1}{3}v(v+5)$.

APPENDIX B: THE COMMUTATION RELATIONS $[p_\mu, p_\nu]$

$$[p_0, p_{\pm 1}] = \pm \frac{2}{\sqrt{3}\sqrt{7}} p_{\pm 1} - \frac{\sqrt{3}\sqrt{5}}{28\sqrt{2}} l_{\pm},$$

$$[p_{\pm 1}, p_{\pm 2}] = \pm \frac{\sqrt{5}}{\sqrt{3}\sqrt{7}} p_{\pm 3},$$

$$[p_0, p_{\pm 2}] = \pm \frac{2}{\sqrt{3}\sqrt{7}} p_{\pm 2},$$

$$[p_{\pm 1}, p_{\pm 3}] = \pm \frac{\sqrt{5}}{2\sqrt{7}} p_{\pm 4},$$

$$[p_0, p_{\pm 3}] = \mp \frac{1}{2\sqrt{3}\sqrt{7}} p_{\pm 3},$$

$$[p_{\pm 1}, p_{\pm 4}] = \mp \frac{3}{2\sqrt{7}} p_{\pm 5},$$

$$[p_0, p_{\pm 4}] = \mp \frac{\sqrt{3}}{\sqrt{7}} p_{\pm 4},$$

$$[p_{\pm 1}, p_{\mp 3}] = \pm \frac{\sqrt{5}}{\sqrt{3}\sqrt{7}} p_{\mp 2},$$

$$[p_0, p_{\pm 5}] = \pm \frac{\sqrt{3}}{2\sqrt{7}} p_{\pm 5},$$

$$[p_{\pm 1}, p_{\mp 4}] = \pm \frac{\sqrt{5}}{2\sqrt{7}} p_{\mp 3},$$

$$[p_{-1}, p_{+1}] = \frac{2}{\sqrt{3}\sqrt{7}} p_0 + \frac{1}{28} l_0,$$

$$[p_{\pm 1}, p_{\mp 5}] = \mp \frac{3}{2\sqrt{7}} p_{\mp 4},$$

$$[p_{-2}, p_{+2}] = \frac{-2}{\sqrt{3}\sqrt{7}} p_0 - \frac{1}{14} l_0,$$

$$[p_{\pm 2}, p_{\pm 3}] = \pm \frac{1}{\sqrt{2}} p_{\pm 5},$$

$$[p_{-3}, p_{+3}] = -\frac{1}{2\sqrt{3}\sqrt{7}} p_0 + \frac{3}{28} l_0,$$

$$[p_{\pm 2}, p_{\mp 5}] = \mp \frac{1}{\sqrt{2}} p_{\mp 3},$$

$$[p_{-4}, p_{+4}] = \frac{\sqrt{3}}{\sqrt{7}} p_0 - \frac{1}{7} l_0,$$

$$[p_{\pm 3}, p_{\mp 5}] = \mp \frac{1}{\sqrt{2}} p_{\mp 2},$$

$$[p_{-5}, p_{+5}] = \frac{\sqrt{3}}{2\sqrt{7}} p_0 + \frac{5}{28} l_0,$$

$$[p_{\pm 1}, p_{\pm 5}] = [p_{\pm 2}, p_{\pm 4}] = 0,$$

$$[p_{\pm 1}, p_{\mp 2}] = \frac{1}{4\sqrt{7}} l_{\mp},$$

$$[p_{\pm 2}, p_{\mp 4}] = [p_{\pm 2}, p_{\pm 5}] = 0, \quad [p_{\pm 3}, p_{\pm 4}] = 0,$$

$$[p_{\pm 2}, p_{\mp 3}] = \pm \frac{\sqrt{5}}{\sqrt{3}\sqrt{7}} p_{\mp 1} - \frac{\sqrt{3}}{14\sqrt{2}} l_{\mp},$$

$$[p_{\pm 3}, p_{\mp 4}] = \mp \frac{\sqrt{5}}{2\sqrt{7}} p_{\mp 1} + \frac{3}{28\sqrt{2}} l_{\mp},$$

$$[p_{\pm 4}, p_{\mp 5}] = \mp \frac{3}{2\sqrt{7}} p_{\mp 1} - \frac{\sqrt{5}}{28\sqrt{2}} l_{\mp},$$

$$[p_{\pm 3}, p_{\pm 5}] = 0, \quad [p_{\pm 4}, p_{\pm 5}] = 0.$$

[1]J. W. B. Hughes, J. Phys. A: Math. Nucl. Gen. **6**, 48 (1973).
[2]J. W. B. Hughes and J. Yadegar, J. Math. Phys. **19**, 2068 (1978), and Refs. quoted therein.
[3]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **21**, 1902 (1980).
[4]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1906 (1980).
[5]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1913 (1980).
[6]S. G. Rohozinski, J. Phys. G: Nucl. Phys. **4**, 1075 (1978).
[7]H. J. Weber, M. G. Huber, and W. Greiner, Z. Phys. **190**, 25 (1966).
[8]B. R. Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw-Hill, New York, 1963).
[9]A. De-Shalit and I. Talmi, *Nuclear Shell Theory* (Academic, New York, 1963).
[10]M. Rotenberg, R. Bivins, N. Metropolis, J. K. Wooten, Jr., *The 3-j and 6-j Symbols* (Tech. P., M.I.T., Cambridge, Mass., 1959).

Shift operator techniques for the classification of multipole-phonon states. V. Properties of shift operators in the $R(7)$ group

H. E. De Meyer[a)] and G. Vanden Berghe
Seminarie voor Wiskundige Natuurkunde, R. U. G., Krijgslaan 271-S9, B-9000 Gent, Belgium

(Received 4 March 1980; accepted for publication 4 April 1980)

With a view to obtaining additional label-generating operators for the classification of octupole-phonon states a set of shift operators O_l^k $(k = 0, \pm 1, \pm 2, \pm 3)$ in the $R(7)$ group is constructed. Expressions which connect quadratic products of these shift operators are given, and it turns out that besides $R(7)$ invariants the expressions also involve the scalar G_2 shift operator P_l^0 previously studied. The opportunity to arrive at an orthogonal solution of the state labelling problem is discussed.

1. INTRODUCTION

In analogy to the quadrupole state labelling problem, which we treated rigorously by means of ths shift operator technique in the first three papers of this series[1-3] (to be referred to as I, II, and III respectively), we investigate the possibility of using $R(3)$ scalar shift operators in the sense of Hughes and Yadegar[4] as labelling operators for the classification of octupole-phonon states. It is well known that the latter may be viewed at as symmetric representation states of the unitary group $U(7)$. Recently, an orthonormalized set of independent octupole N-phonon states has been explicitly constructed for $N \leqslant 5$ by the present authors.[5] However, for classifying in a complete general way all octupole-phonon states, seven quantum numbers are needed. Obviously, four of them are immediately witheld as the ones related to the Casimir operators of the groups $U(7)$, $R(7)$, $R(3)$, and $R(2)$ appearing in the chain $U(7) \supset R(7) \supset G_2 \supset R(3) \supset R(2)$, i.e., the boson number N, the seniority v, the angular momentum l, and its projection, m. The Casimir operators of the special group G_2 in this chain do not provide us with additional labelling operators. Indeed, it has been remarked in the previous paper[6] (to be referred to as IV) that the sixth order G_2-Casimir operator is not independent from the second order one if we are looking only at symmetric representations, whereas we shall demonstrate further that the eigenvalues of the second order operator are, upon a multiplying factor, identical to the eigenvalues of the $R(7)$ Casimir operator and thus also quadratic in the seniority v.

Nevertheless, it was also proved in IV that the scalar shift operator P_l^0 built from G_2 generators p_μ $(\mu = 0, \pm 1, \pm 2, \pm 3, \pm 4, \pm 5)$ and $R(3)$ generators l_+, l_-, l_0, defined in (IV.1.5) and (IV.1.4), respectively, can serve as a suitable fifth label generating operator. As a natural extension we propose also to construct shift operators in the $R(7)$-group, hence containing in the expansion also generators of the type q_μ $(\mu = 0, \pm 1, \pm 2, \pm 3)$, defined by

$$q_\mu = (b_3^+ b_3)^3_\mu \quad (\mu = 0, \pm 1, \pm 2, \pm 3). \tag{1.1}$$

Herein b_{3m}^+ and $(-1)^m b_{3-m}$ are the octupole phonon creation and annihilation operators respectively.

A priori there is a large variety of polynomial forms of the l-, p-, and q-generators on which we can imply the condition of becoming an $R(3)$-scalar—or otherwise stated, the condition of commuting with the operator L^2—such that it also becomes a suitable label generating operator. (This clearly corresponding to the fact that there are also many scalar $Q(k_1k_2k_3k_4k_5)$ operators which have been studied from a numerical point of view elsewhere.[7] However, if we like to keep the analogy with the quadrupole case as much as possible, it is evident we must first analyze the shift operator forms which contain the l- and q-generators only and which in addition are linear in the q-generators. Now, the explicit construction of such shift operators is considerably simplified by remarking that in spite of the fact that the internal structure of the q-generators (1.1) differs from that of the quadrupole q-generators given in (I.1.7), and also in spite of the fact that the l-generators are defined in a slightly different ways, as may be seen by comparing (IV.1.4) to (I.1.7), nevertheless the commuting properties of the l's among each other and of the l's with the q's remain valid—in both the quadrupole and octupole cases the q's form a seven-dimensional tensor representation of the $R(3)$-group indeed. Hence, since the shift operators are found by the use of these commutation properties and without any reference to a possible internal structure, we can formally keep for these operators the same forms O_l^k with $k = 0, \pm 1, \pm 2, \pm 3$ as derived in the quadrupole case, namely in (I.2.1)–(I.2.5). It has to be noticed that these expressions still exhibit the l_0-dependence, but that by using them only acting on states with zero angular momentum projection, as we shall do in the next section, l_0 may be set equal to zero.

Furthermore, it has to be remarked in advance that on account of the commuting properties of the q-generators, properties which are summarized in the Appendix, p-generators can be introduced when calculating quadratic forms of the O_l^k shift operators in the aim of constructing relations among such forms and octupole invariants. Since it turns out that the p's can only occur linearly in relations between quadratic scalar products of the shift operators O_l^k, it is easy to

[a)]Bevoegdverklaard Navorser N.F.W.O. (Belgium).

predict that the operator P_l^0 defined in IV will enter the relations.

2. THE PRODUCT OPERATORS AND THEIR MUTUAL RELATIONS

Since the shift operators O_l^k ($k = 0, \pm 1, \pm 2, \pm 3$) are in form identical to the quadrupole shift operators (I.2.1)–(I.2.5), it is clear that we should be able to take profit of the results obtained in I when setting up relations connecting quadratic products of the shift operators with group invariants. In fact, it is easily verified that we can copy immediately the coefficients occurring with the combinations $O_{l+k}^{-k} O_l^{+k}$ ($k = 0, \pm 1, \pm 2, \pm 3$) in the quadrupole relations (I.3.2)–(I.3.5). Furthermore, by comparing the q-commutators in the Appendix of I with those in the Appendix here, and looking especially at the q-dependency of these commutators, we deduce that the coefficients associated to the terms linear in O_l^0 in the relations of I, have only to be altered by a numerical factor. Looking then at the l_+- and l_--dependency of the q-commutators a similar reasoning applies for the constant terms of the relations. Finally, we can even make use of the dependency of the quadrupole relations on the $R(5)$ Casimir operator V^* by simply substituting now for V^* the analogue $\bar{V}^*$ which is defined as follows:

$$\bar{V}^* = -\tfrac{1}{28} L^2 + \sum_{\mu=-3}^{+3} (-1)^{\mu+1} q_{-\mu} q_\mu. \tag{2.1}$$

However, $\bar{V}^*$ is obviously not the $R(7)$ quadratic Casimir operator since the latter one is given by

$$C_2(R(7)) = \sum_{\mu=-3}^{3} (-1)^{\mu+1} q_{-\mu} q_\mu + \sum_{\mu=-5}^{5} (-1)^{\mu+1} p_{-\mu} p_\mu - \tfrac{1}{28} L^2, \tag{2.2}$$

and therefore contains in addition to (2.1) a part which is bilinear in the p-generators. Hence, it may seen at first sight impossible to introduce in the octupole relations which we are constructing the Casimir operator (2.2) without adding a supplementary term which cancels the quadratic p-terms of this operator. To perform the cancellation, let us remember that the second-order Casimir operator of the group G_2 which is a subgroup of $R(7)$, is, apart from a term in L^2, exactly of the form of a bilinear expression in the p-generators, namely

$$C_2(G_2) = \sum_{\mu=-5}^{5} (-1)^{\mu+1} p_{-\mu} p_\mu - \tfrac{1}{28} L^2, \tag{2.3}$$

and is also an $R(7)$-invariant. Combining the operators (2.2) and (2.3) we deduce that $\bar{V}$ may be written as

$$\bar{V} = -\tfrac{1}{28} L^2 + C_2(R(7)) - C_2(G_2). \tag{2.4}$$

Now, it is a lucky coincidence that the eigenvalues of $C_2(R(7))$ and $C_2(G_2)$ are the same upon a numerical factor. Indeed, when acting on a state with seniority v, $C_2(R(7))$ yields the eigenvalue $-v(v+5)/2$ whereas $C_2(G_2)$ yields the eigenvalue $-v(v+5)/3$. Consequently, the operator $\bar{V}^*$ is an $R(7)$ invariant producing the eigenvalue $-l(l+1)/28 - v(v+5)/6$ on a state with seniority v and angular momentum l. This result allows us to replace the operator $\bar{V}^*$ in the relations under construction formally by the operator $-L^2/28 + [C_2(R(7))]/3$, since both yield the same eigenvalues. For convenience, and to conform with notations elsewhere,[7] we shall use from here on the notation V^* for $C_2(R(7))$, V^* thus representing the $R(7)$-Casimir operator to which the eigenvalues $-v(v+5)/2$ are associated.

Resuming, we can take maximum profit of the relations obtained in the quadrupole situation by replacing some numerical coefficients and by replacing the V^* therein formally by $-L^2/28 + V^*/3$, although both V^* have a different meaning. However, by this we have not finished the construction of the octupole relations yet. Indeed, as we have already remarked at the end of the introductory section, the appearance in the q-commutators of the Appendix of a part which is linear in the p-generators will give rise to the introduction of a new term in these relations which is also linear in the p's. As has been argued before, this term has to be expressible in terms of the G_2-scalar shift operator P_l^0 of IV. By straightforward calculation we have arrived at the following final results:

$$\begin{aligned}
&5(l+3)^2(2l+5)^2(O_l^0)^2 - \sqrt{\tfrac{10}{3}(l+1)(l+2)(l+3)^2(2l+3)(2l+5)^2}\; O_l^0 \\
&\quad = -\frac{15(l+3)^2(l+4)(2l+3)}{(l+1)^2} O_{l+1}^{-1} O_l^{+1} - \frac{3(l+4)(2l+5)(2l+9)}{(l+1)(l+2)^2} O_{l+2}^{-2} O_l^{+2} \\
&\quad\quad - \frac{(2l+3)(l^2+7l+15)}{(l+1)(l+2)^2(l+3)^2} O_{l+3}^{-3} O_l^{+3} - \tfrac{4}{3}(l+1)^2(l+2)^2(l+3)^2(2l+3)^2(2l+5)^2 V^* \\
&\quad\quad - \tfrac{6}{7} l(l+1)^2(l+2)^2(l+3)^2(2l+3)^2(2l+5)^2 + \sqrt{3}(l+1)(l+2)(l+3)(2l+5)(2l+3)\; P_l^0,
\end{aligned} \tag{2.5}$$

$$\begin{aligned}
&(l+2)(2l+1)(8l^2+57l+90)(O_l^0)^2 + \frac{\sqrt{2}}{3\sqrt{15}} l(l+1)(l+2)(2l+1)(2l+3)(4l^3+34l^2+87l+90)\; O_l^0 \\
&\quad = -\frac{(l+2)^2(2l+3)(l^2+7l+15)}{l^2} O_{l-1}^{+1} O_l^{-1} + \frac{l(2l+3)(l^3+5l^2-18l-63)}{(l+1)^2} O_{l+1}^{-1} O_l^{+1} \\
&\quad\quad + \frac{l(2l+1)(2l^2+2l-9)}{(l+1)^2(l+2)^2} O_{l+2}^{-2} O_l^{+2} - \tfrac{4}{3} l^2(l+1)^2(l+2)^2(2l+3)^2(2l+1) V^* \\
&\quad\quad + \tfrac{2}{105} l^2(l+1)^2(l+2)^2(2l+3)^2(2l+1)(2l^3+17l^2+6l+45) \\
&\quad\quad - \frac{1}{\sqrt{3}} l(l+1)(l+2)(l+3)(2l+3)(2l+1)\; P_l^0,
\end{aligned} \tag{2.6}$$

which are the equivalents of (I.3.2) and (I.3.5) respectively. As in the quadrupole case, the replacement of l by $-l-1$ yields two supplementary relations, namely

$$5(l-2)^2(2l-3)^2(O_l^0)^2+\sqrt{\tfrac{10}{3}}\,l(l-1)(l-2)^2(2l-1)(2l-3)^2\,O_l^0$$
$$=-15\frac{(l-2)^2(l-3)(2l-1)}{l^2}O_{l-1}^{+1}O_l^{-1}-3\frac{(l-3)(2l-3)(2l-7)}{l(l-1)^2}O_{l-2}^{+2}O_l^{-2}$$
$$-\frac{(2l-1)(l^2-5l+9)}{l(l+1)^2(l-2)^2}O_{l-3}^{+3}O_l^{-3}-\tfrac{4}{3}l^2(l-1)^2(l-2)^2(2l-1)^2(2l-3)^2V^*$$
$$+\tfrac{6}{7}l^2(l-1)^2(l-2)^2(2l-1)^2(2l-3)^2(l+1)-\sqrt{3}\,l(l-1)(l-2)(2l-1)(2l-3)\;P_l^0 \qquad (2.7)$$

and

$$(l-1)(2l+1)(8l^2-41l+41)(O_l^0)^2+\frac{\sqrt{2}}{3\sqrt{15}}l(l+1)(l-1)(2l-1)(2l+1)(4l^3-22l^2+31l-33)(O_l^0)$$
$$=\frac{(l-1)^2(2l-1)(l^2-5l+9)}{(l+1)^2}O_{l+1}^{-1}O_l^{+1}-\frac{(l+1)(2l-1)(l^3-2l^2-25l+41)}{l^2}O_{l-1}^{+1}O_l^{-1}$$
$$+\frac{(l+1)(2l+1)(2l^2+2l-9)}{l^2(l-1)^2}O_{l-2}^{+2}O_l^{-2}+\tfrac{4}{3}l^2(l+1)^2(l-1)^2(2l-1)^2(2l+1)V^*$$
$$+\tfrac{2}{105}l^2(l+1)^2(l-1)^2(2l-1)^2(2l+1)(2l^3-11l^2-22l-54)-\frac{1}{\sqrt{3}}l(l+1)(l-1)(l-2)(2l-1)(2l+1)\;P_l^0, \qquad (2.8)$$

which are the analogs of (I.3.3) and (I.3.4) respectively. The four independent relations above must allow us to calculate the O_l^0 eigenvalues once the P_l^0 eigenvalues have been found explicitly.

3. DISCUSSION

We have demonstrated how the results of I obtained in the context of the quadrupole state labelling problem can be used to find almost immediately analogous results for the present octupole situation. The most striking difference, however, is that we have to introduce here also the G_2-shift operator P_l^0 besides the $R(7)$-shift operator O_l^0. Hence, one could erroneously conclude that the occurrence of O_l^0 and P_l^0, both $R(3)$-scalars, in the same relation would imply that O_l^0 and P_l^0 commute. Nevertheless, we have numerical confirmation that this is certainly not true, and it is also straightforward to verify this by working out explicitly the commutator $[O_l^0,P_l^0]$ (one can obviously restrict oneself to the calculation of only those terms which have a unique prescribed form, for demonstrating the noncommutativity). As a consequence, we cannot diagonalize O_l^0 and P_l^0 simultaneously, and thus no set of orthogonal phonon states can be generally constructed such that they are always eigenstates of both O_l^0 and P_l^0.

Besides this fact we have to remark that, even putting aside the orthogonalisation difficulty, P_l^0 and O_l^0 are certainly not sufficient to label all the octupole-phonon states (which is certainly in accordance to the well-known fact that three labels are missing). This property can be nicely illustrated by noting that on any state with total angular momentum l equal to 0 or 1, P_l^0 and O_l^0 yield zero eigenvalues. This is verified from their explcit forms given in (IV.2.4) and (I.2.1) respectively. Now, calculating the l-degeneracies for the N-octupole-phonon states which are listed in Ref. 8, we see that there exist for instance two $N=7, l=1$ states. Since neither O_l^0 nor P_l^0 can distinguish between these states, the construction of a third operator will be necessary, and, moreover, we must have the guarantee that this operator should not yield zero eigenvalues on $l=0$ and $l=1$ states.

The construction of orthonormalized octupole phonon-states on acount of the shift-operator technique and the construction of an additional third label generating operator shall be considered in the near future.

APPENDIX: THE COMMUTATION RELATIONS $[q_\mu,q_\nu]$

$$[q_0,q_{\pm1}]=\mp\frac{1}{\sqrt6}q_{\pm1}\mp\frac{\sqrt5}{\sqrt2\sqrt3\sqrt7}p_{\pm1}-\frac{\sqrt3}{28}l_\pm,$$

$$[q_0,q_{\pm2}]=\mp\frac{1}{\sqrt6}q_{\pm2}\mp\frac{1}{\sqrt3}p_{\pm2},$$

$$[q_0,q_{\pm3}]=\pm\frac{1}{\sqrt6}q_{\pm3}\mp\frac{1}{\sqrt3}p_{\pm3},$$

$$[q_{-1},q_{+1}]=-\frac{1}{\sqrt6}q_0-\frac{5}{2\sqrt3\sqrt7}p_0+\tfrac{1}{28}l_0,$$

$$[q_{-2},q_{+2}]=+\frac{1}{\sqrt6}q_0-\frac{2}{\sqrt3\sqrt7}p_0-\tfrac{1}{14}l_0,$$

$$[q_{-3},q_{+3}]=+\frac{1}{\sqrt6}q_0-\frac{1}{2\sqrt3\sqrt7}p_0+\tfrac{3}{28}l_0,$$

$$[q_{\mp1},q_{\pm2}]=\mp\frac{3}{2\sqrt7}p_{\pm1}+\frac{\sqrt5}{28\sqrt2}l_\pm,$$

$$[q_{\mp2},q_{\pm3}]=\pm\frac{1}{\sqrt3}q_{\pm1}\mp\frac{\sqrt5}{\sqrt2\sqrt3\sqrt7}p_{\pm1}-\frac{\sqrt3}{28\sqrt2}l_\pm,$$

$$[q_{\mp1},q_{\pm3}]=\pm\frac{1}{\sqrt3}q_{\pm2}\mp\frac{1}{\sqrt6}p_{\pm2},$$

$$[q_{\mp1},q_{\mp2}]=\pm\frac{1}{\sqrt3}q_{\mp3}\mp\frac{1}{\sqrt6}p_{\mp3},$$

$$[q_{\mp2},q_{\mp3}]=\pm\frac{1}{\sqrt2}p_{\mp5},\quad [q_{\mp1},q_{\mp3}]=\pm\frac{1}{\sqrt2}p_{\mp4}.$$

[1]G. Vanden Berghe and H. De Meyer, J. Math. Phys. **21**, 1902 (1980).
[2]H. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1906 (1980).
[3]H. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1913 (1980).
[4]J. W. B. Hughes and J. Yadegar, J. Math. Phys. **19**, 2068 (1978) and Refs. quoted therein.
[5]G. Vanden Berghe and H. De Meyer, Nucl. Phys. A **323**, 302 (1979).
[6]G. Vanden Berghe and H. De Meyer, J. Math. Phys. **21**, xxxx (1980).
[7]G. Vanden Berghe and H. De Meyer (to be published).
[8]S. G. Rohozinsky, J. Phys. G: Nucl. Phys. **4**, 1075 (1978).

Generalized Bessel functions and the representation theory of U(2)Ⓢ$C^{2\times 2}$

Wayne J. Holman, III
The University of the South, Sewanee, Tennessee 37375
and Department of Mathematics, The University of North Carolina, Chapel Hill, North Carolina 27514

(Received 18 July 1979; accepted for publication 21 September 1979)

We construct the matrix elements of both finite transformations and infinitesimal generators in irreducible representations of the motion group U(2)Ⓢ $\mathbb{C}^{2\times 2}$ with the aid of the contraction limit of the analogous structures of U(4). The matrix elements of finite transformations are found to have a structure similar to that of the classical Bessel function in that they contain two inverse gamma matrices which couple Wigner D functions. An integral representation is established and related to the matrix-values Bessel functions of Gross and Kunze. By means of the representation property of the matrix elements we obtain a new sum rule for classical Bessel functions and an analog of the binomial theorem for the sum of two 2×2 matrices which involves the U(2) gamma matrix instead of the classical gamma function.

I. INTRODUCTION

This paper is part of an ongoing investigation of matrix-valued special functions associated with the representation theory of the conformal group U(2,2). In the first paper which resulted from this investigation[1] K.I. Gross and the present author announced the results of a calculation of the eigenvalues of the matrix-valued gamma function for U(2) and examined the implications for representation spaces for U(2,2). The second paper[2] presented the details of the calculation and some properties of the gamma function which are analogous to those of the classical gamma function. In this paper we shall be concerned with the generalized Bessel function which has already been extensively treated by Gross and Kunze,[3-6] who define it in terms of the integral transform

$$J_\lambda(\alpha,z) = \int_{U(n)} e^{i\mathrm{Re}\,\mathrm{Tr}(z^\dagger u\alpha)}\lambda(u^\dagger)\,du, \tag{1.1}$$

in analogy to the integral representation for the classical Bessel function

$$J_n(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-iz\sin\theta}e^{in\theta}\,d\theta. \tag{1.2}$$

In Eq. (1.1) the integral is taken over the group manifold of U(n), and $\lambda(u)$ denotes the matrix of an irreducible representation of U(n). The quantities α and z are elements of $\mathbb{C}^{n\times n}$, the vector group of complex $n\times n$ matrices under addition. Just as the classical Bessel function (1.2) is the matrix element of an irreducible representation of U(1)Ⓢ$\mathbb{C}^{1\times 1}$, so the generalized Bessel function (1.1) is a particular matrix element of an irreducible representation of the motion group U(n)Ⓢ$\mathbb{C}^{n\times n}$, the semidirect product of U(n) and $\mathbb{C}^{n\times n}$. This generalized Bessel function, however, also appears in the representation theory of the conformal group U(n,n) as the kernel of the integral operators in certain irreducible representations of U(n,n) which correspond to the element

$$\begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}, \tag{1.3}$$

where $\mathbb{1}_n$ denotes the $n\times n$ unit matrix. Hence, the Bessel transform (1.1) is a matrix-valued special function associated with the representation theory of the conformal group as well as with that of U(n)Ⓢ$\mathbb{C}^{n\times n}$.

In this paper we shall study the representation theory of U(2)Ⓢ$\mathbb{C}^{2\times 2}$ in detail; in this case the Wigner–Racah algebra is sufficiently well understood to make such a project feasible. We shall construct the matrix elements of irreducible representations of this group essentially by means of the Wigner–İnönü contraction of those of U(4), and in doing so we shall explore the technical problems involved in the simultaneous contraction both of a group and of a maximal subgroup. The contraction limit of these matrix elements in U(4) will provide us with a set of functions which we shall prove to be the matrix elements of irreducible representations in U(2)Ⓢ$\mathbb{C}^{2\times 2}$ by restricting them to infinitesimal transformations and recovering the correct matrix elements of the infinitesimal generators of the group.

In Sec. II we shall consider the geometry of the group U(4) and the specific contraction procedure which yields the group [U(2)Ⓢ$\mathbb{C}^{2\times 2}$]$\times\mathbb{H}^{2\times 2}$ where $\mathbb{H}^{2\times 2}$ denotes the group of 2×2 Hermitian matrices under addition. We shall also consider the dual object of U(4) and its relation, under contraction, to that of [U(2)Ⓢ$\mathbb{C}^{2\times 2}$]$\times\mathbb{H}^{2\times 2}$ and the classification of all the irreducible representations of U(2)Ⓢ$\mathbb{C}^{2\times 2}$.

In Sec. III we shall treat the matrix elements of the infinitesimal generators of U(2)Ⓢ$\mathbb{C}^{2\times 2}$ and establish them for the case of a general irreducible representation of the group.

In Sec. IV we obtain the matrix elements of finite transformations by a heuristic procedure based on the Wigner–İnönü contraction, the details of which are relegated to Appendix A. The proof that the resulting functions are indeed the required matrix elements of irreducible representations is sketched in Appendix B; it involves the demonstration that these functions have the correct transformation properties under the generators of the group realized as differential operators on the group manifold. We also derive some of the consequences of the representation property of these functions: a new sum rule for the classical Bessel function [Eq. (4.20)] and an analog of the binomial theorem involving the U(2) gamma matrix [Eq. (4.24)]. As a special case of this

 0022-2488/80/081977-34$1.00

generalized binomial theorem we also obtain an expansion of $[\det(\zeta + \eta)]^n$ in powers of $\det\zeta$ and $\det\eta$ with coefficients $\chi^\lambda(\zeta^{-1}\eta)$, where ζ and η are 2×2 matrices and χ^λ is a primitive character of a finite-dimensional representation of GL(2,ℂ) [Eq. (4.25)]. We also establish the relationship between our matrix elements and the Bessel functions of Gross and Kunze. We find that the matrix elements of irreducible representations in $U(2)\circledS\mathbb{C}^{2\times2}$ have a structure analogous to that of the classical Bessel function [the matrix element of an irreducible representation in $U(1)\circledS\mathbb{C}^{1\times1}$], in which inverse U(2) gamma matrices replace the inverse classical gamma functions in the series expansion.

In Sec. V we discuss the realization of the infinitesimal operators of the group $U(2)\circledS\mathbb{C}^{2\times2}$ as differential operators on the group manifold, and we obtain the invariant differential operators of which our matrix elements are eigenfunctions.

The results of this paper will be extended and presented in a more abstract form in a forthcoming paper.[7]

II. THE GROUPS U(4) AND $[U(2)\circledS\mathbb{C}^{2\times2}]\times\mathbb{H}^{2\times2}$ AND THE CONTRACTION PROCESS

We first consider the geometry of the group U(4) and the specific contraction procedure which yields the group $[U(2)\circledS\mathbb{C}^{2\times2}]\times\mathbb{H}^{2\times2}$. We obtain the latter group by restricting the former to transformations of the type

$$\lim_{R\to\infty}\begin{pmatrix} U & \frac{1}{R}Z_1 \\ \frac{1}{R}Z_2 & \mathbb{1}+\frac{iH}{R}\end{pmatrix}, \tag{2.1}$$

where U, Z_1, Z_2, and H all denote 2×2 matrices. The restriction that Eq. (2.1) be the limit of a unitary matrix then implies that

$$\begin{aligned} UU^\dagger &= \mathbb{1},\\ Z_2 &= -Z_1^\dagger U,\\ H &= H^\dagger. \end{aligned} \tag{2.2}$$

The law of group multiplication for the semidirect product group $U(2)\circledS\mathbb{C}^{2\times2}$ is given by

$$(U_2,Z_2)(U_1,Z_1) = (U_2U_1, Z_2 + U_2Z_1), \tag{2.3}$$

where $U_i\in U(2)$ and $Z_i\in\mathbb{C}^{2\times2}$. The matrix multiplication law for the elements of the form (2.1) with the restrictions (2.2) then becomes

$$\lim_{R\to\infty}\begin{pmatrix} U_2 & \frac{1}{R}Z_2 \\ -\frac{1}{R}Z_2^\dagger U_2 & \mathbb{1}+\frac{i}{R}H_2\end{pmatrix}\begin{pmatrix} U_1 & \frac{1}{R}Z_1 \\ -\frac{1}{R}Z_1^\dagger U_1 & \mathbb{1}+\frac{i}{R}H_1\end{pmatrix}$$
$$= \lim_{R\to\infty}\begin{pmatrix} U_2U_1 & \frac{1}{R}(U_2Z_1+Z_2) \\ -\frac{1}{R}(Z_2^\dagger U_2U_1 - Z_1^\dagger U_1) & \mathbb{1}+\frac{i}{R}(H_1+H_2)\end{pmatrix}, \tag{2.4}$$

which tells us that the contracted group is indeed isomorphic to the direct product $[U(2)\circledS\mathbb{C}^{2\times2}]\times\mathbb{H}^{2\times2}$. In deducing the restrictions (2.2) and in deriving the group multiplication law (2.4) we have merely discarded all powers of R^{-1} higher than the first.

We should note that we have an additional degree of freedom available to us. Let Ω denote an arbitrary 2×2 matrix. We can map the group element $[(U,Z),H]$ of $[U(2)\circledS\mathbb{C}^{2\times2}]\times\mathbb{H}^{2\times2}$ onto 4×4 matrices of the form

$$\lim_{R\to\infty}\begin{pmatrix} U & \frac{1}{R}Z\Omega^\dagger \\ -\frac{1}{R}\Omega Z^\dagger U & \mathbb{1}+\frac{i}{R}H\end{pmatrix} \tag{2.5}$$

and obtain a group multiplication law isomorphic to Eq. (2.3). We shall find that the matrix Ω acts as a transformation on the dual object of the Abelian group $\mathbb{C}^{2\times2}$.

The infinitesimal generators of U(4) are simply the matrices E_{ij}, $1\leqslant i,j\leqslant4$, where E_{ij} has unity in the (i,j) place and zeroes elsewhere. The Lie algebra of the generators is therefore specified by the commutation relations

$$[E_{ij},E_{k\ell}] = E_{i\ell}\delta_{jk} - E_{kj}\delta_{i\ell}. \tag{2.6}$$

The group of transformations of the type (2.1) above is generated by the operators (in the limit $R\to\infty$)

$$E_{ij},\ \frac{1}{R}E_{iJ},\ \frac{1}{R}E_{Ij},\ \frac{1}{R}E_{IJ},\quad 1\leqslant i,j\leqslant2,\ 3\leqslant I,J\leqslant4. \tag{2.7}$$

Denoting

$$\begin{aligned} \lim_{R\to\infty}\frac{1}{R}E_{iJ} &\equiv P_{iJ},\\ \lim_{R\to\infty}\frac{1}{R}E_{Ij} &\equiv P_{Ij},\\ \lim_{R\to\infty}\frac{1}{R}E_{IJ} &\equiv P_{IJ}, \end{aligned} \tag{2.8}$$

we find that the operators (2.7) obey the commutation relations (in the limit $R\to\infty$)

$$\begin{aligned} [E_{ij},E_{k\ell}] &= E_{i\ell}\delta_{jk} - E_{kj}\delta_{i\ell},\\ [E_{ij},p_{jK}] &= P_{iK},\\ [E_{ij},p_{Ki}] &= -P_{Kj}, \end{aligned} \tag{2.9}$$

with all other commutators vanishing. We retain the convention that lower case indices range over the values 1 and 2, while capital indices range over the values 3 and 4.

We must now consider the irreducible representations of U(4) and their contraction to those of $[U(2)\circledS\mathbb{C}^{2\times2}]\times\mathbb{H}^{2\times2}$. The irreducible representations of U(4) are specified by sets of four real integers m_{14}, m_{24}, m_{34}, m_{44}, where $m_{14}-m_{24}$, $m_{24}-m_{34}$, and $m_{34}-m_{44}$ are nonnegative integers. In the Gel'fand–Tsetlin basis these representations are reduced by representations of the subgroups in the canonical chain $U(4)\supset U(3)\supset U(2)\supset U(1)$. Hence, each component of a vector in the space of an irreducible representation of U(4) is labeled by a triangular array of numbers, the Gel'fand pattern, in which the jth row (from the bottom) contains j real numbers m_{ij} which obey the "betweenness" conditions

$$m_{i,j+1} \leqslant m_{ij} \leqslant m_{i+1,j+1} \tag{2.10}$$

and which constitute the invariants of the U(j) subgroup of U(4) in the canonical decomposition. The quantities $m_{ij} - m_{i+1,j}$ and $m_{ij} - m_{44}$ are restricted to be nonnegative integers.

The matrix elements of the generators E_{ij} in the Gel'fand basis are well known; we shall give them in the recension of J.D. Louck.[8] Let us consider the $\frac{1}{2}n(n-1)$ generators E_{ij}, $i<j$, of U(n), and let $(m)_n$ denote a Gel'fand state of U(n). Let $(m')_n$ denote a Gel'fand state which can be obtained from $(m)_n$ by application of E_{ij}. Then $(m')_n$ differs from $(m)_n$ only in the $i, i+1, ..., j-1$ rows. In each of these rows $(m')_n$ will differ from the corresponding row in $(m)_n$ in only one entry, say the entry $m_{\tau_k k}$ in the k th row. For this entry we shall have

$$m'_{\tau_k k} = m_{\tau_k k} + 1, \tag{2.11}$$

while

$$m'_{\ell k} = m_{\ell k}, \quad \ell \neq \tau_k . \tag{2.12}$$

Then

$$\begin{aligned} &\langle (m')_n | E_{ij} | (m)_n \rangle \\ &= \left[\left| \frac{\Pi_{s=1}^{j} (p_{sj} - p_{\tau_{j-1}, j-1} - 1)}{\Pi_{s=1, s\neq\tau_{j-1}}^{j-1} (p_{s,j-1} - p_{\tau_{j-1}, j-1} - 1)} \right| \right] \\ &\times \prod_{k=1}^{j-1} \Bigg\{ S(\tau_{k-1} - \tau_k) \\ &\times \left[\left(\prod_{s=1, s\neq\tau_{k-1}}^{k-1} \frac{(p_{s,k-1} - p_{\tau_k, k})}{(p_{s,k-1} - p_{\tau_{k-1}, k-1} - 1)} \right) \right. \\ &\times \left. \left(\prod_{s=1, s\neq\tau_k}^{k} \frac{(p_{s,k} - p_{\tau_{k-1}, k-1} - 1)}{(p_{s,k} - p_{\tau_k, k})} \right) \right]^{1/2} \Bigg\}, \end{aligned} \tag{2.13}$$

where

$$p_{ij} = m_{ij} + j - i \tag{2.14}$$

denote the partial hooks, and $S(p-q)$ is defined to be $+1$ for $p \geqslant q$ and -1 for $p<q$. There is an ambiguity in the expression (2.13), namely, the factor in the product $\{\cdots\}$ corresponding to $k=i$ contains and index τ_{i-1} which does not appear in $(m')_n$ because the labels in row $i-1$ are not shifted by E_{ij}, $i<j$, i.e., $(m')_n$ and $(m)_n$ have identical $(i-1)$th rows. The correct factor for $k=i$ is

$$S(i-\tau_i) \left[\prod_{s=1}^{i=1} (p_{s,i-1} - p_{\tau,i}) \Big/ \prod_{s=1, s\neq\tau_i}^{i} (p_{si} - p_{\tau,i}) \right]^{1/2}. \tag{2.15}$$

The matrix elements of the diagonal generators E_{ii} vanish unless $(m')_n = (m)_n$. We have, then

$$\langle (m)_n | E_{ii} | (m)_n \rangle = \sum_{j=1}^{i} m_{ji} - \sum_{j=1}^{i-1} m_{j,i-1} . \tag{2.16}$$

The matrix elements E_{ij}, $i>j$, can be found from Eq. (2.13) by means of the relation of Hermitian conjugation $E_{ij}^{\dagger} = E_{ji}$.

The contraction process for the Lie algebra of U(4), by which it becomes that of $[\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}] \times \mathbb{H}^{2\times 2}$, leaves the generators of the U(2) subgroup in the canonical decomposition unaffected. This fact [or, rather, choice of uncontracted U(2) subgroup] suggests that we should be able to obtain matrix elements of the infinitesimal operators (2.7) in an irreducible representation of $[\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}] \times \mathbb{H}^{2\times 2}$ by taking matrix elements of these operators between Gel'fand states of the form

$$\left| \left(\begin{array}{ccccccc} Rp_{14} & & Rp_{24} & & -Rp_{34} & & -Rp_{44} \\ & Rp_{13} + m_{13} & & m_{23} & & -Rp_{33} + m_{33} & \\ & & m_{12} & & m_{22} & & \\ & & & m_{11} & & & \end{array} \right) \right\rangle, \tag{2.17}$$

or

$$\left| \left(\begin{array}{ccccccc} Rp_{14} & & Rp_{24} & & -Rp_{34} & & -Rp_{44} \\ & Rp_{14} - m_{13} & & m_{23} & & -Rp_{44} + m_{33} & \\ & & m_{12} & & m_{22} & & \\ & & & m_{11} & & & \end{array} \right) \right\rangle, \tag{2.18}$$

and then taking the limit $R \rightarrow \infty$. In Eqs. (2.17) and (2.18), of course, all entries in the different Gel'fand patterns must maintain the betweenness conditions in the asymptotic limit. The reader may easily devise other asymptotic limits in addition to the two given above. These expressions for the indicated limits of basis states of irreducible representations of

U(4) are merely symbolic; we shall not realize them in any explicit manner. The only asymptotic limits which we shall evaluate explicitly will be those of the matrix elements of the infinitesimal generators or of finite transformations in U(4). In order to evaluate the matrix element of an operator in the contracted group $[\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}] \times \mathbb{H}^{2\times 2}$ we must first determine the corresponding operator in the group U(4), then express it as an operator function of the contraction parameter R ($\mathscr{O} \to \mathscr{O}_R$), then express the Gel'fand labels in terms of R as well ($|(m)\rangle \to |(m),R\rangle$), then read off the asymptotic limit of the U(4) matrix element as the corresponding matrix element in the contracted group. The relation

$$\lim_{R\to\infty} \mathscr{O}_R |(m),R\rangle = \lim_{R\to\infty} \sum_{\oplus(m')} \langle (m'),R|\mathscr{O}_R|(m),R\rangle |(m'),R\rangle \qquad (2.19)$$

defines the symbolic meaning of the expressions (2.17) and (2.18). In all cases of interest to us, only discrete sums, not integrals, will appear on the right of Eq. (2.19), though they will be infinite sums in the case that $\mathscr{O}_R$ is the operator of a finite translation in the Abelian subgroup $\mathbb{C}^{2\times 2}$ and in some other cases. Since we are contracting a compact group to a noncompact group, the spaces of irreducible representations become infinite dimensional.

We must choose an asymptotic limit for the representation labels, either Eqs. (2.17), (2.18), or some other, and we wish our limit to satisfy certain conditions:

(1) The basis state, in the asymptotic limit, must approach the direct product of a basis state of a representation of $\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}$ and a basis state of a representation of $\mathbb{H}^{2\times 2}$. Alternatively, any matrix element must approach a product of a matrix element in a representation of $\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}$ and one in a representation of $\mathbb{H}^{2\times 2}$.

(2) The basis state should approach an eigenstate of the four invariant operators on $\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}$, which are

$$P_{14}P_{41} + P_{24}P_{42}, \qquad (2.20)$$

$$P_{13}P_{31} + P_{23}P_{32}, \qquad (2.21)$$

$$P_{13}P_{41} + P_{23}P_{42}, \qquad (2.22a)$$

$$P_{14}P_{31} + P_{24}P_{32}. \qquad (2.22b)$$

The simultaneous eigenvectors of these operators span the space of an irreducible representation of $\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}$. Note that Eqs. (2.20) and (2.21) are Hermitian operators and have real eigenvalues, while the eigenvalues of Eqs. (2.22a) and (2.22b) must be complex conjugates of each other (since $P^\dagger_{iJ} = P_{Ji}$). Thus, the irreducible representations of $\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}$ are labeled by four (continuous) parameters.

(3) The consistency requirement is as follows: Once we have determined the matrix elements of the infinitesimal operators in an irreducible representation and we have verified that they are consistent with the commutation relations (2.9), we must be able to recover these matrix elements from those of a finite transformation by restricting the latter to an infinitesimal neighborhood of the identity. Also (or rather, alternatively), we must verify that the matrix elements of finite transformations have the proper behavior under the differential operators on the group manifold which realize the algebra of infinitesimal generators. We should note that the validity of the construction, then, can be established completely independently of that of the contraction procedure. We shall use the method of the contraction limit only to obtain matrix elements which we may plausibly conjecture to be those of operators in irreducible representations of the motion group $\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}$; we shall then be able to establish the correctness of this procedure by independent means.

We must now look more closely at the group $\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}$ and its dual object and irreducible representations. This semidirect product is composed of an Abelian subgroup of 2×2 complex matrices under matrix addition. The group U(2) then acts on these matrices from the left. Every complex matrix can be decomposed uniquely into the product UAZ, where U is unitary, A is a real positive semidefinite diagonal matrix, and Z is a complex upper triangular matrix with unit entries on the principal diagonal. Thus, every 2×2 complex matrix can be brought to the form

$$\begin{pmatrix} \rho_1 & \rho_3 e^{i\varphi} \\ 0 & \rho_2 \end{pmatrix} \qquad (2.23)$$

by a unique left unitary transformation, where ρ_i, φ are real and $\rho_i \geqslant 0$, $0 \leqslant \varphi < 2\pi$. Hence, each distinct quartet of real numbers (ρ_i, φ) which satisfies these conditions defines a distinct orbit under the operation of the unitary group U(2) acting on Eq. (2.23) by left multiplication. In the general case, then, there are no nontrivial little groups. In the case $\rho_1 = 0$ or $\rho_2 = 0$ we find nontrivial little groups isomorphic to U(1). If $\rho_1 = \rho_2 = \rho_3 = 0$, then, we find a nontrivial little group isomorphic to U(2). These special cases, of course, occupy sets of measure zero in the space of all orbits. Irreducible representations of the group $\mathrm{U}(2) \circledS \mathbb{C}^{2\times 2}$, then, will be specified in the general case only by quartets $(\rho_1, \rho_2, \rho_3, \varphi)$ with $\rho_i > 0$, $0 \leqslant \varphi < 2\pi$. In the special cases, those for which we have nontrivial little groups, irreducible representations will also be labeled by the discrete invariants of the little groups involved. We shall defer consideration of these degenerate cases until Sec. IV.

Let us now consider the elements of the group U(4) and those of its U(3) subgroup of elements of the form

$$\begin{pmatrix} & & & 0 \\ & U_3 & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (2.24)$$

where $U_3 \in \mathrm{U}(3)$. When we restrict U(4) to transformations of the type (2.1) above, then we also restrict U_3 in Eq. (2.24) to transformations of the form

$$\lim_{R\to\infty} \begin{pmatrix} & & 0 \\ & U_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\times \begin{pmatrix} & & \frac{1}{R}z_1 \\ & \mathbb{1} & \frac{1}{R}z_2 \\ -\frac{1}{R}\bar{z}_1 & -\frac{1}{R}\bar{z}_2 & 1+\frac{i}{R}h \end{pmatrix}, \quad (2.25)$$

where $U_2 \in \mathrm{U}(2)$, $z_i \in \mathbb{C}$, $h \in \mathbb{R}$, and $\mathbb{1}$ denotes the 2×2 unit matrix. The group of transformations (2.25) is isomorphic, then, to $[\mathrm{U}(2) \circledS \mathbb{C}^{2\times 1}]\times \mathbb{R}$. The group $\mathrm{U}(2)\circledS \mathbb{C}^{2\times 1}$ is just the inhomogeneous unitary group IU(2), whose representation theory has been studied by Chakrabarti.[9] In the Gel'fand basis the basis states for the irreducible representations of U(3) are given by triangular arrays of three rows

$$\left| \begin{pmatrix} m_{13} & & m_{23} & & m_{33} \\ & m_{12} & & m_{22} & \\ & & m_{11} & & \end{pmatrix} \right\rangle, \quad (2.26)$$

from which we obtain basis states for the irreducible representations of $[\mathrm{U}(2)\circledS \mathbb{C}^{2\times 1}]\times \mathbb{R}$ by means of the identification $m_{13}\rightarrow Rp_{13}$ and $m_{33}\rightarrow -Rp_{33}$, which p_{13} and p_{33} are real positive numbers, and the limit $R\rightarrow +\infty$. Then the basis state (2.26) becomes merely the lower three rows of Eq. (2.17). The subgroup U(3) of U(4) under consideration [that specified by Eq. (2.24)] becomes under contraction the subgroup $[\mathrm{U}(2)\circledS \mathbb{C}^{2\times 1}]\times \mathbb{R}$ or $[\mathrm{U}(2)\circledS \mathbb{C}^{2\times 2}]\times \mathbb{H}^{2\times 2}$. In the case of this subgroup we find that we obtain irreducible representations of $\mathrm{U}(2)\circledS \mathbb{C}^{2\times 1}$ which are labeled by a discrete invariant m_{23} as well as by a continuous invariant. The Abelian subgroup of 2×1 complex matrices under addition, then, is composed of elements each of which is located on an orbit of the type

$$\begin{pmatrix} \rho \\ 0 \end{pmatrix}, \quad (2.27)$$

where ρ is a nonnegative real number. All points on the orbit of Eq. (2.27) can then be obtained by left multiplication by the elements of U(2). The element (2.27) is left invariant by a nontrivial little group isomorphic to U(1). Hence, the discrete invariant m_{23} in Eq. (2.26) remains a discrete invariant in the contracted group, the invariant which labels the irreducible representations of the U(1) little group. The dual object of the Abelian group of $\mathbb{C}^{2\times 1}$ matrices under addition is isomorphic to the set of elements of the group, and hence to the set of orbits (2.27). Hence, it is parametrized by a single real, positive, continuous parameter ρ. When $\rho = 0$ we have a distinct representation induced from that of the little group U(2) instead of U(1).

When we consider the contraction of the full group U(4) to the group $[\mathrm{U}(2)\circledS \mathbb{C}^{2\times 2}]\times \mathbb{H}^{2\times 2}$, then, we must consider the action of the unitary group U(2) on complex 2×2 matrices. Each $\mathbb{C}^{2\times 2}$ matrix is located on an orbit of the form (2.23) above, and in the general case ($\rho_i \neq 0$) there will be no nontrivial little groups, and hence no discrete invariants for the group $\mathrm{U}(2)\circledS \mathbb{C}^{2\times 2}$. This suggests that we investigate asymptotic limits of the form (2.17) or (2.18) above, i.e., forms in which none of the U(4) invariants becomes a discrete parameter under contraction. We shall defer consideration of the degenerate cases until Sec. IV.

III. THE CONTRACTION OF THE IRREDUCIBLE REPRESENTATIONS OF U(4) TO THOSE OF THE GROUP $[\mathrm{U}(2)\sigma\mathbb{C}^{2\times 2}]\times \mathbb{H}^{2\times 2}$

We shall now consider in detail the contraction of the irreducible representations of the Lie algebra of U(4) to those of the Lie algebra of $[\mathrm{U}(2)\circledS \mathbb{C}^{2\times 2}]\times \mathbb{H}^{2\times 2}$. We operate with the generators (2.8) on the basis states (2.17) and take the limit $R\rightarrow\infty$. Using Eq. (2.13), we find for the $\mathbb{H}^{2\times 2}$ generators that

$$\lim_{R\rightarrow\infty}\left(\frac{1}{R}E_{43}\right)\left| \begin{pmatrix} Rp_{14} & & Rp_{24} & & -Rp_{34} & & -Rp_{44} \\ & Rp_{13}+m_{13} & & m_{23} & & -Rp_{33}+m_{33} & \\ & & m_{12} & & m_{22} & & \\ & & & m_{11} & & & \end{pmatrix} \right\rangle$$

$$= P_{43}\left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle$$

$$= \mathscr{P}_{13} \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}-1, m_{33}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle + \mathscr{P}_{33} \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}-1; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle, \tag{3.1a}$$

$$P_{34} \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle = \mathscr{P}_{13} \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}+1, m_{33}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle + \mathscr{P}_{33} \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}+1; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle, \tag{3.1b}$$

where $\mathscr{P}_{13}$ and $\mathscr{P}_{33}$ are the real numbers

$$\mathscr{P}_{13} = \left[\frac{(p_{14}-p_{13})(p_{13}-p_{24})(p_{13}+p_{34})(p_{13}+p_{44})}{(p_{13}+p_{33})^2} \right]^{1/2},$$

$$\mathscr{P}_{33} = \left[\frac{(p_{14}+p_{33})(p_{24}+p_{33})(p_{33}-p_{34})(p_{44}-p_{33})}{(p_{13}+p_{33})^2} \right]^{1/2}. \tag{3.2}$$

We note, then, that the contraction performed with the limit scheme (2.17) does not yield an irreducible representation of the $\mathbb{H}^{2\times 2}$ subgroup. The contraction yields the following expressions for the representation of the Lie algebra of U(2)Ⓢ$\mathbb{C}^{2\times 2}$:

$$P_{13} \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle = \left[\frac{(m_{12}-m_{23}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)} \right]^{1/2} \mathscr{P}_1 \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}; & m_{12}+1 & & m_{22} \\ & & & m_{11}+1 & \end{pmatrix} \right\rangle - \left[\frac{(m_{23}-m_{22})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})} \right]^{1/2} \mathscr{P}_1 \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}; & m_{12} & & m_{22}+1 \\ & & & m_{11}+1 & \end{pmatrix} \right\rangle, \tag{3.3a}$$

$$P_{23} \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle = \left[\frac{(m_{12}-m_{23}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)} \right]^{1/2} \mathscr{P}_1 \left| \begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}, m_{33}; & m_{12}+1 & & m_{22} \\ & & & m_{11} & \end{pmatrix} \right\rangle$$

$$+\left[\frac{(m_{23}-m_{22})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_1\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13},m_{33};\quad m_{12} & & m_{22}+1 \\ & m_{11} & \end{pmatrix}\right\rangle, \tag{3.3b}$$

$$P_{31}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13},m_{23};\quad m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle$$

$$=\left[\frac{(m_{12}-m_{23})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_1\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13},m_{33};\quad m_{12}-1 & & m_{22} \\ & m_{11}-1 & \end{pmatrix}\right\rangle$$

$$-\left[\frac{(m_{23}-m_{22}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_1\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13},m_{33};\quad m_{12} & & m_{22}-1 \\ & m_{11}-1 & \end{pmatrix}\right\rangle, \tag{3.3c}$$

$$P_{32}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13},m_{33};\quad m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle$$

$$=\left[\frac{(m_{12}-m_{23})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_1\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13},m_{33};\quad m_{12}-1 & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle$$

$$+\left[\frac{(m_{23}-m_{22}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_1\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13},m_{33};\quad m_{12} & & m_{22}-1 \\ & m_{11} & \end{pmatrix}\right\rangle, \tag{3.3d}$$

$$P_{14}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13},m_{33};\quad m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle$$

$$=\left[\frac{(m_{12}-m_{23}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{13}\sqrt{\frac{p_{33}}{p_{13}}}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\quad m_{13}+1,m_{33};\quad m_{12}+1 & & m_{22} \\ & m_{11}+1 & \end{pmatrix}\right\rangle$$

$$
-\left[\frac{(m_{12}-m_{23}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{33}\sqrt{\frac{p_{13}}{p_{33}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}+1; & m_{12}+1 & & m_{22} \\ & & & m_{11}+1 & \end{pmatrix}\right\rangle
$$

$$
-\left[\frac{(m_{23}-m_{22})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{13}\sqrt{\frac{p_{33}}{p_{13}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}+1,m_{33}; & m_{12} & & m_{22}+1 \\ & & & m_{11}+1 & \end{pmatrix}\right\rangle
$$

$$
+\left[\frac{(m_{23}-m_{22})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{33}\sqrt{\frac{p_{13}}{p_{33}}}\left|\begin{pmatrix} & & & m_{23}+1 & \\ p_{ij}; & m_{13},m_{33}+1; & m_{12} & & m_{22}+1 \\ & & & m_{11}+1 & \end{pmatrix}\right\rangle
$$

$$
-\left[\frac{(m_{23}-m_{22}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{2}\left|\begin{pmatrix} & & & m_{23}+1 & \\ p_{ij}; & m_{13},m_{33}; & m_{12}+1 & & m_{22} \\ & & & m_{11}+1 & \end{pmatrix}\right\rangle
$$

$$
-\left[\frac{(m_{12}-m_{23})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{2}\left|\begin{pmatrix} & & & m_{23}+1 & \\ p_{ij}; & m_{13},m_{33}; & m_{12} & & m_{22}+1 \\ & & & m_{11}+1 & \end{pmatrix}\right\rangle, \qquad (3.4a)
$$

$$
P_{24}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix}\right\rangle
$$

$$
=\left[\frac{(m_{12}-m_{23}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{13}\sqrt{\frac{p_{33}}{p_{13}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}+1,m_{33}; & m_{12}+1 & & m_{22} \\ & & & m_{11} & \end{pmatrix}\right\rangle
$$

$$
-\left[\frac{(m_{12}-m_{23}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{33}\sqrt{\frac{p_{13}}{p_{33}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}+1; & m_{12}+1 & & m_{22} \\ & & & m_{11} & \end{pmatrix}\right\rangle
$$

$$
+\left[\frac{(m_{23}-m_{22})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{13}\sqrt{\frac{p_{33}}{p_{13}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}+1,m_{33}; & m_{12} & & m_{22}+1 \\ & & & m_{11} & \end{pmatrix}\right\rangle
$$

$$-\left[\frac{(m_{23}-m_{22})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{33}\sqrt{\frac{p_{13}}{p_{33}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}+1; & m_{12} & & m_{22}+1 \\ & & & m_{11} & \end{pmatrix}\right\rangle\!\!\rangle$$

$$-\left[\frac{(m_{23}-m_{22}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{2}\left|\begin{pmatrix} & & & m_{23}+1 & \\ p_{ij}; & m_{13},m_{33}; & m_{12}+1 & & m_{22} \\ & & & m_{11} & \end{pmatrix}\right\rangle\!\!\rangle$$

$$+\left[\frac{(m_{12}-m_{23})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{2}\left|\begin{pmatrix} & & & m_{23}+1 & \\ p_{ij}; & m_{13},m_{33}; & m_{12} & & m_{22}+1 \\ & & & m_{11} & \end{pmatrix}\right\rangle\!\!\rangle, \tag{3.4b}$$

$$P_{41}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{pmatrix}\right\rangle\!\!\rangle$$

$$=\left[\frac{(m_{12}-m_{23})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{13}\sqrt{\frac{p_{33}}{p_{13}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}-1,m_{33}; & m_{12}-1 & & m_{22} \\ & & & m_{11}-1 & \end{pmatrix}\right\rangle\!\!\rangle$$

$$-\left[\frac{(m_{12}-m_{23})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{33}\sqrt{\frac{p_{13}}{p_{33}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}-1; & m_{12}-1 & & m_{22} \\ & & & m_{11}-1 & \end{pmatrix}\right\rangle\!\!\rangle$$

$$-\left[\frac{(m_{23}-m_{22}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{13}\sqrt{\frac{p_{33}}{p_{13}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13}-1,m_{33}; & m_{12} & & m_{22}-1 \\ & & & m_{11}-1 & \end{pmatrix}\right\rangle\!\!\rangle$$

$$+\left[\frac{(m_{23}-m_{22}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{33}\sqrt{\frac{p_{13}}{p_{33}}}\left|\begin{pmatrix} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}-1; & m_{12} & & m_{22}-1 \\ & & & m_{11}-1 & \end{pmatrix}\right\rangle\!\!\rangle$$

$$-\left[\frac{(m_{23}-m_{22})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{2}\left|\begin{pmatrix} & & & m_{23}-1 & \\ p_{ij}; & m_{13},m_{33}; & m_{12}-1 & & m_{22} \\ & & & m_{11}-1 & \end{pmatrix}\right\rangle\!\!\rangle$$

$$
-\left[\frac{(m_{12}-m_{23}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_2\left|\left(\begin{array}{ccccc} & & & m_{23}-1 & \\ p_{ij}; & m_{13},m_{33}; & m_{12} & & m_{22}-1 \\ & & & m_{11}-1 & \end{array}\right)\right\rangle, \tag{3.4c}
$$

$$
P_{42}\left|\left(\begin{array}{ccccc} & & & m_{23} & \\ p_{ij}; & m_{13},m_{23}; & m_{12} & & m_{22} \\ & & & m_{11} & \end{array}\right)\right\rangle
$$

$$
=\left[\frac{(m_{12}-m_{23})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{13}\sqrt{\frac{p_{33}}{p_{13}}}\left|\left(\begin{array}{ccccc} & & & m_{23} & \\ p_{ij}; & m_{13}-1,m_{33}; & m_{12}-1 & & m_{22} \\ & & & m_{11} & \end{array}\right)\right\rangle
$$

$$
+\left[\frac{(m_{12}-m_{23})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_{33}\sqrt{\frac{p_{13}}{p_{33}}}\left|\left(\begin{array}{ccccc} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}-1; & m_{12}-1 & & m_{22} \\ & & & m_{11} & \end{array}\right)\right\rangle
$$

$$
+\left[\frac{(m_{23}-m_{22}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{13}\sqrt{\frac{p_{33}}{p_{13}}}\left|\left(\begin{array}{ccccc} & & & m_{23} & \\ p_{ij}; & m_{13}-1,m_{33}; & m_{12} & & m_{22}-1 \\ & & & m_{11} & \end{array}\right)\right\rangle
$$

$$
-\left[\frac{(m_{23}-m_{22}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_{33}\sqrt{\frac{p_{13}}{p_{33}}}\left|\left(\begin{array}{ccccc} & & & m_{23} & \\ p_{ij}; & m_{13},m_{33}-1; & m_{12} & & m_{22}-1 \\ & & & m_{11} & \end{array}\right)\right\rangle
$$

$$
-\left[\frac{(m_{23}-m_{22})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_2\left|\left(\begin{array}{ccccc} & & & m_{23}-1 & \\ p_{ij}; & m_{13},m_{33}; & m_{12}-1 & & m_{22} \\ & & & m_{11} & \end{array}\right)\right\rangle
$$

$$
+\left[\frac{(m_{12}-m_{23}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_2\left|\left(\begin{array}{ccccc} & & & m_{23}-1 & \\ p_{ij}; & m_{13},m_{33}; & m_{12} & & m_{22}-1 \\ & & & m_{11} & \end{array}\right)\right\rangle, \tag{3.4d}
$$

where

$$
\mathscr{P}_1=[p_{13}p_{33}]^{1/2},\quad \mathscr{P}_2=\left[\frac{p_{14}p_{24}p_{34}p_{44}}{p_{13}p_{33}}\right]^{1/2}. \tag{3.5}
$$

We note, then, that in the limit scheme of Eq. (2.17) the operators P_{i4} and P_{4i} can shift the indices m_{13} and m_{33}, both when $i=1,2$ and when $i=3$. From Eq. (3.1) we find that m_{13} and m_{33} are state labels for a representation of the $\mathbb{H}^{2\times 2}$ subgroup, but Eq. (3.4) indicates that unless $\mathscr{P}_{13}=\mathscr{P}_{33}=0$ these labels can be shifted by the operators P_{i4}, P_{4i}, $i=1,2$. Hence, the limit scheme of Eq. (2.17) does not yield a decomposition, in the limit $R\to\infty$, of the Gel'fand state into the direct product of basis states for representations, respectively, of the groups $\mathbb{H}^{2\times 2}$ and $U(2)\textcircled{\sigma}\mathbb{C}^{2\times 2}$ except in the special case $\mathscr{P}_{13} = \mathscr{P}_{33} = 0$. In taking the asymptotic limit of the matrix element of a finite transformation between states of the form (2.17), we find that the resulting representation matrix for the group $U(2)\textcircled{\sigma}\mathbb{C}^{2\times 2}$, upon restriction to infinitesimal transformations, yields Eq. (3.4) with the assignments $\mathscr{P}_{13} = \mathscr{P}_{33} = 0$. The states

$$\left|\left(\begin{matrix} p_{ij}; & m_{13}, m_{33}; & m_{12} & \begin{matrix} m_{23} \\ \\ m_{11} \end{matrix} & m_{22} \end{matrix}\right)\right\rangle$$

are indeed eigenstates of the invariant operator (2.21):

$$(P_{13}P_{31} + P_{23}P_{32})\left|\left(\begin{matrix} p_{ij}; & m_{13}, m_{33}; & m_{12} & \begin{matrix} m_{23} \\ \\ m_{11} \end{matrix} & m_{22} \end{matrix}\right)\right\rangle = (\mathscr{P}_1)^2\left|\left(\begin{matrix} p_{ij}; & m_{13}, m_{33}; & m_{12} & \begin{matrix} m_{23} \\ \\ m_{11} \end{matrix} & m_{22} \end{matrix}\right)\right\rangle, \tag{3.6}$$

but they are not eigenstates of the other invariant operators (2.20), (2.22a), and (2.22b) unless $\mathscr{P}_{13} = \mathscr{P}_{33} = 0$. In this case we find

$$(P_{14}P_{41} + P_{24}P_{42})\left|\left(\begin{matrix} p_{ij}; & m_{13}, m_{33}; & m_{12} & \begin{matrix} m_{23} \\ \\ m_{11} \end{matrix} & m_{22} \end{matrix}\right)\right\rangle = (\mathscr{P}_2)^2\left|\left(\begin{matrix} p_{ij}; & m_{13}, m_{33}; & m_{12} & \begin{matrix} m_{23} \\ \\ m_{11} \end{matrix} & m_{22} \end{matrix}\right)\right\rangle, \tag{3.7}$$

$$(P_{13}P_{41} + P_{23}P_{42})\left|\left(\begin{matrix} p_{ij}; & m_{13}, m_{33}; & m_{12} & \begin{matrix} m_{23} \\ \\ m_{11} \end{matrix} & m_{22} \end{matrix}\right)\right\rangle = 0, \tag{3.8a}$$

$$(P_{31}P_{14} + P_{32}P_{24})\left|\left(\begin{matrix} p_{ij}; & m_{13}, m_{33}; & m_{12} & \begin{matrix} m_{23} \\ \\ m_{11} \end{matrix} & m_{22} \end{matrix}\right)\right\rangle = 0, \tag{3.8b}$$

and we have realized the irreducible representation of $U(2)\textcircled{\sigma}\mathbb{C}^{2\times 2}$ which corresponds to the matrix (2.23):

$$\begin{pmatrix} \rho_1 & \rho_3 e^{i\varphi} \\ 0 & \rho_2 \end{pmatrix} = \begin{pmatrix} \mathscr{P}_1 & 0 \\ 0 & \mathscr{P}_2 \end{pmatrix}. \tag{3.9}$$

Alternatively, this representation can be realized directly and immediately by the use of the limit scheme of Eq. (2.18) above. In this case we realize Eq. (3.1), (3.3), and (3.4) with $\mathscr{P}_{13} = \mathscr{P}_{33} = 0, p_{14} = p_{13}$, and $p_{44} = p_{33}$. Hence, the use of the limit scheme (2.18) gives us immediately the direct product of irreducible representations of $\mathbb{H}^{2\times 2}$ and $U(2)\textcircled{\sigma}\mathbb{C}^{2\times 2}$, but the limit in the case (2.18) realizes only the trivial (identity) representation of the subgroups generated by the operators P_{34} and P_{43}.

We can also take asymptotic limits of matrix elements of finite transformations in irreducible representations of U(4) using Eqs. (2.5) as argument and limit scheme (2.17) or (2.18). In each case we find that the matrix element breaks down in the limit into the product of a $U(2)\textcircled{\sigma}\mathbb{C}^{2\times 2}$ matrix element and an $\mathbb{H}^{2\times 2}$ matrix element. Restricting each of these matrix elements to an infinitesimal transformation in the appropriate group [$U(2)\textcircled{\sigma}\mathbb{C}^{2\times 2}$ and $\mathbb{H}^{2\times 2}$], we recover the expressions (3.1) and (3.3) in limit scheme (2.17), but in the limit scheme (2.18) we recover Eq. (3.3), but Eq. (3.1) only with $\mathscr{P}_{13} = \mathscr{P}_{33} = 0$. Setting $\Omega = 1$, we recover Eq. (3.4) in both limit schemes only with $\mathscr{P}_{13} = \mathscr{P}_{33} = 0$.

Relations (3.3) give us the correct matrix elements of the four generators P_{i3}, P_{3i} of the Lie algebra of $U(2) \textcircled{\sigma} \mathbb{C}^{2\times 2}$, and agree with the results of Chakrabarti.[9] The expressions (3.4) do not give us the correct matrix elements for the generators P_{i4}, P_{4i} of $U(2) \textcircled{\sigma} \mathbb{C}^{2\times 2}$ unless $\mathscr{P}_{13} = \mathscr{P}_{33} = 0$, and under these conditions and $\Omega = \mathbb{1}$ in Eq. (3.5) we realize only the special cases (3.9) of irreducible representations of $U(2) \textcircled{\sigma} \mathbb{C}^{2\times 2}$.

In order to realize the most general class of irreducible representations of $U(2) \textcircled{\sigma} \mathbb{C}^{2\times 2}$, we must resort to more general values for the matrix Ω. We shall set $\omega_1^1 = 1$, $\omega_2^1 = 0$, since we already have the correct matrix elements for the generators of the IU(2) subgroup of $U(2) \textcircled{\sigma} \mathbb{C}^{2\times 2}$ [where IU(2) is simply $U(2) \textcircled{\sigma} \mathbb{C}^{2\times 1}$], and since, for convenience, we would like to keep m_{23} as an invariant label for the states of irreducible representations of this subgroup. We shall also set $\omega_2^2 = 1$ and leave ω_1^2 a general complex number, so that

$$\Omega = \begin{pmatrix} 1 & \omega_1^2 \\ 0 & 1 \end{pmatrix}. \qquad (3.10)$$

In place of the relations (3.4), then, we have the following expressions for the generators P_{i4}, P_{4i} of $U(2) \textcircled{\sigma} \mathbb{C}^{2\times 2}$ (where we shall omit the indices m_{13}, m_{33}):

$$P_{14} \left| \begin{pmatrix} & m_{23} & \\ p_{ij};\ m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix} \right\rangle$$

$$= \left[\frac{(m_{12} - m_{23} + 1)(m_{11} - m_{22} + 1)}{(m_{12} - m_{22} + 2)(m_{12} - m_{22} + 1)} \right]^{1/2} \bar{\omega} \mathscr{P}_3 \left| \begin{pmatrix} & m_{23} & \\ p_{ij};\ m_{12} + 1 & & m_{22} \\ & m_{11} + 1 & \end{pmatrix} \right\rangle$$

$$- \left[\frac{(m_{23} - m_{22})(m_{12} - m_{11})}{(m_{12} - m_{22} + 1)(m_{12} - m_{22})} \right]^{1/2} \bar{\omega} \mathscr{P}_3 \left| \begin{pmatrix} & m_{23} & \\ p_{ij};\ m_{12} & & m_{22} + 1 \\ & m_{11} + 1 & \end{pmatrix} \right\rangle$$

$$- \left[\frac{(m_{23} - m_{22} + 1)(m_{11} - m_{22} + 1)}{(m_{12} - m_{22} + 2)(m_{12} - m_{22} + 1)} \right]^{1/2} \mathscr{P}_2 \left| \begin{pmatrix} & m_{23} + 1 & \\ p_{ij};\ m_{12} + 1 & & m_{22} \\ & m_{11} + 1 & \end{pmatrix} \right\rangle$$

$$- \left[\frac{(m_{12} - m_{23})(m_{12} - m_{11})}{(m_{12} - m_{22} + 1)(m_{12} - m_{22})} \right]^{1/2} \mathscr{P}_2 \left| \begin{pmatrix} & m_{23} + 1 & \\ p_{ij};\ m_{12} & & m_{22} + 1 \\ & m_{11} + 1 & \end{pmatrix} \right\rangle, \qquad (3.11a)$$

$$P_{24} \left| \begin{pmatrix} & m_{23} & \\ p_{ij};\ m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix} \right\rangle$$

$$= \left[\frac{(m_{12} - m_{23} + 1)(m_{12} - m_{11} + 1)}{(m_{12} - m_{22} + 2)(m_{12} - m_{22} + 1)} \right]^{1/2} \bar{\omega} \mathscr{P}_3 \left| \begin{pmatrix} & m_{23} & \\ p_{ij};\ m_{12} + 1 & & m_{22} \\ & m_{11} & \end{pmatrix} \right\rangle$$

$$+\left[\frac{(m_{23}-m_{22})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\bar{\omega}\mathscr{P}_3\left|\left(\begin{matrix} & m_{23} & \\ p_{ij};\quad m_{12} & & m_{22}+1 \\ & m_{11} & \end{matrix}\right)\right\rangle\!\!\right\rangle$$

$$-\left[\frac{(m_{23}-m_{22}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_2\left|\left(\begin{matrix} & m_{23}+1 & \\ p_{ij};\quad m_{12}+1 & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle\!\!\right\rangle$$

$$+\left[\frac{(m_{12}-m_{23})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_2\left|\left(\begin{matrix} & m_{23}+1 & \\ p_{ij};\quad m_{12} & & m_{22}+1 \\ & m_{11} & \end{matrix}\right)\right\rangle\!\!\right\rangle, \qquad (3.11b)$$

$$P_{41}\left|\left(\begin{matrix} & m_{23} & \\ p_{ij};\quad m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle\!\!\right\rangle$$

$$=\left[\frac{(m_{12}-m_{23})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\omega\mathscr{P}_3\left|\left(\begin{matrix} & m_{23} & \\ p_{ij};\quad m_{12}-1 & & m_{22} \\ & m_{11}-1 & \end{matrix}\right)\right\rangle\!\!\right\rangle$$

$$-\left[\frac{(m_{23}-m_{22}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\omega\mathscr{P}_3\left|\left(\begin{matrix} & m_{23} & \\ p_{ij};\quad m_{12} & & m_{22}-1 \\ & m_{11}-1 & \end{matrix}\right)\right\rangle\!\!\right\rangle$$

$$\cdot\left[\frac{(m_{23}-m_{22})(m_{11}-m_{22})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_2\left|\left(\begin{matrix} & m_{23}-1 & \\ p_{ij};\quad m_{12}-1 & & m_{22} \\ & m_{11}-1 & \end{matrix}\right)\right\rangle\!\!\right\rangle$$

$$-\left[\frac{(m_{12}-m_{23}+1)(m_{12}-m_{11}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_2\left|\left(\begin{matrix} & m_{23}-1 & \\ p_{ij};\quad m_{12} & & m_{22}-1 \\ & m_{11}-1 & \end{matrix}\right)\right\rangle\!\!\right\rangle, \qquad (3.11c)$$

$$P_{42}\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle$$

$$= \left[\frac{(m_{12}-m_{23})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\omega\mathscr{P}_3\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12}-1 & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle$$

$$+ \left[\frac{(m_{23}-m_{22}+1)(m_{11}-m_{22}+1)}{(m_{12}-m_{22}+2)(m_{12}-m_{22}+1)}\right]^{1/2}\omega\mathscr{P}_3\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12} & & m_{22}-1 \\ & m_{11} & \end{matrix}\right)\right\rangle$$

$$- \left[\frac{(m_{23}-m_{22})(m_{12}-m_{11})}{(m_{12}-m_{22}+1)(m_{12}-m_{22})}\right]^{1/2}\mathscr{P}_2\left|\left(p_{ij};\ \begin{matrix} & m_{23}-1 & \\ m_{12}-1 & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle$$

$$+ \left[\frac{(m_{12}-m_{23}+1)(m_{11}-m_{22}+2)}{(m_{12}-m_{22}+2(m_{12}-m_{22}+1)}\right]^{1/2}\mathscr{P}_2\left|\left(p_{ij};\ \begin{matrix} & m_{23}-1 & \\ m_{12} & & m_{22}-1 \\ & m_{11} & \end{matrix}\right)\right\rangle, \tag{3.11d}$$

where we have used $\mathscr{P}_3$ as a real, nonnegative constant and ω as a complex phase of unit modulus. In terms of $\mathscr{P}_2$ and ω_1^2 we have

$$\omega_1^2\mathscr{P}_2 = \omega\mathscr{P}_3. \tag{3.12}$$

Combining Eqs. (3.3), (3.11), and (3.12), we obtain the following eigenvalues for the invariant operators:

$$(P_{13}P_{31}+P_{23}P_{32})\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle = (\mathscr{P}_1)^2\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle, \tag{3.13a}$$

$$(P_{14}P_{41}+P_{24}P_{42})\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle = [(\mathscr{P}_2)^2+(\mathscr{P}_3)^2]\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle, \tag{3.13b}$$

$$(P_{13}P_{41}+P_{23}P_{42})\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle = \omega\mathscr{P}_1\mathscr{P}_3\left|\left(p_{ij};\ \begin{matrix} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle, \tag{3.13c}$$

$$(P_{31}P_{14}+P_{32}P_{24})\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle = \bar{\omega}\mathscr{P}_1\mathscr{P}_3\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle. \quad (3.13d)$$

Hence, the relations (3.3) and (3.11) realize the equivalence class of the most general irreducible representation of $U(2)\textcircled{\sigma}C^{2\times2}$, along with the usual expressions for the matrix elements of the generators of U(2):

$$E_{11}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle = m_{11}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle, \quad (3.14a)$$

$$E_{22}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle = (m_{12}+m_{22}-m_{11})\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle, \quad (3.14b)$$

$$E_{12}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle = [(m_{11}-m_{22}+1)(m_{12}-m_{11})]^{1/2}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11}+1 & \end{pmatrix}\right\rangle, \quad (3.14c)$$

$$E_{21}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle = [(m_{12}-m_{11}+1)(m_{11}-m_{22})]^{1/2}\left|\begin{pmatrix} & m_{23} & \\ p_{ij};\; m_{12} & & m_{22} \\ & m_{11}-1 & \end{pmatrix}\right\rangle, \quad (3.14d)$$

since we may regard $\mathscr{P}_1$, $\mathscr{P}_2$, $\mathscr{P}_3$, ω as quantities independent of one another.

IV. MATRIX ELEMENTS OF FINITE TRANSFORMATIONS IN IRREDUCIBLE REPRESENTATIONS OF $U(2)\textcircled{\sigma}C^{2\times2}$

Having obtained the matrix elements of the generators of $U(2)\textcircled{\sigma}C^{2\times2}$ in the irreducible representation corresponding to

$$\Pi=\begin{pmatrix}\mathscr{P}_1 & \omega\mathscr{P}_3 \\ 0 & \mathscr{P}_2\end{pmatrix} \quad (4.1)$$

in the dual object, we shall now proceed to examine the matrix elements of finite transformations in the same representation. The contraction procedure, which we follow in deriving these matrix elements, is an informal, heuristic one, and so must be supplemented by a proof that the functions so obtained are indeed the matrix elements which we require. The limit process is described in Appendix A; in Appendix B we sketch the proof that the resulting functions have the right properties under differentiation. We shall now present the results of the calculation of the matrix elements of the irreducible representations of $U(2)\textcircled{\sigma}C^{2\times2}$. Let $\mathscr{O}(U,Z)$ be the operator corresponding to the element (U,Z) of the group as in Eq. (2.5) above, with $H=0$. We shall replace p_{ij} with Π in our notation for the basis states of the representation. Then the matrix element of this operator is given by

$$\left\langle\begin{pmatrix} & m'_{23} & \\ \Pi;\; m'_{12} & & m'_{22} \\ & m'_{11} & \end{pmatrix}\right|\mathscr{O}(U,Z)\left|\begin{pmatrix} & m_{23} & \\ \Pi;\; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}\right\rangle$$

$$=(-1)^{m'_{22}+m'_{23}+m_{22}+m_{23}}[(m'_{12}-m'_{22}+1)(m_{12}-m_{22}+1)]^{1/2}\sum_{\beta_{12},\beta_{22}}\sum_{\alpha_{23}}\sum_{\rho,\sigma}(\beta_{12}-\beta_{22}+1)$$

$$\times D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\;(1/2)(m'_{12}+m'_{22}-\beta_{12}-\beta_{22})-m'_{23}+\alpha_{23}-\sigma}(Z\Pi^\dagger)$$
$$\times N^{-1}_{(1/2)(m'_{12}+m'_{22}-\beta_{12}-\beta_{22})-m'_{23}+\alpha_{23}-\sigma,\alpha_{23}-(1/2)(\beta_{12}+\beta_{22});-m'_{23}+(1/2)(m'_{12}+m'_{22}),\sigma}[m'_{22}-\beta_{12}$$
$$+2|\tfrac{1}{2}(m'_{12}-m'_{22}),\tfrac{1}{2}(\beta_{12}-\beta_{22})]$$
$$\times[\det(Z\Pi^\dagger)]^{m'_{22}-\beta_{12}}D^{(1/2)(\beta_{12}-\beta_{22})}_{\rho\;\sigma}(-\Pi Z^\dagger Z\Pi^\dagger)[\det(-\Pi Z^\dagger)]^{m_{22}-\beta_{12}}$$
$$\times N^{-1}_{(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})-m_{23}+\alpha_{23}-\rho,\alpha_{23}-(1/2)(\beta_{12}+\beta_{22});-m_{23}+(1/2)(m_{12}+m_{22}),\rho}[m_{22}-\beta_{12}$$
$$+2|\tfrac{1}{2}(m_{12}-m_{22}),\tfrac{1}{2}(\beta_{12}-\beta_{22})]$$
$$\times[\det U]^{m_{22}}D^{(1/2)(m_{12}-m_{22})}_{(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})-m_{23}+\alpha_{23}-\rho\;\,m_{11}-(1/2)(m_{12}+m_{22})}(-\Pi Z^\dagger U)$$
$$=(-1)^{m'_{22}+m'_{23}+m_{22}+m_{23}}[(m_{12}-m_{22}+1)(m'_{12}-m'_{22}+1)]^{1/2}\sum_{\beta_{12},\beta_{22}}\sum_{\rho',\tau',\sigma'}\sum_{\rho,\tau,\sigma}(\beta_{12}-\beta_{22}+1)$$
$$\times D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\;\rho'}(Z)D^{(1/2)(m'_{12}-m'_{22})}_{\sigma'\;-m'_{23}+(1/2)(m'_{12}+m'_{22})}(\Pi^\dagger)$$
$$\times N^{-1}_{\rho',\rho'+\tau'-\sigma';\tau',\sigma'}[m'_{22}-\beta_{12}+2|(1/2)(m'_{12}-m'_{22}),(1/2)(\beta_{12}-\beta_{22})]$$
$$\times[\det(Z\Pi^\dagger)]^{m'_{22}-\beta_{12}}D^{(1/2)(\beta_{12}-\beta_{22})}_{\rho'+\tau'-\sigma'\;\rho+\tau-\sigma}(\Pi^\dagger\Pi)D^{(1/2)(\beta_{12}-\beta_{22})}_{\tau\;\tau'}(-Z^\dagger Z)$$
$$\times[\det(-\Pi Z^\dagger)]^{m_{22}-\beta_{12}}N^{-1}_{\rho,\rho+\tau-\sigma;\sigma,\tau}[m_{22}-\beta_{12}+2|(1/2)(m_{12}-m_{22}),(1/2)(\beta_{12}-\beta_{22})]$$
$$\times[\det U]^{m_{22}}D^{(1/2)(m_{12}-m_{22})}_{-m_{23}+(1/2)(m_{12}+m_{22})\;\sigma}(\Pi)D^{(1/2)(m_{12}-m_{22})}_{\rho\;m_{11}-(1/2)(m_{12}+m_{22})}(-Z^\dagger U), \tag{4.2}$$

where we have used the conventions of Appendix A and assumed that the matrix $Z\Pi^\dagger$ is nonsingular. We note here the analogy between Eq. (4.2), the matrix element of a finite transformation in an irreducible representation of U(2) σ $\mathbb{C}^{2\times 2}$, and the classical Bessel function, the matrix element of a transformation in an irreducible representation of U(1) σ $\mathbb{C}$. The latter has a series representation with two classical gamma functions in the denominator of the summand, whereas the former, Eq. (4.2) above, has two inverse U(2) gamma matrices N^{-1}, which are defined in Eq. (A23).

The matrix entries of the irreducible representations of this group can be constructed in different ways. In particular, when we use the standard Mackey construction of induced representations, then the matrix entries can be given in terms of the Bessel functions of Gross and Kunze [Eq. (1.1)].[3–6] We now make the association between our matrix entries and the Gross–Kunze Bessel functions:

$$\left\langle\left(\begin{matrix} & m'_{23} & \\ \Pi;\; m'_{12} & & m'_{22} \\ & m'_{11} & \end{matrix}\right)\middle|\mathscr{O}(1,Z)\middle|\left(\begin{matrix} & m_{23} & \\ \Pi;\; m_{12} & & m_{22} \\ & m_{11} & \end{matrix}\right)\right\rangle$$
$$=[(m'_{12}-m'_{22}+1)(m_{12}-m_{22}+1)]^{1/2}\int_{U(2)}dU e^{2i\,\mathrm{Re}\,\mathrm{Tr}(Z^\dagger U\Sigma_y\Pi)}[\det U]^{m'_{22}}$$
$$\times D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\;m'_{23}-(1/2)(m'_{12}+m'_{22})}(U)[\det U^\dagger]^{m_{22}}D^{(1/2)(m_{12}-m_{22})}_{m_{23}-(1/2)(m_{12}+m_{22})\;m_{11}-(1/2)(m_{12}+m_{22})}(U^\dagger)$$
$$=(-1)^{m'_{22}+m'_{23}+m_{22}+m_{23}}(i)^{m_{12}+m_{22}-m'_{12}-m'_{22}}[(m'_{12}-m'_{22}+1)(m_{12}-m_{22}+1)]^{1/2}\int_{U(2)}dU e^{2i\,\mathrm{Re}\,\mathrm{Tr}(Z^\dagger U\Pi)}$$
$$\times[\det U]^{m'_{22}}D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\;-m'_{23}+(1/2)(m'_{12}+m'_{22})}(U)[\det U^\dagger]^{m_{22}}D^{(1/2)(m_{12}-m_{22})}_{-m_{23}+(1/2)(m_{12}+m_{22})\;m_{11}-(1/2)(m_{12}+m_{22})}(U^\dagger), \tag{4.3}$$

where

$$\Sigma_y=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} \tag{4.4}$$

and dU is the normalized invariant Haar measure on the manifold of the group U(2). The identity of the series representation (4.2) and the integral representation (4.3) can be established by direct application of the differential operators (to be discussed in the following section) which act as generators of the $\mathbb{C}^{2\times 2}$ subgroup. Both Eqs. (4.2) and (4.3) have the same transformation properties under the operators (5.9) and (5.11) below and are eigenfunctions of the second order invariant combinations of these operators with the same eigenvalues. They obey the same boundary condition at the identity element of $\mathbb{C}^{2\times 2}$; hence, all of their partial derivatives are identical. They are both real analytic functions of the eight variables z^I_j and $\bar{z}^I_j$, and therefore they are identical at all elements of the group. Equation (4.3), then, establishes the relation between the Bessel functions defined by Gross and Kunze and the matrix elements of irreducible representations of U(2)Ⓢ$\mathbb{C}^{2\times 2}$ as given by the infinite series (4.2). Specifically, we have the following relation:

$$J_\lambda(\alpha,x)=\int_{U(2)}\langle U^{-1}\alpha|x\rangle\lambda(U)dU=\int_{U(2)}e^{i\,\mathrm{Re}\,\mathrm{Tr}(x^\dagger U^\dagger\alpha)}\lambda(U)dU$$

$$=\int_{U(2)}e^{i\,\mathrm{Re}\,\mathrm{Tr}(x^\dagger U\alpha)}\lambda(U^\dagger)dU=\frac{1}{(m_{12}-m_{22}+1)^{1/2}}\left\langle\!\!\left\langle \begin{pmatrix} & & 0 & \\ \varPi; & 0 & & 0 \\ & & 0 & \end{pmatrix}\right|\mathscr{O}(1,Z)\left|\begin{pmatrix} & & m_{23} & \\ \varPi; & m_{12} & & m_{22} \\ & & m_{11} & \end{pmatrix}\right\rangle\!\!\right\rangle, \tag{4.5}$$

where we make the identifications

$$\begin{aligned} &x=Z,\\ &\alpha=2\Sigma_y\varPi,\\ &\lambda(U^\dagger)=[\det U^\dagger]^{m_{22}}D^{(1/2)(m_{12}-m_{22})}_{m_{23}-(1/2)(m_{12}+m_{22})\ m_{11}-(1/2)(m_{12}+m_{22})}(U^\dagger),\end{aligned} \tag{4.6}$$

or a set of identifications equivalent to these.

We further note that the general matrix element (4.2) for the case of nonsingular $Z\varPi^\dagger$ has the structure of a matrix product. If we define

$$\left\langle\!\!\left\langle \begin{pmatrix} & m'_{23} & \\ m'_{12} & & m'_{22} \\ & m'_{11} & \end{pmatrix}\right|X(Z,\overline{\varPi})\left|\begin{pmatrix} & \beta_{23} & \\ \beta_{12} & & \beta_{22} \\ & \beta_{11} & \end{pmatrix}\right\rangle\!\!\right\rangle$$

$$\begin{aligned} &=\sum_{\rho',\sigma',\tau'}\{(-1)^{m'_{22}+m'_{23}}(m'_{12}-m'_{22}+1)[\det Z]^{m'_{22}}\\ &\times D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\ \rho'}(Z)[\det\overline{\varPi}]^{m'_{22}}D^{(1/2)(m'_{12}-m'_{22})}_{-m'_{23}+(1/2)(m'_{12}+m'_{22})\ \sigma'}(\overline{\varPi})\}\\ &\times N^{-1}_{\rho',\rho'+\tau'-\sigma';\sigma',\tau'}[m'_{22}+\beta_{22}+2|(1/2)(m'_{12}-m'_{22}),(1/2)(\beta_{12}-\beta_{22})]\\ &\times\{(-1)^{\beta_{22}+\beta_{23}}(\beta_{12}-\beta_{22}+1)^{1/2}[\det Z^t]D^{(1/2)(\beta_{12}-\beta_{22})}_{\tau'\ \beta_{11}-1/2)(\beta_{12}+\beta_{22})}(Z^t)\\ &\times[\det\varPi^\dagger]^{\beta_{22}}D^{(1/2)(\beta_{12}-\beta_{22})}_{\rho'+\tau'-\sigma'\ -\beta_{23}+(1/2)(\beta_{12}+\beta_{22})}(\varPi^\dagger)\}\equiv\langle(m')|X(Z,\overline{\varPi})|(\beta)\rangle,\end{aligned} \tag{4.7}$$

then we find that Eq. (4.2) can be written as

$$\left\langle\!\!\left\langle \begin{pmatrix} & & m'_{23} & \\ \varPi; & m'_{12} & & m'_{22} \\ & & m'_{11} & \end{pmatrix}\right|\mathscr{O}(1,Z)\left|\begin{pmatrix} & & m_{23} & \\ \varPi; & m_{12} & & m_{22} \\ & & m_{11} & \end{pmatrix}\right\rangle\!\!\right\rangle=\sum_{(\beta)}(m')|X(Z,\overline{\varPi})|(\beta)\rangle\langle(\beta)|X(-\overline{Z},\varPi)|(m)\rangle. \tag{4.8}$$

We have replaced $-\beta_{12}$ with $+\beta_{22}$ in the argument of the inverse gamma matrix and in the exponents of the determinants, i.e., we take $+\beta_{22}$ and $\frac{1}{2}(\beta_{12}-\beta_{22})$ as our independent indices of summation rather then $-\beta_{12}$ and $\frac{1}{2}(\beta_{12}-\beta_{22})$. In Eq. (4.8) the sum is taken over all real integer values of β_{22}.

We shall now consider the case that $Z\varPi^\dagger$ is singular. We shall restrict our attention to the case that Z is a lower triangular matrix with real elements on the diagonal, since all other cases are unitarily equivalent to this one. Specifically, we shall write

$$Z\varPi^\dagger=\begin{pmatrix}\zeta^1_1 & 0\\ \bar{\zeta}^2_1 & \zeta^2_2\end{pmatrix}\begin{pmatrix}\pi^1_1 & 0\\ \bar{\pi}^2_1 & \pi^2_2\end{pmatrix}=\begin{pmatrix}\zeta^1_1\pi^1_1 & 0\\ \bar{\zeta}^2_1\pi^1_1+\zeta^2_2\bar{\pi}^2_1 & \zeta^2_2\pi^2_2\end{pmatrix}. \tag{4.9}$$

We now have two cases to consider: (1) $\zeta^1_1\pi^1_1=0$ and (2) $\zeta^2_2\pi^2_2=0$. Proceeding from Eq. (A24), we find for case (1)

$$\left\langle\!\!\left\langle \begin{pmatrix} & & m'_{23} & \\ \varPi; & m'_{12} & & m'_{22} \\ & & m'_{11} & \end{pmatrix}\right|\mathscr{O}(1,Z)\left|\begin{pmatrix} & & m_{23} & \\ \varPi; & m_{12} & & m_{22} \\ & & m_{11} & \end{pmatrix}\right\rangle\!\!\right\rangle$$

$$= [(m'_{12} - m'_{22} + 1)(m_{12} - m_{22} + 1)]^{1/2} \sum_{\Lambda,\rho} (-1)^{(1/2)(m'_{12} + m'_{22} + m_{12} + m_{22}) + \Lambda + \rho}$$

$$\times C^{(1/2)(m'_{12} - m'_{22}) \quad (1/2)(m_{12} - m_{22}) \quad \Lambda}_{-m'_{23} + \rho + (1/2)(m'_{12} + m'_{22}) \quad m_{23} - \rho - (1/2)(m_{12} + m_{22}) \quad (1/2)(m'_{12} + m'_{22} - m_{12} - m_{22}) - m'_{23} + m_{23}}$$

$$\times C^{(1/2)(m'_{12} - m'_{22}) \quad (1/2)(m_{12} - m_{22}) \quad \Lambda}_{-m'_{23} + (1/2)(m'_{12} + m'_{22}) \quad m_{23} - (1/2)(m_{12} + m_{22}) \quad (1/2)(m'_{12} + m'_{22} - m_{12} - m_{22}) - m'_{23} + m_{23}}$$

$$\times D^{(1/2)(m'_{12} - m'_{22})}_{m'_{11} - (1/2)(m'_{12} + m'_{22}) \quad -m'_{23} + \rho + (1/2)(m'_{12} + m'_{22})} \left(-\frac{\Sigma_x \Psi_1}{\sqrt{\det \Psi_1}} \right)$$

$$\times D^{(1/2)(m_{12} - m_{22})}_{m_{23} - \rho - (1/2)(m_{12} + m_{22}) \quad m_{11} - (1/2)(m_{12} + m_{22})} \left(\frac{\Psi_1}{\sqrt{\det \Psi_1}} \right) \frac{1}{\sqrt{\det \Psi_1}} J_{2\Lambda + 1}(2\sqrt{\det \Psi_1})\, \delta_{m'_{11} m_{11}}$$

$$= [(m'_{12} - m'_{22} + 1)(m_{12} - m_{22} + 1)]^{1/2} \sum (-1)^{m'_{12} + m_{12} + y + \rho}$$

$$\times D^{(1/2)(m'_{12} - m'_{22})}_{m'_{11} - (1/2)(m'_{12} + m'_{22}) \quad -m'_{23} + \rho + (1/2)(m'_{12} + m'_{22})} \left(-\frac{\Sigma_x \Psi_1}{\sqrt{\det \Psi_1}} \right)$$

$$\times D^{(1/2)(m_{12} - m_{22})}_{m_{23} - \rho - (1/2)(m_{12} + m_{22}) \quad m_{11} - (1/2)(m_{12} + m_{22})} \left(\frac{\Psi_1}{\sqrt{\det \Psi_1}} \right) [\det \Psi_1]^{y + (1/2)(m'_{12} - m'_{22} + m_{12} - m_{22})}$$

$$\times N^{-1}_{-m'_{23} + \rho + (1/2)(m'_{12} + m'_{22}),\, m_{23} - (1/2)(m_{12} + m_{22});\, -m'_{23} + (1/2)(m'_{12} + m'_{22}),\, m_{23} - \rho - (1/2)(m_{12} + m_{22})} [y$$

$$+ 2 | \tfrac{1}{2}(m'_{12} - m'_{22}), \tfrac{1}{2}(m_{12} - m_{22})]\, \delta_{m'_{11}, m_{11}}, \tag{4.10}$$

where

$$\Psi_1 = \begin{pmatrix} \bar{\zeta}^2_1 \pi^1_1 + \zeta^2_1 \bar{\pi}^2_1 & -\zeta^2_2 \pi^2_2 \\ \zeta^2_2 \pi^2_2 & \zeta^2_1 \pi^1_1 + \zeta^2_2 \bar{\pi}^2_1 \end{pmatrix}, \tag{4.11}$$

$$\Sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.12}$$

For cases (2), $\zeta^2_2 \pi^2_2 = 0$, we have

$$\left\langle\!\!\left\langle \begin{pmatrix} & m'_{23} & \\ \Pi; \; m'_{12} & & m'_{22} \\ & m'_{11} & \end{pmatrix} \right| \mathscr{O}(\mathbb{1}, Z) \left| \begin{pmatrix} & m_{23} & \\ \Pi; \; m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix} \right\rangle\!\!\right\rangle = [(m'_{12} - m'_{22} + 1)(m_{12} - m_{22} + 1)]^{1/2}$$

$$\times \sum_{\Lambda,\rho} (-1)^{(1/2)(m'_{12} + m'_{22} + m_{12} + m_{22}) + m_{23} + \Lambda + \rho}\, C^{(1/2)(m'_{12} - m'_{22}) \quad (1/2)(m_{12} - m_{22}) \quad \Lambda}_{-m_{23} + (1/2)(m'_{12} + m'_{22}) \quad m_{23} - (1/2)(m_{12} + m_{22}) \quad (1/2)(m'_{12} + m'_{22} - m_{12} - m_{22})}$$

$$\times C^{(1/2)(m'_{12} - m'_{22}) \quad (1/2)(m_{12} - m_{22}) \quad \Lambda}_{(1/2)(m'_{12} + m'_{22}) - \rho \quad \rho - (1/2)(m_{12} + m_{22}) \quad (1/2)(m'_{12} + m'_{12} - m_{12} - m_{22})}$$

$$\times D^{(1/2)(m'_{12} - m'_{22})}_{m'_{11} - (1/2)(m'_{12} - m'_{22}) \quad +\rho + (1/2)(m'_{12} + m'_{22})} \left(\frac{-\Sigma_x \Psi_2}{\sqrt{\det \Psi_2}} \right) D^{(1/2)(m_{12} - m_{22})}_{\rho - (1/2)(m_{12} + m_{22}) \quad m_{11} - (1/2)(m_{12} + m_{22})} \left(\frac{\Psi_2}{\sqrt{\det \Psi_2}} \right)$$

$$+ \frac{1}{\sqrt{\det \Psi_2}} J_{2\Lambda + 1}(2\sqrt{\det \Psi_2})\, \delta_{m'_{23}, m_{23}}$$

$$= [(m'_{12} - m'_{22} + 1)(m_{12} - m_{22} + 1)]^{1/2} \sum_{y,\rho} (-1)^{m'_{12} + m_{22} + m_{23} + \rho + y} D^{(1/2)(m'_{12} - m'_{22})}_{m'_{11} - (1/2)(m'_{12} + m'_{22}) \quad -\rho + (1/2)(m'_{12} + m'_{22})} \left(\frac{-\Sigma_x \Psi_2}{\sqrt{\det \Psi_2}} \right)$$

$$\times D^{(1/2)(m_{12} - m_{22})}_{\rho - (1/2)(m_{12} + m_{22}) \quad m_{11} - (1/2)(m_{12} + m_{22})} \left(\frac{\Psi_2}{\sqrt{\det \Psi_2}} \right) [\det \Psi_2]^{y + (1/2)(m'_{12} - m'_{22}) + (1/2)(m_{12} - m_{22})}$$

$$\times N^{-1}_{-m'_{23} + (1/2)(m'_{12} + m'_{22}), \rho - (1/2)(m_{12} + m_{22});\, -\rho + (1/2)(m'_{12} + m'_{22}), m_{23} - (1/2)(m_{12} + m_{22})} [y$$

$$+ 2 | \tfrac{1}{2}(m'_{12} - m'_{22}), \tfrac{1}{2}(m_{12} - m_{22})]\, \delta_{m'_{23}, m_{23}}, \tag{4.13}$$

where

$$\Psi_2 = \begin{pmatrix} \bar{\zeta}^2_1 \pi^1_1 + \zeta^2_2 \bar{\pi}^1_1 & -\zeta^1_1 \pi^1_1 \\ \zeta^1_1 \pi^1_1 & \zeta^2_1 \pi^1_1 + \zeta^2_2 \pi^2_1 \end{pmatrix}.$$

Particular cases of these matrix elements for singular $Z\Pi^\dagger$ are also of interest. If we set

$$Z=\begin{pmatrix} z_1^3 & 0 \\ z_2^3 & 0 \end{pmatrix}, \tag{4.14}$$

then we find

$$\left\langle\!\!\left\langle \left(\Pi;\ \begin{array}{ccc} & m'_{23} & \\ m'_{12} & & m'_{22} \\ & m'_{11} & \end{array}\right) \right| \mathscr{O}(1,Z) \left| \left(\Pi;\ \begin{array}{ccc} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{array}\right) \right\rangle\!\!\right\rangle$$

$$= [(m'_{12}-m'_{22}+1)(m_{12}-m_{22}+1)]^{1/2}$$

$$\times \sum_{\Lambda,\rho} (-1)^{(1/2)(m'_{12}-m'_{22})+(1/2)(m_{12}-m_{22})-\Lambda+m_{23}+\rho}\, C^{(1/2)(m'_{12}-m'_{22})\quad (1/2)(m_{12}-m_{22})\quad \Lambda}_{-\rho+(1/2)(m'_{12}+m'_{22})\quad \rho-(1/2)(m_{12}+m_{22})\quad (1/2)(m'_{12}+m'_{22}-m_{12}-m_{22})}$$

$$\times C^{(1/2)(m'_{12}-m'_{22})\quad (1/2)(m_{12}-m_{22})\quad \Lambda}_{-m_{23}+(1/2)(m'_{12}+m'_{22})\quad m_{23}-(1/2)(m_{12}+m_{22})\quad (1/2)(m'_{12}+m'_{22}-m_{12}-m_{22})}$$

$$\times D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\quad -\rho+(1/2)(m'_{12}+m'_{22})}\left(\frac{\zeta_3}{\sqrt{\det\zeta_3}}\right) D^{(1/2)(m_{12}-m_{22})}_{-\rho+(1/2)(m_{12}+m_{22})\quad m_{11}-(1/2)(m_{12}+m_{22})}\left(\frac{-\zeta_3^\dagger}{\sqrt{\det\zeta_3}}\right)$$

$$\times \frac{1}{\pi_1^1\sqrt{\det\zeta_3}} J_{2\Lambda+1}(2\pi_1^1\sqrt{\det\zeta_3})\delta_{m'_{23},m_{23}}$$

$$= [(m'_{12}-m'_{22}+1)(m_{12}-m_{22}+1)]^{1/2} \sum_{\rho,y}(-1)^{m_{23}+\rho+y}(\pi_1^1)^{m'_{12}-m'_{22}+m_{12}-m_{22}}[(\pi_1^1)^2(\det\zeta_3)]^y$$

$$\times N^{-1}_{-\rho+(1/2)(m'_{12}+m'_{22}),m_{23}-(1/2)(m_{12}+m_{22}),-m_{23}+(1/2)(m'_{12}+m'_{22}),\rho-(1/2)(m_{12}+m_{22})}[y+2|\tfrac{1}{2}(m'_{12}-m'_{22}),\tfrac{1}{2}(m_{12}-m_{22})]$$

$$\times D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\quad -\rho+(1/2)(m'_{12}+m'_{22})}(\zeta_3)\, D^{(1/2)(m_{12}-m_{22})}_{\rho-(1/2)(m_{12}+m_{22})\quad m_{11}-(1/2)(m_{12}+m_{22})}(-\zeta_3^\dagger)\delta_{m'_{23},m_{23}}, \tag{4.15}$$

where

$$\zeta_3=\begin{pmatrix} z_1^3 & -\bar z_2^3 \\ z_2^3 & \bar z_1^3 \end{pmatrix}. \tag{4.16}$$

Similarly, if

$$Z=\begin{pmatrix} 0 & z_1^4 \\ 0 & z_2^4 \end{pmatrix}, \tag{4.17}$$

then we find

$$\left\langle\!\!\left\langle \left(\Pi;\ \begin{array}{ccc} & m'_{23} & \\ m'_{12} & & m'_{22} \\ & m'_{11} & \end{array}\right) \right| \mathscr{O}(1,Z) \left| \left(\Pi;\ \begin{array}{ccc} & m_{23} & \\ m_{12} & & m_{22} \\ & m_{11} & \end{array}\right) \right\rangle\!\!\right\rangle$$

$$= [(m'_{12}-m'_{22}+1)(m_{12}-m_{22}+1)]^{1/2}$$

$$\times \sum_{\Lambda,\rho,\sigma} (-1)^{(1/2)(m'_{12}-m'_{22})+(1/2)(m_{12}-m_{22})-\Lambda+\rho+\sigma}\, C^{(1/2)(m_{12}-m_{22})\quad (1/2)(m'_{12}-m'_{22})\quad \Lambda}_{(1/2)(m_{12}+m_{22})-\sigma\quad \sigma-(1/2)(m'_{12}+m'_{22})\quad (1/2)(m_{12}+m_{22}-m'_{12}-m'_{22})}$$

$$\times C^{(1/2)(m_{12}-m_{22})\quad (1/2)(m'_{12}-m'_{22})\quad \Lambda}_{(1/2)(m_{12}+m_{22})-\rho\quad \rho-(1/2)(m'_{12}+m'_{22})\quad (1/2)(m_{12}+m_{22}-m'_{12}-m'_{22})}$$

$$\times D^{(1/2)(m'_{12}-m'_{22}}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\quad -\sigma+(1/2)(m'_{12}+m'_{22})}\left(\frac{\zeta_4}{\sqrt{\det\zeta_4}}\right) D^{(1/2)(m'_{12}-m'_{22})}_{-m'_{23}+(1/2)(m'_{12}+m'_{22})\quad -\rho+(1/2)(m'_{12}+m'_{22})}\left(\frac{\hat\Pi}{\sqrt{\det\hat\Pi}}\right)$$

$$\times D^{(1/2)(m_{12}-m_{22})}_{-\sigma+(1/2)(m_{12}+m_{22})\quad m_{11}-(1/2)(m_{12}+m_{22})}\left(\frac{-\zeta_4^\dagger}{\sqrt{\det\zeta_4}}\right) D^{(1/2)(m_{12}-m_{22})}_{-\rho+(1/2)(m_{12}+m_{22})\quad -m_{23}+(1/2)(m_{12}+m_{23})}\left(\frac{\hat\Pi^\dagger}{\sqrt{\det\hat\Pi}}\right)$$

$$\times \frac{1}{[\det\zeta_4)(\det\widehat{\Pi})]^{1/2}} J_{2\Lambda+1}\{2[(\det\zeta_4)(\det\widehat{\Pi})]^{1/2}\}$$

$$= [(m'_{12} - m'_{22} + 1)(m_{12} - m_{22} + 1)]^{1/2} \sum_{\rho,\sigma,y} (-1)^{y+\rho+\sigma}[(\det\zeta_4)(\det\widehat{\Pi})]^y$$
$$\times N^{-1}_{(1/2)(m_{12}+m_{22})-\sigma,\rho-(1/2)(m'_{12}+m'_{22});(1/2)(m_{12}+m_{22})-\rho,\sigma-(1/2)(m'_{12}+m'_{22})} [y + 2\,|\tfrac{1}{2}(m_{12} - m_{22}),\tfrac{1}{2}(m'_{12} - m'_{22})]$$
$$\times D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\;\;-\sigma+(1/2)(m'_{12}+m_{22})}(\zeta_4)$$
$$\times D^{(1/2)(m'_{12}-m'_{22})}_{-m'_{23}+(1/2)(m'_{12}+m'_{22})\;-\rho+(1/2)(m'_{12}+m'_{22})}(\widehat{\Pi})$$
$$\times D^{(1/2)(m_{12}-m_{22})}_{-\sigma+(1/2)(m_{12}+m_{22})\;\;m_{11}-(1/2)(m_{12}+m_{22})}(-\zeta_4^{\dagger}) D^{(1/2)(m_{12}-m_{22})}_{(1/2)(m_{12}+m_{22})-\rho\;-m_{23}+(1/2)(m_{12}+m_{22})}(\widehat{\Pi}^{\dagger}), \tag{4.18}$$

where

$$\zeta_4 = \begin{pmatrix} z_1^4 & -\bar{z}_2^4 \\ z_2^4 & \bar{z}_1^4 \end{pmatrix},$$
$$\widehat{\Pi} = \begin{pmatrix} \bar{\pi}_1^2 & \pi_2^2 \\ -\pi_2^2 & \pi_1^2 \end{pmatrix}. \tag{4.19}$$

These matrix elements for the case of singular $Z\Pi^{\dagger}$ are simpler than those for the nonsingular case, and it is easier to derive new relations for the special functions involved from them. We shall note some of the consequences of the representation property of the matrix elements (4.18). Taking two successive translations, through the matrices ζ and η [as in Eq. (4.19), with the dual object labeled by the matrix $\widehat{\Pi}$], we find

$$J_{2\Lambda''+1}\{2[\det(\zeta+\eta)\det\widehat{\Pi}]^{1/2}\} = \left[\frac{\det(\zeta+\eta)}{(\det\zeta)(\det\eta)(\det\widehat{\Pi})}\right]^{1/2}$$
$$\times \sum_{\Lambda,\Lambda',\rho,\sigma} (-1)^{\Lambda+\Lambda'-\Lambda''+\rho+\sigma}(2\Lambda+1)(2\Lambda''+1) C^{\Lambda'\;\;\Lambda\;\;\Lambda''}_{\rho-m'\;m-\rho\;m-m'}$$
$$\times C^{\Lambda'\;\;\Lambda\;\;\Lambda''}_{\sigma-m'\;m-\sigma\;m-m'}$$
$$\times D^{\Lambda'}_{m'-\sigma\;m'-\rho}\left(\frac{\eta^{\dagger}(\zeta+\eta)}{[(\det\eta)(\det(\zeta+\eta))]^{1/2}}\right) D^{\Lambda}_{m-\rho\;m-\sigma}\left(\frac{(\zeta^{\dagger}+\eta^{\dagger})\zeta}{\{(\det\zeta)[\det(\zeta+\eta)]\}^{1/2}}\right)$$
$$\times J_{2\Lambda+1}\{2[(\det\zeta)(\det\widehat{\Pi})]^{1/2}\} J_{2\Lambda'+1}\{2[(\det\eta)(\det\widehat{\Pi})]^{1/2}\}. \tag{4.20}$$

Setting

$$\zeta = \begin{pmatrix} v_1 & -\bar{v}_2 \\ v_2 & \bar{v}_1 \end{pmatrix}, \quad \eta = \begin{pmatrix} w_1 & -\bar{w}_2 \\ w_2 & \bar{w}_1 \end{pmatrix},$$
$$\det\zeta = |v_1|^2 + |v_2|^2, \quad \det\eta = |w_1|^2 + |w_2|^2, \tag{4.21}$$
$$\det(\zeta+\eta) = |v_1 + w_1|^2 + |v_2 + w_2|^2,$$

we find that Eq. (4.20) gives us a combination of classical Bessel functions which effect the addition of the moduli of vectors in two complex planes at once, in the sense that

$$J_{2\Lambda''+1}\{2\sqrt{\det\widehat{\Pi}}\,[|v_1 + w_1|^2 + |v_2 + w_2|^2]^{1/2}\}$$
$$= \sum_{\Lambda,\Lambda'} (C^{\Lambda''}_{\Lambda,\Lambda'}) J_{2\Lambda+1}\{2\sqrt{\det\widehat{\Pi}}\,[|v_1|^2 + |v_2|^2]^{1/2}\} J_{2\Lambda'+1}(2\sqrt{\det\widehat{\Pi}}\,[|w_1|^2 + |w_2|^2]^{1/2}), \tag{4.22}$$

whereas the representation property of the Bessel function as matrix element of a finite transformation in the Euclidean group of the plane gives us the addition of vectors in only one complex plane:

$$J_{m'-m}[p(z_1 + z_2)] = \sum_{m''=-\infty}^{+\infty} J_{m'-m''}(pz_1) J_{m''-m}(pz_2). \tag{4.23}$$

The sums in Eqs. (4.20) and (4.22) are carried out over all integral and half-integral values of Λ and Λ' such that the conditions of the Clebsch–Gordan series for SU(2) are fulfilled.

The representation property of the functions (4.18) allows us to formulate an analog of the binomial theorem for the U(2) gamma matrix. If we represent a translation of the type (4.17) through the matrix $\zeta_4 + \eta_4$ as two successive translations through ζ_4 and η_4, using the representation property of the functions (4.17), then equate the coefficients of the same power of $\det\hat{\Pi}$ on both sides, and eliminate common factors, we obtain the relation

$$\sum_{\lambda,\sigma,\sigma',x,\eta} (-1)^{2\lambda+\sigma+\sigma'+\rho}(2\lambda+1)(\det\zeta_4)^x(\det\eta_4)^{n-x-2\lambda} D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\ \ -\sigma'+(1/2)(m'_{12}+m'_{22})}(\zeta_4)$$
$$\times N^{-1}_{\rho-(1/2)(m'_{12}+m'_{22}),\mu-\sigma';\sigma'-(1/2)(m'_{12}+m'_{22}),\mu-\rho}[x+2|(1/2)(m'_{12}-m'_{22}),\lambda\ ]$$
$$\times D^{\lambda}_{\mu-\sigma'\mu-\sigma}(-\zeta_4^\dagger\eta_4)N^{-1}_{\rho-\mu,(1/2)(m_{12}+m_{22})-\sigma;\sigma-\mu,(1/2)(m_{12}+m_{22})-\rho}[n-x-2\lambda+2|\lambda,\tfrac{1}{2}(m_{12}-m_{22})]$$
$$\times D^{(1/2)(m_{12}-m_{22})}_{(1/2)(m_{12}+m_{22})-\sigma\ m_{11}-(1/2)(m_{12}+m_{22})}(-\eta_4^\dagger)$$
$$=\sum_\tau(-1)^\tau[\det(\zeta_4+\eta_4)]^n D^{(1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\ \ -\tau+(1/2)(m'_{12}+m'_{22})}(\zeta_4+\eta_4)$$
$$\times D^{(1/2)(m_{12}-m_{22})}_{(1/2)(m_{12}+m_{22})-\tau\ m_{11}-(1/2)(m_{12}+m_{22})}(-\zeta_4^\dagger-\eta_4^\dagger)$$
$$\times N^{-1}_{\rho-(1/2)(m'_{12}+m'_{22}),(1/2)(m_{12}+m_{22})-\tau;\tau-(1/2)(m'_{12}+m'_{22}),(1/2)(m_{12}+m_{22})-\rho}[n+2|\tfrac{1}{2}(m'_{12}-m'_{22}),\tfrac{1}{2}(m_{12}-m_{22})], \quad (4.24)$$

where the sum is taken over both integral and half-integral values of λ. In particular, we can move the Wigner D functions and the gamma matrix from the right side to the left side by matrix multiplication and obtain an expansion for $\det(\zeta_4 + \eta_4)$, as announced in Eq. (9.3) of Ref. 2. In the simplest case, setting $\tfrac{1}{2}(m_{12} - m_{22}) = \tfrac{1}{2}(m'_{12} - m'_{22}) = 0$, we find

$$[\det(\zeta_4+\eta_4)]^n = \sum_\lambda \binom{n}{2\lambda}\chi^\lambda\left(\frac{\zeta_4^\dagger\eta_4}{\sqrt{\det\zeta_4}\sqrt{\det\eta_4}}\right)(\det\zeta_4)^\lambda(\det\eta_4)^{n-\lambda}{}_2F_1\left(-n+2\lambda,-n-1|2\lambda+2|\frac{\det\zeta_4}{\det\eta_4}\right), \quad (4.25)$$

where $\chi^\lambda(u)$ denotes the primitive character of the $(2\lambda + 1)$-dimensional representation of SU(2) corresponding to the group element u. Again, the sum is taken over all integral and half-integral values of λ such that $0 \leqslant 2\lambda \leqslant n$. This result can be generalized to matrices other than those of the form specified in Eqs. (4.16) and (4.19) by analytic continuation. We note that $\zeta_4 = (\det\zeta_4)\zeta_4^{-1}$; making this replacement, we can immediately extend Eq. (4.25) to all pairs of 2×2 matrices at least one of which is nonsingular. We can express Eq. (4.25) more succinctly by making use of the generalization of the hypergeometric series defined by Louck and Biedenharn.[10] With them, we define

$${}_p\mathscr{F}_q(a_1,\ldots,a_q;b_1,\ldots,b_q;z) = \sum_{[\mu]_t}\langle {}_p\mathscr{F}_q(a_1,\ldots,a_p;b_1,\ldots,b_q)|\mu\rangle\langle\mu|z\rangle, \quad (4.26)$$

where $\langle\mu|z\rangle$ is the Schur function (or primitive character) with argument $z\in\mathbb{C}^{t\times t}$ belonging to the $[\mu]_t = [\mu_{1t},\ldots,\mu_{tt}]$ irreducible representation of $GL(t,\mathbb{C})$ and

$$\langle {}_p\mathscr{F}_q(a_1,\ldots,a_p;b_1,\ldots,b_q)|\mu\rangle = \frac{1}{M(\mu)}\prod_{i=1}^{p}\prod_{s=1}^{t}(a_i-s+1)_{\mu_{st}}\prod_{j=1}^{q}\prod_{s=1}^{t}\frac{1}{(b_j-s+1)_{\mu_{st}}}, \quad (4.27)$$

where

$$M(\mu) = \frac{\Pi_{s=1}^t(\mu_{st}+t-s)!}{\Pi_{r<s}(\mu_{rt}-\mu_{st}+s-r)} \quad (4.28)$$

and, as usual,

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}. \quad (4.29)$$

In this notation, then, Eq. (4.25) becomes

$$[\det(\mathbb{1}+\zeta)]^n = {}_1\mathscr{F}_0(-n;;\zeta), \quad t=2, \quad (4.30)$$

in exact analogy to the classical binomial theorem

$$(1+z)^n = {}_1F_0(-n;;-z). \quad (4.31)$$

Louck and Biedenharn have established Saalschütz' theorem and the Euler transform for their generalized hypergeometric series for all t; in Eq. (4.30) we obtain a third analog of a classical hypergeometric theorem for the case $t = 2$. We conjecture its validity for all t, with a minus sign before the argument in the case of odd t.

V. THE LIE ALGEBRA OF $U(2) \textcircled{s} C^{2\times 2}$ REALIZED AS AN ALGEBRA OF DIFFERENTIAL OPERATORS ON THE GROUP MANIFOLD

We must now realize the Lie algebra of the group $U(2)\textcircled{s}C^{2\times 2}$ in terms of differential operators on the group manifold. We shall relegate to Appendix B the verification that the representation matrices which we obtained in the previous section do indeed have the correct transformation properties under the Lie algebra of differential operators. The derivation of the required operators can be achieved through the contraction process, so we shall first examine the comparable set of operators in U(4), following the treatment of Louck.[8]

The generators of U(4) can be realized as differential operators on the group manifold in two distinct, mutually isomorphic ways. We perform the mappings

$$E_{ij} \to \mathscr{E}_{ij} = \sum_{k=1}^{4} u_i^k \frac{\partial}{\partial u_j^k}, \tag{5.1a}$$

$$E_{ij} \to \mathscr{E}^{ij} = \sum_{k=1}^{4} u_k^i \frac{\partial}{\partial u_k^j}, \tag{5.1b}$$

and the resulting operator obeys the commutation rules

$$[\mathscr{E}_{ij}, \mathscr{E}_{k\ell}] = \mathscr{E}_{i\ell}\delta_{jk} - \mathscr{E}_{kj}\delta_{i\ell}, \tag{5.2a}$$

$$[\mathscr{E}^{ij}, \mathscr{E}^{k\ell}] = \mathscr{E}^{i\ell}\delta_{jk} - \mathscr{E}_{kj}\delta_{i\ell}, \tag{5.2b}$$

$$[\mathscr{E}_{ij}, \mathscr{E}^{k\ell}] = 0. \tag{5.2c}$$

These sets of differential operators, then, realize two distinct copies of the Lie algebra of U(4) which are "kinematically independent" in the sense of Eq. (5.2c). In Eqs. (5.1) and (5.2) all indices range over the integers from 1 to 4, and the variables u_i^k denote the matrix entries of the defining (4×4) representation of the group U(4).

The matrix elements of a general finite transformation in an irreducible representation of U(4), i.e.,

$$D^{[m]_4}_{(m')_3(m)_3}(u) \tag{5.3}$$

in the notation of Louck, transform as basis states of an irreducible representation of the Lie algebra in the following manner:

$$\mathscr{E}_{ij} D^{[m]_4}_{(m')_3(m)_3}(u) = \sum_{(m'')_3} \left\langle \begin{pmatrix} [m]_4 \\ (m'')_3 \end{pmatrix} \middle| E_{ij} \middle| \begin{pmatrix} [m]_4 \\ (m')_3 \end{pmatrix} \right\rangle \times D^{[m]_4}_{(m'')_3(m)_3}(u), \tag{5.4a}$$

$$\mathscr{E}^{ij} D^{[m]_4}_{(m')_3(m)_3}(u) = \sum_{(m'')_3} \overline{\left\langle \begin{pmatrix} [m]_4 \\ (m'')_3 \end{pmatrix} \middle| E_{ij} \middle| \begin{pmatrix} [m]_4 \\ (m)_3 \end{pmatrix} \right\rangle} \times D^{[m]_4}_{(m')_3(m'')_3}(u), \tag{5.4b}$$

where the bar denotes complex conjugation. In U(4) the matrix elements of the operators E_{ij} may be taken to be real by a suitable choice of phase conventions. Hence, the matrix elements (5.3) are simultaneous eigenfunctions of the corresponding two distinct copies of the invariant operators of U(4) and its subgroups and the operators of the Cartan subalgebras $\mathscr{E}_{ii}$, $\mathscr{E}^{ii}$, $1\leqslant i\leqslant 4$. Under the operators

$$I_k^{(4)} \equiv \sum_{i_1,\dots,i_k} \mathscr{E}_{i_1 i_2}\mathscr{E}_{i_2 i_3}\cdots\mathscr{E}_{i_k i_1}, \quad 1\leqslant k\leqslant 4, 1\leqslant i_\ell \leqslant 4, \tag{5.5a}$$

$$I_k^{(j)} \equiv \sum_{i_1,\dots,i_k} \mathscr{E}_{i_1 i_2}\mathscr{E}_{i_2 i_3}\cdots\mathscr{E}_{i_k i_1}, \quad 1\leqslant k\leqslant j, 1\leqslant i_\ell \leqslant j, \tag{5.5b}$$

the matrix elements (5.3) are eigenfunctions which yields eigenvalues determined by the invariants $[m]_4$ [in the case of Eq. (5.5a)] and the entries in the jth row from the bottom of the left Gel'fand pattern $(m')_3$ [in the case of Eq. (5.5b)]. Under the corresponding invariant operators composed out of the generators $\mathscr{E}^{ij}$ instead of $\mathscr{E}_{ij}$, the matrix elements (5.3) belong to eigenvalues determined by $[m]_4$ and the entries $(m)_3$ of the right Gel'fand pattern.

We must now examine the contraction of the operators (5.1) to those of the group $[U(2)\textcircled{s}C^{2\times 2}]\times \mathbb{H}^{2\times 2}$ as specified in Eq. (2.8) above. We recall that the general element of the contracted group is obtained by restriction of some of the entries in the defining representation to infinitesimal quantities or quantities infinitesimally close to unity:

$$\begin{pmatrix} u_1^1 & u_1^2 & u_1^3 & u_1^4 \\ u_2^1 & u_2^2 & u_2^3 & u_2^4 \\ u_3^1 & u_3^2 & u_3^3 & u_3^4 \\ u_4^1 & u_4^2 & u_4^3 & u_4^4 \end{pmatrix} \underset{R\to\infty}{\longrightarrow} \begin{pmatrix} u_1^1 & u_1^2 & \frac{1}{R}z_1^3 & \frac{1}{R}z_1^4 \\ u_2^1 & u_2^2 & \frac{1}{R}z_2^3 & \frac{1}{R}z_2^4 \\ \frac{1}{R}z_3^1 & \frac{1}{R}z_3^2 & 1+\frac{1}{R}z_3^3 & \frac{1}{R}z_3^4 \\ \frac{1}{R}z_4^1 & \frac{1}{R}z_4^2 & \frac{1}{R}z_4^3 & 1+\frac{1}{R}z_4^4 \end{pmatrix}, \tag{5.6}$$

where the condition that the above matrix be unitary (to order R^{-1}) implies the conditions

$$\begin{pmatrix} z_3^1 & z_3^2 \\ z_4^1 & z_4^2 \end{pmatrix} = -\begin{pmatrix} \bar z_1^3 & \bar z_2^3 \\ \bar z_1^4 & \bar z_2^4 \end{pmatrix}\begin{pmatrix} u_1^1 & u_1^2 \\ u_2^1 & u_2^2 \end{pmatrix} \equiv -Z^\dagger U, \tag{5.7a}$$

$$\begin{pmatrix} z_3^3 & z_3^4 \\ z_4^3 & z_4^4 \end{pmatrix} = -\begin{pmatrix} \bar z_3^3 & \bar z_4^3 \\ \bar z_3^4 & \bar z_4^4 \end{pmatrix} \equiv iH, \tag{5.7b}$$

where

$$U = \begin{pmatrix} u_1^1 & u_1^2 \\ u_2^1 & u_2^2 \end{pmatrix} \tag{5.8}$$

is unitary. Hence, by the prescription (2.8), these generators become under contraction

$$\mathscr{E}_{ij} \to u_i^1 \frac{\partial}{\partial u_j^1} + u_i^2 \frac{\partial}{\partial u_j^2} + z_i^3 \frac{\partial}{\partial z_j^3} + z_i^4 \frac{\partial}{\partial z_j^4}, \quad 1\leqslant i,j\leqslant 2, \tag{5.9a}$$

$$P_{iJ} = u_i^1 \frac{\partial}{\partial z_J^1} + u_i^2 \frac{\partial}{\partial z_J^2} = -\frac{\partial}{\partial \bar{z}_i^J}, \tag{5.9b}$$

$$P_{Ij} = \frac{\partial}{\partial z_j^I}, \tag{5.9c}$$

and

$$P_{IJ} = \frac{\partial}{\partial z_J^I}, \tag{5.10}$$

where we maintain the prescription that lower case indices take the values 1,2, and capital indices take the values 3,4.

We can realize the operators of the Lie algebra in the isomorphic form

$$\mathscr{E}^{ij} \to u_1^i \frac{\partial}{\partial u_1^j} + u_2^i \frac{\partial}{\partial u_2^j} + z_3^i \frac{\partial}{\partial z_3^j} + z_4^i \frac{\partial}{\partial z_4^j}$$

$$= \sum_{\ell} u_\ell^i \frac{\partial}{\partial u_\ell^j} + \sum_{\ell,\ell',K} u_\ell^i \bar{u}_{\ell'}^j \bar{z}_\ell^K \frac{\partial}{\partial z_{\ell'}^K}, \tag{5.11a}$$

$$P^{iJ} = u_1^i \frac{\partial}{\partial z_1^J} + u_2^i \frac{\partial}{\partial z_2^J}, \tag{5.11b}$$

$$P^{Ij} = \frac{\partial}{\partial z_I^j} = -\bar{u}_1^j \frac{\partial}{\partial \bar{z}_1^I} - \bar{u}_2^j \frac{\partial}{\partial \bar{z}_2^I}, \tag{5.11c}$$

and

$$P^{IJ} = \frac{\partial}{\partial z_I^J}. \tag{5.12}$$

The invariant operators of the group $U(2) \circledS C^{2\times 2}$ are then given by

$$P_{13}P_{31} + P_{23}P_{32} = -\frac{\partial^2}{\partial \bar{z}_1^3 \partial z_1^3} - \frac{\partial^2}{\partial \bar{z}_2^3 \partial z_2^3}, \tag{5.13a}$$

$$P_{14}P_{41} + P_{24}P_{42} = -\frac{\partial^2}{\partial \bar{z}_1^4 \partial z_1^4} - \frac{\partial^2}{\partial \bar{z}_2^4 \partial z_2^4}, \tag{5.13b}$$

$$P_{13}P_{41} + P_{23}P_{42} = -\frac{\partial^2}{\partial \bar{z}_1^3 \partial z_1^4} - \frac{\partial^2}{\partial \bar{z}_2^3 \partial z_2^4}, \tag{5.13c}$$

$$P_{31}P_{14} + P_{32}P_{24} = -\frac{\partial^2}{\partial z_1^3 \partial \bar{z}_1^4} - \frac{\partial^2}{\partial z_2^3 \partial \bar{z}_2^4}. \tag{5.13d}$$

The invariant operators realized in terms of the generators with upper indices (5.11) are identical with those given in Eqs. (5.13). It may be verified by a tedious but straightforward calculation that the three invariant operators $I_k^{(4)}$, $2 \leqslant k \leqslant 4$, given in Eq. (5.5a), become polynomials in the operators (5.10) and (5.13) in the contraction limit.

ACKNOWLEDGMENT

The author would like to thank Dr. Kenneth I. Gross for generous encouragement and helpful criticism throughout the course of this work, which was directly inspired by his own work on the generalization of the Bessel function.

APPENDIX A: THE CONTRACTION OF THE MATRICES OF IRREDUCIBLE REPRESENTATIONS OF U(4) TO REPRESENTATIONS OF $[U(2) \sigma C^{2\times 2}] \times H^{2\times 2}$

Our starting point is the matrix element of an irreducible representation of U(4); Holman[11] has constructed the matrix elements of all irreducible representations of all $U(n)$ in the canonical Gel'fand basis in the form

$$D^{[m]_n}_{(m')_{n-1},(m)_{n-1}}(u) = \mathscr{M}^{1/2}([m]_n) \sum_{[\alpha]_{n-1}} \sum_{(\alpha')_{n-2}} \sum_{(\alpha)_{n-2}} \mathscr{M}^{-1/2}[\alpha]_{n-1} D^{[\alpha]_{n-1}}_{(\alpha')_{n-2},(\alpha)_{n-2}}(u_{n-1})$$

$$\times \left\langle \begin{pmatrix} [m']_{n-1} \\ (m')_{n-2} \end{pmatrix} \middle| \left\langle \sum_{i=1}^{n-1} (m'_{in-1} - \alpha_{in-1}) 0 \cdots 0 \right\rangle \middle| \begin{pmatrix} [\alpha]_{n-1} \\ (\alpha')_{n-2} \end{pmatrix} \right\rangle \prod_{i=1}^{n-1} \frac{(u_n^i)^{q_i}}{[(q_i)!]^{1/2}}$$

$$\times \left\langle \begin{pmatrix} [m]_{n-1} \\ (m)_{n-2} \end{pmatrix} \middle| \left\langle \sum_{i=1}^{n-1} (m_{in-1} - \alpha_{in-1}) 0 \cdots 0 \right\rangle \middle| \begin{pmatrix} [\alpha]_{n-1} \\ (\alpha)_{n-2} \end{pmatrix} \right\rangle \prod_{i=1}^{n-1} \frac{(u_i^n)^{p_i}}{[(p_i)!]^{1/2}}$$

$$\times \left\langle \begin{pmatrix} [m]_n \\ (m)_{n-1} \end{pmatrix} \middle| \left[\left(\sum_{i=1}^{n} m_{in} - \sum_{i=1}^{n-1} m'_{in-1} \right) 0 \cdots 0 \right] \middle| \begin{pmatrix} m'_{in-1} \cdots m'_{n-in-1} & 0 \\ (\alpha)_{n-1} & \end{pmatrix} \right\rangle \frac{(u_n^n)^{p_n}}{[(p_n)!]^{1/2}}, \tag{A1}$$

where the u_j^i are matrix elements of the $n \times n$ defining representation, u_{n-1} denotes the $(n-1)\times(n-1)$ submatrix of u which results from the removal of the nth column and the nth row, and the normalization constant is given by

$$\mathscr{M}([m]_n) = \frac{\prod_{i=1}^n (m_{in} + n - i)!}{\prod_{i<j}^n (m_{in} - m_{jn} + j - i)}. \tag{A2}$$

The coefficients $\langle | \langle\ \rangle | \rangle$ denote matrix elements of totally symmetric unit tensor operators in $U(n-1)$ while the coefficient $\langle | [\ ] | \rangle$ is the matrix element of a reduced $U(n){:}U(n-1)$ tensor operator. In our notation for these matrix elements we follow the conventions of Louck[8] rather than those of Chacón, Ciftan, and Biedenharn.[12] We shall omit the upper and lower patterns of the tensor operators since, in the case of totally symmetric operators, these are uniquely determined by the initial and final states. The exponents in Eq. (A1) are given by

$$q_1 = m'_{11} - \alpha'_{11}; \quad q_i = \sum_{j=1}^{i} (m'_{ji} - \alpha'_{ji}) - \sum_{j=1}^{i-1} (m'_{ji-1} - \alpha'_{ji-1}), \quad 2 \leqslant i \leqslant n-1;$$

$$p_1 = m_{11} - \alpha_{11};\quad p_i = \sum_{j=1}^{i} (m_{ji} - \alpha_{ji}) - \sum_{j=1}^{i-1} (m_{ji-1} - \alpha_{ji-1}),\quad 2\leqslant i \leqslant n-1;$$

$$p_n = \sum_{i=1}^{n} m_{in} - \sum_{i=1}^{n-1} (m'_{in-1} + m_{in-1} - \alpha_{in-1}). \tag{A3}$$

The relation (A1), then, gives us a recursive construction of the matrix elements of all irreducible representations of all U(n). The matrix elements of reduced totally symmetric tensor operators are necessary for the construction and have been evaluated by Chacón, Ciftan, and Biedenharn.[12] The result is

$$\left\langle\!\left(\begin{matrix}[m]_n\\(m)_{n-1}\end{matrix}\right)\right|\left[\left(\sum_{i=1}^{n} m_{in} - \sum_{i=1}^{n-1} m'_{in-1}\right)0\cdots 0\right]\left|\left(\begin{matrix}m'_{1n-1}\cdots m'_{n-1n-1}\,0\\(\alpha)_{n-1}\end{matrix}\right)\right\rangle$$

$$= \left[\left(\sum_{i=1}^{n} m_{in} - \sum_{i=1}^{n-1} (m_{in-1} + m'_{in-1} - \alpha_{in-1})\right)!\right]^{1/2} \frac{S_{nn}(m_n;m_n)S_{nn-1}(m'_{n-1};\alpha_{n-1})}{S_{nn}(m_n;m'_{n-1})S_{nn-1}(m_n;m_{n-1})}$$

$$\times S_{n-1n-1}(m_{n-1};\alpha_{n-1})S_{n-1n-1}(\alpha_{n-1};\alpha_{n-1})$$

$$\times \sum_{\rho_1\cdots\rho_{n-1}} (-1)^{\rho_1+\cdots+\rho_{n-1}} \left[\frac{S_{nn-1}(m_n;\bar q_{n-1})S_{n-1n-1}(\bar q_{n-1};\bar q_{n-1})}{S_{nn-1}(m'_{n-1};\bar q_{n-1})S_{n-1n-1}(m_{n-1};\bar q_{n-1})S_{n-1n-1}(\bar q_{n-1};\alpha_{n-1})}\right]^2, \tag{A4}$$

where

$$S_{nr}(m_n;\mu_r) = \left[\frac{\Pi_{k=1}^{r}\Pi_{s=1}^{k}(m_{sn} - \mu_{kr} + k - s)!}{\Pi_{k=1}^{n-1}\Pi_{k+1}^{s}(\mu_{kr} - m_{sn} + s - k - 1)!}\right]^{1/2},\quad r = n, n-1. \tag{A5}$$

Here, m_n denotes the n-tuple of Gel'fand labels $[m_{1n},\cdots,m_{nn}]$, m'_{n-1} denotes the n-tuple $[m'_{1n-1},\cdots,m'_{n-1n-1},0]$ and m_{n-1} and α_{n-1} denote the corresponding $(n-1)$-tuples. The $(n-1)$-tuple $\bar q_{n-1}$ is given by $\alpha_{1n-1} + \rho_1,\cdots,\alpha_{n-1n-1} + \rho_{n-1}$.

We recall the relation

$$D^{[m]_n}_{(m')_{n-1}(m)_{n-1}}(u) = (\det u)^{m_{nn}} D^{[\mu]_n}_{(\mu')_{n-1}(\mu)_{n-1}}(u), \tag{A6}$$

where $\mu_{ij} = m_{ij} - m_{nn}$, $\mu'_{ij} = m'_{ij} - m_{nn}$. We shall consider the asymptotic limit of the D function on the right of Eq. (A6), taking as our argument the 4×4 matrix

$$\begin{pmatrix} u_1^1 & u_1^2 & \frac{1}{R}\zeta_1^3 & \frac{1}{R}\zeta_1^4 \\ u_2^1 & u_2^2 & \frac{1}{R}\zeta_2^3 & \frac{1}{R}\zeta_2^4 \\ \frac{1}{R}\zeta_3^1 & \frac{1}{R}\zeta_3^2 & 1+\frac{i}{R}h_3^3 & \frac{i}{R}h_3^4 \\ \frac{1}{R}\zeta_4^1 & \frac{1}{R}\zeta_4^2 & \frac{i}{R}h_4^3 & 1+\frac{i}{R}h_4^4 \end{pmatrix} = \begin{pmatrix} U & \frac{1}{R}\bar\zeta \\ \frac{1}{R}\zeta & 1+\frac{i}{R}H \end{pmatrix} \tag{A7}$$

and initial and final states of the form

$$\left|\left(\begin{matrix}[m]_4\\(m)_3\end{matrix}\right)\right\rangle$$

$$= \left|\begin{pmatrix} R(p_{14}+p_{44}) & & R(p_{24}+p_{44}) & & R(p_{44}-p_{34}) & & 0 \\ & R(p_{13}+p_{44})+m_{13} & & Rp_{44}+m_{23} & & R(p_{44}-p_{33})+m_{33} & \\ & & Rp_{44}+m_{12} & & Rp_{44}+m_{22} & & \\ & & & Rp_{44}+m_{11} & & & \end{pmatrix}\right\rangle \tag{A8}$$

The factor $(\det u)^{m_{nn}}$ in Eq. (A6) now has the form

$$(\det u)^{-Rp_{44}} = [\det U]^{-Rp_{44}}\left(1+\frac{i}{R}h_3^3\right)^{-Rp_{44}}\left(1+\frac{i}{R}h_4^4\right)^{-Rp_{44}} \underset{R\to\infty}{\longrightarrow} [\det U]^{-Rp_{44}} e^{i(h_3^3+h_4^4)p_{44}}. \tag{A9}$$

We now wish to take the asymptotic limit $R\to\infty$. In order to do so we shall make an assumption about the form which this asymptotic limit will take and, consequently, about what happens to the internal indices of summation in Eq. (A1). We shall assume that the exponent of any factor, such as

$$\frac{\zeta_j^i}{R}, \frac{h_j^i}{R}, \tag{A10}$$

which occurs with R in the denominator, will remain finite and contain only discrete indices of summation. We shall allow the exponents of factors of the form

$$\left(1+\frac{i}{R}h_j^j\right) \tag{A11}$$

to become infinite, i.e.,

$$\left(1+\frac{i}{R}h_j^j\right)^{Rq} \to e^{ih_j^j q}. \tag{A12}$$

With this assumption we can write down the internal state $[\alpha]_3$ immediately as

$$\left|\begin{pmatrix}[\alpha]_3\\(\alpha)_2\end{pmatrix}\right\rangle = \left|\begin{pmatrix} R(p_{13}+p_{44})+\alpha_{13} & & R(p_{44}+\alpha_{23}) & & R(p_{44}-p_{33})+\alpha_{33} \\ & Rp_{44}+\alpha_{12} & & Rp_{44}+\alpha_{22} & \\ & & Rp_{44}+\alpha_{11} & & \end{pmatrix}\right\rangle, \tag{A13}$$

and for the substates $(\alpha')_2$ we find labels of the same form as those given on the right of Eq. (A13) with $\alpha_{ij}, j\leqslant 2$, replaced by the corresponding primed quantities. We also find that a transformation of the form (A7) shifts the Gel'fand labels in Eq. (A8) only by a finite amount, i.e., that the transformation leaves all the p_{ij} invariant.

We must also apply Eq. (A1) in order to evaluate

$$\begin{aligned} D^{[\alpha]_3}_{(\alpha')_2(\alpha)_2}(u) = {} & \mathscr{M}^{1/2}([\alpha]_3) \sum_{[\beta]_2}\sum_{[\beta']_1,[\beta]_1} \mathscr{M}^{-1/2}([\beta]_2) D^{[\beta]_2}_{(\beta')_1(\beta)_1}(u_2) \\ & \times \left\langle \begin{pmatrix}\alpha'_{12} & \alpha'_{22}\\ & \alpha'_{11}\end{pmatrix}\right| \left\langle \alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22} \;\; 0\right\rangle \left|\begin{pmatrix}\beta_{12} & \beta_{22}\\ & \beta'_{11}\end{pmatrix}\right\rangle \prod_{i=1}^{2} \frac{(u_3^i)^{q_i}}{[(q_i)!]^{1/2}} \\ & \times \left\langle \begin{pmatrix}\alpha_{12} & \alpha_{22}\\ & \alpha_{11}\end{pmatrix}\right| \left\langle \alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22} \;\; 0\right\rangle \left|\begin{pmatrix}\beta_{12} & \beta_{22}\\ & \beta_{11}\end{pmatrix}\right\rangle \prod_{i=1}^{2} \frac{(u_i^3)^{p_i}}{[(p_i)!]^{1/2}} \\ & \times \left\langle \begin{pmatrix}\alpha_{13} & \alpha_{23} & \alpha_{33}\\ & \alpha_{12} & \alpha_{22}\end{pmatrix}\right| \left[\sum_{i=1}^{3}\alpha_{i3}-\alpha'_{12}-\alpha'_{22} \;\; 0 \;\; 0\right] \left|\begin{pmatrix}\alpha'_{12} & \alpha'_{22} & 0\\ & \beta_{12} & \beta_{22}\end{pmatrix}\right\rangle \frac{(u_3^3)^{p_3}}{[(p_3)!]^{1/2}}. \end{aligned} \tag{A14}$$

In order to take the asymptotic limit we insert the arguments prescribed by Eq. (A7) and the states of the form (A13) into Eq. (A14). Our assumption about the exponents of factors of the form (A10) remaining finite then tells us to replace $[\beta]_2$ as follows:

$$\left|\begin{pmatrix}\beta_{12} & & \beta_{22}\\ & \beta_{11} & \end{pmatrix}\right\rangle \to \left|\begin{pmatrix}Rp_{44}+\beta_{12} & & Rp_{44}+\beta_{22}\\ & Rp_{44}+\beta_{11} & \end{pmatrix}\right\rangle. \tag{A15}$$

From now on the quantities α_{ij}, β_{ij}, and the corresponding primed quantities will be those indicated on the right side of Eqs. (A13) and (A15).

The asymptotic limit $R\to\infty$ can now be taken. The process is tedious but straightforward, since the asymptotic forms of all the matrix elements of totally symmetric tensor operators that occur in Eqs. (A1) and (A14) can be evaluated by means of Stirling's approximation alone from Eq. (A4). Some of the gamma functions in the normalization and denominator factors in Eqs. (A1) and (A14) have arguments which become infinite in the limit $R\to\infty$, but these are canceled by factors which occur within the matrix elements of tensor operators. The final result contains only finite powers of R.

We note that the D function on the right of Eq. (A14) becomes

$$[\det U]^{Rp_{44}+\beta_{22}} D^{(1/2)(\beta_{12}-\beta_{22})}_{\beta'_{11}(1/2)(\beta_{12}+\beta_{22})\;\beta_{11}-(1/2)(\beta_{12}+\beta_{22})}(U), \tag{A16}$$

where the Wigner function is given by

$$D^{\ell}_{m'm}(u) = \sum_n \frac{[(\ell+m')!(\ell-m')!(\ell+m)!(\ell-m)!]^{1/2}}{n!(\ell-m'-n)!(\ell-m-n)!(m'+m+n)!}(u_1^1)^{m'+m+n}(u_1^2)^{\ell-m-n}(u_2^1)^{\ell-m'-n}(u_2^2)^n. \tag{A17}$$

The factor $(\det U)^{Rp_{44}}$ in Eq. (A17) is canceled by its reciprocal, which occurs in Eq. (A9).

In the final result we find factors of the form

$$\frac{R^{-m'_{23}+\alpha_{23}+m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}}}{[(-\alpha_{23}+m'_{23}-m'_{12}-m'_{22}+\alpha'_{12}+\alpha'_{22})!]^{1/2}}, \frac{R^{-m_{23}+\alpha_{23}+m_{12}+m_{22}-\alpha_{12}-\alpha_{22}}}{[(-\alpha_{23}+m_{23}-m_{12}-m_{22}+\alpha_{12}+\alpha_{22})!]^{1/2}}, \tag{A18}$$

which vanish except when

$$\alpha_{23} = m'_{23} - m'_{12} - m'_{22} + \alpha'_{12} + \alpha'_{22} = m_{23} - m_{12} - m_{22} + \alpha_{12} + \alpha_{22}. \tag{A19}$$

With this restriction on the summations in the next equation, we can express our final result as

$$\sum_{(\alpha')_2}\sum_{(\alpha)_2}\sum_{(\beta)_2} \frac{1}{[(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})!(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22})!(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})!(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})!]^{1/2}}$$

$$\times C^{(1/2)(\alpha_{12}-\alpha_{22}) \quad (1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}) \quad (1/2)(\beta_{12}-\beta_{22})}_{-\alpha_{23}+(1/2)(\alpha_{12}+\alpha_{22}) \quad -(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}) \quad -\alpha_{23}+(1/2)(\beta_{12}+\beta_{22})}$$

$$\times C^{(1/2)(m_{12}-m_{22}) \quad (1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}) \quad (1/2)(\alpha_{12}-\alpha_{22})}_{-m_{23}+(1/2)(m_{12}+m_{22}) \quad (1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}) \quad -\alpha_{23}+(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times C^{(1/2)(\alpha'_{12}-\alpha'_{22}) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \quad (1/2)(\beta_{12}-\beta_{22})}_{-\alpha_{23}+(1/2)(\alpha'_{12}+\alpha'_{22}) \quad -(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \quad -\alpha_{23}+(1/2)(\beta_{12}+\beta_{22})}$$

$$\times C^{(1/2)(m'_{12}-m'_{22}) \quad (1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}) \quad (1/2)(\alpha'_{12}-\alpha'_{22})}_{-m'_{23}+(1/2)(m'_{12}+m'_{22}) \quad (1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}) \quad -\alpha_{23}+(1/2)(\alpha'_{12}+\alpha'_{22})}$$

$$\times D^{[\beta]_2}_{(\beta')_1(\beta)_1}(U) C^{(1/2)(\beta_{12}-\beta_{22}) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \quad (1/2)(\alpha'_{12}-\alpha'_{22})}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22}) \quad \alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22})}$$

$$\times C^{(1/2)(\alpha'_{12}-\alpha'_{22}) \quad (1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}) \quad (1/2)(m'_{12}-m'_{22})}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}) \quad m'_{11}-\alpha'_{11}-(1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}) \quad m'_{11}-(1/2)(m'_{12}+m'_{22})}$$

$$\times \frac{(\mathscr{P}_1\zeta_1^3)^{\alpha'_{11}-\beta'_{11}}(\mathscr{P}_1\zeta_2^3)^{\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11}}(\mathscr{P}_2\zeta_1^4)^{m'_{11}-\alpha'_{11}}(\mathscr{P}_2\zeta_2^4)^{m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11}}}{[(\alpha'_{11}-\beta'_{11})!(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11})!(m'_{11}-\alpha'_{11})!(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11})!]^{1/2}}$$

$$\times C^{(1/2)(\beta_{12}-\beta_{22}) \quad (1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}) \quad (1/2)(\alpha_{12}-\alpha_{22})}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22}) \quad \alpha_{11}-\beta_{11}-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}) \quad \alpha_{11}-(1/2)(\alpha_{12}-\alpha_{22})}$$

$$\times C^{(1/2)(\alpha_{12}-\alpha_{22}) \quad (1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}) \quad (1/2)(m_{12}-m_{22})}_{\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22}) \quad m_{11}-\alpha_{11}-(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}) \quad m_{11}-(1/2)(m_{12}+m_{22})}$$

$$\times \frac{(\mathscr{P}_1\zeta_3^1)^{\alpha_{11}-\beta_{11}}(\mathscr{P}_1\zeta_3^2)^{\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11}}(\mathscr{P}_2\zeta_4^1)^{m_{11}-\alpha_{11}}(\mathscr{P}_2\zeta_4^2)^{m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11}}}{[(\alpha_{11}-\beta_{11})!(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11})!(m_{11}-\alpha_{11})!(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11})!]^{1/2}}$$

$$\times \left\{ e^{i(p_{13}-p_{33})h_3^3} e^{i(p_{14}+p_{24}-p_{34}-p_{44}-p_{13}+p_{33})h_4^4} \left(-\frac{h_3^4}{h_4^3}\right)^{(1/2)(m'_{13}+m'_{33}-m_{13}-m_{33})} \right.$$

$$\left. \times J_{m'_{13}-m_{13}}(2\mathscr{P}_{13}\sqrt{h_4^3 h_3^4})J_{m'_{33}-m_{33}}(2\mathscr{P}_{33}\sqrt{h_4^3 h_3^4}) \right\}. \tag{A20}$$

The sums are taken over all positive and negative integer values of α_{ij}, α'_{ij}, β_{ij} such that the betweenness conditions for the Gel'fand labels are obeyed. The coefficients $C^{\dots}_{\dots}$ are just the SU(2) Wigner coefficients. The quantities $\mathscr{P}_{13}$, $\mathscr{P}_{33}$, $\mathscr{P}_1$, and $\mathscr{P}_2$ are given in Eq. (3.2) and (3.5) above. The factor in curly brackets is just the matrix element of the (reducible) representation of $\mathbb{H}^{2\times 2}$ which we obtain from the contraction. Discarding it, we obtain the matrix element of a finite transformation in an irreducible representation of U(2)Ⓢ$\mathbb{H}^{2\times 2}$. In Appendix B we shall sketch the method of proof for this assertion.

We now wish to simplify the expression (A20). We set

$$\bar{\zeta}\mathscr{P} = \begin{pmatrix} \mathscr{P}_1\zeta_1^3 & \mathscr{P}_2\zeta_1^4 \\ \mathscr{P}_1\zeta_2^3 & \mathscr{P}_2\zeta_2^4 \end{pmatrix} = Z\Pi^\dagger, \tag{A21a}$$

$$\mathscr{P}\underline{\zeta} = \begin{pmatrix} \mathscr{P}_1\zeta_3^1 & \mathscr{P}_1\zeta_3^2 \\ \mathscr{P}_2\zeta_4^1 & \mathscr{P}_2\zeta_4^2 \end{pmatrix} = -\Pi Z^\dagger U. \tag{A21b}$$

When these matrices are nonsingular we use the identity

$$\sum_{\alpha_{11}} C^{(1/2)(\beta_{12}-\beta_{22}) \quad (1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}) \quad (1/2)(\alpha_{12}-\alpha_{22})}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22}) \quad \alpha_{11}-\beta_{11}-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}) \quad \alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times C^{(1/2)(\alpha_{12}-\alpha_{22}) \quad (1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}) \quad (1/2)(m_{12}-m_{22})}_{\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22}) \quad m_{11}-\alpha_{11}-(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}) \quad m_{11}-(1/2)(m_{12}+m_{22})}$$

$$\times\frac{(\mathscr{P}_1\zeta^1_3)^{\alpha_{11}-\beta_{11}}(\mathscr{P}_1\zeta^2_3)^{\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11}}(\mathscr{P}_2\zeta^1_4)^{m_{11}-\alpha_{11}}(\mathscr{P}_2\zeta^2_4)^{m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11}}}{[(\alpha_{11}-\beta_{11})!(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11})!(m_{11}-\alpha_{11})!(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11})!]^{1/2}}$$

$$=\sum_{\Lambda,\rho}(-1)^{m_{12}+\beta_{12}+\beta_{22}+\beta_{11}}\left[\frac{(m_{12}-m_{22}+1)(\alpha_{12}-\alpha_{22}+1)(2\Lambda+1)}{[(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})-\Lambda]![(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})+\Lambda+1]!}\right]^{1/2}$$

$$\times\begin{Bmatrix}(1/2)(\beta_{12}-\beta_{22}) & (1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}) & (1/2)(\alpha_{12}-\alpha_{22})\\ (1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}) & (1/2)(m_{12}-m_{22}) & \Lambda\end{Bmatrix}$$

$$\times D^{(1/2)(\beta_{12}-\beta_{22})}_{(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})-\rho\ \ -\beta_{11}+(1/2)(\beta_{12}+\beta_{22})}(\mathscr{P}\underline{\zeta})$$

$$\times D^{(1/2)(m_{12}-m_{22})}_{(1/2)(\alpha_{12}+\alpha_{22}-m_{12}-m_{22})-\rho\ \ m_{11}-(1/2)(m_{12}+m_{22})}(\mathscr{P}\underline{\zeta})(\det\mathscr{P}\underline{\zeta})^{m_{22}-\beta_{12}}$$

$$\times C^{(1/2)(m_{12}-m_{22})\ \ (1/2)(\beta_{12}-\beta_{22})\ \ \Lambda}_{(1/2)(\alpha_{12}+\alpha_{22}-m_{12}-m_{22})+\rho\ \ (1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})-\rho\ \ \alpha_{12}+\alpha_{22}-(1/2)(m_{12}+m_{22}+\beta_{12}+\beta_{22})} \tag{A22}$$

We now perform the sum over $\frac{1}{2}(\alpha_{12}-\alpha_{22})$, eliminating the Racah coefficient. When we have done so we find that the sum over Λ can be expressed in terms of the inverse $U(2)$ gamma matrix discussed in Ref. 2:

$$\sum_{\Lambda}\frac{1}{[(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})-\Lambda]![(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})+\Lambda+1]!}$$

$$\times C^{(1/2)(m_{12}-m_{22})\ \ (1/2)(\beta_{12}-\beta_{22})\ \ \Lambda}_{-m_{23}+(1/2)(m_{12}+m_{22})\ \ \alpha_{23}-(1/2)(\beta_{12}+\beta_{22})\ \ \alpha_{12}+\alpha_{22}-(1/2)(m_{12}+m_{22}+\beta_{12}+\beta_{22})}$$

$$\times C^{(1/2)(m_{12}-m_{22})\ \ (1/2)(\beta_{12}-\beta_{22})\ \ \Lambda}_{(1/2)(\alpha_{12}+\alpha_{22}-m_{12}-m_{22})+\rho\ \ (1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})-\rho\ \ \alpha_{12}+\alpha_{22}-(1/2)(m_{12}+m_{22}+\beta_{12}+\beta_{22})}$$

$$\equiv N^{-1}_{(1/2)(\alpha_{12}+\alpha_{22}-m_{12}-m_{22})+\rho,\alpha_{23}-(1/2)(\beta_{12}+\beta_{22});-m_{23}+(1/2)(m_{12}+m_{22}),(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})-\rho}[m_{22}-\beta_{12}$$

$$+2|(1/2)(m_{12}-m_{22}),(1/2)(\beta_{12}-\beta_{22})] \tag{A23}$$

This inverse gamma matrix then appears in Eq. (4.2) above (with a slightly different definition of the summation index ρ).

In the case that the matrices (A21) are singular we cannot use the identity (A22). Instead we have

$$C^{(1/2)(\beta_{12}-\beta_{22})\ \ (1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})\ \ (1/2)(\alpha_{12}-\alpha_{22})}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22})\ \ \alpha_{11}-\beta_{11}-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})\ \ \alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times C^{(1/2)(\alpha_{12}-\alpha_{22})\ \ (1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})\ \ (1/2)(m_{12}-m_{22})}_{\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})\ \ m_{11}-\alpha_{11}-(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})\ \ m_{11}-(1/2)(m_{12}+m_{22})}$$

$$\times\frac{(\mathscr{P}_1\zeta^1_3)^{\alpha_{11}-\beta_{11}}(\mathscr{P}_1\zeta^2_3)^{\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11}}(\mathscr{P}_2\zeta^1_4)^{m_{11}-\alpha_{11}}(\mathscr{P}_2\zeta^2_4)^{m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11}}}{[(\alpha_{11}-\beta_{11})!(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11})!(m_{11}-\alpha_{11})!(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11})!]^{1/2}}$$

$$=(-1)^{m_{12}-m_{22}}\left[\frac{(\alpha_{12}-\alpha_{22}+1)}{(m_{12}+m_{22}-\beta_{12}-\beta_{22})!}\right]^{1/2}C^{(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})\ \ (1/2)(\beta_{12}-\beta_{22})\ \ (1/2)(m_{12}-m_{22})}_{m_{11}-\beta_{11}-(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})\ \ \beta_{11}-(1/2)(\beta_{12}+\beta_{22})\ \ m_{11}-(1/2)(m_{12}+m_{22})}$$

$$\times\begin{Bmatrix}(1/2)(\beta_{12}-\beta_{22}) & (1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}) & (1/2)(\alpha_{12}-\alpha_{22})\\ (1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}) & (1/2)(m_{12}-m_{22}) & (1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})\end{Bmatrix}$$

$$\times D^{(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})}_{\alpha_{12}+\alpha_{22}-(1/2)(m_{12}+m_{22}+\beta_{12}+\beta_{22})\ \ m_{11}-\beta_{11}-(1/2)(m_{12}+m_{22}-\beta_{12}-\beta_{22})}(\mathscr{P}\underline{\zeta}). \tag{A24}$$

We note that since its argument is a singular matrix the Wigner D function on the right of Eq. (A24) can be written as a monomial. For singular u the D function becomes

$$D^{\ell}_{m'm}(u)=(2l)!\,\frac{(u^1_1)^{m'+m}(u^2_1)^{\ell-m}(u^1_2)^{l-m'}}{[(\ell+m')!(\ell-m')!(l+m)!(\ell-m)!]^{1/2}} \tag{A25}$$

in place of Eq. (A17). When the sum over $\frac{1}{2}(\alpha_{12}-\alpha_{22})$ is performed the recoupling process which leads to the expressions (4.10) and (4.13) is straightforward.

APPENDIX B: THE PROPERTIES OF THE MATRIX ELEMENTS OF IRREDUCIBLE REPRESENTATIONS OF $U(2)\textcircled{\small{s}}\mathbb{C}^{2\times 2}$ UNDER DIFFERENTIATION

We have adopted a heuristic procedure in Appendix A for the contraction of the matrices of irreducible representations of U(4) to representations of $[U(2)\textcircled{\small{s}}\mathbb{C}^{2\times 2}]\times\mathbb{H}^{2\times 2}$. The construction was nonrigorous, based only on a plausible assumption as to the behavior of the matrix elements under contraction. It remains, then, to provide a proof that the matrix element which we have obtained as a result, given in Eq. (A20) above, is indeed the desired matrix element of a representation of $[U(2)\textcircled{\small{s}}\mathbb{C}^{2\times 2}]\times\mathbb{H}^{2\times 2}$, and that the representation matrix of $U(2)\textcircled{\small{s}}\mathbb{C}^{2\times 2}$ is irreducible. We shall provide such a proof by demonstrating that the matrix element proposed in Eq. (A20) as that of an irreducible representation of $U(2)\textcircled{\small{s}}\mathbb{C}^{2\times 2}$ does indeed have the correct transformation properties under the differential operators which realize the Lie algebra of the group. Specifically, the matrix element (4.2) must obey the relations

$$P_{iJ}\langle(\Pi;\langle m'\rangle)|\mathscr{O}(U,Z)|(\Pi;\langle m\rangle)\rangle = \sum_{\langle m''\rangle} \overline{\langle(\Pi;\langle m''\rangle)|\, P_{iJ}|(\Pi;\langle m'\rangle)\rangle}\, \langle(\Pi;\langle m''\rangle)|\mathscr{O}(U,Z)|(\Pi;\langle m\rangle)\rangle, \tag{B1a}$$

$$P^{iJ}\langle(\Pi;\langle m'\rangle)|\mathscr{O}(U,Z)|(\Pi;\langle m\rangle)\rangle = \sum_{\langle m''\rangle} \langle\Pi;\langle m''\rangle)|\, P_{iJ}|(\Pi;\langle m\rangle)\rangle\, \langle(\Pi;\langle m'\rangle)|\mathscr{O}(U,Z)|(\Pi;\langle m''\rangle)\rangle, \tag{B1b}$$

which hold, *mutatis mutandis*, for the generators P_{Ij} and P^{Ij}. It is evident that Eqs. (4.2) and (A20) obey the boundary condition at the identity element of the group. The matrix elements of the generators are given in Eqs. (3.3) and (3.11) above, and the realization of the generators as differential operators on the group manifold is given in Eqs. (5.9) and (5.11). We shall not carry out the program of verification in detail but shall merely sketch its procedure in two representative cases. Specifically, we shall apply the operators $P_{42}=\partial/\partial z_2^4$ and $P_{24}=-\partial/\partial\bar{z}_2^4$ to Eq. (A20). To facilitate comparison we shall note the correspondences

$$\begin{pmatrix} \mathscr{P}_1\zeta_1^3 & \mathscr{P}_2\zeta_1^4 \\ \mathscr{P}_1\zeta_2^3 & \mathscr{P}_2\zeta_2^4 \end{pmatrix} = \begin{pmatrix} \mathscr{P}_1 z_1^3+\bar{\omega}\mathscr{P}_3 z_1^4 & \mathscr{P}_2 z_1^4 \\ \mathscr{P}_1 z_2^3+\bar{\omega}\mathscr{P}_3 z_2^4 & \mathscr{P}_2 z_2^4 \end{pmatrix},$$

$$\begin{pmatrix} \mathscr{P}_1\zeta_3^1 & \mathscr{P}_1\zeta_3^2 \\ \mathscr{P}_2\zeta_4^1 & \mathscr{P}_2\zeta_4^2 \end{pmatrix}$$

$$= \begin{pmatrix} -\mathscr{P}_1 u_1^1\bar{z}_1^3-\omega\mathscr{P}_3 u_1^1\bar{z}_1^4-\mathscr{P}_1 u_2^1\bar{z}_2^3-\omega\mathscr{P}_3 u_2^1\bar{z}_2^4 & -\mathscr{P}_1 u_1^2\bar{z}_1^3-\omega\mathscr{P}_3 u_1^2 z_1^4-\mathscr{P}_1 u_2^2\bar{z}_2^3-\omega\mathscr{P}_3 u_2^2\bar{z}_2^4 \\ -\mathscr{P}_2 u_1^1\bar{z}_1^4-\mathscr{P}_2 u_2^1\bar{z}_2^4 & -\mathscr{P}_2 u_1^2\bar{z}_1^4-\mathscr{P}_2 u_2^2\bar{z}_2^4 \end{pmatrix}. \tag{B2}$$

The easier of the two cases is that of P_{42}. Applying it to Eq. (A20), we find

$$\frac{\partial}{\partial z_2^4}\sum_{\alpha'_{11}}\Big\{ C^{(1/2)(\beta_{12}-\beta_{22})\quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})\quad (1/2)(\alpha'_{12}-\alpha'_{22})}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22})\quad \alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})\quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22})}$$

$$\times C^{(1/2)(\alpha'_{12}-\alpha'_{22})\quad (1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\beta'_{22})\quad (1/2)(m'_{12}-m'_{22})}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22})\quad m'_{11}-\alpha'_{11}-(1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22})\quad m'_{11}-(1/2)(m'_{12}+m'_{22})}\Big\}$$

$$\times\frac{(\mathscr{P}_1 z_1^3+\bar{\omega}\mathscr{P}_3 z_1^4)^{\alpha'_{11}-\beta'_{11}}(\mathscr{P}_1 z_2^3+\bar{\omega}\mathscr{P}_3 z_2^4)^{\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11}}(\mathscr{P}_2 z_1^4)^{m'_{11}-\alpha'_{11}}(\mathscr{P}_2 z_2^4)^{m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11}}}{[(\alpha'_{11}-\beta'_{11})!(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11})!(m'_{11}-\alpha'_{11})!(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11})!]^{1/2}}$$

$$=\sum_{\alpha'_{11}}\{C^{\cdots}_{\cdots}\,C^{\cdots}_{\cdots}\}\Bigg[\bar{\omega}\mathscr{P}_3(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11})^{1/2}\left(\frac{\mathscr{P}_2 z_2^4}{(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11})^{1/2}}\right)$$

$$+\mathscr{P}_2(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11})^{1/2}\left(\frac{(\mathscr{P}_1 z_2^3+\bar{\omega}\mathscr{P}_3\bar{z}_2^4)}{(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11})^{1/2}}\right)\Bigg]$$

$$\times\frac{(\mathscr{P}_1 z_1^3+\bar{\omega}\mathscr{P}_3 z_1^4)^{\alpha'_{11}-\beta'_{11}}(\mathscr{P}_1 z_2^3+\bar{\omega}\mathscr{P}_3 z_2^4)^{\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11}}(\mathscr{P}_2 z_1^4)^{m'_{11}-\alpha'_{11}}(\mathscr{P}_2 z_2^4)^{m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11}}}{[(\alpha'_{11}-\beta'_{11})!(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11})!(m'_{11}-\alpha'_{11})!(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11})!]^{1/2}}. \tag{B3}$$

The first of the two terms in square brackets on the right of Eq. (B3) corresponds to the first two terms on the right of Eq. (3.11d), and the second corresponds to the last two terms of Eq. (3.11d). We note from Eq. (B1a) that we expect the complex conjugation of the matrix elements given in Eq. (3.11d); hence, the complex conjugate $\bar{\omega}\mathscr{P}_3$ of the eigenvalue $\omega\mathscr{P}_3$ appears in the first term in square brackets on the right of Eq. (B3), whereas $\omega\mathscr{P}_3$ appears in the first two terms of Eq. (3.11d). We shall now limit our attention to the first term in square brackets in Eq. (B3) and sketch the derivation of the first two terms of Eq. (3.11d) from it. A completely analogous procedure holds for the second term in Eq. (B3). The first term on the right of Eq. (B3) can be written as

$$-\bar{\omega}\mathscr{P}_3\sum_{\alpha'_{11}}\{C^{\cdots}_{\cdots}\,C^{\cdots}_{\cdots}\}$$

$$\times(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1)^{1/2}C^{(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})\quad (1/2)\quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-1)}_{\alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})\quad (1/2)\quad \alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-1)}$$

$$\times\frac{(\mathscr{P}_1 z_1^3+\bar{\omega}\mathscr{P}_3 z_1^4)^{\alpha'_{11}-\beta'_{11}}(\mathscr{P}_1 z_2^3+\bar{\omega}\mathscr{P}_3 z_2^4)^{\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11}}(\mathscr{P}_2 z_1^4)^{m'_{11}-\alpha'_{11}}(\mathscr{P}_2 z_2^4)^{m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11}}}{[(\alpha'_{11}-\beta'_{11})!(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11})!(m'_{11}-\alpha'_{11})!(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11})!]^{1/2}}. \tag{B4}$$

We shall now relabel the index of summation α'_{12}, replacing α'_{12} with $\alpha'_{12}+1$. Making this substitution in Eq. (B4) and including the other relevant structures of Eq. (A20), we find ourselves dealing with the expression

$$-\bar{\omega}\mathscr{P}_3\sum_{[\alpha']_2}\sum_{(\alpha')_1}\left(\frac{(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+2)}{(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-1)!(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1)!}\right)^{1/2}$$

$$\times C^{(1/2)(\beta_{12}-\beta_{22}) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad (1/2)(\alpha'_{12}-\alpha'_{22}+1)}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22}) \quad \alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha_{22})}$$
$$\times C^{(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad (1/2) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})}_{\alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad (1/2) \quad \alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})}$$
$$\times C^{(1/2)(\alpha'_{12}-\alpha'_{22}+1) \quad (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \quad (1/2)(m'_{12}-m'_{22})}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1) \quad m'_{11}-\alpha'_{11}-(1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \quad m'_{11}-(1/2)(m'_{12}+m_{22})}$$
$$\times C^{(1/2)(m'_{12}-m'_{22}) \quad (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \quad (1/2)(\alpha'_{12}-\alpha'_{22}+1)}_{-m'_{23}+(1/2)(m'_{12}+m'_{22}) \quad (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}+\alpha'_{22}) \quad -m'_{23}+m'_{12}+m'_{22}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)}$$
$$\times C^{(1/2)(\alpha'_{12}-\alpha'_{22}+1) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad (1/2)(\beta_{12}-\beta_{22})}_{-m'_{23}+m'_{12}+m'_{22}-(1/2)(\alpha'_{12}+\alpha'_{22}+1) \quad -(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad -m'_{23}+m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-1+(1/2)(\beta_{12}+\beta_{22})}$$
$$\times \frac{(\mathscr{P}_1 z_1^3+\overline{\omega}\mathscr{P}_3 z_1^4)^{\alpha'_{11}-\beta'_{11}}(\mathscr{P}_1 z_2^3+\overline{\omega}\mathscr{P}_3 z_2^4)^{\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11}}(\mathscr{P}_2 z_1^4)^{m'_{11}-\alpha'_{11}}(\mathscr{P}_2 z_2^4)^{m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11}}}{[(\alpha'_{11}-\beta'_{11})!(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}-\alpha'_{11}+\beta'_{11})!(m'_{11}-\alpha'_{11})!(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}-m'_{11}+\alpha'_{11})!]^{1/2}}. \tag{B5}$$

The normalized arguments, the factor in curly brackets in Eq. (B5), now have the correct form prescribed by Eq. (3.11d): The operator P_{42} shifts the sum $m'_{12}+m'_{22}$ to $m'_{12}+m'_{22}-1$. It remains only to perform recoupling operations on the Wigner coefficients in Eq. (B5) in order to obtain an expression of the form (B1a) and verify the correspondence with Eq (3.11d). We use the following identities:

$$C^{(1/2)(\beta_{12}-\beta_{22}) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad (1/2)(\alpha'_{12}-\alpha'_{22}+1)}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22}) \quad \alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)}$$
$$\times C^{(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad (1/2) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})}_{\alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad (1/2) \quad \alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})}$$
$$= \sum_{A'}(-1)^{\alpha'_{12}-\beta_{22}}[(2A'+1)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1)]^{1/2}$$
$$\times\begin{Bmatrix} (1/2)(\alpha'_{12}-\alpha'_{22}+1) & (1/2)(\beta_{12}-\beta_{22}) & (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \\ (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) & (1/2) & A' \end{Bmatrix}$$
$$\times C^{(1/2)(\beta_{12}-\beta_{22}) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \quad A'}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22}) \quad \alpha'_{11}-\beta'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22})}$$
$$\times C^{A' \quad (1/2) \quad (1/2)(\alpha'_{12}-\alpha'_{22}+1)}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}) \quad -(1/2) \quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)}, \tag{B6}$$

and then

$$C^{A' \quad (1/2) \quad (1/2)(\alpha'_{12}-\alpha'_{22}+1)}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}) \quad -(1/2) \quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)}$$
$$\times C^{(1/2)(\alpha'_{12}-\alpha'_{22}+1) \quad (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \quad (1/2)(m'_{12}-m'_{22})}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1) \quad m'_{11}-\alpha'_{11}-(1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}+\alpha'_{22}) \quad m'_{11}-(1/2)(m'_{12}+m'_{22})}$$
$$= \sum_{M'}(-1)^{M'-\alpha'_{22}+(1/2)(m'_{12}+m'_{22}+1)}[(2M'+1)(\alpha'_{12}-\alpha'_{22}+1)]^{1/2}$$
$$\times\begin{Bmatrix} (1/2) & A' & (1/2)(\alpha'_{12}-\alpha'_{22}+1) \\ (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) & (1/2)(m'_{12}-m'_{22}) & M' \end{Bmatrix}$$
$$\times C^{A' \quad (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \quad M'}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}) \quad m'_{11}-\alpha'_{11}-(1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \quad m'_{11}-(1/2)(m'_{12}+m'_{22}-1)}$$
$$\times C^{M' \quad (1/2) \quad (1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22}-1) \quad -(1/2) \quad m'_{11}-(1/2)(m'_{12}+m'_{22})}. \tag{B7}$$

We now wish the eliminate the Racah coefficients in Eqs. (B6) and (B7) and perform the summation over $\frac{1}{2}(\alpha'_{12}-\alpha'_{22}+1)$ by means of the completeness relation for the SU(2) Wigner coefficient. Thus, the sum over $\frac{1}{2}(\alpha'_{12}-\alpha'_{22}+1)$ becomes

$$\sum_{(1/2)(\alpha'_{12}-\alpha'_{22}+1)}(-1)^{M'+\alpha'_{12}+\alpha'_{22}-\beta_{22}+(1/2)(m'_{12}+m'_{22}+1)}$$
$$\times[(2A'+1)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1)(2M'+1)(\alpha'_{12}-\alpha'_{22}+2)]^{1/2}$$
$$\times C^{(1/2)(m'_{12}-m'_{22}) \quad (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \quad (1/2)(\alpha'_{12}-\alpha'_{22}+1)}_{-m'_{23}+(1/2)(m'_{12}+m'_{22}) \quad (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}+\alpha'_{22}) \quad -m'_{23}+m'_{12}+m'_{22}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)}$$
$$\times\begin{Bmatrix} (1/2) & A' & (1/2)(\alpha'_{12}-\alpha'_{22}+1) \\ (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) & (1/2)(m'_{12}-m'_{22}) & M' \end{Bmatrix}$$
$$\times C^{(1/2)(\alpha'_{12}-\alpha'_{22}+1) \quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad (1/2)(\beta_{12}-\beta_{22})}_{-m'_{23}+m'_{12}+m'_{22}-(1/2)(\alpha'_{12}+\alpha'_{22}+1) \quad -(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \quad -m'_{23}+m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}-1+(1/2)(\beta_{12}+\beta_{22})}$$

$$\times\begin{Bmatrix} (1/2)(\alpha'_{12}-\alpha'_{22}+1) & (1/2)(\beta_{12}-\beta_{22}) & (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \\ (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) & (1/2) & A' \end{Bmatrix}$$

$$= \sum_{(1/2)(\alpha'_{12}-\alpha'_{22}+1)} (-1)^{M'+\beta_{12}+(1/2)(m'_{12}+m'_{22}-1)}\left(\frac{(2M'+1)}{(m'_{12}-m'_{22}+1)}\right)^{1/2}$$

$$\times\sum_{\rho} C^{(1/2)\;\; M' \;\; (1/2)(m'_{12}-m'_{22})}_{\rho\;\; -m'_{23}-\rho+(1/2)(m'_{12}+m'_{22}) \;\; -m'_{23}+(1/2)(m'_{12}+m'_{22})}$$

$$\times C^{A' \;\; (1/2) \;\; (1/2)(\alpha'_{12}-\alpha'_{22}+1)}_{-m'_{23}+m'_{12}+m'_{22}-\rho-(1/2)(\alpha'_{12}+\alpha'_{22}+1) \;\; \rho \;\; -m'_{23}+m'_{12}+m'_{22}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)}$$

$$\times C^{(1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \;\; M' \;\; A'}_{(1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \;\; -m'_{23}-\rho+(1/2)(m'_{12}+m'_{22}) \;\; -m'_{23}+m'_{12}+m'_{22}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)-\rho}$$

$$\times\sum_{\sigma} C^{(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \;\; (1/2) \;\; (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})}_{-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \;\; \sigma \;\; \sigma-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1)}$$

$$\times C^{A' \;\; (1/2) \;\; (1/2)(\alpha'_{12}-\alpha'_{22}+1)}_{-m'_{23}+m'_{12}+m'_{22}-\sigma-(1/2)(\alpha'_{12}+\alpha'_{22}+1) \;\; \sigma \;\; -m'_{23}+m'_{12}+m'_{22}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)}$$

$$\times C^{(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \;\; A' \;\; (1/2)(\beta_{12}-\beta_{22})}_{\sigma-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \;\; -m'_{23}+m'_{12}+m'_{22}-\sigma-(1/2)(\alpha'_{12}+\alpha'_{22}+1) \;\; -m'_{23}+m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}+(1/2)(\beta_{12}+\beta_{22})-1}$$

$$= \left[\frac{(2M'+1)}{(m'_{12}-m'_{22}+1)}\right]^{1/2} C^{(1/2) \;\; M' \;\; (1/2)(m'_{12}-m'_{22})}_{(1/2) \;\; -m'_{23}+(1/2)(m'_{12}+m'_{22}-1) \;\; -m'_{23}+(1/2)(m'_{12}+m'_{22})}$$

$$\times C^{(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \;\; (1/2) \;\; (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})}_{-(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}+1) \;\; (1/2) \;\; -(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})}$$

$$\times C^{M' \;\; (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \;\; A'}_{-m'_{23}+(1/2)(m'_{12}+m'_{22}-1) \;\; (1/2)(m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}) \;\; -m'_{23}+m'_{12}+m'_{22}-1-(1/2)(\alpha'_{12}+\alpha'_{22})}$$

$$\times C^{A' \;\; (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \;\; (1/2)(\beta_{12}-\beta_{22})}_{-m'_{23}+m'_{12}+m'_{22}-1-(1/2)(\alpha'_{12}+\alpha'_{22}) \;\; -(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) \;\; -m'_{23}+m'_{12}+m'_{22}-1-\alpha'_{12}-\alpha'_{22}+(1/2)(\beta_{12}+\beta_{22})}. \qquad \text{(B8)}$$

We are now in a position to collate the results of Eqs. (B4)–(B8) compare the final expression with Eq. (B1a). We have already established that $m''_{12}+m''_{22}=m'_{12}+m'_{22}-1$. We now find that $m''_{23}=m'_{23}$ and that $M'=\frac{1}{2}(m''_{12}-m''_{22})$ can take on the two values $\frac{1}{2}(m'_{12}-m'_{22}-1)$ and $\frac{1}{2}(m'_{12}-m'_{22}+1)$. The former corresponds to the first term on the right of Eq. (3.11d) and the latter to the second term. We set $A'=\frac{1}{2}(\alpha'_{12}-\alpha'_{22})$, i.e., we allow A' to serve as the corresponding index of summation in Eq. (A20). The final result is

$$\overline{\left\langle\left(\Pi;\begin{pmatrix} & m'_{23} & \\ m'_{12}-1 & & m'_{22} \\ & m'_{11} & \end{pmatrix}\right)\right|} P_{42} \left|\left(\Pi;\begin{pmatrix} & m'_{23} & \\ m'_{12} & & m'_{22} \\ & m'_{11} & \end{pmatrix}\right)\right\rangle$$

$$= \overline{\omega}\mathscr{P}_3 C^{(1/2)(m'_{12}-m'_{22}) \;\; (1/2) \;\; (1/2)(m'_{12}-m'_{22}-1)}_{-m'_{23}+(1/2)(m'_{12}+m'_{22}) \;\; -(1/2) \;\; -m'_{23}+(1/2)(m'_{12}+m'_{22}-1)} C^{(1/2)(m'_{12}-m'_{22}-1) \;\; (1/2) \;\; (1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22}-1) \;\; -(1/2) \;\; m'_{11}-(1/2)(m'_{12}+m'_{22})},$$

$$\overline{\left\langle\left(\Pi;\begin{pmatrix} & m'_{23} & \\ m'_{12} & & m'_{22}-1 \\ & m'_{11} & \end{pmatrix}\right)\right|} P_{42} \left|\left(\Pi;\begin{pmatrix} & m'_{23} & \\ m'_{12} & & m'_{22} \\ & m'_{11} & \end{pmatrix}\right)\right\rangle$$

$$= \overline{\omega}\mathscr{P}_3 C^{(1/2)((m'_{12}-m'_{22}) \;\; (1/2) \;\; (1/2)(m'_{12}-m'_{22}+1)}_{-m'_{23}+(1/2)(m'_{12}+m'_{22}) \;\; -(1/2) \;\; -m'_{23}+(1/2)(m'_{12}+m'_{22}-1)} C^{(1/2)(m'_{12}-m'_{22}+1) \;\; (1/2) \;\; (1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22}-1) \;\; -(1/2) \;\; m'_{11}-(1/2)(m'_{12}+m'_{22})},$$

(B9)

which correspond exactly to the first two terms of Eq. (3.11d).

We have now consider the operation of $P_{24}=-(\partial/\partial\overline{z}_2^4)$ on the representation matrix (A20). In our treatment of the case of P_{42} above we have given a paradigm of the recoupling process needed for the verification of Eq. (B1) for all the generators P_{Ij} and P^{Ij}. The generators P_{iJ} and P^{iJ} require a distinct and, unfortunately, more complicated paradigm. From Eqs. (A20) and (B2) we have immediately

$$-\frac{\partial}{\partial \bar{z}_2^4}(\mathscr{P}_1\zeta_3^1)^{\alpha_{11}-\beta_{11}}(\mathscr{P}_1\zeta_3^2)^{\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11}}(\mathscr{P}_2\zeta_4^1)^{m_{11}-\alpha_{11}}(\mathscr{P}_2\zeta_4^2)^{m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11}}$$
$$\overline{[(\alpha_{11}-\beta_{11})!(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11})!(m_{11}-\alpha_{11})!(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11})!]^{1/2}}$$

$$\begin{aligned}
&= [\omega\mathscr{P}_3 u_2^1(\alpha_{11}-\beta_{11})^{1/2} \\
&\quad\times\left(\frac{(\mathscr{P}_1\zeta_3^2)(\mathscr{P}_2\zeta_4^1)(\mathscr{P}_2\zeta_4^2)}{[(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11})(m_{11}-\alpha_{11})(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11})]^{1/2}}\right) \\
&\quad+\omega\mathscr{P}_3 u_2^2(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11})^{1/2} \\
&\quad\times\left(\frac{(\mathscr{P}_1\zeta_3^1)(\mathscr{P}_2\zeta_4^1)(\mathscr{P}_2\zeta_4^2)}{[(\alpha_{11}-\beta_{11})(m_{11}-\alpha_{11})(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11})]^{1/2}}\right)+\mathscr{P}_2 u_2^1(m_{11}-\alpha_{11})^{1/2} \\
&\quad\times\left(\frac{(\mathscr{P}_1\zeta_3^1)(\mathscr{P}_1\zeta_3^2)(\mathscr{P}_2\zeta_4^2)}{[(\alpha_{11}-\beta_{11})(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11})(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11})]^{1/2}}\right) \\
&\quad+\mathscr{P}_2 u_2^2(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11})^{1/2} \\
&\quad\times\left(\frac{(\mathscr{P}_1\zeta_3^1)(\mathscr{P}_1\zeta_3^2)(\mathscr{P}_2\zeta_4^1)}{[(\alpha_{11}-\beta_{11})(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11})(m_{11}-\alpha_{11})]^{1/2}}\right)] \\
&\quad\times(\mathscr{P}_1\zeta_3^1)^{\alpha_{11}-\beta_{11}-1}(\mathscr{P}_1\zeta_3^2)^{\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11}-1}(\mathscr{P}_2\zeta_4^1)^{m_{11}-\alpha_{11}-1}(\mathscr{P}_2\zeta_4^2)^{m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11}-1} \\
&\quad\times[(\alpha_{11}-\beta_{11}-1)!(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-\alpha_{11}+\beta_{11}-1)! \\
&\quad\times(m_{11}-\alpha_{11}-1)!(m_{12}+m_{22}-\alpha_{12}-\alpha_{22}-m_{11}+\alpha_{11}-1)!]^{-1/2}
\end{aligned} \tag{B10}$$

The first two terms in square brackets on the right of Eq. (B10) correspond to the first two terms on the right of Eq. (3.11b) and the second two terms of Eq. (B10) to the second two terms of Eq. (3.11b). Again, we shall limit our attention to the first two terms of Eq. (B10), and we shall carry out the recoupling process in explicit detail only for the first term and merely present the final result for the second. It is our goal to express the operation on the left of Eq. (B1a) in terms of the expression on the right. The matrix element of the finite transformation on the right of Eq. (B1a) will be called the "reconstituted matrix element." As its structures emerge from the recoupling process they will be denoted by insertion in boxes for easy comparison with Eq. (A20). The matrix element of the finite transformation on the left of Eq. (B1a) will be called the "original matrix element." We shall proceed as before, introducing the structures of the original matrix element as they are needed in the sequence of the recoupling process, then recoupling, and then labeling the structures of the reconstituted matrix element as they emerge. Once the reconstituted matrix element is completely assembled we shall be able to compare its coefficients in the expansion on the right of Eq. (B1a) with those given in Eq. (3.11b).

We now consider the first term on the right of Eq (B10) and make the replacements

$$\begin{aligned}(\alpha_{11}-\beta_{11})^{1/2} &= (\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}+1)^{1/2} \\ &\quad\times C^{(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})}_{-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})+\alpha_{11}-\beta_{11}}\,{}^{(1/2)}_{-(1/2)}\,{}^{(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-1)}_{-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}+1)+\alpha_{11}-\beta_{11}}\end{aligned} \tag{B11}$$

and

$$u_2^1 = D^{(1/2)}_{-(1/2)\,(1/2)}(u). \tag{B12}$$

Hence,

$$\begin{aligned}
&u_2^1(\det u)^{\beta_{22}}D^{(1/2)(\beta_{12}-\beta_{22})}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22})\;\beta_{11}-(1/2)(\beta_{12}+\beta_{22})}(u) \\
&\quad=(\det u)^{\beta_{22}}D^{(1/2)}_{-(1/2)\,(1/2)}(u)D^{(1/2)(\beta_{12}-\beta_{22})}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22})\;\beta_{11}-(1/2)(\beta_{12}+\beta_{22})}(u)=\sum_B \boxed{(\det u)^{(1/2)(\beta_{12}+\beta_{22}+1)-B}} \\
&\quad\times C^{(1/2)(\beta_{12}-\beta_{22})}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22})}\,{}^{(1/2)}_{-(1/2)}\,{}^{B}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22}+1)} \\
&\quad\times\boxed{D^{B}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22}+1)\;\beta_{11}-(1/2)(\beta_{12}+\beta_{22}-1)}(u)}\,C^{(1/2)(\beta_{12}-\beta_{22})}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22})}\,{}^{(1/2)}_{(1/2)}\,{}^{B}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22}-1)}.
\end{aligned} \tag{B13}$$

Bringing in a Wigner coefficient from the original matrix element, we then use the identity

$$\begin{aligned}
&C^{(1/2)(\beta_{12}-\beta_{22})}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22})}\,{}^{(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})}_{-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})+\alpha_{11}-\beta_{11}}\,{}^{(1/2)(\alpha_{12}-\alpha_{22})}_{\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})}\;C^{(1/2)(\beta_{12}-\beta_{22})}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22})}\,{}^{(1/2)}_{(1/2)}\,{}^{B}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22}-1)} \\
&\quad\times C^{(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})}_{-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})+\alpha_{11}-\beta_{11}}\,{}^{(1/2)}_{-(1/2)}\,{}^{(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-1)}_{-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}+1)+\alpha_{11}-\beta_{11}}
\end{aligned}$$

$$= \sum_{K,A} [(\alpha_{12} - \alpha_{22} + 1)(2K+1)(2B+1)(\alpha_{12} + \alpha_{22} - \beta_{12} - \beta_{22})]^{1/2}$$

$$\times \begin{Bmatrix} (1/2)(\beta_{12} - \beta_{22}) & (1/2)(\alpha_{12} + \alpha_{22} - \beta_{12} - \beta_{22}) & (1/2)(\alpha_{12} - \alpha_{22}) \\ (1/2) & (1/2) & K \\ B & (1/2)(\alpha_{12} + \alpha_{22} - \beta_{12} - \beta_{22} - 1) & A \end{Bmatrix}$$

$$\times C^{(1/2)\;\;(1/2)\;\;K}_{(1/2)\;\;-(1/2)\;\;0}\, C^{(1/2)(\alpha_{12}-\alpha_{22})\;\;K\;\;A}_{\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})\;\;0\;\;\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times \boxed{C^{B\;\;(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-1)\;\;A}_{\beta_{11}-(1/2)(\beta_{12}+\beta_{22}-1)\;\;-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}+1)+\alpha_{11}-\beta_{11}\;\;\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})}}\,. \qquad \text{(B14)}$$

Taking the second Wigner coefficient on the right of Eq. (B14) and introducing another Wigner coefficient from the original matrix element, we obtain

$$C^{(1/2)(\alpha_{12}-\alpha_{22})\;\;K\;\;A}_{\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})\;\;0\;\;\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times C^{(1/2)(\alpha_{12}-\alpha_{22})\;\;(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})\;\;(1/2)(m_{12}-m_{22})}_{\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})\;\;-(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})+m_{11}-\alpha_{11}\;\;m_{11}-(1/2)(m_{12}+m_{22})}$$

$$= \sum_{M} (-1)^{m_{12}-m_{22}+K} [(2A+1)(2M+1)]^{1/2}$$

$$\times \begin{Bmatrix} K & A & (1/2)(\alpha_{12} - \alpha_{22}) \\ (1/2)(m_{12} + m_{22} - \alpha_{12} - \alpha_{22}) & (1/2)(m_{12} - m_{22}) & M \end{Bmatrix}$$

$$\times (-1)^{A+(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})-M} C^{M\;\;K\;\;(1/2)(m_{12}-m_{22})}_{m_{11}-(1/2)(m_{12}+m_{22})\;\;0\;\;m_{11}-(1/2)(m_{12}+m_{22})}$$

$$\times \boxed{C^{A\;\;(1/2)(m_{12}-m_{22}-\alpha_{12}-\alpha_{22})\;\;M}_{\alpha_{11}-(1/2)(\alpha_{12}+\alpha_{22})\;\;-(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})+m_{11}-\alpha_{11}\;\;m_{11}-(1/2)(m_{12}+m_{22})}}\,. \qquad \text{(B15)}$$

However, then

$$C^{(1/2)(m_{12}-m_{22})\;\;(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})\;\;(1/2)(\alpha_{12}-\alpha_{22})}_{-m_{23}+(1/2)(m_{12}+m_{22})\;\;(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})\;\;-\alpha_{23}+(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times \begin{Bmatrix} K & A & (1/2)(\alpha_{12} - \alpha_{22}) \\ (1/2)(m_{12} + m_{22} - \alpha_{12} - \alpha_{22}) & (1/2)(m_{12} - m_{22}) & M \end{Bmatrix}$$

$$= \frac{(-1)^{M+(1/2)(m_{12}+m_{22}+\alpha_{12}+\alpha_{22})-A}}{[(2A+1)(m_{12}-m_{22}+1)]^{1/2}} \sum_{\kappa} C^{K\;\;M\;\;(1/2)(m_{12}-m_{22})}_{\kappa\;\;-m_{23}-\kappa+(1/2)(m_{12}+m_{22})\;\;-m_{23}+(1/2)(m_{12}+m_{22})}$$

$$\times C^{A\;\;K\;\;(1/2)(\alpha_{12}-\alpha_{22})}_{-\alpha_{23}-\kappa+(1/2)(\alpha_{12}+\alpha_{22})\;\;\kappa\;\;-\alpha_{23}+(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times \boxed{C^{M\;\;(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})\;\;A}_{-m_{23}-\kappa+(1/2)(m_{12}+m_{22})\;\;(1/2)(m_{12}+m_{22}-\alpha_{12}-\alpha_{22})\;\;-\alpha_{23}-\kappa+(1/2)(\alpha_{12}+\alpha_{22})}}\,. \qquad \text{(B16)}$$

We now introduce the last Wigner coefficient of the original matrix element which contains $\frac{1}{2}(\alpha_{12} - \alpha_{22})$ and sum over this parameter, eliminating the (9-j) symbol constructed in Eq. (B14):

$$\sum_{(1/2)(\alpha_{12}-\alpha_{22})} (\alpha_{12} - \alpha_{22} + 1)^{1/2} C^{(1/2)(\alpha_{12}-\alpha_{22})\;\;(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})\;\;(1/2)(\beta_{12}-\beta_{22})}_{-\alpha_{23}+(1/2)(\alpha_{12}+\alpha_{22})\;\;-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22})\;\;-\alpha_{23}+(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times C^{A\;\;K\;\;(1/2)(\alpha_{12}-\alpha_{22})}_{-\alpha_{23}-\kappa+(1/2)(\alpha_{12}+\alpha_{22})\;\;\kappa\;\;-\alpha_{23}+(1/2)(\alpha_{12}+\alpha_{22})}$$

$$\times \begin{Bmatrix} (1/2)(\beta_{12} - \beta_{22}) & (1/2)(\alpha_{12} + \alpha_{22} - \beta_{12} - \beta_{22}) & (1/2)(\alpha_{12} - \alpha_{22}) \\ (1/2) & (1/2) & K \\ B & (1/2)(\alpha_{12} + \alpha_{22} - \beta_{12} - \beta_{22} - 1) & A \end{Bmatrix}$$

$$= \frac{(-1)^{B+K-(1/2)(\beta_{12}-\beta_{22}+1)}}{[(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}+1)(2K+1)(2B+1)]^{1/2}}$$

$$\times C^{(1/2)\;\;(1/2)\;\;K}_{\kappa-(1/2)\;\;(1/2)\;\;\kappa}\, C^{(1/2)\;\;B\;\;(1/2)(\beta_{12}-\beta_{22})}_{\kappa-(1/2)\;\;-\alpha_{23}-\kappa+(1/2)(\beta_{12}+\beta_{22}+1)\;\;-\alpha_{23}+(1/2)(\beta_{12}+\beta_{22})}$$

$$\times \boxed{C^{A\;\;(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-1)\;\;B}_{-\alpha_{23}-\kappa+(1/2)(\alpha_{12}+\alpha_{22})\;\;-(1/2)(\alpha_{12}+\alpha_{22}-\beta_{12}-\beta_{22}-1)\;\;-\alpha_{23}-\kappa+(1/2)(\beta_{12}+\beta_{22}+1)}}\,. \qquad \text{(B17)}$$

The rest is nothing but a straightforward recoupling process which involves only Wigner and Racah coefficients. We introduce a Wigner coefficient from the original matrix element and pick up a hitherto unused Wigner coefficient from Eq. (B13) to form

$$
\begin{aligned}
&C^{(1/2)(\beta_{12}-\beta_{22})\quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})\quad (1/2)(\alpha'_{12}-\alpha'_{22})}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22})\quad -(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})+\alpha'_{11}-\beta_{11}\quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22})} \\
&\quad\times C^{(1/2)(\beta_{12}-\beta_{22})\quad (1/2)\quad B}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22})\quad -(1/2)\quad \beta'_{11}-(1/2)(\beta_{12}+\beta_{22}+1)} \\
&\quad=\sum_{A'}(-1)^{A'+(1/2)(\alpha'_{12}+\alpha'_{22}+\beta_{12}+\beta_{22})-B}[(2B+1)(2A'+1)]^{1/2} \\
&\quad\times\begin{Bmatrix} (1/2) & B & (1/2)(\beta_{12}-\beta_{22}) \\ (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22}) & (1/2)(\alpha'_{12}-\alpha'_{22}) & A' \end{Bmatrix} \\
&\quad\times C^{A'\quad (1/2)\quad (1/2)(\alpha'_{12}-\alpha'_{22})}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)\quad (1/2)\quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22})} \\
&\quad\times\boxed{C^{B\quad (1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})\quad A'}_{\beta'_{11}-(1/2)(\beta_{12}+\beta_{22}+1)\quad -(1/2)(\alpha'_{12}+\alpha'_{22}-\beta_{12}-\beta_{22})+\alpha'_{11}-\beta_{11}\quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)}}\,.
\end{aligned}
\tag{B18}
$$

Then

$$
\begin{aligned}
&C^{A'\quad (1/2)\quad (1/2)(\alpha'_{12}-\alpha'_{22})}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)\quad (1/2)\quad \alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22})} \\
&\quad\times C^{(1/2)(\alpha'_{12}-\alpha'_{22})\quad (1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22})\quad (1/2)(m'_{12}-m'_{22})}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22})\quad -(1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22})+m'_{11}-\alpha'_{11}\quad m'_{11}-(1/2)(m'_{12}+m'_{22})} \\
&\quad\times\sum_{M'}(-1)^{M'+\alpha'_{22}+(1/2)(m'_{12}+m'_{22}+1)}[(\alpha'_{12}-\alpha'_{22}+1)(2M'+1)]^{1/2} \\
&\quad\times\begin{Bmatrix} (1/2) & A' & (1/2)(\alpha'_{12}-\alpha'_{22}) \\ (1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22}) & (1/2)(m'_{12}-m'_{22}) & M' \end{Bmatrix} \\
&\quad\times C^{M'\quad (1/2)\quad (1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22}+1)\quad (1/2)\quad m'_{11}-(1/2)(m'_{12}+m'_{22})} \\
&\quad\times\boxed{C^{A'\quad (1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22})\quad M'}_{\alpha'_{11}-(1/2)(\alpha'_{12}+\alpha'_{22}+1)\quad -(1/2)(m'_{12}+m'_{22}-\alpha'_{12}-\alpha'_{22})+m'_{11}-\alpha'_{11}\quad m'_{11}-(1/2)(m'_{12}-m'_{22}+1)}}
\end{aligned}
\tag{B19}
$$

It is now a straightforward matter to introduce the remaining two Wigner coefficients from the original matrix element, eliminate the Racah coefficients in Eqs. (B18) and (B19), and perform the summations over $\frac{1}{2}(\beta_{12}-\beta_{22})$ and $\frac{1}{2}(\alpha'_{12}-\alpha'_{22})$. We are then able to assemble the whole reconstituted matrix element simply by redefining the following indices of summation:

$$
\begin{aligned}
&\beta_{12}+\beta_{22}+1\to\beta_{12}+\beta_{22}, \\
&\beta_{11}+1\to\beta_{11}, \\
&\alpha'_{12}+\alpha'_{22}+1\to\alpha'_{12}+\alpha'_{22}.
\end{aligned}
\tag{B20}
$$

Once we have made these redefinitions to bring the reconstituted matrix element into the form (A20) and assembled our final results, we find that the right side Eq. of (B1a) which results from the terms corresponding to the first two terms on the right of Eq. (B10) has the form

$$
\begin{aligned}
&\omega\mathscr{P}_3\sum_{K,\kappa}\sum_{M',M}[1+(-1)^K](-1)^{M'-M-(1/2)(m'_{12}-m'_{22}-1)+(1/2)(m_{12}-m_{22})}\left[\frac{(2M+1)(2M'+1)}{(m_{12}-m_{22}+1)(m'_{12}-m'_{22}+1)}\right]^{1/2} \\
&\quad\times C^{(1/2)\quad (1/2)\quad K}_{-(1/2)\quad (1/2)\quad 0}\,C^{(1/2)\quad (1/2)\quad K}_{\kappa-(1/2)\quad (1/2)\quad \kappa}\,C^{M\quad K\quad (1/2)(m_{12}-m_{22})}_{m_{11}-(1/2)(m_{12}+m_{22}\quad 0\quad m_{11}-(1/2)(m_{12}+m_{22})} \\
&\quad\times C^{M\quad K\quad (1/2)(m_{12}-m_{22})}_{-m_{23}-\kappa+(1/2)(m_{12}+m_{22})\quad \kappa\quad -m_{23}+(1/2)(m_{12}+m_{22})}\,C^{M'\quad (1/2)\quad (1/2)(m'_{12}-m'_{22})}_{m'_{11}-(1/2)(m'_{12}+m'_{22}+1)\quad (1/2)\quad m'_{11}-(1/2)(m'_{12}+m'_{22})} \\
&\quad\times C^{M'\quad (1/2)\quad (1/2)(m'_{12}-m'_{22})}_{-m'_{23}-\kappa+(1/2)(m'_{12}+m'_{22}+1)\quad \kappa-(1/2)\quad -m'_{23}+(1/2)(m'_{12}+m'_{22})} \\
&\left\langle\left[\Pi;\begin{pmatrix} & m'_{23}+\kappa & \\ M'+(1/2)(m'_{12}+m'_{22}+1) & & -M'+(1/2)(m'_{12}+m'_{22}+1) \\ & m'_{11} & \end{pmatrix}\right]\right. \\
&\qquad\left.|\,\mathscr{O}(U,Z)\,|\left[\Pi;\begin{pmatrix} & m_{23}+\kappa & \\ M+(1/2)(m_{12}+m_{22}) & & -M+(1/2)(m_{12}+m_{22}) \\ & m_{11} & \end{pmatrix}\right]\right\rangle.
\end{aligned}
\tag{B21}
$$

Let us consider the factor $[1+(-1)^K]$ which occurs in the summand. The first term corresponds to the first term on the right of Eq. (B10), which we have been working out in detail; the second term $(-1)^K$ comes from the second term on the right of Eq. (B10). Since K can take on only the values $+1$ and 0, it follows that the $K = 1$ term vanishes identically. This necessarily implies that $\kappa = 0$ and $M = \frac{1}{2}(m_{12} - m_{22})$, i.e., as expected, the operator P_{24} does not shift the labels on the right side of the original matrix element. Writing Eq. (B21) in the form (B1a), we find

$$\left\langle\left\langle \Pi; \begin{pmatrix} & m'_{23} & \\ m'_{12}+1 & & m'_{22} \\ & m'_{11} & \end{pmatrix} \right|\right| P_{24} \left|\left| \Pi; \begin{pmatrix} & m'_{23} & \\ m'_{12} & & m'_{22} \\ & m'_{11} & \end{pmatrix} \right\rangle\right\rangle$$

$$= \omega \mathscr{P}_3 C^{(1/2)(m'_{12}-m'_{22}+1)\quad (1/2)\quad (1/2)(m'_{12}-m'_{22})}_{-m'_{23}+(1/2)(m'_{12}+m'_{22}+1)\quad -(1/2)\quad -m'_{23}+(1/2)(m'_{12}+m'_{22})} C^{(1/2)(m'_{12}-m'_{22})\quad (1/2)\quad (1/2)(m'_{12}-m'_{22}+1)}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\quad -(1/2)\quad m'_{11}-(1/2)(m'_{12}+m'_{22}+1)},$$

$$\left\langle\left\langle \Pi; \begin{pmatrix} & m'_{23} & \\ m'_{12} & & m'_{22}+1 \\ & m'_{11} & \end{pmatrix} \right|\right| P_{24} \left|\left| \Pi; \begin{pmatrix} & m'_{23} & \\ m'_{12} & & m'_{22} \\ & m'_{11} & \end{pmatrix} \right\rangle\right\rangle$$

$$= \omega \mathscr{P}_3 C^{(1/2)(m'_{12}-m'_{22}-1)\quad (1/2)\quad (1/2)(m'_{12}-m'_{22})}_{-m'_{23}+(1/2)(m'_{12}+m'_{22}+1)\quad -(1/2)\quad -m'_{23}+(1/2)(m'_{12}+m'_{22})} C^{(1/2)(m'_{12}-m'_{22})\quad (1/2)\quad (1/2)(m'_{12}-m'_{22}-1)}_{m'_{11}-(1/2)(m'_{12}+m'_{22})\quad -(1/2)\quad m'_{11}-(1/2)(m'_{12}+m'_{22}+1)}, \tag{B22}$$

which is in accordance both with the first two terms of Eq. (3.11b) and with Eq. (B9) and the requirement $P_{24} = (P_{42})^{\dagger}$.

[1]K.I. Gross, W.J. Holman, III, and R.A. Kunze, Bull. Am. Math. Soc. **83**, 412 (1977).

[2]K.I. Gross and W.J. Holman, III, Matrix-valued special functions and representation theory of the conformal group, I: the generalized gamma function (to be published).

[3]K.I. Gross and R.A. Kunze, "Fourier Bessel transforms and unitary representations," in *Conference on Harmonic Analysis*, Lecture Notes in Mathematics (Springer, Berlin and New York, 1971), Vol. 266, pp. 79–122.

[4]K.I. Gross and R.A. Kunze, Generalized Bessel transforms and unitary representations, in "Harmonic Analysis on Homogeneous Spaces," American Mathematical Society Symposium XXVI, Providence, Rhode Island, 1973, pp. 347–350.

[5]K.I. Gross and R.A. Kunze, J. Funct. Anal. **22**, 73 (1976).

[6]K.I. Gross and R.A. Kunze, J. Funct. Anal. **25**, 1 (1977).

[7]K.I. Gross and W.J. Holman, III, Matrix-valued special functions and representation theory of the conformal group, II: the generalized Bessel function (in preparation).

[8]J.D. Louck, Am. J. Phys. **38**, 3 (1970).

[9]A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968).

[10]J.D. Louck and L.C. Biedenharn, J. Math. Anal. Appl. **59**, 423 (1977).

[11]W.J. Holman, III, Nuovo Cimento A **4**, 904 (1971).

[12]E. Chacon, M. Ciftan, and L.C. Biedenharn, J. Math. Phys. **13**, 577 (1972).

Selection rules for type II Shubnikov space groups

R. Dirl
Institut fr Theoretitsche Physick, TU Wien, A-1040 Wien, Karlsplatz 13; Austria

(Received 18 July 1979; accepted for publication 26 October 1979)

Clebsch–Gordan coefficients for type II Shubnikov space groups, containing the inversion as point group operation are expressed by simple formulas in terms of convenient Clebsch–Gordan coefficients for the unitary subgroup.

INTRODUCTION

This paper continues a series of articles which dealt with the problem of decomposing Kronecker products of antiunitary representations into their irreducible constituents. Within the present work we transfer the general formulas which have been derived in Refs. 1–6 to type II Shubnikov space groups containing the inversion I as point group operation.

Due to the general approach given in Refs. 1–6, convenient Clebsch–Gordan coefficients (CG coefficients) for the unitary subgroup (= ordinary space group) have to be computed at first, which are then used to calculate unitary transformations linking CG coefficients for type II Shubnikov space groups with them for the unitary subgroup.

The material is organized as follows: In Sec. I we recall briefly the basic definitions and notations concerning ordinary space groups and their unitary irreducible representations (unirreps). In the following section we summarize the main properties of type II Shubnikov space groups and write down their co-unirreps in standard form. Assuming for the following considerations that the inversion I is contained as point group operation in the unitary subgroup, conditions are derived under which a given space group unirrep can be a constituent of a co-unirrep of type I or type II. Type III co-unirreps are easily distinguishable, since they contain two inequivalent unirreps from the unitary subgroup. Section III is devoted to the problem of determining those unitary matrices which link CG coefficients for type II Shubnikov space groups with convenient ones for the unitary subgroup. Utilizing the fact that for nearly all cases the multiplicity problem for ordinary space group CG coefficients can be solved in a very special way,[7] the corresponding coefficients are given by special expressions. These coefficients are then used to compute for all possible cases unitary matrices which lead immediately to the desired CG coefficients for type II Shubnikov space groups. Since restricting our considerations to space groups which contain the inversion I as the point group operation, we are able to give simple solutions for the multiplicity problem (containing among others the "wave vector selection rules") and the corresponding unitary matrices.

I. DEFINITIONS AND NOTATIONS

Let $M = G\times\{E,\theta\} = \{G,\theta G\}$ be a type II Shubnikov space group,[8] where G is the corresponding ordinary space group:

$$M=\{(\alpha\,|\,\tau(\alpha)+\mathbf{t}\,|\,\theta^{k}):\alpha\in P,\ \mathbf{t}\in T,\ k=0,1\},\tag{I.1}$$

$$(\alpha\,|\,\tau(\alpha)+\mathbf{t}\,|\,\theta^{k})(\alpha'\,|\,\tau(\alpha')+\mathbf{t}'\,|\,\theta^{k'})$$
$$=(\alpha\alpha'\,|\,\tau(\alpha\alpha')+\mathbf{t}+D(\alpha)\mathbf{t}'+\mathbf{t}(\alpha,\alpha')\,|\,\theta^{k+k'}),\tag{I.2}$$

$$\mathbf{t}(\alpha,\alpha')=\tau(\alpha)+D(\alpha)\tau(\alpha')-\tau(\alpha\alpha').\tag{I.3}$$

The symbols $\mathbf{t}$ denote primitive lattice translations, $\tau(\alpha)$ nonprimitive lattice translations, $D=\{D(\alpha):\alpha\in P\}$ is a faithful representation[8] of the point group $P\simeq G/T$ of the crystal, and θ the time reversal operation. Since M is a direct product, we write sometimes the elements of the subgroup G for the sake of simplicity as $(\alpha|\tau(\alpha)+\mathbf{t})$ instead of $(\alpha|\tau(\alpha)+\mathbf{t}|E)$ and θ instead of $(e|\mathbf{0}|\theta)$.

The matrix elements of the vector unirreps of G can be written in the following form[9]:

$$D^{(\mathscr{K},\mathbf{q})\uparrow G}_{\sigma a;\sigma' a'}(\beta\,|\,\tau(\beta)+\mathbf{t})$$
$$=\Delta^{\mathbf{q}}(\sigma,\beta\sigma')e^{-i\mathbf{q}(\sigma)\cdot\mathbf{t}}B^{\mathbf{q}}_{\sigma,\sigma'}(\beta)\mathbb{R}^{\mathscr{K}}_{aa'}(\sigma^{-1}\beta\sigma'),$$
$$\mathbf{q}\in\Delta BZ,\quad \mathscr{K}\in A_{P^{\mathbf{q}}(S^{\mathbf{q}})},\ \sigma,\sigma'\in P:P^{\mathbf{q}},\ a,a'=1,2,\ldots,n_{\mathscr{K}},\tag{I.4}$$

$$P^{\mathbf{q}}=\{\alpha:D(\alpha)\mathbf{q}=\mathbf{q}+\mathbf{Q}\{\mathbf{q}(\alpha)\};\ \alpha\in P\},\tag{I.5}$$

$$\Delta^{\mathbf{q}}(\gamma,\gamma')=\delta_{\gamma P^{\mathbf{q}},\gamma' P^{\mathbf{q}}},\quad\text{for all}\ \gamma,\gamma'\in P,\tag{I.6}$$

$$\mathbf{q}(\gamma)=D(\gamma)\mathbf{q},\tag{I.7}$$

$$B^{\mathbf{q}}_{\sigma,\sigma'}(\beta)=\exp[-i\mathbf{q}(\sigma)\cdot\{\tau(\beta)+D(\beta)\tau(\sigma')-\tau(\sigma)\}],\quad\text{for all}\ \beta\in P.\tag{I.8}$$

Thereby ΔBZ denotes the fundamental (representation) domain of the Brillouin zone BZ, $G^{\mathbf{q}}$ the group of the $\mathbf{q}$ vector, and $\mathbf{Q}\{\mathbf{q}(\alpha)\}$ reciprocal lattice vectors; $\sigma,\sigma'\in P:P^{\mathbf{q}}$ left coset representatives of $P^{\mathbf{q}}\simeq G^{\mathbf{q}}/T$ with respect to P and $\mathbb{R}^{\mathscr{K}}=\{\mathbb{R}^{\mathscr{K}}(\alpha):\alpha\in P^{\mathbf{q}}\}$ $n_{\mathscr{K}}$-dimensional projective unirreps of $P^{\mathbf{q}}$ which belong to the standard factor system

$$S^{\mathbf{q}}(\alpha,\beta)=\exp\{-i\mathbf{q}\cdot(D(\alpha)-\mathbb{1})\tau(\beta)\},\quad\text{for all}\ \alpha,\beta\in P^{\mathbf{q}}.\tag{I.9}$$

(In this connection we have to note that for symmorphic space groups the projective unirreps $\mathbb{R}^{\mathscr{K}}$ reduce to ordinary vector unirreps, whereas for nonsymmorphic space groups this proposition only holds for $\mathbf{q}$'s not lying on the "surface" of ΔBZ.)

II. CO-UNIRREPS OF M

In order to be able to write down the three different types of co-unirreps of M, it is necessary to consider

$$D^{(\mathscr{K},\mathbf{q})\uparrow G}((e\,|\,\mathbf{0}\,|\,\theta)(\alpha\,|\,\tau(\alpha)+\mathbf{t}\,|\,E)(e\,|\,\mathbf{0}\,|\,\theta))$$

 0022-2488/80/082011-08$1.00

$$= D^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}), \quad \text{(II.1)}$$

which is an obvious consequence of the direct product structure of M. Thus, it suffices to compute a $|P:P^{\mathbf{q}}|n_{\mathscr{K}}$-dimensional unitary matrix $U^{(\mathscr{K},\mathbf{q})}$ satisfying

$$D^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})^* = U^{(\mathscr{K},\mathbf{q})\dagger}D^{\overline{(\mathscr{K},\mathbf{q})\uparrow G}}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})U^{(\mathscr{K},\mathbf{q})}, \quad \text{for all } (\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})\in G, \quad \text{(II.2)}$$

where the same notation for identical equivalence classes is introduced as in Ref. 10, i.e.,

$$\{(\mathscr{H},\mathbf{q})\uparrow G\}^* = (\overline{\mathscr{H}},\mathbf{q}')\uparrow G, \quad \mathbf{q}'\in \Delta BZ \text{ and } \overline{\mathscr{H}}\in A_{P^{\mathbf{q}}(S^{\mathbf{q}})}. \quad \text{(II.3)}$$

Hence, it follows that the first step must be to determine for a given equivalence class $(\mathscr{H},\mathbf{q})\uparrow G$ the corresponding equivalence class $(\overline{\mathscr{H}},\mathbf{q}')\uparrow G$. This can be done by means of the character test given by Dimmock and Wheeler,[8] or by investigating the formulas (III.3), (III.8), or (III.16) of Ref. 10, which however presuppose that the inversion I is a symmetry operation for the crystal in question, and to consider Eq. (V.3) of Ref. 8 for the general case.

Provided this task has been solved we are in the position to write down the three different types of co-unirreps of M in their standard form:

type I: $\{(\mathscr{H},\mathbf{q})\uparrow G\}^* = (\mathscr{H},\mathbf{q})\uparrow G$,

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}) = D^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}), \quad \text{for all } (\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})\in G, \quad \text{(II.4)}$$

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\theta) = U^{(\mathscr{K},\mathbf{q})}, \quad \text{(II.5)}$$

$$U^{(\mathscr{K},\mathbf{q})}U^{(\mathscr{K},\mathbf{q})*} = +\mathbb{1}_{(\mathscr{K},\mathbf{q})\uparrow G}; \quad \text{(II.6)}$$

type II: $\{(\mathscr{H},\mathbf{q})\uparrow G\}^* = (\mathscr{H},\mathbf{q})\uparrow G$,

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}) = \begin{bmatrix} D^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}) & 0 \\ 0 & D^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}) \end{bmatrix}, \quad \text{(II.7)}$$

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\theta) = \begin{bmatrix} 0 & U^{(\mathscr{K},\mathbf{q})} \\ -U^{(\mathscr{K},\mathbf{q})} & 0 \end{bmatrix}, \quad \text{(II.8)}$$

$$U^{(\mathscr{K},\mathbf{q})}U^{(\mathscr{K},\mathbf{q})*} = -\mathbb{1}_{(\mathscr{K},\mathbf{q})\uparrow G}; \quad \text{(II.9)}$$

type III: $\{(\mathscr{H},\mathbf{q})\uparrow G\}^* \neq (\mathscr{H},\mathbf{q})\uparrow G$,

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}) = \begin{bmatrix} D^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}) & 0 \\ 0 & D^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})^* \end{bmatrix}, \quad \text{(II.10)}$$

$$D^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})^* = U^{(\mathscr{K},\mathbf{q})\dagger}D^{(\overline{\mathscr{K}},\mathbf{q}')\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})U^{(\mathscr{K},\mathbf{q})}, \quad \text{(II.11)}$$

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\theta) = \begin{bmatrix} 0 & \mathbb{1}_{(\mathscr{K},\mathbf{q})\uparrow G} \\ \mathbb{1}_{(\mathscr{K},\mathbf{q})\uparrow G} & 0 \end{bmatrix}, \quad \text{(II.12)}$$

$$U^{(\mathscr{K},\mathbf{q})}U^{(\mathscr{K},\mathbf{q})\dagger} = \mathbb{1}_{(\mathscr{K},\mathbf{q})\uparrow G}. \quad \text{(II.13)}$$

These matrices are satisfying well known properties which are inherent to corepresentations, namely,

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}((e|\mathbf{0}|\theta)(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})(e|\mathbf{0}|\theta))^* = \mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\theta)^\dagger\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\theta), \quad \text{(II.14)}$$

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\theta)\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\theta)^* = \mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(e|\mathbf{0}|E), \quad \text{(II.15)}$$

$$\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\theta)\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})^* = \mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t}|\theta), \quad \text{for all } (\alpha|\boldsymbol{\tau}(\alpha)+\mathbf{t})\in G. \quad \text{(II.16)}$$

For the following we restrict our considerations to space groups which contain the inversion I as point group operation. The reason for this restriction is that only for such space groups are we in the position to give closed expressions for the unitary matrices $U^{(\mathscr{K},\mathbf{q})}$ (see Ref. 11). The first consequence is

$$\{(\mathscr{H},\mathbf{q})\uparrow G\}^* = (\overline{\mathscr{H}},\mathbf{q})\uparrow G, \quad \mathbf{q}\in \Delta BZ \text{ and } \mathscr{H},\overline{\mathscr{H}}\in A_{P^{\mathbf{q}}(S^{\mathbf{q}})} \quad \text{(II.17)}$$

[compare Eq. (II.5) of Ref. 10], which implies that inequivalence of $\{(\mathscr{H},\mathbf{q})\uparrow G\}^*$ and $(\mathscr{H},\mathbf{q})\uparrow G$ occurs only if $\overline{\mathscr{H}}\neq\mathscr{H}$. Furthermore, we derived in Ref. 11 two different types of unitary matrices $U^{(\mathscr{K},\mathbf{q})}$, depending on whether the inversion I belongs to $P^{\mathbf{q}}$ or not, namely,

$$U^{(\mathscr{K},\mathbf{q})}_{\sigma,\sigma'} = \delta_{\sigma,\sigma'}U^{\mathscr{K}}, \quad I\in P^{\mathbf{q}}, \quad \text{(II.18)}$$

$$\mathbb{R}^{\mathscr{K}}(\alpha)^* = e^{i\mathbf{Q}\{\mathbf{q}(I)\}\cdot\boldsymbol{\tau}(\alpha)}U^{\mathscr{K}\dagger}\mathbb{R}^{\overline{\mathscr{K}}}(\alpha)U^{\mathscr{K}}, \quad \text{for all } \alpha\in P^{\mathbf{q}}, \quad \text{(II.19)}$$

$$U^{(\mathscr{K},\mathbf{q})}_{\sigma,\sigma'} = \delta_{\sigma',I\sigma}e^{i\mathbf{q}(\sigma)\cdot\mathbf{t}(\sigma,I)}U^{\mathscr{K}}, \quad \text{(II.20)}$$

$$\mathbb{R}^{\mathscr{K}}(\alpha)^* = e^{i\mathbf{Q}\{\mathbf{q}(\alpha^{-1})\}\cdot\boldsymbol{\tau}(I)}U^{\mathscr{K}\dagger}\mathbb{R}^{\overline{\mathscr{K}}}(\alpha)U^{\mathscr{K}}, \quad \text{for all } \alpha\in P^{\mathbf{q}}, \quad \text{(II.21)}$$

where we have partly used matrix notation for $U^{(\mathscr{K},\mathbf{q})}$. Obviously, Eq. (II.18) can only be realized for $\mathbf{q}$'s lying on the surface of ΔBZ, whereas Eq. (II.20) is either possible for $\mathbf{q}$'s which are elements of the surface of ΔBZ or lying inside of ΔBZ. In this connection we have to note that the first situation is rather rare whereas the second possibility will be the usual one. Nevertheless, we are forced to discuss both situations. Furthermore, for the last case (i.e., $\mathbf{q}\notin$surface of ΔBZ) the unimodular factor appearing in Eq. (II.21) reduces to 1 and Eq. (II.21) represents the usual equivalence relation for ordinary vector unirreps.

Provided $\overline{\mathscr{H}} = \mathscr{H}$ is valid we can readily verify for both cases [Eqs. (II.18) and (II.20)]

$$\sum_{\sigma} U^{(\mathscr{K},\mathbf{q})}_{\sigma',\sigma}U^{(\mathscr{K},\mathbf{q})*}_{\sigma,\sigma''} = \delta_{\sigma',\sigma''}U^{\mathscr{K}}U^{\mathscr{K}*}, \quad \text{(II.22)}$$

which implies that the property

$$U^{(\mathscr{K},\mathbf{q})}U^{(\mathscr{K},\mathbf{q})*} = \pm\mathbb{1}_{(\mathscr{K},\mathbf{q})\uparrow G}, \quad \text{(II.23)}$$

characterizing type I or type II co-unirreps, can only be a consequence of

$$U^{\mathscr{K}}U^{\mathscr{K}*} = \pm\mathbb{1}_{\mathscr{K}}. \quad \text{(II.24)}$$

Thus, we arrive at the important result that the property of a space group unirrep to be a constituent of a type I or type II co-unirrep [of a type II Shubnikov space group which contains the inversion $(I|\boldsymbol{\tau}(I))$] originates only from the projective point group unirreps of $P^{\mathbf{q}}$. (This holds of course also for type III corepresentations.) Now let us consider in more detail Eq. (II.24) in order to see whether we can make general predictions with respect to the type of co-unirreps or not.

For this purpose we recall that the matrix elements of the $n_{\mathscr{K}}$-dimensional unitary matrix $U^{\mathscr{K}}$ can be written as

$$U^{\mathscr{K}}_{d'd} = \sqrt{\frac{n_{\mathscr{K}}}{|P^{\mathbf{q}}|}} \left\{ \sum_{\alpha\in P^{\mathbf{q}}} e^{i\mu(\alpha)} \mathbb{R}^{\mathscr{K}}_{11}(\alpha) \mathbb{R}^{\mathscr{K}}_{c_0'c_0'}(\alpha) \right\}^{-1/2} \times \sum_{\beta\in P^{\mathbf{q}}} e^{i\mu(\beta)} \mathbb{R}^{\mathscr{K}}_{d1}(\beta) \mathbb{R}^{\mathscr{K}}_{d'c_0'}(\beta), \tag{II.25}$$

where the unimodular factors $e^{i\mu(\alpha)}$ are defined by

$$e^{i\mu(\alpha)} = \begin{cases} e^{i\mathbf{Q}[\mathbf{q}(I)]\cdot\tau(\alpha)}, & \text{for } I\in P^{\mathbf{q}}, \\ e^{i\mathbf{Q}[\mathbf{q}(\alpha^{-1})]\cdot\tau(I)}, & \text{for } I\notin P^{\mathbf{q}} \end{cases} \tag{II.26}$$

[compare Eqs. (II.7), (II.20), (II.33), and (II.45) of Ref. 11]. Introducing the following definitions (see Ref. 12):

$$\{\mathbf{B}_{aa'}\}_{dd'} = \delta_{ad}\delta_{a'd'}, \quad a,a',d,d' = 1,2,\dots,n_{\mathscr{K}}, \tag{II.27}$$

$$\mathbb{E}^{\mathscr{K}\mathscr{K}^*;0} = \frac{1}{|P^{\mathbf{q}}|} \sum_{\alpha\in P^{\mathbf{q}}} \mathbb{R}^{\mathscr{K}}(\alpha) \otimes U^{\mathscr{K}}\mathbb{R}^{\mathscr{K}}(\alpha)^* U^{\mathscr{K}\dagger}, \tag{II.28}$$

$$\mathbf{B}^{0(1c_0')} = \mathbb{E}^{\mathscr{K}\mathscr{K}^*;0}\mathbf{B}_{1c_0'}, \tag{II.29}$$

we obtain immediately

$$\begin{aligned} \sum_{\alpha} e^{i\mu(\alpha)} \mathbb{R}^{\mathscr{K}}_{11}(\alpha) \mathbb{R}^{\mathscr{K}}_{c_0'c_0'}(\alpha) &= |P^{\mathbf{q}}| \langle \mathbf{B}_{1c_0'}, \mathbb{E}^{\mathscr{K}\mathscr{K}^*;0}\mathbf{B}_{1c_0'} \rangle \\ &= |P^{\mathbf{q}}|\, \|\mathbf{B}^{0(1c_0')}\|^2 = |P^{\mathbf{q}}|\, \|\mathbf{B}^{0(c_0'1)}\|^2 \end{aligned} \tag{II.30}$$

$$\mathbf{B}^{0(1c_0')} = e^{i\rho}\mathbf{B}^{0(c_0'1)} \tag{II.31}$$

since $m_{\mathscr{K}\mathscr{K}^*;0} = 1$. On the other hand, we obtain for

$$\begin{aligned} \sum_{d} U^{\mathscr{K}}_{d'd} U^{\mathscr{K}*}_{dd''} &= \delta_{d'd''} |P^{\mathbf{q}}|^{-1} \|\mathbf{B}^{0(1c_0')}\|^{-2} \sum_{\alpha} e^{i\mu(\alpha)} \mathbb{R}^{\mathscr{K}}_{c_0'1}(\alpha) \mathbb{R}^{\mathscr{K}}_{1c_0'}(\alpha) \\ &= \delta_{d'd''} \|\mathbf{B}^{0(1c_0')}\|^{-2} \langle \mathbf{B}_{c_0'1}, \mathbb{E}^{\mathscr{K}\mathscr{K}^*;0}\mathbf{B}_{1c_0'} \rangle \\ &= \delta_{d'd''} \|\mathbf{B}^{0(1c_0')}\|^{-2} \langle \mathbf{B}_{1c_0'}, \mathbb{E}^{\mathscr{K}\mathscr{K}^*;0}\mathbf{B}_{c_0'1} \rangle \\ &= e^{i\rho}\delta_{d'd''}, \end{aligned} \tag{II.32}$$

which has as a consequence that the phase factor $e^{i\rho}$ can take only the values ± 1, i.e.,

$$e^{i\rho} = \pm 1, \tag{II.33}$$

since $\mathbb{E}^{\mathscr{K}\mathscr{K}^*;0}$ is a Hermitian matrix. In order to be able to verify the first line of Eq. (II.32) one has to use among others the equivalence relations (III.9) and (III.17) of Ref. 10. Thus, we have shown generally that the unitary matrices $U^{\mathscr{K}}$ satisfy Eq. (II.24), but we cannot predict for the general case whether the phase factor $e^{i\rho}$ is $+1$ or -1. Only for such cases where we can choose $c_0' = c_0 = 1$ must the corresponding co-unirreps be of type I.

III. CG COEFFICIENTS FOR TYPE II SHUBNIKOV SPACE GROUPS

Due to the general approach given in Refs. 1–6 the first task is to determine convenient CG coefficients for the normal subgroup G. Obviously, it suffices to compute only those columns of the corresponding CG matrices, which refer to a fixed column index a_0 of the considered unirrep, since the remaining column vectors are immediately obtained by projection techniques as was described in Refs. 13 and 7. Provided such vectors have been determined, one has to calculate with the aid of them the unitary matrices B and C which link CG coefficients for M with those for G. Thereby it is very useful to utilize the additional symmetry properties of the submatrices of F, which are the nontrivial constituents of B and C.

As already pointed out, we have shown in Ref. 7 that the multiplicity problem for space group CG coefficients can be solved in a very special way for nearly all cases by taking special column indices of the considered Kronecker product as the multiplicity index. This has the consequence that we are able to give closed expressions for the corresponding CG coefficients without reference to a special space group. This implies that it suffices to transfer special formulas (which have been derived in Refs. 1–6) to space group representations, which are needed to calculate the $m_{\mu_1\mu_2;\mu}$-dimensional submatrices $F^{\mu(\mu_1\mu_2)}$ of F.

In order to simplify the notation let us use sometimes the following symbols:

$$\begin{aligned} &\mu \leftrightarrow (\mathscr{H}_0,\mathbf{q}_0)\uparrow G, \\ &\mu_1 \leftrightarrow (\mathscr{H},\mathbf{q})\uparrow G, \\ &\mu_2 \leftrightarrow (\mathscr{H}',\mathbf{q}')\uparrow G, \\ &i \leftrightarrow (\sigma,c), \\ &i_v \leftrightarrow (\sigma_v,c_v), \quad \sigma,\sigma_v \in P:P^{\mathbf{q}}, \quad c,c_v = 1,2,\dots,n_{\mathscr{K}}, \\ &j \leftrightarrow (\sigma',c'), \\ &j_v \leftrightarrow (\sigma_v',c_v'), \quad \sigma',\sigma_v' \in P:P^{\mathbf{q}'}, \quad c',c_v' = 1,2,\dots,n_{\mathscr{K}'}, \\ &a_0 \leftrightarrow (e,1), \\ &v \leftrightarrow (i_v;j_v) \leftrightarrow (\sigma_v,c_v,\sigma_v',c_v'), \\ &\qquad v = 1,2,\dots,m_{(\mathscr{K},\mathbf{q})(\mathscr{K}',\mathbf{q}');(\mathscr{K}_0,\mathbf{q}_0)}, \end{aligned} \tag{III.1}$$

$$A_G = A_{\rm I} \cup A_{\rm II} \cup A_{\rm III}. \tag{III.2}$$

Now we are in the position to write down the columns of the convenient CG matrices for G, which are needed for the following considerations.

Case $A \otimes B$:C: $A,B\in A_{\rm I} \cup A_{\rm II}$ and $C\in A_G$. According to Eqs. (III.24), (III.48), and (III.90) of Ref. 1 and Eqs. (II.30), (II.67), and (II.112) of Ref. 2, and Eqs. (II.29), (II.68), and (II.114) of Ref. 4, we have to consider in principle the vectors $\mathbf{M}^{\mu_1\mu_2;\mu(i_vj_v)}_{a_0}$, $\mu\in A_G$, $\mu_1,\mu_2\in A_{\rm I}\cup A_{\rm II}$, whose components are given by

$$\{\mathbf{M}^{\mu_1\mu_2;\mu(i_vj_v)}_{a_0}\}_{ij} = \|\mathbf{B}^{\mu_1\mu_2;\mu(i_vj_v)}_{a_0}\|^{-1} \delta_{\mathbf{q}(\sigma)+\mathbf{q}'(\sigma'),\mathbf{q}_0+\mathbf{Q}[\mathbf{q}(\sigma)+\mathbf{q}'(\sigma')]} \frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|} \sum_{\beta\in P} \Delta^{\mathbf{q}}(\sigma,\beta\sigma_v) B^{\mathbf{q}}_{\sigma,\sigma_v}(\beta) \mathbb{R}^{\mathscr{K}}_{cc_v}(\sigma^{-1}\beta\sigma_v) \Delta^{\mathbf{q}'}(\sigma',\beta\sigma_v')$$
$$\times B^{\mathbf{q}'}_{\sigma',\sigma_v'}(\beta) \mathbb{R}^{\mathscr{K}'}_{c'c_v'}(\sigma'^{-1}\beta\sigma_v') \Delta^{\mathbf{q}_0}(e,\beta) B^{\mathbf{q}_0}_{e,e}{}^*(\beta) \mathbb{R}^{\mathscr{K}_0}_{11}(\beta)^*, \quad \mu_1,\mu_2\in A_{\rm I}\cup A_{\rm II}, \quad \mu\in A_G. \tag{III.3}$$

Thereby we have to note that the multiplicity indices

$$v\leftrightarrow(i_v;j_v)\leftrightarrow(\sigma_v,c_v;\sigma'_v,c'_v),\quad v=1,2,...,m_{\mu_1\mu_2;\mu} \tag{III.4}$$

of course depend on the considered case. The remaining vectors $\mathbf{M}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_vj_v)}$, $\bar{\mu}\in A_{\rm III}$ and $\mu_1,\mu_2\in A_{\rm I}\cup A_{\rm II}$, which are defined by Eq. (III.91) of Ref. 1, Eq. (II.113) of Ref. 2, and Eq. (II.115) of Ref. 4, take the following form:

$$\{\mathbf{M}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_vj_v)}\}_{ij}=\|\mathbf{B}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_vj_v)}\|^{-1}\delta_{\mathbf{q}(\sigma)+\mathbf{q}'(\sigma'),-\mathbf{q}_0+\mathbf{Q}[\cdots]}$$
$$\times\frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}\sum_{\beta\in P}\Delta^{\mathbf{q}}(\sigma,\beta\sigma_v)B^{\mathbf{q}}_{\sigma,\sigma_v}(\beta)\mathbb{R}^{\mathscr{K}}_{cc_v}(\sigma^{-1}\beta\sigma_v)\Delta^{\mathbf{q}'}(\sigma',\beta\sigma'_v)B^{\mathbf{q}'}_{\sigma',\sigma'_v}(\beta)\mathbb{R}^{\mathscr{K}'}_{c'c'_v}(\sigma'^{-1}\beta\sigma'_v)\Delta^{\mathbf{q}_0}(e,\beta)B^{\mathbf{q}_0}_{e,e}(\beta)\mathbb{R}^{\mathscr{K}_0}_{11}(\beta),$$
$$\mu_1,\mu_2\in A_{\rm I}\cup A_{\rm II},\quad \bar{\mu}\in A_{\rm III}. \tag{III.5}$$

Comparing the wave vector selection rules, which are contained in Eqs. (III.3) and (III.5) for the case $\mu\in A_{\rm III}\Longleftrightarrow\bar{\mu}\in A_{\rm III}$, we are confronted with two different situations, depending on whether the complex conjugation of $D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}$ is governed by Eq. (II.18) or (II.20). Provided

$$\mathbf{q}(\sigma_v)+\mathbf{q}'(\sigma'_v)=\mathbf{q}_0+\mathbf{Q}[\mathbf{q}(\sigma_v)+\mathbf{q}'(\sigma'_v)],\quad \mathbf{q}_0\in\Delta BZ \tag{III.6}$$

is valid and Eq. (II.18) connects $D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G*}$ with $D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}$, i.e.,

$$\{U^{(\mathscr{K},\mathbf{q})\dagger}D^{(\bar{\mathscr{K}},\mathbf{q})\uparrow G}(\beta|\tau(\beta)+\mathbf{t})U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_1,\sigma_2}$$
$$=\Delta^{\mathbf{q}}(\sigma_1,\beta\sigma_2)e^{-i\mathbf{q}(\sigma_1)\cdot\mathbf{t}}B^{\mathbf{q}}_{\sigma_1,\sigma_2}(\beta)U^{\mathscr{K}\dagger}\mathbb{R}^{\mathscr{K}}(\sigma_1^{-1}\beta\sigma_2)U^{\mathscr{K}}, \tag{III.7}$$

it follows that

$$\mathbf{q}(\sigma_v)+\mathbf{q}'(\sigma'_v)=-\mathbf{q}_0+\mathbf{Q}[\cdots]$$
$$=\mathbf{q}_0+\mathbf{Q}\{\mathbf{q}_0(I)\}+\mathbf{Q}[\cdots], \tag{III.8}$$

where the reciprocal lattice vector $\mathbf{Q}\{\mathbf{q}_0(I)\}$ arises from the definition of the corresponding little co-group $P^{\mathbf{q}_0}$. Equation (III.8) has as a consequence that for both situations ($\mathbf{M}_{a_0}^{\mu v}$ and $\mathbf{M}_{a_0}^{\bar{\mu}v}$) the same left coset representatives (σ_v,σ'_v) can be used as a part of the multiplicity index, whereas the indices c_v and c'_v are not necessarily the same. On the other hand, if Eq. (II.20) connects $D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G*}$ with $D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}$, i.e.,

$$\{U^{(\mathscr{K},\mathbf{q})\dagger}D^{(\bar{\mathscr{K}},\mathbf{q})\uparrow G}(\beta|\tau(\beta)+\mathbf{t})U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_1,\sigma_2}$$
$$=\Delta^{\mathbf{q}}(\sigma_1,\beta\sigma_2)e^{i\mathbf{q}(\sigma_1)\cdot\mathbf{t}}e^{-i\mathbf{q}(\sigma_1)\cdot\mathbf{t}(\sigma_1,I)+i\mathbf{q}(\sigma_2)\cdot\mathbf{t}(\sigma_2,I)}$$
$$\times B^{\mathbf{q}}_{I\sigma_1,I\sigma_2}(\beta)U^{\mathscr{K}\dagger}\mathbb{R}^{\mathscr{K}}(\sigma_1^{-1}\beta\sigma_2)U^{\mathscr{K}}, \tag{III.9}$$

we have

$$\mathbf{q}(I\sigma_v)+\mathbf{q}'(I\sigma'_v)=\mathbf{q}_0-\mathbf{Q}[\mathbf{q}(\sigma_v)+\mathbf{q}'(\sigma'_v)] \tag{III.10}$$

since $I\notin P^{\mathbf{q}_0}$. Thus, we arrive in this case at the result that at least one of the two left coset representatives σ_v and σ'_v has to be replaced by $I\sigma_v$ and/or $I\sigma'_v$, since it is impossible that $I\in\sigma_vP^{\mathbf{q}}\sigma_v^{-1}$ and/or $I\in\sigma'_vP^{\mathbf{q}'}\sigma'^{-1}_v$. That $I\sigma_v$ and/or $I\sigma'_v$ can be chosen as the left coset representative has been shown in Ref. 10. This implies that $D^{(\mathscr{K},\mathbf{q})\uparrow G*}$ and/or $D^{(\mathscr{K}',\mathbf{q}')\uparrow G*}$ are related by matrices of type (II.20) with $D^{(\mathscr{K},\mathbf{q})\uparrow G*}$ and $D^{(\mathscr{K}',\mathbf{q}')\uparrow G}$, respectively.

Case A ⊗ III: *C*: $A\in A_{\rm I}\cup A_{\rm II}$ and $C\in A_G$. According to Eq. (II.35), (II.36), (II.77), (II.78), (II.131) and (II.132) of Ref. 3, and Eqs. (II.37), (II.38), (II.85), (II.86), (II.143), and (II.144) of Ref. 5, we have to consider the vectors $\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_vj_v)}$ and $\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_vj_v)}$, $\mu_1\in A_{\rm I}\cup A_{\rm II}$, $\mu_2\in A_{\rm III}$, and $\mu\in A_G$, whose components are given by

$$\{\mathbf{M}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_vj_v)}\}_{ij}$$
$$=\|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_vj_v)}\|^{-1}\delta_{\mathbf{q}(\sigma)+\mathbf{q}'(\sigma'),\mathbf{q}_0+\mathbf{Q}[\cdots]}$$
$$\times\frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}\sum_{\beta}\Delta^{\mathbf{q}}(\sigma,\beta\sigma_v)B^{\mathbf{q}}_{\sigma,\sigma_v}(\beta)\mathbb{R}^{\mathscr{K}}_{cc_v}(\sigma^{-1}\beta\sigma_v)\Delta^{\mathbf{q}'}(\sigma',\beta\sigma'_v)B^{\mathbf{q}'}_{\sigma',\sigma'_v}(\beta)\mathbb{R}^{\mathscr{K}'}_{c'c'_v}(\sigma'^{-1}\beta\sigma'_v)\Delta^{\mathbf{q}_0}(e,\beta)B^{\mathbf{q}_0*}_{e,e}(\beta)\mathbb{R}^{\mathscr{K}_0}_{11}(\beta)^*, \tag{III.11}$$

$$\{\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_vj_v)}\}_{ij}=\|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_vj_v)}\|^{-1}\delta_{\mathbf{q}(\sigma)-\mathbf{q}'(\sigma'),\mathbf{q}_0+\mathbf{Q}[\cdots]}\frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}\sum_{\beta}\Delta^{\mathbf{q}}(\sigma,\beta\sigma_v)B^{\mathbf{q}}_{\sigma,\sigma_v}(\beta)\mathbb{R}^{\mathscr{K}}_{cc_v}(\sigma^{-1}\beta\sigma_v)\Delta^{\mathbf{q}'}(\sigma',\beta\sigma'_v)B^{\mathbf{q}'}_{\sigma',\sigma'_v}(\beta)^*$$
$$\times\mathbb{R}^{\mathscr{K}'*}_{c'c'_v}(\sigma'^{-1}\beta\sigma'_v)\Delta^{\mathbf{q}_0}(e,\beta)B^{\mathbf{q}_0*}_{e,e}(\beta)\mathbb{R}^{\mathscr{K}}_{11}(\beta)^*,\quad \mu_1\in A_{\rm I}\cup A_{\rm II},\ \mu_2\in A_{\rm III};\mu\in A_G, \tag{III.12}$$

whereas the remaining vectors $\mathbf{M}_{a_0}^{\bar{\mu}v}$ and $\mathbf{N}_{a_0}^{\bar{\mu}v}$ are not written down, since they are immediately obtained from Eqs. (III.11) and (III.12), respectively by replacing the wave vector selection rules $\mathbf{q}(\sigma_v)\pm\mathbf{q}'(\sigma'_v)=\mathbf{q}_0+\mathbf{Q}[\cdots]$ through $\mathbf{q}(\sigma_v)\pm\mathbf{q}'(\sigma'_v)=-\mathbf{q}_0+\mathbf{Q}[\cdots]$ and $\{B^{\mathbf{q}_0}_{e,e}(\beta)\mathbb{R}^{\mathscr{K}_0}_{11}(\beta)\}^*$ through their complex conjugate values. Now, comparing the wave vector selection rules which are contained in Eqs. (III.11) and (III.12), we can use the same left coset representatives (σ_v,σ'_v) as a part of the multiplicity index, if $D^{(\mathscr{K}',\mathbf{q}')\uparrow G*}$ is linked by a unitary matrix of type (II.18) with $D^{(\bar{\mathscr{K}}',\mathbf{q}')\uparrow G}$ and we have to take $(\sigma_v,I\sigma'_v)$, if $D^{(\mathscr{K}',\mathbf{q}')\uparrow G*}$ is connected by a unitary matrix of type (II.20) with $D^{(\bar{\mathscr{K}}',\mathbf{q}')\uparrow G}$. By similar arguments as for Eq. (III.8) or (III.10) it is possible to predict appropriated left coset representatives appearing in $\mathbf{M}_{a_0}^{\mu v}$ and $\mathbf{M}_{a_0}^{\bar{\mu}v}$, $\mu\in A_{\rm III}$ ($\mathbf{N}_{a_0}^{\mu v}$, and $\mathbf{N}_{a_0}^{\bar{\mu}v}$, respectively, $\mu\in A_{\rm III}$).

Case III ⊗ III: *C*: $C\in A_G$. Finally, in accordance with Eqs. (II.46)–(II.49), (II.101)–(II.104), and (II.166)–(II.169) of Ref. 6, we express the corresponding vectors $\mathbf{K}_{a_0}^{\mu v}$, $\mathbf{L}_{a_0}^{\mu v}$, $\mathbf{M}_{a_0}^{\mu v}$, and $\mathbf{N}_{a_0}^{\mu v}$ by a single formula, where the vectors **K**, **L**, **M**, and **N** are abbreviated by the symbol **P**:

$$\{\mathbf{P}_{a_0}^{\mu_1\mu_2;\mu(i_vj_v)}\}_{ij}=\|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_vj_v)}\|^{-1}\delta_{\pm\mathbf{q}(\sigma)\pm\mathbf{q}'(\sigma'),\mathbf{q}_0+\mathbf{Q}[\cdots]}\frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}$$

$$\times \sum_{\beta\in P} \Delta^{\mathbf{q}}(\sigma,\beta\sigma_v) \begin{Bmatrix} B^{\mathbf{q}}_{\sigma,\sigma_v}(\beta)\mathbb{R}^{\mathscr{K}}_{cc_v}(\sigma^{-1}\beta\sigma_v) \\ B^{\mathbf{q}*}_{\sigma,\sigma_v}(\beta)\mathbb{R}^{\mathscr{K}*}_{cc_v}(\sigma^{-1}\beta\sigma_v)\end{Bmatrix} \Delta^{\mathbf{q}'}(\sigma',\beta\sigma'_v) \begin{Bmatrix} B^{\mathbf{q}'}_{\sigma',\sigma'_v}(\beta)\mathbb{R}^{\mathscr{K}'}_{c'c'_v}(\sigma'^{-1}\beta\sigma'_v) \\ B^{\mathbf{q}'*}_{\sigma',\sigma'_v}(\beta)\mathbb{R}^{\mathscr{K}'*}_{c'c'_v}(\sigma'^{-1}\beta\sigma'_v)\end{Bmatrix}$$
$$\times \Delta^{\mathbf{q}_0}(e,\beta) B^{\mathbf{q}_0}_{e,e}{}^*(\beta)\mathbb{R}^{\mathscr{K}_0}_{11}(\beta)^*, \qquad \mu\in A_G, \tag{III.13}$$

Thereby it is obvious what combinations have to be taken in order to obtain the corresponding vectors $\mathbf{K}^{\mu v}_{a_0}, \mathbf{L}^{\mu v}_{a_0}, \mathbf{M}^{\mu v}_{a_0}$, and $\mathbf{N}^{\mu v}_{a_0}$, Predictions concerning appropriated left coset representatives (σ_v, σ'_v) are obtained by similar arguments as in the previous cases. Finally, the remaining column vectors $\mathbf{P}^{\mu_1\mu_2;\bar{\mu}(i_v j_v)}_{a_0}$ are readily derived from Eq. (III.13), if the wave vector selection rules $\pm \mathbf{q}(\sigma_v) \pm \mathbf{q}'(\sigma'_v) = \mathbf{q}_0 + \mathbf{Q}[\cdots]$ are replaced by $\pm \mathbf{q}(\sigma_v) \pm \mathbf{q}'(\sigma'_v) = -\mathbf{q}_0 - \mathbf{Q}[\cdots]$, and $\{B^{\mathbf{q}_0}_{e,e}(\beta)\mathbb{R}^{\mathscr{K}_0}_{11}(\beta)\}^*$ by their complex conjugate values.

Now we are in the position to give closed expressions for the unitary submatrices $F^{\mu(\mu_1\mu_2)}$ of F, which are the nontrivial constituents of B and C. According to Eq. (II.18) and (II.20) we have to take either

$$\{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta)+\mathbf{t})U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_1,\sigma_2}$$
$$= \Delta^{\mathbf{q}}(\sigma_1,\beta\sigma_2)e^{-i\mathbf{q}(\sigma_1)\cdot\mathbf{t}}B^{\mathbf{q}}_{\sigma_1,\sigma_2}(\beta)\mathbb{R}^{\mathscr{K}}(\sigma_1^{-1}\beta\sigma_2)U^{\mathscr{K}}$$
$$= e^{-i\mathbf{q}(\sigma_1)\cdot\mathbf{t}}\{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta))U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_1,\sigma_2} \tag{III.14}$$

into account or

$$\{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta)+\mathbf{t})U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_1,\sigma_2}$$
$$= \Delta^{\mathbf{q}}(\sigma_1,\beta I\sigma_2)e^{-i\mathbf{q}(\sigma_1)\cdot\mathbf{t}+i\mathbf{q}(\sigma_2)\cdot\mathbf{t}(\sigma_2,I)}B^{\mathbf{q}}_{\sigma_1,I\sigma_2}(\beta)$$
$$\times \mathbb{R}^{\mathscr{K}}(\sigma_1^{-1}\beta I\sigma_2)U^{\mathscr{K}}$$
$$= e^{-i\mathbf{q}(\sigma_1)\cdot\mathbf{t}}\{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta))U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_1,\sigma_2}, \tag{III.15}$$

where we have partly used matrix notation for $D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta)+\mathbf{t})U^{(\mathscr{K},\mathbf{q})}$. Thereby we have to remember that Eq. (II.18) and therefore Eq. (III.14) are rather rare, whereas Eq. (II.20) and therefore Eq. (III.15) will be realized for nearly all cases. Nevertheless, both situations, depending on the case in question, are possible and are therefore considered.

Case $A \otimes B$:C: $A,B,C\in A_{\rm I}\cup A_{\rm II}$. Corresponding to the general formulas

$$F^{\mu(\mu_1\mu_2)}_{wv} = \left\langle \mathbf{M}^{\mu w}_{a_0}, U^{\mu_1}\otimes U^{\mu_2}\left\{\sum_l U^{\mu}_{a_0 l}\mathbf{M}^{\mu v}_l\right\}^*\right\rangle,$$
$$\text{and } \mu_1,\mu_2,\mu\in A_{\rm I}\cup A_{\rm II} \tag{III.16}$$

which cover the cases (III.23) and (III.47) of Ref. 1, Eqs. (II.24) and (II.61) or Ref. 2, and Eqs. (II.24) and (II.62) of Ref. 4, we have to compute

$$F^{\mu(\mu_1\mu_2)}_{wv} = \|\mathbf{B}^{\mu_1\mu_2;\mu(i_w j_w)}_{a_0}\|^{-1}\;\|\mathbf{B}^{\mu_1\mu_2;\mu(i_v j_v)}_{a_0}\|^{-1}$$
$$\times \delta_{\mathbf{q}(\sigma_w)+\mathbf{q}'(\sigma'_w),\mathbf{q}_0+\mathbf{Q}[\cdots]}\frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_\beta \{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta))U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_w c_w;\sigma_v c_v}$$
$$\times \{D^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta))U^{(\mathscr{K}',\mathbf{q}')}\}_{\sigma'_w c'_w;\sigma'_v c'_v}$$
$$\times \{D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta))U^{(\mathscr{K}_0,\mathbf{q}_0)}\}^*_{e1;e1} \tag{III.17}$$

since we assume that the corresponding CG coefficients for G can be expressed by Eq. (III.3). Obviously, Eq. (III.14) or (III.15) has to be taken into account. Equations (III.17) represent the formulas (III.25) and (III.49) of Ref. 1, Eqs. (II.31) and (II.67) of Ref. 2, and Eqs. (II.30) and (II.69) of Ref. 4. In calculating the matrix elements (III.17), it is very useful to utilize the property of the matrices $F^{\mu(\mu_1\mu_2)}$ to be either symmetric or antisymmetric [compare Eqs. (III.17) and (III.39) of Ref. 1, Eqs. (II.27) and (II.64) of Ref. 2, and Eqs. (II.27) and (II.65) or Ref. 4]. In this connection it is reasonable to make some remarks concerning the row and column indices of the matrices $F^{\mu(\mu_1\mu_2)}$. These indices are composed on the one hand by the left coset representatives $(\sigma_v,\sigma'_v)\in P(\mathbf{q},\mathbf{q}')$ [compare definition (III.1) of Ref. 9], and on the other hand by special column indices c_v and c'_v of the projective unirreps $\mathbb{R}^{\mathscr{K}}$ and $\mathbb{R}^{\mathscr{K}'}$. Although for many cases $|P(\mathbf{q},\mathbf{q}';\mathbf{q}_0)| = 1$ is realized, we cannot exclude the possibility $|P(\mathbf{q},\mathbf{q}';\mathbf{q}_0)| > 1$ (compare remarks in Refs. 9 and 14). For both cases the left coset representatives $(\sigma_v,\sigma'_v)\in P(\mathbf{q},\mathbf{q}')$ are linked by special formulas [see Eqs. (II.4), (II.12), (II.25), (II.35), (II.42), and (II.65). of Ref. 7] with the elements $(\sigma_a,\sigma'_a)\in P(\mathbf{q},\mathbf{q}';\mathbf{q}_0)$, whose definitions are given through Eqs. (III.75) and (III.76) of Ref. 9. For the second case we may expect a block structure of $F^{\mu(\mu_1\mu_2)}$ originating from $(\sigma_a,\sigma'_a)\in P(\mathbf{q},\mathbf{q}';\mathbf{q}_0)$, respectively, there arises the question whether the matrix elements of $F^{\mu(\mu_1\mu_2)}$ can be different from zero for different pairs $(\sigma_a,\sigma'_a)\in P(\mathbf{q},\mathbf{q}';\mathbf{q}_0)$. In order to be able to answer this question it is necessary to decide each case on its own merits. Provided such a matrix element would be different from zero, the corresponding CG coefficients for M would link CG coefficients for G which belong to different wave vector selection rules. Obviously, similar arguments will hold for all following cases.

Provided the matrix elements (III.17) have been calculated for the various cases, the corresponding CG coefficients for M are immediately obtained from the general formulas which have been derived in Refs. 1, 2, and 4. Apart from the first two cases, where one has to find solutions of $FB^* = B$ [see Eq. (III.20) of Ref. 1] and $FB^* = BG^T$ [see Eq. (III.44) of Ref. 1], we summarize our results as follows: (I ⊗ II : I) CG coefficients for M are given by Eqs. (II.38) and (II.39) of Ref. 2 by taking the definitions (II.11) of Ref. 2 into account. (I ⊗ II : II) CG coefficients follow from Eqs. (II.73) and (II.74) of Ref. 2 by using the corresponding definitions (II.47) of Ref. 2. (II ⊗ II : I) CG coefficients are obtained by means of Eqs. (II.37)–(II.40) of Ref. 4, where Eq. (II.12) of Ref. 4 has to be used. Finally, (II ⊗ II :II) CG coefficients are given by Eqs. (II.76) and (II.77) of Ref. 4, where the definitions (II.49) of Ref. 4 have to be taken into account. Besides this the nontrivial components of $\mathbf{Q}^{\mu v\cdots}_{a_0}$ are fixed through Eq. (III.3).

Case $A \otimes B$:III: $A,B\in A_{\rm I}\cup A_{\rm II}$. Due to the general formulas

$$F^{\bar{\mu}(\mu_1\mu_2)}_{wv} = \langle \mathbf{M}^{\bar{\mu}w}_{a_0}, U^{\mu_1}\otimes U^{\mu_2}\mathbf{M}^{\mu v*}_{a_0}\rangle, \quad \mu_1,\mu_2\in A_{\rm I}\cup A_{\rm II} \tag{III.18}$$

which represent the cases (III.89) of Ref. 1, Eq. (II.110) of Ref. 2, and Eq. (II.112) of Ref. 4, we have to calculate for the present approach the following expressions:

$$F_{wv}^{\bar{\mu}(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}\|^{-1} \delta_{\mathbf{q}(\sigma_u) + \mathbf{q}'(\sigma'_u), -\mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_\beta \{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\tau(\beta))U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_u c_u;\sigma_v c_v} \{D^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta|\tau(\beta))U^{(\mathscr{K}',\mathbf{q}')}\}_{\sigma'_u c'_u;\sigma'_v c'_v} D_{e1;e1}^{(\mathscr{K}_0\mathbf{q}_0)\uparrow G}(\beta|\tau(\beta)), \tag{III.19}$$

where depending on the considered case either Eq. (III.14) or (III.15) has to be used. Equations (III.19) coincide with the formulas (III.92) of Ref. 1, Eq. (II.114) of Ref. 2, and Eq. (II.116) of Ref. 4. Concerning the row and column indices of $F^{\bar{\mu}(\mu_1\mu_2)}$, one has to be very careful, since they are not necessarily identical [compare remarks to Eqs. (III.8) and (III.10) of the present paper]. Presupposing the matrix elements (III.19) have been computed, the corresponding CG coefficients for M are obtained as follows: (I ⊗ I : III) CG coefficients are given by Eqs. (III.87) and (III.88) of Ref. 1. (I ⊗ II : III) CG coefficients are defined by Eqs. (II.115) and (II.116) of Ref. 2 where the definitions (II.87) and (II.88) of Ref. 2 have to be taken into account. Finally, (II ⊗ II : III) CG coefficients are immediately obtained from Eqs. (II.117) and (II.118) of Ref. 4 by using the definitions (II.89) and (II.90) of Ref. 4. Analogously to the previous cases, the nontrivial components of the vectors $\mathbf{Q}_{a_0}^{\mu v\cdots}$ and $\mathbf{Q}_{a_0}^{\bar{\mu}v\cdots}$ are given by Eqs. (III.3) and (III.5), respectively.

Case $A \otimes$ III : C $A, C \in A_{\rm I} \cup A_{\rm II}$. In this case it suffices to consider the formulas

$$F_{wv}^{\mu(\mu_1\mu_2)} = \left\langle \mathbf{M}_{a_0}^{\mu w}, U^{\mu_1} \otimes \mathbf{1}_{\mu_2} \left\{ \sum_l U_{a_0 l}^{\mu} \mathbf{N}_l^{\mu v} \right\}^* \right\rangle, \quad \mu, \mu_1 \in A_{\rm I} \cup A_{\rm II}, \mu_2 \in A_{\rm III} \tag{III.20}$$

which cover the cases (II.28) and (II.70) of Ref. 3 and Eqs. (II.30) and (II.78) of Ref. 5, since the remaining matrices $F^{\mu(\mu_1\bar{\mu}_2)}$ are obtained by symmetry relations of the kind $F^{\mu(\mu_1\mu_2)T} = \pm F^{\mu(\mu_1\bar{\mu}_2)}$ [compare Eqs. (II.33) and (II.75) of Ref. 3 and Eqs. (II.35) and (II.85) of Ref. 5] and $D^{\mu_2}(\theta^2) = \mathbf{1}_{\mu_2}$. Thus, we have to calculate the following matrix elements:

$$F_{wv}^{\mu(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \delta_{\mathbf{q}(\sigma_u) + \mathbf{q}'(\sigma'_u), \mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_\beta \{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\tau(\beta))U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_u c_u;\sigma_v c_v} D_{\sigma'_u c'_u;\sigma'_v c'_v}^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta|\tau(\beta)) \{D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta|\tau(\beta))U^{(\mathscr{K}_0,\mathbf{q}_0)}\}_{e1;e1}^*, \tag{III.21}$$

where for the unitary matrices $U^{(\cdot,\cdot)}$ either Eq. (III.14) or (III.15) has to be inserted. Equations (III.21) represent Eqs. (II.37) and (II.79) of Ref. 3 and Eq. (II.39) and (II.87) of Ref. 5. Provided the matrix elements (III.21) have been determined for the various cases, the corresponding CG coefficients for M are immediately obtainable as follows: (I ⊗ III : I) CG coefficients follow from Eq. (II.45) and (II.46) of Ref. 3 by using Eqs. (II.14) and (II.15) of Ref. 3. On the other hand, (I ⊗ III :II) CG coefficients are given by Eqs. (II.86) and (II.87) of Ref. 3, where the definitions (II.55) and (II.56) have to be used. (II ⊗ III :I) CG coefficients are defined by Eqs. (II.49)–(II.52), (II.14) and (II.15) of Ref. 5. Finally (II ⊗ III : II) CG coefficients follow from Eqs. (II.96)–(II.98), and (II.61), and (II.62) of Ref. 5. Thereby we have to note that the nontrivial components of the vectors $\mathbf{Q}_{a_0}^{\mu v\cdots}$ are fixed through Eqs. (III.11) and (III.12).

Case $A \otimes$ III :III $:A \in A_{\rm I} \cup A_{\rm II}$. In this case we have to consider

$$F_{wv}^{\bar{\mu}(\mu_1\mu_2)} = \langle \mathbf{M}_{a_0}^{\bar{\mu}w}, U^{\mu_1} \otimes \mathbf{1}_{\mu_2} \mathbf{N}_{a_0}^{\mu v *} \rangle, \tag{III.22}$$

$$F_{wv}^{\bar{\mu}(\mu_1\bar{\mu}_2)} = \langle \mathbf{N}_{a_0}^{\bar{\mu}w}, U^{\mu_1} \otimes \mathbf{1}_{\mu_2} \mathbf{M}_{a_0}^{\mu v *} \rangle, \quad \mu_1 \in A_{\rm I} \cup A_{\rm II}, \tag{III.23}$$

which coincide with Eqs. (II.127) and (II.128) of Ref. 3 and Eqs. (III.139) and (III.140) of Ref. 5. Due to our approach it is only necessary to compute

$$F_{wv}^{\bar{\mu}(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \delta_{\mathbf{q}(\sigma_u) + \mathbf{q}'(\sigma'_u), -\mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_\beta \{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\tau(\beta))U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_u c_u;\sigma_v c_v} D_{\sigma'_u c'_u;\sigma'_v c'_v}^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta|\tau(\beta)) D_{e1;e1}^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta|\tau(\beta)), \tag{III.24}$$

$$F_{wv}^{\bar{\mu}(\mu_1\bar{\mu}_2)} = \|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\bar{\mu}(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}\|^{-1} \delta_{\mathbf{q}(\sigma_u) - \mathbf{q}'(\sigma'_u), -\mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{K}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_\beta \{D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\tau(\beta))U^{(\mathscr{K},\mathbf{q})}\}_{\sigma_u c_u;\sigma_v c_v} D_{\sigma'_u c'_u;\sigma'_v c'_v}^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta|\tau(\beta))^* D_{e1;e1}^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta|\tau(\beta)), \tag{III.25}$$

which are in accord to Eqs. (II.135) and (II.136) of Ref. 3, and Eqs. (II.147) and (II.148) of Ref. 5. In this connection we have to note that the definition of row and column indices of the matrix $F^{\bar{\mu}(\mu_1\mu_2)}$ and $F^{\bar{\mu}(\mu_1\bar{\mu}_2)}$ should not be confused. Provided the calculation of the matrix elements (III.24) and (III.25) has been carried out, the corresponding CG coefficients for M are obtained as follows: (I ⊗ III :III) CG coefficients follow from Eqs. (II.137)–(II.139) and (II.102)–(II.105) of Ref. 3, whereas (II ⊗ III :III) CG coefficients are defined by Eqs. (II.149)–(II.151) and (II.112)–(II.115) of Ref. 5. For both cases the nontrivial components of the vectors $\mathbf{Q}_{a_0}^{\mu v\cdots}$ and $\mathbf{Q}_{a_0}^{\bar{\mu}v\cdots}$ are given by Eqs. (III.11) and (III.12) and analogously defined expressions.

Case III ⊗ III : C: $C \in A_{\rm I} \cup A_{\rm II}$. In this case it suffices to consider

$$F_{wv}^{\mu(\mu_1\mu_2)} = \left\langle \mathbf{K}_{a_0}^{\mu w}, \left\{ \sum_l U_{a_0 l}^{\mu} \mathbf{N}_l^{\mu v} \right\}^* \right\rangle, \tag{III.26}$$

$$F_{wv}^{\mu(\mu_1\bar{\mu}_2)} = \left\langle \mathbf{L}_{a_0}^{\mu w}, \left\{ \sum_l U_{a_0 l}^{\mu} \mathbf{M}_l^{\mu v} \right\}^* \right\rangle \quad \mu \in A_{\mathrm{I}} \cup A_{\mathrm{II}} \tag{III.27}$$

since the remaining matrices $F^{\mu(\bar{\mu}_1\bar{\mu}_2)}$ and $F^{\mu(\bar{\mu}_1\mu_2)}$ are obtained by relations of the kind $F^{\mu(\mu_1\mu_2)T} = \pm F^{\mu(\bar{\mu}_1\bar{\mu}_2)}$ [see Eqs. (II.42) and (II.97) of Ref. 6] and $F^{\mu(\mu_1\bar{\mu}_2)T} = \pm F^{\mu(\bar{\mu}_1\mu_2)}$ [see Eqs. (II.43) and (II.98) of Ref. 6]. Obviously Eq. (III.26) coincides with Eqs. (II.32) and (II.87) of Ref. 6 and Eq. (III.27) is identical with Eqs. (II.33) and (II.88) of Ref. 6. Provided the corresponding column vectors can be expressed by Eq. (III.13), we obtain

$$F_{wv}^{\mu(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\bar{\mu}_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \delta_{\mathbf{q}(\sigma_u) + \mathbf{q}'(\sigma'_u), -\mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{H}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_{\beta} D^{(\mathscr{H},\mathbf{q})\uparrow G}_{\sigma_u c_u;\sigma_v c_v}(\beta|\tau(\beta)) D^{(\mathscr{H}',\mathbf{q}')\uparrow G}_{\sigma'_u c'_u;\sigma'_v c'_v}(\beta|\tau(\beta)) \{D^{(\mathscr{H}_0,\mathbf{q}_0)\uparrow G}(\beta|\tau(\beta)) U^{(\mathscr{H}_0,\mathbf{q}_0)}\}^*_{e1;e1}, \tag{III.28}$$

$$F_{wv}^{\mu(\mu_1\bar{\mu}_2)} = \|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\bar{\mu}_1\mu_2;\mu(i_v j_v)}\|^{-1} \delta_{\mathbf{q}(\sigma_u) - \mathbf{q}'(\sigma'_u), \mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{H}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_{\substack{\beta \\ P}} D^{(\mathscr{H},\mathbf{q})\uparrow G}_{\sigma_u c_u;\sigma_v c_v}(\beta|\tau(\beta)) D^{(\mathscr{H}',\mathbf{q}')\uparrow G*}_{\sigma'_u c'_u;\sigma'_v c'_v}(\beta|\tau(\beta)) \{D^{(\mathscr{H}_0,\mathbf{q}_0)\uparrow G}(\beta|\tau(\beta)) U^{(\mathscr{H}_0,\mathbf{q}_0)}\}^*_{e1;e1}, \tag{III.29}$$

which are just the formulas (II.50) and (II.105) and (II.51) and (II.106), respectively, of Ref. 6. Hence, the corresponding (III ⊗ III : I) CG coefficients for M are given by Eqs. (II.60)–(II.63) and (II.17)–(II.20) of Ref. 6, whereas the (III ⊗ III : II) CG coefficients follow from Eqs. (II.115)–(II.117) and (II.71)–(II.74) of Ref. 6. For both cases the nontrivial components of the vectors $\mathbf{Q}_{a_0}^{\mu v \cdots}$ are given by the corresponding vectors (III.13).

Case III ⊗ III : III: In the last case we are forced to compute

$$F_{wv}^{\bar{\mu}(\mu_1\mu_2)} = \langle \mathbf{K}_{a_0}^{\bar{\mu}w}, \mathbf{N}_{a_0}^{\mu v *} \rangle, \tag{III.30}$$
$$F_{wv}^{\bar{\mu}(\mu_1\bar{\mu}_2)} = \langle \mathbf{L}_{a_0}^{\bar{\mu}w}, \mathbf{M}_{a_0}^{\mu v *} \rangle, \tag{III.31}$$
$$F_{wv}^{\bar{\mu}(\bar{\mu}_1\mu_2)} = \langle \mathbf{M}_{a_0}^{\bar{\mu}w}, \mathbf{L}_{a_0}^{\mu v *} \rangle, \tag{III.32}$$
$$F_{wv}^{\bar{\mu}(\bar{\mu}_1\bar{\mu}_2)} = \langle \mathbf{N}_{a_0}^{\bar{\mu}w}, \mathbf{K}_{a_0}^{\mu v *} \rangle, \tag{III.33}$$

which are identical with Eqs. (II.157)–(II.160) of Ref. 6. Presupposing the column vectors $\mathbf{K}_{a_0}^{\mu v}, \mathbf{L}_{a_0}^{\mu v}, \mathbf{M}_{a_0}^{\mu v}$, and $\mathbf{N}_{a_0}^{\mu v}$ and $\mathbf{K}_{a_0}^{\bar{\mu} v}, \mathbf{L}_{a_0}^{\bar{\mu} v}, \mathbf{M}_{a_0}^{\bar{\mu} v}$, and $\mathbf{N}_{a_0}^{\bar{\mu} v}$ can be expressed by Eq. (III.13), and analogously defined vectors, we have to calculate

$$F_{wv}^{\bar{\mu}(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\bar{\mu}_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \delta_{\mathbf{q}(\sigma_u) + \mathbf{q}'(\sigma'_u), -\mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{H}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_{\beta} D^{(\mathscr{H},\mathbf{q})\uparrow G}_{\sigma_u c_u;\sigma_v c_v}(\beta|\tau(\beta)) D^{(\mathscr{H}',\mathbf{q}')\uparrow G}_{\sigma'_u c'_u;\sigma'_v c'_v}(\beta|\tau(\beta)) D^{(\mathscr{H}_0,\mathbf{q}_0)\uparrow G}_{e1;e1}(\beta|\tau(\beta)), \tag{III.34}$$

$$F_{wv}^{\bar{\mu}(\mu_1\bar{\mu}_2)} = \|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\bar{\mu}(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\bar{\mu}_1\mu_2;\mu(i_v j_v)}\|^{-1} \delta_{\mathbf{q}(\sigma_u) - \mathbf{q}'(\sigma'_u), -\mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{H}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_{\beta} D^{(\mathscr{H},\mathbf{q})\uparrow G}_{\sigma_u c_u;\sigma_v c_v}(\beta|\tau(\beta)) D^{(\mathscr{H}',\mathbf{q}')\uparrow G*}_{\sigma'_u c'_u;\sigma'_v c'_v}(\beta|\tau(\beta)) D^{(\mathscr{H}_0,\mathbf{q}_0)\uparrow G}_{e1;e1}(\beta|\tau(\beta)), \tag{III.35}$$

$$F_{wv}^{\bar{\mu}(\bar{\mu}_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_u j_u)}\|^{-1} \|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \delta_{-\mathbf{q}(\sigma_u) + \mathbf{q}'(\sigma'_u), -\mathbf{q}_0 + \mathbf{Q}[\cdots]} \frac{n_{\mathscr{H}_0}}{|P^{\mathbf{q}_0}|}$$
$$\times \sum_{\beta} D^{(\mathscr{H},\mathbf{q})\uparrow G*}_{\sigma_u c_u;\sigma_v c_v}(\beta|\tau(\beta)) D^{(\mathscr{H}',\mathbf{q}')\uparrow G}_{\sigma'_u c'_u;\sigma'_v c'_v}(\beta|\tau(\beta)) D^{(\mathscr{H}_0,\mathbf{q}_0)\uparrow G}_{e1;e1}(\beta|\tau(\beta)), \tag{III.36}$$

which coincide with Eqs. (II.170)–(II.172) of Ref. 6. Note that for this case $F_{wv}^{\bar{\mu}(\bar{\mu}_1\bar{\mu}_2)} = \delta_{wv}$ [compare Eq. (II.173) of Ref. 6] is valid. Furthermore, one has to note the different definition of row and column indices of $F^{\bar{\mu}(\cdots)}$ and that the dimensions of the four matrices $F^{\bar{\mu}(\cdots)}$ will be in general not equal. Provided the matrix elements (III.34)–(III.36) have been computed, the corresponding (III ⊗ III : III) CG coefficients for M are given by Eqs. (II.174)–(II.178) and (II.129)–(II.136) of Ref. 6.

SUMMARY

The aim of this paper was to specialize the general results of Refs. 1–6 to type II Shubnikov space groups, which contain the inversion as the point group operation. Assuming that convenient CG coefficients for the unitary subgroup can be determined by means of the method given in Ref. 7 (what is really true for nearly all cases), we succeeded in deriving simple formulas for unitary transformations which link CG coefficients for type II Shubnikov space groups with such ones for the unitary subgroup. The crucial point of the present method is that the dimensions of these unitary transformations are essentially smaller than that of the considered Kronecker products.

In particular when calculating the matrix elements (II.17), (II.19), (II.21), (II.24), (II.25), (II.28), (II.29), (II.34)–(II.36) for the various cases, it is useful to use directly the special solutions (II.4), (II.12), (II.25), (II.35), (II.42), (II.65) of the multiplicity problem, which have been discussed extensively in Ref. 7. However, we must be aware that for cases where not all multiplicity indices can be traced back to special column indices of the considered Kronecker product, one is forced to determine the corresponding CG coefficients for the unitary subgroup by projection techniques (see Ref. 13) and to compute with them the general expressions (III.16), (III.18), (III.20), (III.23), (III.26), (III.27), (III.30)–(III.33). Concluding this paper we mention that the utility of the present method will be demonstrated in the following paper on the type II Shubnikov space group $Pn3'n$ by considering a series of different examples.

[1]R. Dirl, J. Math. Phys. **21**, 961 (1980).
[2]R. Dirl, J. Math. Phys. **21**, 968 (1980).
[3]R. Dirl, J. Math. Phys. **21**, 975 (1980).
[4]R. Dirl, J. Math. Phys. **21**, 983 (1980).
[5]R. Dirl, J. Math. Phys. **21**, 989 (1980).
[6]R. Dirl, J. Math. Phys. **21**, 997 (1980).
[7]R. Dirl, J. Math. Phys. **20**, 671 (1979).
[8]C.J. Bradley, A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[9]R. Dirl, J. Math. Phys. **20**, 664 (1979).
[10]R. Dirl, J. Math. Phys. **20**, 1566 (1979).
[11]R. Dirl, J. Math. Phys. **20**, 1570 (1979).
[12]R. Dirl, J. Math. Phys. **20**, 1562 (1979).
[13]R. Dirl, J. Math. Phys. **20**, 659 (1979).
[14]D.B. Litvin and J. Zak, J. Math. Phys. **9**, 212 (1968).

Selection rules for *Pn* 3′*n*

R. Dirl

Institut für Theoretische Physik, TU Wien, A-1040 Wien, Karlsplatz 18; Austria

(Received 18 July 1979; accepted for publication 26 October 1979)

Clebsch–Gordan coefficients for the type II Shubnikov space group *Pn* 3′*n* are calculated in terms of such coefficients for the unitary subgroup *Pn* 3*n*.

I. INTRODUCTION

In the previous paper we have specialized our general approach to type II Shubnikov space groups. In particular, we were able to calculate by simple formulas unitary matrices which link CG coefficients for type II Shubnikov space groups with convenient ones for the unitary subgroup. This approach, however, requires a special solution of the multiplicity problem. In this paper we shall give the results of calculation of CG coefficients for the type II Shubnikov space group *Pn*3′*n* by considering a series of examples. The only demerit of *Pn*3′*n* is that type II co-unirreps do not occur and that (III ⊗ III : III) CG coefficients are impossible. The organization of the material of this article is as follows: In Sec. I we recall briefly some definitions and notations concerning the unitary subgroup *Pn*3*n* with its unirreps and list a series of examples which shall be considered in the following. Convenient CG coefficients for *Pn*3*n* are determined for these examples in Sec. II, at which for each case the multiplicity indices can be traced back to special column indices of the corresponding Kronecker products. In the last section we compute for our examples unitary matrices which link CG coefficients for *Pn*3′*n* with them for *Pn*3*n*. Finally, CG coefficients for *Pn*3′*n* are listed in full detail for the considered examples.

I. DISCUSSION OF VARIOUS EXAMPLES

Starting from the nonsymmorphic space group *Pn*3*n*, the definition of the corresponding type II Shubnikov space group *Pn*3′*n* is obvious. Furthermore, throughout this paper we use the same definitions and notations concerning *Pn*3*n* as was introduced in Refs. 1 and 2. For convenience some of them are recalled. Nontrivial lattice translations $\boldsymbol{\tau}(\alpha)$; $\alpha \in \mathscr{O}_h = \mathscr{O} \times \{e, I\}$ are defined by

$$\boldsymbol{\tau}(n) = \mathbf{0}\,, \tag{I.1}$$

$$\boldsymbol{\tau}(In) = (1/2,\ 1/2,\ 1/2), \quad \text{for all } n \in \mathscr{O}\,, \tag{I.2}$$

where the group element I denotes the inversion and where for the sake of simplicity the lattice constant is chosen as 1. The fundamental domain ΔBZ (being identical with the basis domain) for *Pn*3*n* is defined by Eq. (I.3) of Ref. 1. The orthogonal matrices $D(\alpha)$; $\alpha \in \mathscr{O}_h$ are readily obtainable from Table 1.4 of Ref. 3. Complete sets of projective unirreps together with their factor systems of the little co-groups $P^{\mathbf{q}} \simeq G^{\mathbf{q}}/T$ are listed in full detail in Ref. 4 for all points of the "surface" of ΔBZ and for some $\mathbf{q}$'s lying inside of ΔBZ. Finally, special sets $P : P^{\mathbf{q}}$ of left coset representatives will be used within this paper (compare Sec. I of Ref. 5).

In the following we list a series of examples. Thereby the following quantities are written down explicitly: Starting from a $\mathbf{q} \in \Delta BZ$, we specify the corresponding group of the $\mathbf{q}$ vector, namely, $P^{\mathbf{q}}$, an appropriate set of left coset representatives $P : P^{\mathbf{q}}$, by taking Eq. (I.20) of Ref. 5 into account, if $I \notin P^{\mathbf{q}}$. Furthermore, those projective unirreps of $P^{\mathbf{q}}$ with their corresponding unitary matrices $U^{(\mathscr{K},\mathbf{q})}$ (which have been calculated in Refs. 2 and 6) are listed which shall be considered. Concerning the equivalence classes $\mathscr{K} \in A_{P^{\mathbf{q}}(S^{\mathbf{q}})}$, we use throughout this paper the same notation as in Refs. 1 and 2. However, for each case where

$$P^{\mathbf{q}} = \{e\} \Longleftrightarrow P : P^{\mathbf{q}} = \mathscr{O}_h \Rightarrow I \notin P^{\mathbf{q}}\,, \tag{I.3}$$

the trivial unirrep is not written down. Due to Eq. (II.7) of Ref. 6 the corresponding co-unirrep of *Pn*3′*n* must be of type I. This can be seen from

$$U^{(0,\mathbf{q})}_{\sigma,\sigma'} = \delta_{\sigma',I\sigma}\, e^{i\mathbf{q}(\sigma)\cdot\mathbf{t}(\sigma,I)}, \quad \sigma,\sigma' \in \mathscr{O}_h \tag{I.4}$$

since $I \in P = \mathscr{O}_h$ and Eq. (II.22) of Ref. 7 gives $U^{(0,\mathbf{q})}\, U^{(0,\mathbf{q})*} = \mathbb{1}_{(0,\mathbf{q}) \uparrow G}$. Finally, the sets $P(\mathbf{q}, \mathbf{q}', \mathbf{q}_0)$ are specified whose definition was stated in Ref. 8 [see Eqs. (III.75) and (III.76) of Ref. 8].

Example 1: (i) I ⊗ I : I, (ii) I ⊗ I : III

$$\mathbf{q} = \pi(x, y, z) \in \Delta BZ, \quad 1 > y > x > z > 0\,, \tag{I.5}$$

$$\mathbf{q}' = \pi(1 - x,\ 1 - z,\ 1 - y) \in \Delta BZ\,, \tag{I.6}$$

$$\mathbf{q} + \mathbf{q}'(\sigma_{df}) = \mathbf{q}_0 = \pi(1,1,1) = \mathbf{q}_R\,, \tag{I.7}$$

$$P^{\mathbf{q}_0} = \mathscr{O}_h \Longleftrightarrow P : P^{\mathbf{q}_0} = \{e\} \Rightarrow I \in P^{\mathbf{q}_0}\,, \tag{I.8}$$

$$D(\sigma_{df}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \tag{I.9}$$

$$P(\mathbf{q}, \mathbf{q}', \mathbf{q}_0) = \{(e, \sigma_{df})\}\,; \tag{I.10}$$

(i) $\mathbb{R}^{\mathscr{K}_0}$: $\mathscr{K}_0 = (\mu = 0) \uparrow \mathscr{O}_h = \{(\mu = 0) \uparrow \mathscr{O}_h\}^*$,

$$C_{2x} \rightarrow \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

$$C_{2z} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

$$C_{31}^{+} \rightarrow \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

$$C_{2d} \rightarrow \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix},$$

$$I \rightarrow \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad (I.11)$$

$$U^{(\mathscr{H}_0,\mathbf{q}_0)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} = U^{\mathscr{H}_0}; \qquad (I.12)$$

(ii) $\mathbb{R}^{\mathscr{H}_0}$: $\mathscr{H}_0 = (0,\mu = 2) \uparrow \mathscr{O}_h = \{(1,\mu = 2) \uparrow \mathscr{O}_h\}^*$,

$$C_{2x} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, C_{2z} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$C_{31}^{+} \rightarrow \begin{bmatrix} \omega & 0 \\ 0 & \omega^2 \end{bmatrix} \quad (\omega = e^{-i2\pi/3}),$$

$$C_{2d} \rightarrow \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, I \rightarrow \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \qquad (I.13)$$

$$U^{(\mathscr{H}_0,\mathbf{q}_0)} = U^{\mathscr{H}_0} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \qquad (I.14)$$

Equations (I.12) and (I.14) are identical with Eqs. (III.14) and (III.12), respectively, of Ref. 2.

Example 2: I $\otimes$ III : I

$$\mathbf{q} = \pi\,(x,y,z) \in \Delta BZ, \quad 1 > y > x > z > 0\,, \qquad (I.15)$$

$$\mathbf{q}' = \pi\,(0,1,0) = \mathbf{q}_x\,, \qquad (I.16)$$

$$P^{\mathbf{q}'} = \{e,I\} \times C_{4v} \Rightarrow P : P^{\mathbf{q}'} = \{e, C_{31}^{+}, C_{31}^{-}\} \Rightarrow I \in P^{\mathbf{q}'}; \qquad (I.17)$$

$$\mathbb{R}^{\mathscr{H}'} : \mathscr{H}' = (0,5) \uparrow P^x = \{(1,5) \uparrow P^x\}^*,$$

$$C_{4y}^{+} \rightarrow \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}, \sigma_x \rightarrow \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, I \rightarrow \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \qquad (I.18)$$

$$U_{\sigma',\sigma}^{(\mathscr{H}',\mathbf{q}')} = \delta_{\sigma',\sigma} U^{\mathscr{H}'}, U^{\mathscr{H}'} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \sigma,\sigma' \in P : P^{\mathbf{q}'}, \qquad (I.19)$$

$$\mathbf{q}(S_{63}^{+}) + \mathbf{q}' = \mathbf{q}_0 = \pi(y, 1 - z, x) \in \Delta BZ$$
$$(1 > 1 - z > y > x > 0)\,, \qquad (I.20)$$

$$D(S_{63}^{+}) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad (I.21)$$

$$P(\mathbf{q},\mathbf{q}',\mathbf{q}_0) = \{(S_{63}^{+}, e)\}\,. \qquad (I.22)$$

Equation (I.19) is identical with Eq. (IV.2) of Ref. 2, where for $U^{\mathscr{H}'}$ Eq. (III.7) of Ref. 2 has to be inserted.

Example 3: (i) I $\otimes$ III : I, (ii) I $\otimes$ III : III

$$\mathbf{q} = \pi(1,1,0) = \mathbf{q}_M \qquad (I.23)$$

$$P^{\mathbf{q}} = D_{4h} = \{e,\sigma_x\} \circledS (\{e,I\} \times C_{2v}) \Rightarrow P : P^{\mathbf{q}}$$
$$= \{e, \sigma_{de}, \sigma_{df}\} \Rightarrow I \in P^{\mathbf{q}}\,, \qquad (I.24)$$

$$\mathbb{R}^{\mathscr{H}} : \mathscr{H} = (0,0,0) \uparrow P^M = \{(0,0,0) \uparrow P^M\}^*,$$

$$C_{2a} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_z \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$I \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \sigma_x \rightarrow \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad (I.25)$$

$$U_{\sigma',\sigma}^{(\mathscr{H},\mathbf{q})} = \delta_{\sigma',\sigma} U^{\mathscr{H}}, \; U^{\mathscr{H}} = \mathbb{1}_{\mathscr{H}}, \; \sigma,\sigma' \in P : P^{\mathbf{q}}, \qquad (I.26)$$

$$\mathbf{q}' = \pi(0,1,0)\,, \qquad (I.27)$$

$$\mathbb{R}^{\mathscr{H}'} : \mathscr{H}' = (0,5) \uparrow P^x = \{(1,5) \uparrow P^x\}^*, \qquad (I.28)$$

$$\mathbf{q} + \mathbf{q}'(C_{31}^{-}) = \mathbf{q}_0 + \mathbf{Q}[\ldots] \doteq \mathbf{q}_0 = \pi(0,1,0)\,, \qquad (I.29)$$

$$D(C_{31}^{-}) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad (I.30)$$

$$P(\mathbf{q},\mathbf{q}',\mathbf{q}_0) = \{(e, C_{31}^{-})\}\,; \qquad (I.31)$$

(i) $\mathbb{R}^{\mathscr{H}_0}$: $\mathscr{H}_0 = (\mu = 0) \uparrow P^x = \{(\mu = 0) \uparrow P^x\}^*$,

$$C_{4y}^{+} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \sigma_x \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, I \rightarrow \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad (I.32)$$

$$U_{\sigma',\sigma}^{(\mathscr{H}_0,\mathbf{q}_0)} = \delta_{\sigma',\sigma} U^{\mathscr{H}_0}, \; U^{\mathscr{H}_0} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma,\sigma' \in P : P^{\mathbf{q}_0}; \qquad (I.33)$$

(ii) $\mathbb{R}^{\mathscr{H}_0}$: $\mathscr{H}_0 = (0,5) \uparrow P^x = \{(1,5) \uparrow P^x\}^*$. $\qquad$ (I.34)

Equation (I.26), to be more precise $U^{\mathscr{H}} = \mathbb{1}_{\mathscr{H}}$, can be readily verified by means of Eq. (II.36) of Ref. 6, since the factor system $S^{\mathbf{q}_0}$ reduces to one for the generating elements. Equation (I.33) is identical with Eq. (III.8) of Ref. 2.

Example 4: III $\otimes$ III : I

$$\mathbf{q} = \mathbf{q}' = \pi\,(0,1,0)\,, \qquad (I.35)$$

$$\mathscr{H} = \mathscr{H}' = (0,5) \uparrow P^x = \{(1,5) \uparrow P^x\}^*, \qquad (I.36)$$

$$\mathbf{q} + \mathbf{q}'(C_{31}^{-}) = \mathbf{q}_0 = \pi\,(1,1,0) = \mathbf{q}_M\,, \qquad (I.37)$$

$$\mathscr{H}_0 = (0,0,0) \uparrow P^M = \{(0,0,0) \uparrow P^M\}^*. \qquad (I.38)$$

II. CG COEFFICIENTS FOR *Pn*3*n*

The first step of the present approach is concerned with the task of determining convenient CG coefficients for *Pn*3*n*, whose corresponding multiplicity indices are identifiable with special column indices of the considered Kronecker products. Thus, the first problem will be whether for our examples the multiplicity problem can be solved indeed in this special manner. Provided this can be done, the corresponding columns of the CG matrices in question for $Pn3n$ are readily obtained from the general formulas given in Ref. 9. Of course, for those cases which have been already solved in Ref. 1 we only summarize our results.

Example 1 (i): This example has been extensively discussed in Ref. 1. Thereby we obtained for the multiplicity indices

$$v \longleftrightarrow (x_v;x_v\sigma_{df}), \quad v=1,2,...,6, \tag{II.1}$$

where the group elements $x_v \in \mathscr{O}_h$ are fixed through

$$\begin{aligned} x_1 &= e, \\ x_2 &= C_{31}^+, \\ x_3 &= I, \\ x_4 &= C_{31}^+ I, \\ x_5 &= C_{2d}C_{31}^+, \\ x_6 &= C_{2d}C_{31}^+ I. \end{aligned} \tag{II.2}$$

Inserting the special values (II.1) and (II.2) into Eq. (II.18) of Ref. 9 and using the abbreviated notation (as in the previous paper) for the CG coefficients, we obtain

$$\begin{aligned} \{\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}\}_{ij} &= \{\mathbf{M}_{a_0}^{\mu v}\}_{ij} \\ &= \frac{1}{\sqrt{8}}\delta_{\sigma,\sigma'\sigma_{df}} B_{e,e}^{\mathbf{q}_0}{}^*(\sigma x_v^{-1}) B_{\sigma,x_v}^{\mathbf{q}}(\sigma x_v^{-1}) \\ &\quad \times B_{\sigma',x_v\sigma_{df}}^{\mathbf{q}'}(\sigma x_v^{-1}) \mathbb{R}_{11}^{\mathscr{H}_0}{}^*(\sigma x_v^{-1}), \\ \mu_1 &\longleftrightarrow (0,\mathbf{q}) \uparrow G, \\ \mu_2 &\longleftrightarrow (0,\mathbf{q}') \uparrow G, \\ \mu &\longleftrightarrow (\mathscr{H}_0,\mathbf{q}_0) \uparrow G, \\ i,j &\longleftrightarrow \sigma,\sigma' \in \mathscr{O}_h, \\ a_0 &\longleftrightarrow e,1, \\ v &\longleftrightarrow (i_v;j_v) \longleftrightarrow (x_v;x_v\sigma_{df}), \quad v=1,2,...,6, \end{aligned} \tag{II.3}$$

which represent convenient columns of the corresponding 48^2-dimensional CG matrix M for $Pn3n$.

Example 1 (ii): Before starting our considerations let us recall that we have to compute not only $\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}$ but also $\mathbf{M}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_v j_v)}$ (compare the previous paper). This forces us to investigate the general formula (II.17) of Ref. 9 for both cases separately. For the first case we obtain by using Eq. (I.13)

$$v \longleftrightarrow (x_v;x_v\sigma_{df}), \quad v=1,2, \tag{II.4}$$

where the group elements $x_v \in \mathscr{O}_h$ are given by

$$\begin{aligned} x_1 &= e, \\ x_2 &= C_{2d}. \end{aligned} \tag{II.5}$$

Obviously, the result (II.4) is in accordance with the corresponding multiplicity $m_{(0,\mathbf{q})(0,\mathbf{q}');(\mathscr{H}_0,\mathbf{q}_0)} = n_{\mathscr{H}_0} = 2$. For the second case, i.d., $\mu \rightarrow \bar{\mu}$ we can choose the same multiplicity indices (II.4) and (II.5), since the complex conjugation of $D^{(\mathscr{H}_0,\mathbf{q}_0)} \uparrow G$ is governed by Eq. (I.14), which is of type (II.18) of Ref. 7. However, the corresponding CG coefficients for $Pn3n$ are quite different, namely,

$$\begin{aligned} \{\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}\}_{ij} &= \{\mathbf{M}_{a_0}^{\mu v}\}_{ij} \\ &= \frac{1}{\sqrt{24}}\delta_{\sigma,\sigma'\sigma_{df}} \mathbf{B}_{e,e}^{\mathbf{q}_0}{}^*(\sigma x_v^{-1}) B_{\sigma,x_v}^{\mathbf{q}}(\sigma x_v^{-1}) \\ &\quad \times B_{\sigma',x_v\sigma_{df}}^{\mathbf{q}'}(\sigma x_v^{-1}) \mathbb{R}_{11}^{\mathscr{H}_0}{}^*(\sigma x_v^{-1}), \\ \mu &\longleftrightarrow (\mathscr{H}_0,\mathbf{q}_0) \uparrow G, \\ v &\longleftrightarrow (i_v;j_v) \longleftrightarrow (x_v;x_v\sigma_{df}), \quad v=1,2; \end{aligned} \tag{II.6}$$

$$\begin{aligned} \{\mathbf{M}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_v j_v)}\}_{ij} &= \{\mathbf{M}_{a_0}^{\bar{\mu} v}\}_{ij} \\ &= \frac{1}{\sqrt{24}}\delta_{\sigma,\sigma'\sigma_{df}} \mathbf{B}_{e,e}^{\mathbf{q}_0}(\sigma x_v^{-1}) B_{\sigma,x_v}^{\mathbf{q}}(\sigma x_v^{-1}) \\ &\quad \times B_{\sigma',x_v\sigma_{df}}^{\mathbf{q}'}(\sigma x_v^{-1}) \mathbb{R}_{11}^{\mathscr{H}_0}(\sigma x_v^{-1}), \\ \bar{\mu} &\longleftrightarrow \overline{(\mathscr{H}_0,\mathbf{q}_0) \uparrow G} = (\bar{\mathscr{H}}_0,\mathbf{q}_0) \uparrow G, \\ v &\longleftrightarrow (i_v,j_v) \longleftrightarrow (x_v;x_v\sigma_{df}), \quad v=1,2, \end{aligned} \tag{II.7}$$

since $\mathbb{R}^{\mathscr{H}_0}$ and $\mathbb{R}^{\mathscr{H}_0 *}$ are inequivalent projective unirreps of $P^{\mathbf{q}_0}$.

Example 2: In this case we are confronted with the task of determining the multiplicity indices v for $\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}$ and $\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}$. Due to the general solution (II.25) of Ref. 9 and Eq. (I.22) we find for both cases

$$v \longleftrightarrow (S_{63}^+;e,v), \quad v(=c_v')=1,2,...,n_{\mathscr{H}'}, n_{\mathscr{H}'}=2, \tag{II.8}$$

where, however, $c_v' = v$ for the second case is defined by

$$\mathbb{R}_{d'c_v'}^{\mathscr{H}'}{}^*(e) = \delta_{d',c_v'} = \mathbb{R}_{d'c_v'}^{\bar{\mathscr{H}}'}(e), \tag{II.9}$$

i.e., the column indices c_v' originate from $\mathbb{R}^{\mathscr{H}'*}$ and not from $\mathbb{R}^{\mathscr{H}'}$ as for the first case. Thereby we have to note that Eq. (I.19) is the reason that Eq. (II.9) has to be investigated for the second case, since Eq. (I.19) is of the type (II.18) of Ref. 7. Provided the complex conjugation of $D^{(\mathscr{H}',\mathbf{q}')\uparrow G}$ would be governed by unitary matrices of the type (II.20) of Ref. 7, the left coset representatives of $(\sigma;\sigma',v)$ would have to be replaced by $(\sigma;I\sigma',v)$. However this situation is not realized for $Pn3n$. Returning to our problem, we find for both cases the same vectors, namely,

$$\begin{aligned} \{\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}\}_{ij} &= \{\mathbf{M}_{a_0}^{\mu v}\}_{ij} = \{\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}\}_{ij} \\ &= \{\mathbf{N}_{a_0}^{\mu v}\}_{ij} = \delta_{\sigma,S_{63}^+}\delta_{\sigma',e}\delta_{d',v}, \\ \mu_1 &\longleftrightarrow (0,\mathbf{q}) \uparrow G, \\ \mu_2 &\longleftrightarrow (\mathscr{H}',\mathbf{q}') \uparrow G, \\ \mu &\longleftrightarrow (0,\mathbf{q}_0) \uparrow G, \\ i;j &\longleftrightarrow \sigma;\sigma',d', \\ a_0 &\longleftrightarrow e, \\ v &\longleftrightarrow (i_v;j_v) \longleftrightarrow (S_{63}^+;e,v), \quad v=1,2. \end{aligned} \tag{II.10}$$

Equation (II.10) is a direct consequence of Eq. (II.26) of Ref. 9.

Example 3 (i): Obviously, this example belongs to the most complicated ones, since $P_{\sigma,\sigma'}^{\mathbf{q}\mathbf{q}';\mathbf{q}_0}$ is nontrivial, namely,

$$\begin{aligned} P_{\sigma,\sigma'}^{\mathbf{q}\mathbf{q}';\mathbf{q}_0} &= P^{\mathbf{q}} \cap C_{31}^- P^{\mathbf{q}'} C_{31}^+ \cap P^{\mathbf{q}_0} \\ &= \{e,I\} \times \{e,\sigma_x\} \times \{e,\sigma_z\}. \end{aligned} \tag{II.11}$$

The first task which has to be solved is to determine the multiplicities. If taking among others Eq. (I.31) into account, it follows from Eq. (III.83) of Ref. 8 that

$$m_{\mu_1\mu_2;\mu} = 1\,, \tag{II.12}$$

at which the characters of the respective projective unireps are immediately obtained from Eqs. (I.25), (I.18), and (I.32). Since $(\mathscr{H}_0,\mathbf{q}_0)\uparrow G$ is of type I, it is only necessary to determine suitable multiplicity indices for $\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}$ and $\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}$ in terms of special column indices of the considered Kronecker products. When applying the general formulas (I.5) and (I.6) of Ref. 9 to the first case, it can be shown by simple calculations that

$$v \longleftrightarrow (e,1;C_{31}^-,1),\quad v=1\ (c_v = c_v' = 1) \tag{II.13}$$

can be chosen as multiplicity indices. For the second case, i.e., $\mu_2 \to \bar{\mu}_2$, which implies that the matrix elements $D^{(\mathscr{H}',\mathbf{q}')\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta)+\mathbf{t})$ have to be replaced by their complex conjugate values, we can use the same multiplicity indices (II.13). By similar arguments as previously, the left coset representative C_{31}^- remains because of Eq. (I.19) unchanged. Consequently, the corresponding columns of the CG matrices M and N are given by

$$\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)} = \mathbf{M}_{a_0}^{\mu v}\quad v=1\,, \tag{II.14}$$

$$\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)} = \mathbf{N}_{a_0}^{\mu v},\quad v=1\,, \tag{II.15}$$

where the special values $\mu_1 \longleftrightarrow (\mathscr{H},\mathbf{q})\uparrow G$; $\mu_2 \longleftrightarrow (\mathscr{H}',\mathbf{q}')\uparrow G$; $\mu \longleftrightarrow (\mathscr{H}_0,\mathbf{q}_0)\uparrow G$; $i \longleftrightarrow \sigma,c$; $j \longleftrightarrow \sigma',c'$; $a_0 \longleftrightarrow e,1$; and $v \longleftrightarrow (e,1;C_{31}^-,1)$ have to be inserted into Eqs. (III.11) and (III.12) of Ref. 7.

Example 3 (ii): As in the foregoing case one obtains for the multiplicities

$$m_{\mu_1\mu_2;\mu} = m_{\mu_1\bar{\mu}_2;\mu} = 1\,, \tag{II.16}$$

where $\mu \longleftrightarrow (\mathscr{H}_0,\mathbf{q}_0)\uparrow G$ is now of type III. Therefore, one has to determine for four different cases the corresponding multiplicity indices v. We obtain for each case ($\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}$, $\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}$, $\mathbf{M}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_v j_v)}$, and $\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\bar{\mu}(i_v j_v)}$) in principle the same multiplicity indices v, namely,

$$v \longleftrightarrow (e,1;C_{31}^-,2),\quad v=1\,, \tag{II.17}$$

where similar arguments concerning C_{31}^- hold as before. The corresponding columns of the CG matrices

$$\mathbf{M}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)} = \mathbf{M}_{a_0}^{\mu v},\quad v=1\,, \tag{II.18}$$

$$\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)} = \mathbf{N}_{a_0}^{\mu v}\quad v=1, \tag{II.19}$$

$$\mathbf{M}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_v j_v)} = \mathbf{M}_{a_0}^{\bar{\mu} v},\quad v=1\,, \tag{II.20}$$

$$\mathbf{N}_{a_0}^{\mu_1\bar{\mu}_2;\bar{\mu}(i_v j_v)} = \mathbf{N}_{a_0}^{\bar{\mu} v},\quad v=1\,, \tag{II.21}$$

are readily obtainable from Eqs. (III.11) and (III.12) of Ref. 7 by taking the special values for μ_1, μ_2, μ, i, j, a_0, and v into account.

Example 4: Analoguously to the previous cases we are because of

$$P_{\sigma,\sigma'}^{\mathbf{qq}';\mathbf{q}_0} = P^{\mathbf{q}}\cap C_{31}^-\ P^{\mathbf{q}'}C_{31}^+\cap P^{\mathbf{q}_0} = \{e,I\}\times\{e,\sigma_x\}\times\{e,\sigma_z\} \tag{II.22}$$

confronted with the most complicated situation. Simple calculations yield

$$m_{\mu_1\mu_2;\mu} = m_{\mu_1\bar{\mu}_2;\mu} = 1 \tag{II.23}$$

and lead us to the same solution for the four different multiplicity problems, namely,

$$v \longleftrightarrow (e,1;C_{31}^-,2),\quad v=1\,. \tag{II.24}$$

Hence, the corresponding columns of the CG matrices K, L, M, and N, i.e.,

$$\mathbf{K}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_v j_v)} = \mathbf{K}_{a_0}^{\mu v},\quad v=1\,, \tag{II.25}$$

$$\mathbf{L}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)} = \mathbf{L}_{a_0}^{\mu v},\quad v=1\,, \tag{II.26}$$

$$\mathbf{M}_{a_0}^{\bar{\mu}_1\mu_2;\mu(i_v j_v)} = \mathbf{M}_{a_0}^{\mu v},\quad v=1\,, \tag{II.27}$$

$$\mathbf{N}_{a_0}^{\bar{\mu}_1\bar{\mu}_2;\mu(i_v j_v)} = \mathbf{N}_{a_0}^{\mu v},\quad v=1\,, \tag{II.28}$$

are readily obtained from Eq. (III.13) of Ref. 7 by inserting the special values for μ_1, μ_2, μ, i, j, a_0, and v. Concluding this example we summarize without proof the following results:

$$m_{\mu_1\mu_2;(\mathscr{K},\mathbf{q}_M)} = m_{\mu_1\bar{\mu}_2;(\mathscr{K},\mathbf{q}_M)} = 1,\quad \text{for all } \mathscr{H}\in A_{P^M(S^{\mathbf{q}_M})}\,, \tag{II.29}$$

$$m_{\mu_1\mu_2;((\mu,\sigma),0)} = \begin{cases}\delta_{\mu,\sigma+1}, & \mu=0,1 \text{ and } \sigma=0,1,\\ 1, & \mu=2 \text{ and } \sigma=0,1,\quad (\text{mod } 2),\\ \delta_{\mu,\sigma+1}, & \mu=4,3 \text{ and } \sigma=0,1,\end{cases} \tag{II.30}$$

$$m_{\mu_1\bar{\mu}_2;((\mu,\sigma),0)} = \begin{cases}\delta_{\mu,\sigma}, & \mu=0,1 \text{ and } \sigma=0,1,\\ 1, & \mu=2 \text{ and } \sigma=0,1,\quad (\text{mod } 2),\\ \delta_{\mu,\sigma}, & \mu=4,3 \text{ and } \sigma=0,1,\end{cases} \tag{II.31}$$

whose knowledge is necessary in order to be able to compute the whole CG matrix W. Thereby we use the same notation as in Ref. 4 for equivalence classes of projective unirreps of $P^{\mathbf{q}_0}$ [see Eqs. (5.12)–(5.15) of Ref. 4], where the unirreps of $\{e,I\}$ are characterized by $\sigma(=0,1)$. A simple inspection of

$$4\{|P:P^{\mathbf{q}}|n_{\mathscr{K}}\}^2 = 144 = \sum_{\substack{\mathbf{q}_0\in\Delta BZ\\ \mathscr{K}_0\in A_{P^{\mathbf{q}_0}(S^{\mathbf{q}_0})}}} M_{\mu_1\mu_2;(\mathscr{K}_0,\mathbf{q}_0)}\,|P:P^{\mathbf{q}_0}|\,n_{\mathscr{K}_0} \tag{II.32}$$

show its correctness. In this connection we have to note that some of the multiplicities $m_{\mu_1\mu_2;((\mu,\sigma),0)}$ and $m_{\mu_1\bar{\mu}_2;((m,\sigma),0)}$ are not equal.

III. CG COEFFICIENTS FOR *Pn3'n*

Due to our approach the next step requires us to compute the unitary matrices $F^{\mu(\mu_1\mu_2)}$, which are constituents of those unitary matrices which link CG coefficients for $Pn3'n$ with convenient ones for $Pn3n$. In this connection we remark that the knowledge of the multiplicity indices $v \longleftrightarrow (\sigma_v, c_v; \sigma_v', c_v')$ suffices in order to be able to compute the matrices $F^{\mu(\mu_1\mu_2)}$, i.e., it is not necessary to calculate the corresponding columns $\mathbf{P}_{a_0}^{\mu v}$ (apart from their norm) of the considered CG matrices for $Pn3n$.

Example 1 (i): Specializing Eq. (III.17) of Ref. 7 to this example, we have to calculate

$$F_{wv}^{\mu(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_w j_w)}\|^{-1}\|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}\|^{-1} \times \tfrac{1}{8}\sum_{\beta}\mathbb{D}_{x_w;x_v}^{(0,\mathbf{q})\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta)\,|\,\theta)\,\mathbb{D}_{x_w\sigma_{d_j};x_v\sigma_{d_j}}^{(0,\mathbf{q}')\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta)\,|\,\theta) \times \mathbb{D}_{11}^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta\,|\,\boldsymbol{\tau}(\beta)\,|\,\theta)^*\,, \tag{III.1}$$

where we have partly used an abbreviated notation (as in Ref. 7), i.e., Eq. (II.16) of Ref. 7 and Eqs. (II.1) and (II.2), and superfluous indices are omitted. Inserting Eq. (I.4) twice and Eq. (I.12), it is readily verified that only the matrix elements F_{13}, F_{24}, and F_{56} can be different from zero, at which the remaining nonzero matrix elements of F are obtainable by utilizing the symmetry relation $F_{wv} = F_{vw}$ [see Eq. (III.17) of Ref. 10]. Simple calculations yield

$$\begin{aligned} F_{13} = F_{24} &= \exp\{2i\mathbf{q}'(C_{31}^{+}\sigma_{df})\cdot\boldsymbol{\tau}(I)\} \\ &= e^{2i\mathbf{q}'\cdot\boldsymbol{\tau}(I)} = e^{i\mathbf{q}'\cdot(\mathbf{t}_1+\mathbf{t}_2+\mathbf{t}_3)} = \omega\,, \end{aligned} \tag{III.2}$$

$$\{\mathbf{t}_i\}_j = \delta_{ij}, \quad i,j = 1,2,3\,, \tag{III.3}$$

$$\begin{aligned} F_{56} &= \exp\{-i\mathbf{q}(C_{2d}C_{31}^{+})\cdot[1 - D(C_{2d}C_{31}^{+})]\boldsymbol{\tau}(I)\} \\ &= \exp\{i\mathbf{q}\cdot\mathbf{t}_1 + i\mathbf{q}'(\sigma_{df})\cdot(\mathbf{t}_2+\mathbf{t}_3)\} = \omega'\,, \end{aligned} \tag{III.4}$$

where Table 1.4 of Ref. 3 and Eqs. (I.1) and (I.2) have to be used. Thus, our matrix F takes the form

$$F^{\mu(\mu_1\mu_2)} = \begin{bmatrix} 0 & 0 & \omega & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega & 0 & 0 \\ \omega & 0 & 0 & 0 & 0 & 0 \\ 0 & \omega & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \omega' \\ 0 & 0 & 0 & 0 & \omega' & 0 \end{bmatrix}, \tag{III.5}$$

where the row and columns of $F^{\mu(\mu_1\mu_2)}$ are enumerated by the multiplicity indices (II.1) and (II.2). Due to Eq. (III.20) of Ref. 10 we have to find a solution of $FB^* = B$ for some unitary B. Obviously,

$$B = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & -i & 0 & 0 & 0 \\ 0 & 1 & 0 & -i & 0 & 0 \\ \omega & 0 & -i\omega & 0 & 0 & 0 \\ 0 & \omega & 0 & -i\omega & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -i \\ 0 & 0 & 0 & 0 & \omega' & -i\omega' \end{bmatrix} \tag{III.6}$$

represents a solution of $FB^* = B$. Hence, it follows that the corresponding columns of the CG matrix W are given by

$$\mathbf{W}_{a_0}^{\mu v} = \sum_{w=1}^{6} B_{wv}\mathbf{M}_{a_0}^{\mu w}, \quad v = 1,2,\ldots 6 \tag{III.7}$$

[see Eq. (III.15) of Ref. 10], where the vectors $\mathbf{M}_{a_0}^{\mu v}$ are defined by Eq. (II.3). A comparison of $\dim B = 6$ with $\dim M = \dim W = 2304$ demonstrates the utility of the present method.

Example 1 (ii): Since $D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}$ is of type III, we have to calculate Eq. (III.18) of Ref. 7, i.e.,

$$\begin{aligned} F_{wv}^{\bar{\mu}(\mu_1\mu_2)} = {} & \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_w j_w)}\|^{-1}\,\|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_v j_v)}\|^{-1} \\ & \times \tfrac{1}{24}\sum_{\beta}\mathbb{D}_{x_w;x_v}^{(0,\mathbf{q})\uparrow G}(\beta|\boldsymbol{\tau}(\beta)|\theta)\,\mathbb{D}_{x_w\sigma_{df};x_v\sigma_{df}}^{(0,\mathbf{q}')\uparrow G}(\beta|\boldsymbol{\tau}(\beta)|\theta) \\ & \times D_{11}^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta|\boldsymbol{\tau}(\beta))\,, \end{aligned} \tag{III.8}$$

where the special indices (II.4) and (II.5) have been already inserted. We obtain after simple calculations

$$F_{i,i+1} = 0, \quad \text{for } i = 1,2\,, \tag{III.9}$$

$$F_{11} = e^{2i\mathbf{q}'(\sigma_{df})\cdot\boldsymbol{\tau}(I)} = e^{i\mathbf{q}'\cdot(\mathbf{t}_1+\mathbf{t}_2+\mathbf{t}_3)} = \omega_1\,, \tag{III.10}$$

$$\begin{aligned} F_{22} = {} & \exp\{i\mathbf{q}(C_{2d})\cdot(D(C_{2d}-1)\,\boldsymbol{\tau}(I) \\ & + i\mathbf{q}'(C_{2d}\sigma_{df})\cdot[1 + D(C_{2d}\sigma_{df})]\,\boldsymbol{\tau}(I)\} \\ = {} & \exp\{i\mathbf{q}\cdot\mathbf{t}_1 + i\mathbf{q}'(\sigma_{df})\cdot(\mathbf{t}_2+\mathbf{t}_3)\} = \omega_2\,, \end{aligned} \tag{III.11}$$

which implies the following for the matrix $F^{\bar{\mu}(\mu_1\mu_2)}$:

$$F^{\bar{\mu}(\mu_1\mu_2)} = \begin{bmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{bmatrix}, \tag{III.12}$$

where row and columns are enumerated by Eqs. (II.4) and (II.5). Hence, the corresponding columns of the CG matrix W are immediately obtained from Eqs. (III.87) and (III.88) of Ref. 10, namely,

$$\mathbf{W}_{1,a_0}^{\mu v} = \mathbf{M}_{a_0}^{\mu v}, \quad v = 1,2\,, \tag{III.13}$$

$$\mathbf{W}_{2,a_0}^{\mu v} = \omega_v\mathbf{M}_{a_0}^{\bar{\mu} v}, \quad v = 1,2\,, \tag{III.14}$$

where the vectors $\mathbf{M}_{a_0}^{\mu v}$ and $\mathbf{M}_{a_0}^{\bar{\mu} v}$ are defined by Eqs. (II.6) and (II.7).

Example 2: In this case it suffices to calculate the matrix elements (III.21) of Ref. 7. These matrix elements turn out to be

$$\begin{aligned} F_{wv}^{\mu(\mu_1\mu_2)} = {} & \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_w j_w)}\|^{-1}\,\|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \\ & \times\mathbb{D}_{S_{63}^{+};S_{63}^{+}}^{(0,\mathbf{q})\uparrow G}(I|\boldsymbol{\tau}(I)|\theta)\,D_{e,w;\,e,v}^{(\mathscr{K}',\mathbf{q}')\uparrow G}(I|\boldsymbol{\tau}(I)) \\ & \times\mathbb{D}_{e;e}^{(0,\mathbf{q}_0)\uparrow G}(I|\boldsymbol{\tau}(I)|\theta)\,, \end{aligned} \tag{III.15}$$

where we have already used Eq. (II.8) and the fact that, for example, the matrix elements $\mathbb{D}_{e;e}^{(0,\mathbf{q}_0)\uparrow G}(\beta|\boldsymbol{\tau}(\beta)|\theta)$, $\beta\in\mathscr{O}_h$ can be different from zero only for $\beta = I$. Utilizing this fact, we obtain for Eq. (III.15)

$$\begin{aligned} F_{wv}^{\mu(\mu_1\mu_2)} &= \exp\{i\mathbf{q}(S_{63}^{+})\cdot\mathbf{t}(S_{63}^{+},I) - i\mathbf{q}'\cdot\boldsymbol{\tau}(I)\}\mathbb{R}_{wv}^{\mathscr{K}'}(I) \\ &= \delta_{wv}\exp\{i\mathbf{q}(S_{63}^{+})\cdot(\mathbf{t}_1+\mathbf{t}_3) - i\mathbf{q}'\cdot\boldsymbol{\tau}(I)\}\mathbb{R}_{ww}^{\mathscr{K}'}(I) \\ &= e^{i\mathbf{q}\cdot(\mathbf{t}_1+\mathbf{t}_2)}(-1)^{v+1}\delta_{wv}, \quad w,v = 1,2\,. \end{aligned} \tag{III.16}$$

It can be shown by a simple calculation that $F_{wv}^{\mu(\mu_1\mu_2)} = F_{vw}^{\mu(\mu_1\bar{\mu}_2)}$ is indeed satisfied, which is in accord with Eq. (II.33) of Ref. 11. Introducing the abbreviation $\exp\{i\mathbf{q}\cdot(\mathbf{t}_1+\mathbf{t}_2)\} = \rho$, we arrive at the final results

$$F = \begin{bmatrix} 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & -\rho \\ \rho & 0 & 0 & 0 \\ 0 & -\rho & 0 & 0 \end{bmatrix} \tag{III.17}$$

and

$$B = \frac{1}{\sqrt{2}} \begin{bmatrix} i & 0 & 1 & 0 \\ 0 & i & 0 & 1 \\ -i\rho & 0 & \rho & 0 \\ 0 & i\rho & 0 & -\rho \end{bmatrix}, \tag{III.18}$$

where B represents a solution of $FB^* = B$ [see Eq. (II.43) of Ref. 11]. Consequently, the corresponding (I ⊗ III : I) CG coefficients for $Pn3'n$ read as [compare Eqs. (II.45) and (II.46) of Ref. 11]

$$\mathbf{W}_{a_0}^{\mu(1v)} = \frac{i}{\sqrt{2}}\{\mathbf{Q}_{a_0}^{\mu v1} + \rho(-1)^{v}\mathbf{Q}_{a_0}^{\mu v2}\}, \quad v = 1,2\,, \tag{III.19}$$

$$\mathbf{W}_{a_0}^{\mu(2v)} = \frac{1}{\sqrt{2}}\{\mathbf{Q}_{a_0}^{\mu v1} - \rho(-1)^{v}\mathbf{Q}_{a_0}^{\mu v2}\}, \quad v = 1,2\,, \tag{III.20}$$

where the nontrivial components of [see Eqs. (II.14) and

(II.15) of Ref. 11]

$$\{\mathbf{Q}_{a_0}^{\mu v 1}\}_{i,bj} = \delta_{b1}\{\mathbf{M}_{a_0}^{\mu v}\}_{ij}\,, \tag{III.21}$$

$$\{\mathbf{Q}_{a_0}^{\mu v 2}\}_{i,bj} = \delta_{b2}\{\mathbf{N}_{a_0}^{\mu v}\}_{ij}\,, \tag{III.22}$$

are given by Eq. (II.10). Note that $\dim F^{\mu(\mu_1\mu_2)} = 2$, whereas $\dim M = \dim N = 288 = \frac{1}{2}\dim W$.

Example 3 (i): This example is analogous to the previous one, but with the only difference that both **q** vectors **q** and **q**′ belong to stars of higher symmetry. Equation (III.21) of Ref. 7 turns out to be for this case

$$F_{wv}^{\mu(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_w j_w)}\|^{-1}\,\|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \times \tfrac{1}{8}\sum_{\beta} D_{e,1;e,1}^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\boldsymbol{\tau}(\beta))\, D_{C_{31}^-,1;C_{31}^-,1}^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta|\boldsymbol{\tau}(\beta)) \times \mathbb{D}_{e,1;e,1}^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta|\boldsymbol{\tau}(\beta)|\theta)^*\,, \tag{III.23}$$

where the summation about $\beta\in\mathscr{O}_h$ reduces because of Eqs. (I.32) and (I.18) to the single group element I. Furthermore, we have to note that $\mathbb{D}^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\boldsymbol{\tau}(\beta)|\theta) = D^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\boldsymbol{\tau}(\beta))$ has already been taken into account [see Eq. (I.26) of the present paper]. A simple calculation yields

$$F_{wv}^{\mu(\mu_1\mu_2)} = i = F_{vw}^{\mu(\mu_1\bar{\mu}_2)}, \quad w = v = 1\,, \tag{III.24}$$

since $B_{e,e}^{\mathbf{q}}(I)\,B_{C_{31}^-,C_{31}^-}^{\mathbf{q}'}(I)\,B_{e,e}^{\mathbf{q}_0}(I)^* = \exp\{-i\mathbf{Q}[\ldots]\cdot\boldsymbol{\tau}(\beta)\}$. Therefore we have

$$F = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \tag{III.25}$$

and

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}. \tag{III.26}$$

which give rise to the following (I ⊗ III : I) CG coefficients;

$$\mathbf{W}_{a_0}^{\mu(1v)} = \frac{i}{\sqrt{2}}\{\mathbf{Q}_{a_0}^{\mu v 1} - i\mathbf{Q}_{a_0}^{\mu v 2}\}, \quad v = 1\,. \tag{III.27}$$

$$\mathbf{W}_{a_0}^{\mu(2v)} = \frac{1}{\sqrt{2}}\{\mathbf{Q}_{a_0}^{\mu v 1} + i\mathbf{Q}_{a_0}^{\mu v 2}\}, \quad v = 1, \tag{III.28}$$

where the nontrivial components of $\mathbf{Q}_{a_0}^{\mu v a}$ are given by Eqs. (II.14) and (II.15).

Example 3 (ii): In contrast to the foregoing example $D^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}$ is of type III. Therefore, we have to compute Eqs. (III.24) and (III.25) of Ref. 7. The first formula specifies to

$$F_{wv}^{\bar{\mu}(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\bar{\mu}(i_w j_w)}\|^{-1}\|\mathbf{B}_{a_0}^{\mu_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \times \tfrac{1}{8}\sum_{\beta} D_{e,1;e,1}^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\boldsymbol{\tau}(\beta)) D_{C_{31}^-,2;C_{31}^-,2}^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta|\boldsymbol{\tau}(\beta)) \times D_{e,1;e,1}^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta|\boldsymbol{\tau}(\beta))\,, \tag{III.29}$$

whereas for the second formula the matrix elements $D^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta|\boldsymbol{\tau}(\beta)), \beta\in\mathscr{O}_h$ have to be replaced by their complex conjugate values. A simple inspection of $\mathbb{R}_{11}^{\mathscr{K}}(\beta)$, $\mathbb{R}_{22}^{\mathscr{K}'}(C_{31}^+\beta C_{31}^-)$, and $\mathbb{R}_{11}^{\mathscr{K}_0}(\beta)$ shows that the summation about β reduces to e and I. Carrying out the corresponding calculations we obtain for both cases

$$F_{wv}^{\bar{\mu}(\mu_1\mu_2)} = F_{wv}^{\bar{\mu}(\mu_1\bar{\mu}_2)} = 1, \quad w,v = 1\,. \tag{III.30}$$

Using Eqs. (II.137) – (II.139) of Ref. 11, we arrive immediately to the final results

$$\mathbf{W}_{1,a_0}^{\mu(av)} = \mathbf{Q}_{a_0}^{\mu v a}, \quad a = 1,2 \text{ and } v = 1\,, \tag{III.31}$$

$$\mathbf{W}_{2,a_0}^{\mu(av)} = \mathbf{Q}_{a_0}^{\bar{\mu} v a+1}, \quad a = 1,2 \text{ and } v = 1\,, \tag{III.32}$$

where the nontrivial components of

$$\{\mathbf{Q}_{a_0}^{\mu v 1}\}_{i,bj} = \delta_{b1}\{\mathbf{M}_{a_0}^{\mu v}\}_{ij}\,, \tag{III.33}$$

$$\{\mathbf{Q}_{a_0}^{\mu v 2}\}_{i,bj} = \delta_{b2}\{\mathbf{N}_{a_0}^{\mu v}\}_{ij}\,, \tag{III.34}$$

$$\{\mathbf{Q}_{a_0}^{\bar{\mu} v 1}\}_{i,bj} = \delta_{b1}\{\mathbf{M}_{a_0}^{\bar{\mu} v}\}_{ij}\,, \tag{III.35}$$

$$\{\mathbf{Q}_{a_0}^{\bar{\mu} v 2}\}_{i,bj} = \delta_{b2}\{\mathbf{N}_{a_0}^{\bar{\mu} v}\}_{ij}\,, \tag{III.36}$$

are given by Eqs. (II.18)–(II.21). Comparing the dimensions of the unitary matrices $F^{\bar{\mu}(\mu_1\mu_2)}$ and $F^{\bar{\mu}(\mu_1\bar{\mu}_2)}$ ($\dim F^{\bar{\mu}(\ldots)} = 1$) with those of the CG matrices M, N, and W ($\dim M = \dim N = \frac{1}{2}\dim W = 36$), we realize once more the utility of the present method.

Example 4: This case is of course the most complicated one which shall be considered and requires only the calculation of the matrix elements (III.28) and (III.29) of Ref. 7, since the remaining are obtained by the symmetry relations (II.42) and (II.43) of Ref. 12. Equation (III.28) of Ref. 7 turns out to be

$$F_{wv}^{\mu(\mu_1\mu_2)} = \|\mathbf{B}_{a_0}^{\mu_1\mu_2;\mu(i_w j_w)}\|^{-1}\|\mathbf{B}_{a_0}^{\bar{\mu}_1\bar{\mu}_2;\mu(i_v j_v)}\|^{-1} \times \tfrac{1}{8}\sum_{\beta} D_{e,1;e,1}^{(\mathscr{K},\mathbf{q})\uparrow G}(\beta|\boldsymbol{\tau}(\beta))\, D_{C_{31}^-,2;C_{31}^-,2}^{(\mathscr{K}',\mathbf{q}')\uparrow G}(\beta|\boldsymbol{\tau}(\beta)) \times D_{e,1;e,1}^{(\mathscr{K}_0,\mathbf{q}_0)\uparrow G}(\beta|\boldsymbol{\tau}(\beta))^*\,; \tag{III.37}$$

the analog to Eq. (III.29) of Ref. 7 is readily obtainable from Eq. (III.37) by replacing the matrix elements of $D^{(\mathscr{K}',\mathbf{q}')\uparrow G}$ through their complex conjugate values. Simple calculations yield for both cases

$$F_{wv}^{\mu(\mu_1\mu_2)} = F_{wv}^{\mu(\mu_1\bar{\mu}_2)} = 1, \quad w,v = 1\,, \tag{III.38}$$

which leads immediately to

$$F = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{III.39}$$

and to a possible solution of $FB^* = B$, namely,

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i & 0 & 0 & 1 \\ 0 & i & 1 & 0 \\ 0 & -i & 1 & 0 \\ -i & 0 & 0 & 1 \end{bmatrix} \tag{III.40}$$

[compare Eqs. (II.41) and (II.58) of Ref. 12]. Hence, the corresponding (III ⊗ III : I) CG coefficients for $Pn3'n$ take the form

$$\mathbf{W}_{a_0}^{\mu(abv)} = \frac{1}{\sqrt{2}}\begin{cases} i[\mathbf{Q}_{a_0}^{\mu v ab} - \mathbf{Q}_{a_0}^{\mu v a+1\, b+1}], & a = 1, b = 1,2, \text{ and } v = 1, \\ [\mathbf{Q}_{a_0}^{\mu v a+1\, b+1} + \mathbf{Q}_{a_0}^{\mu v ab}], & a = 2, b = 1,2, \text{ and } v = 1 \end{cases} \tag{III.41}$$

where the nontrivial components of

$$\{\mathbf{Q}^{\mu\nu 11}_{a_0}\}_{ai,bj} = \delta_{a1}\delta_{b1}\{\mathbf{K}^{\mu\nu}_{a_0}\}_{ij}\,, \quad \text{(III.42)}$$

$$\{\mathbf{Q}^{\mu\nu 12}_{a_0}\}_{ai,bj} = \delta_{a1}\delta_{b2}\{\mathbf{L}^{\mu\nu}_{a_0}\}_{ij}\,, \quad \text{(III.43)}$$

$$\{\mathbf{Q}^{\mu\nu 21}_{a_0}\}_{ai,bj} = \delta_{a2}\delta_{b1}\{\mathbf{M}^{\mu\nu}_{a_0}\}_{ij}\,, \quad \text{(III.44)}$$

$$\{\mathbf{Q}^{\mu\nu 22}_{a_0}\}_{ai,bj} = \delta_{a2}\delta_{b2}\{\mathbf{N}^{\mu\nu}_{a_0}\}_{ij}\,, \quad \text{(III.45)}$$

are given by Eqs. (II.25)–(II.28). Concluding this section, we mention once more that, also for this example, there exists a remarkable difference between the dimensions of $F^{\mu(\ldots)}$ ($\dim F^{\mu(\ldots)} = 1$) and that of K, L, M, N, and W ($\dim K = \dim L = \dim M = \dim N = \frac{1}{4}\dim W = 36$).

IV. CONCLUDING REMARKS

It was the aim of this article to demonstrate the utility of the present method on hand of the type II Shubnikov space group *Pn*3′*n* by discussing a series of examples. Thereby we have shown that even for the most complicated examples the determination of CG coefficients for *Pn*3′*n* requires the computation of simple formulas. The only demerit of the group *Pn*3′*n* is that type II co-unirreps are not realized and that all possible type III corepresentations are characterized among others by the fact that the corresponding little co-groups P^q always contain the inversion. Apart from this we can summarize that the main points of our method consist of determining convenient CG coefficients for *Pn*3*n* (which belong to a fixed column index of the considered unirrep of *Pn*3*n*) and to compute with them unitary matrices which link CG coefficients for *Pn*3′*n* with those for *Pn*3*n*. In this connection we remark that these unitary matrices can be computed without explicit knowledge of the corresponding CG coefficients for *Pn*3*n*, presupposing their multiplicity indices can be traced back to special column indices of the considered Kronecker products.

[1]R. Dirl, J. Math Phys. **20**, 679 (1979).
[2]R. Dirl, J. Math Phys. **20**, 1575 (1979).
[3]C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[4]R. Dirl, J. Math Phys. **18**, 2065 (1977).
[5]R. Dirl, J. Math Phys. **20**, 1566 (1979).
[6]R. Dirl, J. Math Phys. **20**, 1570 (1979).
[7]R. Dirl, J. Math Phys. **21**, 2011 (1980).
[8]R. Dirl, J. Math Phys. **20**, 664 (1979).
[9]R. Dirl, J. Math Phys. **20**, 671 (1979).
[10]R. Dirl, J. Math Phys. **21**, 961 (1980).
[11]R. Dirl, J. Math Phys. **21**, 975 (1980).
[12]R. Dirl, J. Math Phys. **21**, 997 (1980).

Congruence number, a generalization of SU(3) triality [a]

F. Lemire
Department of Mathematics, University of Windsor, Windsor, Ontario, Canada
J. Patera
Centre de Recherche de Mathématiques Appliquées, Université de Montréal, Montréal, Québec, Canada

(Received 3 December 1979; accepted for publication 14 March 1980)

Congruence classes of finite-dimensional representations of semisimple Lie groups are defined. Each class is characterized by a congruence number. For the group SU(3) the concept reduces to the familiar triality number.

The well-known triality number[1] of irreducible representations of SU(3) and, more generally, "n-ality" number for SU(n), is a label which facilitates many of the standard computations on such representations. These labels also have importance as quantum numbers of elementary particles. Only particles having triality 0 should be observable according to some models. Similar labels for representations of other semisimple Lie groups are not generally known although particular properties of this type have been part of particle physics folklore for some time.

In this note we define the general concept of congruence classes characterized by congruence numbers of irreducible representations of a semisimple Lie group G which reduces to the concept on n-ality when G = SU(n). Particularly interesting and apparently unknown are the results for orthogonal groups O($2n$); there are four congruence classes of irreducible representations of O($2n$) labelled by two-component congruence numbers. These congruence numbers add as vectors under the tensor product of representation and furthermore, nonequivalent spinor representations of equal dimensions belong to different congruence classes. [In particular, the three representations of O(8) of dimension eight are all distinguished by their congruence numbers.]

We say that two irreducible representations Λ and Λ' of a semisimple Lie group G belong to the same congruence class or are mutually congruent iff the difference of any weight from the representation Λ and any weight from the representation Λ' is a linear combination of simple roots of G with integer coefficients.

An irreducible representation Λ of G of rank n is specified by n nonnegative integers

$$(a_1 a_2 \cdots a_n), \quad a_j = 2(\Lambda,\alpha_j)/(\alpha_j,\alpha_j),$$

where the subscript j of a_j refers to the jth simple root α_j of G; we adopt Dynkin's numbering of simple roots.[2]

Let us first summarize the results.

The congruence number c of an irreducible representation $(a_1 a_2 \cdots a_n)$ of a simple Lie group G is given by

$$c = \sum_{k=1}^{n-1} k a_k \pmod n \quad \text{for} \quad G = SU(n),$$

$$c = a_n \pmod 2 \quad \text{for } G = O(2n+1),$$

$$c = a_1 + a_3 + a_5 \cdots \pmod 2 \quad \text{for} \quad G = Sp(2n),$$

$$c = a_1 - a_2 + a_4 - a_5 \pmod 3 \quad \text{for} \quad G = E_6,$$

$$c = a_4 + a_6 + a_7 \pmod 2 \quad \text{for} \quad G = E_7,$$

$c = 0$ for all representations of E_8, F_4 and G_2 groups.

There are four congruence classes of irreducible representation of O($2n$) groups. Each class is labeled by a two-component vector

$$c = (a_{n-1} + a_n, 2a_1 + 2a_3 + \cdots + 2a_{n-2} + (n-2)a_{n-1} + na_n) \quad \text{for } n \text{ odd}$$

$$\pmod{2, \bmod 4}.$$

$$c = (a_{n-1} + a_n, 2a_1 + 2a_3 + \cdots + 2a_{n-3} + (n-2)a_{n-1} + na_n) \quad \text{for } n \text{ even}$$

That is, the first component is calculated mod 2, the second one mod 4. When n is even the only vectors c which may occur are (0,0), (0,2), (1,0), and (1,2). For n odd, the four different congruence vectors are (0,0), (0,2), (1,1), and (1,3).

If the group G is not simple, the congruence number c is a vector with components corresponding to each simple ideal of G.

The above classification of irreducible representations of semisimple Lie group into congruence classes and their characterization by congruence number c is a straightforward generalization of a theorem of Dynkin.[3]

[a] Work supported in part by the Ontario–Québec Exchange Programme, by the National Science and Engineering Research Council of Canada, and by the Ministère de l'Education du Québec.

We now draw particular attention to two properties of these congruence classes which prove to have practical application to computations. First, if R_i denote irreducible representations of the simple Lie group G and $c(R_i)$ denote their corresponding congruence numbers then $R_1 \otimes R_2 = \cdots = R_3 \oplus \cdots \oplus R_k$ implies that

$$c(R_3) = \cdots = c(R_k) = c(R_1) + c(R_2),$$

where equalities and addition are to be interpreted relative to the appropriate modulo of the congruence numbers.

We illustrate this property with some examples. First a completely trivial example of SU(3) representations,

$$\begin{array}{ccccccccc} (1,0) & \otimes & (0,1) & = & (1,1) & \oplus & (0,0) & & \\ 3 & \times & 3 & = & 8 & + & 1, & & \\ 1 & + & 2 & = & 0 & = & 0 & & (\text{mod } 3). \end{array}$$

The first line describes the tensor product and its decomposition, the second line gives the corresponding dimensions, and the final line provides the appropriate relation of the congruence numbers of the representations.

A less trivial example involves representations of E_6.

$$\begin{array}{cccccccccccc} (1,0,0,0,0,0) & \otimes & (0,1,0,0,0,0) & = & (1,1,0,0,0,0) & \oplus & (0,0,1,0,0,0) & \oplus & (1,0,0,0,1,0) & \oplus & (0,0,0,0,0,1), & \\ 27 & \times & 351 & = & 5{,}824 & + & 2{,}925 & + & 650 & + & 78, & \\ 1 & + & 2 & = & 0 & = & 0 & = & 0 & = & 0 & (\text{mod } 3). \end{array}$$

Subsequent examples concern the O(2*n*) groups. For O(8) one has

$$\begin{array}{cccccccc} (1,0,0,0) & \otimes & (0,0,0,1) & = & (1,0,0,1) & \oplus & (0,0,1,0) & \\ 8 & \times & 8 & = & 56 & + & 8, & \\ (0,2) & + & (1,0) & = & (1,2) & = & (1,2) & (\text{mod } 2, \text{mod } 4). \end{array}$$

And finally, for O(10) we have

$$\begin{array}{cccccccccc} (0,0,0,1,0) & \otimes & (0,0,0,0,1) & = & (0,0,0,1,1) & \oplus & (0,1,0,0,0) & \oplus & (0,0,0,0,0), & \\ 16 & \times & 16 & = & 210 & + & 45 & + & 1, & \\ (1,3) & + & (1,1) & = & (0,0) & = & (0,0) & = & (0,0) & (\text{mod } 2, \text{mod } 4). \end{array}$$

A second important property of these congruence classes relates to their application to the decomposition of irreducible representations of a group G with respect to a subgroup G′. For each group-subgroup pair (G,G′) one can specify rules as to which congruence classes of irreducible representation of G′ can occur in the reduction of a given irreducible representation of G. This is particularly simple when G′ is an *integral*[4] subgroup of G, i.e., the adjoint representation of G decomposes into irreducible representations of G′ which all belong to the zero congruence class. If G′ is an integral subgroup G then any representation of G considered as a representation of G′ decomposes into irreducible G′ representations which all must belong to the same congruence class.

[1]G. E. Baird and L. C. Biedenharn, in *Symmetry Principles at High Energy*, edited by B. Korsunoglu and A. Perlmutter (Freeman, San Francisco, 1964).

[2]Table 1 of Ref. 4.

[3]Theorem 3.1 of Ref. 4. Note that the statement of the theorem concerning the groups O(2*n*) requires a correction.

[4]E. B. Dynkin, "Semisimple Subalgebras of Semisimple Lie Algebras," Am. Math. Soc. Transl. Ser. 2, **6**, 111 (1957).

Simple derivation of the Newton–Wigner position operator

Thomas F. Jordan
Department of Physics, University of Minnesota, Duluth, Duluth, Minnesota 55812

(Received 6 November 1979; accepted for publication 28 March 1980)

The kind of operator algebra familiar in ordinary quantum mechanics is used to show formally that in an irreducible unitary representation of the Poincaré group for positive mass, the Newton–Wigner position operator is the only Hermitian operator with commuting components that transforms as a position operator should for translations, rotations, and time reversal and does not behave in a singular way that contradicts what can be learned from Lorentz transformations in the nonrelativistic limit.

1. INTRODUCTION

In basing the description of a particle on an irreducible unitary representation of the Poincaré group, we begin with the operators for transformations of states corresponding to transformations of space–time coordinates.[1] From these we have to derive the operators representing variables such as position and velocity. The velocity operator can be identified very simply by its transformations under the Poincaré group, together with the assumption that it is a Hermitian operator whose components commute with each other.[2,3] Our goal here is a simple derivation of the Newton–Wigner position operator.

The corresponding problem for the Galilei group is very simple. Suppose we want to identify the position operator $\mathbf{R}$ in an irreducible unitary representation of the Galilei group for mass m, spin s, and lowest energy E_0. The generators are $\mathbf{P}$ for space translations,

$$H = E_0 + (1/2m)\mathbf{P}^2 \tag{1.1}$$

for time translations,

$$\mathbf{J} = \mathbf{Q}\times\mathbf{P} + \mathbf{S} \tag{1.2}$$

for rotations, and

$$\mathbf{K} = m\mathbf{Q} \tag{1.3}$$

for Galilei transformations.[4] We use $(2s+1)$-component momentum–space wavefunctions on which $\mathbf{P}$ is multiplication, $\mathbf{Q}$ is $i\nabla$, and the three components of $\mathbf{S}$ are the usual $(2s+1)\times(2s+1)$ irreducible spin matrices for spin s. For a position operator $\mathbf{R}$ that transforms right for space translations and Galilei transformations we must have[5]

$$[R_j, P_k] = i\delta_{jk}, \tag{1.4}$$

and

$$[R_j, K_k] = 0 \tag{1.5}$$

for $j,k = 1,2,3$. [See Eqs. (5.5)–(5.8).] Let

$$\mathbf{R} = \mathbf{Q} + \mathbf{D}.$$

Then $\mathbf{D}$ must commute with $\mathbf{Q}$ and $\mathbf{P}$. If $\mathbf{R}$ is to be Hermitian and rotate as a vector, then $\mathbf{D}$ must be Hermitian and rotate as a vector. The only possibility is

$$\mathbf{D} = b\,\mathbf{S},$$

with b a real number. If $\mathbf{R}$ is to be invariant for time reversal, then $\mathbf{D}$ must be invariant, which implies $\mathbf{D}$ is zero (because to fit in with the unitary transformations of the Galilei group, the antiunitary time-reversal operator must commute with H and $\mathbf{K}$ and anticommute with $\mathbf{P}$ and $\mathbf{J}$, which means it commutes with $\mathbf{Q}$ and anticommutes with $\mathbf{S}$).[5] We may choose to not assume $\mathbf{R}$ is invariant for time reversal but instead assume the components of $\mathbf{R}$ commute with each other. Then again $\mathbf{D}$ must be zero. Either way, we find that $\mathbf{R}$ is $\mathbf{Q}$.

For the Poincaré group, the classic and rigorous derivations of the Newton–Wigner position operator are not so direct.[6,7] Here we give a formal derivation, along the same lines as for the Galilei group, using the same kind of operator algebra that is familiar in ordinary quantum mechanics.

We show that in an irreducible unitary representation of the Poincaré group for positive mass, the Newton–Wigner position operator is the unique Hermitian operator with commuting components that transforms as a position operator should for translations, rotations, and time reversal and does not behave in a singular way that contradicts what we know of Lorentz transformations in the nonrelativistic limit. The assumptions we make about the latter are discussed in a separate section at the end of the paper. They replace the regularity conditions used by Newton and Wigner[6] and Wightman.[7]

We write the generators of the Poincaré group in two different forms, but use the Foldy form at the beginning and end of the derivation and the Moses form for just one technical step, so the idea of the derivation can be followed using only the Foldy form.

2. FOLDY FORM

We consider an irreducible unitary representation of the Poincaré group with positive mass m, spin s, and positive energy. We use units such that c and $\hbar$ are 1. Let $\mathbf{P}$ denote the generator for space translations. The generator for time translations is

$$H = (P^2 + m^2)^{1/2}, \tag{2.1}$$

where P is $(\mathbf{P}^2)^{1/2}$. In the Foldy form[8] the generators for rotations and Lorentz transformations are

$$\mathbf{J} = \mathbf{Q}\times\mathbf{P} + \mathbf{S}, \tag{2.2}$$

 0022-2488/80/082028-05$1.00

$$\mathbf{K} = (1/2)(H\mathbf{Q} + \mathbf{Q}H) + (H + m)^{-1}\mathbf{P}\times\mathbf{S}, \tag{2.3}$$

Here $\mathbf{Q}$ is $i\nabla$, and S_1, S_2, S_3 are the usual $(2s+1)\times(2s+1)$ irreducible spin matrices for spin s, on momentum–space wave functions with $2s+1$ components labeled by the eigenvalues $\lambda = -s, -s+1, \ldots, s$ of the diagonal matrix S_3, with inner product

$$(\psi,\phi) = \sum_{\lambda=-s}^{s} \int d^3p\, \psi_\lambda(\mathbf{p})^* \phi_\lambda(\mathbf{p}). \tag{2.4}$$

The Pauli–Lubanski four-vector is

$$W_0 = \mathbf{P}\cdot\mathbf{J} = \mathbf{P}\cdot\mathbf{S}, \tag{2.5}$$

$$\mathbf{W} = H\mathbf{J} + \mathbf{P}\times\mathbf{K} = (H+m)^{-1}(\mathbf{P}\cdot\mathbf{S})\mathbf{P} + m\mathbf{S}. \tag{2.6}$$

From this we see that in terms of the generators, the spin operator is

$$\mathbf{S} = m^{-1}(H\mathbf{J} + \mathbf{P}\times\mathbf{K}) - m^{-1}(H+m)^{-1}(\mathbf{P}\cdot\mathbf{J})\mathbf{P}. \tag{2.7}$$

The Newton–Wigner position operator[6] is $\mathbf{Q}$. In terms of the generators it is

$$\mathbf{Q} = H^{-1}(\mathbf{K} - i\mathbf{P}/2H) - H^{-1}m^{-1}(H+m)^{-1}\mathbf{P}\times(H\mathbf{J} + \mathbf{P}\times\mathbf{K}). \tag{2.8}$$

Clearly $\mathbf{Q}$ is Hermitian, and its components commute with each other. It transforms as a position operator should for space translations because

$$[Q_j, P_k] = i\delta_{jk} \tag{2.9}$$

for $j,k = 1,2,3$. That $\mathbf{Q}$ rotates as a vector is evident from the form (2.2) of $\mathbf{J}$ or the formula (2.8) for $\mathbf{Q}$ in terms of the generators. From the latter we can also see that $\mathbf{Q}$ is invariant for time reversal, because to fit in with the other transformations the time-reversal operator must commute with H and $\mathbf{K}$, anticommute with $\mathbf{P}$ and $\mathbf{J}$, and be antilinear (anticommute with i). We can see similarly that $\mathbf{Q}$ changes sign for parity, and from the form (2.1) of H and the commutation relations (2.9) we see that the commutator of $\mathbf{Q}$ with H gives the correct[2,3] velocity operator $\mathbf{P}/H$, but these two properties are not needed to identify the Newton–Wigner position operator. We will consider Lorentz transformations later.

If instead of the inner product (2.4) we were to use the inner product defined with the invariant $(\mathbf{p}^2 + m^2)^{-1/2}\, d^3p$, we would change from momentum–space wave functions $\psi(\mathbf{p})$ to $(\mathbf{p}^2 + m^2)^{1/4}\psi(\mathbf{p})$. Then the Newton–Wigner position operator would appear as[6]

$$(P^2 + m^2)^{1/4}\mathbf{Q}(P^2 + m^2)^{-1/4} = \mathbf{Q} - i(1/2)(P^2 + m^2)^{-1}\mathbf{P},$$

where $\mathbf{Q}$ is still $i\nabla$.

3. MOSES FORM

In the "standard helicity" form of H. E. Moses,[9] which can be obtained from the Foldy form by a unitary change of basis, the generators for rotations and Lorentz transformations are

$$\mathbf{J} = \mathbf{Q}\times\mathbf{P} + \mathbf{M}S, \tag{3.1}$$

$$\mathbf{K} = (1/2)(H\mathbf{Q} + \mathbf{Q}H) + (H/P)\mathbf{N}S - (m/P)(\mathbf{E}_2T_1 - \mathbf{E}_1T_2), \tag{3.2}$$

where

$$\mathbf{M} = (P_1/(P+P_3), P_2/(P+P_3), 1) = (P\hat{z} + \mathbf{P})/(P + \mathbf{P}\cdot\hat{z}), \tag{3.3}$$

$$\mathbf{N} = (P_2/(P+P_3), -P_1/(P+P_3), 0) = \mathbf{P}\times\hat{z}/(P + \mathbf{P}\cdot\hat{z}), \tag{3.4}$$

$$\mathbf{E}_1 = (P_1P_2/P(P+P_3), -P_1^2/P(P+P_3) - P_3/P, P_2/P), \tag{3.5}$$

$$\mathbf{E}_2 = (P_2^2/P(P+P_3) + P_3/P, -P_1P_2/P(P+P_3), -P_1/P), \tag{3.6}$$

again $\mathbf{Q}$ is $i\nabla$ on $(2s+1)$-component momentum–space wave functions with the inner product (2.4), but now T_1, T_2, S are the usual $(2s+1)\times(2s+1)$ spin matrices for spin s, the same as used for S_1, S_2, S_3 in the Foldy form, and the components of the wavefunctions are labeled by the eigenvalues $\lambda = -s, -s+1, \ldots, s$ of the diagonal matrix S. The vectors $\mathbf{E}_1, \mathbf{E}_2$ and $\hat{P} = \mathbf{P}/P$ are orthonormal and $\hat{P}\times\mathbf{E}_1$ is $\mathbf{E}_2$, etc.

The point is that S represents the helicity. The Pauli–Lubanski four-vector is

$$W_0 = \mathbf{P}\cdot\mathbf{J} = PS, \tag{3.7}$$

$$\mathbf{W} = H\mathbf{J} + \mathbf{P}\times\mathbf{K} = H\hat{P}S + m(\mathbf{E}_1T_1 + \mathbf{E}_2T_2). \tag{3.8}$$

Substituting these into Eq. (2.7) we find that the spin operator $\mathbf{S}$ of the Foldy form is

$$\mathbf{E}_1T_1 + \mathbf{E}_2T_2 + \hat{P}S.$$

Making the same substitution in the formula (2.8) for the Newton–Wigner position operator in terms of the generators we find that in the Moses form the Newton–Wigner position operator is

$$\mathbf{R} = \mathbf{Q} + P^{-1}\mathbf{N}S - P^{-1}(\mathbf{E}_2T_1 - \mathbf{E}_1T_2). \tag{3.9}$$

4. UNIQUENESS PROOF

We can see rather directly that the Newton–Wigner position operator is unique. Our proof uses the Foldy form at the beginning and end and uses the Moses form for just one technical step that is easy to accept intuitively, so the idea of the proof can be followed using only the Foldy form.

We will show that in the irreducible unitary representation of the Poincaré group there is no other Hermitian operator $\mathbf{R}$ with commuting components that transforms as a position operator should for translations, rotations, and time reversal and does not behave in a singular way that contradicts what can be learned from Lorentz transformations in the nonrelativistic limit.

Suppose $\mathbf{R}$ is such a position operator. Using the Foldy form now, we let

$$\mathbf{R} = \mathbf{Q} + \mathbf{D}.$$

Then $\mathbf{D}$ must be Hermitian, invariant for translations, a vector for rotations, and invariant for time reversal.

Consider $\mathbf{D}\cdot\hat{P}$. It is invariant for translations and rotations; it commutes with $\mathbf{P}$ and $\mathbf{J}$. Therefore it commutes with $\mathbf{P}$ and $\hat{P}\cdot\mathbf{S} = \hat{P}\cdot\mathbf{J}$. Now $\mathbf{P}$, $\hat{P}\cdot\mathbf{S}$ is a complete set of commuting operators, so since $\mathbf{D}\cdot\hat{P}$ commutes with them, $\mathbf{D}\cdot\hat{P}$ must be a function of $\mathbf{P}$, $\hat{P}\cdot\mathbf{S}$. Since $\mathbf{D}\cdot\hat{P}$ and $\hat{P}\cdot\mathbf{S}$ are rotation invariant, $\mathbf{D}\cdot\hat{P}$ must be a function of P, $\hat{P}\cdot\mathbf{S}$. The time-reversal operator anticommutes with **P** and **J**, so it commutes with P and $\hat{P}\cdot\mathbf{S} = \hat{P}\cdot\mathbf{J}$. The time-reversal operator must anticommute with $\mathbf{D}\cdot\hat{P}$, so $\mathbf{D}\cdot\hat{P}$ must be an imaginary function of P and $\hat{P}\cdot\mathbf{S}$. But $\mathbf{D}\cdot\hat{P}$ is Hermitian, because **D** and $\hat{P}$ commute with each other. Therefore $\mathbf{D}\cdot\hat{P}$ is zero.

For zero spin, our proof is easily completed. To be invariant for translations and a vector for rotations, **D** must be of the form

$$\mathbf{D} = F(P)\mathbf{P}$$

for some function F of P. Then, since $\mathbf{D}\cdot\hat{P}$ is zero, **D** is zero. Therefore **R** is **Q**.

In general, since **D** is invariant for translations and a vector for rotations, and $\mathbf{D}\cdot\hat{P}$ is zero, we might expect **D** to be of the form

$$\mathbf{D} = -B\hat{P}\times(\hat{P}\times\mathbf{S}) + C\hat{P}\times\mathbf{S},$$

with B and C functions of P and $\hat{P}\cdot\mathbf{S}$. Our next step is to prove that. If you believe it, you can skip to Eqs. (4.9) and (4.10).

For this step, we transform everything to the Moses form. For the Newton–Wigner position operator we now have Eq. (3.9) instead of **Q**, so we have

$$\mathbf{R} = \mathbf{Q} + P^{-1}\mathbf{N}S - P^{-1}(\mathbf{E}_2T_1 - \mathbf{E}_1T_2) + \mathbf{D}. \qquad (4.1)$$

Since $\mathbf{D}\cdot\hat{P}$ is zero, we can let

$$\mathbf{D} = B_1\mathbf{E}_1 + B_2\mathbf{E}_2. \qquad (4.2)$$

From Eq. (3.8) we have

$$(1/m)\hat{P}\times\mathbf{W} = \mathbf{E}_2T_1 - \mathbf{E}_1T_2, \qquad (4.3)$$

$$(-1/m)\hat{P}\times(\hat{P}\times\mathbf{W}) = \mathbf{E}_1T_1 + \mathbf{E}_2T_2. \qquad (4.4)$$

Then

$$\mathbf{D}\cdot(1/m)\hat{P}\times\mathbf{W} = B_2T_1 - B_1T_2,$$

$$\mathbf{D}\cdot(-1/m)\hat{P}\times(\hat{P}\times\mathbf{W}) = B_1T_1 + B_2T_2.$$

These operators are invariant for translations and rotations; they commute with **P** and **J**. Then they commute with $S = \hat{P}\cdot\mathbf{J}$. It follows that they are functions of $\mathbf{P},S$, because they commute with $\mathbf{P},S$ and $\mathbf{P},S$ is a complete set of commuting operators. The representation of the Poincaré group is spanned by eigenkets $|\mathbf{p},\lambda\rangle$ of $\mathbf{P},S$, in terms of which we have

$$(B_1 \pm iB_2)(T_1 \mp iT_2)|\mathbf{p},\lambda\rangle$$
$$= [B_1T_1 + B_2T_2 \pm i(B_2T_1 - B_1T_2)]|\mathbf{p},\lambda\rangle$$
$$\propto |\mathbf{p},\lambda\rangle.$$

We can see similarly that

$$(T_1 \mp iT_2)(B_1 \pm iB_2)|\mathbf{p},\lambda\rangle \propto |\mathbf{p},\lambda\rangle.$$

Of course

$$(T_1 \pm iT_2)|\mathbf{p},\lambda\rangle \propto |\mathbf{p},\lambda \pm 1\rangle, \qquad (4.5)$$

with a proportionality factor that is zero only when λ is the maximum or minimum eigenvalue of S and there is no $|\mathbf{p},\lambda \pm 1\rangle$ in the representation. It follows that

$$B_1 \pm iB_2 = f_\pm(T_1 \pm iT_2) + \pi_\pm,$$

or

$$\mathbf{D} = B(\mathbf{E}_1T_1 + \mathbf{E}_2T_2) + C(\mathbf{E}_2T_1 - \mathbf{E}_1T_2)$$
$$+ (1/2)(\pi_+ + \pi_-)\mathbf{E}_1 - (i/2)(\pi_+ - \pi_-)\mathbf{E}_2, \qquad (4.6)$$

where $f_\pm$ and B,C are functions of $\mathbf{P},S$, and $\pi_\pm$ are operators restricted such that

$$\pi_\pm|\mathbf{p},\lambda\rangle \propto \delta_{\lambda,\pm s}|\mathbf{p},\lambda = \mp s\rangle.$$

Then

$$\pi_\pm S = \pm s\pi_\pm, \quad S\pi_\pm = \mp s\pi_\pm,$$
$$[\pi_\pm, S] = \pm 2s\pi_\pm.$$

Since **D** is translation invariant and rotates as a vector, we must have

$$[\mathbf{D},\mathbf{P}\cdot\mathbf{J}] = i\mathbf{P}\times\mathbf{D}$$

or

$$[\mathbf{D},S] = i\hat{P}\times\mathbf{D}.$$

This implies

$$s(\pi_+ - \pi_-)\mathbf{E}_1 - is(\pi_+ + \pi_-)\mathbf{E}_2$$
$$= i(1/2)(\pi_+ + \pi_-)\mathbf{E}_2 - (1/2)(\pi_+ - \pi_-)\mathbf{E}_1,$$

so $\pi_\pm$ are zero.

From Eqs. (4.3), (4.4), and

$$[T_1,T_2] = iS,$$

$$T_1^2 + T_2^2 = s(s+1) - S^2,$$

we get

$$\mathbf{D}\cdot(1/m)\hat{P}\times\mathbf{W} = -iBS + C[s(s+1) - S^2], \qquad (4.7)$$

$$\mathbf{D}\cdot(-1/m)\hat{P}\times(\hat{P}\times\mathbf{W}) = B[s(s+1) - S^2] + iCS. \qquad (4.8)$$

To the extent that these equations determine B and C, they imply that B and C are rotation invariant. Solving Eqs. (4.7) and (4.8) does determine B and C except for the value of $B + iC$ at the minimum eigenvalue $\lambda = -s$ of S and $B - iC$ at the maximum eigenvalue $\lambda = s$ which are superfluous because they do not occur in

$$\mathbf{D} = (1/2)(B + iC)(\mathbf{E}_1 - i\mathbf{E}_2)(T_1 + iT_2)$$
$$+ (1/2)(B - iC)(\mathbf{E}_1 + i\mathbf{E}_2)(T_1 - iT_2).$$

Therefore we can regard B and C as functions of P and S.

Now we transform everything back to the Foldy form. Using Eqs. (4.3) and (4.4), we see from Eq. (4.6) that with $\pi_\pm$ being zero we have

$$\mathbf{D} = -B(1/m)\hat{P}\times(\hat{P}\times\mathbf{W}) + C(1/m)\hat{P}\times\mathbf{W}.$$

From Eq. (2.6) we also see that in the Foldy form

$$(1/m)\hat{P}\times\mathbf{W} = \hat{P}\times\mathbf{S},$$

$$(1/m)\hat{P}\times(\hat{P}\times\mathbf{W}) = \hat{P}\times(\hat{P}\times\mathbf{S}),$$

so in the Foldy form we have

$$\mathbf{D} = -B\hat{P}\times(\hat{P}\times\mathbf{S}) + C\hat{P}\times\mathbf{S}, \qquad (4.9)$$

$$\mathbf{R} = \mathbf{Q} + \mathbf{D}. \qquad (4.10)$$

Since B and C are functions of P and S in the Moses form, from Eqs. (3.7) and (2.5) we see that in the Foldy form B and C are functions of P and $\hat{P}\cdot\mathbf{S}$.

Clearly

$$\mathbf{R}\cdot\mathbf{P} = \mathbf{Q}\cdot\mathbf{P}.$$

Since we assume that the components of **R** commute with each other and that

$$(1/i)[R_j,P_k] = \delta_{jk}, \qquad (4.11)$$

for $j,k = 1,2,3$, we have

$$(1/2)[\mathbf{R},\mathbf{Q}\cdot\mathbf{P}] = (1/i)[\mathbf{R},\mathbf{R}\cdot\mathbf{P}] = \mathbf{R}. \tag{4.12}$$

For a function f of $\mathbf{P}$,

$$(1/i)[f(\mathbf{P}),\mathbf{Q}\cdot\mathbf{P}] = -\mathbf{P}\cdot\nabla f(\mathbf{P}) = -P\,\partial f(\mathbf{P})/\partial P.$$

These equations determine the P dependence of B and C; they must be proportional to P^{-1}.

For nonzero spin there is more than one Hermitian operator $\mathbf{R}$ with commuting components that transforms as a position operator should for translations, rotations, and time reversal, because there are unitary transformations generated by $\hat{P}\cdot\mathbf{J}$ that preserve all these properties.[10,11] To get a unique position operator we have to consider Lorentz transformations.

From Eq. (2.3) we see that

$$\mathbf{K}\cdot\mathbf{P} = (1/2)(H\,\mathbf{Q}\cdot\mathbf{P} + \mathbf{Q}\cdot\mathbf{P}H).$$

Since $\mathbf{D}$ commutes with $\mathbf{P}$, it must commute with H, so $[\mathbf{R},H]$ is the same as $[\mathbf{Q},H]$, which is $i\mathbf{P}/H$. From this and Eq. (4.12) we get

$$(1/i)[\mathbf{R},\mathbf{K}\cdot\mathbf{P}] = (1/2)\{(\mathbf{P}/H)\mathbf{Q}\cdot\mathbf{P} + \mathbf{Q}\cdot\mathbf{P}(\mathbf{P}/H) + H\mathbf{R} + \mathbf{R}H\}.$$

Then using Eqs. (4.11), (4.10), and (2.3) we find that

$$(1/i)\sum_{k=1}^{3}[\mathbf{R},K_k]P_k = (1/i)[\mathbf{R},\mathbf{K}\cdot\mathbf{P}] - \mathbf{K} = (1/2)\{(\mathbf{P}/H)\mathbf{Q}\cdot\mathbf{P} + \mathbf{Q}\cdot\mathbf{P}(\mathbf{P}/H)\} + H\mathbf{D} - (H+m)^{-1}\mathbf{P}\times\mathbf{S}. \tag{4.13}$$

Since B and C are proportional to P^{-1}, we see from Eq. (4.9) that $\mathbf{D}$ is proportional to P^{-1}, and from Eq. (4.13) we see that if $\mathbf{D}$ is not zero, the $[\mathbf{R},K_k]$ contain terms proportional to P^{-2}. This allows us to conclude that $\mathbf{D}$ is zero, and thus complete our proof that $\mathbf{R}$ is the Newton–Wigner position operator, because this singular dependence of $\mathbf{D}$ and $[\mathbf{R},K_k]$ on P would contradict what we know of Lorentz transformations in the nonrelativistic limit, which we discuss next.

5. NONRELATIVISTIC LIMIT

We cannot say exactly what $[\mathbf{R},K_k]$ should be. The classical Poisson-bracket equation[12]

$$[\mathbf{R},K_k] = R_k[\mathbf{R},H], \tag{5.1}$$

is ambiguous for quantization because the position and velocity on the right do not commute. Lorentz transformations are not simple for the Newton–Wigner position operator and its localized wavefunctions.[6] A unitary transformation generated by $\mathbf{K}$ takes a wavefunction for time $t = 0$ for one frame to the corresponding wavefunction for $t' = 0$ for another frame, and a wave function localized in space at $t = 0$ is spread out in space at the different times t corresponding to $t' = 0$.

However, we need to consider Lorentz transformations only in the nonrelativistic limit, that is for states where the velocity is small. We can identify the velocity operator as being

$$(1/i)[\mathbf{R},H] = (1/i)[\mathbf{Q},H] = \mathbf{P}/H. \tag{5.2}$$

In fact we can show the velocity operator is $\mathbf{P}/H$ without even considering the position; we just need to assume it is a Hermitian operator with commuting components that transforms as a velocity should under the Poincaré group.[2,3] Thus small velocity means small P/m.

From Eq. (5.1) we expect $[\mathbf{R},K_k]$ to be small compared to R_k when the velocity is small. We found that if $\mathbf{D}$ is not zero, the $[\mathbf{R},K_k]$ contain terms proportional to P^{-2}. Since that could not give the right nonrelativistic limit, we conclude that $\mathbf{D}$ is zero.

For small velocities, the leading terms of the Poincaré generators in the Foldy form are the Galilei generators (1.1)–(1.3). In the same order of approximation, Lorentz transformations of position are described by Eq. (1.5). Indeed, if we neglect the terms of H with higher powers of P/m than those of Eq. (1.1), then to preserve the commutation relation

$$(1/i)[\mathbf{K},H] = \mathbf{P}, \tag{5.3}$$

which holds for both the Poincaré and Galilei groups, we must also neglect terms of $\mathbf{K}$ with higher powers than that of Eq. (1.3). We take $\mathbf{S}$ to be of the same order as $\mathbf{Q}\times\mathbf{P}$. Then $(H+m)^{-1}\mathbf{P}\times\mathbf{S}$ is of order P^2/m^2 compared to $m\mathbf{Q}$. With Eq. (1.3) for $\mathbf{K}$, we have

$$(1/i)[K_j,P_k] = m\delta_{jk}, \tag{5.4}$$

for $j,k = 1,2,3$. Since m is just a number, the unitary operators for infinitesimal translations and Lorentz transformations commute with each other to within a phase factor, which means the corresponding transformations of space–time coordinates commute. That happens only in an approximation where transformations of time coordinates are neglected. Then Lorentz transformations of position to first order in the transformation velocity $\boldsymbol{\beta}$ are the Galilei transformations

$$e^{i\boldsymbol{\beta}\cdot\mathbf{K}}\mathbf{R}(t)e^{-i\boldsymbol{\beta}\cdot\mathbf{K}} = \mathbf{R}'(t) = \mathbf{R}(t) - \boldsymbol{\beta}t. \tag{5.5}$$

For $t = 0$ we have

$$e^{i\boldsymbol{\beta}\cdot\mathbf{K}}\mathbf{R}e^{-i\boldsymbol{\beta}\cdot\mathbf{K}} = \mathbf{R}, \tag{5.6}$$

from which we obtain Eq. (1.5). For nonzero t, the transformations (5.5) result from the Hamiltonian (1.1) giving

$$(1/i)[\mathbf{R},H] = (1/i)[\mathbf{Q},H] = \mathbf{P}/m,$$
$$\mathbf{R}(t) = \mathbf{R}(t=0) + (\mathbf{P}/m)t, \tag{5.7}$$

and the commutation relation (5.4) giving

$$e^{i\boldsymbol{\beta}\cdot\mathbf{K}}(\mathbf{P}/m)e^{-i\boldsymbol{\beta}\cdot\mathbf{K}} = \mathbf{P}/m - \boldsymbol{\beta}. \tag{5.8}$$

For the Galilei group, using Eqs. (1.3) and (1.5), we showed that the position operator is $\mathbf{Q}$. Thus we expect that in the Foldy form the difference between the position operator and $\mathbf{Q}$ will be small when the velocity is small. We found that $\mathbf{D}$ is proportional to P^{-1}. That can not give the right nonrelativistic limit unless $\mathbf{D}$ is zero.

[1]E. P. Wigner, Ann. Math. **40**, 149 (1939); V. Bargmann and E. P. Wigner, Proc. Nat. Acad. Sci. U.S.A. **34**, 211 (1948).

[2]S. M. W. Ahmad and E. P. Wigner, Nuovo Cimento A **28**, 1 (1975).

[3]T. F. Jordan, J. Math. Phys. **18**, 608 (1977).

[4]E. Inonu and E. P. Wigner, Nuovo Cimento **9**, 705 (1952); V. Bargmann,

Ann. Math. **59**, 1 (1954); M. Hamermesh, Ann. Phys. (N.Y.) **9**, 518 (1960); J.-M. Levy-Leblond, J. Math. Phys. **4**, 776 (1963); J. Voisin, J. Math. Phys. **6**, 1519 (1965).
[5]T. F. Jordan, *Linear Operators for Quantum Mechanics* (Wiley, New York, 1969), Secs. 35 and 36.
[6]T. D. Newton and E. P. Wigner, Rev. Mod. Phys. **21**, 400 (1949). For some more recent work on position operators see C. Bachas, Phys. Lett. B **79**, 61 (1978); C. Bénard, J. Math. Phys. **16**, 710 (1975); B. Durand, Phys. Rev. D **14**, 1554 (1976); G. C. Hegerfeldt, Phys. Rev. D **10**, 3320 (1974); B. Jancewicz, J. Math. Phys. **18**, 2487 (1977); A. J. Kálnay, *Studies in the Foundations, Methodology and Philosophy of Science* (Springer, New York, 1971), Vol. 4, p. 93; A. J. Kálnay and P. L. Torres, Phys. Rev. D **9**, 1670 (1974); M. Lorente and P. Roman, J. Math. Phys. **15**, 70 (1974); R. F. O'Connell and E. P. Wigner, Phys. Lett. A **67**, 319 (1978); J. F. Perez and I. F. Wilde, Phys. Rev. D **16**, 315 (1977); B. K. Skagerstam, Int. J. Theor. Phys. **15**, 213 (1976).
[7]A. S. Wightman, Rev. Mod. Phys. **34**, 845 (1962).
[8]L. L. Foldy, Phys. Rev. **102**, 568 (1956); **122**, 275 (1961).
[9]H. E. Moses, J. Math. Phys. **9**, 2039 (1968).
[10]A. Galindo, Nuovo Cimento **37**, 413 (1965).
[11]T. F. Jordan, Am. J. Phys. **43**, 1089 (1975).
[12]D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963).

On six-dimensional canonical realizations of the so(4,2) algebra

Mircea Iosifescu and Horia Scutaru
Department of Fundamental Physics, Central Institute of Physics, Bucharest, P.O.B. 5206, Romania

(Received 25 May 1979; accepted for publication 19 October 1979)

On each six-dimensional symplectic manifold a coordinate-free realization of the so(4,2) algebra can be constructed, the generators of which satisfy the polynomial relations fulfilled by the so(4,2) generators associated with the Kepler problem. This realization contains as a particular case several realizations of so(4,2) known in the literature. An expression of the symplectic form on a six-dimensional symplectic manifold, in terms of the so(4,2) generators defined on this manifold, is obtained. In particular, on the six-dimensional orbit of the SO(4,2) group in so(4,2)* this symplectic form coincides with the symplectic form introduced by Kirillov Kostant and Souriau. The symplectic form is given a Darboux expression with the aid of three pairs of canonically conjugated variables, which are a generalization of the Delaunay elements defined in the Kepler problem.

1. INTRODUCTION

A canonical realization of a Lie algebra G is a realization of G as an algebra of functions defined on a symplectic manifold $\mathfrak{M}$, the Lie bracket being the Poisson bracket associated to the symplectic 2-form on $\mathfrak{M}$.[1,2]

Canonical realizations of Lie algebras are an essential tool in the study of classical dynamical systems which possess a group of symmetry, as a preliminary step for the canonical realizations of this symmetry group. One of the best-known examples of canonical realizations of Lie algebras is the algebra so(4) formed by the constants of motion of the nonrelativistic three-dimensional Kepler problem (the invariance algebra of the Kepler problem).[3–5]

Another interesting canonical realization related to the Kepler problem is its noninvariance so(4,2) algebra pointed out by Györgyi.[6–9] The generators of this algebra satisfy a number of homogeneous second-degree polynomial relations, called in Ref. 9 kinematic identities, which have been obtained in Refs. 6 and 9 as consequences of the equations related to the Kepler problem (trajectory and hodograph equations) and of the specific construction of the canonical realization considered.

The aim of the present paper is to investigate the information contained in these polynomial relations. From this analysis the following results have emerged:

(a) The set of polynomial relations satisfied by the so(4,2) Györgyi generators is constituted by three distinct ad-invariant subsets. The first two subsets have the property that under the adjoint action the polynomials of each subset transform into polynomials of the same subset. The third subset of relations contains only one element, the equation obtained by putting the Casimir invariant of the so(4,2) algebra equal to zero.

(b) Among these polynomial relations nine are functionally independent.

(c) Using properties (a) and (b), a coordinate-free realization of the so(4,2) algebra, induced from the realization of its so(4) subalgebra, has been obtained. It has also been proved that such an induced canonical realization of the so(4,2) algebra satisfying the above polynomial relations exists on any six-dimensional manifold (Sec. 3).

(d) The coordinate-free canonical realization of the so(4,2) algebra contains as particular cases the realizations associated in Ref. 6 to the Kepler problem as well as other canonical realizations of the so(4,2) algebra known in the literature.[10–13] The generators of each realization satisfies all three sets of polynomial relations; as a consequence of a theorem of Lie these realizations are all canonically equivalent.

This situation is of general validity for Lie algebras.[14] The inspection of a canonical realization R of a Lie algebra G reveals the presence of a set of functional relations among the generators. To each couple (G,R) corresponds a given set of such relations. As a consequence of the same theorem of Lie, two realizations, the generators of which satisfy the same set of functional relations, are canonically equivalent.

(e) In Sec. 5 a coordinate-free expression of the symplectic 2-form on a six-dimensional symplectic manifold $\mathfrak{M}$ is obtained in terms of the generators of the canonical realizations of so(4,2) which can be constructed on $\mathfrak{M}$, as stated above.

(f) A canonical expression for this symplectic form is obtained in Sec. 6. A system of coordinates called action-angle variables, in terms of which the symplectic form takes the canonical expression given by the Darboux theorem,[1] is defined. These coordinates, which are functions only of the generators of the so(4,2) algebra, represent a generalization of the Delaunay elements,[15] which are of use in the study of the Kepler problem. The Delaunay elements acquire thereby a definite Lie algebraic meaning.

(g) The Györgyi polynomial relations characterize a class of canonically equivalent six-dimensional canonical realizations of the so(4,2) algebra. Canonical equivalence is an equivalence relation which divides the set of canonical realizations of so(4,2) into equivalence classes: Each class may be represented modulo canonical equivalence by any of its elements. Nevertheless, a deeper insight into the whole class may be obtained by considering a representative element which is naturally connected with the structure of the so(4,2) algebra. We denote as "natural" a canonical realization of the so(4,2) algebra defined on an orbit of the SO(4,2) group in the dual of the so(4,2) algebra, in the sense of Kirillov,

 0022-2488/80/082033-13$1.00

Kostant, and Souriau (KKS).[16,17] It is known[18] that there is a one-to-one correspondence between the canonical realizations of a Lie algebra and the set of orbits of the corresponding group in the dual algebra. It may now be observed that the nine independent equations extracted from the three sets of polynomial relations [points (a), (b)] are precisely the equations of the six-dimensional SO(4,2)-orbit given in Refs. 19 and 20. This orbit possesses a natural symplectic structure given by the KKS 2-form ω. On the other side, we can particularize to this six-dimensional orbit the results described above. The coordinate functions in so(4,2)* form a realization of the so(4,2) algebra and a symplectic form $\widetilde{\omega}$ can be thus defined in terms of them, as stated in (e). It is proved in Sec. 8 that on the six-dimensional SO(4,2) orbit the two symplectic forms ω and $\widetilde{\omega}$ coincide. By proving that, we have obtained an explicit expression of the KKS symplectic form on the six-dimensional orbit of SO(4,2).

For the sake of consistency we have presented in Sec. 7 several elements from the KKS orbit theory, which have been given a local formulation refering to Lie algebras only.

2. THE CANONICAL REALIZATIONS OF THE so(4) AND so(4,2) ALGEBRAS RELATED TO THE KEPLER PROBLEM

A. Realization of the so(4) algebra[3-5]

Let us consider the three-dimensional Kepler problem, the Hamiltonian of which is ($m = Z = e = 1$)

$$H = p^2/2 - 1/q, \tag{2.1}$$

where[21] $q = (q_i q_i)^{1/2}$ and $p = (p_i p_i)^{1/2}$.

The set formed by the following six functions defined on $R^3 \times R^3$

$$L_i = (\mathbf{q}\times\mathbf{p})_i = \epsilon_{jki} q_j p_k \quad (i = 1,2,3) \tag{2.2}$$

(the components of the angular momentum) and

$$A_i = (1/\sqrt{-2H}\,)(\mathbf{L}\times\mathbf{p} + \mathbf{q}/q)_i \quad (i = 1,2,3) \tag{2.3}$$

with $H < 0$ (the components of the Laplace–Runge–Lenz vector) form a canonical realization of the so(4) algebra, i.e., they satisfy the so(4) Poisson bracket relations

$$\{L_i, L_j\} = \epsilon_{ijk} L_k,$$
$$\{L_i, A_j\} = \epsilon_{ijk} A_k, \quad \{A_i, A_j\} = \epsilon_{ijk} L_k, \tag{2.4}$$

where

$$\{\,,\,\} = \frac{\partial}{\partial q_i}\frac{\partial}{\partial p_i} - \frac{\partial}{\partial p_i}\frac{\partial}{\partial q_i} \tag{2.5}$$

is the standard Poisson bracket in $R^3 \times R^3$.

The generators of this realization of the so(4) Lie algebra satisfy also the homogeneous polynomial identity

$$\mathbf{L}\cdot\mathbf{A} = L_i A_i = 0. \tag{2.6}$$

The elements of this so(4) Lie algebra are constants of the motion for the Kepler Hamiltonian (2.1), i.e., $\{L_i, H\} = \{A_i, H\} = 0$. This may be seen from the relation

$$2H = -\frac{1}{L_i L_i + A_i A_i}, \tag{2.7}$$

where $L_i L_i + A_i A_i$ is the so(4) Casimir invariant.

Canonical realizations of the so(4,1) algebra have been obtained in Ref. 22.

B. Realization of the so(4,2) Lie algebra[6-9]

The set of 15 functions defined on $R^3 \times R^3$ by the equalities (2.2), (2.3), and by the following nine equalities:

$$B_i \equiv q p_i \cos(\sqrt{-2H}\,u) + (1/\sqrt{-2H}\,) \times (q_i/q - u p_i)\sin(\sqrt{-2H}\,u), \tag{2.8}$$

$$C_i \equiv -q p_i \sin(\sqrt{-2H}\,u) + (1/\sqrt{-2H}\,) \times (q_i/q - u p_i)\cos(\sqrt{-2H}\,u), \quad (i = 1,2,3), \tag{2.9}$$

$$B_4 \equiv -(1/\sqrt{-2H}\,)(qp^2 - 1)\cos(\sqrt{-2H}\,u) - u\sin(\sqrt{-2H}\,u), \tag{2.10}$$

$$C_4 \equiv (1/\sqrt{-2H}\,)(qp^2 - 1)\sin(\sqrt{-2H}\,u) - u\cos(\sqrt{-2H}\,u), \tag{2.11}$$

$$M \equiv 1/\sqrt{-2H}, \tag{2.12}$$

$$u = q_i p_i \tag{2.13}$$

form a canonical realization of the so(4,2) algebra; i.e., these functions satisfy the so(4,2) Poisson bracket relations given in Table I. [We see from this table that each set of functions B_i ($i = 1,2,3$) and C_i ($i = 1,2,3$) forms a vector with respect to O(3). We shall therefore use also the notations $\mathbf{B}$ and C instead of (B_1, B_2, B_3) and (C_1, C_2, C_3) respectively.]

The generators of this realization of the so(4,2) algebra satisfy, in addition to relation (2.6), the following homogeneous polynomial identities [Ref. 6, relations (3.26), (3.27) or (3.30), (3.31)]:

$$\mathbf{L}\cdot\mathbf{B} = 0, \tag{2.14}$$
$$\mathbf{L}\cdot\mathbf{C} = 0, \tag{2.15}$$
$$\mathbf{A}\times\mathbf{B} - B_4\mathbf{L} = 0, \tag{2.16}$$
$$\mathbf{A}\times\mathbf{C} - C_4\mathbf{L} = 0, \tag{2.17}$$
$$\mathbf{B}\times\mathbf{C} + M\mathbf{L} = 0, \tag{2.18}$$
$$M\mathbf{A} + C_4\mathbf{B} - B_4\mathbf{C} = 0 \tag{2.19}$$

and

$$B_\alpha C_\alpha \equiv \mathbf{B}\cdot\mathbf{C} + B_4 C_4 = 0, \tag{2.20}$$
$$\mathbf{C}\cdot\mathbf{A} - B_4 M = 0, \tag{2.21}$$
$$\mathbf{A}\cdot\mathbf{B} + C_4 M = 0, \tag{2.22}$$
$$B_\alpha B_\alpha - C_\alpha C_\alpha \equiv \mathbf{B}^2 + B_4^2 - \mathbf{C}^2 - C_4^2 = 0, \tag{2.23}$$
$$\mathbf{C}^2 + \mathbf{A}^2 - B_4^2 - M^2 = 0, \tag{2.24}$$
$$\mathbf{A}^2 + \mathbf{B}^2 - C_4^2 - M^2 = 0, \tag{2.25}$$
$$\mathbf{A}\times\mathbf{L} + B_4\mathbf{B} + C_4\mathbf{C} = 0, \tag{2.26}$$
$$\mathbf{B}\times\mathbf{L} + B_4\mathbf{A} - M\mathbf{C} = 0, \tag{2.27}$$
$$\mathbf{C}\times\mathbf{L} + C_4\mathbf{A} + M\mathbf{B} = 0, \tag{2.28}$$
$$\hat{T} \equiv \mathbf{A}\otimes\mathbf{A} - \mathbf{L}\otimes\mathbf{L} - \mathbf{B}\otimes\mathbf{B} - \mathbf{C}\otimes\mathbf{C} + \hat{1}(\mathbf{L}^2 - \mathbf{A}^2 + B_4^2 + C_4^2) = 0 \tag{2.29}$$
$$\hat{U} \equiv \mathbf{L}\otimes\mathbf{L} - \mathbf{A}\otimes\mathbf{A} + \mathbf{B}\otimes\mathbf{B} + \mathbf{C}\otimes\mathbf{C} - \hat{1}(\mathbf{L}^2 + \mathbf{B}^2 + B_4^2 - M^2) = 0 \tag{2.31}$$
$$\hat{V} \equiv \mathbf{L}\otimes\mathbf{L} - \mathbf{A}\otimes\mathbf{A} + \mathbf{B}\otimes\mathbf{B} + \mathbf{C}\otimes\mathbf{C} - \hat{1}(\mathbf{L}^2 + \mathbf{C}^2 + C_4^2 + M^2) = 0 \tag{2.31}$$

where

$\mathbf{X}\cdot\mathbf{Y}$ denotes the scalar product,

TABLE I. Lie multiplication table of the so(4,2) algebra. The multiplication law is

	Y
X	$\{X,Y\}$

	L_1	L_2	L_3	A_1	A_2	A_3	B_1	B_2	B_3	C_1	C_2	C_3	B_4	C_4	M
L_1	0	L_3	$-L_2$	0	A_3	$-A_2$	0	B_3	$-B_2$	0	C_3	$-C_2$	0	0	0
L_2	$-L_3$	0	L_1	$-A_3$	0	A_1	$-B_3$	0	B_1	$-C_3$	0	C_1	0	0	0
L_3	L_2	$-L_1$	0	A_2	$-A_1$	0	B_2	$-B_1$	0	C_2	$-C_1$	0	0	0	0
A_1	0	A_3	$-A_2$	0	L_3	$-L_2$	B_4	0	0	C_4	0	0	$-B_1$	$-C_1$	0
A_2	$-A_3$	0	A_1	$-L_3$	0	L_1	0	B_4	0	0	C_4	0	$-B_2$	$-C_2$	0
A_3	A_2	$-A_1$	0	L_2	$-L_1$	0	0	0	B_4	0	0	C_4	$-B_3$	$-C_3$	0
B_1	0	B_3	$-B_2$	$-B_4$	0	0	0	$-L_3$	L_2	$-M$	0	0	$-A_1$	0	$-C_1$
B_2	$-B_3$	0	B_1	0	$-B_4$	0	L_3	0	$-L_1$	0	$-M$	0	$-A_2$	0	$-C_2$
B_3	B_2	$-B_1$	0	0	0	$-B_4$	$-L_2$	L_1	0	0	0	$-M$	$-A_3$	0	$-C_3$
C_1	0	C_3	$-C_2$	$-C_4$	0	0	M	0	0	0	$-L_3$	$-L_2$	0	$-A_1$	B_1
C_2	$-C_3$	0	C_1	0	$-C_4$	0	0	M	0	L_3	0	$-L_1$	0	$-A_2$	B_2
C_3	C_2	$-C_1$	0	0	0	$-C_4$	0	0	M	$-L_2$	L_1	0	0	$-A_3$	B_3
B_4	0	0	0	B_1	B_2	B_3	A_1	A_2	A_3	0	0	0	0	$-M$	$-C_4$
C_4	0	0	0	C_1	C_2	C_3	0	0	0	A_1	A_2	A_3	M	0	B_4
M	0	0	0	0	0	0	C_1	C_2	C_3	$-B_1$	$-B_2$	$-B_3$	C_4	$-B_4$	0

$\mathbf{X}\times\mathbf{Y}$ denotes the vector product,
$\mathbf{X}\otimes\mathbf{Y}$ denotes the tensor product,

of the vectors $\mathbf{X}$ and $\mathbf{Y}$ and where $\hat{1}$ denotes the Kronecker tensor.

It may also be observed that the generators of this realization of the so(4,2) algebra satisfy the identity

$$L_iL_i + A_i A_i + M^2 - B_\alpha B_\alpha - C_\alpha C_\alpha = 0, \qquad (2.32)$$

i.e., the quadratic invariant of the so(4,2) algebra vanishes.

3. A COORDINATE-FREE CANONICAL REALIZATION OF THE so(4,2) ALGEBRA

In this section we shall construct a coordinate-free canonical realization of the so(4,2) algebra, the generators of which satisfy the polynomial identities (2.6), (2.14)–(2.32), which are verified by the genrators (2.2), (2.3), (2.8)–(2.12) associated with the Kepler problem.

Let us first examine some properties of the polynomial relations (2.6), (2.14)–(2.19), and (2.20)–(2.31). The following proposition shows that the two sets of polynomial relations (2.6), (2.14)–(2.19), and (2.20)–(2.31) among the generators of an so(4,2) algebra may be obtained by applying the adjoint action to one relation of the set.

Proposition 1 (cf. Refs. 23, 24): Let $\mathfrak{M}$ be a symplectic manifold and let the following functions, defined on $\mathfrak{M}$, L_i, A_i $(i = 1,2,3)$, B_α, C_α $(\alpha = 1,...,4)$, and M, be the generators of an so(4,2) algebra with respect to the Poisson bracket (induced on $\mathfrak{M}$ by the symplectic 2-form associated with $\mathfrak{M}$), i.e., let them satisfy the Poisson bracket relations given in Table I.

(1) The set of functions formed by the lhs of relations (2.6), (2.14)–(2.19) is closed with respect to the adjoint action of the so(4,2) algebra. In particular, if relation (2.6) is valid, then relations (2.14)–(2.19) are valid.

(2) The set of functions formed by the lhs of relations (2.20)–(2.31) is closed with respect to the adjoint action of the so(4,2) algebra. In particular, if relation (2.20) is valid, then relations (2.21)–(2.31) are valid.

Proof: By taking the Poisson bracket between the so(4,2) generators and the function L_iA_i, we obtain functions which vanish due to relation (2.6). Taking now the Poisson bracket between the so(4,2) generators and the functions obtained in the previous step, we obtain again vanishing functions. The iteration of the process leads to all the relations (2.14)–(2.19). The system of functions formed by the lhs of relations (2.6), (2.14)–(2.19) can be reached by starting from any function of the system; this system is thus closed vs. the adjoint action. These facts are shown in Table II.

A similar procedure leads from relation (2.20) to the relations (2.21)–(2.31). The system of functions (2.20)–(2.31) is also closed vs. the adjoint action. These properties are shown inTable III. □

The following two propositions show that the polynomial relations (2.6), (2.14)–(2.31) are not algebraically independent:

Proposition 2: If relations (2.18), (2.19) are satisfied, then relations (2.6), (2.14)–(2.17) are satisfied.

Proof: Scalar multiplication of (2.18) with $\mathbf{B}$ and $\mathbf{C}$ leads to (2.14) and (2.15), respectively. Scalar multiplication of (2.19) with $\mathbf{L}$ and use of (2.14), (2.15) leads to (2.6). Vector multiplication of (2.19) with $\mathbf{B}$ and $\mathbf{C}$ and use of (2.18) leads to (2.16) and (2.17) respectively. □

Proposition 3: If relations (2.18), (2.19), (2.20), (2.23), and (2.32) are verified, then (2.6), (2.14)–(2.31) are verified.

Proof: From (2.18)–(2.20), (2.23), and (2.32), we get

$$M^2 = \mathbf{L}^2 + \mathbf{A}^2 = B_\alpha B_\alpha = C_\alpha C_\alpha, \qquad (3.1)$$

$$\mathbf{A}^2 = B_4^2 + C_4^2, \qquad (3.2)$$

and from these relations, we obtain (2.24)–(2.25). Vector multiplication of $M\mathbf{L} = \mathbf{C}\times\mathbf{B}$ with $M\mathbf{A} = B_4\mathbf{C} - C_4\mathbf{B}$ (or with $C_4\mathbf{B} = B_4\mathbf{C} - M\mathbf{A}$ or $B_4\mathbf{C} = M\mathbf{A} + C_4\mathbf{B}$) and use of (2.20) leads to (2.26) or (2.27) or (2.28), respectively.

The nondiagonal components of Eqs. (2.29), (2.30), (2.31) are verified if B_i, C_i, A_i, L_i are replaced by their values taken from (2.28), (2.27), (2.19), and (2.18), respectively.

TABLE II. Adjoint action of the so(4,2) algebra on the set of polynomials (2.6), (2.14)–(2.19). The multiplication law is

	X_i
$P_j(X_1,...,X_{15})$	$\{X_i, P_j(X_1,...,X_{15})\}$

The following notations have been used:

(1) If $\mathbf{X}(X_1,X_2,X_3)$, $\mathbf{Y}(Y_1,Y_2,Y_3)$, $\mathbf{Z}(Z_1,Z_2,Z_3)$ are vectors with respect to O(3), then $\{\mathbf{X},\mathbf{Y}\}=\mathbf{Z}$ stands for the set of equations $\{X_i,Y_j\}=\epsilon_{ijk}Z_k$ $(i,j=1,2,3)$.

(2) If $\mathbf{X}$, $\mathbf{Y}$ are vectors and Z is a scalar, then $\{\mathbf{X},\mathbf{Y}\}=Z$ stands for $\{X_i,Y_j\}=\delta_{ij}Z$ $(i,j=1,2,3)$.

(3) If $\mathbf{X}$, $\mathbf{Z}$ are vectors and Y is a scalar, then $\{\mathbf{X},Y\}=\mathbf{Z}$ stands for $\{X_i,Y\}=Z_i$ $(i=1,2,3)$.

The polynomials $B_4C_i-C_4B_i-MA_i$ $(i=1,2,3)$, $(\mathbf{C}\times\mathbf{B}-M\mathbf{L})_i$ $(i=1,2,3)$, $(\mathbf{C}\times\mathbf{A}+C_4\mathbf{L})_i$ $(i=1,2,3)$, $(\mathbf{A}\times\mathbf{B}-B_4\mathbf{L})_i$ $(i=1,2,3)$, $-L_iC_i$, L_iB_i, $-L_iA_i$ represent, as it may be seen from the Table, the components of a vector in so(4,2), i.e., denoting these polynomials by $P_j(X_1,...,X_{15})$ $(j=1,...,15)$, they satisfy the relation $\{X_i,P_j\}=\Sigma_k C_{ijk}P_k$, where X_i $(i=1,...,15)$ and c_{ijk} are the generators and structure constants of the so(4,2) algebra. Vectors of this type can be constructed for any so(n,2) algebra. Their expression for $n=4$ has been given in Ref. 23. These vectors represent the pseudo-orthogonal analog of the vector defined in Ref. 30, relation (7.15), for orthogonal algebras.

	$\mathbf{L}$	$\mathbf{A}$	$\mathbf{B}$	$\mathbf{C}$	B_4	C_4	M
$B_4\mathbf{C}-C_4\mathbf{B}-M\mathbf{A}$	$B_4\mathbf{C}-C_4\mathbf{B}-M\mathbf{A}$	$\mathbf{C}\times\mathbf{B}-M\mathbf{L}$	$\mathbf{C}\times\mathbf{A}+C_4\mathbf{L}$	$\mathbf{A}\times\mathbf{B}-B_4\mathbf{L}$	0	0	0
$\mathbf{C}\times\mathbf{B}-M\mathbf{L}$	$\mathbf{C}\times\mathbf{B}-M\mathbf{L}$	$B_4\mathbf{C}-C_4\mathbf{B}-M\mathbf{A}$	$\mathbf{L}\cdot\mathbf{C}$	$-\mathbf{L}\cdot\mathbf{B}$	$\mathbf{C}\times\mathbf{A}-C_4\mathbf{L}$	$\mathbf{A}\times\mathbf{B}-B_4\mathbf{L}$	0
$\mathbf{C}\times\mathbf{A}+C_4\mathbf{L}$	$\mathbf{C}\times\mathbf{A}+C_4\mathbf{L}$	$-\mathbf{L}\cdot\mathbf{C}$	$-(B_4\mathbf{C}-C_4\mathbf{B}-M\mathbf{A})$	$-\mathbf{L}\cdot\mathbf{A}$	$\mathbf{C}\times\mathbf{B}-M\mathbf{L}$	0	$\mathbf{A}\times\mathbf{B}-B_4\mathbf{L}$
$\mathbf{A}\times\mathbf{B}-B_4\mathbf{L}$	$\mathbf{A}\times\mathbf{B}-B_4\mathbf{L}$	$\mathbf{L}\cdot\mathbf{B}$	$\mathbf{L}\cdot\mathbf{A}$	$-(B_4\mathbf{C}-C_4\mathbf{B}-M\mathbf{A})$	0	$\mathbf{C}\times\mathbf{B}-M\mathbf{L}$	$-(\mathbf{C}\times\mathbf{A}+C_4\mathbf{L})$
$-\mathbf{L}\cdot\mathbf{C}$	0	$-(\mathbf{C}\times\mathbf{A}+C_4\mathbf{L})$	$-(\mathbf{C}\times\mathbf{B}-M\mathbf{L})$	0	0	$-\mathbf{L}\cdot\mathbf{A}$	$\mathbf{L}\cdot\mathbf{B}$
$\mathbf{L}\cdot\mathbf{B}$	0	$-(\mathbf{A}\times\mathbf{B}-B_4\mathbf{L})$	0	$-(\mathbf{C}\times\mathbf{B}-M\mathbf{L})$	$\mathbf{L}\cdot\mathbf{A}$	0	$\mathbf{L}\cdot\mathbf{C}$
$-\mathbf{L}\cdot\mathbf{A}$	0	0	$-(\mathbf{A}\times\mathbf{B}-B_4\mathbf{L})$	$\mathbf{C}\times\mathbf{A}+C_4\mathbf{L}$	$-\mathbf{L}\cdot\mathbf{B}$	$-\mathbf{L}\cdot\mathbf{C}$	0

TABLE III. Adjoint action of the so(4,2) algebra on the set of polynomials (2.20)–(2.31). The multiplication law is the same as in Table II. We have denotd by $\hat{T}$, $\hat{U}$, $\hat{V}$ the tensors (2.29), (2.30), and (2.31), respectively. In order to avoid too large dimensions of this Lie multiplication table, the following notations have been used in addition to those used in Table II:

(4) If $\mathbf{X}(X_1,X_2,X_3)$, $\mathbf{Y}(Y_1,Y_2,Y_3)$ are vectors and $\hat{Z}(Z_{11},Z_{12},...,Z_{33})$ is a tensor with respect to O(3), then $\{\mathbf{X},\mathbf{Y}\}=\hat{Z}$ stands for the set of equations $\{X_i,Y_j\}=Z_{ij}$ $(i,j=1,2,3)$.

(5) If $\mathbf{X}$ is a vector and $\hat{Y}$, $\hat{Z}$ are tensors, then the equality $\{\mathbf{X},\hat{Y}\}=\hat{Z}$ stands for the set of equations $\{X_i,Y_{jk}\}=\epsilon_{ijm}Z_{km}+\epsilon_{ikm}Z_{jm}$ $(i,j,k=1,2,3)$.

(6) If $\mathbf{X}$, $\mathbf{Z}$ are vectors and $\hat{Y}$ is a tensor, then:

(i) $\{\mathbf{X},\hat{Y}\}=\mathbf{Z}$ stands for

$\{X_i,Y_{jk}\}=\delta_{ij}Z_k+\delta_{ik}Z_j+2\delta_{jk}Z_i$ $(i,j,k=1,2,3)$.

(ii) $\{\mathbf{X},\hat{Y}\}=\mathbf{Z}$ (*) stands for

$\{X_i,Y_{jk}\}=\delta_{ij}Z_k+\delta_{ik}Z_j$ $(i,j,k=1,2,3)$.

(7) If X, Z are scalars and $\hat{Y}$ is a tensor, then $\{X,\hat{Y}\}=Z$ stands for $\{X,Y_{ij}\}=\delta_{ij}Z$ $(i,j=1,2,3)$

	$\mathbf{L}$	$\mathbf{A}$	$\mathbf{B}$	$\mathbf{C}$	B_4	C_4	M
$\mathbf{BC}+B_4C_4$	0	0	$\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}-\mathbf{MB}$	$\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-\mathbf{MC}$	$\mathbf{CA}-B_4\mathbf{M}$	$\mathbf{AB}+C_4\mathbf{M}$	$\mathbf{B}^2+B_4^2-\mathbf{C}-C_4^2$
$\mathbf{CA}-B_4M$	0	$\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B}$	0	$-(\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C})$	$\mathbf{BC}+B_4C_4$	$\mathbf{C}^2+\mathbf{A}^2-B_4^2-M^2$	$-(\mathbf{AB}+C_4M)$
$\mathbf{AB}+C_4M$	0	$\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-M\mathbf{C}$	$-(\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C})$	0	$\mathbf{A}^2+\mathbf{B}^2-C_4^2-M^2$	$\mathbf{BC}+B_4C_4$	$\mathbf{CA}-B_4M$
$\mathbf{B}^2+B_4^2-\mathbf{C}^2-C_4^2$	0	0	$-2(\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-M\mathbf{C})$	$2(\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B})$	$2(\mathbf{AB}+C_4M)$	$-2(\mathbf{CA}-B_4M)$	$4(\mathbf{BC}+B_4C_4)$
$\mathbf{C}^2+\mathbf{A}^2-B_4^2-M^2$	0	$2(\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C})$	0	$-2(\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B})$	$2(\mathbf{AB}+C_4M)$	$4(\mathbf{CA}+B_4M)$	$-2(\mathbf{BC}+B_4C_4)$
$\mathbf{A}^2+\mathbf{B}^2-C_4^2-M^2$	0	$2(\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C})$	$-2(\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-M\mathbf{C})$	0	$4(\mathbf{AB}+C_4M)$	$2(\mathbf{CA}+B_4M)$	$2(\mathbf{BC}+B_4C_4)$
$\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C}$	$\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C}$	$\hat{T}$	$-(\mathbf{AB}+C_4M)$	$-(\mathbf{CA}-B_4M)$	$\mathbf{B}\times\mathbf{L}+\mathbf{B}_4-\mathbf{A}-M\mathbf{C}$	$\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B}$	0
$\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}+M\mathbf{C}$	$\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-M\mathbf{C}$	$-(\mathbf{AB}+C_4M)$	$\hat{U}$	$-(\mathbf{BC}+B_4C_4)$	$\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C}$	0	$\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B}$
$\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B}$	$\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B}$	$-(\mathbf{CA}-B_4M)$	$-(\mathbf{BC}+B_4C_4)$	$\hat{V}$	0	$\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C}$	$-(\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-M\mathbf{C})$
$\hat{T}$	$\hat{T}$	$-(\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C})$	$-(\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-M\mathbf{C})$	$-(\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}-M\mathbf{B})$ (*)	$-2(\mathbf{AB}+C_4M)$ (*)	$-2(\mathbf{CA}-B_4M)$	0
$\hat{U}$	$\hat{U}$	$\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C}$ (*)	$\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-M\mathbf{C}$ (*)	$\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B}$ (*)	$-2(\mathbf{AB}+C_4M)$	0	$-2(\mathbf{BC}+B_4C_4)$
$\hat{V}$	$\hat{V}$	$\mathbf{A}\times\mathbf{L}+B_4\mathbf{B}+C_4\mathbf{C}$ (*)	$\mathbf{B}\times\mathbf{L}+B_4\mathbf{A}-M\mathbf{C}$ (*)	$\mathbf{C}\times\mathbf{L}+C_4\mathbf{A}+M\mathbf{B}$	0	$-2(\mathbf{CA}-B_4M)$	$2(\mathbf{BC}+B_4C_4)$

Similarly, introducing instead of one factor in each square $B_i^2, C_i^2, A_i^2, L_i^2$ the above values of B_i, C_i, A_i, L_i, we obtain that the diagonal components of Eqs. (2.29), (2.30), (2.31) are verified. Finally, relations (2.21) and (2.22) result from the scalar multiplication of (2.27) and (2.28) with **B** and **C** respectively and use of (2.20).

The following proposition states the existence of a coordinate-free realization of the so(4,2) algebra, induced by a realization of its so(4) subalgebra.

Proposition 4 (cf. Refs. 24, 25): Let $\mathfrak{M}$ be a symplectic manifold and let us assume that there exists on $\mathfrak{M}$ a canonical realization of the so(4) algebra (with respect to the Poisson bracket associated with the symplectic 2-form on $\mathfrak{M}$), the generators L_i, A_i ($i = 1,2,3$) of which satisfy the Poisson bracket relations (2.4) and the polynomial relation (2.6).

There exists then on $\mathfrak{M}$ a canonical realization of an so(4,2) algebra with the following properties

(i) It contains the so(4) algebra generated by L_i, A_i ($i = 1,2,3$) as a subalgebra.

(ii) Its generators satisfy the so(4,2) Lie bracket relations from Table I and the polynomial relations (2.20) and (2.32).

The generators of this realization of the so(4,2) algebra are

$$\mathbf{B} = \eta\left[(L^2 + A^2)^{1/2}\cos\psi\frac{\mathbf{A}}{A} + \sin\psi\,\mathbf{L}\times\frac{\mathbf{A}}{A}\right], \tag{3.3}$$

$$B_4 = \eta A\sin\psi, \tag{3.4}$$

$$\mathbf{C} = -\epsilon\eta\left[-(L^2 + A^2)^{1/2}\sin\psi\frac{\mathbf{A}}{A} + \cos\psi\,\mathbf{L}\times\frac{\mathbf{A}}{A}\right], \tag{3.5}$$

$$C_4 = -\epsilon\eta A\cos\psi, \tag{3.6}$$

$$M = \epsilon(L^2 + A^2)^{1/2}, \tag{3.7}$$

where $\epsilon = \pm 1$, $\eta = \pm 1$, $L = (L_i L_i)^{1/2}$, $A = (A_i A_i)^{1/2}$, and ψ is a scalar function

$$\{\mathbf{L},\psi\} = 0, \tag{3.8}$$

satisfying the equations

$$\{\psi,\mathbf{A}\} = \frac{M}{A^2}\mathbf{A}. \tag{3.9}$$

Proof: We shall prove the existence of a coordinate-free realization of so(4,2) by showing that conditions (i) and (ii) lead to the explicit expressions (3.3)–(3.7) of the generators B_α, C_α, and M as functions of **L** and **A** which are submitted only to (2.6) and of the function ψ which can always be determined as a solution of (3.8), (3.9).

Let us first remark that (Proposition 1) conditions (2.6) and (2.20) lead to (2.14)–(2.31). On the other side (Proposition 3) the conditions (2.6), (2.14)–(2.32) are algebraically dependent and can be deduced from the subset of conditions (2.18)–(2.20), (2.23), and (2.32). Multiplication of (2.26) with B_4 and of (2.19) with C_4 and addition of the results leads to

$$(B_4^2 + C_4^2)\mathbf{B} = -C_4 M\mathbf{A} + B_4\mathbf{L}\times\mathbf{A}. \tag{3.10}$$

Multiplication of (2.26) and (2.19) with C_4 and $-B_4$, respectively, and addition of the results gives

$$(B_4^2 + C_4^2)\mathbf{C} = B_4 M\mathbf{A} + C_4\mathbf{L}\times\mathbf{A}. \tag{3.11}$$

Taking (3.1) and (3.2) into account and introducing the notations

$$C_4/A = \cos[(\epsilon - \eta)\pi/2 + \epsilon\psi],$$
$$B_4/A = \sin[(\epsilon - \eta)\pi/2 + \epsilon\psi], \tag{3.12}$$
$$\text{if } M = \epsilon(L^2 + A^2)^{1/2},$$

i.e.,

$$C_4/A = \cos\psi, \quad B_4/A = \sin\psi, \tag{3.13a}$$

$$-C_4/A = \cos(\psi + \pi) = -\cos\psi,$$
$$B_4/A = \sin(\psi + \pi) = -\sin\psi, \tag{3.13b}$$
$$\text{if } M = +(L^2 + A^2)^{1/2}$$

$$-C_4/A = \cos(\pi - \psi) = -\cos\psi, \tag{3.13a'}$$
$$B_4/A = \sin(\pi - \psi) = \sin\psi,$$

$$-C_4/A = \cos(-\psi) = \cos\psi,$$
$$B_4/A = \sin(-\psi) = -\sin\psi, \tag{3.13b'}$$

$$\text{if } M = -(L^2 + A^2)^{1/2},$$

we obtain from (3.10) relation (3.3) and from (3.11) relation (3.5). Relations (3.4), (3.6), and (3.7) are equivalent to the notation (3.12).

The verification of the so(4,2) Poisson bracket relations (Table I) by the generators L_i, A_i and the generators B_α, C_α, M (3.3)–(3.7) is equivalent to the validity of relations (3.8) and (3.9). Relations (3.8) and (3.9) may indeed be obtained from the Poisson bracket relations $\{\mathbf{L},B_4\} = 0$ and $\{B_4,\mathbf{A}\} = \mathbf{B}$, respectively. Conversely, if the expressions (3.3)–(3.7) of the nine generators B_α, C_α, ($\alpha = 1,2,3,4$), M are used and if the function ψ satisfies relations (3.8) and (3.9), then for any **L** and **A** satisfying (2.4) and (2.6) all so(4,2) Poisson bracket relations of Table I are verified. The proof of the validity of these Poisson bracket relations is a matter of straightforward calculation requiring only the properties (2.4) and (2.6) of the generators **L** and **A** and (3.8), (3.9) of the function ψ. □

Propositions 1 and 3 lead to

Corollary 1: One any symplectic manifold $\mathfrak{M}$ on which a canonical realization of the so(4) algebra exists, with the generators **L**, **A** satisfying $\mathbf{L}\cdot\mathbf{A} = 0$, there exists a canonical realization of the so(4,2) algebra, containing this so(4) algebra as a subalgebra, and such that its generators satisfy all the conditions (2.14)–(2.32) satisfied by the generators of the realization (2.2), (2.3), (2.8)–(2.12) associated with the Kepler problem.

The restrictive condition of the existence on $\mathfrak{M}$ of a canonical realization of the so(4) algebra, satisfying (2.4), (2.6) is in fact always satisfied if $\mathfrak{M}$ is six-dimensional. Indeed, let ω be the symplectic form on $\mathfrak{M}$; we have

Proposition 5: Given a six-dimensional symplectic manifold $(\mathfrak{M},\omega)$ and a point $m\in\mathfrak{M}$, there exists a neighborhood U of m and a set of functions L_i, A_i ($i = 1,2,3$) defined on U which satisfy the relation $L_i A_i = 0$ and which generate a canonical realization of the so(4) algebra. The functions L_i, A_i are defined by the following relations:

$$L_i = (\mathbf{q}\times\mathbf{p})_i \quad (i = 1,2,3),$$

$$A_i = A\left(\frac{q_i}{q}\cos\xi + \frac{(\mathbf{L}\times\mathbf{q})_i}{Lq}\sin\xi\right) \quad (i = 1,2,3), \tag{3.14}$$

in which the scalar functions $A(q,L,u)$, $\xi(q,L,u)$ satisfy the equation

$$\left(\frac{\partial A^2}{\partial q}\frac{\partial \xi}{\partial u} - \frac{\partial A^2}{\partial u}\frac{\partial \xi}{\partial q}\right)q = 2L + \frac{\partial A^2}{\partial L}. \tag{3.15}$$

Proof: From the Darboux theorem[1,2] there exists a chart (U,φ), $m\in U$, such that $\varphi(m) = 0$,

$$\varphi(u) = (q_1(u),q_2(u),q_3(u);p_1(u),p_2(u),p_3(u))$$

and (3.16)

$$\omega_{|U} = dq_i \wedge dp_i.$$

The Poisson bracket associated with the 2-form (3.16) is (2.5); with respect to it the functions L_i $(i = 1,2,3)$ given by (2.2) generate an so(3) algebra, i.e., satisfy $\{\mathbf{L},\mathbf{L}\} = \mathbf{L}$. The general expression of a set of functions A_1, A_2, A_3, which are the components of a vector with respect to O(3) ($\{\mathbf{L},\mathbf{A}\} = \mathbf{A}$), which is orthogonal to $\mathbf{L}$ ($\mathbf{LA} = 0$), is given by relation (3.14), where ξ is a scalar function ($\{\mathbf{L},\xi\} = 0$); the condition for the validity of $\{\mathbf{A},\mathbf{A}\} = \mathbf{L}$ with $\mathbf{A}$ given by (3.14) is (3.15), which is always solvable. Hence all the Lie bracket relations (2.4) which define the so(4) algebra are satisfied. □

The solvability of Eqs. (3.15) and of Eq. (3.9) for any $\mathbf{A}$ leads to

Corollary 2: On any six-dimensional manifold $(\mathfrak{M},\omega)$ a canonical realization of the so(4,2) algebra can be determined, the generators of which satisfy the Poisson bracket relations of Table I and the polynomial relations (2.18)–(2.20), (2.23), (2,32) [or, what amounts to the same, relations (2.6), (2.14)–(2.32)]. □

Equations (3.6) and (3.7) satisfied by the function ψ lead to the following properties of this function.

Proposition 6: The function ψ is determined only modulo the addition of a function of the so(4) Casimir operator $\mathbf{L}^2 + \mathbf{A}^2$. □

Proposition 7: The function ψ is functionally independent on the functions L_i and A_i.

Proof: Let us suppose the contrary, i.e., that

$$\psi = \psi(L_1,L_2,L_3,A_1,A_2,A_3). \tag{3.17}$$

Relations (3.8) become

$$0 = \frac{\partial\psi}{\partial L_i}\{L_i,L_j\} + \frac{\partial\psi}{\partial A_i}\{A_i,L_j\}$$
$$= \epsilon_{ijk}\left(\frac{\partial\psi}{\partial L_i}L_k + \frac{\partial\psi}{\partial A_i}A_k\right) \quad (j = 1,2,3). \tag{3.18}$$

Multiplication with L_j and summation with respect to j give

$$(\mathbf{L}\times\mathbf{A})_i\frac{\partial\psi}{\partial A_i} = 0. \tag{3.19}$$

Relations (3.9) become

$$\frac{(L^2+A^2)^{1/2}}{A^2}A_j = \frac{\partial\psi}{\partial L_i}\{L_i,A_j\} + \frac{\partial\psi}{\partial A_i}\{A_i,A_j\}$$
$$= \epsilon_{ijk}\left(\frac{\partial\psi}{\partial L_i}A_k + \frac{\partial\psi}{\partial A_i}L_k\right) \quad (j = 1,2,3). \tag{3.20}$$

Multiplication with A_j and summation with respect to j give

$$(\mathbf{L}\times\mathbf{A})_i\frac{\partial\psi}{\partial A_i} = -(L^2+A^2)^{1/2} \tag{3.21}$$

in contradiction with (3.19). This proves the proposition. □

Definition:[26,27] A k-tuple of functions $(f_1,\ldots,f_k)$ on the symplectic manifold $(\mathfrak{M},\omega)$ is called *complete* if:

(1) $df_1,\ldots,df_k$ are linearly independent.

(2) There exists a set of functions $U_{ij}: R^k \to R$ such that

$$\{f_i,f_j\} = U_{ij}(f_1,\ldots,f_k), \quad 1\leqslant i,j\leqslant k. \tag{3.22}$$

The matrix of functions (U_{ij}) is called the *structural matrix* of $(f_1,\ldots,f_k)$.

A set of functions which may be expressed as functions only of the elements of a complete k-tuple are said to form a *group of functions of order k* on $(\mathfrak{M},\omega)$.[27,28]

The expressions (3.3)–(3.7) and Proposition 4 tells us the following:

Proposition 8: The generators of the realization of the so(4,2) algebra defined in Proposition 3 form a group of functions of order 6.

Proof: From (2.6), (3.7)–(3.9), and Proposition 7 it results that, for instance, $(L_1,L_2,L_3,A_1,A_2,\psi)$ form a complete sextuple. The 15 so(4,2) generators form [relations (3.3)–(3.7)] a group of functions of order 6. This results also by writing

$$\mathbf{L} = \epsilon(1/\sqrt{B_\alpha B_\alpha})\mathbf{C}\times\mathbf{B}, \tag{3.23}$$

$$\mathbf{A} = \epsilon(1/\sqrt{B_\alpha B_\alpha})(B_4\mathbf{C} - C_4\mathbf{B}), \tag{3.24}$$

$$M = \epsilon\sqrt{B_\alpha B_\alpha}. \tag{3.25}$$

The eight functions B_α, C_α $(\alpha = 1,2,3,4)$ are submitted to the two independent constraints $B_\alpha C_\alpha = 0$, $B_\alpha B_\alpha = C_\alpha C_\alpha$ and contain thus a complete sextuple of functions, in terms of which all the so(4,2) generators may be expressed.

The following theorem, due to Lie, on complete K-tuples of functions will be applicable to the realizations of the so(4,2) algebra. We present this theorem in the formulation of Roels and Weinstein:

Lie's Theorem[28,29]: Let $(f_1,\ldots,f_k)$ and $(f'_1,\ldots,f'_k)$ be complete k-tuples of functions defined on the symplectic manifolds $(\mathfrak{M},\omega)$ and $(\mathfrak{M}',\omega')$, respectively. Suppose that $f_i(x) = f'_i(x)$ $(i = 1,\ldots,k)$ for some points $x\in\mathfrak{M}$ and $x'\in\mathfrak{M}'$. Then there exists a diffeomorphism φ from a neighborhood of x onto a neighborhood of x' such that $\varphi^*\omega' = \omega$ and $\varphi^*f'_i = f_i$ if and only if $(f_1,\ldots,f_k)$ and $(f'_1,\ldots,f'_k)$ have the same structural matrix and $\dim\mathfrak{M} = \dim\mathfrak{M}'$. □

As established by Corollary 2, on any six-dimensional manifold a realization of the so(4,2) algebra can be defined the generators of which form a group of functions of order six. Let us now assume that, for instance, the sextuple of functions $(L_1,L_2,L_3,A_1,A_2,\psi)$ has been given two different specifications $(L^1_1,L^1_2,L^1_3,A^1_1,A^1_2,\psi^1)$, $(L^2_1,L^2_2,L^2_3,A^2_1,A^2_2,\psi^2)$ on two six-dimensional symplectic manifolds $(\mathfrak{M}_1,\omega_1)$ and $(\mathfrak{M}_2,\omega_2)$, respectively, both sextuples having thus the same structural matrix as the sextuple $(L_1,L_2,L_3, A_1, A_2,\psi)$; i.e., let us assume that two different realizations of the so(4,2) algebra have been defined, the generators of which satisfy the same set of polynomial identities (2.18)–(2.20), (2.23), (2.32). Lie's theorem tells us that these two realizations are canonically equivalent.

We shall use this result in the following section, in which different realizations of the so(4,2) algebra on six-dimensional manifolds will be obtained, by giving to the sextuple of functions $(L_1,L_2,L_3, A_1, A_2,\psi)$ different specific expressions.

4. DIFFERENT PARTICULAR CASES OF THE CANONICAL REALIZATION OF THE so(4,2) ALGEBRA GIVEN IN SEC. 3

In this section we shall show that several canonical realizations of the so(4,2) algebra on a six-dimensional symplectic manifold quoted in the literature[6,10–13] can be obtained from the canonical realization (3.3)–(3.7) by using appropriate expressions of the sextuple of functions $(\mathbf{L},\mathbf{A},\psi)$; all these realizations satisfy all the polynomial identities (2.14)–(2.32) and are thus, as a consequence of Lie's theorem, canonically equivalent. These realizations are:

(1) The realization (2.2), (2.3), (2.8)–(2.12) obtained by Györgyi[6] is obtained from the general so(4,2) realization of Sec. 3 if **L** and **A** are given by (2.2) and (2.3) respectively and if

$$\psi = w - \epsilon \sin w - \pi/2 = (2\pi/T)t - \pi/2, \tag{4.1}$$

where

w = eccentric anomaly of the elliptic motion,
ϵ = numerical eccentricity of the ellipse,
T = period of the elliptic motion,
t = time.

This expression of the function ψ may be obtained by observing that if **A** is the Laplace Runge–Lenz vector (2.3) and if $u = \mathbf{qp}$ and $A = (A_i A_i)^{1/2}$, then

$$qp^2 - 1 = A\sqrt{-2H}\cos w \tag{4.2}$$

$$u = A \sin w. \tag{4.3}$$

Introducing (4.2) and (4.3) in the expression (2.10) of B_4, we obtain

$$B_4 = A \sin(w - \sqrt{-2H}\, u - \pi/2), \tag{4.4}$$

which, if compared to (3.2) and if we observe that for the Kepler problem we have

$$\sqrt{-2H}\, u = \epsilon \sin w, \tag{4.5}$$

leads to relation (4.1).

The generators of this canonical realization satisfy the Györgyi relations (2.14)–(2.31) (which have been established precisely for this realization) and relation (2.32).

We shall denote, in the rest of this section, by **L**, **A**, B_α, C_α, M the generators of the realizations of so(4,2), (2.2), (2.3), (2.8)–(2.12) and shall express in terms of them the generators of some of the realizations of the so(4,2) algebra considered in this section.

(2) The canonical realization obtained by Dothan [Ref. 10, formulae (66)–(70)] reduces to the above realization if:

(i) The constants m, Z, e in the Kepler Hamiltonian $H = p^2/2m - Ze^2/q$ are all put equal to 1 (convention adopted in the present paper);

(ii) The following identifications are made,

$$\mathbf{J}\equiv\mathbf{L},\quad \mathbf{A}\equiv\mathbf{A},\quad \mathbf{U}\equiv\mathbf{B},\quad U_4 = B_4,$$
$$\mathbf{V} = \mathbf{C},\quad V_4 = C_4,\quad M_{56} = M; \tag{4.6}$$

(iii) The argument $\sqrt{-2Hu}$ of the trigonometric functions in B_α, C_α, (2.8)–(2.11) is replaced by

$$\varphi = \sqrt{-2H}\,(u + 2Ht). \tag{4.7}$$

We have seen (Proposition 2) that the function ψ in (3.3)–(3.7) is determined modulo the addition of a function of the so(4) Casimir invariant. Due to relation (3.7) between this invariant and the Kepler Hamiltonian, the function ψ is determined up to the addition of a function of the Kepler Hamiltonian.

The complete sextuple of functions which determine the canonical realization [Ref. 10, (66)–(70)] is formed by the functions **L** and **A** [(2.2), (2.3)] and by

$$\psi = w - \sqrt{-2H}\,(u + 2Ht) - \pi/2. \tag{4.8}$$

It may be also verified by direct calculation that if we denote by B_α, C_α the generators (2.8)–(2.11) and by B'_α, C'_α the corresponding functions in which the argument $\sqrt{-2Hu}$ is replaced by $\sqrt{-2Hu} + f(H)$, then for any f we have $\{B'_\alpha,C'_\alpha\} = \{B_\alpha,C_\alpha\}$. Similar equalities are obtained if in the so(4,2) Poisson bracket relations the functions B_α, C_α are replaced by B'_α, C'_α, respectively. The functions of the new argument form thus also a set of generators of the so(4,2) algebra. This new set of generators, which differ only in the expression of the argument of the trigonometric functions, from the generators of the realization (2.2), (2,3), (2.8)–(2.12) satisfy, therefore, the Györgyi identities (2.18)–(2.20), (2.23), (2.32) too.

(3) Similar remarks are valid for the first of the two canonical realizations of the so(4,2) algebra constructed by Tripathy, Gupta, and Anand [Ref. 11, formulae (2.11), (2.25a), (2.27), (2.29), (2.30)]. This realization reduces also to (2.2), (2.3), (2.8)–(2,12) (Sec. 2) if

(i) The following identifications are made:

$$\mathbf{L}\equiv\mathbf{L},\quad \mathbf{f}\equiv\mathbf{A},\quad \mathbf{M}\equiv\mathbf{B},\quad \mathbf{\Gamma}\equiv -\mathbf{C},$$
$$\Gamma_4\equiv B_4,\quad T\equiv -C_4,\quad \Gamma_0\equiv -M. \tag{4.9}$$

(ii) The argument $\sqrt{-2Hu}$ of the trigonometric functions in (2.8)–(2.12) is replaced by $\beta = \sqrt{-2H}\,(u - 2Ht)$.

The complete sextuple is given by **L**, **A** [(2.2)–(2.3)] and by

$$\psi = w - \sqrt{-2H}\,(u - 2Ht) - \pi/2. \tag{4.10}$$

The polynomial identities (2.13);(2.32) are verified.

(4) The generators of the second canonical realization given in Ref. 11 [formulae (2,11), (2.25b), (2.27), (2.30), (2.31)] can be expressed in terms of the generators (2.2), (2.3), (2.8)–(2.12) by the formulae:

$$\mathbf{L} = \mathbf{L},\quad \mathbf{f} = -\frac{\mathbf{L}}{L}\times\mathbf{A},$$

$$\mathbf{M} = -\frac{\mathbf{L}}{L}\times\mathbf{B}, \quad \mathbf{\Gamma} = \frac{\mathbf{L}}{L}\times\mathbf{C}, \tag{4.11}$$
$$\Gamma^4 = B_4, \quad T = -C_4, \quad \Gamma_0 = -M,$$

in which a substitution of the argument $\sqrt{-2Hu}$ by $\beta = \sqrt{-2H}(u - 2Ht)$ has to be performed.

It is easy to verify that if the set of generators $\mathbf{L}$, $\mathbf{A}$, $\mathbf{B}$, B_4, $\mathbf{C}$, C_4, M satisfy the polynomial identities (2.14)–(2.32) then the new set of generators (4.11) satisfy the same identities.

The complete sextuple is given by $\mathbf{L}$, $-(\mathbf{L}/L)\times\mathbf{A}$ and the same ψ as above (4.10).

(5) Another canonical realization of the so(4,2) algebra on a six-dimensional space is the realization obtained by Barut and Bornzin [Ref. 12, Formulae (A14), (A1)].

The generators of this realization, which will be denoted by $\mathbf{L}'$, $\mathbf{A}'$, B'_α, C'_α, M' in order to suggest the correspondence with the notations of Sec. 2, are

$$\mathbf{L}' = \mathbf{q}\times\mathbf{p}, \quad \mathbf{A}' = \tfrac{1}{2}(p^2 - 1)\mathbf{q} - u\mathbf{p},$$
$$\mathbf{B}' = \tfrac{1}{2}(p^2 + 1)\mathbf{q} - u\mathbf{p}, \quad \mathbf{C} = q\mathbf{p}, \tag{4.12}$$
$$B_4 = u, \quad C_4 = \tfrac{1}{2}(p^2 - 1)q, \quad M = -\tfrac{1}{2}(p^2 + 1)q.$$

They satisfy the polynomial identities (2.18)–(2.20) (2.23), (2.32). The realization (4.12) admits the complete sextuple of functions $\mathbf{L}'$, $\mathbf{A}'$ and

$$\Psi = \arctan(B_4/C_4) = \arctan[2u/q(p^2 - 1)]. \tag{4.13}$$

(6) The canonical realization of the so(4,2) algebra obained by Serebrennikov and Shabad[13] in their study of the central problems in mechanics is obtained from the general expression (3.3)–(3.7) if the following expressions are ascribed to $\mathbf{L}$, $\mathbf{A}$, and ψ:

$$\mathbf{L} = \mathbf{r}\times\mathbf{p}, \tag{4.14}$$
$$\mathbf{A} = A\left(\frac{\mathbf{r}}{r}\cos\xi + \frac{\mathbf{r}\times\mathbf{L}}{rL}\sin\xi\right), \tag{4.15}$$
$$\psi = \frac{2\pi}{T}\, t(r_{\min}, r), \tag{4.16}$$

where

$$\xi = \int_{r_{min}}^{r} \frac{L - Fr^2}{r}\,[2(H - V(r))r^2 - L^2]^{-1/2}\, dr$$

and

$$H = \tfrac{1}{2} p_i\, p_i + V(r)$$

is the Hamiltonian of a central motion submitted to the Poisson bracket condition,

$$\{\mathbf{A}^2 + \mathbf{L}^2, H\} = 0,$$

$r_{\min}$ is the perigee distance of the central motion, and

$$F = \frac{\partial(L^2 + A^2)}{\partial L}\left[\frac{\partial(L^2 + A^2)}{\partial H}\right]^{-1},$$
$$T = \frac{2\pi\sqrt{2}}{(L^2 + A^2)^{1/2}}\,\frac{\partial(L^2 + A^2)}{\partial H},$$
$$t(r_{\min}, r) = \int_{r_{\min}}^{r}\left(2H - 2V - \frac{L^2}{r^2}\right)^{-1/2} dr. \tag{4.17}$$

For these expressions of $\mathbf{L}$, $\mathbf{A}$, and ψ, Serebrennikov and Shabad obtained the dependence (3.3)–(3.7) on $\mathbf{L}$, $\mathbf{A}$, and ψ of the so(4,2) generators B_α, C_α ($\alpha = 1,2,3,4$), and M.

5. COORDINATE-FREE EXPRESSION OF THE SYMPLECTIC 2-FORM OF A SIX-DIMENSIONAL SYMPLECTIC MANIFOLD

In this section we shall obtain a coordinate-free expression of the symplectic 2-form $\widetilde{\omega}$ on the six-dimensional symplectic manifold $\mathfrak{M}$ in terms of the canonical realization of the so(4,2) algebra that can be determined on $\mathfrak{M}$ (Corollary 2).

This expression will result as a corollary of the following:

Proposition 9: Let the functions L_i, A_i, and ψ be defined on a six-dimensional symplectic manifold and let them satisfy the relations (2.4), (2.6), and (3.7)–(3.9). Let q_i, p_i ($i = 1,2,3$) be the local coordinates on this manifold and let L_i have the expression $L_i = (\mathbf{q}\times\mathbf{p})_i$ with respect to this system of coordinates. Then, taking in (3.3)–(3.7) $\epsilon = \eta = +1$.

$$\theta \equiv \frac{1}{M} B_\alpha dC_\alpha = p_i dq_i + dF, \tag{5.1}$$

where B_α, C_α ($\alpha = 1,\ldots,4$), and M are given by (3.3)–(3.7) and dF is the differential of a function $F(q, p)$.

Proof: Using relations (3.3)–(3.7) and (2.20), we obtain

$$\frac{1}{M} B_\alpha dC_\alpha = Md\psi - L\frac{\mathbf{A}}{A}\, d\left(\frac{\mathbf{L}}{L}\times\frac{\mathbf{A}}{A}\right). \tag{5.2}$$

A vector function $\mathbf{A}$ orthogonal (2.6) to the vector $\mathbf{L} = \mathbf{q}\times\mathbf{p}$ has the form

$$\frac{\mathbf{A}}{A} = \frac{\mathbf{q}}{q}\cos\xi + \frac{\mathbf{L}\times\mathbf{q}}{Lq}\sin\xi. \tag{5.3}$$

Using (5.3), we obtain

$$\frac{\mathbf{A}}{A}\, d\left(\frac{\mathbf{L}}{L}\times\frac{\mathbf{A}}{A}\right) = -d\xi - \frac{1}{L}\mathbf{p}d\mathbf{q} + \frac{u}{qL}dq \tag{5.4}$$

and hence

$$\frac{1}{M} B_\alpha dC_\alpha = p_i dq_i + Md\psi + Ld\xi + \frac{u}{q}dq. \tag{5.5}$$

We have to prove the existence of a function F [which has to be a scalar with respect to O(3), i.e., $F = F(q,L,u)$] such that

$$dF = Md\psi + Ld\xi - \frac{u}{q}dq. \tag{5.6}$$

Introducing in (5.6) the differentials of the scalar functions $\psi(q,L,u)$ and $\xi(q,L,u)$, we obtain

$$dF = \left(M\frac{\partial\psi}{\partial q} + L\frac{\partial\xi}{\partial q} - \frac{u}{q}\right)dq + \left(M\frac{\partial\psi}{\partial L} + L\frac{\partial\xi}{\partial L}\right)dL + \left(M\frac{\partial\psi}{\partial u} + L\frac{\partial\xi}{\partial u}\right)du. \tag{5.7}$$

The rhs of (5.7) is a total differential if and only if the following conditions are satisfied:

$$\frac{\partial M}{\partial q}\frac{\partial\psi}{\partial L} - \frac{\partial M}{\partial L}\frac{\partial\psi}{\partial q} = \frac{\partial\xi}{\partial q}, \tag{5.8}$$
$$\frac{\partial M}{\partial u}\frac{\partial\psi}{\partial L} - \frac{\partial M}{\partial L}\frac{\partial\psi}{\partial u} = \frac{\partial\xi}{\partial u}, \tag{5.9}$$
$$-\frac{\partial M}{\partial u}\frac{\partial\psi}{\partial q} + \frac{\partial M}{\partial q}\frac{\partial\psi}{\partial u} = -\frac{1}{q}. \tag{5.10}$$

Using the relation $\{q,u\} = q$, we obtain from (5.10) the following relation equivalent to (3.9):

$$\{M,\psi\} = \frac{D(M,\psi)}{D(q,u)}\{q,u\} = -1. \tag{5.11}$$

Let us now observe that from relations (3.1) and (3.15) we get

$$\frac{\partial M}{\partial q}\frac{\partial \xi}{\partial u} - \frac{\partial M}{\partial u}\frac{\partial \xi}{\partial q} = \frac{1}{q}\frac{\partial M}{\partial L}. \tag{5.12}$$

On the other side, from relation $\{\psi,\mathbf{A}/A\} = 0$, which is a consequence of (3.9), we obtain, using (5.3),

$$\{\psi,\xi\} = \frac{\partial \psi}{\partial L}. \tag{5.13}$$

Let us now prove that relations (5.8) and (5.9) are verified. In order to do that we first multiply them with $\partial\xi/\partial u$ and $-\partial\xi/\partial q$, respectively. The relation obtained by adding the results is verified as a consequence of relations (5.12)–(5.13). A second relation is obtained from (5.8) and (5.9) by multiplying them with $\partial M/\partial u$ and $-\partial M/\partial q$, respectively, and adding the results; we obtain, up to a factor, relation (5.11) which has already been proved. Relations (5.8) and (5.9) are thus verified and the proof is complete. □

By exterior differentiation of expression (5.1), we get the following:

Corollary: The symplectic form $\widetilde{\omega}$ of the six-dimensional symplectic manifold $(\mathfrak{M},\widetilde{\omega})$ admits the following expression in terms of the generators of the so(4,2) realization on $\mathfrak{M}$:

$$\widetilde{\omega} = d\theta = \frac{1}{M}dB_\alpha \wedge dC_\alpha - \frac{B_\alpha}{M^2}dM\wedge dC_\alpha. \tag{5.14}$$

This expression of the symplectic form $\widetilde{\omega}$ is independent of the local chart (U,φ).

6. ACTION-ANGLE VARIABLES ON THE SUBMANIFOLD $\mathfrak{M}$

In this section we point out a system of coordinates in which the symplectic 2-form $\widetilde{\omega}$ [(5.14)] takes the canonical form given by the Darboux theorem.[1]

Let us assume that $A \neq 0$ and define the angular variable

$$\psi = \begin{cases} -\arctan(B_4/C_4) & \text{if } C_4\neq 0,\\ \arctan(C_4/B_4) + \pi/2 & \text{if } B_4\neq 0.\end{cases} \tag{6.1}$$

We have

$$B_4/A = \sin\psi, \quad C_4/A = -\cos\psi, \tag{6.2}$$

and Proposition 4 tells us that from relations (2.18)–(2.20), (2.23), (2.32), among the so(4,2) generators we get the expressions (3.3)–(3.7) with $\epsilon = \eta = 1$ of the generators B_α, C_α ($\alpha = 1,2,3,4$), and M in terms of the variables ψ, $\mathbf{L}$, and $\mathbf{A}$ with $\mathbf{L}\cdot\mathbf{A} = 0$.

By straightforward calculations we obtain that, with respect to the variables ψ, $\mathbf{L}$, and $\mathbf{A}$, the 1-form θ [(5.1)] becomes

$$\theta = M\,d\psi + \left(\mathbf{L}\times\frac{\mathbf{A}}{A}\right)\cdot d\left(\frac{\mathbf{A}}{A}\right). \tag{6.3}$$

Let us now define the new angular variables Ω and Γ by

$$\Omega = \begin{cases} -\arctan L_1/L_2 & \text{if } L_2\neq 0,\\ \arctan(L_2/L_1) + \pi/2 & \text{if } L_1\neq 0,\end{cases} \tag{6.4}$$

$$\Gamma = \begin{cases} \arctan\left(\frac{LA_3}{L_1A_2 - A_1L_2}\right) & \text{if } L_1A_2 - A_1L_2\neq 0,\\ \arctan\left(\frac{L_1A_2 - A_1L_2}{LA_3}\right) + \frac{\pi}{2} & \text{if } A_3\neq 0.\end{cases} \tag{6.5}$$

With the aid of these variables and using the relation (2.6), we obtain

$$\begin{aligned}
\frac{L_1}{L} &= \left[1 - \left(\frac{L_3}{L}\right)^2\right]^{1/2}\sin\Omega,\\
\frac{L_2}{L} &= -\left[1 - \left(\frac{L_3}{L}\right)^2\right]^{1/2}\cos\Omega,\\
\frac{A_1}{A} &= \cos\Omega\cos\Gamma - \frac{L_3}{L}\sin\Omega\sin\Gamma,\\
\frac{A_2}{A} &= \sin\Omega\cos\Gamma + \frac{L_3}{L}\cos\Omega\sin\Gamma,\\
\frac{A_3}{A} &= \left[1 - \left(\frac{L_3}{L}\right)^2\right]^{1/2}\sin\Gamma,\\
\left(\frac{\mathbf{L}}{L}\times\frac{\mathbf{A}}{A}\right)_1 &= -\cos\Omega\sin\Gamma - \frac{L_3}{L}\sin\Omega\cos\Gamma,\\
\left(\frac{\mathbf{L}}{L}\times\frac{\mathbf{A}}{A}\right)_2 &= -\sin\Omega\sin\Gamma + \frac{L_3}{L}\cos\Omega\cos\Gamma,\\
\left(\frac{\mathbf{L}}{L}\times\frac{\mathbf{A}}{A}\right)_3 &= \left[1 - \left(\frac{L_3}{L}\right)^2\right]^{1/2}\cos\Gamma.
\end{aligned} \tag{6.6}$$

We can now prove

Proposition 10: The pairs of variables (M,ψ), (L,Ω), and (L_3,Γ) are canonically conjugated pairs, i.e., the following relations are valid:

$$\begin{aligned}
&\{M,L_3\} = \{M,L\} = \{L_3,L\} = 0,\\
&\{\psi,\Omega\} = \{\psi,\Gamma\} = \{\Omega,\Gamma\} = 0,\\
&\{M,\psi\} = \{L_3,\Omega\} = \{L,\Gamma\} = -1.
\end{aligned} \tag{6.7}$$

In terms of these variables the differential 1-form θ defined by Eq. (5.1) takes the following canonical expression:

$$\theta = M\,d\psi + L_3 d\Omega + L\,d\Gamma. \tag{6.8}$$

Proof: The proof of the first assertion is obtained by straightforward calculations, using the fact that the functions $\mathbf{L}$, $\mathbf{A}$, $\mathbf{B}$, B_4, $\mathbf{C}$, C_4, M satisfy the so(4,2) Lie bracket relations and the polynomial relations (2.18)–(2.20), (2.23), (2.32). The proof of the second assertion results from the equality

$$\left(\frac{\mathbf{L}}{L}\times\frac{\mathbf{A}}{A}\right)\cdot d\left(\frac{\mathbf{A}}{A}\right) = \frac{L_3}{L}d\Omega + d\Gamma, \tag{6.9}$$

which is a consequence of relations (6.6). □

From relations (3.3)–(3.7) with $\epsilon = \eta = 1$ and (6.6) we obtain the expressions of the so(4,2) generators in terms of the canonically conjugated variables M, ψ, L_3, Ω, L, Γ.

The variables M, ψ, L_3, Ω, L, Γ are a generalization of the Delaunay elements defined in classical dynamics,[15] which are shown in this way to possess a Lie algebraic signification.

7. CANONICAL REALIZATIONS OF A LIE ALGEBRA ON SUBMANIFOLDS OF ITS DUAL

Let G be a Lie algebra and let G^* be its dual. The action of G and G, defined by

$$(x,y)\in G\times G\to[x,y]\in G \tag{7.1}$$

($[\,,\,]$ = Lie bracket in G) and denoted by

$$\mathrm{ad}x(y)=[x,y], \tag{7.2}$$

is called the adjoint representation of G on G. For any $x\in G$ let us consider the differentiable function f_x on G^* defined by

$$f_x: u\in G^*\to f_x(u)\equiv u(x)\in R. \tag{7.3}$$

Let $\{x_1,...,x_n\}$ be a basis of G; we have

$$[x_i,x_j]=\sum_{j=1}^{n}c_{ijk}\,x_k, \tag{7.4}$$

where the c_{ijk} are the structure constants with respect to this basis; the dual basis $\{u_1,...,u_n\}$ in G^* is defined by

$$u_i(x_j)=\delta_{ij}. \tag{7.5}$$

An element $u\in G^*$ has the expression

$$u=\sum_{i=1}^{n}\xi_i(u)u_i, \tag{7.6}$$

where the coordinate functions $\xi_i(u)$ are given by

$$\xi_i(u)=u(x_i)=f_{x_i}(u). \tag{7.7}$$

Let us now associate with each element x of G a vector field on G^* defined (in a point $u\in G^*$ of coordinates ξ_i) by

$$X_x(u)=\sum_{j=1}^{n}u([x,x_j])\frac{\partial}{\partial\xi_j}. \tag{7.8}$$

In particular, we associate with each element x_i of the basis in G the field

$$X_{x_i}(u)=\sum_{j,k=1}^{n}c_{ijk}\xi_k\frac{\partial}{\partial\xi_j}. \tag{7.9}$$

The system of vectors X_{x_i} $(i=1,...,n)$ forms a realization of the Lie algebra G with the commutator as Lie bracket

$$[X_{x_i},X_{x_j}]=X_{x_i}X_{x_j}-X_{x_j}X_{x_i}=\sum_k c_{ijk}X_{x_k}. \tag{7.10}$$

Another realization of the Lie algebra G may be obtained by introducing the following bracket defined for any $f,g\in C^\infty(G^*)$ by

$$\{f,g\}=\sum_{i,j,k=1}^{n}c_{ijk}\xi_k\frac{\partial f}{\partial\xi_i}\frac{\partial g}{\partial\xi_j}. \tag{7.11}$$

This bracket is antisymmetric in f and g and satisfies the Jacobi identity: It is thus a Lie bracket. With respect to it the coordinate functions ξ_i of a point in G^* generate a realization of the Lie algebra G

$$\{\xi_i,\xi_j\}=\sum_k c_{ijk}\xi_k. \tag{7.12}$$

The action of the field X_{x_i} on a function $f\in C^\infty(G^*)$ is given by

$$X_{x_i}\mathrm{f}=\{\xi_i,\mathrm{f}\}. \tag{7.13}$$

The realization (7.10) and (7.12) of the Lie algebra G are defined in every point of G^*. They are not, however, canonical realizations on G^*, i.e., they are not associated with (and related to) a nondegenerate symplectic form.

Let us now point out a submanifold of G^* on which the Kirillov–Kostant–Souriau symplectic form can be defined; on this submanifold the realizations (7.10) and (7.12) are canonical with respect to the KKS 2-form.

Let $\mathfrak{M}_m$ be an m-dimensional submanifold of G^* defined by a set of $n-m$ functionally independent relations

$$f_i(\xi_1,...,\xi_n)=0\quad(i=1,...,n-m) \tag{7.14}$$

satisfying the following conditions:

$$(1)\quad X_{x_i}f_j=0\quad(i=1,...,n,\ j=1,...,n-m), \tag{7.15}$$

$$(2)\quad \text{if } u\in\mathfrak{M}_m,\quad\text{then}\quad \mathrm{rank}\left\|\sum_k c_{ijk}\xi_k(u)\right\|\geqslant m. \tag{7.16}$$

Then a symplectic form ω can be defined on $\mathfrak{M}_m$ by the KKS condition

$$\omega_u(X_x(u),X_y(u))=-u([x,y]). \tag{7.17}$$

The Poisson bracket associated with the KKS symplectic form ω on $\mathfrak{M}_m$ has the expression (7.11) in terms of the coordinates ξ_i.

Let us observe that condition (7.15) expresses the fact that the vector fields X_{x_i} are tangent to $\mathfrak{M}_m$; it is the local expression of the property of $\mathfrak{M}_m$ to be an orbit of the group generated by X_{x_i}. Condition (7.16) tells us that among the n vector fields, tangent to $\mathfrak{M}_m$ in u, at least m are linearly independent.

The existence in each point $u\in\mathfrak{M}_m$ of a system of m linearly independent tangent vectors X_{x_i} is sufficient to define on $\mathfrak{M}_m$ the KKS 2-form ω. Let us show that ω is nondegenerate and closed. The kernel of ω_u

$$\ker\omega_u=\{X\in T_u\mathfrak{M}_m;\ \omega_u(X,Y)=0, Y\in T_u\mathfrak{M}_m\} \tag{7.18}$$

has dimension zero. Indeed, let $X_{x_1}(u),...,X_{x_m}(u)$ be m linearly independent vectors in $T_u\mathfrak{M}_m$; we have

$$\begin{aligned}\ker\omega_u&=\Big\{\sum_{i=1}^{m}\lambda_iX_{x_i};\ \sum_{i,j=1}^{m}\lambda_i\,\mu_j\omega_u(X_{x_i},X_{x_j})=0,\forall\,\mu_1,...,\mu_m\Big\}\\&=\Big\{\sum_{i=1}^{m}\lambda_iX_{x_i};\ \sum_{i,k=1}^{m}\lambda_ic_{ijk}\xi_k(u)=0,\ j=1,...,m\Big\}.\end{aligned} \tag{7.19}$$

The dimension of $\ker\omega_u$ is thus the number of linearly independent solutions $\{\lambda_i\}$ of the system

$$\sum_{i=1}^{m}\lambda_i\sum_{k=1}^{m}c_{ijk}\xi_k(u)=0,\quad j=1,...,m \tag{7.20}$$

and from (7.16) it results that on $\mathfrak{M}_m$

$$\dim\ker\omega=0. \tag{7.21}$$

The 2-form ω is thus nondegenerate on $\mathfrak{M}_m$.

The property of ω to be closed, i.e., to satisfy $d\omega=0$ results from the formula of exterior differentiation

$$\begin{aligned}d\omega(X_0,X_1,X_2)&=X_0\omega(X_1,X_2)-X_1\omega(X_0,X_2)+X_2\omega(X_0,X_1)\\&\quad-\omega([X_0,X_1],X_2)-\omega([X_1,X_2],X_0)+\omega([X_0,X_2],X_1)\end{aligned} \tag{7.22}$$

and from the expression (7.9) of X_{x_i}, the definition (7.17) of ω_u, and the identity of Jacobi.

ω is thus a symplectic form on $\mathfrak{M}_m$. This allows to associate with any function $f \in C^\infty(\mathfrak{M}_m)$ a vector field X_f on $\mathfrak{M}_m$ defined by

$$X_f(u) \lrcorner \omega_u = -df(u). \qquad (7.23)$$

The Poisson bracket of any two functions $f, g \in C^\infty(\mathfrak{M}_m)$ is now defined by

$$\{f, g\}(u) = -\omega_u(X_f(u), X_g(u)). \qquad (7.24)$$

Let us show that the Poisson bracket admits the expression (7.11) in terms of the coordinates ξ. We have, using (7.9),

$$\omega(X_f, X_{x_i}) = -X_{x_i} f = -\sum_{j,k} c_{ijk} \xi_k \frac{\partial f}{\partial \xi_j}. \qquad (7.25)$$

The expression

$$X_f = \sum_j \frac{\partial f}{\partial \xi_j} X_{x_j} \qquad (7.26)$$

is a solution for X_f which is, in general, not unique. On $\mathfrak{M}_m$, however, due to nondegeneracy of ω the solution is unique. This expression of X_f leads to the expression (7.11) for the Poisson bracket on $\mathfrak{M}_m$.

Let us observe that

$$X_{f_y} = X_y. \qquad (7.27)$$

Indeed

$$\omega_u(X_{f_y}, X_x(u)) = -df_y(X_x)(u) = -(X_x f_y)(u) = \omega(X_y, X_x). \quad (7.28)$$

Using this remark, we get

$$\{f_x, f_y\}(u) = -\omega_u(X_x, X_y) = u([x,y]) = f_{[x,y]}(u), \quad (7.29)$$

i.e., the correspondence $x \to f_x$ gives a canonical realization of the Lie algebra G. Taking in (7.29) $x = x_i, y = y_j$, and using (7.4), (7.7), we obtain (7.12).

Let us look for an explicit expression of the KKS symplectic form ω on $\mathfrak{M}_m$

$$\omega_u = \tfrac{1}{2} \sum_{s,t=1}^{n} \omega_{st}(u)\, d\xi_s \wedge d\xi_t, \quad \omega_{st} = -\omega_{ts} \qquad (7.30)$$

in terms of the coordinate functions ξ_i $(i = 1,...,n)$. From the definition of the KKS form we obtain the following equations for the coefficients ω_{st}:

$$\begin{aligned} \omega_u(X_{x_i}, X_{x_j}) &= \sum_{s,t,p,q=1}^{n} \omega_{st}(u) c_{isp} c_{jtq} \xi_p(u) \xi_q(u) \\ &= -\sum_{k=1}^{n} c_{ijk} \xi_k(u) \quad (i,j = 1,...,n). \end{aligned} \qquad (7.31)$$

If we know an expression for the coefficients ω_{st} in terms of the coordinates ξ_i, then the set of equations (7.31) gives the equations of the manifold $\mathfrak{M}_m$ on which the symplectic form ω is defined.

As an example let us write down Eqs. (7.31) when the algebra is so(3). The vector space so(3)* is three-dimensional, the so(3) orbits are spheres, and the expression of ω on a sphere is

$$\omega = -\epsilon_{str} \xi_r\, d\xi_s \wedge d\xi_t. \qquad (7.32)$$

Hence $\omega_{st} = -\epsilon_{str}\xi_r$ and, taking into account that for so(3) the structure constants have the form $c_{ijk} = \epsilon_{ijk}$, Eqs. (7.31) become

$$(\epsilon_{ijk}\xi_k)\xi_p\, \xi_p = \epsilon_{ijk}\xi_k, \qquad (7.33)$$

whence $\xi_p\, \xi_q = 1$, which is the equation of a sphere.

In the next section, the correspondence between the equations of an orbit, and the symplectic form on that orbit will be deduced in the particular case of the six-dimensional orbit of the SO(4,2) group.

8. A SIX-DIMENSIONAL SYMPLECTIC SUBMANIFOLD OF so(4,2)*

Let us specify the Lie algebra G of the previous section, taking $G = \text{so}(4,2)$, and let us introduce for the $n = 15$ coordinate functions ξ_i in so(4,2)* the following notations:

$$\begin{aligned} &\xi_i = L_i, \quad \xi_{3+i} = A_i \quad (i = 1,2,3) \\ &\xi_{6+j} = B_j, \quad \xi_{10+j} = C_j \quad (j = 1,2,3,4), \quad \xi_{15} = M \end{aligned} \qquad (8.1)$$

assuming that these coordinate functions, which generate a realization (7.11)–(7.12) of the so(4,2) algebra, verify the Lie multiplication table (Table I) of this algebra.

Let us consider in so(4,2)* the subset $\mathfrak{M}_6$ defined by the following nine relations: the two vector relations (2.18), (2.19) and the three scalar relations (2.20), (2.23), and (2.32); this subset is a six-dimensional submanifold of so(4.2)* as it results from the proof of Proposition 8.

The set of polynomials which define $\mathfrak{M}_6$ satisfy condition (7.15). Indeed, according to Proposition 1, the adjoint action of the so(4,2) algebra transforms any of the polynomials on the lhs of the relations (2.6), (2.14)–(2.19) into a polynomial of the same set; similarly, the set of polynomials (2.20)–(2.31) is closed with respect to the adjoint action. But, according to Proposition 3, any of these polynomials vanishes on $\mathfrak{M}_6$; hence, by using Eqs. (7.13), it results that condition (7.15) is satisfied.

Let us now prove that on $\mathfrak{M}_6$ condition (7.16) is satisfied.

Proposition 11: If the coordinate functions of so(4,2)* satisfy the system (2.18)–(2.20), (2.23), (2.32), then [with notations (8.1)]

$$\text{rank} \left\| \sum_{k=1}^{15} c_{ijk}\xi_k \right\| \geqslant 6. \qquad (8.2)$$

Proof: Let us suppose that the inequality (8.2) is false, i.e., that all 6×6 minors of the matrix $\|\Sigma_{k=1}^{15} c_{ijk}\xi_k\|$ $(i,j = 1,...,15)$ vanish and that the coordinate functions ξ_k (8.1) are a nonvanishing solution of the equations (2.18)–(2.20), (2.23), (2.32). We shall consider the following 6×6 minors of the matrix $\|\Sigma_{k=1}^{15} c_{ijk}\xi_k\|$ $(i,j = 1,...,15)$:

$$M_1 = \left\| \sum_{k=1}^{15} c_{ijk}\xi_k \right\| \quad (i,j = 4,5,6,7,8,9) \qquad (8.3)$$

$$M_2 = \left\| \sum_{k=1}^{15} c_{ijk}\xi_k \right\| \quad (i,j = 4,5,6,11,12,13) \qquad (8.4)$$

$$M_{1\alpha} = \left\| \sum_{k=1}^{15} c_{ijk}\xi_k \right\| \quad (i,j = 1,2,3,\alpha,5,6) \qquad (8.5)$$

$$M_{2\alpha} = \left|\left| \sum_{k=1}^{15} c_{ijk}\xi_k \right|\right| \quad (i,j = 1,2,3,4,\alpha,6), \quad \alpha = 7,8,9 \tag{8.6}$$

$$M_{3\alpha} = \left|\left| \sum_{k=1}^{15} c_{ijk}\xi_k \right|\right| \quad (i,j = 1,2,3,4,5,\alpha) \tag{8.7}$$

we have

$$\det M_1 = B_4^2(B_4^2 + L_1^2 + L_2^2 + L_3^2)^2, \tag{8.8}$$

$$\det M_2 = C_4^2(C_4^2 + L_1^2 + L_2^2 + L_3^2)^2. \tag{8.9}$$

Hence $\det M_1 = 0$ if and only if $B_4 = 0$ and $\det M_2 = 0$ if and only if $C_4 = 0$.

The equations $B_4 = C_4 = 0$ and (2.18)–(2.20), (2.23), (2.32) lead to the following relations:

$$\mathbf{A} = 0, \quad L^2 = B^2 = C^2 = M^2, \quad \mathbf{B}\cdot\mathbf{C} = 0. \tag{8.10}$$

If $\mathbf{A} = 0$, then

$$\det M_{i\alpha} = L_i^2(\mathbf{L}\cdot\mathbf{B})_{\alpha-6} \quad (i = 1,2,3, \ \alpha = 7,8,9), \tag{8.11}$$

and from (8.10) and $\det M_{i\alpha} = 0$ ($i = 1,2,3$, $\alpha = 7,8,9$), we obtain that either $L^2 = 0$, i.e., $L^2 = B^2 = C^2 = M^2 = 0$, or $\mathbf{L}\times\mathbf{B} = 0$. The last alternative together with relations (2.14) and (8.10) leads again to the vanishing of all the coordinate functions. This contradicts the assumption and proves the proposition. □

Conditions (7.15) and (7.16) are thus satisfied; a symplectic form defined by the KKS condition exists therefore on $\mathfrak{M}_6$.

Let us now write down an explicit expression for the symplectic form on $\mathfrak{M}_6$. It has been proved in Sec. 3 that on any six-dimensional symplectic manifold a realization of the so(4,2) algera can be constructed, the generators of which satisfy the relations (2.18)–(2.20), (2.23), (2.32). The symplectic form on this manifold can be expressed in terms of the generators of this algebra by the relation (5.14). The coordinate functions (8.1) generate a canonical realization of the so(4,2) algebra on $\mathfrak{M}_6$ and satisfy the polynomial relations quoted above which define $\mathfrak{M}_6$; thus on $\mathfrak{M}_6$ a symplectic form (5.14) can be defined.

The problem which remains to be solved is whether on $\mathfrak{M}_6$ the 2-form (5.14) coincides with the KKS 2-form, i.e., whether on $\mathfrak{M}_6$ we have

$$\widetilde{\omega}_u(X_{x_i},X_{x_j}) = \omega_u(X_{x_i},X_{x_j}) \quad \text{for} \quad i,j = 1,\dots,15, \tag{8.12}$$

where the vector fields X_{x_i} (7.9) have to be expressed in terms of the coordinates (8.1). In order to prove relation (8.12), let us first specify the vector fields (7.9) which correspond to the so(4,2) algebra:

$$\begin{aligned}
X_{L_i} &= \epsilon_{ijk}L_k \frac{\partial}{\partial L_j} + \epsilon_{ijk}A_k\frac{\partial}{\partial A_j} + \epsilon_{ijk}B_k\frac{\partial}{\partial B_j} + \epsilon_{ijk}C_k\frac{\partial}{\partial C_j} \quad (i = 1,2,3),\\
X_{A_i} &= \epsilon_{ijk}L_k\frac{\partial}{\partial A_j} - \epsilon_{ijk}A_k\frac{\partial}{\partial L_j} + B_4\frac{\partial}{\partial B_i} - B_i\frac{\partial}{\partial B_4} + C_4\frac{\partial}{\partial C_i} - C_i\frac{\partial}{\partial C_4} \quad (i = 1,2,3),\\
X_{B_i} &= -\epsilon_{ijk}L_k\frac{\partial}{\partial B_j} + \epsilon_{ijk}B_k\frac{\partial}{\partial L_j} - B_4\frac{\partial}{\partial A_i} - A_i\frac{\partial}{\partial B_4} - M\frac{\partial}{\partial C_i} - C_i\frac{\partial}{\partial M} \quad (i = 1,2,3),\\
X_{C_i} &= -\epsilon_{ijk}L_k\frac{\partial}{\partial C_j} + \epsilon_{ijk}C_k\frac{\partial}{\partial L_j} - C_4\frac{\partial}{\partial A_i} - A_i\frac{\partial}{\partial C_4} + M\frac{\partial}{\partial B_i} + B_i\frac{\partial}{\partial M} \quad (i = 1,2,3),\\
X_{B_4} &= B_k\frac{\partial}{\partial A_k} + A_k\frac{\partial}{\partial B_k} - M\frac{\partial}{\partial C_4} - C_4\frac{\partial}{\partial M},\\
X_{C_4} &= C_k\frac{\partial}{\partial A_k} + A_k\frac{\partial}{\partial C_k} + M\frac{\partial}{\partial B_4} + B_4\frac{\partial}{\partial M},\\
X_M &= C_k\frac{\partial}{\partial B_k} - B_k\frac{\partial}{\partial C_k} + C_4\frac{\partial}{\partial B_4} - B_4\frac{\partial}{\partial C_4}.
\end{aligned} \tag{8.13}$$

Proposition 12: Relations (8.12) are satisfied if and only if $u\in\mathfrak{M}_6$.

Proof: Let us calculate the difference between

$$\omega_u(X_{x_i},X_{x_j}) = -u([x_i,x_j]) = -\sum_k c_{ijk}\xi_k \tag{8.14}$$

and $\widetilde{\omega}_u(X_{x_i},X_{x_j})$ [(5.14), (8.13)]:

$$(\omega - \widetilde{\omega})(X_{L_i},X_{L_j}) = (1/M)\epsilon_{ijk}(\mathbf{C}\times\mathbf{B} - M\mathbf{L})_k, \tag{8.15}$$

$$(\omega - \widetilde{\omega})(X_{L_i},X_{A_j}) = -(1/M)\epsilon_{ijk}(M\mathbf{A} + C_4\mathbf{B} - B_4\mathbf{C})_k, \tag{8.16}$$

$$(\omega - \widetilde{\omega})(X_{L_i},X_{B_j}) = -(1/M)\delta_{ij}\mathbf{L}\cdot\mathbf{C} - \frac{C_j}{M^2}(\mathbf{C}\times\mathbf{B} - M\mathbf{L})_i, \tag{8.17}$$

$$(\omega - \widetilde{\omega})(X_{L_i},X_{B_j}) = -(1/M)(\mathbf{A}\times\mathbf{C} - C_4\mathbf{L})_i - (C_4/M^2)(\mathbf{C}\times\mathbf{B} - M\mathbf{L})_i, \tag{8.18}$$

$$(\omega - \widetilde{\omega})(X_{L_i},X_{C_j}) = (1/M)\delta_{ij}\mathbf{L}\cdot\mathbf{B} + (B_j/M^2)(\mathbf{C}\times\mathbf{B} - M\mathbf{L})_i, \tag{8.19}$$

$$(\omega - \widetilde{\omega})(X_{L_i},X_{C_4}) = -(1/M)(\mathbf{A}\times\mathbf{B} - B_4\mathbf{L})_i + (B_4/M^2)(\mathbf{C}\times\mathbf{B} - M\mathbf{L})_i, \tag{8.20}$$

$$(\omega - \widetilde{\omega})(X_{L_i},X_M) = 0, \tag{8.21}$$

$$(\omega - \tilde{\omega})(X_{A_i}, X_{A_j}) = \epsilon_{ijk}(1/M)(\mathbf{C}\times\mathbf{B} - M\mathbf{L})_k, \tag{8.22}$$
$$(\omega - \tilde{\omega})(X_{A_i}, X_{B_j}) = -\epsilon_{ijk}(1/M)(\mathbf{A}\times\mathbf{C} - C_4\mathbf{L})_k + (C_j/M^2)(M\mathbf{A} + C_4\mathbf{B} - B_4\mathbf{C})_i, \tag{8.23}$$
$$(\omega - \tilde{\omega})(X_{A_i}, X_{B_4}) = (C_4/M^2)(M\mathbf{A} + C_4\mathbf{B} - B_4\mathbf{C})_i, \tag{8.24}$$
$$(\omega - \tilde{\omega})(X_{A_i}, X_{C_j}) = \epsilon_{ijk}(1/M)(\mathbf{A}\times\mathbf{B} - B_4\mathbf{L})_k - (B_j/M^2)(M\mathbf{A} + C_4\mathbf{B} - B_4\mathbf{C})_i, \tag{8.25}$$
$$(\omega - \tilde{\omega})(X_{A_i}, X_{C_4}) = -(B_4/M^2)(M\mathbf{A} + C_4\mathbf{B} - B_4\mathbf{C})_i, \tag{8.26}$$
$$(\omega - \tilde{\omega})(X_{A_i}, X_M) = 0, \tag{8.27}$$
$$(\omega - \tilde{\omega})(X_{B_i}, X_{B_j}) = \epsilon_{ijk}(1/M)(\mathbf{C}\times\mathbf{B} - M\mathbf{L})_k, \tag{8.28}$$
$$(\omega - \tilde{\omega})(X_{B_i}, X_{B_4}) = -(1/M)(M\mathbf{A} + C_4\mathbf{B} - B_4\mathbf{C})_i, \tag{8.29}$$
$$(\omega - \tilde{\omega})(X_{B_i}, X_{C_j}) = (1/M)(L_iL_j - A_iA_j + B_iB_j + C_iC_j) - \delta_{ij}(1/M)L_kL_k - (C_i/M^2)(\mathbf{B}\times\mathbf{L} + B_4\mathbf{A} - M\mathbf{C})_j, \tag{8.30}$$
$$(\omega - \tilde{\omega})(X_{B_i}, X_{C_4}) = (1/M)(\mathbf{A}\times\mathbf{L} + B_4\mathbf{B} + C_4\mathbf{C})_i - (C_i/M^2)(\mathbf{A}\mathbf{B} + C_4M), \tag{8.31}$$
$$(\omega - \tilde{\omega})(X_{B_i}, X_M) = -(1/M)(\mathbf{B}\times\mathbf{L} + B_4\mathbf{A} - M\mathbf{C})_i + (C_i/M^2)(B_\alpha B_\alpha - M^2), \tag{8.32}$$
$$(\omega - \tilde{\omega})(X_{B_4}, X_{C_i}) = (1/M)(\mathbf{A}\times\mathbf{L} + B_4\mathbf{B} + C_4\mathbf{C})_i + (C_4/M^2)(\mathbf{B}\times\mathbf{L} + B_4\mathbf{A} - M\mathbf{C})_i, \tag{8.33}$$
$$(\omega - \tilde{\omega})(X_{B_4}, X_{C_4}) = (1/M)(B_4^2 + C_4^2 - \mathbf{A}^2) - (C_4/M^2)(\mathbf{A}\cdot\mathbf{B} + C_4M), \tag{8.34}$$
$$(\omega - \tilde{\omega})(X_{B_4}, X_M) = (1/M)(\mathbf{A}\cdot\mathbf{B} + C_4M) + (C_4/M^2)(B_\alpha B_\alpha - M^2), \tag{8.35}$$
$$(\omega - \tilde{\omega})(X_{C_i}, X_{C_j}) = (B_j/M^2)(\mathbf{B}\times\mathbf{L} + B_4\mathbf{A} - M\mathbf{C})_i - (B_i/M^2)(\mathbf{B}\times\mathbf{L} + B_4\mathbf{A} - M\mathbf{C})_j + (1/M)\epsilon_{ijk}(\mathbf{C}\times\mathbf{B} - M\mathbf{L})_k, \tag{8.36}$$
$$(\omega - \tilde{\omega})(X_{C_i}, X_{C_4}) = (B_4/M^2)(\mathbf{B}\times\mathbf{L} + B_4\mathbf{A} - M\mathbf{C})_i + (B_i/M^2)(\mathbf{A}\cdot\mathbf{B} + C_4M) - (1/M)(M\mathbf{A} + C_4\mathbf{B} - B_4\mathbf{C})_i, \tag{8.37}$$
$$(\omega - \tilde{\omega})(X_{C_i}, X_M) = -(1/M)(\mathbf{C}\times\mathbf{L} + C_4\mathbf{A} + M\mathbf{B})_i - (B_i/M^2)(B_\alpha B_\alpha - M^2), \tag{8.38}$$
$$(\omega - \tilde{\omega})(X_{C_4}, X_M) = (1/M)(\mathbf{A}\cdot\mathbf{C} - B_4M) - (B_4/M^2)(B_\alpha B_\alpha - M^2). \tag{8.39}$$

Equations (2.6), (2.14), (2.32), which result (Proposition 3) from the equations defining the manifold $\mathfrak{M}_6$, lead to the equality of the two symplectic forms $\tilde{\omega}$ and ω on $\mathfrak{M}_6$.

To prove the converse assertion, we have to show that from $\tilde{\omega} = \omega$ it follows that the relations defining the submanifold $\mathfrak{M}_6$ are verified; it is even sufficient to show that one relation among each set of relations (2.6), (2.14)–(2.19) and (2.20)–(2.31) is verified and that (2.32) is true. Relations (2.18) and (2.19) result, if $\tilde{\omega} = \omega$, directly from relations (8.15) and (8.16), respectively. Using the equalities (2.19), (8.31), (8.33), (8.37) and discarding the trivial solution, we obtain (2.22). Finally, relation (2.32) is a consequence of (2.18)–(2.20), (2.22), and (8.35). □

[1]R. Abraham and J. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1967).
[2]V.I. Arnold, *Méthodes mathématiques de la Mécanique Classique* (Mir, Moscow, 1976).
[3]W. Pauli, Z. Phys. **36**, 336 (1926).
[4]V. Bargmann, Z. Phys. **99**, 578 (1936).
[5]H. Bacry, H. Ruegg, and J.M. Souriau, Commun. Math. Phys. **3**, 323 (1966).
[6]G. Györgyi, Nuovo Cimento A **53**, 717 (1968).
[7]G. Györgyi, Nuovo Cimento A **62**, 449 (1969).
[8]G. Györgyi, Magyar. Fiz. Foly. **17**, 39 (1969).
[9]G. Györgyi, Magyar Fiz. Foly. **20**, 163 (1972).
[10]Y. Dothan, Phys. Rev. D **2**, 2944 (1970).
[11]K.C. Tripathy, R. Gupta, and J.D. Anand, J. Math. Phys. **16**, 1139 (1975).
[12]A.O. Barut and G.L. Bornzin, J. Math. Phys. **12**, 841 (1971).
[13]V.B. Serebrennikov, A.E. Shabad, Intern. J. Theoret. Phys. **16**, 517 (1977).
[14]For the $o(n+2)$ algebra, a set of polynomial relations among the generators has been obtained in N. Mukunda, J. Math. Phys. **8**, 1069 (1967).
[15]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Cambridge, Mass. 1950).
[16]A.A. Kirillov, *Elementy Teorii Predstavlenii* (Nauka, Moscow, 1972).
[17]D.J. Simms, "Projective Representations, Symplectic Manifolds and Extensions of Lie Algebras," Preprint CNRS, Marseille (1969).
[18]A. Joseph, "Realizations in Classical and Quantum Mechanics," Third International Colloquium on Group Theoretical Methods in Physics, Marseille 1974, Preprint IHES.
[19]E. Onofri and M. Pauri, J. Math. Phys. **13**, 533 (1972).
[20]E. Onofri, J. Math. Phys. **17**, 401 (1976).
[21]The Einstein summation convention will be used in summations over indices which take the values 1,2,3 or 1,2,3,4. We shall use, to distinguish these two cases, Latin and Greek indices, respectively.
[22]H. Bacry, Nuovo Cimento A **41**, 222 (1966).
[23]M. Iosifescu and H. Scutaru, "On Six-Dimensional Canonical Realizations of the so(4,2) Algebra," Preprint FT-168-79, Bucharest (1979).
[24]M. Iosifescu, "Canonical Realizations of the so(4,2) Algebra," Preprint FT-163-1978, Bucharest (1978).
[25]M. Iosifescu, "so(4,2) Algebra and Central Motion," Preprint FT-150-1978, Bucharest (1978).
[26]A. Simoni, F. Zaccaria, and B. Vitale, Nuovo Cimento A **51**, 448 (1967).
[27]B. Vitale, *Lectures in Theoretical Physics* (Colorado U. P., Boulder, 1968), Vol. XIA, p. 405.
[28]S. Lie, Math. Ann. **8**, 214 (1875).
[29]J. Roels and A. Weinstein, J. Math. Phys. **12**, 1482 (1971).
[30]S. Okubo, J. Math. Phys. **18**, 2382 (1977).

Symmetries of differential equations. II

F. González-Gascón
Departamento de Metodos Matematicos de la Universidad Complutense e Instituto de Estructura de la Materia (CSIC, GIFT), Madrid 6, Spain

(Received 19 July 1979; accepted for publication 7 February 1980)

The importance of the non-pointlike transformations of symmetry is vindicated in relation with the first integrals. A new first integral of a broad class of systems of second order differential equations is obtained out of a symmetry of them *without* having to impose the restriction that the system be equivalent to a Lagrangian system. The existence of a reciprocal relationship among the local infinitesimal symmetries (l.i.s.) of a Newtonian system of differential equations and the pseudosymmetries of the associated dynamical system is proved. Several applications are developed, and some open problems concerning the dynamical systems of constant divergence are proposed.

I. INTRODUCTION

In a recent paper,[1] hereafter referred to as I, we studied (from a local point of view) the important role played by the local families of monoparametric transformations of symmetry (l.i.s.) of systems of differential equations in classical mechanics. In this paper we stress further the importance of this kind of symmetry transformations in connection with the important practical task of finding first integrals of the differential equations (Sec. II). We connect the l.i.s. with either the pseudosymmetries of the dynamical system associated with the differential equations (Sec. III) or, and more importantly, with the symmetries of it (Sec. IV). A reciprocal relationship is established among the l.i.s. of a Newtonian set of differential equations and the pseudosymmetries of the associated dynamical system (Secs. III and IV). Finally, and as the more important application, we obtain, when the dynamical system is of constant divergence, a first integral out of an l.i.s. of the original set of differential equations. An l.i.s. out of a first integral of Lagrangian set of differential equations is obtained as well (Sec. IV).

II. THE NECESSITY OF THE LOCAL NON-POINTLIKE SYMMETRIES IN RELATION WITH THE LOCAL FIRST INTEGRALS

Throughout this paper, and exclusively by reasons of symplicity in the notation, we shall deal with Newtonian systems of differential equations, i.e., with sets of differential equations of the kind

$$\ddot{q}_i = f_i(t,q_1,\dots,q_n,\dot{q}_1,\dots,\dot{q}_n), \quad i = 1,\dots,n. \tag{1}$$

We discussed in I that, in general, Eqs. (1) do not admit infinitesimal pointlike symmetries of the kind

$$\bar{q}_i = q_i + \epsilon g_i(t,q_1,\dots,q_n), \qquad \bar{t} = t + \epsilon g_0(t,q_1,\dots,q_n). \tag{2}$$

On the other hand, it was shown that Eqs. (1) do always admit local infinitesimal symmetries (l.i.s.) of the more general kind:

$$\bar{q}_i = q_i + \epsilon h_i(t,q_1,\dots,q_n;\dot{q}_1,\dots,\dot{q}_n), \qquad \bar{t} = t + \epsilon h_0(t,q_1,\dots,q_n,\dot{q}_1,\dots,\dot{q}_n), \tag{3}$$

transforming *any* solution of Eqs. (1) into another solution of them.

It is well known that the local first integrals of (1), defined as the solutions of the linear partial differential equation,

$$\frac{\partial I}{\partial t} + \sum_{i=1}^{n} \frac{\partial I}{\partial q_i}\dot{q}_i + \sum_{i=1}^{n} \frac{\partial I}{\partial \dot{q}_i} f_i = 0 \tag{4}$$

do *always* exist, the number of the independent ones being just equal to $2n$.

Therefore, if we desire to establish connections among the first integrals and the symmetries of Eqs. (1) [i.e., discover rules permitting us to obtain a first integral of (1) out of a symmetry of them, or obtain a symmetry of (1) out of a first integral of Eqs. (1)], it is clear that we cannot restrict the set of infinitesimal symmetries to those of the kind (2), since if this were the case for most of the Eqs. (1), the connection would be impossible, as the set of symmetries of the kind (2) is, for most of the equations, the void set.

Accordingly the possible discovery of new connections among the symmetries and first integrals of Eqs. (1) oblige us to consider the symmetries of the more general kind (3), studied in I for other reasons, that always exist for any system of the kind (1). Therefore, the theoretical importance of the local non-pointlike symmetries (3) is, once more, vindicated here.

From the practical point of view we shall see, in Sec. IV, that, under certain conditions restricting the form of Eqs. (1) it is possible to get a direct relationship between the l.i.s. and the first integrals, thus permitting us to obtain a first integral of Eqs. (1) from an l.i.s. of them.

III. THE l.i.s. OF A SET OF DIFFERENTIAL EQUATIONS AND THE PSEUDOSYMMETRIES OF THE ASSOCIATED DYNAMICAL SYSTEM

It is well known that the system (1) can be written in the equivalent form

$$\frac{\partial q_i}{dt} = \dot{q}_i, \qquad \frac{\partial \dot{q}_i}{dt} = f_i(t,q_1,\dots,q_n,\dot{q}_1,\dots,\dot{q}_n) \tag{5}$$

 0022-2488/80/082046-06$1.00

or in the equivalent form (the parameter s does not appear now in the right-hand side of the equations)

$$\frac{\partial q_i}{ds}=\lambda \dot q_i,\quad \frac{d\dot q_i}{ds}=\lambda f_i(t,q_1,...,q_n,\dot q_1,...,\dot q_n),\quad \frac{dt}{ds}=\lambda, \tag{6}$$

λ being any function of $t,q_1,...,q_n,\dot q_1,...,\dot q_n$.

In relation with Eqs. (6) it is clear that a variation of the function λ only amounts to a different parametrization of the variables $(t,q_i,\dot q_i)$ in terms of s, but the functionals relations expressing $q_i,\dot q_i$ in terms of t [and given by Eqs. (5)] remain unaltered. Therefore, it is natural to expect that the important concept corresponding to the symmetries of Eqs. (1), when they are written in the form (6) or in the form

$$\frac{\partial q_i}{ds}=\dot q_i,\quad \frac{\partial \dot q_i}{ds}=f_i(t,q_1,...,q_n;\dot q_1,...,\dot q_n),\quad \frac{dt}{ds}=1, \tag{7}$$

is the concept of pseudosymmetry of Eqs. (7).

We recall here that, given an autonomous set **X** of first order differential equations,

$$\frac{dx_i}{ds}=X_i(x_1,...,x_r),\quad i=1,...,r, \tag{8}$$

we call **S** a pseudosymmetry set of first order differential equations,

$$\frac{dx_i}{ds}=S_i(x_1,...,x_r),\quad i=1,...,r, \tag{9}$$

of (8), when the local pseudogroup of transformations associated with **S**[2] transforms any trajectory (not necessarily the solutions) of (8) into another trajectory of (8). Necessary and sufficient conditions for this are

$$[\mathbf S,\mathbf X]=\rho(x_1,...,x_r)\mathbf X \tag{10}$$

[,] being the Lie–Jacobi bracket of vector fields.[2]

Now, assume that

$$\bar q_i=q_i+\epsilon h_i(t,q_1,...,q_n;\dot q_1,...,\dot q_n),\quad \bar t=t+\epsilon h_0(t,q_1,...,q_n;\dot q_1,...,\dot q_n) \tag{11}$$

is a symmetry of Eqs. (1). Then, and following the techniques of I, we can induce the following transformation of the $\dot q_i$ variables:

$$\dot{\bar q}_i=\dot q_i+\epsilon\left(\frac{dh_i}{dt}-\dot q_i\frac{dh_0}{dt}\right)\overset{\text{def}}{=}\dot q_i+\epsilon(\dot h_i-\dot q_i\dot h_0), \tag{12}$$

where it is implicitly assumed that in Eqs. (12) we have substituted the terms $\ddot q_i$ appearing there by their values

$$f_i(t,q_1,...,q_n,\dot q_1,...,\dot q_n),$$

obtained from Eqs. (1).

Therefore, Eqs. (11) and (12) define an infinitesimal transformations **S** in the $(t,q_i,\dot q_i)$-space, which is usually represented by the vector field

$$\mathbf S=h_0\cdot\frac{\partial}{\partial t}+\sum_{i=1}^n h_i\cdot\frac{\partial}{\partial q_i}+\sum_{i=1}^n(\dot h_i-\dot q_i\dot h_0)\frac{\partial}{\partial \dot q_i}. \tag{13}$$

Let us now show that this vector field **S** is, indeed, a pseudosymmetry of the vector field **X**,

$$\mathbf X\overset{\text{def}}{=}1\cdot\frac{\partial}{\partial t}+\sum_{i=1}^n\dot q_i\cdot\frac{\partial}{\partial q_i}+\sum_{i=1}^n f_i\cdot\frac{\partial}{\partial\dot q_i}, \tag{14}$$

associated with EQs. (7).

In fact, substituting (13) and (14) into (10), we get

$$\begin{aligned}&-\frac{\partial h_0}{\partial t}-\sum_i\dot q_i\frac{\partial h_0}{\partial q_i}-\sum_i f_i\frac{\partial h_0}{\partial\dot q_i}=\rho,\\ &(\dot h_i-\dot q_i\dot h_0)-\left(\frac{\partial h_i}{\partial t}+\sum_j\dot q_j\cdot\frac{\partial h_i}{\partial q_j}+\sum_j f_j\frac{\partial h_i}{\partial\dot q_j}\right)=\rho\dot q_i,\\ &h_0\frac{\partial f_i}{\partial t}+\sum_j h_j\cdot\frac{\partial f_i}{\partial q_j}+\sum_j(\dot h_j-\dot q_j\dot h_0)\frac{\partial f_i}{\partial\dot q_j}\\ &\quad-\frac{\partial}{\partial t}(\dot h_i-\dot q_i\dot h_0)-\sum_j\dot q_j\frac{\partial}{\partial q_j}(\dot h_i-\dot q_i\dot h_0)\\ &\quad-\sum_j f_j\frac{\partial}{\partial\dot q_j}(\dot h_i-\dot q_i\dot h_0)=\rho f_i,\end{aligned} \tag{15}$$

or, equivalently,

$$-\dot h_0=\rho,\quad -\dot q_i\dot h_0=\rho\dot q_i, \tag{16}$$

$$h_0\frac{\partial f_i}{\partial t}+\sum_j h_j\frac{\partial f_i}{\partial q_j}+\sum_j(\dot h_j-\dot q_j\dot h_0)\frac{\partial f_i}{\partial\dot q_j}-\frac{d}{dt}(\dot h_i-\dot q_i\dot h_0)=(-\dot h_0)f_i. \tag{17}$$

But Eqs. (17) are just the equations that one obtains if (11) is considered as a symmetry of Eqs. (1). Indeed, the transformations induced by (11) and (12) on the $\ddot q_i$ are given by

$$\begin{aligned}\ddot{\bar q}_i&=\ddot q_i+\epsilon\left[\frac{d}{dt}(\dot h_i-\dot q_i\dot h_0)-\ddot q_i\cdot\dot h_0\right]\\ &=\ddot q_i+\epsilon\left[\frac{d}{dt}(\dot h_i-\dot q_i\dot h_0)-f_i\cdot\dot h_0\right],\end{aligned} \tag{18}$$

and, therefore, the conditions to be satisfied in order that (11), (12), and (18) be a symmetry of equations (1) are[1]

$$\begin{aligned}&\ddot q_i+\epsilon\left[\frac{d}{dt}(\dot h_i-\dot q_i\dot h_0)-f_i\dot h_0\right]\\ &\quad=f_i+\epsilon\left[\frac{\partial f_i}{\partial t}\cdot\dot h_0+\sum_j\frac{\partial f_i}{\partial q_j}h_j+\sum_j\frac{\partial f_i}{\dot q_j}(\dot h_j-\dot q_j\dot h_0)\right];\end{aligned}$$

that is,

$$\frac{d}{dt}(\dot h_i-\dot q_i\cdot\dot h_0)-f_i\cdot\dot h_0=\frac{\partial f_i}{\partial t}h_0+\sum_j\frac{\partial f_i}{\partial q_j}h_j+\sum_j\frac{\partial f_i}{\dot q_j}(\dot h_j-\dot q_j\cdot\dot h_0),$$

which are simply Eqs. (17).

The importance of this result is, for the moment, purely geometrical, since the result permits passage from the non-pointlike symmetries (11) of (1) to the pointlike pseudosymmetries **S** of (14), given by (13). In the next sections we shall get a slight modification of this result and make full use of it.

IV. THE l.i.s. OF A SET OF DIFFERENTIAL EQUATIONS AND THE SYMMETRIES OF THE ASSOCIATED DYNAMICAL SYSTEM. APPLICATIONS

We have seen in Sec. III that if (11) is to be a symmetry of Eqs. (1), Eqs. (17) must be satisfied. Moreover, the pseudosymmetry factor ρ of the corresponding pseudosymmetry (13) of the dynamical system (14) is given by Eq. (16):

$$\rho = -\frac{dh_0}{dt}. \tag{19}$$

Therefore, if we begin with an l.i.s. of (1) such that $h_0(t,q_i,\dot{q}_i)$ is a constant function then, according to (19), the pseudosymmetry factor ρ would be equal to zero. But in this case (10) can be written

$$[\bar{\mathbf{S}},\bar{\mathbf{X}}] = 0, \tag{20}$$

that is, the associated **S** is, in this case, a symmetry vector field of **X**, transforming not only the trajectories of (14) into themselves but also the solutions of (7) into themselves.[2]

Now, we are now going to show that out of any symmetry (11) of Eqs. (1) we can get the following one:

$$\bar{q}_i = q_i + \epsilon(h_i - \dot{q}_i \cdot h_0), \quad \bar{t} = t + \epsilon \cdot 0, \tag{21}$$

in which t is not transformed at all.

Accordingly the associated vector field $\tilde{\mathbf{S}}$, corresponding to (21) in the $(t,q_i,\dot{q}_i)$ space and given [see Eq. (13)] by

$$\tilde{\mathbf{S}} = 0\frac{\partial}{\partial t} + \sum_i (h_i - \dot{q}_i \cdot h_0)\frac{\partial}{\partial q_i} + \sum_i (\dot{h}_i - f_i h_0 - \dot{q}_i \dot{h}_0)\frac{\partial}{\partial \dot{q}_i}, \tag{22}$$

is a symmetry vector of the dynamical system (7).

We recall again here that in (22) $\dot{h}_i$ and $\dot{h}_0$ are to be considered as functions of $(t,q_i,\dot{q}_i)$obtained by substituting, where necessary, for $\ddot{q}_i, f_i(t,q_i,\dot{q}_i)$.

We show now that (21) is indeed a symmetry of Eqs. (1) where (11) is a symmetry of them. In fact, from (21) and by the standard prolongation procedure, we get

$$\begin{aligned}\dot{\bar{q}}_i &= \dot{q}_i + \epsilon(\dot{h}_i - \ddot{q}_i h_0 - \dot{q}_i \dot{h}_0)\\ &= \dot{q}_i + \epsilon(\dot{h}_i - f_i h_0 - \dot{q}_i \dot{h}_0),\end{aligned} \tag{23}$$

$$\ddot{\bar{q}}_i = \ddot{q}_i + \epsilon\left(\frac{d}{dt}(\dot{h}_i - \dot{q}_i \dot{h}_0) - \frac{d}{dt}(f_i \cdot h_0)\right),$$

and, therefore, the conditions in order that (21) and (23) be a symmetry of (1) are

$$\begin{aligned}&\frac{d}{dt}(\dot{h}_i - \dot{q}_i \dot{h}_0) - \frac{d}{dt}(f_i \cdot h_0)\\ &= \left(\frac{\partial f_i}{\partial t}\right)\cdot 0 + \sum_j \frac{\partial f_i}{\partial q_j}(h_j - \dot{q}_j \cdot h_0)\\ &\quad + \sum_j \frac{\partial f_i}{\partial \dot{q}_j}(\dot{h}_j - f_j h_0 - \dot{q}_j \cdot \dot{h}_0)\end{aligned}$$

or, equivalently,

$$\begin{aligned}&\frac{d}{dt}(\dot{h}_i - \dot{q}_i \dot{h}_0) - f_i \dot{h}_0 - h_0\frac{df_i}{dt}\\ &= \sum_j \frac{\partial f_i}{\partial q_j} h_j + \sum_j \frac{\partial f_i}{\partial \dot{q}_j}(\dot{h}_j - \dot{q}_j \dot{h}_0)\\ &\quad - h_0\left(\sum_j \frac{\partial f_i}{\partial q_j}\dot{q}_j + \sum_j \frac{\partial f_i}{\partial \dot{q}_j} f_j\right),\end{aligned}$$

which after simplification become

$$\begin{aligned}&\frac{d}{dt}(\dot{h}_i - \dot{q}_i \dot{h}_0) - f_i \dot{h}_0\\ &= h_0\frac{\partial f_i}{\partial t} + \sum_j \frac{\partial f_i}{\partial q_j} h_j + \sum_j \frac{\partial f_i}{\partial \dot{q}_j}(\dot{h}_j - \dot{q}_j \dot{h}_0),\end{aligned} \tag{24}$$

which Eqs. (24) are exactly equal to Eqs. (17).

Some applications

The importance of the above results is not purely academic. In fact, it has been shown elsewhere[3] that if a certain dynamical system **X** possesses a constant divergence [i.e., for the case of (8) when $\Sigma_1^r \partial X_i/\partial x_i$ is a constant function] and if we are able to find a symmetry vector **S** of **X**, then the divergence of **S** is a first integral of **X**.

Assume, therefore, that the vector field **X** (14) associated with Eqs. (1) has a constant divergence, that is,

$$\text{div}\mathbf{X} = \sum_{i=1}^{n} \frac{\partial f_i}{\partial \dot{q}_i} = \kappa \tag{25}$$

[note that a quite broad class of nonlinear differential systems (1) does satisfy this condition]. In that case, if (11) is an l.i.s. of (1), then $\tilde{\mathbf{S}}$ is a symmetry of **X** and, accordingly, the function

$$\text{dim}\tilde{\mathbf{S}} = \frac{\partial}{\partial t}(0) + \sum_j \frac{\partial}{\partial q_i}(h_i - \dot{q}_i h_0) + \sum_i \frac{\partial}{\partial \dot{q}_i}(\dot{h}_i - f_i \cdot h_0 - \dot{q}_i \cdot \dot{h}_0) \tag{26}$$

would be a first integral of **X** associated with the l.i.s. (11). We obtain, therefore, without any need of knowing that Eqs. (1) are equivalent to a Lagrangian system,[4] a direct connection among the l.i.s. and the first integrals of (14) [or what is the same, of the first integrals of Eqs. (1)].

On the other hand, if we are able to *find* a Lagrangian function L such that Eqs. (1) are equivalent to the Euler–Lagrange associated with such a function L, then it is well known that one can introduce the new local coordinates $(t,q_1,...,q_n,p_1,...,p_n)$, the p_i being *defined* by

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \tag{27}$$

which permit us to obtain locally $\dot{q}_i$ in terms of $p_1,..., p_n$, when L is nondegenerate (i.e., $\det(\partial^2 L/\partial \dot{q}_i \partial \dot{q}_j) \neq 0$ on a certain open domain). In these local coordinates the dynamical system **X′** whose solutions provide the solutions of Eqs. (1), adopts the form

$$\mathbf{X}' = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i} - \sum_{i=1}^{n} \frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i} + 1\cdot\frac{\partial}{\partial t}, \tag{28}$$

H being the standard Hamiltonian function, defined by

$H(t,q_i,p_i)$

$$= \sum_{i=1}^{n} p_i \cdot \dot q_i(t,q_1,\ldots,q_n;p_1,\ldots,p_n) - L(t,q_i,\dot q_i(t,q,p)). \quad (29)$$

In these coordinates it is clear that the divergence of $\mathbf{X}'$ is equal to zero since

$$\sum_i \frac{\partial}{\partial q_i}\left(\frac{\partial H}{\partial p_i}\right) + \sum_i \frac{\partial}{\partial p_i}\left(-\frac{\partial H}{\partial q_i}\right) + \frac{\partial}{\partial t}(1) = 0. \quad (30)$$

On the other hand, if $\mathbf{S}$ is the symmetry of (1) corresponding to a given symmetry (11) of Eqs. (1) and since the symmetry condition

$$[\mathbf{X},\tilde{\mathbf{S}}] = \bar 0, \quad (31)$$

is maintained under any change of the coordinates, calling $\tilde{\mathbf{S}}'$ the vector field $\tilde{\mathbf{S}}$ when it is expressed in the coordinates (t,q,p) and writing

$$\tilde{\mathbf{S}} = \sum_{i=1}^{n} S_i \frac{\partial}{\partial q_i} + \sum_{i=1}^{n} T_i \frac{\partial}{\partial p_i} + S_0 \frac{\partial}{\partial t}, \quad (32)$$

the divergence of $\tilde{\mathbf{S}}'$ *will be a first integral of* $\mathbf{X}'$. That is,

$$\sum_{i=1}^{n} \frac{\partial S_i}{\partial q_i} + \sum_{i=1}^{n} \frac{\partial T_i}{\partial p_i} + \frac{\partial S_0}{\partial t} \quad (33)$$

is a first integral of $\mathbf{X}'$. Now since the first integrals of a dynamical system $\mathbf{X}$ are the solutions of

$$\mathbf{X}(I) = 0, \quad (34)$$

which equation is independent of any coordinate change, then writing (33) in terms of the old variables $(t,q,\dot p)$, we would obtain a first integral of the original dynamical system $\mathbf{X}$ and of the original set of Eqs. (1) as well. Note that the symmetry used in obtaining the first integral (33), written in the $q,\dot q$ coordinates, was a symmetry of the differential equations and, accordingly, has nothing to do (in general) with the symmetries of the action

$$\int_{t_0}^{t_1} L\,dt. \quad (35)$$

For this reason, the connection established here has little to do with the standard Noether theorem[5] in which the connection among symmetries and first integrals is based on the symmetries of the action and not of the differential equations associated with L.

The explicit obtention of the functions S_i and T_i of (33) from $\tilde{\mathbf{S}}$ as given in (22) offers no difficulty. Recalling that the general expression giving the new components of a vector field [i.e., the vector field $\mathbf{X}$ of equations (8)] in the new coordinates x_i' is

$$x_i' = \sum_j X_i \frac{\partial x_i'}{\partial x_j}, \quad (36)$$

we get

$$S_i = h_i - \dot q_i - h_0,$$
$$T_i = \sum_{j=1}^{n} (\dot h_j - f_j h_0 - \dot q_j \dot h_0)\frac{\partial p_i}{\partial \dot q_j}, \quad (37)$$
$$S_0 = 0,$$

and therefore the first integral (33) can be written as

$$\sum_{i=1}^{n} \frac{\partial}{\partial q_i}[h_i - \dot q_i h_0]_{(q,p)} + \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial}{\partial p_i}\left[(\dot h_j - f_j h_0 - \dot q_j \dot h_0)\frac{\partial^2 L}{\partial \dot q_i \partial \dot q_j}\right]_{(q,p)}. \quad (38)$$

We recall again that in (38) the previous substitution of $\dot q_i$ by its values in function of (t,q,p) has to be made before the calculation of the partial derivatives $\partial/\partial q_i$ and $\partial/\partial p_i$ in the terms $h_i - \dot q_i h_0$ and $(\dot h_j - f_j h_0 - \dot q_j \dot h_0)\partial^2 L/\partial \dot q_i \partial \dot q_j$. This is indicated by the subindexes (q,p) appearing in (38). If we desire to omit these subindexes, we should write (38) in the more involved form

$$\sum_{i=1}^{n} \frac{\partial}{\partial q_i}(h_i - \dot q_i h_0) + \sum_{i,k=1}^{n} \frac{\partial}{\partial p_k}(h_i - \dot q_i h_0)\frac{\partial p_k}{\partial q_i} + \sum_{i,j,k}^{n} \frac{\partial}{\partial \dot q_k}\left((\dot h_j - f_j h_0 - \dot q_j - \dot h_0)\frac{\partial^2 L}{\partial \dot q_i \partial \dot q_j}\frac{\partial \dot q_k}{\partial p_i}\right)$$
$$= \sum_{i=1}^{n} \frac{\partial}{\partial q_i}(h_i - \dot q_i h_0) + \sum_{i,k,l}^{n} \frac{\partial}{\partial \dot q_p}(h_i - \dot q_i h_0)\frac{\partial \dot q_l}{\partial p_k}\frac{\partial^2 L_\partial}{\dot q_k \partial q_i} \times \sum_{i,j,k}^{n} \frac{\partial}{\partial \dot q_k}\left((\dot h_j - f_j \cdot h_0 - \dot q_j \dot h_0)\frac{\partial^2 L}{\partial \dot q_i \partial \dot q_j}\right)\cdot\frac{\partial \dot q_k}{\partial p_i}, \quad (39)$$

where the derivatives $\partial \dot q_l/\partial p_k$ and $\partial \dot q_k/\partial p_i$ are given by

$$\frac{\partial \dot q_l}{\partial p_k} = \mathscr{H}_{lk}^{-1}, \quad \frac{\partial \dot q_k}{\partial p_i} = \mathscr{H}_{ki}^{-1} \quad (40)$$

these last expressions have been obtained from the equalities

$$\delta_{jl} = \frac{\partial p_j}{\partial p_l} = \sum_i \frac{\partial p_j}{\partial \dot q_i}\frac{\partial \dot q_i}{\partial p_l} = \sum_i \frac{\partial^2 L}{\partial \dot q_j \partial \dot q_i}\frac{\partial \dot q_i}{\partial p_l}, \quad (41)$$

which can be written in the matrix form

$$I = \mathscr{H}.\dot Q, \quad (42)$$

I being the identity matrix and $\mathscr{H}$ and $\dot Q$ the matrices defined by:

$$\mathscr{H}_{ij} = \left(\frac{\partial^2 L}{\partial \dot q_i \partial \dot q_j}\right), \quad \dot Q_{ij} = \frac{\partial \dot q_i}{\partial p_j}. \quad (43)$$

Finally, and although less interesting from the practical point of view, it is also of theoretical interest to note the reciprocal relationship that can be established among the first integrals of a Lagrangian system of Eqs. (1) when L is known and the symmetries of it. Concretely: With any given first integral of Eqs. (1), one can associate a symmetry of them of type (11). Before proceeding to show this, we first show that for any pseudosymmetry $\mathbf{S}$ of Eqs. (7),

$$\mathbf{S} \overset{\text{def}}{=} S \cdot \frac{\partial}{\partial t} + \sum_{i=1}^{n} S_i \frac{\partial}{\partial q_i} + \sum_{i=1}^{n} T_i \frac{\partial}{\partial \dot q_i}, \quad (44)$$

the monoparametric family of l.i.s. given by

$$\bar{q}_i = q_i + \epsilon S_i(t,q_1,...,q_n;\dot{q}_1,...,\dot{q}_n),$$
$$\bar{t} = t + \epsilon S_0(t,q_1,...,q_n;\dot{q}_1,...,\dot{q}_n) \tag{45}$$

is a symmetry of Eqs. (1).

Indeed, writing in explicit form the conditions expressed by

$$[\mathbf{S},\mathbf{X}] = \rho \mathbf{X},$$

for the vector field **X** given by (7) and for the **S** given by (44), we get the equations

$$-\frac{dS_0}{dt} = \rho, \tag{46}$$

$$T_i - \frac{dS_i}{dt} = \rho \cdot \dot{q}_i \tag{47}$$

$$\sum_j S_j \frac{\partial f_i}{\partial \dot{q}_j} + T_i - \frac{\partial S_i}{dt} = \rho f_i. \tag{48}$$

From (46) and (47) we obtain

$$T_i = \frac{dS_i}{dt} + \dot{q}_i \frac{dS_0}{dt} \tag{49}$$

which is exactly condition (12), indicating that the functions T_i of (44) are obtained by the standard procedure (see I) of extending the transformations (45) to the $(t,q,\dot{q})$-space.

On the other hand, conditions (48) are exactly conditions (17), as we discussed in Sec. III, showing that (45) is a symmetry of Eqs. (1), as we desired to prove.

We proceed now to prove that with any first integral of a Lagrangian system of differential equations (1) we can associate an l.i.s. of type (11).

Indeed, assume that we write the associated dynamical system **X** given by Eqs. (7) in the form (28). Let $f(t,q,\dot{q})$ be the first integral in the original $(t,q,\dot{q})$ variables and $F(t,q,p)$ its functional form in the (t,q,p) variables. Then it is quite easy to check out that the vector field:

$$\mathbf{S}_F = 0\frac{\partial}{\partial t} + \sum_{i=1}^{n} \frac{\partial F}{\partial p_i}\frac{\partial}{\partial q_i} - \sum_{i=1}^{n} \frac{\partial F}{\partial q_i}\frac{\partial}{\partial p_i} \tag{50}$$

is a symmetry of (28). Therefore, as we have just discussed the l.i.s. induced by $\mathbf{S}_F$ on the (t,q) space will be an l.i.s. of the original set of Lagrangian equations (1). Note that $\mathbf{S}_F$ is still a symmetry of (28) when F is no longer a first integral of (28) but satisfies the weaker condition

$$\{H,F\} + \frac{\partial F}{\partial t} = K, \tag{51}$$

$\{\,,\,\}$ being the standard Poisson bracket and K a constant. Note, as well, that we could think of obtaining an additional first integral of equations (1), out of the l.i.s.

$$\bar{q}_i = q_i + \epsilon \frac{\partial F}{\partial p_i}, \quad \bar{t} = t + \epsilon \cdot 0 \tag{52}$$

associated with $\mathbf{S}_F$, when F satisfies Eq. (51). But for this particular kind of l.i.s. we have

$$\text{div}(\mathbf{S}_F) = 0,$$

and, therefore, one only obtains, from a function F satisfying Eq. (51), a banal first integral, i.e., a constant function. For reasons of completeness we write now the form adopted by the l.i.s. (52) in terms of the original form $f(t,q,\dot{q})$ of the first integral given. This form is given by:

$$\bar{q}_i = q_i + \epsilon \sum_j \frac{\partial f}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial p_i}$$
$$\bar{t} = t + \epsilon \cdot 0$$

or, equivalently, using Eqs. (40),

$$\bar{q}_i = q_i + \epsilon \sum_{j=1}^{n} \frac{\partial f}{\partial \dot{q}_j} \mathscr{H}_{ji}^{-1}$$
$$\bar{t} = t + \epsilon \cdot 0. \tag{53}$$

Note that when L is a polynomial in the $\dot{q}$ of degree 2 and f is a polynomial of degree 1 in the $\dot{q}$, then (53) is a pointlike or geometrical transformation of type (2). Apart from this case the symmetry (53) associated with a first integral will be, in general, an l.i.s. of the general kind (11), that is, a dynamical symmetry instead of a geometrical symmetry.

V. CONCLUSIONS, FINAL REMARKS, AND OPEN QUESTIONS

We have obtained, for a quite broad class of systems of Newton-like differential equations (1), a first integral out of an l.i.s. of these systems. It is important to stress here that the method followed is not based on existence theorems, but it is of *computational* value. It may happen that in particular circumstances the first integral obtained be a constant (see the example of the symmetry $\mathbf{S}_F$ at the end of Sec. IV). It is of interest to remark here that our procedure is general enough to permit the extension of it to more general kinds of systems of differential equations than those of kind (1), but such that the divergence of the associated dynamical system is a constant function. The crucial step was the construction of a symmetry $\tilde{\mathbf{S}}$ of the associated dynamical system **X** out of a symmetry vector of the differential equations (1).

Additional first integrals of Eqs. (1), for instance those of the kind

$$\tilde{\mathbf{S}}(\text{div}\tilde{\mathbf{S}}), \quad \tilde{\mathbf{S}}(\tilde{S}(\text{div}\tilde{\mathbf{S}})),\cdots \tag{54}$$

(see Ref. 3), can be also obtained by our procedure. We have not written in the text the explicit coordinate expression of them since this expression is quite complicated.

It should be stressed as well that, although it has been asserted (see the paper by Havas, 1973, in Ref. 4) that any system of equations like (1) is always equivalent to a Lagrangian system, this conclusion is of theoretical value. Therefore, even if this is an important conclusion, its practical value cannot be fully exploited if one is *not* able to develop a *constructive* procedure in order to find a *concrete* Lagrangian function L from which Eqs. (1) could be obtained. Only *when* this L (at least *one* of the possible L's that exist according to Havas[4]) has been obtained, is it possible to *compute* the first integrals by the procedures delineated in this paper.

On the other hand the results obtained in Secs. III and IV are important since what they say amounts to the fact that in relation to the symmetries no new advantages can be obtained in the passage from (1) to (7), i.e., by increasing the number of the independent variables of the problem. What is gained, in any case, is a more geometrical picture of the symmetries since the dynamical symmetries (11) of (1) are converted into geometrical symmetries of the trajectories of (7).

As open problems connected with the paper we quote the following:

(1) To find conditions less restrictive than the condition (25).

(2) If (25) is *not* satisfied in the coordiantes $(t,q,\dot{q})$, it is necessary to find *computational* rules in order to pass from the coordinates $(t,q,\dot{q})$ to new local coordinates (T,Q,P) (such that in the new local coordinates the dynamical system $\mathbf{X}'$ obtained is now of constant divergence).

Note that when (1) is equivalent to a known Lagrangian system L, the computational rules, in order to achieve this goal, are given by the equations

$$T = t, \quad Q_i = q_i, \quad p_i = \frac{\partial L}{\partial \dot{q}_i}. \tag{55}$$

Note, as well, that from a theoretical point of view the canonical form theorem[6] assures the existence of local coordinated (T,Q,P) in which the vector field $\mathbf{X}$ associated with (7) adopts the form

$$\bar{\mathbf{X}}' = \sum_i 0 \cdot \frac{\partial}{\partial Q_i} + \sum_i 0 \cdot \frac{\partial}{\partial p_i} + 1 \cdot \frac{\partial}{\partial T}, \tag{56}$$

which has, obviously, zero divergence. The difficulty is that no *computational* rule [other than solving (7) completely] is known (in general) in order to achieve the reduction of (7) either to the above canonical form or to a form in which the vector field has constant divergence. Accordingly, the practical problem of finding *computationally* the appropriate coordinates in which a given vector field has constant divergence arises. *If* this problem were solved, then, and following the procedure indicated in this paper, one could construct for every set of differential equations (1), and out of any l.i.s. of them, a first integral of the original set of differential equations. The restriction (25) would be eliminated.

(3) A third, and more difficult problem, is the classification of the l.i.s. such that the associated first integral constructed above is a constant function. We have, so far, no idea how this problem could be attacked.

Additional applications of the results obtained here in order to solve the "inverse" problem concerning the symmetries of the set of differential equations, i.e., the problem of finding Eqs. (1), if any, admitting a certain *given* set of l.i.s. more or less physically motivated, are in progress.

ACKNOWLEDGMENTS

It is a pleasure to express here my gratitude to Professor C. Ruiz and to F. Moreno-Insertis for their continuous assistance. It is also a pleasure to acknowledge the constant encouragement given by M. C. Hidalgo-Brinquis. I thank as well John McCleary for his collaboration in the correction of the English.

[1]F. González-Gascón, J. Math. Phys. **18**, 1763 (1977).
[2]R. Hermann, *Differential Geometry and the Calculus of Variations* (Benjamin, New York, 1968), Part I. See, as well, F. González-Gascón, Lett. Nuovo Cimento **19**, 366 (1977) and the references therein.
[3]F. González-Gascón, Phys. Lett. A **61**, 375 (1977).
[4]P. Havas, Suppl. Nuovo Cimento **5**, 363 (1957); E. Engels, Nuovo Cimento B **26**, 481 (1975); P. Havas, Acta Physica Austr. **38**, 145 (1973).
[5]I. Gel'fand and S. Fomin, *Calculus of Variations*(Prentice-Hall, Englewood Cliffs, N.J., 1968).
[6]See. Ref. 2 (R. Hermann), pp. 39, 63–72.

A variational principle for resonances

A. G. Ramm [a)]
Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109

(Received 11 March 1980; accepted for publication 14 April 1980)

A stationary variational principle for calculation of the complex poles of Green functions is given.

1. INTRODUCTION

Consider the problem

$$(-\nabla^2 - k^2)u = 0 \text{ in } \Omega, \quad u|_\Gamma = 0, \tag{1}$$

where Ω is an exterior domain, Γ is its closed smooth boundary, $D = \mathbb{R}^3 \setminus \Omega$ is bounded. Problem (1) has nontrivial solution iff ($\equiv$ if and only if) k is a complex pole k_q of the Green function of the exterior Dirichlet problem. The nontrivial solution has the following asymptotic near infinity

$$u = r^{-1}\exp(ikr)\sum_{j=0}^{\infty} f_j(n,k)r^{-j}, \tag{2}$$

$$n = x/|x|, \quad r = |x|, f_0 \not\equiv 0.$$

If u and v are of the form (2) with $k = k_1$ and $k = k_2$, respectively, $\text{Re}(k_1 + k_2) \neq 0$, $\pi < \arg k_j < 2\pi$, $j = 1,2$, then the following limit exists

$$\langle u,v\rangle = \lim_{\epsilon\to+0}\int \exp(-\epsilon r \ln r)u(x)v(x)\,dx, \quad \int = \int_\Omega. \tag{3}$$

This will be proved in Sec. 2. From (1) and (3) it follows that

$$\text{st}K(u) = \text{st}\{\langle\nabla u,\nabla u\rangle/\langle u,u\rangle\} = k^2, \tag{4}$$

where st means the stationary value and the admissible functions vanish on Γ and are of the form (2) near infinity. The stationary principle (4) looks like Rayleigh–Ritz quotient but is actually different in the following respects: (i) the functional (4) is complex-valued, the variational principle is a stationary one and not an extremal as for the usual Rayleigh–Ritz functionals; (ii) the functions which give stationary values to K are growing exponentially at infinity.

The variational principle (4) can be used for calculations as follows:

(1) Take a test function of the form

$$u_N = r^{-1}\exp(ikr)\sum_{j=0}^{N}\sum_{|m|\leqslant j} r^{-j}Y_{jm}(n)c_{jm}(k)\cdot g(x), \tag{5}$$

where $Y_{jm}(n)$ are the spherical harmonics, $Y_{jm}(n) = P_{j,m}(\cos\theta)\cdot\exp(im\phi)$, $n = (\theta,\phi)$; $P_{j,m}(\cos\theta)$ are the associated Legendre polynomials; $c_{jm}(k)$ do not depend on r,n; N is a fixed number; $g(x)\geqslant 0$ is a fixed smooth function which is equal to 1 outside of a ball which contains D and which is equal to zero on Γ.

(2) Put (5) in (4) and use the necessary conditions for K to be stationary: $\partial K/\partial c_{jm} = 0$.

Because the numerator and denominator in (4) are quadratic forms in c_{jm} we can write the above condition as

$$\sum_{q=0}^{Q}[a_{sq}(k) - k^2 b_{sq}]\,c_q(k) = 0, \quad 0\leqslant s\leqslant Q, \tag{6}$$

where for brevity we denote by one index q the double index jm and by Q the maximal value of q which is defined by N. We took into account also that the Lagrange multiplier is equal to k^2 according to (4). The elements $a_{sq}(k)$ and $b_{sq}(k)$ can be explicitly expressed in the form

$$a_{sq}(k) = \langle\nabla\{b(x,k)F_s(r,n)\}, \nabla\{b(x,k)F_q(r,n)\}\rangle, \tag{7}$$

$$b_{sq}(k) = \langle b(x,k)F_s(r,n), b(x,k)F_q(r,n)\rangle, \tag{8}$$

where

$$b(x,k) = g(x)r^{-1}\exp(ikr), \quad F_q(r,n) = r^{-j}Y_{jm}(n). \tag{9}$$

The elements (7), (8) are entire functions of k of exponential type, i.e., the inequalities $|a_{sq}(k)|\leqslant c\exp[A|k|]$ hold, where $c = \text{const} > 0$ and $A = \text{const}$ which depends on D but does not depend on N. The system (6) has a nontrivial solution iff

$$\det\{a_{sq}(k) - k^2 b_{sq}(k)\} = 0. \tag{10}$$

This equation has infinitely many roots $k_l^{(Q)}$, $l = 1,2,\cdots$, generally speaking, since its left-hand side is an entire function of k.

(3) The mathematical question related to this numerical scheme can be formulated as follows: is it true that $k_l^{(Q)}\to k_l$ as $Q\to\infty$, where k_l are the complex poles of the Green function?

2. EXISTENCE OF THE LIMIT (3)

First let us note that it is enough to prove that for sufficiently large $R > 0$ the following limit exists

$$\lim_{\epsilon\to+0}\int_{|x|\geqslant R}\exp(-\epsilon r\ln r)u(x)v(x)\,dx.$$

For $|x|\geqslant R$ we can use series (2) representing u and v. These series converge absolutely and uniformly in n and r, $r\geqslant R$. Therefore it is enough to prove existence of the limit

$$\lim_{\epsilon\to+0}\int_R^\infty \exp(-\epsilon r\ln r)r^{-j}\exp(br + iar)\,dr,$$

where

$$j\geqslant -2,$$

$$b = -\text{Im}(k_1 + k_2) > 0, \quad a = \text{Re}(k_1 + k_2)\neq 0.$$

Suppose that $a > 0$. Let

$$C_N = \{z:|z - R| = N, \quad 0\leqslant\arg(z - R)\leqslant\theta\},$$

$$C_{\theta N} = \{z:\arg(z - R) = \theta, \quad 0\leqslant|z - R|\leqslant N\},$$

$$C_R = \{z:R\leqslant z\leqslant R + N\}, \quad C_\theta = C_{\theta\infty}, \quad C = C_N\cup C_{\theta N}\cup C_R.$$

It is clear that

[a)] Supported by AFOSR F4962079C0128.

 0022-2488/80/082052-02$1.00

$$\int_{C_N} \exp(-\epsilon r \ln r) r^{-j} \exp(br + iar)\, dr \to 0$$

as $N \to \infty$, $\forall \epsilon > 0$. Thus

$$\int_R^\infty \exp(-\epsilon r \ln r) r^{-j} \exp(br + iar)\, dr$$
$$= \int_{C_\theta} \exp(-\epsilon r \ln r) r^{-j} \exp(br + iar)\, dr.$$

Let us take $0 < \theta \leqslant \pi/2$ such that a $\sin\theta > n \cos\theta$. Then the integral over C_θ converges absolute for $\epsilon > 0$ and its limit as $\epsilon \to +0$ exists.

The case $a < 0$ can be considered similarly, with $-\theta$ instead of θ. This completes the proof.

Remark 1: The limit (3) was used by B. Vainberg[1] in connection with the orthogonality of the generalized eigen and root functions, corresponding to different complex poles of the Green function, but our argument differs from the argument in Ref. 1. In Ref. 2 the Green function of the Schrödinger operator with a compactly supported potential was considered and the limit with the weight function $\exp(-\epsilon r^2)$ instead of $\exp(-\epsilon r \ln r)$ was considered. There is a mistake in calculation in Ref. 1: the authors claim that the limit $\lim_{\epsilon \to +0} \int_\Omega \exp(-\epsilon r^2)\, uv\, dx$ exists for any k_1, k_2, but this limit does not exist for $5\pi/4 < \arg(k_1 + k_2) < 7\pi/4$. In Refs. 3 and 6 a method for calculation of the complex poles of the Green functions in diffraction problems and in the potential scattering problem was given and justified. This method used Galerkin-type procedure. In Refs. 4, 5, and 7 some facts about the location and properties of the complex poles are given.

Remark 2: Justification of the numerical approach suggested in Sec. 1 is an open mathematical problem. For some other variation principles in nonselfadjoint problems the importance of the mathematical analysis of the situation was mentioned in Ref. 9. In Refs. 3 and 6 the mathematical justification of the numerical approach described in Refs. 3 and 6 was based on the compactness of the integral operators in the equations to which the problem of calculation of the complex poles was reduced in Refs. 3 and 6. In the situation described in this paper the operator is essentially $-\nabla^2 - k^2$ with Dirichlet boundary condition and it is noncompact.

[1] B. Vainberg, "On eigenfunctions of an operator corresponding to the poles of the analytic continuation of the resolvent through the continuous spectrum," Math. USSR, Sb. **16**, 302–22 (1972).

[2] A. Baz, Ja. Zeldovich, and A. Perelomov, *Diffusion, Reaction and Decay in Nonrelativistic Quantum Mechancis* (Nauka, Moscow, 1966). (Russian).

[3] A. G. Ramm, "Computation of quasistationary states in nonrelativistic quantum mechanics,"Dokl. Akad. Nauk SSSR **204**, 1071–74 (1972) [Sov. Phys. Dokl. **17**, 522–4 (1972)].

[4] A. G. Ramm, "Regions free of resonance poles in the scattering problem for a three-dimensional potential, Dokl. Akad. Nauk SSSR **166**, 114–6 (1966), I [Sov. Phys. Dokl. **11**, 114–116 (1966)].

[5] A. G. Ramm, "Expansions in eigenfunctions of the nonselfadjoint Schrödinger operator, Dokl. Akad. Nauk SSSR **191**, 50–3 (1970) [Sov. Phys. Dokl. **15**, 231–4 (1970)].

[6] A. G. Ramm, "Exterior problems of diffraction," Radio Eng. Electron. Phys. (USSR) **17**, 1064–7 (1972).

[7] A. G. Ramm, "Eigenfunction expansion of a discrete spectrum in diffraction problems," Radio Eng. Electron. Phys. (USSR) **18**, 364–9 (1973).

[8] A. G. Ramm, "Nonselfadjoint operators in diffraction and scattering," Math. Methods Appl. Sci. **2** (1980).

[9] V. Voitovich, B. Kacenelenbaum, and A. Sivov, *Generalized Method of Eigenoscillations in Diffraction Theory* (Nauka, Moscow, 1977) (Russian).

WKB method for systems of integral equations [a)]

H. L. Berk and D. Pfirsch [b)]
University of California, Lawrence Livermore Laboratory, Livermore CA 94550

(Received 25 January 1980; accepted for publication 8 March 1980)

The WKB theory for vector systems of integral equations is developed herein. A variational technique is used to derive the equations for the WKB amplitudes in x-space or its dual k-space. Compact, explicit solutions are obtained in one dimension. When a solution breaks down at a turning point, the dual-space representation can be used to derive the connection formulas between WKB solutions. These connection formulas are equivalent to the rules of the Furry method. The Furry method is used to show how general golbal-dispersion relations can be constructed.

I. INTRODUCTION

The Wentzel–Kramers–Brillouin (WKB) method to approximate solutions of differential equations is well known in quantum mechanics[1] and geometrical optics.[2] It has usually been applied to solve differential equations and it has occasionally been used to solve integral equations.[3-5] The purpose of this paper is to present a systematic derivation of the WKB method for determining normal modes of vector systems of integral equations. These equations arise from the linear perturbations of physical systems obeying integral equations where one dimension is spatially varying. The technique described here is a general asymptotic method which includes as a subset differential equations of arbitrary order. An important application of this technique is in the field of plasma physics, where the linearized Vlasov–Maxwell equations lead to a three-component field equation satisfying integral equations for which there is a spectrum of complex eigenvalues.

The geometrical optics theory for a vector system varying in three dimensions has been described by Bernstein.[6] However, Bernstein's paper is limited to wave propagation and does not treat the eigenvalue problem which is readily formulated when there is only one dimension of inhomogeneity. We shall also see that for an inhomogeneity in one dimension compact expressions for the wave amplitudes can be derived. WKB treatments for one component of polarization have been described in Refs. 3–5. Here we extend the technique to an arbitrary number of components, and present a careful analysis of the theory.

II. SUMMARY OF RESULTS

In infinite-medium theory, waves with an amplitude ζ_j will satisfy a set of equations given by

$$\sum_j \int d\mathbf{r}' G_{ij}(\mathbf{r}-\mathbf{r}',\omega)\zeta_j(\mathbf{r}') = 0. \tag{1}$$

The solution for this system is

$$\zeta_j(\mathbf{r}) = \hat{\zeta}_j \exp(i\mathbf{k}\cdot\mathbf{r}), \tag{2}$$

where $\mathbf{k}$ is usually real, and ω satisfies the relation

$$|\Lambda(\mathbf{k},\omega)| = 0, \tag{3}$$

where $|\Lambda(\mathbf{k},\omega)|$ is the determinant of

$$\Lambda_{ij}(\mathbf{k},\omega) = \int_{-\infty}^{\infty} d\mathbf{r} \exp(-i\mathbf{k}\cdot\mathbf{r}) G_{ij}(\mathbf{r},\omega). \tag{4}$$

Further, the polarization amplitudes $\hat{\zeta}_i$ satisfy the relation

$$\frac{\hat{\zeta}_i}{\hat{\zeta}_j} = \frac{\Lambda^{ir}(\mathbf{k},\omega)}{\Lambda^{jr}(\mathbf{k},\omega)}, \tag{5}$$

for arbitrary r where Λ^{ij} is the cofactor of Λ_{ij}, i.e.,

$$\sum_j \Lambda_{ij}\Lambda^{jr} = |\Lambda|\delta_{ir}. \tag{6}$$

To generalize these results to a system which varies in one dimension, we consider an integral equation of the form

$$\sum_j \int_{-\infty}^{\infty} d\mathbf{r}'\, G_{ij}\Big(\mathbf{r}-\mathbf{r}', \epsilon\frac{x+x'}{2}, \omega\Big)\zeta_j(\mathbf{r}') = 0. \tag{7}$$

We define

$$\Lambda_{ij}(\mathbf{k},\epsilon x,\omega) = \int d\mathbf{z} \exp(-i\mathbf{k}\cdot\mathbf{z}) G_{ij}(\mathbf{z},\epsilon x,\omega). \tag{8}$$

The WKB solution is then of the form

$$\zeta_j(x) = \sum_p \hat{\zeta}_j^p(\epsilon x) \exp\Big[i\int_{x_0}^{x} k_x^{(p)}(\epsilon x')\,dx' + ik_y y + ik_z z\Big], \tag{9}$$

where x_0 is an arbitrary reference point

$$\sum_j \Lambda_{ij}[\mathbf{k}^{(p)}(\epsilon x),\epsilon x,\omega]\hat{\zeta}_j^p(\epsilon x) = 0, \tag{10}$$

$\mathbf{k}^p(\epsilon x) = k_x(\epsilon x)\hat{\mathbf{x}} + k_y\hat{\mathbf{y}} + k_z\hat{\mathbf{z}}$ satisfies

$$|\Lambda[\mathbf{k}^{(p)}(\epsilon x),\epsilon x,\omega]| = 0, \tag{11}$$

where $k_x(\epsilon x)$ is a complex function, and

$$\frac{\hat{\zeta}_j^p(\epsilon x)}{\hat{\zeta}_i^p(\epsilon x)} = \frac{\Lambda^{jr}[\mathbf{k}^{(p)}(\epsilon x),\epsilon x,\omega]}{\Lambda^{ir}[\mathbf{k}^{(p)}(\epsilon x),\epsilon x,\omega]}, \tag{12}$$

for arbitrary r. We now suppress the dependence on k_y and

[a)] Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore Laboratory under Contract No. W-7405-ENG-48.

[b)] Permanent address: Max Planck Institute, Garching, German Federal Republic.

 0022-2488/80/082054-13$1.00

k_z, and $k(\epsilon x)$ refers to $k_x(\epsilon x)$. Further, the p index will be dropped to simplify notation.

At this stage the solutions are similar to the infinite-medium-theory problem, but the overall variation of $\hat{\xi}_j(\epsilon x)$ from point to point needs to be given. We find that it is given by

$$\hat{\xi}_v(\epsilon x) = \frac{C_s \langle^v_{x_0}\{\Lambda\,[k(\epsilon x),\epsilon x]\}\rangle}{\{(\partial/\partial k)|\Lambda\,[k(\epsilon x),\epsilon x,\omega)]|\}^{1/2}} \times \Lambda^{vs}[k(\epsilon x_0),\epsilon x_0], \tag{13}$$

where C_s is a space-independent constant and the polarization component given in the numerator is

$$\begin{aligned}\langle^v_{x_0}\{\Lambda\,[k(\epsilon x),\epsilon x]\}\rangle &= \frac{\Lambda^{vs}[k(\epsilon x),\epsilon x]}{\Lambda^{vs}[k(\epsilon x_0),\epsilon x_0]}\left(\frac{\Lambda^{rs}[k(\epsilon x_0),\epsilon x_0]}{\Lambda^{rs}[k(\epsilon x),\epsilon x]}\right)^{1/2}\cdot\exp\Big[-\int_{x_0}^{x}\epsilon\,dx' \\ &\times\sum_{i,j}\frac{[(\partial/\partial y)\Lambda^{ri}(k,y)\Lambda_{ij}(k,y)(\partial/\partial k)\Lambda^{js}(k,y)-(\partial/\partial k)\Lambda^{ri}(k,y)\Lambda_{ij}(k,y)(\partial/\partial y)\Lambda^{is}(k,y)]}{2(\partial/\partial k)|\Lambda(k,y)|\Lambda^{rs}(k,y)}\bigg|_{\substack{k=k(\epsilon x')\\ y=\epsilon x'}}\Big],\end{aligned} \tag{14}$$

and can be interpreted as a propagator of the polarization amplitude from the point x_0 to x for arbitrary r, s, and x_0. The most compact form of this solution is $r = s = v$. For symmetric self-adjoint kernels which have the properties that $G_{ij}(x,x') = G_{ji}(x',x)$ and $G_{ij}(x,x') = G_{ij}(x',x)$, the argument of the exponential is zero when $s = r$, and the amplitude factor reduces to

$$\begin{aligned}\langle^v_{x_0}\{\Lambda\,[k(\epsilon x),\epsilon x]\}\rangle &= \frac{\Lambda^{vr}[k(\epsilon x),\epsilon x]}{\Lambda^{vr}[k(\epsilon x_0),\epsilon x_0]}\left(\frac{\Lambda^{rr}[k(\epsilon x_0),\epsilon x_0]}{\Lambda^{rr}[k(\epsilon x),\epsilon x]}\right)^{1/2} \\ &= \left(\frac{\Lambda^{vv}[k(\epsilon x),\epsilon x]}{\Lambda^{vv}[k(\epsilon x_0),\epsilon x_0]}\right)^{1/2}.\end{aligned} \tag{15}$$

(Self-adjointness is not equivalent to Hermitian systems; e.g., in plasma physics, a suitable basis can be found where the kernel is symmetric-self-adjoint and dissipation is present from Landau damping.)

The WKB solution is consistent with a concept of action. We define the action of mode p as

$$A^p = \sum_{i,j}\hat{\xi}_i^{+p}(\epsilon x)\frac{\partial}{\partial\omega}\Lambda_{ij}[k(\epsilon x),\epsilon x,\omega]\hat{\xi}_j^p(\epsilon x), \tag{16}$$

where

$$\xi_i^{+p}(\epsilon x) = \hat{\xi}_i^{+p}(\epsilon x)\exp\left[-i\int^x k(\epsilon x')\,dx'\right]$$

is the WKB solution of the adjoint equation

$$\sum_j\int_{-\infty}^{\infty}dx'\,G_{ji}\left(x'-x,\epsilon\frac{x'+x}{2},\omega\right)\xi_j^+(x') = 0.$$

We find that the action flow, $A^p v_g = \text{const}$, where

$$v_g = -\frac{(\partial/\partial k)|\Lambda\,[k(\epsilon x),\epsilon x,\omega]|}{(\partial/\partial\omega)|\Lambda\,[k(\epsilon x),\epsilon x,\omega]|} \tag{17}$$

is the group velocity, which, in general, is complex.

The group velocity vanishes where $|\Lambda|_k = 0$, which is also the condition for roots of $k(\epsilon x)$ to merge and the point where the WKB solution fails. In the vicinity of the turning point the problem is solved an alternate way. The alternate method is to develop the WKB solution in the Fourier transformed space (k-space). We find that the equations in k-space are the "dual" to the x-space equations as identical equations arise but with the variables $x,k(x)\Rightarrow k,x(k)$. Such duality has previously been observed by Percival.[7] Thus, the solution in k-space is exactly the same as the x-space solutions with the x and k parameters interchanged. By Fourier-transforming back to x-space, an accurate solution near the x-space turning point is obtained. One can then find rules of how to connect the wave amplitudes of interacting waves (i.e., waves whose roots merge at the turning point). These rules can also be derived by generalizing the Furry method as described by Heading.[8]

From the connection rules of merging waves one can derive global dispersion relations for the eigenvalue ω. The most common dispersion relation arises from the two-turning-point problem of two waves, which reduces to the Bohr–Sommerfeld quantum rule[1] if the two k roots are of opposite sign. When wave tunnelling can be neglected, the more general turning point problem reduces to the form

$$\oint_C [k(z,\omega) + \delta k(z,\omega)]\,dz = (2n+1)\pi, \tag{18}$$

where n is an integer and the contour C encloses all the turning points in the complex z-plane which are branch points of the function $k(z,\omega)$, and $\delta k(z,\omega)$ is a correction to the usual answer that can arise for systems that are not symmetric self-adjoint [see Eq. (78) of the text].

More generally, the integral equation has multiple roots and turning points, and we shall demonstrate how global dispersion relations, more general than the Bohr–Sommerfeld relation, can be constructed. Besides satisfying the dispersion relation, the WKB waves have to be well behaved at infinity.

We also briefly discuss wave propagation in a n-dimensional system and derive equations for the wave trajectories. For more than one dimension, the usefulness of this technique is limited to nearly Hermitian systems where trajectories can be defined along real space coordiantes.

III. SOLUTION FOR WAVE AMPLITUDE

Let us consider an integral equation of the form

$$\sum_j\int d\mathbf{r}'G_{ij}\left(\mathbf{r}-\mathbf{r}',\epsilon\frac{\mathbf{r}+\mathbf{r}'}{2},\omega\right)\xi_j(\mathbf{r}') = 0, \tag{19}$$

where for the moment we consider the space of n-dimensions. Associated with the integral is the adjoint equation,

$$\sum_j\int d\mathbf{r}'G_{ji}\left(\mathbf{r}'-\mathbf{r},\epsilon\frac{\mathbf{r}+\mathbf{r}'}{2},\omega\right)\xi_j^+(\mathbf{r}') = 0. \tag{20}$$

These systems of equations are equivalent to the variation of the quadratic form,

$$I = \sum_{ij} \int d\mathbf{r}\, d\mathbf{r}' \zeta_i^+(\mathbf{r}') G_{ij}\left(\mathbf{r}-\mathbf{r}', \epsilon \frac{\mathbf{r}+\mathbf{r}'}{2}, \omega\right) \zeta_j(\mathbf{r}'). \tag{21}$$

The variation with respect to ζ^+ reproduces Eq. (19), and the variation with respect to ζ reproduces Eq. (20). We now derive, in a compact way, an approximate set of equations by first approximating the quadratic form and then taking the variation, a procedure used in Ref. 9.

If we substitute the Fourier representations

$$\zeta_i(\mathbf{r}) = \int d\mathbf{k} \exp(-i\mathbf{k}\cdot\mathbf{r})\phi_i(k)$$

and

$$\zeta_i^+(\mathbf{r}) = \int d\mathbf{k} \exp(-i\mathbf{k}\cdot\mathbf{r})\phi_i^+(\mathbf{k}),$$

and introduce the variables $\mathbf{R} = (\mathbf{r}+\mathbf{r}')/2$, $\mathbf{z} = \mathbf{r}-\mathbf{r}'$, we find

$$\begin{aligned} I &= \int \frac{d\mathbf{k}}{(2\pi)^n} \int \frac{d\mathbf{k}'}{(2\pi)^n} \phi_i^+(-\mathbf{k})\phi_j(\mathbf{k}') \\ &\quad \times \int d\mathbf{r}'\, d\mathbf{r} \exp\left[i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R} - i(\mathbf{k}+\mathbf{k}')\cdot\frac{\mathbf{z}}{2}\right] \\ &\quad \cdot G_{ij}(\mathbf{z},\epsilon\mathbf{R},\omega) \\ &= \int d\mathbf{R} \frac{d\mathbf{k}}{(2\pi)^n} \exp(-i\mathbf{k}\cdot\mathbf{R})\phi_i^+(-\mathbf{k}) \\ &\quad \cdot \int \frac{d\mathbf{k}'}{(2\pi)^n} \exp(i\mathbf{k}'\cdot\mathbf{R})\phi_j(\mathbf{k}')\Lambda_{ij}\left(\frac{\mathbf{k}+\mathbf{k}'}{2}, \epsilon\mathbf{R}, \omega\right), \end{aligned} \tag{22}$$

where repeated indices imply summation unless otherwise stated, and

$$\Lambda_{ij}(\mathbf{k},\epsilon\mathbf{R},\omega) = \int d\mathbf{z} \exp(-i\mathbf{k}\cdot\mathbf{z}) G_{ij}(\mathbf{z},\epsilon\mathbf{R},\omega). \tag{23}$$

If we now expand $\mathbf{k}$ and $\mathbf{k}'$ in Λ_{ij} about $\mathbf{k}(\mathbf{R})$, we find

$$\begin{aligned} I &= \int d\mathbf{R} \{ \hat{\zeta}_i^+(\epsilon\mathbf{R})\hat{\zeta}_j(\epsilon\mathbf{R})\Lambda_{ij}[\mathbf{k}(\epsilon\mathbf{R}),\epsilon\mathbf{R},\omega] \} \\ &\quad + \frac{i\epsilon}{2}\left[\left\{\frac{\partial}{\partial \epsilon R}\hat{\zeta}_i^+(\epsilon\mathbf{R})\hat{\zeta}_j(\epsilon\mathbf{R}) - \hat{\zeta}_i^+(\epsilon\mathbf{R})\frac{\partial}{\partial\epsilon\mathbf{R}}\hat{\zeta}_j(\epsilon\mathbf{R})\right\} \right. \\ &\quad \left.\left. \cdot \frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}(\mathbf{k},\epsilon\mathbf{R},\omega)\right]\right|_{\mathbf{k}=\mathbf{k}(\epsilon\mathbf{R})} \\ &\quad - \frac{\epsilon^2}{8}\left[\frac{\partial^2\hat{\zeta}_i^+(\epsilon\mathbf{R})}{\partial\epsilon\mathbf{R}\,\partial\epsilon\mathbf{R}}\hat{\zeta}_j(\epsilon\mathbf{R}) - 2\frac{\partial\hat{\zeta}_i^+}{\partial\epsilon\mathbf{R}}\frac{\partial\hat{\zeta}_j}{\partial\epsilon\mathbf{R}} + \hat{\zeta}_i^+ \right. \\ &\quad \left. \times \frac{\partial^2\hat{\zeta}_j}{\partial\epsilon\mathbf{R}\,\partial\epsilon\mathbf{R}}\right] : \left.\frac{\partial^2}{\partial\mathbf{k}\,\partial\mathbf{k}}\Lambda_{ij}(\mathbf{k},\epsilon\mathbf{R},\omega)\right|_{\mathbf{k}=\mathbf{k}(\epsilon\mathbf{R})} \\ &\quad + \mathscr{O}(\epsilon^3), \end{aligned} \tag{24}$$

where

$$\hat{\zeta}_i(\epsilon\mathbf{r}) = \zeta_i(\mathbf{r})\exp[-iS(\mathbf{r})],$$
$$\hat{\zeta}_i^+(\epsilon\mathbf{r}) = \zeta_i^+(\mathbf{r})\exp[iS(\mathbf{r})], \tag{25}$$

and

$$\frac{\partial S}{\partial\mathbf{r}} = \mathbf{k}(\epsilon\mathbf{r}).$$

We see that Eq. (24) is an explicit expansion in WKB amplitudes $\hat{\zeta}_i(\epsilon\mathbf{R})$ with a natural clustering of orders $\epsilon^n \simeq 1/(kR)^n$.

We now keep terms correct to order $n = 1$. We then find the WKB approximation of the integral equations for $\hat{\zeta}_i$ and $\hat{\zeta}_i^+$ by taking the variation of Eq. (24) with respect to $\hat{\zeta}_i^+$ and $\hat{\zeta}_i$ respectively. We obtain

$$\begin{aligned} &\left\{\Lambda_{ij}[\mathbf{k}(\epsilon\mathbf{r}),\mathbf{r},\omega]\hat{\zeta}_j(\epsilon\mathbf{r}) - i\frac{\hat{\zeta}_j(\epsilon\mathbf{r})}{2}\frac{d}{d\mathbf{r}}\cdot\frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r},\omega] \right. \\ &\quad \left. - i\frac{d\hat{\zeta}_j(\epsilon\mathbf{r})}{d\mathbf{r}}\cdot\frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r},\omega]\right\} = 0, \end{aligned} \tag{26}$$

$$\begin{aligned} &\{\Lambda_{ji}[\mathbf{k}(\epsilon\mathbf{r},\mathbf{r},\omega)]\hat{\zeta}_j^+(\epsilon\mathbf{r}) + i\frac{\hat{\zeta}_j^+(\epsilon\mathbf{r})}{2}\frac{d}{d\mathbf{r}}\cdot\frac{\partial}{\partial k}\Lambda_{ji}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r},\omega] \\ &\quad \left. + i\frac{d}{d\mathbf{r}}\zeta_j^+(\epsilon\mathbf{r})\cdot\frac{\partial}{\partial k}\Lambda_{ji}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r},\omega]\right\} = 0, \end{aligned} \tag{27}$$

where

$$\frac{d}{d\mathbf{r}}F[k(\epsilon\mathbf{r}),\epsilon\mathbf{r}] = \left[\frac{\partial}{\partial\mathbf{r}} + \left(\frac{d\mathbf{k}}{d\mathbf{r}}\right)\cdot\frac{\partial}{\partial k}\right]F[\mathbf{k}(\epsilon\mathbf{r}),\mathbf{r}].$$

To find the solution to Eq. (26), we shall assume and ultimately verify that it takes the following form:

$$\hat{\zeta}_j(\epsilon\mathbf{r}) = \Lambda^{js}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r}]\psi_s(\epsilon\mathbf{r}),$$

where s is arbitrary and fixed, and $\Lambda^{js}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r}]$ is the cofactor of the matrix

$$\Lambda_{ij}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r}],$$

i.e.,

$$\Lambda_{ij}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r}]\Lambda^{js}[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r}] = \det|\Lambda_{ij}|\delta_{is} \equiv |\Lambda|\delta_{is}, \tag{28}$$

and we choose $\mathbf{k}(\epsilon\mathbf{r})$ so that $|\Lambda[\mathbf{k}(\epsilon\mathbf{r}),\epsilon\mathbf{r}]| = 0$.

Then, using $\mathbf{y} = \epsilon\mathbf{r}$, Eq. (26) becomes,

$$\begin{aligned} &\left\{\frac{\partial}{\partial k}\Lambda_{ij}[\mathbf{k}(\mathbf{y}),\mathbf{y}]\cdot\frac{d}{d\mathbf{y}}\Lambda^{jr}[\mathbf{k}(\mathbf{y}),\mathbf{y}] + \Lambda^{jr}[\mathbf{k}(\mathbf{y}),\mathbf{y}]\right. \\ &\quad \times \frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}[\mathbf{k}(\mathbf{y}),\mathbf{y}]\cdot\frac{d}{d\mathbf{y}} + \frac{1}{2}\frac{d}{d\mathbf{y}} \Bigg\} \\ &\quad \left.\cdot\frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}[\mathbf{k}(\mathbf{y}),\mathbf{y}]\Lambda^{jr}(\mathbf{k},\mathbf{y})\right\}\psi_r(\mathbf{y}) = 0. \end{aligned} \tag{29}$$

Now multiplying by $\Lambda^{vi}[\mathbf{k}(\mathbf{y}),\mathbf{y}]$ and summing over i yields

$$\begin{aligned} &\left[\Lambda^{vi}\frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}\cdot\frac{d}{d\mathbf{y}}\Lambda^{jr} + \Lambda^{vi}\left(\frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}\right)\Lambda^{jr}\cdot\frac{d}{d\mathbf{y}} \right. \\ &\quad \left. + \frac{1}{2}\Lambda^{vi}\left(\frac{d}{d\mathbf{y}}\cdot\frac{\partial\Lambda_{ij}}{\partial\mathbf{k}}\right)\Lambda^{jr}\right]\psi_r = 0. \end{aligned} \tag{30}$$

We now use the following identities for arbitrary r:

$$\begin{aligned} &\Lambda^{vi}\frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}\cdot\frac{d}{d\mathbf{y}}\Lambda^{jr} \\ &= \frac{\partial}{\partial\mathbf{k}}|\Lambda|\cdot\frac{d}{d\mathbf{y}}\Lambda^{vr} - \frac{\partial\Lambda^{vi}}{\partial\mathbf{k}}\Lambda_{ij}\cdot\frac{d}{d\mathbf{y}}\Lambda^{jr}, \end{aligned} \tag{31}$$

$$\begin{aligned} &\Lambda^{vi}\left(\frac{\partial}{\partial\mathbf{k}}\Lambda_{ij}\right)\Lambda^{jr} \\ &= \frac{\partial|\Lambda|}{\partial\mathbf{k}}\Lambda^{vr} - \frac{\partial\Lambda^{vi}}{\partial\mathbf{k}}\Lambda_{ij}\Lambda^{jr} \\ &= \frac{\partial|\Lambda|}{\partial\mathbf{k}}\Lambda^{vr} - \frac{\partial}{\partial\mathbf{k}}\Lambda^{vr}|\Lambda| = \frac{\partial|\Lambda|}{\partial\mathbf{k}}\Lambda^{vr}, \end{aligned} \tag{32}$$

and

$$\Lambda^{vi}\left(\frac{d}{dy}\cdot\frac{\partial}{\partial \mathbf{k}}\Lambda_{ij}\right)\Lambda^{jr}$$
$$= -\Lambda^{vi}\frac{\partial}{\partial \mathbf{k}}\Lambda_{ij}\cdot\frac{d}{dy}\Lambda^{jr} + \Lambda^{vi}\frac{d}{dy}\cdot\left(\frac{\partial}{\partial \mathbf{k}}\Lambda_{ij}\Lambda^{jr}\right)$$
$$= -\frac{\partial|\Lambda|}{\partial \mathbf{k}}\cdot\frac{d}{dy}\Lambda^{vr} + \frac{\partial}{\partial \mathbf{k}}\Lambda^{vi}\Lambda_{ij}\cdot\frac{d}{dy}\Lambda^{jr} + \frac{d}{dy}\cdot\left(\Lambda^{vi}\frac{\partial\Lambda_{ij}}{\partial \mathbf{k}}\Lambda^{jr}\right) - \left(\frac{d}{dy}\Lambda^{vi}\right)\cdot\frac{\partial\Lambda_{ij}}{\partial \mathbf{k}}\Lambda^{jr}$$
$$= -\frac{\partial|\Lambda|}{\partial \mathbf{k}}\cdot\frac{\partial\Lambda^{vr}}{dy} + \frac{\partial}{\partial \mathbf{k}}\Lambda^{vi}\Lambda_{ij}\cdot\frac{d}{dy}\Lambda^{jr} + \left(\frac{d}{dy}\Lambda^{vi}\right)\Lambda_{ij}\cdot\frac{\partial}{\partial \mathbf{k}}\Lambda^{jr} + \left(\frac{d}{dy}\cdot\frac{\partial|\Lambda|}{\partial \mathbf{k}}\right)\Lambda^{vr}. \tag{33}$$

Now using Eqs. (31)–(33) in Eq. (30) gives,

$$\left(\frac{1}{2}\frac{d}{dy}\cdot\frac{\partial}{\partial \mathbf{k}}|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]| + {}^{``v"}\{\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\}^r\right)\psi_r + \frac{\partial}{\partial \mathbf{k}}|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]|\cdot\frac{d}{dy}\psi_r = 0, \tag{34}$$

where

$${}^{``v"}\{\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\}^r = \left[\frac{1}{2}\frac{\partial}{\partial \mathbf{k}}|\Lambda|\cdot\frac{d}{dy}\Lambda^{vr} - \frac{1}{2}\left(\frac{\partial}{\partial \mathbf{k}}\Lambda^{vi}\right)\Lambda_{ij}\cdot\frac{d}{dy}\Lambda^{jr} + \frac{1}{2}\left(\frac{d}{dy}\Lambda^{vi}\right)\Lambda_{ij}\cdot\frac{\partial}{\partial \mathbf{k}}\Lambda^{jr}\right]\Big/\Lambda^{vr}$$
$$= \left(\frac{1}{2}|\Lambda|_k\cdot\frac{d\Lambda^{vr}}{dy} - \frac{1}{2}\frac{\partial\Lambda^{vi}}{\partial \mathbf{k}}\Lambda_{ij}\cdot\frac{\partial\Lambda^{jr}}{\partial \mathbf{y}} + \frac{1}{2}\frac{\partial\Lambda^{vi}}{\partial \mathbf{y}}\Lambda_{ij}\cdot\frac{\partial\Lambda^{jr}}{\partial \mathbf{k}}\right)\Big/\Lambda^{vr}. \tag{35}$$

In obtaining Eq. (34) we assumed ψ_r is independent of any other index. Hence for consistency we need ${}^v\{\Lambda\}^r$ to be independent of v. To prove this, we note that

$$\frac{\Lambda^{ri}[\mathbf{k}(\mathbf{y}),\mathbf{y}]}{\Lambda^{rj}[\mathbf{k}(\mathbf{y}),\mathbf{y}]} = \frac{\Lambda^{si}[\mathbf{k}(\mathbf{y}),\mathbf{y}]}{\Lambda^{sj}[\mathbf{k}(\mathbf{y}),\mathbf{y}]}, \tag{36}$$

a result of linear algebra that follows from $|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]| = 0$. Then we can write

$${}^{``v"}\{\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\}^r = \frac{1}{2}\frac{\partial}{\partial \mathbf{k}}|\Lambda|\cdot\frac{d}{dy}\ln\Lambda^{vr} - \frac{1}{2}\left(\frac{\partial}{\partial \mathbf{k}}\Lambda^{vi}\Lambda_{ij}\cdot\frac{d}{dy}\Lambda^{jr} - \frac{d\Lambda^{vi}}{dy}\cdot\Lambda_{ij}\frac{\partial}{\partial \mathbf{k}}\Lambda^{jr}\right)\frac{1}{\Lambda^{vr}}$$
$$= \left[\Lambda^{vi}\frac{\partial}{\partial \mathbf{k}}\Lambda_{ij}\cdot\frac{d}{dy}\Lambda^{jr} - \Lambda^{vi}\frac{d\Lambda_{ij}}{dy}\cdot\frac{\partial\Lambda^{jr}}{\partial \mathbf{k}} + \frac{d|\Lambda|}{dy}\cdot\frac{\partial}{\partial \mathbf{k}}\Lambda^{vr}\right]\Big/2\Lambda^{vr}. \tag{37}$$

Since $(d/dy)|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]| = 0$, we have

$${}^{``v"}\{\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\}^r = \frac{\Lambda^{vi}}{2\Lambda^{vr}}\left(\frac{\partial}{\partial \mathbf{k}}\Lambda_{ij}\cdot\frac{d}{dy}\Lambda^{jr} - \frac{d\Lambda_{ij}}{dy}\cdot\frac{\partial\Lambda^{jr}}{\partial \mathbf{k}}\right). \tag{38}$$

Now as $\Lambda^{vi}/\Lambda^{vr}$ is independent of v, we have proven our contention, and hereafter suppress the symbol "v". Similarly, one can show

$${}^v\{\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\}^{``r"} \equiv -\frac{1}{2}\frac{\partial}{\partial \mathbf{k}}|\Lambda|\cdot\frac{d}{dy}\ln\Lambda^{vr} - \frac{1}{2\Lambda^{vr}}\left(\frac{\partial}{\partial \mathbf{k}}\Lambda^{vi}\Lambda_{ij}\cdot\frac{d}{dy}\Lambda^{jr} - \frac{d}{dy}\Lambda^{vr}\Lambda_{ij}\cdot\frac{\partial}{\partial \mathbf{k}}\Lambda^{vr}\right)$$
$$= -\frac{1}{2}\frac{\partial}{\partial \mathbf{k}}|\Lambda|\cdot\frac{d}{dy}\ln\Lambda^{vr} - \frac{1}{2\Lambda^{vr}}\left(\frac{\partial}{\partial \mathbf{k}}\Lambda^{vi}\Lambda_{ij}\cdot\frac{\partial}{\partial \mathbf{y}}\Lambda^{jr} - \frac{\partial}{\partial \mathbf{y}}\Lambda^{vi}\Lambda_{ij}\cdot\frac{\partial}{\partial \mathbf{k}}\Lambda^{jr}\right), \tag{39}$$

is independent of r, and we can suppress "r". Subtracting Eq. (39) and (35), we have

$$\{\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\}^r = {}^v\{\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\} + \frac{\partial}{\partial \mathbf{k}}|\Lambda|\cdot\frac{d}{dy}\ln\Lambda^{vr}. \tag{40}$$

Hence, Eq. (34) can be written

$$\left(\frac{1}{2}\frac{d}{dy}\cdot\frac{\partial}{\partial \mathbf{k}}|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]| + \{\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\}^r\right)\psi_r + \frac{\partial}{\partial \mathbf{k}}|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]|\cdot\frac{d}{dy}\psi_r = 0. \tag{41}$$

At this stage we limit our discussion to spatial variations in one dimension. Further discussion for the multidimensional case will be given in Appendix I. The symbol $k(x)$ refers to $k_x(x)$; k_y and k_z are fixed parameters, and y now refers to ϵx.

For one dimension, the equation of ψ_r can be integrated straightforwardly to yield

$$\psi_r = C_r \exp\left(-\int_{y_0}^{y}\frac{dy'\{\Lambda\,[k(y'),y']\}^r}{\partial|\Lambda|/\partial k}\right)\left\{\frac{\partial}{\partial k}|\Lambda\,[k(y),y]|\right\}^{-1/2}, \tag{42}$$

where C_r is a space-independent constant dependent on the reference basis r. Then $\hat{\xi}^v = \Lambda^{vr}\psi_r$ becomes

$$\hat{\xi}_v(y) = C_r\Lambda^{vr}[k(y),y]\exp\left(-\int_{y_0}^{y}\frac{dy'}{\partial|\Lambda|/\partial k}\{\Lambda\,[k(y'),y']\}^r\right)\left\{\frac{\partial}{\partial k}|\Lambda\,[k(y),y]|\right\}^{-1/2}$$
$$= C_r\Lambda^{vr}[k(y_0),y_0]\exp\left(-\int_{y_0}^{y}\frac{dy'}{\partial|\Lambda|/\partial k}\{\Lambda\,[k(y'),y']\}^r - \frac{d\Lambda^{vr}/dy'}{\Lambda^{vr}}\right)\left(\frac{\partial}{\partial k}|\Lambda\,[k(y),y]|\right)^{-1/2}$$

$$= \frac{C_r \Lambda^{\nu r}[k(y_0), y_0]}{\{(\partial/\partial k)|\Lambda[k(y),y]|\}^{1/2}} \exp\left(-\int_{y_0}^{y} \frac{dy'\,{}^{\nu}\{\Lambda[k(y'),y']\}}{(\partial/\partial k)|\Lambda[k(y'),y']|}\right), \tag{43}$$

where we have used Eq. (40). We see that the spatial variation of $\hat{\xi}_\nu(y)$ is independent of r to within a global constant, and the r dependence is only necessary to establish the correct polarization vectors at one point. We shall define as the polarization propagator,

$$\langle {}^{\nu}_{y_0}\{\Lambda[k(y),y]\}\rangle \equiv \exp\left(-\int_{y_0}^{y} \frac{dy'\{\Lambda[k(y'),y']\}}{(\partial/\partial k)|\Lambda[k(y'),y']|}\right), \tag{44}$$

as it propagates the amplitude of polarization ν from y_0 to y.

The polarization propagator simplifies considerably when the kernal is symmetric-self-adjoint, i.e.,

$$G_{ij}\left(x-x',\epsilon\frac{x+x'}{2}\right) = G_{ji}\left(x'-x,\epsilon\frac{x+x'}{2}\right)$$

and

$$G_{ij}\left(x-x',\epsilon\frac{x+x'}{2}\right) = G_{ij}\left(x'-x,\epsilon\frac{x+x'}{2}\right).$$

It then readily follows that $\Lambda_{ij}[k(y),y] = \Lambda_{ji}[k(y),y]$, from which we infer $\Lambda^{ji}[k(y),y] = \Lambda^{ij}[k(y),y]$. Then, in Eq. (39), the last two terms cancel when $r=\nu$, and we find

$$\langle {}^{\nu}_{y_0}\{\Lambda[k(y),y]\}\rangle = \left\{\frac{\Lambda^{\nu\nu}[k(y),y]}{\Lambda^{\nu\nu}[k(y_0),y_0]}\right\}^{1/2}.$$

Then, from Eq. (43),

$$\hat{\xi}_\nu(y) = \frac{C_r \Lambda^{\nu r}[k(y_0),y_0]}{[(\partial/\partial k)|\Lambda|]^{1/2}} \left\{\frac{\Lambda^{\nu\nu}[k(y),y]}{\Lambda^{\nu\nu}[k(y_0),y_0]}\right\}^{1/2}. \tag{45}$$

The general solution for the adjoint equation can be constructed in a manner similar to the methods above. We find,

$$\xi_\nu^+(y) = \frac{C_r^+ \Lambda^{r\nu}[k(y_0),y_0]}{\{(\partial/\partial k)|\Lambda[k(y),y]|\}^{1/2}} \langle {}^{\nu}_{y_0}\{\Lambda[k(y),y]\}^+\rangle \times \exp\left(\frac{-i}{\epsilon}\int_{y_0}^{y} k(y)\,dy\right), \tag{46}$$

where

$$\langle {}^{\nu}_{y_0}\{\Lambda[k(y),y]\}^+\rangle = \frac{1}{\langle\{\Lambda[k(y),y]\}^{\nu}_{y_0}\rangle}, \tag{47}$$

and from Eq. (40) we have, for arbitrary s and t,

$$\langle\{\Lambda[k(y),y]\}^{r}_{y_0}\rangle \equiv \exp\left(-\int_{y_0}^{y} \frac{dy'\{\Lambda[k(y'),y']\}^{r}}{(\partial/\partial k)|\Lambda[k(y'),y']|}\right)$$
$$= \langle {}^{s}_{y_0}\{\Lambda[k(y),y]\}\rangle \Lambda^{sr}[k(y_0),y_0]/\Lambda^{sr}[k(y),y]$$
$$= \langle\{\Lambda[k(y),y]\}^{t}_{y_0}\rangle \frac{\Lambda^{sr}[k(y_0),y_0]\Lambda^{st}[k(y),y]}{\Lambda^{sr}[k(y),y]\Lambda^{st}[k(y_0),y_0]}. \tag{48}$$

We now define action as

$$A \equiv \hat{\xi}_i^+(y)\frac{\partial}{\partial\omega}\Lambda_{ij}[k(y),y,\omega]\hat{\xi}_j(y).$$

Using Eqs. (46), (47), (42), and (43) gives, for arbitrary r,

$$A = \text{const}\,\Lambda^{ri}[k(y_0),y_0](\partial/\partial\omega)\Lambda_{ij}[k(y),y,\omega] \times \Lambda^{jr}[k(y),y]\langle\{\Lambda[k(y),y]\}^{r}_{y_0}\rangle \times (\langle\{\Lambda[k(y),y]\}^{i}_{y_0}\rangle(\partial/\partial k)|\Lambda[k(y),y]|)^{-1}.$$

Substituting Eq. (48) for the denominator then yields

$$A = \text{const}\,\frac{(\partial/\partial\omega)|\Lambda[k(y),y]|}{(\partial/\partial k)|\Lambda[k(y),y]|} \times \Lambda^{rr}[k(y_0),y_0] = \frac{\text{const}'}{v_g}, \tag{49}$$

where

$$v_g = -\frac{(\partial/\partial k)|\Lambda[k(y),y,\omega]|}{(\partial/\partial\omega)|\Lambda[k(y),y,\omega]|},$$

is the complex group velocity. Hence the WKB solution satisfies an action flow conservation, i.e., $Av_g = \text{const}$.

Our solutions [Eqs. (43) and (46)] diverge where the group velocity, $\propto |\Lambda|_k$, vanishes. We shall define as the turning point $x = x_T$, where $|\Lambda[k(x),x]| = |\Lambda[k(x),x]|_k = 0$. It is also the condition for the coalescence of two k roots, as in the vicinity of $x = x_T$, $k = k(x_T) = k_T$, we have

$$|\Lambda| = \tfrac{1}{2}\delta k^2|\Lambda|_{kkT} + \frac{\delta k^3}{6}|\Lambda|_{kkkT} + \epsilon\delta x|\Lambda|_{xT} + \epsilon\delta x\delta k\,|\Lambda|_{kxT} + \cdots, \tag{50}$$

where the T refers to evaluation at $x = x_T$ and $k = k_T$, and $\delta k = k(x) - k_T$ and $\delta x = x - x_T$. Solving Eq. (50) yields

$$\delta k = \pm\left(\frac{-2\epsilon\delta x|\Lambda|_{xT}}{|\Lambda|_{kkT}}\right)^{1/2} \times \frac{-\epsilon\delta x|\Lambda|_{xT}}{|\Lambda|_{kkT}} \times \left(\frac{|\Lambda|_{xkT}}{|\Lambda|_{xT}} - \frac{1}{3}\frac{|\Lambda|_{kkkT}}{|\Lambda|_{kkT}}\right) + \cdots, \tag{51}$$

which demonstrates the two solutions for k_0 that coalesce at x_T.

IV. *k*-SPACE DUAL AND SOLUTION NEAR TURNING POINT

To solve the equation near a turning point it is first convenient to consider the integral equation in the dual k-space. A WKB solution in the dual space can be found and the Fourier transform of this solution gives an x-space representation that is an accurate approximation in the vicinity of the x-space turning point. This x-space representation can then be used to obtain connection rules of WKB modes in x-space.

By using the variational method we readily obtain the equations for the dual representation in k-space. From Eq. (22) we have

$$I = \int \frac{d\mathbf{k}}{(2\pi)^n}\frac{d\mathbf{k}'}{(2\pi)^n} d\mathbf{R}\, dz'\phi_i^+(-\mathbf{k})\phi_j(\mathbf{k}') \cdot \exp\left[i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R} - i(\mathbf{k}+\mathbf{k}')\frac{\mathbf{z}}{2}\right]G_{ij}(\mathbf{z},\epsilon\mathbf{R},\omega)$$
$$= \int d\mathbf{k}'\, d\mathbf{k} Q_{ij}\left(\frac{\mathbf{k}+\mathbf{k}'}{2}, \frac{\mathbf{k}'-\mathbf{k}}{\epsilon}, \omega\right)\phi_i^+(-\mathbf{k})\phi_j(\mathbf{k}'), \tag{52}$$

where

$$Q_{ij}(\mathbf{s},\mathbf{t},\omega) = \int \frac{d\mathbf{R}^3 d^3\mathbf{z}}{(2\pi)^{2n}} \exp(-i\mathbf{s}\cdot\mathbf{z} - i\mathbf{t}\cdot\epsilon\mathbf{R})G_{ij}(\mathbf{x},\epsilon\mathbf{R},\omega).$$

Equation (52) is of the same structure as the r-space quadratic form given by Eq. (21). Now proceeding along exactly analogous lines as in the previous section, the variational method leads us to consider

$$\phi_j(\mathbf{k}) = \hat{\phi}_j(\mathbf{k}) \exp\left[\frac{-i}{\epsilon}\int_{k_0}^{k} \epsilon\mathbf{r}(\mathbf{k})\cdot d\mathbf{k}\right],$$

and we find $\hat{\phi}_j(\mathbf{k})$ satisfies the equation

$$\Lambda_{ij}[\mathbf{k},\epsilon\mathbf{r}(\mathbf{k}),\omega]\hat{\phi}_j(\mathbf{k}) + i\epsilon\left\{\frac{\phi_j(\mathbf{k})}{2}\frac{d}{d\mathbf{k}}\cdot\frac{\partial}{\epsilon\partial\mathbf{r}}\Lambda_{ij}[\mathbf{k},\epsilon\mathbf{r}(\mathbf{k}),\omega] + \frac{d\hat{\phi}_j(\mathbf{k})}{d\mathbf{k}}\cdot\frac{\partial}{\epsilon\partial\mathbf{r}}\Lambda_j[\mathbf{k},\epsilon\mathbf{r}(\mathbf{k}),\omega]\right\} = 0. \tag{53}$$

If we choose $\mathbf{r}(\mathbf{k})$ such that $|\Lambda[\mathbf{k},\epsilon\mathbf{r}(\mathbf{k}),\omega]| = 0$, then in one dimension the solution to Eq. (53) is essentially the same as the solution to Eq. (26) but with $\mathbf{r}$ and $\mathbf{k}$ interchanged. Hence the solution in one dimension is given by

$$\phi_v(k) = C_r\frac{\Lambda^{vr}[k_0, y(k_0)]}{|\Lambda[k,y(k)]|_y^{1/2}} \times\exp\left(-\int_{k_0}^{k} dk'\frac{{}^v\{\Lambda[y(k'),k']\}}{2|\Lambda|_y}\right) \times\exp\left(-i\int_{k_0}^{k}\frac{dk'}{\epsilon}y(k')\right). \tag{54}$$

Note that

$$\frac{{}^v\{\Lambda[y(k),k)]\}}{|\Lambda|_y} = -\frac{1}{2}\frac{d}{dk}\Lambda^{vr} - [(\partial/\partial y)\Lambda^{vi}\Lambda_{ij}(\partial/\partial k)\Lambda^{jr} - (\partial/\partial k)\Lambda^{vi}\Lambda_{ij}(\partial/\partial y)\Lambda^{jr}]\{2\Lambda^{vr}|\Lambda|_y\}^{-1}$$
$$= \frac{dy}{dk}\left[-\frac{1}{2}\frac{d\Lambda^{vr}}{dy(k)} - [(\partial\Lambda^{vi}/\partial k)\Lambda_{ij}(\partial/\partial y)\Lambda^{jr} - (\partial\Lambda^{vi}/\partial y)\Lambda_{ij}(\partial/\partial k)\Lambda^{jr}]\{|\Lambda|_k\}^{-1}\right]\{\Lambda^{vr}\}^{-1}$$
$$= \frac{dy}{dk}\frac{{}^v\{\Lambda[k,y(k)]\}}{|\Lambda|_k}. \tag{55}$$

Hence, Eq. (54) becomes

$$\phi_v(k) = \frac{C_r\Lambda^{vr}[k_0,y(k_0)]}{|\Lambda[k,y(k)]|_y^{1/2}}\langle {}^v_{y(k_0)}\{\Lambda[k,y(k)]\}\rangle \times\exp\left(-\frac{i}{\epsilon}\int_{k_0}^{k} dk\, y(k)\right). \tag{56}$$

To obtain the solution in the vicinity of y_T we transform Eq. (56) back to x-space (for convenience we choose $k_0 = k_T$):

$$\zeta_v(x) = \frac{1}{2\pi}\int dk\phi_v(k)\exp(ikx)$$
$$= C_r\Lambda^{vr}(k_T, y_T) \times\int_{-\infty}^{\infty} dk\exp\left\{i\frac{1}{\epsilon}\int_{k_T}^{k} dk'[y - y(k)] + ik_Tx\right\} \times\frac{\langle {}^v_{y_T}\{\Lambda[k,y(k)]\}\rangle}{|\Lambda[k,y(k)]|_y^{1/2}}. \tag{57}$$

For $y - y_T$ sufficiently large, Eq. (57) can be evaluated by the method of stationary phase. The phase function, $\int_{k_T}^{k}[y(k) - y]\, dk$ can be expressed as

$$\int_{k_t}^{k(x)} dk'[x - x(k')]$$
$$= \int_{x_T}^{x} dx\frac{dk'}{dx}[x - x(k')]$$
$$= k_T(x_T - x) + \int_{x_T}^{x} dx'k(\epsilon x'). \tag{58}$$

We then find

$$\zeta_v(x) \simeq \frac{C_r\Lambda^{vr}(k_T, y_T)\exp(ik_Tx_T)}{(2\pi i[dk(y)/dy]|\Lambda(k,y)|_y)^{1/2}} \times\exp\left[i\int_{x_T}^{x}k(\epsilon x')dx'\right]\langle {}^v_{y_T}\{\Lambda[k(x),x]\}\rangle$$
$$= \frac{C_r'\Lambda^{vr}(k_T, y_T)}{\{(\partial/\partial k)|\Lambda[k(y),y)]|\}^{1/2}} \times\exp\left[+i\int_{x_T}^{x}k(\epsilon x')\,dx'\right]\langle {}^v_{y_T}\{\Lambda[k(y),y]\}\rangle. \tag{59}$$

This is just the WKB solution in x-space. In the vicinity of $x = x_T$ the stationary phase aproximation cannot be applied but the integral representation given by Eq. (57) is still accurate. Note that the k-space solution breaks down where $|\Lambda[k_T, y_T(k)]| = |\Lambda[k_T, y_T(k)]|_y = 0$, and the Fourier transform of the x-space solution is an accurate representation of the dual solution in the vicinity of $k_T, y_T(k)$.

We now discuss the important point that there are two solutions to the stationary phase problem corresponding to the two $k(x)$ solutions that merge at $k(x) = k_T$ when $x = x_T$. To relate the two amplitudes somewhat away from $x - x_T$, it suffices to calculate $x(k)$ as

$$\delta x \equiv x(k) - x_T = \frac{-|\Lambda|_{kkT}\delta k^2}{2\epsilon|\Lambda|_{xT}} + \mathscr{O}\left(\frac{\delta k^3}{\epsilon}\right). \tag{60}$$

The phase $\int dkx(k) = -|\Lambda|_{kk}\delta k^3/6\epsilon|\Lambda|_x$ is greater than unity for $\delta k > \epsilon^{1/3}$, while the correction to the phase $1/\epsilon\int\delta k^3\, d\delta k \simeq \mathscr{O}(\delta k^4/\epsilon)$ is less than unity for $\delta k < \epsilon^{1/4}$. Hence, in the regime $\epsilon^{1/3} < \delta k < \epsilon^{1/4}$, the stationary phase and the simple expression for the phase is accurate, then $\xi_v(x)$ can be written, to within a global constant, as

$$\xi_v(x) = \frac{\Lambda^{vr}(y_T)}{(2\pi)} \exp(ik_T\delta x) \int \frac{dk}{|\Lambda(k_T, y_T|_y^{1/2}} \times \exp\left[ik\delta x + \frac{i(k-k_T)^3}{6\epsilon}\frac{|\Lambda|_{kk}}{|\Lambda|_y}\right] = \frac{\Lambda^{vr}(y_T)}{2\pi|\Lambda|_y^{1/2}} \exp(ik_T\delta x)\left(\frac{2\epsilon|\Lambda|_y}{|\Lambda|_{kk}}\right)^{1/3} \times \int dz \exp\left[iz\left(\frac{2\epsilon|\Lambda|_y}{|\Lambda|_{kk}}\right)^{1/3}\delta x + \frac{iz^3}{/3}\right]. \quad (61)$$

The integral in Eq. (61) is an Airy function. In the subsequent section we will be interested in the case of connecting the solution in the region $x - x_T \equiv r\exp(i\theta)$, where $\xi_v(x)\exp(-ik_T\delta x)$ is exponentially small, to the solution in the region $-r\exp(i\theta)$. Let

$$se^{i\psi} = \left(\frac{2\epsilon|\Lambda|_y}{|\Lambda|_{kk}}\right)^{1/3}.$$

The solution that is exponentially small for

$$sr \equiv \left|\delta x\left(\frac{2\epsilon|\Lambda|_y}{|\Lambda|_{kk}}\right)\right|^{1/3} \gg 1,$$

is

$$\xi_v(x) = \frac{\Lambda^{vr}(y_T)}{|\Lambda|_y^{1/2}} \exp(ik_T\delta x)\left(\frac{2\epsilon|\Lambda|_y}{|\Lambda|_{kk}}\right)^{1/3} \cdot \mathrm{Ai}\left[\delta x\left(\frac{2\epsilon|\Lambda|_y}{|\Lambda|_{kk}}\right)^{1/3}\right] \xrightarrow[sr\gg1]{} \frac{\epsilon^{1/2}\Lambda^{vr}(y_T)\exp(ik_T\delta x)}{(2\pi|\Lambda|_{kkT})^{1/2}s^{3/4}r^{1/4}\exp(i\theta/4 + 3i\psi/4)} \times \exp\left\{-\tfrac{2}{3}(sr)^{3/2}\exp\left[\frac{3i}{2}(\theta+\psi)\right]\right\}, \quad (62)$$

provided $-2\pi/3 < \theta + \psi < 2\pi/3$ (the solution is exponentially small where $|\theta + \psi| < \pi/3$).

In the region $2\pi/3 < |\theta + \psi| < 4\pi/3$ we use the asymptotic solution to the Airy function to find that $\xi_v^{\pm}(x)$ ($\pm$ refers to the sign chosen for $\theta + \psi$) is given by,

$$\xi_v^{\pm}(x) \xrightarrow[sr\gg1]{} \frac{\epsilon^{1/2}\Lambda^{vr}(y_T)\exp(ik_T\delta x)}{(2\pi|\Lambda|_{kk})^{1/2}s^{3/4}r^{1/4}\exp(i\theta/4 + 3i\psi/4)} \cdot\left(\exp\left\{-\tfrac{2}{3}(sr)^{3/2}\exp\left[\frac{3i}{2}(\theta+\psi)\right]\right\} \pm i\exp\left\{\tfrac{2}{3}(sr)^{3/2}\exp\left[\frac{3i}{2}(\theta+\psi)\right]\right\}\right). \quad (63)$$

Examination of Eq. (63) shows that $\xi_v^+(x) = \xi_v^-(x)$. However, there appears to be a formal difference in the two waves depending on whether one follows the wave in the counterclockwise (+ solution) or clockwise (− solution) direction. One wave is the analytic continuation of Eq. (34) and the other wave is induced with a relative phase shift $\pm\pi/2\exp(i\pi/2) = \pm i$. In the next section we describe a more heuristic approach due to Furry, that allows one to readily find the relation of two WKB waves in a manner consistent with Eq. (63).

V. THE FURRY METHOD

An equivalent and convenient way to treat the turning point problem is to determine the Stokes multipliers of the sub-dominant solutions along a Stokes line. This terminology will now be explained and the Furry method will be used to obtain the Stokes multipliers. We have established that an approximate WKB solution for mode $k(z)$ (z is the complex variable and we now supress the ϵ-dependence), is of the form

$$\xi_v = \frac{\Lambda_T^{vr}\langle{}^{v}_{z_t}\{\Lambda[k(z),z]\}\rangle}{(|\Lambda[k(z),z]|_k)^{1/2}} \exp\left[i\int_{z_T}^{z} k(z')\,dz'\right]. \quad (64)$$

Near the turning point we can set $\langle{}^{v}_{z_t}\{\Lambda[k(z),z]\}\rangle = \langle{}^{v}_{z_t}\{\Lambda[k(z_T),z_T]\}\rangle = 1$, and the phase function is approximated as,

$$\int_{z_t}^{z} k(z')\,dz' = k_T(z-z_T) \pm \tfrac{2}{3}(z-z_T)^{3/2}(-2|\Lambda|_z/|\Lambda|_{kk})^{1/2}, \quad (65)$$

where the Λ's are evaluated at $z = z_T$.

Let us denote $k_1(z)$ as the solution with the plus sign and $k_2(z)$ as the solution with the minus sign. We define a function

$$G(z,z_T) = i\int_{z_t}^{z}[k_1(z') - k_2(z')]\,dz' \xrightarrow[z\to z_t]{} -\tfrac{4}{3}(z-z_T)^{3/2}(2|\Lambda|_z/|\Lambda|_{kk})^{1/2}. \quad (66)$$

Using the following definitions,

$$se^{i\psi} = (2|\Lambda|_z/|\Lambda|_{kk})^{1/3}, \quad z - z_T = re^{i\theta},$$

we rewrite $G(z,z_T)$ in the neighborhood of $z = z_T$ as,

$$G(z,z_T) = -\tfrac{4}{3}(sr)^{3/2}e^{i3(\theta+\psi)/2}.$$

Then ξ_v is of the form

$$\xi_v = \frac{A}{r^{1/4}\exp[i(\theta+\psi)/4]} \times \exp[-\tfrac{2}{3}(sr)^{3/2}e^{i3(\theta+\psi)/2} + ik_T(z-z_T)] \quad (67)$$

and

$$A = \frac{\Lambda_T^{vr}}{[\exp(i\psi/2 + i\pi/4)s^{3/4}|\Lambda(k_T,z_T)|_{kk}^{1/2}]}.$$

The k_1 mode corresponds to the function $\xi_1(r,\theta)$ with $0 \leqslant \theta + \psi < 2\pi$ and the k_2 mode corresponds to $\xi_2(r,\theta) = \xi_1(r,\theta + 2\pi)$ and we have $\xi_1(r,\theta + 4\pi) = -\xi_1(r,\theta)$. We define the Stokes lines (anti-Stokes lines) of the turning point z_T, as those curves in the complex z-plane for which $G(z,z_T)$ is real (imaginary). Near $z = z_T$ the Stokes lines form equally oriented stars with three rays in each of the two Riemannian planes as G is real for

$$\tfrac{3}{2}(\theta + \psi) = n\pi, \quad \text{or} \quad \theta = \theta_n = -\psi + n2\pi/3,$$

$$n = 0,1,\ldots,5, \quad n \equiv n \bmod 6.$$

If we replace n by $n + \frac{1}{2}$, then G is purely imaginary on the rays of the new stars forming the anti-Stokes lines emerging from z_T. The topology of these stars is shown on Fig. 1.

G is negative for $n = 0,2,4$ and positive for $n = 1,3,5$.

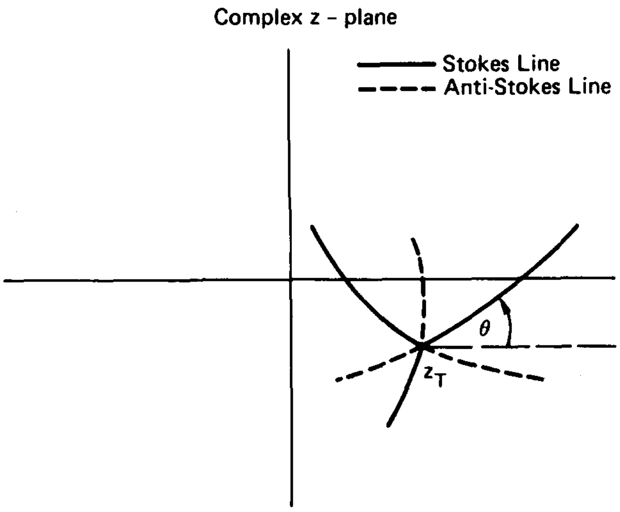


FIG. 1. Stokes and anti-stokes lines emanating from a turning point z_T in the complex z-plane.

Thus $\xi(r,\theta)$ is exponentially small compared with $\xi(r,\theta+2\pi)$ for the first set of θ_n's for r sufficiently large and is therefore called subdominant with respect to $\xi(r,\theta+2\pi)$. For the second set of θ_n's, $\xi(r,\theta)$ is large compared with $\xi(r,\theta+2\pi)$ and is therefore called dominant with respect to $\xi(r,\theta+2\pi)$.

Suppose now we propagate $\xi(r,\theta)$ from the Stokes line $n=0$ on a half circle in the clockwise direction. The result will then be different from the one obtained when propagating on a half circle in the counterclockwise direction. However close to a turning point, the governing integral equation can only lead to a single-valued solution. The failure to obtain single-valuedness is due to the fact that the analytical continuation of the asymptotic solution over a finite distance can deviate appreciably from the analytical continuation of the exact solution. We can resolve the discrepany if we obseve that, when we arrive at the $n=1$ line [where $\xi(r,\theta)$ is dominant over $\xi(r,\theta+2\pi)$], only a slight change of $\xi(r,\theta)$ is obtained if we add to the solution $\xi(r,\theta+2\pi)=-\xi(r,\theta-2\pi)$ multiplied by a factor of order 1. Thus we can assume that the asymptotic function in the region $\theta_1<\theta<\theta_2$ that is the continuation from the solution $\xi(r,\theta)$ is the region $\theta_0<\theta<\theta_1$ is

$$\xi^+=\xi(r,\theta)-\delta_+\xi(r,\theta+2\pi)$$
$$=\xi(r,\theta)+\delta_+\xi(r,\theta-2\pi). \tag{68}$$

Going the other way around and arriving at $n=-1$ (which is equivalent to $n=5$), we can also add $\xi(r,\theta+2\pi)$ multiplied by a factor δ_-. Then we obtain

$$\xi^-=\xi(r,\theta)+\delta_-\xi(r,\theta+2\pi),\quad \theta_{-2}<\theta<\theta_{-1}. \tag{69}$$

Uniqueness of ξ requires that

$$\xi^+(\theta=\theta_0+\pi+\delta\theta)=\xi^-(\theta=\theta_0-\pi+\delta\theta).$$

From this we obtain

$$\xi(r,\theta_0+\pi+\delta\theta)+\delta_+\xi(r,\theta_0-\pi+\delta\theta)$$
$$=\xi(r,\theta_0-\pi+\delta\theta)+\delta_-\xi(r,\theta_0+\pi+\delta\theta),$$

and therefore

$$\delta_+=1,\quad \delta_-=1. \tag{70}$$

Equation (70) can be interpreted as follows. If we follows the propagation of a WKB wave, $\xi(r,\theta)$, from region A to region B, where A and B are separated by a Stokes line emanating from a turning point z_T, and if $\xi(r,\theta)$ is dominant on that Stokes line, then in region B we need to add to $\xi(r,\theta)$ the subdominant contribution which is the analytic continuation of $\xi(r,\theta)$ obtained by encircling z_T in the counter direction to the original sense of propagation.

If we write

$$\xi_1(r,\theta)\equiv\xi_2(r,\theta),$$
$$\xi_2(r,\theta)=i\xi_2(r,\theta+2\pi)=-i\xi_2(r,\theta-2\pi), \tag{71}$$

we obtain the relations in the more usual form,

$$\xi^+=\xi_1+i\xi_2,\quad \xi^-=\xi_1-i\xi_2,\quad \theta_1<|\theta|<\theta_2. \tag{72}$$

Note that this form is convenient since in the ξ^+ form we have $\lim_{r\to 0}\xi_2=\xi_1$ and similarly in ξ^-, $\lim_{r\to 0}\xi_2=\xi_1$.

VI. CONSTRUCTION OF GLOBAL DISPERSION RELATIONS

We will now consider how global dispersion relations can be constructed in a somewhat intuitive manner. We assume that our boundary conditions are that the waves have to vanish for $|\mathrm{Re}z|=|x|\to-\infty$. Hence, the allowable WKB waves are the ones with $\mathrm{Im}k>0$ for $x\to+\infty$ and $\mathrm{Im}k<0$ for $x\to-\infty$. If there exists a solution at all, then there exists at leat one $k(z)$ having $\mathrm{Im}k<0$ for $x\to-\infty$ and a path from there is the complex x-plane to $x\to+\infty$ such that $\mathrm{Im}k>0$ at this boundary. Propagation of the mode from $x=-\infty$ along this path leads to additional waves being picked up whenever it crosses a Stokes line on which either one of the original or new waves are dominant. The requirement that waves with $\mathrm{Im}k<0$ cancel when going to $x=+\infty$ forms the global dispersion relation.

A. Two-turning-point problem

To be more quantitative let us consider the two-turning-point eigenvalue problem. This is the most common WKB problem. Suppose we have wavenumbers k_1 and k_2 where $\mathrm{Im}k_1<0$ as $|x|\to\infty$ and $\mathrm{Im}k_2>0$ as $|x|\to\infty$, and

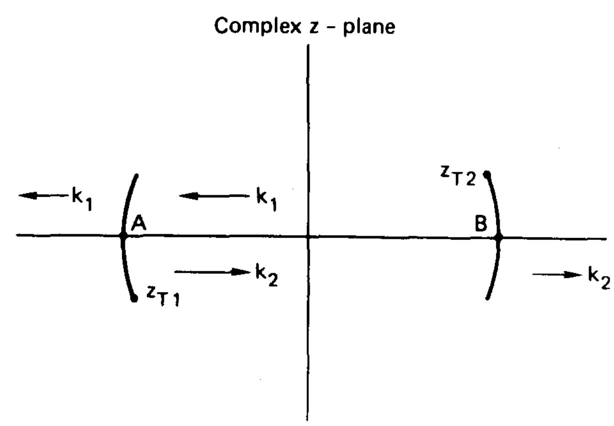


FIG. 2. Schematic diagram of Stokes lines that induce additional waves in a two wave, two-turning point problem.

there are two turning points z_{T1} and z_{T2} where k_1 and k_2 merge, say in the configuration shown in Fig. 2. To the left of point A the wave is of the form,

$$\phi = C \exp\left(i \int_0^z q_1(z')\,dz'\right) \equiv C\,\langle q_1|0,z\rangle, \tag{73}$$

where

$$q_1 = k_1 - \frac{1}{2}\frac{d}{dz}\ln|\Lambda\ |_k + i\frac{{}^{\nu}\{\Lambda\ [k(z),z]\}}{(\partial/\partial k)|\Lambda\ [k(z),z]\,|},$$

as only this mode is allowable as $x \to -\infty$. Note the relation $\langle q_1|z_{T1},z\rangle = \langle q_1|z_{T1},z_{T2}\rangle\,\langle q_1|z_{T2},z\rangle$. If the mode corresponding to q_1 is dominant with respect to q_2 along the Stokes line (z_{T1},A), then to the right of point A the wave has the form,

$$\begin{aligned}\phi &= \langle q_1|0,z\rangle - iC\,\langle q_1|0,z_{T1}\rangle\,\langle q_2|z_{T1},z\rangle,\\ &= C\,\langle q_1|0,z\rangle - iC\,\langle q_1|0,z_{T1}\rangle\,\langle q_2|z_{T1},0\rangle\,\langle q_2|0,z\rangle.\end{aligned} \tag{74}$$

If as we pass point B, we assume that the mode with wavenumber q_2 is dominant with respect to q_1 along the Stokes line (z_{T2},B), the eigenfunction then has the form,

$$\begin{aligned}\phi &= C\,\langle q_1|0,z\rangle - iC\,\langle q_1|0,z_{T1}\rangle\,\langle q_2|z_{T1},0\rangle\,\langle q_2|0,z\rangle\\ &\quad + C\,\langle q_1|0,z_{T1}\rangle\,\langle q_2|z_{T1},0\rangle\,\langle q_2|0,z_{T2}\rangle\,\langle q_1|z_{T2},z\rangle\\ &= C\,\langle q_1|0,z\rangle\,[1 + \langle q_1|0,z_{T1}\rangle\,\langle q_2|z_{T1},0\rangle\,\langle q_2|0,z_{T2}\rangle\\ &\quad \times \langle q_1|z_{T2},0\rangle - iC\,\langle q_1|0,z_{T1}\rangle\,\langle q_2|z_{T1},z\rangle\,].\end{aligned} \tag{75}$$

As only the q_2 mode is allowed as $x \to \infty$, we then demand that the two waves constituting q_1 cancel each other. This leads to the dispersion relation,

$$-1 = \exp\left\{i \int_{z_{t1}}^{z_{t2}} [q_2(z',\omega) - q_1(z',\omega)]\,dz'\right\}. \tag{76}$$

If we use Eq. (35) for ${}^{\nu}\{\Lambda\ \}$ we find,

$$\begin{aligned}&\int_{z_{t1}}^{z_{t2}} (q_2 - q_1)\,dz'\\ &= \int_{z_{t1}}^{z_{t2}} dz'\Big(k_2(z') - k_1(z') + \delta k_2^{\nu\mu}(z') - \delta k_1^{\nu\mu}(z')\\ &\quad - \frac{i}{2}\frac{d}{dz'}\ln\left\{\frac{\Lambda^{\nu\mu}[k_2(z'),z']\,|\Lambda\ [k_2(z'),z']\,|}{\Lambda^{\nu\mu}[k_1(z'),z']\,|\Lambda\ [k_1(z'),z']\,|}\right\}\Big),\end{aligned} \tag{77}$$

where

$$\begin{aligned}\delta k^{\nu\mu} = i[&(\partial/\partial z)\Lambda^{\nu i}\Lambda_{ij}(\partial/\partial k)\Lambda^{j\mu} - (\partial/\partial k)\Lambda^{\nu i}\\ &\times \Lambda_{ij}(\partial/\partial z)\Lambda^{j\mu}]\{2\Lambda^{\mu\nu}(\partial/\partial k)|\Lambda\ |\}^{-1}.\end{aligned} \tag{78}$$

The total logarithmic derivative term vanishes in Eq. (77) as the endpoint value is zero. Further, it can be shown that

$$\begin{aligned}I &= \int_{z_{t1}}^{z_{t2}} dz'\,[(\delta k_2^{\nu\mu} - \delta k_1^{\nu\mu}) - (\delta k_2^{sr} - \delta k_1^{sr})],\\ &= \int_{z_{t1}}^{z_{t2}} dz'\,\frac{d}{dz'}\ln\left(\frac{\Lambda_2^{\mu s}\Lambda_1^{\nu r}}{\Lambda_1^{\mu s}\Lambda_2^{\nu r}}\right) = 0.\end{aligned}$$

Thus the $\delta k^{\nu\mu}$ contribution to Eq. (77) is independent of ν and μ and the superscripts can therefore be neglected.

In general we have found a correction to the general WKB dispersion relation which is now of the form

$$1 + \exp\left[\int_{z_{T1}}^{z_{t2}} (k_2 + \delta k_2 - k_1 - \delta k_1)\,dz'\right] = 0 \tag{79}$$

or

$$\int_{z_{t1}}^{z_{t2}} (k_2 + \delta k_2 - k_1 - \delta k_1)\,dz' = (2n+1)\pi, \quad n = 0,1,\cdots. \tag{80}$$

In Appendix B the reason for the δk correction is discussed. It arises because given two equivalent systems whose governing kernals, $G\,[x - x',(x + x')/2]$ and $\tilde{G}\,[x - x',(x + x')/2]$, which are related by a similarity transformation $\hat{G} = U(x)GU^{-1}(x')$, then the lowest order solution for $k(x)$ will differ for the two systems. The WKB correction derived here is the compensation for the discrepancy. We note that $\delta k^{\nu\nu} = 0$ for a symmetric self-adjoint system, and hence there is no additional WKB correction for this case. From Appendix B it readily follows that the corrections vanish for all kernals that can be obtained from a spatially constant similarity transformation of a symmetric self-adjoint matrix.

If we delete from k all logarithmic derivatives terms that ultimately vanish, it is then convenient to write

$$\int_{z_{t1}}^{z_{t2}} dz'(k_{21} + \delta k_2 - k_1 - \delta k_1) = \oint_{(z_{t1},z_{t2})} dz'(k + \delta k),$$

where the right-hand side is a loop integral (or action integral) that encloses the turning points z_{T1} and z_{T2}. The reduction in the final algorithm from $\oint q\,dz$ to $\oint(k + \delta k)\,dz$ is quite general. For simplicity in subsequent examples we shall not differentiate between these forms, but simply write $\exp(i \int^z k\,dz')$, with the implication that the final form is $\exp[i \oint (k + \delta k)\,dz']$.

Equation (80) leads to the most common form of the WKB dispersion relation when $k_2 = -k_1$ and $\delta k = 0$ (as is the case for a second order differential equation).

In general, any number of modes can be involved in the global dispersion relation. For definiteness we shall consider some possible examples where three or four modes are involved in determining a global dispersion relation.

B. Three-mode problem

Suppose we have three modes with wave numbers k_1, k_2, and k_3 with $\mathrm{Im}k_1 < 0$, $\mathrm{Im}k_2 > 0$, and $\mathrm{Im}k_3 > 0$ as $|x| \to \infty$. Suppose z_{13} and z_{13}^* are turning points of the k_1 and k_3 modes and z_{12} and z_{12}^* the turning points of the k_1 and k_2 modes. Further we suppose that the configuration of these turning points and the relevant Stokes lines are as in Fig. 3. We shall assume mode 1 is dominant with respect to mode 3 along (z_{13}, A), mode 1 is dominant with respect to mode 2 along (z_{12},B), mode 2 is dominant with respect to mode 1 along (z_{12}^*,B^*), and mode 3 is dominant with respect to mode 1 along (z_{13}^*,A^*). Then to the left of A only mode 1 is allowable, and the wave has the form:

$$\phi = \langle k_1|0,z\rangle. \tag{81}$$

Between A and B mode 3 is present, and the waveform is

$$\phi = \langle k_1|0,z\rangle + i\langle k_1|0,z_{13}\rangle\,\langle k_3|z_{13},z\rangle. \tag{82}$$

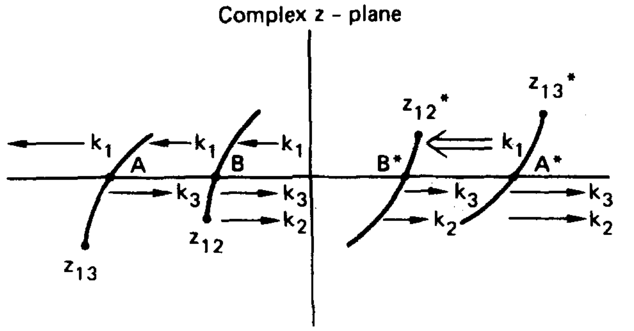


FIG. 3. Schematic diagram of Stokes lines that induce additional waves in a three-wave, four-turning-point problem. Double arrow indicates superposition of wavelets.

Between B and B^* mode 2 is present and the waveform is

$$\phi = \langle k_1|0,z\rangle + i\langle k_1|0,z_{13}\rangle \langle k_3|z_{13},z\rangle + i\langle k_1|0,z_{12}\rangle \langle k_2|z_{12},z\rangle. \quad (83)$$

Between B^* and A^* the k_1 mode obtains an additional component induced from the k_2 wave. The k_1 component has the form

$$\langle k_1|0,z\rangle [1 + \langle k_1|0,z_{12}\rangle \langle k_2|z_{12},z_{12}^*\rangle \langle k_1|z_{12}^*,0\rangle]. \quad (84)$$

Finally to the right of A^* the k_1 wave obtains an additional component induced from the k_3 wave. The total k_2 component is then

$$\langle k_1|0,z\rangle [1 + \langle k_1|0,z_{12}\rangle \langle k_2|z_{12},z_{12}^*\rangle \langle k_1|z_{12}^*,0\rangle + \langle k_1|0,z_{13}\rangle \langle k_3|z_{13},z_{13}^*\rangle \langle k_1|z_{13}^*,0\rangle]. \quad (85)$$

However, the k_1 components has to vanish to the right of A^*. Hence the overall dispersion relation is

$$1 + \exp\left\{i\int_{z_{12}}^{z_{12}^*} [k_2(z',\omega) - k_1(z',\omega)]\,dz'\right\} + \exp\left[i\int_{z_{13}}^{z_{13}^*} [k_3(z',\omega) - k_1(z',\omega)]\,dz'\right] \equiv 1 + \exp\left[i\oint_{(z_{12},z_{12}^*)} k(z')\,dz'\right] + \exp\left[i\oint_{(z_{13},z_{13}^*)} k(z')\,dz'\right] = 0. \quad (86)$$

Dispersion relations of this form have been derived in several previous works.[4,10]

Frequently one of the exponentials is exponentially small, and the three-mode problem reduces to the two-mode problem. However, as some external parameter in a problem varies, it is possible that the exponentially small component gets larger and is eventually significant.

C. Four-wave problem

We now consider a four-wave problem for the case where symmetry exists about $z = 0$. For such a case it can be shown that eigenfunctions are either even or odd about the midplane. This allows some simplification in solving the global dispersion relation. From symmetry it follows that if $k(z)$ is a solution of the local dispersion relation, then $-k(z)$ is also a solution. We shall consider waves $k_1(z)$, $-k_1(z)$, $k_2(z)$, and $-k_2(z)$, where $\mathrm{Im}k_1(z) > 0$ and $\mathrm{Im}k_2(z) < 0$. Let $z = \pm z_{T1}$ be the turning point for $\pm k_1(z)$ and $\pm z_{T12}$ be the turning points for k_1, k_2 modes (by symmetry $\pm z_{T12}$ is also the turning point for $-k_1$, $-k_2$ modes). The configuration of turning points is shown in Fig. 4 for $z < 0$.

The wave to the left of pont A is of the form

$$\phi = \langle -k_1|0,z\rangle + \alpha\langle k_2|0,z\rangle, \quad (87)$$

where α is a constant to be determined. We assume that k_2 is dominant with respect to k_1 along the Stokes line $(-z_{T12}, A)$then ($-k_1$ is dominant with respect to $-k_2$ as well) and $-k_1$ is dominant with respect to k_1 along the ray (z_{T1}, B). Then, between (A,B) the wave has the form

$$\phi = \langle -k_1|0,z\rangle + i\alpha\langle k_2|0,-z_{T12}\rangle \langle k_1|-z_{T12},z\rangle + \alpha\langle k_2|0,z\rangle + i\langle -k_1|0,-z_{T12}\rangle \times\langle -k_2|-z_{T12},z\rangle. \quad (88)$$

Between $(B,0)$ the wave has the form

$$\phi = \langle -k_1|0,z\rangle + i\langle k_1|0,z\rangle \times(\alpha\langle k_2|0,-z_{T12}\rangle \langle k_1|-z_{T12},0\rangle - \langle -k_1|0,-z_{T1}\rangle \times\langle k_1|-z_{T1},0\rangle) + \alpha\langle k_2|0,z\rangle + i\langle -k_1|0,-z_{T12}\rangle \times\langle -k_2|-z_{T12},0\rangle \langle -k_2|0,z\rangle. \quad (89)$$

The evenness or oddness of the eigenfunction then demands that the coefficient of $\langle -k_1|0,z\rangle$ be within a sign the same as the coefficient of $\langle k_1|0,z\rangle$ and the same restriction applies to $\langle -k_2|0,z\rangle$ and $\langle k_2|0,z\rangle$. Hence we have

$$1 = \pm i\left(\alpha \exp\left\{i\int_{-z_{T12}}^{0} [k_1(z') - k_2(z')]\,dz'\right\} - \exp\left[2i\int_{-z_{T1}}^{0} k_1(z',\omega)\,dz'\right]\right), \quad (90)$$

$$\alpha = \pm i \exp\left\{i\int_{-z_{T12}}^{0} [k_1(z',\omega) - k_2(z',\omega)]\,dz'\right\}.$$

Eliminating α then yields

$$1 = -\exp\left\{i\int_{-z_{T12}}^{z_{T12}} [k_1(z',\omega) - k_2(z',\omega)]\,dz'\right\} + i\exp\left[i\int_{-z_{T1}}^{z_{T1}} k_1(z',\omega)\,dz'\right], \quad (91)$$

or its equivalent,

$$\left(1 + \exp\left\{i\int_{-z_{T12}}^{z_{T12}} [k_1(z',\omega) - k_2(z',\omega)]\,dz'\right\}\right)^2$$

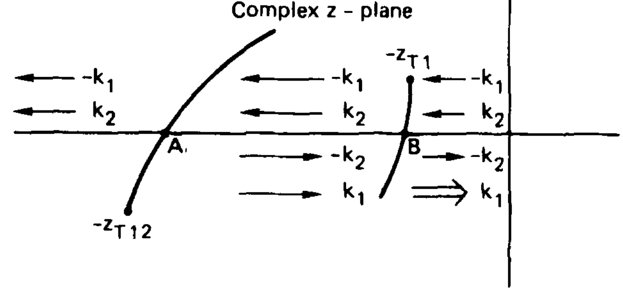


FIG. 4. Schematic diagram of Stokes lines that induce additional waves in four-wave problem.

$$= -\exp\left[2i\int_{-z_{T1}}^{z_{T1}} k_1(z',\omega)\,dz'\right]. \tag{92}$$

There are three terms in the dispersion relation above. A quantization rule, similar to that of Bohr–Sommerfeld, can only arise in the approximation that two of the terms are dominant. For example if the exponential in the large parantheses is small, the dispersion relation is equivalent to

$$\oint k dz = (2n+1)\pi, \tag{93}$$

where the countour integral encloses only the turning points $\pm z_{T1}$. If the exponential on the right hand side is small, the dispersion relation also becomes Eq. (93) with only the turning points z_{T12} and $-z_{T12}$ enclosed. The integral can be interpreted as k taking on the value k_1 which converts to k_2 on going around z_{T12} which converts to k_1 again on going around $-z_{T12}$. Because of the degeneracy of our problem we can also interpret Eq. (93) as to loop integral for $-k_2 \Rightarrow -k_1$ at $z_{T12} \Rightarrow -k_2$ at $-z_{T12}$.

Now if 1 is the small term in Eq. (92) and is neglected, Eq. (93) applies with the following interpretation. The reader is referred to the diagram in Fig. 5 to help himself follow this discussion. Starting from the $-k_2$ wave on the lower loop of the figure we have:

$-k_2$ propagates to z_{T12}, then converts to
$-k_1$ propagates to z_{T1}, then converts to
k_1 propagates to z_{T12}, then converts to
k_2 propagates to $-z_{T12}$, then converts to
k_1 propagates to $-z_{T1}$, then converts to
$-k_1$ propagates to $-z_{T12}$, then converts to
$-k_2$.

In Fig. 5 we see that the wave can go directly from $-z_{T1}$ to z_{T1} by tunnelling. If the 1 in Eq. (92) is included, the tunnelling term is roughly taken into account. A more precise evaluation is discussed in Ref. 3 where a specific example of this four-wave problem is evaluated.

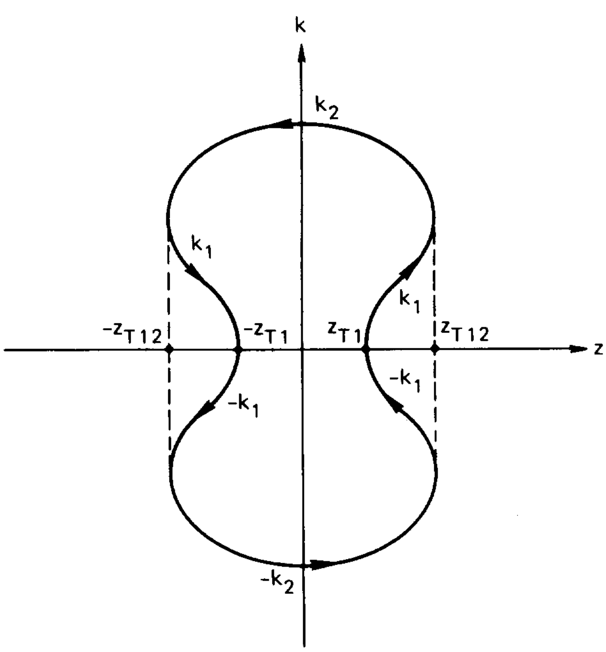


FIG. 5. Schematic closed curve in z-k plane.

The general WKB method we have outlined here is deficient for an integral equation in that, in principle, an infinite number of modes need to be followed. However, the eigenmode may be dominated by the interaction of just two modes, while the interaction with other modes is exponentially small. If a global dispersion relation is obtained with two modes, one can then ascertain whether the interaction with other modes is large or small. If it is large, a more complicated global dispersion relation can be obtained by using the methods indicated above.

ACKNOWLEDGMENTS

We gratefully acknowledge discussions with Keith Roberts, James Hammer, Rene Dominguez, and Harold Weitzer, and we would like to thank Gary Smith for aid in editing this manuscript.

APPENDIX A: SOLUTION IN n-DIMENSIONS

Equation (41) is of the form,

$$\frac{\partial}{\partial \mathbf{k}} |\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]| \cdot \frac{d}{d\mathbf{y}}\,\psi_r(\mathbf{y}) + G\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]\psi_r(\mathbf{y}) = 0, \tag{A1}$$

where $\mathbf{k}(\mathbf{y})$ is constrained to satisfy,

$$|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y}]| = 0. \tag{A2}$$

If $\mathbf{k}(\mathbf{y}_0) \equiv \mathbf{k}_0$ and $\psi(\mathbf{y}_0)$ are given at $y = y_0$, Eq. (A1) can be integrated by using the method of characteristics, which allows $\psi(y)$ to be evaluated along some curve of a complex parameter τ. The method of characteristics demands that $y(\tau)$ satisfy the differential equation,

$$\frac{d\mathbf{y}}{d\tau} = \alpha[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]\,\frac{\partial}{\partial \mathbf{k}}\,|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]|. \tag{A3}$$

As Eq. (A2) must be satisfied, we must simultaneously satisfy

$$\frac{d\mathbf{k}}{d\tau} = -\alpha[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]\,\frac{\partial}{\partial \mathbf{y}}\,|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]|. \tag{A4}$$

Equation (A3) and (A4) justify the usual assumption of the equation of motion of geometrical optics. To obtain an even closer correspondence we choose

$$\alpha[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega] = -1\Big/\frac{\partial}{\partial\omega}|\Lambda\,[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]|,$$

and the parameter τ is time (which can take on complex values).

Equation (A1) can then be written as

$$\frac{d}{d\tau}\left\{\psi_r[\mathbf{y}(\tau)]\exp\left[-\int_{\tau_0}^{\tau} d\tau' \frac{G\,[\mathbf{k}(\tau),\mathbf{y}(\tau')]}{(\partial/\partial\omega)|\Lambda\,[\mathbf{k}(\tau'),\mathbf{y}(\tau')]|}\right]\right\} = 0, \tag{A5}$$

where $\mathbf{y}(\tau_0) = \mathbf{y}_0$, $\mathbf{k}[\mathbf{y},(\tau_0)] = k_0$. The solution is

$$\psi_r[\mathbf{y}(\tau)] = \exp\left[\int_0^\tau \frac{d\tau' G[\mathbf{k}(\tau'),\mathbf{y}(\tau')]}{(\partial/\partial\omega)|\Lambda[\mathbf{k}(\tau'),\mathbf{y}(\tau')]|}\right]\psi_r(\mathbf{y}_0). \tag{A6}$$

The problem with this solution is that the coordinates $\mathbf{y}(\tau)$ are not real. We can choose a complex τ path so that one component of $\mathbf{y}(\tau)$ is real, but in general the other components are complex. Hence this method of integration fails to determine $\psi_r[\mathbf{y},\mathbf{k}(\mathbf{y})]$ at physical coordinates. An exception to this rule occurs when $\Lambda_{ij}(\mathbf{k},\mathbf{y},\omega)$ is Hermitian for real ω. For this case the equations of motion of $\mathbf{y}$ and $\mathbf{k}$ are Hamiltonian equations for a real Hamiltonian ω when $\mathbf{y}$ and $\mathbf{k}$ are real. The trajectories are also real and the method of characteristics then determines the field at real-space coordinates.

If the anti-Hermitian part of $\Lambda_{ij}(\mathbf{k},\mathbf{y},\omega)\equiv\Lambda_{Rij}(\mathbf{k},\mathbf{y},\omega) + i\epsilon\Lambda_{Aij}(\mathbf{k},\mathbf{y},\omega)$ is small, i.e., $\epsilon\Lambda_{Aij}\ll\Lambda_{Rij}$ (as assumed in Ref. 6) then the anti-Hermitian part of Λ_{ij} can be considered as first order in ϵ, and the WKB procedure can be straightforwardly modified to yield for arbitrary r,

$$\frac{\partial}{\partial\mathbf{k}}|\Lambda_R[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]|\cdot\frac{d}{d\mathbf{y}}\psi_r(y) + \psi_r(\mathbf{y})\{G_R[\mathbf{k}(\mathbf{y}),\mathbf{y}]\} + \frac{\Lambda_R^{ri}[\mathbf{k}(\mathbf{y}),\mathbf{y}]\Lambda_{Aij}[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]\Lambda_R^{jr}[\mathbf{k}(\mathbf{y}),\mathbf{y}]}{\Lambda_R^{rr}[\mathbf{k}(\mathbf{y}),\mathbf{y}]} = 0, \tag{A7}$$

where $\mathbf{k}(\mathbf{y})$ satisfies $|\Lambda_R[\mathbf{k}(\mathbf{y}),\mathbf{y},\omega]|=0$ and

$$G_R[\mathbf{k}(\mathbf{y}),\mathbf{y}] = \frac{1}{2}\frac{d}{d\mathbf{y}}\cdot\frac{\partial}{\partial\mathbf{k}}|\Lambda_R[\mathbf{k}(\mathbf{y}),\mathbf{y}]| + \{\Lambda_R[\mathbf{k}(\mathbf{y}),\mathbf{y}]\}^r.$$

The method of characteristics now leads to real trajectories for real ω, as $|\Lambda_R|$ is Hermitian. The solution is then

$$\psi_r[\mathbf{y}(\tau)] = \psi_r(\mathbf{y}_0)\exp\left(-\int_{\tau_0}^{\tau} d\tau'\left\{\omega_I(\tau') - \frac{G_R[\mathbf{k}(\tau'),\mathbf{y}(\tau')]}{(\partial/\partial\omega)|\Lambda_R[\mathbf{k}(\tau'),\mathbf{y}(\tau')]|}\right\}\right),$$

with

$$\omega_I(\tau') = \frac{-\Lambda_R^{ri}[\mathbf{k}(\tau),\mathbf{y}(\tau)]\Lambda_{Aij}[\mathbf{k}(\tau),\mathbf{y}(\tau)]\Lambda_R^{jr}[\mathbf{k}(\tau),\mathbf{y}(\tau)]}{\Lambda_R^{rr}[\mathbf{k}(\tau),\mathbf{y}(\tau)](\partial/\partial\omega)|\Lambda_R[\mathbf{k}(\tau),\mathbf{y}(\tau)]|}.$$

The term $\exp(-\int^\tau \omega_I\, d\tau')$ is the attenuation factor along the path of the trajectory.

APPENDIX B: INTERPRETATION OF WKB CORRECTION FACTOR

Consider two systems that are given by the integral equations

$$\int dx' G\left(x-x',\epsilon\frac{x+x'}{2}\right)\xi(x') = 0,$$
$$\int dx' \tilde{G}\left(x-x',\epsilon\frac{x+x'}{2}\right)\tilde{\xi}(x') = 0, \tag{B1}$$

where G and $\tilde{G}$ are square matrices, $\xi(x)$ and $\hat{\xi}(x')$ are column matrices, and

$$\tilde{G} = U(\epsilon x)G(x-x',\epsilon(x+x')/2)U^{-1}(\epsilon x'). \tag{B2}$$

Clearly, as $\tilde{G}$ is a similarity transformation of G, the two systems are equivalent with identical eigenvalues ω, with their solutions related by

$$\tilde{\xi}(x) = U\xi(x). \tag{B3}$$

However, the WKB solutions for $k(x)$ of these two systems differ as,

$$\Lambda(k,y) = \int dz\exp(-ikz)G(z,y), \tag{B4}$$

$$\begin{aligned}\tilde{\Lambda}(k,y) &= \int dz\exp(-ikz)U(y+\epsilon z/2)G(z,x)\times U(y-\epsilon z/2)\\ &= \int dz\exp(-ikz)\Big\{U(y)G(z,y)U^{-1}(y) + \epsilon\frac{z}{2}[U_y(y)G(z,y)U^{-1}(y)] - U(y)G(z,y)U_y^{-1}(y)\Big\} + O(\epsilon^2),\end{aligned} \tag{B5}$$

where $y=\epsilon x$.

The two k values for each system are denoted as $k(y)$ and $\tilde{k}(y)$ which are respectively the solutions of

$$|\Lambda(k,y)| = 0,\quad |\tilde{\Lambda}(k,y)| = 0. \tag{B6}$$

To find $\tilde{k}(y)$ in terms of $k(y)$ (in the remaining text we suppress the argument "y") we multiply Eq. (B5) on the left and right by the matrix $\mathrm{cof}\tilde{\Lambda}(\tilde{k},y)$ whose matrix elements are $\tilde{\Lambda}^{rs}$. The left-hand side of Eq. (B5) then vanishes as $|\tilde{\Lambda}(\tilde{k},y)|=0$, and from the remaining terms we obtain

$$\mathrm{cof}\tilde{\Lambda}(\tilde{k},y)U(y)\Lambda(\tilde{k},y)U^{-1}(y)\,\mathrm{cof}\tilde{\Lambda}(\tilde{k},y) = \frac{-i\epsilon}{2}\mathrm{cof}\tilde{\Lambda}(U_y\Lambda_k U^{-1} - U\Lambda_k U_y^{-1})\,\mathrm{cof}\tilde{\Lambda}. \tag{B7}$$

Now $\mathrm{cof}\tilde{\Lambda}(\tilde{k},y) = \mathrm{cof}\tilde{\Lambda}(k,y) + \epsilon[\mathrm{cof}\tilde{\Lambda}(k,y)]_1$, and since $|\Lambda(k,y)|=0$, we find

$$\begin{aligned}2i\Delta k\,|\Lambda|_k\,\mathrm{cof}\tilde{\Lambda}(k,y) &= \mathrm{cof}\Lambda(k,y)[U_y\Lambda_k(k,y)U^{-1} - U\Lambda_k(k,y)U^{-1}]\\ &\quad\times\mathrm{cof}\Lambda(k,y) + O(\epsilon),\end{aligned} \tag{B8}$$

where $\epsilon\Delta k = \tilde{k} - k$. Additional algebraic manipulation of Eq. (B8) yields [See Eq. (78) for notation]

$$\Delta k = -\delta k^{\nu\mu} + [UM(\delta k)U^{-1}]_{\nu\mu}/\Lambda^{\nu\mu} + \frac{i}{2}\frac{(U_y\,\mathrm{cof}\Lambda\, U^{-1} - U\,\mathrm{cof}\Lambda\, U_y^{-1})_{\nu\mu}}{\Lambda^{\nu\mu}}, \tag{B9}$$

where $M(\delta k)$ is a matrix whose matrix elements are $\delta k^{rs}\Lambda^{rs}$ (no summation implied).

Equation (B9) determines the lowest order WKB shift of the local k value of systems related by a similarity transformation. We now wish to indicate how the correction of the WKB phase integral given in Eq. (79) allows for the compensation of this shift, so that the predicted WKB eigenvalues of these two systems agree to (ϵ^2). More precisely we need to show that

$$\oint dz(\tilde{k}+\delta\tilde{k}) \equiv \oint dz(k+\Delta k+\delta\tilde{k}) = \oint dz(k+\delta k),\quad (\text{let } \epsilon=1),\ \text{or}$$

$$\oint dz(\Delta k + \delta\tilde{k} - \delta k) = 0. \tag{B10}$$

From Eq. (B3) we expect that the WKB solutions of the equivalent solutions are related by,

$$\tilde{\xi} \exp\left(i \int_{y_0}^{y} \Delta k \, dy'\right) = U\hat{\xi}. \tag{B11}$$

From Eq. (43) of the text it follows that

$$\frac{d}{dy}\left[\hat{\xi}^{\,\nu}(y)\left(\frac{\partial|\Lambda|}{\partial k}\right)^{1/2}\right] = -\frac{{}^{\nu}\{\Lambda\}\hat{\xi}^{\,\nu}(y)}{(\partial|\Lambda|/\partial k)^{1/2}}, \tag{B12}$$

$$\frac{d}{dy}\left[\hat{\tilde{\xi}}^{\,\nu}(y)\left(\frac{\partial|\Lambda|}{\partial k}\right)^{1/2}\right] = -\frac{{}^{\nu}\{\tilde{\Lambda}\}\tilde{\xi}^{\,\nu}(y)}{(\partial|\Lambda|/\partial k)^{1/2}}. \tag{B13}$$

If Eq. (B11) is substituted into the left hand side of Eq. (B13), one establishes, after some algebra, that the two sides of Eq. (B13) will be equal as a consequence of Eq. (B9). Thus, the conjecture of Eq. (B11) is verified). We now use,

$$\begin{aligned}
\hat{\tilde{\xi}}^{\,\nu}(y,y_0) &\equiv \hat{\tilde{\xi}}^{\,\nu}(y)\left(\frac{|\Lambda(k,y)|_k}{|\Lambda(k_0,y_0)|_k}\right)^{1/2} \\
&= \hat{\tilde{\xi}}^{\,\nu}(y_0)\exp\left[-\int_{y_0}^{y}\frac{dy'\,{}^{\nu}\{\Lambda\}}{|\Lambda|_k}\right] \\
&= \hat{\tilde{\xi}}^{\,\nu}(y_0)\left(\frac{\tilde{\Lambda}^{\,\nu\nu}(k,y)}{\tilde{\Lambda}^{\,\nu\nu}(k_0,y_0)}\right)^{1/2}\exp\left[i\int_{y_0}^{y}\delta k^{\,\nu\nu}\,dz'\right],
\end{aligned} \tag{B14}$$

$$\begin{aligned}
\tilde{\xi}^{\,r}(y,y_0) &= \hat{\xi}^{\,r}(y_0)\exp\left[-\int_{y_0}^{y}dy'\,{}^{r}\{\Lambda\}\Big/|\Lambda|_k\right] \\
&= \hat{\xi}^{\,r}(y_0)\left(\frac{\Lambda^{\,rr}(k,y)\Lambda^{\,r\nu}(k,y)\Lambda^{\,\nu r}(k_0,y_0)}{\Lambda^{\,rr}(k_0,y_0)\Lambda^{\,r\nu}(k_0,y_0)\Lambda^{\,\nu r}(k,y)}\right)^{1/2} \\
&\quad\times\exp\left[i\int_{y_0}^{y}\delta k^{\,\nu\nu}\,dz\right].
\end{aligned} \tag{B15}$$

Further, referring to the two WKB solutions merging at $z = z_{T1}$ as "1" and "2", we have

$$\lim_{y_0\to y_{T1}} \frac{\hat{\tilde{\xi}}^{\,\nu}_2(y_0)}{\tilde{\xi}^{\,\nu}_1(y_0)} = \frac{\hat{\xi}^{\,r}_2(y_0)}{\hat{\xi}^{\,r}_1(y_0)} = \pm i.$$

Then if we construct

$$\lim_{\substack{y\to y_{T2}\\ y_0\to y_{T1}}}\left[\frac{\hat{\tilde{\xi}}^{\,\nu}_2(y,y_0)\exp(i\int_{y_0}^{y}\Delta k_2\,dy')}{\hat{\tilde{\xi}}^{\,\nu}(y,y_0)\exp(i\int_{y_0}^{y}\Delta k_1\,dy')}\right] = \frac{\Sigma_s\, U_{\nu s}(y)\hat{\xi}^{\,r}_2(y,y_0)}{\Sigma_s\, U_{\nu s}(y)\hat{\xi}_1(y,y_0)},$$

we find that the amplitude factors cancel and the phase dependence demands,

$$\int_{y_{T1}}^{y_{T2}}[(\delta\tilde{k}_2 + \Delta k_2 - \delta k_1) - (\delta k_1 + \Delta k_1 - \delta k_1)]\,dz' \equiv \oint(\delta k + \Delta k - \delta k)\,dz' = 0,$$

which is the relation we desired to prove.

[1] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New Yrok, 1958), p. 157.

[2] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, New York 1960), p. 284.

[3] H. L. Berk and D. L. Book, Phys. Fluids **12**, 649 (1969).

[4] D. Watson, *Alfven Ion-Cyclotron Instability in Mirror Machines*, Lawrence Livermore Laboratory, UCRL-82068 (to be published in Phys. Fluids).

[5] T. Watanabe, H. Sanuki, and M. Watanabe, "Treatment of Eigenmode Analysis of Vlasov Equation," (to be published J. Phys. Soc. Jpn.).

[6] I. B. Bernstein, Phys. Fluid, **19**, 320 (1975).

[7] I. C. Percival, Adv. Chem. Phys. **36**, 1 (1977).

[8] J. Heading, *An Introduction to Phase Integral Methods* (Wiley, New York, 1962).

[9] H. L Berk and R. R. Dominguez, J. Plasma Phys. **18**, 31 (1977).

[10] J. Knoll and R. Schaeffer, Ann. Phys. (N.Y.) **97**, 307 (1976).

Partial inner product spaces. IV. Topological considerations

J.-P. Antoine
Institut de Physique Théorique, Université Catholique de Louvain, Belgium

(Received 30 January 1980; accepted for publication 20 March 1980)

Whereas the third paper in this series dealt with the algebraic structure of partial inner product (PIP) spaces, the present one explores systematically their topological properties. A slightly more restricted object is introduced, that we call an indexed PIP-space: it consists of a PIP-space together with a distinguished family of assaying subspaces. The upshot of the analysis is the characterization of two types of indexed PIP-spaces, called type (B) and type (H), respectively, as the most likely candidates for practical applications; they are simply lattices of Banach, resp. Hilbert spaces. Operators on indexed PIP-spaces are discussed and conditions are given that guarantee that the domain of any such operator is a vector subspace. Finally, we examine the question of the existence of a central Hilbert space, in the case of a positive definite partial inner product.

1. INTRODUCTION

This paper is the fourth in a series devoted to the systematic study of partial inner product (PIP) spaces (see Refs. 1–3; these will be called I, II, and III, respectively, in the sequel). Parts I and II presented the basic definitions about the spaces and operators between them. Paper III in a sense reversed the perspective: Whereas the general definition of I used as only input a linear compatibility on a vector space V, it was shown in III that the same object may be obtained from a suitable lattice of subspaces of V. This made contact with earlier approaches like Gel'fand's[4] and Grossmann's.[5] In the present paper we shall continue the analysis of III. Whereas the latter was concerned with the algebraic structure only (i.e., the compatibility relation), we will discuss here primarily the topological stucture given by the partial inner product (pip) itself. The aim is to tighten the definitions so as to eliminate as many pathologies as possible. The picture that emerges is reassuringly simple: Only two types of PIP-spaces seem sufficiently regular to have any practical use. Roughly speaking they consist of lattices of Hilbert or Banach spaces.

Let us now sketch the contents of the paper in more detail. We will keep throughout the notations and terminology of I–III, and our standard reference on topological vector spaces will be the textbook of Köthe.[6] In addition, for the convenience of the reader, we have collected in an Appendix most of the necessary, but not so familiar, notions needed in the text.

The first step is to analyze a single dual pair $\langle V_r, V_{\bar{r}}\rangle$ of assaying subspaces (Sec. 2). It turns out, the useful situation is that of a *reflexive* dual pair.[6] Next we investigate the lattice generated, by intersection and vector sum, from two such reflexive pairs. First, in Sec. 3, we restrict our attention to pairs of Banach spaces. This is the typical framework of (abstract) interpolation theory,[7] which has in fact a great role to play in the development of PIP-space theory. But, when we try, in Sec. 4, to generalize the results to arbitrary reflexive pairs, no general conclusion can be reached, for too many pathologies are possible in the general setup. There is a way out, however. As the examples indeed show, the complete lattice $\mathscr{F}$ of *all* assaying subsets is in general extremely large. Fortunately, as shown in III, the whole structure can be reconstructed from a fairly small sublattice $\mathscr{I}$ of $\mathscr{F}$, and on such a family of subsets it makes sense to impose additional restrictions, typically to contain only topological vector of the same type. In this way, we are led to a slightly more restricted structure, called an *indexed PIP-space*. Technically it consists of a PIP-space with a distinguished rich sublattice $\mathscr{I}\subset\mathscr{F}$. Various conditions, mostly topological, can be imposed on that family $\mathscr{I}$, which we explore systematically in Sec. 5. Two useful structures emerge: Indexed PIP-spaces of type (B), resp. type (H), where every $V_r \in \mathscr{I}$ is a reflexive Banach space, resp. a Hilbert space. Several examples, mostly taken from III, illustrate the new concepts. At this point, of course, we have essentially returned to the philosophy of Grossmann's nested Hilbert spaces[5] except for the condition of positivity of the pip. Before looking into that matter, we discuss briefly (Sec. 6) the notion of operator on an indexed PIP-space. In fact, it differs only in a simple and obvious way from the object defined in II: Its domain of definition is a union of elements of $\mathscr{I}$, instead of general assaying subsets from $\mathscr{F}$. A useful result is a sufficient condition for that domain to be a vector subspace of V. Interestingly enough this condition is automatically satisfied for spaces of type (B) or (H).

The last section is devoted to the so-called central Hilbert space. If the pip is positive definite, whenever defined, a unique central subspace $\mathscr{H}$ may be obtained as the Hilbert space completion of $V^{\#}$, in the norm defined by the pip, provided a simple condition of completeness is satisfied. Moreover, in most cases, either the resulting subspace is assaying and self-dual, or the same result can be obtained by slightly refining (in the technical sense of III) the compatibility.

Here also, this is automatic for spaces of type B and H, which shows once again that these are the most useful struc-

[+]Postal Address: Institut de Physique Théorique UCL, Chemin du Cyclotron, 2, B-1348-Louvain-La-Neuve (Belgium).

tures for practical applications. Those results had actually been obtained previously in a particular case, in a joint work with W. Karwowski.[8]

One may say, at this stage the groundwork is finished and the time has come for a deeper study of operators on indexed PIP-spaces, especially those operators that arise in quantum contexts, such as scattering theory or quantum field theory. Work in that direction is in progress and will be reported on later. Also interpolation theory finds here, in our opinion, a natural framework, which deserves further study.

2. TOPOLOGIES ON DUAL PAIRS OF ASSAYING SUBSETS

Let $(V, \#, \langle\cdot|\cdot\rangle)$ be a nondegenerate PIP-space. In this section we will focus our attention to a single dual pair $\langle V_r, V_{\bar{r}}\rangle$ of assaying subsets. (In fact, we should speak of an *antidual* pair, but it makes no difference; see Schwartz[9] for a full discussion). What are the possible topologies on V_r, $V_{\bar{r}}$? In answering this question, we will follow mostly the terminology of Köthe[6]; however, for a given dual pair $\langle E, F\rangle$, $\sigma(E, F)$ will denote the weak topology on E, $\tau(E,F)$ its Mackey topology, and $\beta(E,F)$ its strong topology (see the Appendix).

From the compatibility (or lattice) point of view, there is perfect symmetry between V_r and $V_{\bar{r}}$. In any natural scheme, this feature should be preserved at the topological level as well.

As stated in I, every assaying subset V_r is endowed with its canonical Mackey topology $\tau(V_r,V_{\bar{r}})$ and this will always be understood in the sequel, unless stated otherwise. We will write $V_r|_\tau$ or even $V_r[\tau(V_r,V_{\bar{r}})]$ if a danger of confusion arises. From this choice (perfectly symmetric with respect to $V_r,V_{\bar{r}}$) it follows already that (see I):

(i) the dual of V_r is $V_{\bar{r}}$;

(ii) whenever $V_p \subset V_q$, the injection $E_{qp}: V_p \rightarrow V_q$ is continuous and has dense range;

(iii) $V^{\#}$ is dense in every V_r, and every V_r is dense in V.

However, that is not sufficient for eliminating all pathologies. For practical purposes, indeed, the Mackey topology $\tau(V_r,V_{\bar{r}})$ is rather awkward, or at least unfamiliar, unless it coincides with a norm or a metric topology. If a locally convex space $E[T]$ with topology T is metrizable, then T coincides with the canonical Mackey topology on E, i.e., $T=\tau(E,E')$ (but not necessarily with the strong topology $\beta(E,E')$ if $E[T]$ is not complete). Let us give two examples to emphasize the point.

(i) Take the dual pair $\langle\varphi,l^2\rangle$, with respect to the l^2 inner product (φ is the space of all finite sequences); then $\tau(\varphi,l^2)$ is the l^2-norm topology on φ, but it is coarser than $\beta(\varphi,l^2)$ (see Ref. 6, Sec. 21.5), whereas $\tau(l^2,\varphi)$, although metrizable, is not a norm topology.

(ii) Take $\langle l^1,l^\infty\rangle$. Then $\tau(l^1,l^\infty)$ coincides with $\beta(l^1,l^\infty)$ and the l^1-norm topology, but $\tau(l^\infty,l^1)$ is weaker than the l^∞-norm topology, and, indeed, is not metrizable.

The origin of the difficulty is clear: $\varphi[\tau(\varphi,l^2)]$ is a noncomplete normed space, whereas $l^1[\tau(l^1,l^\infty)]$ is a nonreflexive Banach space. Such pathologies are avoided if the dual pair $\langle V_r,V_{\bar{r}}\rangle$ is *reflexive*, in the sense of Köthe,[6] i.e., if the dual of $V_r|_\beta$ coincides with $V_{\bar{r}}$ and vice versa. This is indeed equivalent with either (hence both) $V_r|_\tau$ or $V_{\bar{r}}|_\tau$ being reflexive (a locally convex space E is called *reflexive* if it coincides with the strong dual of its strong dual; see the Appendix for further details, as well as for other, equivalent, characterizations of reflexive dual pairs). In addition, each space of a reflexive dual pair is quasicomplete (i.e., closed bounded sets are complete) for the weak, the Mackey and the strong topology (the two last ones in fact coincide). We shall make use of this fact in Sec. 7. Typical instances of reflexive dual pairs are the following:

(i) V_r is a Hilbert space; so is then $V_{\bar{r}}$:

(ii) V_r is a *reflexive* Banach space; so is then $V_{\bar{r}}$;

(iii) V_r is a *reflexive* Fréchet space; $V_{\bar{r}}$ is then a reflexive complete (DF)-space[6];

(iv) V_r is a Montel space[6]; so is then $V_{\bar{r}}$.

As we shall see in the sequel, these cases cover already most spaces of practical interest, in particular all spaces of distributions.

Actually cases (i) and (ii) play a special role in the theory, for they have specially nice properties; we will study them systematically in Sec. 3 below. What distinguishes them from the others is metrizability. Indeed:

Proposition 2.1: Let $\langle V_r,V_{\bar{r}}\rangle$ be a reflexive dual pair. Then $V_r|_\tau$ and $V_{\bar{r}}|_\tau$ are reflexive Banach spaces if and only if they are both metrizable.

Proof: The "only if" part if obvious. Let $V_r|_\tau$ and $V_{\bar{r}}|_\tau$ be metrizable. By reflexivity, they are strong duals of each other, since $\tau(V_r,V_{\bar{r}})=\beta(V_r,V_{\bar{r}})$ and $\tau(V_{\bar{r}},V_r) =\beta(V_{\bar{r}},V_r)$. Hence they are both normable since the strong dual of a metrizable locally convex space can only be metrizable if both are normable. Finally V_r and $V_{\bar{r}}$ are both normed spaces and reflexive, hence Banach spaces. ■

3. INTERPLAY BETWEEN TOPOLOGICAL AND LATTICE PROPERTIES: THE BANACH CASE

Let $\langle V_p,V_{\bar{p}}\rangle$, $\langle V_q,V_{\bar{q}}\rangle$ be two reflexive dual pairs of assaying subsets, consisting of (reflexive) Banach spaces. What can be said about the pair $\langle V_{p\wedge q},V_{\bar{p}\vee\bar{q}}\rangle$? By definition, $V_{p\wedge q}$ is the vector space $V_p \cap V_q$, and $V_{\bar{p}\vee\bar{q}}$ is $(V_{\bar{p}}+V_{\bar{q}})^{\#\#}$, which *a priori* could be larger that $V_{\bar{p}}+V_{\bar{q}}$ (usually they coincide, see Lemma 3.4 and Proposition 3.7 below). In order to appreciate the situation, we will introduce two auxiliary spaces $V_{[p,q]}$ and $V_{(\bar{p},\bar{q})}$ (in the notation of Ref. 5), using a standard construction from interpolation theory[7] which we sketch now.

Let X_a, X_b be two Banach spaces, which are both continuously embedded in a Hausdorff TVS X (two such Banach spaces are called an *interpolation couple*; see Ref. 7). Let $X_a \oplus X_b$ denote their direct sum; it is also a Banach space, with norm $\|(f,g)\|=\|f\|_a+\|g\|_b$, $(f\in X_a, g\in X_b)$. We consider the subspace $X_{[a,b]}$ of $X_a \oplus X_b$ which consists of all pairs of the form $(f,-f)$ for some $f\in X_a \cap X_b$. This subspace is obviously isomorphic to $X_a\cap X_b$ and it is closed in $X_a \oplus X_b$. With the induced topology, we will denote $X_{[a,b]}$ by $(X_a \cap X_b)_{\text{proj}}$. Indeed, the induced topology is precisely the projective limit (Appendix) of the two norm (= Mackey) topologies on X_a, X_b. Thus $(X_a \cap X_b)_{\rm proj}$ is again a Banach space, with norm:

$$\|f\|_{[a,b]} \equiv \|f\|_a + \|f\|_b \quad (f \in X_a \cap X_b). \tag{3.1}$$

Next we define the quotient $X_{(a,b)} \equiv (X_a \oplus X_b)/X_{[a,b]}$. As a vector space $X_{(a,b)}$ is isomorphic to the vector sum $X_a + X_b$. Equipped with the quotient topology, $X_{(a,b)}$ will be denoted by $(X_a + X_b)_{\rm ind}$, for it is precisely the inductive limit (Appendix) of X_a, X_b with respect to the mappings $X_a \to X_a + X_b$, $X_b \to X_a + X_b$. It is again a Banach space, with norm (that the following expression is indeed a norm results from the continuous embedding of X_a, X_b into X[7]):

$$\|f\|_{(a,b)} \equiv \inf_{f = g + h} (\|g\|_a + \|h\|_b), \tag{3.2}$$

where the infimum is taken over all possible decompositions $f = g + h$, $g \in X_a$, $h \in X_b$; such decompositions are non unique as soon as $X_a \cap X_b \neq \{0\}$.

Proposition 3.1: Let X_a, X_b be two Banach spaces, both continuously embedded in a Hausdorff space X. Then:

(i) The two spaces $(X_a \cap X_b)_{\rm proj}$ and $(X_a + X_b)_{\rm ind}$ are Banach spaces and the following inclusions hold, where $\twoheadrightarrow$ denotes a continuous injection:

$$(X_a \cap X_b)_{\rm proj} \twoheadrightarrow \begin{Bmatrix} X_a \\ X_b \end{Bmatrix} \twoheadrightarrow (X_a + X_b)_{\rm ind}. \tag{3.3}$$

(ii) The norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are *consistent* on $X_a \cap X_b$: if $\{f_n\} \in X_a \cap X_b$ is Cauchy in both norms and $f_n \to 0$ in X_a, then $f_n \to 0$ in X_b also.

Proof: Part (i) is clear from the discussion above. As for (ii) let $X_{ab|a}$ (resp. $X_{ab|b}$) be the image of $X_a \cap X_b$ in X_a (resp. X_b) under the identity map. Denote by E_{ba} the identity map $X_{ab|a} \to X_{ab|b}$. The graph of E_{ba} is exactly the set $X_{[a,b]}$, which is closed in $X_a \oplus X_b$; thus E_{ba} is a closed map, which means precisely that the norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are consistent on $X_a \cap X_b$. ■

Remarks:

(i) If X_a and X_b are reflexive, so are $(X_a \cap X_b)_{\rm proj}$ and $(X_a + X_b)_{\rm ind}$.

(ii) If X_a and X_b are Hilbert spaces, the same construction goes through, using squared norms everywhere (see Ref. 9).

(iii) The construction above and Proposition 3.1 remain valid if X_a, X_b are assumed to be Fréchet spaces.

(iv) It is shown in Ref. 10 (Appendix to IX. 4) that the norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are consistent on $X_a \cap X_b$ iff the expression $\|\cdot\|_{(a,b)}$ is a norm on $X_a \cap X_b$ and the identity map on $X_a \cap X_b$ extends to continuous injections of $\widehat{X}_a$, resp. $\widehat{X}_b$, into $\widehat{X}_{(a,b)}$, defined as the completion of $X_a \cap X_b$ under $\|\cdot\|_a$, resp. $\|\cdot\|_b$, $\|\cdot\|_{(a,b)}$. When $X_a \cap X_b$ is dense in X_a, X_b, then $\widehat{X}_a = X_a$, $\widehat{X}_b = X_b$, $\widehat{X}_{(a,b)} = (X_a + X_b)_{\rm ind}$.

Let now $X_{\bar a}$, $X_{\bar b}$ be the duals of X_a, X_b respectively. We assume now that $X_a \cap X_b$ is dense in X_a and in X_b. It follows that $X_{\bar a}$ and $X_{\bar b}$ can be embedded into $(X_a \cap X_b)'_{\rm proj}$, i.e., $\{X_{\bar a}, X_{\bar b}\}$ is also an interpolation couple (see Ref. 7). Then we have:

Lemma 3.2: Let X_a, X_b as above. Then:

(i) The dual of $(X_a \cap X_b)_{\rm proj}$ is $X_{\bar a} + X_{\bar b}$.

(ii) The dual of $(X_a + X_b)_{\rm ind}$ is $X_{\bar a} \cap X_{\bar b}$.

Proof: The proof of both assertions results from the following two observations:

(i) The dual of $X_a \oplus X_b$ is $X_{\bar a} \oplus X_{\bar b}$.

(ii) The closed subspace $X_{[\bar a,\bar b]}$ of $X_{\bar a} \oplus X_{\bar b}$ is isomorphic, as a vector space, to the orthogonal space of $X_{[a,b]}$, the isomorphism being $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ in $X_{\bar a} \oplus X_{\bar b}$. ■

We return now to the PIP-space V and the two pairs of Banach assaying subsets $\langle V_p, V_{\bar p} \rangle$, $\langle V_q, V_{\bar q} \rangle$ (they are necessarily reflexive, by Proposition 2.1, but the whole discussion that follows is independent of this fact). All four assaying subsets are continuously embedded into V, and $V_p \cap V_q$ is dense in V_p and V_q. Hence $\{V_p, V_q\}$ and $\{V_{\bar p}, V_{\bar q}\}$ are interpolation couples and the construction above goes through. Thus, we get the following scheme (and the corresponding one for the duals, taking Lemma 3.2 into account), where all injections are continuous and have dense range:

$$V_{p\wedge q}|_\tau \twoheadrightarrow (V_p \cap V_q)_{\rm proj} \twoheadrightarrow \begin{Bmatrix} V_p \\ V_q \end{Bmatrix} \twoheadrightarrow (V_p + V_q)_{\rm ind} \twoheadrightarrow V_{p \vee q}|_\tau. \tag{3.4}$$

It will prove useful to introduce the following conditions for the couple V_p V_q:

(PROJ) $V_{p\wedge q}|_\tau \simeq (V_p \cap V_q)_{\rm proj}$ (TVS isomorphism),

(ADD) $V_{\bar p \vee \bar q} = V_{\bar p} + V_{\bar q}$ (as vector spaces),

(IND) $V_{\bar p \vee \bar q}|_\tau \simeq (V_{\bar p} + V_{\bar q})_{\rm ind}$ (TVS isomorphism).

Lemma 3.3: Let V_p, V_q be Banach assaying subsets. Then (ADD) $\Leftrightarrow$ (IND).

Proof: Let $V_{\bar p \vee \bar q} = V_{\bar p} + V_{\bar q}$. Then we have:

$$\langle V_{p\wedge q}, V_{\bar p \vee \bar q} \rangle \equiv \langle V_p \cap V_q, V_{\bar p} + V_{\bar q} \rangle.$$

Since Mackey topologies are inherited both by direct sums and by quotients, it follows that $\tau(V_{\bar p} + V_{\bar q}, V_p \cap V_q)$ is the quotient topology inherited from $V_{\bar p} \oplus V_{\bar q}$, that is the inductive topology on $V_{\bar p} + V_{\bar q}$; this means that $V_{\bar p \vee \bar q}|_\tau \simeq (V_{\bar p} + V_{\bar q})_{\rm ind}$. The converse assertion is obvious. ■

Lemma 3.4: Let again V_p, V_q be Banach assaying subsets. Then (PROJ) implies (ADD).

Proof: The dual of $V_{p\wedge q}|_\tau$ is, by definition, $V_{\bar p \vee \bar q}$. Then the result follows from Lemma 3.2. ■

Lemma 3.5: Let again V_p, V_q be Banach assaying subsets. Then (ADD) implies (PROJ).

Proof: V_p and V_q being Banach spaces, so is $V_p \cap V_q$ under the projective topology. Therefore it carries its Mackey topology, namely, $\tau(V_p \cap V_q, (V_p \cap V_q)'_{\rm proj}) = \tau(V_p \cap V_q, V_{\bar p} + V_{\bar q})$, which is $\tau(V_{p\wedge q}, V_{\bar p \vee \bar q})$ by the condition (ADD). ■

Summarizing Lemmas 3.3 to 3.5, we get:

Proposition 3.6: Let V_p, V_q be Banach assaying subsets. Then all the conditions (PROJ), (ADD), (IND) are equivalent. ■

Remark: Proposition 3.5 remains true if V_p, V_q are only assumed to be Fréchet spaces. Notice also that reflexivity has *not* been assumed although it will be satisfied in practice (compare Proposition 2.1).

At this point, we must return to the general discussion of III. There we observed that, quite often, a linear compatibility relation on a vector space V is defined in terms of an involutive covering $\mathscr{I}$ of V, that is, an involutive lattice of subspaces of V (with set intersection as infimum). It is convenient to work with this family $\mathscr{I}$ only and forget the complete lattice $\mathscr{F}$ of all assaying subsets it generates. Operators, in particular, are best studied in this restricted setup. This we will do systematically in Sec. 5 and 6 below. As in the general cases described in II, an operator A can be studied through its representatives $A_{qp}: V_p \to V_q$, but this is helpful only if the corresponding assaying subsets V_p, V_q are sufficiently simple, like Banach or Hilbert spaces. This will lead us to PIP-spaces defined by involutive coverings that consist of Banach or Hilbert spaces only.

A further element of simplicity is that $\mathscr{I}$ be a sublattice of the lattice $\mathscr{L}(V)$ of all subspaces of V, i.e., that (ADD) hold for *every* couple of elements of $\mathscr{I}$:

$$V_{p\vee q} = V_p + V_q, \ \forall V_p, V_q \in \mathscr{I}. \tag{3.5}$$

As we will see in Sec. 6 below, condition (3.5) implies that the domain $\mathscr{D}(A)$ of *every* operator A be a vector subspace of V. Thus it is desirable to find conditions that guarantee the validity of Eq. (3.5). In view of Proposition 3.6, this is equivalent to requiring that (PROJ) hold for every couple of elements of $\mathscr{I}$. It turns out that the condition is always satisfied when $\mathscr{I}$ consists of Banach spaces only. Indeed:

Proposition 3.7: Let V_p, V_q and $V_{p\wedge q}|_\tau$ be complete metrizable (i.e., Fréchet spaces). Then $V_{p\wedge q}|_\tau$ is isomorphic to $(V_p \cap V_q)_{\text{proj}}$, that is, (PROJ) holds.

Proof: As vector spaces, $V_{p\wedge q} = V_p \cap V_q$. Hence $V_p \cap V_q$ carries two distinct topologies for which it is metrizable and complete, namely, its Mackey topology (by assumption) and the projective topology. The latter is coarser; hence they coincide [Ref 6, Sec. 15.12(7)]. ■

Corollary 3.8: Let V_p, V_q be reflexive Banach spaces, and $V_{p\wedge q}|_\tau$ complete metrizable. Then $\langle V_{p\wedge q}, V_{\bar p \vee \bar q}\rangle = \langle V_p \cap V_q, V_{\bar p} + V_{\bar q}\rangle$ is a reflexive dual pair of Banach spaces. ■

4. INTERPLAY BETWEEN TOPOLOGICAL AND LATTICE PROPERTIES: THE GENERAL CASE

Let now $\langle V_p, V_{\bar p}\rangle$, $\langle V_q, V_{\bar q}\rangle$ be two arbitrary dual pairs of assaying subsets. What about the pair $\langle V_{p\wedge q}, V_{\bar p\vee \bar q}\rangle$?

We proceed exactly as in the Banach case, by constructing first the auxiliary pair $\langle V_p \cap V_q, V_{\bar p} + V_{\bar q}\rangle$ with the projective and inductive topologies, respectively.[11] The dual pair $\langle V_p \oplus V_q, V_{\bar p} \oplus V_{\bar q}\rangle$ is again reflexive. The subspace $V_{[p,q]} \equiv V_p \cap V_q$ is closed in $V_p \oplus V_q$, as before. With the induced topology, we denote it again by $(V_p \cap V_q)_{\text{proj}}$ for the same reason. Similarly, $(V_{\bar p} + V_{\bar q})_{\text{ind}}$ is the quotient of $V_{\bar p} \oplus V_{\bar q}$ by its closed subspace $V_{\bar p} \cap V_{\bar q}$ with the quotient, i.e., inductive, topology. Thus, we have the following continuous embeddings:

$$V_{p\wedge q}[\tau(V_p \cap V_q, V_{\bar p \vee \bar q})] \rightarrowtail (V_p \cap V_q)[\tau(V_p \cap V_q, V_{\bar p} + V_{\bar q})] \rightarrowtail (V_p \cap V_q)_{\text{proj}}, \tag{4.1a}$$

$$(V_{\bar p} + V_{\bar q})_{\text{ind}} \rightarrowtail V_{\bar p\vee \bar q}[\tau(V_{\bar p \vee \bar q}, V_p \cap V_q)]. \tag{4.1b}$$

We can now repeat Lemmas 3.3 and 3.4, but not 3.5 in general. So, we get a weaker result.

Proposition 4.1: Let $\langle V_p, V_{\bar p}\rangle$, $\langle V_q, V_{\bar q}\rangle$ be two arbitrary dual pairs of assaying subsets. Then, the following implications are true:

(PROJ) $\Rightarrow$ (ADD) $\Leftrightarrow$ (IND). ■

The difference with the Banach case lies in the fact that Mackey topologies are inherited by direct sums and by quotients, hence by inductive limits, but they are *not* inherited by subspaces (unless these are metrizable or dense), hence they are in general not inherited by projective limits. In other words, Lemma 3.5 holds in the general case only if $(V_p \cap V_q)_{\text{proj}}$ is metrizable, and then all three conditions (PROJ), (ADD), and (IND) are equivalent.

Next we assume the two pairs $\langle V_p, V_{\bar p}\rangle$, $\langle V_q, V_{\bar q}\rangle$ to be reflexive and look at the pair $\langle V_p \cap V_q, V_{\bar p} + V_{\bar q}\rangle$. $(V_p \cap V_q)_{\text{proj}}$ is semireflexive, as a closed subspace of the reflexive space $V_p \oplus V_q$ (Ref. 6, Sec. 23.3). Consequently its *strong* dual $(V_{\bar p} + V_{\bar q})_\beta$ is barreled (Appendix). Hence we get the following picture (all topologies refer to the dual pair we consider here):

$$(V_p \cap V_q)_\beta \rightarrowtail (V_p \cap V_q)_\tau \rightarrowtail (V_p \cap V_q)_{\text{proj}}, \tag{4.2a}$$

$$(V_{\bar p} + V_{\bar q})_\beta = (V_{\bar p} + V_{\bar q})_{\text{ind}} = (V_{\bar p} + V_{\bar q})_\tau. \tag{4.2b}$$

In general nothing more can be said: It is quite possible that $(V_p \cap V_q)_\tau$ be semireflexive but not reflexive, and $V_{\bar p} + V_{\bar q}$ not even semireflexive (Ref. 6, Sec. 23.6). We will exhibit an example below. Of course this cannot happen if V_p, V_q are Fréchet spaces.

As for the pair $\langle V_{p\wedge q}, V_{\bar p \vee \bar q}\rangle$, no general conclusion can be drawn, since the right-hand side depends explicitly (as a vector space on the compatibility and cannot be characterized *a priori* when condition (ADD) fails. This again suggests that the structure of PIP-space that we have used so far is too general.

What about the pair $\langle V^{\#}, V\rangle$ itself? Here one more piece of information is available, namely, a (generalized) condition (ADD), since we have

$$V^{\#} = \bigcap_{V_r \in \mathscr{I}} V_r, \quad V = \sum_{V_r \in \mathscr{I}} V_r. \tag{4.3}$$

Each V_r is, as usual, assumed to carry its Mackey topology $\tau(V_r, V_{\bar r})$. Then $V^{\#}$ carries three natural topologies: the strong topology $\beta(V^{\#}, V)$, the Mackey topology $\tau(V^{\#}, V)$ and the projective limit of all the $\tau(V_r, V_{\bar r})$ defined exactly as in the previous section. All three are in general distinct, but the last two give the same dual, namely, V itself, whereas $V^{\#}|_\beta$ could have a larger dual. V also carries three natural topologies, namely, $\beta(V, V^{\#})$, $\tau(V, V^{\#})$, and the inductive limit of the $\tau(V_r, V_{\bar r})$, but the last two always coincide since Mackey topologies are inherited by inductive limits. Thus the general picture is the following (with the same notation as above):

$$V^{\#}|_\beta \rightarrowtail V^{\#}|_\tau \rightarrowtail V^{\#}|_{\text{proj}}, \tag{4.4a}$$

$$V|_\beta \rightarrowtail V|_{\text{ind}} = V|_\tau. \tag{4.4b}$$

Of course, this by no means implies that the pair $\langle V^{\#}, V\rangle$ be reflexive, since $V|_\tau$ need not even be semireflexive. Similiar pathologies have been noticed by Friedrich and Lassner, in their study of rigged Hilbert spaces generated by algebras of unbounded operators.[12]

We will now conclude this section with an example, taken again from Ref. 6 (Sec. 30.4), which illustrates how bad the situation can be in general. Let, as usual, $\omega = \Pi_{i=1}^{\infty} \mathbb{C}_{(i)}$ be the space of all complex sequences, with the product topology, and $\varphi = \Sigma_{i=1}^{\infty} \mathbb{C}_{(i)}$ the space of all finite sequences, with the direct sum topology. Then $\langle \omega,\varphi \rangle$ is a reflexive dual pair, where ω is Fréchet, φ is complete (DF), and both are Montel spaces. Then one considers the space of arbitrary double sequences (a_{ij}), $\omega\omega = \Pi_{i=1}^{\infty} \omega_{(i)}$; one defines similarly $\omega\varphi = \Pi_{i=1}^{\infty} \varphi_{(i)}$, $\varphi\omega = \Sigma_{i=1}^{\infty} \omega_{(i)}$ and $\varphi\varphi = \Sigma_{i=1}^{\infty} \varphi_{(i)}$. Exactly as ω, the space $\omega\omega$ carries a natural PIP-space structure $[(a_{ij})\#(b_{ij}) \Leftrightarrow \Sigma_{i,j} |a_{ij} b_{ij}| < \infty]$, for which $(\omega\omega)^{\#} = \varphi\varphi$, $(\omega\varphi)^{\#} = \varphi\omega$, $(\varphi\omega)^{\#} = \omega\varphi$, $(\varphi\varphi)^{\#} = \omega\omega$. The intersection $\varphi\omega \cap \omega\varphi$ coincides with $\varphi\varphi$ as vector space, hence $(\varphi\omega\cap\omega\varphi)^{\#} = \omega\omega$. On the other hand, $(\varphi\omega)^{\#} + (\omega\varphi)^{\#} = \omega\varphi + \varphi\omega \neq \omega\omega$. Thus condition (ADD) fails. Condition (PROJ) fails also, as can be seen easily: on $\varphi\varphi$ the Mackey topology $\tau(\varphi\varphi,\omega\omega) = \beta(\varphi\varphi,\omega\omega)$ is strictly finer than the projective topology induced by $\varphi\omega \oplus \omega\varphi$. Finally it can be seen that $(\varphi\omega\cap\omega\varphi)_{\text{proj}}$ is semireflexive as closed subspace of the reflexive space $\varphi\omega \oplus \omega\varphi$, but not reflexive and not barreled (hence not Montel), whereas $(\omega\varphi + \varphi\omega)_{\text{ind}}$ is barreled but not semireflexive, *a fortiori* not Montel (Ref. 6, Sec. 31.5).

The lesson of the example is clear. The PIP-space ω (or $\omega\omega$, since they are isomorphic) contains a rich lattice $\mathscr{I}$ of Hilbert spaces, the family of all weighted l^2-spaces described in (III). Sec. 3.A, for which all conditions (PROJ), (ADD), (IND) hold. But it also contains bad assaying subsets $\varphi\omega$ and $\omega\varphi$ for which all three conditions fail and the regularity properties are lost. So why not exclude such pathological assaying subsets and concentrate instead on the nice rich lattice $\mathscr{I}$? The important point is that nothing is lost in this restriction, since the compatibility is fully recovered from $\mathscr{I}$, by the very definition of richness. In fact, concentrating on a fixed rich subset is exactly like describing a topology in terms of a fixed, convenient, basis of neighborhoods instead of considering *explicitly* arbitrary open sets. These considerations motivate the next section.

5. INDEXED PIP-SPACES

In III we have established the basic equivalence between a compatibility relation $\#$ on a vector space V and an involutive covering $\mathscr{I}$ of V, i.e., a rich involutive sublattice of $\mathscr{F}(V, \#)$. We have also exhibited several examples of such involutive coverings, consisting entirely of Hilbert spaces. We will now systematize this idea as a way of eliminating pathologies such a those of the example above. First, we define the basic concept.

Definition 5.1: By an *indexed partial inner product space* we mean a triple $(V, \mathscr{I}, \langle\cdot|\cdot\rangle)$, where V is a vector space, $\mathscr{I}$ an involutive covering of V and $\langle\cdot|\cdot\rangle$ a Hermitian form defined on those pairs of vectors of V that are compatible for the associated compatibility $\#_{\mathscr{I}}$.

Equivalently, an indexed PIP-space consists of a PIP-space $(V, \#, \langle\cdot|\cdot\rangle)$ together with a rich involutive sublattice $\mathscr{I}$ of $\mathscr{F}(V, \#)$. For convenience, we will denote the indexed PIP-space $(V, \mathscr{I}, \langle\cdot|\cdot\rangle)$ simply as V_I, where I is the isomorphy class of $\mathscr{I}$, i.e., $\mathscr{I}$ considered as an abstract partially ordered set (this notation was already introduced in I). Thus $\mathscr{I} = \{V_r, r\in I\}$.

The two concepts are closely related. Given an indexed PIP-space V_I, it defines a unique PIP-space, namely, $(V, \#_{\mathscr{I}}, \langle\cdot|\cdot\rangle)$, with $\mathscr{F}(V,\#_{\mathscr{I}})$ the lattice completion of $\mathscr{I}$. And a PIP-space is a particular indexed PIP-space for which $\mathscr{I}$ happens to be a *complete* involutive lattice.

Remark: (1) Although an indexed PIP-space generates a unique PIP-space, the converse is not true. Let $\#$ be a compatibility on V. Then each involutive sublattice of $\mathscr{F}(V,\#)$, if it is rich, defines an indexed PIP-space that generates the same PIP-space. Actually, the set of *all* involutive sublattices of $\mathscr{F}$ is itself a complete lattice for the following operations:

$$\mathscr{I}_1 \wedge \mathscr{I}_2 = \mathscr{I}_1 \cap \mathscr{I}_2,$$

$$\mathscr{I}_1 \vee \mathscr{I}_2 = \text{sublattice generated by } \mathscr{I}_1 \text{ and } \mathscr{I}_2.$$

If $\mathscr{I}_1$ and $\mathscr{I}_2$ are distinct and both rich, $\mathscr{I}_1 \wedge \mathscr{I}_2$ might be nonrich, even empty [see Remark (2) below], but $\mathscr{I}_1 \vee \mathscr{I}_2$ is rich *a fortiori*, hence corresponds to another indexed PIP-space.

(2) In general an involutive convering of V need not contain the extreme elements of $\mathscr{F}$, i.e., $V^{\#}$ and V. These, however, are always implicitly present, since they can be recovered from $\mathscr{I} = \{V_r, r\in I\}$:

$$V^{\#} = \bigcap_{r\in I} V_r, \quad V = \sum_{r\in I} V_r. \tag{5.1}$$

Thus it may happen that the intersection $\mathscr{I}_1 \cap \mathscr{I}_2$ of two rich sublattices $\mathscr{I}_1$ and $\mathscr{I}_2$ is empty, that is, they have no common element, although all elements of $\mathscr{I}_1$ and $\mathscr{I}_2$ contain $V^{\#}$ and verify Eq. (5.1).

The case of interest is when a given involutive covering $\mathscr{I}$ consists of topological spaces of the same type. Then relations (5.1) will imply better properties for $V^{\#}$ and V, equipped with their Mackey or projective, resp. inductive, topology; similarly, we will then be able to improve the results of Sec. 4 about a given couple V_p, V_q $(p, q\in I)$.

For that purpose it is useful to introduce some additional terminology.

Definitions 5.2: An indexed PIP-space V_I is said to be:

(i) *additive*, if condition (ADD) holds throughout I, $V_{p\vee q} = V_p + V_q$, $\forall p,q\in I$ [or $\mathscr{I}$ is a sublattice of $\mathscr{L}(V)$];

(ii) *projective* or *tight*, if condition (PROJ) holds throughout I,

$$V_{p\wedge q}|_\tau \simeq (V_p \cap V_q)_{\text{proj}}, \quad \forall p, q\in I;$$

(iii) *reflexive*, if $\langle V_p, V_{\bar{p}} \rangle$ is a reflexive dual pair, for every $p\in I$.

As we know already by Proposition 4.1, a projective indexed PIP-space is always additive. Next we draw some easy consequences of reflexivity.

Proposition 5.3: Let V_I be a reflexive indexed PIP-space. Then $V^{\#}$, with either its projective topology or its Mackey topology, is semireflexive and weakly quasicomplete, and V is barreled.

Proof: Every V_r is reflexive, *a fortiori* semireflexive. $V^{\#}|_{\text{proj}}$ is semireflexive as projective limit of semireflexive spaces (Ref. 6, Sec. 23.3), thus $V^{\#}|_\tau$ is also semireflexive; semireflexivity is equivalent to weak quasicompleteness and also to the (common) Mackey dual $V|_\tau$ being barreled: $V|_{\text{ind}} = V|_\tau = V|_\beta$. ■

Notice that reflexivity of V_I is not sufficient to imply that $\langle V^{\#}, V\rangle$ be a reflexive dual pair, for $V^{\#}|_\tau$ could still be semireflexive and nonreflexive. What is missing is $V^{\#}|_\tau$ being barreled, i.e., $V^{\#}|_\tau = V^{\#}|_\beta$. One way of avoiding the difficulty is to require $V^{\#}$ to be metrizable.

Proposition 5.4: Let V_I be reflexive and $V^{\#}|_{\text{proj}}$ be metrizable. Then $\langle V^{\#}, V\rangle$ is a reflexive dual pair, with $V^{\#}$ a Fréchet space.

Proof: The assumptions imply that $V^{\#}|_{\text{proj}}$ is both semireflexive and metrizable. That is possible only if it is complete, hence a Fréchet space. A semireflexive Fréchet space is necessarily reflexive, and so is its strong dual. ■

For the case described in Proposition 5.4, all three topologies of relations (4.4) coincide on $V^{\#}$, and similarly for V; in addition, both spaces are complete. However, from the fact that $\langle V^{\#}, V\rangle$ is a reflexive dual pair, we can conclude only that $V^{\#}|_\tau$ is barreled, i.e., $V^{\#}|_\tau = V^{\#}|_\beta$; the projective topology on $V^{\#}$ could still be coarser: Such was the case in the reflexive pair $\langle\varphi\varphi, \omega\omega\rangle$ in the example of Sec. 4. Also $V^{\#}|_\tau$ and $V|_\tau$ are then quasicomplete, but not necessarily complete. Fortunately, quasicompleteness of $V|_\tau$ is sufficient for the two arguments where a completeness result is needed: The existence of a central Hilbert space, to be discussed in Sec. 7, and the identification of the algebra of good operators with an algebra of unbounded operators in that Hilbert space, discussed in Ref. 13.

Thus reflexivity of an indexed PIP-space is not sufficient by itself. Actually the discussion of Secs. 2 and 3 shows that, among reflexive dual pairs, those consisting of reflexive Banach spaces (in particular, Hilbert spaces) are the only ones that are really compatible with the lattice structure. Hence it is worthwhile to give a separate name to the corresponding indexed PIP-spaces.

5.5 *Definitions*: A reflexive indexed PIP-space V_I is said to be of *type* (*B*) if every V_r, $r\in I$, is a reflexive Banach space; of *type* (*H*), if every V_r, $r\in I$, is a Hilbert space.

Applications show, in fact, that only these two types are useful in practice. They enjoy much better properties, as follows from the next proposition.

Proposition 5.6: Let V_I be an indexed PIP-space of type (B). Then:

(i) V_I is projective, hence additive;

(ii) If, in addition, I is countable, $V^{\#}|_{\text{proj}}$ is metrizable and $\langle V^{\#}, V\rangle$ is a reflexive dual pair.

Proof: Part (i) results from Propositions 3.6 and 3.7. As for (ii), it follows from the fact that the projective limit of a countable family of Banach spaces is a Fréchet space, and Proposition 5.4. ■

If V_I is an indexed PIP-space of type (H), with I countable, then V is a *countably Hilbert space*, in the terminology of Gel'fand–Vilenkin.[4] This particular case has been discussed in a joint work with W. Karwowski.[8] For V_I of type (H), with I arbitrary, $V^{\#}$ is a *quasi–Hilbert space*, in the sense of Hirschfeld.[14]

We will now conclude this section with a number of examples of indexed PIP-spaces of type (H) or (B).

A. Examples: Indexed PIP-spaces of type (H)

1. Chains of Hilbert spaces

The standard concept of a chain of Hilbert spaces, developed by Krein and Petunin[15] and by Palais,[16] obviously fits in here: A chain $\{V_k, k\in\mathbb{Z} \text{ or } \mathbb{R}\}$ of Hilbert spaces, such that $V_k \subset V_l$ for $k>l$ with continuous injection, and V_{-k} being the antidual of V_k. Typical examples of this structure are the Sobolev spaces, the space s' of slowly increasing sequences and the space $\mathscr{S}'$ of tempered distributions, for instance in Bargmann's realization $\mathscr{E}' = \lim\text{ind}_{\rho\in\mathbb{R}}\mathscr{F}^{\rho}$ [see III or Ref. 17]. In the continuous case ($k\in\mathbb{R}$), the projective topology on $V^{\#}$ may clearly be defined by a cofinal countable subset of $\mathbb{R}$, such as $\mathbb{Z}$, so that $V^{\#}|_{\text{proj}}$ is still metrizable and Proposition 5.4 applies. These chains, as well as the corresponding PIP-spaces, have been discussed in III, see Sec. 3.C.

2. Sequence spaces: Weighted l^2-spaces

A very simple example of indexed PIP-space of type (H) is given by the lattice of weighted l^2-spaces, described in III, Examples 3.A and 4.A. Let $V=\omega$, $\mathscr{I} = \{l^2(r)|r=(r_n)_{1<n<\infty}, r_n>0\}$. The lattice operations and the involution on $\mathscr{I}$ are defined by the relations

$$l^2(r)\wedge l^2(s) = l^2(p), \quad \text{with } p_n = \min\{r_n,s_n\},$$
$$l^2(r)\vee l^2(s) = l^2(q), \quad \text{with } q_n = \max\{r_n,s_n\},$$
$$[l^2(r)]^{\#} = l^2(\bar r), \quad \text{with } \bar r_n = r_n^{-1}.$$

All conditions are manifestly satisfied, and V_I is of type (H), in particular it is projective and additive. One has indeed

$$l^2(p) = l^2(r)\cap l^2(s),$$
$$l^2(q) = l^2(r) + l^2(s).$$

Straightforward estimates show the equivalence of the relevant norms:

$$\|x\|_p^2 \sim \|x\|_{r\wedge s}^2 = \|x\|_r^2 + \|x\|_s^2$$
$$\|x\|_q^2 \sim \|x\|_{r\vee s}^2 = \inf_{x=y+z}(\|y\|_r^2 + \|z\|_s^2)$$
$$= \sum_n |x|_n^2 (r_n^2 + s_n^2)^{-1}.$$

3. Spaces of locally integrable functions: Weighted L^2-spaces

A similar analysis holds in $L^1_{\text{loc}}(X,\mu)$ for the family $\mathscr{I} = \{L^2(r)|r\ \mu\text{-measurable and a.e. positive}, r \text{ and } r^{-1} \in L^1_{\text{loc}}(X,\mu)\}$ discussed in III, Example 4.B.

4. Nested Hilbert spaces

A nested Hilbert space, as defined by Grossmann,[5] is an indexed PIP-space of type (H), which satisfies two additional conditions: The partial inner product is positive definite, and there exists a central, self-dual Hilbert space. We will study these conditions in detail in Sec. 7 below.

B. Examples: Indexed PIP-spaces of type (B)

1. Chains of Banach spaces

As in Sec. 5.A.1 above, for the case of reflexive chains of Banach spaces, discussed at length in III, Sec. 3.C.

2. Normed Köthe spaces

A nontrivial (i.e., not a chain) example of indexed PIP-space of type (B) has been studied by Luxemburg and Zaanen[18] under the name of normed Köthe spaces. Since it is highly instructive, we feel it worthwhile to discuss it extensively. We will begin by repeating the basic definitions.

Let (X,μ) be a σ-finite measure space, M^+ the set of all measurable, non-negative functions on X, where two functions are identified if they differ at most on a μ-null set. A *function norm* is a mapping $\rho: M^+ \to \bar{\mathbb{R}}$ such that:

(i) $0 \leqslant \rho(f) \leqslant \infty$, $\forall f \in M^+$ and $\rho(f) = 0$ iff $f = 0$;

(ii) $\rho(f_1 + f_2) \leqslant \rho(f_1) + \rho(f_2)$, $\forall f_1, f_2 \in M^+$;

(iii) $\rho(af) = a\,\rho(f)$, $\forall f \in M^+$, $\forall a \geqslant 0$;

(iv) $f_1 \leqslant f_2 \Rightarrow \rho(f_1) \leqslant \rho(f_2)$, $\forall f_1, f_2 \in M^+$.

A function norm ρ is said to have the *Fatou property* iff $0 \leqslant f_1 \leqslant f_2 \leqslant \cdots, f_n \in M^+$ and $f_n \to f$ pointwise, implies $\rho(f_n) \to \rho(f)$.

Given a function norm ρ, it can be extended to all complex measurable functions on X by defining $\rho(f) = \rho(|f|)$. Denote by L_ρ the set of all measurable f such that $\rho(f) < \infty$. With the norm $\|f\| = \rho(f)$, L_ρ is a normed space and a subspace of the vector space V of all measurable, μ-a.e. finite, functions on X. Furthermore, if ρ has the Fatou property, L_ρ is complete, i.e., a Banach space. This is a generalisation of the spaces $L^p(X,\mu)$, which correspond to $\rho(f) = (\int |f|^p\, d\mu)^{1/p}$ for $1 \leqslant p < \infty$ and $\rho(f) = \sup|f|$ for $p = \infty$.

A function norm ρ is said to be *saturated* if, for any measurable set $E \subset X$ of positive measure, there exists a measurable subset $F \subset E$ such that $\mu(F) > 0$ and $\rho(\chi_F) < \infty$ (χ_F is the characteristic function of F).

Let ρ be a saturated function norm with the Fatou property. Define:

$$\rho'(f) = \sup\left\{ \int |fg|\, d\mu; \rho(g) \leqslant 1 \right\}. \tag{5.2}$$

Then ρ' is a saturated function norm with the Fatou property and $\rho'' \equiv (\rho')' = \rho$. Moreover, one has also:

$$\rho'(f) = \sup\left\{ \left| \int fg\, d\mu \right|; \rho(g) \leqslant 1 \right\}. \tag{5.3}$$

In our language these results can be restated as follows. The vector space V of all measurable, a.e.-finite functions on X carries a natural PIP-space structure, with compatibility

$$f \# g \Longleftrightarrow \int |fg|\, d\mu < \infty, \tag{5.4}$$

and partial inner product

$$\langle f | g \rangle = \int \bar{f} g\, d\mu. \tag{5.5}$$

V is clearly the largest space on which the pip (5.5) may be defined, but it is too large: indeed, $V^\# = \{0\}$ and the pip is degenerate. However, there are plenty of subspaces of V which are nondegenerate, such as $L^1_{\rm loc}$, $L^2_{\rm loc}$ or the space of Gould[19] to be defined below. Furthermore, for each ρ as above, L_ρ is a Banach space and $L_{\rho'} = (L_\rho)^\#$, i.e., each L_ρ is assaying. The pair $\langle L_\rho, L_{\rho'} \rangle$ is actually a dual pair, although $\langle V^\#, V \rangle$ is not.

However, $L_{\rho'}$ is in general only a closed subspace of the Banach dual $(L_\rho)'$, thus the Mackey topology $\tau(L_\rho, L_{\rho'})$ is coarser than the ρ–norm topology, which is $\tau(L_\rho, (L_\rho)')$. This defect can be remedied by further restricting ρ. A function norm ρ is called *absolutely continuous* if $\rho(f_n) \searrow 0$ for every sequence $f_n \in L_\rho$ such that $f_1 \geqslant f_2 \geqslant \cdots \searrow 0$ pointwise a.e. on X. For instance, the Lebesgue L^p-norm is absolutely continuous for $1 \leqslant p < \infty$ but the L^∞-norm is *not*! Also, even if ρ is absolutely continuous, ρ' need not be. Yet, this is the appropriate concept, in view of the following results:

(i) $L_{\rho'} = (L_\rho)'$ iff ρ is absolutely continuous;

(ii) L_ρ is reflexive iff ρ *and* ρ' are absolutely continuous and have the Fatou property.

Let us denote by J the set of saturated, absolutely continuous function norms ρ on X, with the Fatou property and such that ρ' is also absolutely continuous. Then, for every $\rho \in J$, $\langle L_\rho, L_{\rho'} \rangle$ is a reflexive dual pair of Banach assaying subspaces of $(V, \#, \langle\cdot|\cdot\rangle)$. All that remains to do in order to get an indexed PIP-space of type (B) is to restrict the pip to a nondegenerate subspace and perform the lattice construction of Sec. 3. Now the last point is in fact already done:

Lemma 5.7: The set J is an involutive lattice with respect to the partial order $\rho_1 \leqslant \rho_2$ iff $\rho_1(f) \leqslant \rho_2(f)$, $\forall f \in V$.

The lattice operations are the following:

$$(\rho_1 \vee \rho_2)(f) = \max\{\rho_1(f), \rho_2(f)\},$$

$$(\rho_1 \wedge \rho_2)(f) = \inf\{\rho_1(f_1) + \rho_2(f_2); f_1, f_2 \in M^+,\ f_1 + f_2 = |f|\},$$

$$\text{involution: } \rho \longleftrightarrow \rho'$$

Proof: Let $\rho_1, \rho_2 \in J$; so are ρ_1', ρ_2'. First we show that $\rho_1 \vee \rho_2$ is a saturated norm. It is obviously a norm. Suppose it is not saturated, i.e., there exists a measurable set E of positive measure, such that $(\rho_1 \vee \rho_2)(\chi_F) = \infty$ for every measurable subset $F \subset E$ of positive measure. Thus for every such $F \subset E$, $\rho_1(\chi_F) = \infty$ or $\rho_2(\chi_F) = \infty$. Since ρ_1 is saturated, there is a set $G \subset E$ such that $\rho_1(\chi_G) < \infty$ and for every $G_1 \subset G$, $\rho_1(\chi_{G_1}) < \infty$. This implies that $\rho_2(\chi_{G_1}) = \infty$ for every such G_1 and this is impossible for a saturated ρ_2.

Next is it always true (Ref. 18, Problem 71.2) that $(\rho_1 \wedge \rho_2)' = \rho_1' \vee \rho_2'$, although $\rho_1 \wedge \rho_2$ as defined could be only a function seminorm [i.e., $\rho(f) = 0 \not\Rightarrow f = 0$]. However, since $\rho_1' \vee \rho_2'$ is a saturated norm, it follows from this equality that $\rho_1 \wedge \rho_2$ is one also (Ref. 18, Theorem 71.4). Since ρ_1 and ρ_2 are absolutely continuous, so are all the others. Thus $L_{(\rho_1 \wedge \rho_2)'} = (L_{\rho_1 \wedge \rho_2})'$ is reflexive, and, therefore, $L_{\rho_1 \wedge \rho_2}$ is reflexive also, which implies that $\rho_1 \wedge \rho_2$ has the Fatou property. Since $\rho_1 \vee \rho_2$ also has the Fatou property, like any supremum (Ref. 18, Theorem 65.4), the proof is complete. ■

It is clear from the construction that we have recovered the general situation, for we have the relations

$$L_{\rho_1 \vee \rho_2} = (L_{\rho_1} \cap L_{\rho_2})_{\text{proj}}, \quad L_{\rho_1 \wedge \rho_2} = (L_{\rho_1} + L_{\rho_2})_{\text{ind}} .$$

It is interesting to notice that for any $\rho_1, \rho_2 \in J$, the ρ_1 norm and the ρ_2 norm are always consistent on $L_{\rho_1} \cap L_{\rho_2}$, since $\rho_1 \wedge \rho_2$ is a norm.

Finally, for any sublattice I of J, define the space $V_I \equiv \Sigma_{\rho \in I} L_\rho$; thus $V_I{}^{\#} = \cap_{\rho \in I} L_\rho$. Then we have:

Proposition 5.8: Let V be the vector space of all measurable, a.e. finite functions on the σ-finite measure space (X,μ). With the compatibility (5.4) and pip (5.5), V becomes a degenerate PIP-space. Denote by J the involutive lattice of all saturated, absolutely continuous function norms ρ on X, which have the Fatou property and are such that ρ' is also absolutely continuous. Let I be any involutive sublattice of J such that:

(i) $V_I \equiv \sum_{\rho \in I} L_\rho$ is an assaying subset of V;

(ii) $(V_I{}^{\#})^\perp = \{0\}$.

Then V_I with the PIP-space structure induced by V is a nondegenerate indexed PIP-space of type (B). ■

An interesting example is the following. Take (X,μ) to be $\mathbb{R}^n$ with Lebesgue measure, as in III Example 4.B, and put $V_I = L^1_{\text{loc}}$. Then $V_I{}^{\#} = L^\infty_{\text{comp}}$ and the two conditions above are verified. The corresponding set I is easily characterized: $\rho \in J$ belongs to I iff $L_\rho \subset L^1_{\text{loc}}$ with continuous injection. If we write $X = \cup_j K_j$, K_j compact, then L^1_{loc} can be represented as

$$L^1_{\text{loc}}(\mathbb{R}^n, d^n x) = \bigcap_{j=1}^{\infty} L^1(K_j, d^n x).$$

With the projective topology, L^1_{loc} becomes a Fréchet space, with dual

$$L^\infty_{\text{comp}}(\mathbb{R}^n, d^n x) = \bigcup_{j=1}^{\infty} L^\infty(K_j, d^n x).$$

Then $L_\rho \subset L^1_{\text{loc}}$ with continuous injection, iff ρ satisfies the following set of conditions:

$$\int_{K_j} |f| \, d^n x \leqslant c_j \, \rho(f) \quad \text{for each } j = 1,2,\cdots .$$

For instance, a weighted L^2-space $L^2(r)$ satisfies this condition if $r \in L^1_{\text{loc}}$, with $c_j = [\int_{K_j} r \, d^n x]^{1/2}$. It is thus clear that the family $\{L_\rho, \rho \in I\}$ is rich in L^1_{loc}, since it contains the rich subfamily $\{L^2(r), r \text{ and } r^{-1} \in L^1_{\text{loc}}\}$.

3. L^p-spaces

Another example of the construction given in Proposition 5.8 is the lattice generated by the spaces $L^p(X,\mu)$, $1 < p < \infty$. Indeed, each L^p norm is saturated and absolutely continuous and has the Fatou property.

Corollary 5.9: Let (X,μ) be a σ-finite measure space. Then the family $\{L^p(X,\mu), 1 < p < \infty\}$ generates an indexed PIP-space of type (B) with the compatibility (5.4) and the L^2 inner product. ■

Remember that if $\mu(X) = \infty$ and μ has no atoms, no two L^p spaces are contained in each other, but $L^p \cap L^r \subset L^q$ for all q such that $p \leqslant q \leqslant r$. Hence, we get a genuine lattice in that case.

What about the spaces L^1 and L^∞? They are not reflexive since the L^∞-norm is not absolutely continuous; hence they do not belong to any V_I. But if they are added by hand, an interesting result follows, which is due to Gould.[18,19] He introduces the following norm:

$$\rho_0(f) = \sup_E \left\{ \int_E |f| \, d\mu \, ; \mu(E) = 1 \right\}.$$

Then:

(i) ρ_0 has the Fatou propety (since it is a supremum of function norms that have it), hence L_{ρ_0} is a Banach space;

(ii) $L_{\rho_0} = (L^1 + L^\infty)_{\text{ind}}$, $(L_{\rho_0'} = L^1 \cap L^\infty)_{\text{proj}}$;

(iii) $\bigcup_{1 \leqslant p \leqslant \infty} L^p$ is properly included in L_{ρ_0}.

So, exactly as for the chain $\{L^p[0,1], 1 \leqslant p \leqslant \infty\}$ discussed in III, Example 3.B, the family $\{L^p(X,\mu), 1 \leqslant p \leqslant \infty\}$ generates a lattice with extreme elements L_{ρ_0} and $L_{\rho_0'}$. The completion of that lattice can easily be described along the same line, using again the results of Davis *et al.* quoted in III.

6. OPERATORS ON INDEXED PIP-SPACES

In II we have defined operators on PIP-spaces. That definition can be adapted in an obvious fashion to indexed PIP-spaces. For the convenience of the reader, we state the new definition in full.

Definition 6.1: Let V_I and Y_K be two nondegenerate indexed PIP-spaces, A a map from a subset $\mathscr{D}(A) \subset V$ into Y. We say that A is an *operator* if:

(i) $\mathscr{D}(A)$ is a nonempty union of elements of $\mathscr{I}$:

$$\mathscr{D}(A) = \bigcup_{r \in D(A)} V_r, \qquad D(A) \subset I.$$

(ii) For every $r \in D(A)$, there exists $b \in K$ such that the restriction of A to V_r is a continuous linear map into Y_b.

(iii) A has no proper extension satisfying (i) and (ii).

Exactly as before we denote by $Op(V_I, Y_K)$ the set of all such operators. For a given $A \in Op(V_I, Y_K)$, $J(A)$ is the set of all pairs $\{r,b\} \in I \times K$ such that A maps V_r linearly and continuously into Y_K. The domain of A is $D(A) = \{r \in I \mid \exists b \in K$ such that $\{r,b\} \in J(A)\}$, its range is $R(A) = \{b \in K \mid \exists r \in I$ such that $\{r,b\} \in J(A)\}$. For each $\{r,b\} \in J(A)$, the continuous linear map $A_{br}: V_r \to Y_b$ is the corresponding *representative* of A. Thus the whole machinery of representatives can be developed, as was done in II and previously by Grossmann[5] for nested Hilbert spaces. The only difference with II is that here the extreme spaces $V^{\#}$, Y are omitted, since they usually do not belong to $\mathscr{I}$, resp. $\mathscr{K}$. This does not change anything: since the embeddings $V^{\#}|_\tau \to V_r$ and $Y_b \to Y|_\tau$ are continuous for every $r \in I$, $b \in K$, the extreme representative $A: V^{\#} \to Y$ always exists.

Remark: The use of the set $J(A)$ is in fact much older, especially in the context of L^p spaces; a systematic presentation can be found in the monograph of Krasnoselski *et al.*,[20] where $J(A)$ is called the L-characteristic of A.

In this section we will study the lattice properties of the three sets $J(A)$, $D(A)$, $R(A)$, for a given $A \in Op(V_I, Y_K)$. First we consider $I \times K$. Given any two involutive lattices L and L', their Cartesian product carries a natural partial order:

$$\{x,x'\} \geqslant \{y,y'\} \quad \text{iff} \quad x \leqslant y \quad \text{and} \quad x' \leqslant y'.$$

With that order, $L \times L'$ is in fact an involutive lattice with respect to the following operations:

$$\{x,x'\} \wedge \{y,y'\} = \{x \vee y, x' \wedge y'\},$$
$$\{x,x'\} \vee \{y,y'\} = \{x \wedge y, x' \vee y'\},$$
$$\overline{\{x,x'\}} = \{\bar{x},\bar{x}'\}.$$

Thus given to indexed PIP-spaces V_I, Y_K, the product $I \times K$ is an involutive lattice. From the definition we conclude immediately:

Lemma 6.2: Let $A \in Op(V_I, Y_K)$. Then:

(i) $J(A)$ is a final subset of $I \times K$.

(ii) $D(A)$ is an initial subset of I.

(iii) $R(A)$ is a final subset of K. ■

It follows that $J(A)$ and $R(A)$ are always $\vee$-stable and directed to the right, whereas $D(A)$ is always $\wedge$-stable and directed to the left. This property of $J(A)$ is to be contrasted with the behavior of the set $J(f) \subset I$ for a given $f \in V$: $J(f) = \{r \in I \mid f \in V_r\}$. $J(f)$ is always a sublattice of I, in particular it is always $\wedge$-stable, whereas $J(A)$ is not. Indeed, let A map V_r into Y_b and V_q into Y_c continuously, that is $\{r,b\}$ and $\{q,c\}$ are in $J(A)$; this of course does *not* imply that A maps $V_{r\vee q}$ into $Y_{b\wedge c}$, which is $\{r,b\} \wedge \{q,c\} \in J(A)$. On the contrary, $\{r,b\}$ and $\{q,c\}$ in $J(A)$ implies $\{r \wedge q, b \vee c\} = \{r,b\} \vee \{q,c\} \in J(A)$. In fact $J(A)$ is *never* a sublattice of $I \times K$, even if the only assaying subsets are $V^{\#}$ and V. Indeed let A map continuosuly $V^{\#}$ into itself, and V into itself. Then *a fortiori*, it maps $V^{\#} \wedge V = V^{\#}$ into $V^{\#} \vee V = V$ but not the converse! Yet $J(A)$ characterizes the behavior of A, as $J(f)$ does for f: the bigger $J(A)$, the more regular the operator A.

For the domain $D(A)$ and the range $R(A)$, however, the situation can be improved. According to Definition 6.1, the operator A is defind on the set $\mathscr{D}(A) = \cup_{r \in D(A)} V_r$ and such a union of vector subspaces need not be a vector subspace of V. A sufficient condition is that $D(A)$ be directed to the right, *a fortiori* that $D(A)$ be $\vee$-stable. Indeed, let $f, g \in \mathscr{D}(A)$, with $f \in V_p$, $g \in V_q$, $p,q \in D(A)$. If $D(A)$ is directed to the right, p and q have a common successor $z \in D(A)$, i.e., $V_p \subseteq V_z$ and $V_q \subseteq V_z$. Hence $V_p + V_q \subseteq V_z$, or $\lambda f + \mu g \in V_z \subset \mathscr{D}(A)$, $\forall \lambda, \mu \in \mathbb{C}$.

It is of course desirable that *every* operator on the PIP-space have a vector subspace as domain of definition. A sufficient condition is found easily with the results of Sec. 5. Let $A \in Op(V_I, Y_K)$ and $\{p,a\}$, $\{q,b\} \in J(A)$. *A fortiori* $\{p, a \vee b\}$ and $\{q, a \vee b\}$ belong to $J(A)$, that is, A maps both V_p and V_q continuously into $Y_{a\vee b}$. Thus it can be extended by linearity to a *continuous* map from their inductive limit $(V_p + V_q)_{\rm ind}$ into $Y_{a\vee b}$. Since $(V_p + V_q)$ need not be assaying, we cannot conclude, unless of course $V_p + V_q = V_{p\vee q}$. In a similar fashion, $\{p \wedge q, a\}$ and $\{p \wedge q, b\}$ belong to $J(A)$. Thus A maps $V_{p\wedge q}$ continuously into Y_a and Y_b, hence also into their projective limit $(Y_a \cap Y_b)_{\rm proj}$. Again, since the Mackey topology on $Y_a \cap Y_b$ might be strictly finer than the projective topology, this does not imply that A: $V_{p\wedge q} \to Y_{a\wedge b}$ is continuous, unless the two topologies coincide on $Y_a \cap Y_b$. So we have proved:

Proposition 6.3: Let V_I, Y_K be two indexed PIP-spaces, A any operator from V_I into Y_K. Then:

(i) If V_I is additive, the domain $D(A)$ is a sublattice of I; in particular, the set $\mathscr{D}(A)$ is a vector subspace of V.

(ii) If Y_K is projective, the range $R(A)$ is a sublattice of K. ■

Corollary 6.4: Let V_I be of type (B) or (H), and Y_K arbitrary. Then the domain of definition $\mathscr{D}(A)$ of any operator A from V_I into Y_K is a vector subspace of V. ■

Remark: Let again $\{p,a\}$, $\{q,b\} \in J(A)$. Then V_I additive implies $\{p \vee q, a \vee b\} \in J(A)$ and Y_K projective implies $\{p \wedge q, a \wedge b\} \in J(A)$, but this does not mean that $J(A)$ is a sublattice of $I \times K$. Assume now that all four pairs $\{p,a\}$, $\{p,b\}$, $\{q,a\}$, $\{q,b\}$ belong to $J(A)$. If V_I is additive and Y_K is projective, it follows that A maps $V_{p\vee q} = (V_p + V_q)_{\rm ind}$ continuously into $Y_{a\wedge b} = (Y_a \cap Y_b)_{\rm proj}$, i.e., $\{p \vee q, a \wedge b\} \in J(A)$. Of course, this still does not imply that $J(A)$ be $\wedge$-stable and a sublattice. What this argument does give, in fact, is a three-line proof of Proposition 4.4 of Ref. 5!

7. THE CENTRAL HILBERT SPACE

In most examples studied so far, the (indexed) PIP-space V has two additional properties:

(i) The pip $\langle\cdot|\cdot\rangle$ is positive definite, in the sense that $\langle f|f\rangle > 0$ for every nonzero, self-compatible $f \in V$; equivalently the identity operator $\mathbb{1}$ is positive.

(ii) There exists a unique, self-dual assaying subset $V_0 = V_{\bar{0}}$, which is a Hilbert space, namely the completion of $V^{\#}$ in the norm $\|f\| = \langle f|f\rangle^{1/2}$ (which we will call the pip-norm).

In this section, we will assume the positivity condition (i) and see to what extent it implies (ii). First, it follows from (i) that $(V^{\#}, \|\cdot\|)$ is a pre-Hilbert space (the same is true for every assaying subset V_r such that $V_r \subseteq V_{\bar{r}}$). As such it admits a unique completion $\mathscr{H}$, but *a priori* there is no reason why $\mathscr{H}$ could be identified with a subspace of V, and even less that it should be an assaying subset. Conditions ensuring these two properties will be given below, for a general (indexed) PIP-space. These results generalize those obtained previously[8] for the case where V is a countably Hilbert space, in the sense of Gel'fand.[4]

As in the latter case, the argument is based on a result of L. Schwartz,[9] namely, Hilbertian subspaces of V correspond one-to-one to the so-called positive kernels. In this terminology, a *kernel* is a weakly continuous (hence Mackey continuous) linear map from $V^{\#}$ into V, i.e., the restriction to $V^{\#}$ of an operator on V. A *Hermitian kernel* is the restriction of a symmetric operator ($A = A^*$). A *positive kernel* corresponds to a positive operator. The first result concerns the existence of the central Hilbert space $\mathscr{H}$.

Proposition 7.1: Let V be an (indexed) PIP-space with positive definite pip. Assume there exists an assaying subset

V_r such that $V_r \subseteq V_{\bar{r}}$ and $V_{\bar{r}}$ is quasicomplete in its Mackey topology. Then, the completion $\mathscr{H}$ of $V^{\#}$ in the pip-norm is a dense subspace of V, and we have

$$V^{\#} \subseteq V_r \subseteq \mathscr{H} \subseteq V_{\bar{r}} \subseteq V.$$

Proof: The spaces V_r and $V_{\bar{r}}$ are dual to each other and $V_{\bar{r}}$ is quasicomplete. Hence the pair $\langle V_r, V_{\bar{r}} \rangle$ satisfies the general conditions of Ref. 9. By assumption the identity map $1: V_r \to V_{\bar{r}}$ is a positive kernel. Hence, by Proposition 10 and Sec. 10 of Ref. 9, it corresponds to a unique Hilbertian subspace $\mathscr{H}$ of $V_{\bar{r}}$, which is the completion of the pre-Hilbert space $(V_r, \|\cdot\|)$ and is dense in $V_{\bar{r}}$. *A fortiori*, we get $V^{\#} \subseteq V_r \subseteq \mathscr{H} \subseteq V_{\bar{r}} \subseteq V$, and the statement follows. ■

Remark: Proposition 7.1 certainly holds if $V_{\bar{r}}$ is a Hilbert space; therefore, condition (I_4) of Ref. 5 is always satisfied.

Before going further let us mention a few cases where the proposition applies.

(1) $\langle V^{\#}, V \rangle$ is a reflexive dual pair, which implies that V is Mackey-quasicomplete. Typical examples are PIP-spaces of distributions, such as $V = \mathscr{S}'$ or $\mathscr{D}'$. Notice that under the same condition of Mackey-quasicompleteness of V, the *-algebra of good operators on V (a *good operator* is an operator that maps both $V^{\#}$ and V continuously into themselves) is isomorphic to the canonical *-algebra of unbounded operators $\mathscr{L}^{+}(V^{\#})$, as defined by Lassner (see Refs. 12 and 13).

(2) V_I is reflexive, in particular of type (B) or (H): given any assaying subset V_p $(p \in I)$, the pair $\langle V_{p\wedge\bar{p}}, V_{p\vee\bar{p}} \rangle$ is reflexive and verifies the conditions of Proposition 7.1. Plenty of examples have been given.

(3) The following counterexample is instructive. Let (X,μ) be a measure space such that $\mu(X) = \infty$ and μ has no atoms, as in Example 5.B.3. The spaces L^p are not comparable to each other, but the family $\{L^1 \cap L^p, 1 \leqslant p \leqslant \infty\}$ is a chain. With the compatibility inherited from L^2, we get $(L^1 \cap L^p)^{\#} = L^1 \cap L^q$ and $(L^1 \cap L^2)^{\#} = L^1 \cap L^2$. However none of them is Mackey-quasicomplete, nor even sequentially complete. Indeed, on $L^1 \cap L^p$ the topology $\tau(L^1 \cap L^p, L^1 \cap L^q)$ is strictly coarser than $\tau(L^1 \cap L^p, L^q)$; but this is just the L^p-norm topology, for which $L^1 \cap L^p$ is not (sequentially) complete, its completion being L^p. Let $(f_n) \in L^1 \cap L^p$ be a sequence that converges, in the L^p norm, to an element $f \in L^p$ which does not belong to $L^1 \cap L^p$. Thus (f_n) converges also for the coarser topology $\tau(L^1 \cap L^p, L^1 \cap L^q)$. Now, if $L^1 \cap L^p$ were sequentially complete for that topology, the limit f of (f_n) would belong to $L^1 \cap L^p$, contrary to the assumption. Hence $(L^1 \cap L^p)$ $[\tau(L^1 \cap L^p, L^1 \cap L^q)]$ is not sequentially complete, *a fortiori* not quasicomplete or complete. Indeed, the completion of $V^{\#} = L^1 \cap L^\infty$ in the pip-norm, i.e., the L^2-norm, is L^2, which is not contained in $V = L^1$.

Assume now $\mathscr{H}$ exists as a dense subspace of V. When is it assaying? For answering that question, it is useful to consider, as was done in Ref. 8, the set Λ of all self-compatible vectors, introduced by Popowicz[21]:

$$\Lambda = \{ f \in V \mid f \# f \}.$$

Let now $f \in \Lambda$. This means, there exists $r \in I$ such that $f \in V_r \cap V_{\bar{r}} = V_{r\wedge\bar{r}}$. Since $V_{r\wedge\bar{r}}$ is a pre-Hilbert space with respect to the pip, we have $V_{r\wedge\bar{r}} \subseteq \mathscr{H}$, and therefore $\Lambda \subseteq \mathscr{H}$. To go further, we need one more assumption.

Proposition 7.2: Let V_I be a positive indexed PIP-space, Λ the subset of self-compatible vectors, $\mathscr{H}$ the completion of $V^{\#}$ in the pip-norm. Assume that, for every $r \in I$, the assaying subset $V_{r\vee\bar{r}}$ is quasicomplete in its Mackey topology. Then one has

$$\Lambda \subseteq \mathscr{H} \subseteq \Lambda^{\#}. \tag{7.1}$$

Proof: From the discussion above, we have (as sets)

$$\Lambda = \bigcup_{r \in I} V_{r\wedge\bar{r}}$$

and therefore

$$\Lambda^{\#} = \bigcap_{r \in I} V_{r\vee\bar{r}}.$$

Now, for every $r \in I$, one has $V_{r\wedge\bar{r}} \subseteq V_{r\vee\bar{r}}$ and $V_{r\vee\bar{r}}$ is Mackey-quasicomplete by assumption. Hence, by Proposition 7.1, we have

$$V_{r\wedge\bar{r}} \subseteq \mathscr{H} \subseteq V_{r\vee\bar{r}}.$$

Thus we get

$$\Lambda = \bigcup_{r \in I} V_{r\wedge\bar{r}} \subseteq \mathscr{H} \subseteq \Lambda^{\#} = \bigcap_{r \in I} V_{r\vee\bar{r}}. \qquad ■$$

Corollary 7.3: Under the assumptions of Proposition 7.2, Λ is an assaying subset: $\Lambda = \Lambda^{\#\#} \in \mathscr{F}(V,\#)$.

Proof: Let S be any family of self-compatible vectors, i.e., $S \subseteq S^{\#}$. Then $S \subseteq S^{\#\#} \subseteq S^{\#}$. This implies $S^{\#\#} \subseteq \Lambda$. The result follows by taking $S = \Lambda$. ■

The corollary implies in particular that Λ is a vector subspace of V, and so $\Lambda = \Sigma_{r \in I} V_{r\wedge\bar{r}}$. Also any self-compatible assaying subspace, no matter whether it belongs to $\mathscr{I}$ or not, is a pre-Hilbert space with respect to the pip $\langle \cdot | \cdot \rangle$. Thus Λ is also the largest pre-Hilbert subspace of V.

Proposition 7.4: Let V_I, Λ, $\mathscr{H}$ be as in Proposition 7.2. Then the following inclusions hold:

$$\Lambda = \mathscr{H}^{\#} \subseteq \mathscr{H} \subseteq \Lambda^{\#} = \mathscr{H}^{\#\#}. \tag{7.2}$$

Proof: From Proposition 7.2 and the involution, we get $\Lambda = \Lambda^{\#\#} \subseteq \mathscr{H}^{\#}$. Next we show: $\mathscr{H}^{\#} \subseteq \mathscr{H}$. Indeed $\langle \mathscr{H}^{\#}, \mathscr{H} \rangle$ is a dual pair. The Mackey topology $\tau(\mathscr{H}^{\#}, \mathscr{H})$ induces on the dense subspace $V^{\#}$ the topology $\tau(V^{\#}, \mathscr{H})$, which coincides with the one given by the pip-norm. Since $\mathscr{H}$ is the completion of $V^{\#}$ in the latter topology, it follows that $\mathscr{H}^{\#} \subseteq \mathscr{H}$. So $\mathscr{H}^{\#}$ is a self-compatible assaying subset and therefore $\mathscr{H}^{\#} \subseteq \Lambda$. The statement follows. ■

Here again, the result applies if V_I is reflexive, for then every $V_{r\vee\bar{r}}$ is reflexive, hence τ-quasicomplete. It applies in particular if V_I is of type (B) or (H).

If all three spaces in Eq. (7.2) coincide, we will say that V_I has a *central Hilbert space*. This happens if any of the following equivalent conditions is satisfied:

(i) Λ is self-dual: $\Lambda = \Lambda^{\#}$;

(ii) I contains a self-dual element o, i.e., $V_o = V_o^{\#}$ (for then $V_o \subseteq \Lambda \subseteq \Lambda^{\#} \subseteq V_o^{\#} = V_o$);

(iii) Λ is complete in the pip-norm.

There are three main cases of interest for applications:

(1) *Type* (H): In this case, V_I is a *nested Hilbert space*, in the sense of Grossmann.[5] Several examples have been given in I–III, and in Sec. 5.A. In each case, the central Hilbert space $\mathscr{H}$ is the natural one, which, so to speak, generates the PIP-space by extension of the natural inner product to a larger domain, e.g., l^2 for $\{l^2(r)\}$, L^2 for $\{L^2(r)\}$, etc.

(2) *Type* (B): The structure obtained consists of a lattice of reflexive Banach spaces, the duality $V_r \longleftrightarrow V_{\bar r}$ being taken with respect to the inner product of $\mathscr{H}$. Examples are $\mathscr{H} = L^2[0,1]$ for the chain $\{L^p[0,1]\}$ or $\mathscr{H} = L^2(X,\mu)$ for the lattice of Köthe normed spaces (cf. Sec. 5.B). This structure has been called a *Dirac space* in Ref. 22.

(3) *Reflexive indexed PIP-spaces*: The cases of interest are essentially those spaces for which $\mathscr{I}$ is a lattice of type (B) or (H) together with the extreme elements $V^{\#}$ and V, where $\langle V^{\#}, V\rangle$ is a reflexive dual pair. Such are, for instance, rigged Hilbert spaces,[4] or more generally, PIP-spaces of distributions, like $\mathscr{S}'$ or Bargmann's space $\mathscr{E}'$ discussed in III, Sec. 3.C.

However, in the general case, the inclusions in Eq. (7.2) are strict and $\mathscr{H}$ is not assaying. Examples can be given readily, e.g., simply by omitting the central element in any lattice I of the previous type, but they tend to be artificial.

In fact, the missing central Hilbert space may often be reconstructed by hand. This will be done below for the case where V_I is additive. The key is the following observation.

$\mathscr{H}^{\#} \subsetneq \mathscr{H}$ means the given compatibility is too coarse, it does not admit sufficiently many compatible pairs. Thus adding $\mathscr{H}$ to the assaying subsets amounts to refining the compatibility (in the sense of III).

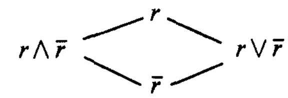
Proposition 7.5: Let V_I be a positive PIP-space for which every $V_{r\vee\bar r}$, $r \in I$ is quasicomplete but $\mathscr{H}$ is not assaying. Assume V_I is additive. Then V carries a finer PIP-space structure for which $\mathscr{H}$ is assaying and thus a central Hilbert space.

Proof: V_I additive means than $\mathscr{I}$ is an involutive sublattice of the lattice $\mathscr{L}(V)$ of all vector subspaces of V. Define $\mathscr{I}_1$ as the sublattice of $\mathscr{L}(V)$ generated by $\mathscr{I}$ and $\mathscr{H}$. $\mathscr{I}_1$ consists of elements of $\mathscr{I}$, plus $\mathscr{H}$ itself and additional elements of the form $V_p \cap \mathscr{H}$, $V_p + \mathscr{H}$, $(V_p + \mathscr{H}) \cap V_q$, etc. An involution $\#_1$ can be defined on $\mathscr{I}_1$ as follows:

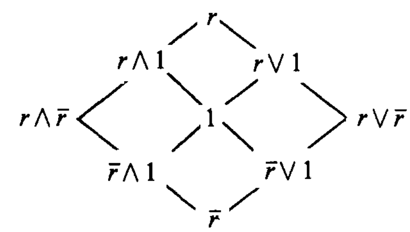
- for $V_p \in \mathscr{I}$, $\quad V_p^{\#_1} = V_p^{\#} = V_{\bar p}$,
- $\mathscr{H}^{\#_1} = \mathscr{H}$,
- $(V_p \cap \mathscr{H})^{\#_1} = V_{\bar p} + \mathscr{H}$,
- $(V_p + \mathscr{H})^{\#_1} = V_{\bar p} \cap \mathscr{H}$, and so on.

By this construction, $\#_1$ is a lattice dual isomorphism on $\mathscr{I}_1$ and so $\mathscr{I}_1$ is an involutive sublattice of $\mathscr{L}(V)$, and an involutive covering of V. Thus $\#_1$ defines a linear compatibility on V, finer than $\#$, in the technical sense of III. Obviously the new indexed PIP-space V_{I_1} has $\mathscr{H}$ as central Hilbert space. ■

It is instructive to consider a simple example of the construction just described. Let $l^2(r)$ be a weighted l^2-space (see Sec. 5.A.2, and $\mathscr{I}$ be the four-element involutive lattice it generates. Denoting each space $l^2(p)$ simply by its weight p, we obtain the following picture (smaller spaces stand on the left):



Following the general construction, the space $\mathscr{H} = l^2 \equiv l^2(1)$ is obtained as the completion of $l^2(r \wedge \bar r)$ in the pip-norm. The lattice $\mathscr{I}_1$ generated by $\mathscr{I}$ and l^2 has nine elements, and is described by the following diagram:



This is indeed an involutive lattice. One verifies easily, for instance, the following relations: $r \wedge (\bar r \vee 1) = r \wedge 1$ and $\overline{r \wedge (\bar r \vee 1)} = \bar r \vee (r \wedge 1) = \bar r \vee 1$.

The crucial fact for the construction of Proposition 7.5 is additivity: Since $\mathscr{I}$ and $\mathscr{H}$ are then embedded in the lattice $\mathscr{L}(V)$, the supremum in $\mathscr{I}_1$ is defined independently of the involution. If additivity fails, there is no obvious way of enlarging $\mathscr{I}$, since both the supremum and the involution have to be defined on $\mathscr{I}_1$. This is consistent with the discussion given at the end of III: compatibilities may always be coarsened, but not always refined.

As a final remark, we notice that the construction always works for indexed PIP-spaces of type (B) or (H): Each of these can be embedded canonically in a Dirac space, resp. a nested Hilbert space [in other words, condition (I_3) of Ref. 5 may be satisfied automatically]. Thus one may as well assume the existence of the central Hilbert space $\mathscr{H} = \mathscr{H}^{\#}$ from the beginning. Here again we see how simple this class of indexed PIP-spaces is.

ACKNOWLEDGMENTS

It is a pleasure to thank Alex Grossmann and Françoise Debacker-Mathot for many conversations on the questions discussed here. Also Brian Davies deserves credit for asking the right questions, thus stimulating the work reported here.

APPENDIX: SOME FACTS ABOUT LOCALLY CONVEX SPACES

In this appendix, as well as throughout the paper, we are concerned with locally convex topological vector spaces (LCS), i.e., topological vector spaces (TVS) which have a base of neighborhoods of zero consisting of convex sets, or equivalently, spaces with a topology that can be defined in terms of a family of seminorms. Our reference is the textbook of Köthe,[6] except for the notation of the different topologies, where we follow Schaefer.[23]

1. *Completeness*: A LCS $E[T]$ is *complete* if every

Cauchy net has a limit in E; it is *quasicomplete* if every closed bounded set in $E[T]$ is complete; it is *sequentially complete* if every Cauchy sequence has a limit in E.

Of course, completeness $\Rightarrow$ quasicompleteness $\Rightarrow$ sequential completeness, and the three notions coincide for metrizable spaces, i.e., Banach or Fréchet spaces.

2. *Dual pairs and canonical topologies*: Two vector spaces E, F form a *dual pair* $\langle E,F\rangle$ if there is a bilinear form $\langle\cdot|\cdot\rangle$ on $E\times F$, separating in both arguments: $\langle e|f\rangle=0$, $\forall f\in F$ implies $e=0$, $\langle e|f\rangle=0$. $\forall e\in E$, implies $f=0$. For any LCS E, with dual E', $\langle E,E'\rangle$ is a dual pair.

Given the dual pair $\langle E,F\rangle$, the *weak topology* $\sigma(E,F)$ is the coarsest topology on E for which the linear forms $e\rightarrow\langle e|f\rangle, f\in F$, are continuous, and in fact, for which the dual of E is F. It is locally convex and Hausdorff.

A basis of neighborhoods of zero consists of the sets S^0, where S runs over all finite subsets of F, and $S^0=\{e\in E\mid |\langle e|f\rangle|\leqslant 1, \forall f\in S\}$ is the absolute polar of $S\subset F$.

The weak topology on F, $\sigma(F,E)$, is defined similarly.

The *Mackey topology* $\tau(E,F)$ can be defined as the finest topology on E such that the dual is F (its existence is the content of the Mackey–Arens theorem); a basis of neighborhoods of zero is given by the sets T^0, where T runs over all absolutely convex, $\sigma(F,E)$-compact subsets of F.

The *strong topology* $\beta(E,F)$ is defined by the basis of neighborhoods of zero $\{U^0\}$ where U runs over all absolutely convex $\sigma(F,E)$-closed and bounded subsets of F.

A topology $T(E)$ on E is called a *topology of the dual pair* $\langle E,F\rangle$ if the dual of $E[T(E)]$ is F. Then one has, for any topology $T(E)$ of the dual pair:

$$\sigma(E,F)<T(E)<\tau(E,F)<\beta(E,F).$$

If we start with a given topology $T(E)$ on E we have the same inclusions with $F=E'$.

In a dual pair $\langle E,F\rangle$, several classes of subsets of E depend only on the dual pair and *not* on the topology of E, i.e., they coincide for all topologies of the dual pair. Such are: closed subspaces, convex closed subsets, dense and total subsets, bounded subsets.

A LCS $E[T]$ is *barreled* if $T(E)=\tau(E,E')=\beta(E,E')$. A metrizable LCS always carries its Mackey topology, $T(E)=\tau(E,E')$, but need not be barreled. A complete metrizable LCS, i.e., a Banach or a Fréchet space, is always barreled.

A barreled LCS $E[T]$ is called a *Montel space* if every bounded subset of E is relatively compact. A Montel space is necessarily quasicomplete and reflexive (see below), and its strong dual is also a Montel space. An infinite dimensional Banach space cannot be Montel. Typical examples are ω, φ, $\mathscr{S}(\mathbb{R}^n)$, $\mathscr{S}'(\mathbb{R}^n)$.

3. *Reflexivity*: Given an LCS E, the canonical topologies $\sigma(E',E)$, $\tau(E',E)$, $\beta(E',E)$ are defined in the same way; thus, with the notation of Sec. 4:

$$E'|_\beta \rightarrow E'|_\tau \rightarrow E'|_\sigma.$$

By definition $(E'|_\tau)'=E$, but the dual of the strong dual, called the *bidual*, $E''=(E'|_\beta)'$ may be strictly larger than E. A LCS E is called *semireflexive* if E'' coincides with E as a vector space. E is called *reflexive* if, in addition, the strong bidual $E''[\beta(E'',E')]\equiv(E'|_\beta)'|_\beta$ coincides with E as a TVS.

The two notions are different in general, but they coincide for Fréchet spaces. In fact, a Fréchet or Banach space is reflexive iff it is semireflexive, iff it is weakly quasicomplete, or iff its strong dual is reflexive. Notice that an incomplete normed or metrizable space can never be reflexive.

A dual pair $\langle E,F\rangle$ is *reflexive* if each space is the strong dual of the other: $(E|_\beta)'=F$, $(F|_\beta)'=E$. Equivalently, if $E|_\tau$ and $F|_\tau$ are both semireflexive, or if they are both barreled: $\tau(E,F)=\beta(E,F)$ and $\tau(F,E)=\beta(F,E)$. In a reflexive pair, both spaces are reflexive and quasicomplete for their weak and their Mackey (= strong) topology.

4. *Projective limits*: Let be given a vector space E, a family $\{E_\alpha\}$ of LCS and maps $i_\alpha:E\rightarrow E_\alpha$ such that for every nonzero $x\in E$ there is some α with $i_\alpha(x)\neq 0$. Then there is a coarsest topology on E that makes all the maps i_α continuous; it is called the *projective topology* and E with this topology, $E_{\rm proj}$, is called the *projective limit* of $\{E_\alpha\}$ with respect to the maps i_α. The projective limit is said to be *reduced* if $i_\alpha(E)$ is dense in E_α for each α (this can always be achieved without restriction of generality). The following properties are useful:

- $E_{\rm proj}$ is complete (resp. quasicomplete, sequentially complete) if every E_α is.
- Given any LCS Y, a linear map $t:Y\rightarrow E_{\rm proj}$ is continuous iff each composed map $t_\alpha=i_\alpha\circ t: Y\rightarrow E_\alpha$ is continuous.

The following examples are important:

(i) Let E be a LCS, H a subspace of E. The *subspace topology* on H is the projective topology with respect to the embedding $i:H\rightarrow E$.

(i) Let $\{E_\alpha\}$ be as above and $E=\Pi E_\alpha$. *The product topology* on E is the projective topology with respect to the projection maps $p_\alpha:E\rightarrow E_\alpha$.

(iii) In a general projective limit, $E_{\rm proj}$ is isomorphic to a closed subspace of the product ΠE_α. In the case considered in this paper $\{E_\alpha\}$ is a family of vector subspaces of a given vector space V, each of which is itself a LCS. Then $E_{\rm proj}$ $=$ proj lim E_α is the subspace $\cap E_\alpha$ with the projective topology. $E_{\rm proj}$ is metrizable iff the family $\{E_\alpha\}$ contains a cofinal countable subfamily of metrizable spaces (this makes sense since the subspaces are partially ordered by inclusion).

5. *Inductive limits*: Let be given a vector space F, a family $\{F_\gamma\}$ of LCS and maps $j_\gamma:F_\gamma\rightarrow F$. Then there is a finest topology on F that makes all the maps j_γ continuous; it is called the *inductive topology* and F with this topology, denoted $F_{\rm ind}$, is called the *inductive limit* of $\{F_\gamma\}$ with respect to the maps j_γ. Given any LCS, Y, a linear map $t:F_{\rm ind}\rightarrow Y$ is continuous iff each composed map $t_\gamma=t\circ j_\gamma:F_\gamma\rightarrow Y$ is continuous. Again three cases are worth mentioning.

(i) If E is a LCS, H a closed subspace, the *quotient topology* on E/H is the inductive topology with respect to the canonical surjection $\pi:E\rightarrow E/H$.

(ii) For any family $\{F_\gamma\}$, let $F=\Sigma F_\gamma$; the *direct sum topology* on F is the inductive topology with respect to the embeddings $j_\gamma:F_\gamma\rightarrow F$.

(iii) For a general inductive limit, $F_{\rm ind}$ is isomorphic to a quotient of ΣF_γ.

6. *Duality and hereditary properties*: Let E be a LCS, H a closed subspace. Let $H^\perp \subset E'$ be the orthogonal space of H: $H^\perp = \{ f \in E' | \langle f | h \rangle = 0, \forall h \in H \}$. $H^\perp$ is a closed subspace of E'. Then the dual of H is E'/H', and the dual of E/H is $H^\perp$.

As for canonical topologies, the hereditary properties are the following:

- *The Mackey topology* is inherited by quotients, but not by closed subspaces in general:

$$\tau(E,E') \upharpoonright_{E/H} = \tau(E/H, H^\perp), \tau(E,E') \upharpoonright_H < \tau(H, E'/H^\perp).$$

we do get equality for subspaces in two cases: if $\tau(E,E') \upharpoonright_H$ is metrizable, or if H is a dense subspace (hence not closed).

- *The weak topology* is inherited both by quotients and closed subspaces, whereas the *strong topology* is inherited by neither of them, in general.

Direct sums and products are dual to each other:

$$\Big(\prod E_\alpha\Big)'_{\text{proj}} = \sum E'_\alpha,$$

$$\Big(\sum F_\gamma\Big)'_{\text{ind}} = \prod F'_\gamma,$$

and Mackey topologies go through:

$$\Big(\prod E_\alpha\Big)\Big|_\tau = \Big(\prod E_\alpha \Big|_\tau\Big)_{\text{proj}}, \quad \Big(\sum F_\gamma\Big)\Big|_\tau = \Big(\sum F_\gamma \Big|_\tau\Big)_{\text{ind}}.$$

Reduced projective limits and inductive limits are also dual to each other:

$$(\text{proj lim } E_\alpha)' = \text{ind lim } E'_\alpha \quad \text{(if the lhs is reduced).}$$

$$(\text{ind lim } F_\gamma)' = \text{proj lim } F'_\gamma$$

Combining all the above results, we get finally, for the Mackey topologies:

$$\text{proj lim}(E_\alpha|_\tau) < (\text{proj lim } E_\alpha)|_\tau \quad \text{(equality if lhs is metrizable).}$$

$$\text{ind lim } (F_\gamma|_\tau) = (\text{ind lim } F_\gamma)|_\tau$$

Weak topologies go through projective limits only, and there is no general result for strong topologies.

[1] J. -P. Antoine and A. Grossmann, J. Funct. Anal. **23**, 369–78 (1976).
[2] J. -P. Antoine and A. Grossmann, J. Funct. Anal. **23**, 379–91 (1976).
[3] J. -P. Antoine, J. Math. Phys. **21**, 268–79 (1980).
[4] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. IV.
[5] A. Grossmann, Commun. Math. Phys. **2**, 1–30 (1966).
[6] G. Köthe, *Topological Vector Spaces. I.* (Springer-Verlag, Berlin, 1969).
[7] See, e.g., H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators* (North-Holland, Amsterdam, 1978), or J. Bergh and J. Löfström, *Interpolation Spaces, an Introduction* (Springer-Verlag, Berlin, 1976).
[8] J. -P. Antoine and W. Karwowski, Bull. Acad. Pol. Sci. Ser. Sci. Phys. (in press).
[9] L. Schwartz, J. Anal. Math. (Jerusalem) **13**, 115–256 (1964).
[10] M. Reed and B. Simon, *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-Adjointness* (Academic, New York, 1975).
[11] C. Goulaouic, Ann. Inst. Fourier (Grenoble) **18**, 1–98 (1968).
[12] M. Friedrich and G. Lassner, Wiss. Z. Karl-Marx-Univ. Leipzig, Math. Naturwiss. Reihe. **27**, 245–51 (1978); M. Friedrich and G. Lassner, Rep. Math. Phys. (to be published).
[13] J. -P. Antoine, *Proc. Intern. Conf. on Operator Algebras, Ideals and their Applications in Theoretical Physics*, H. Baumgärtel *et al.* (Teubner, Leipzig, 1978), pp. 279–289.
[14] R. A. Hischfeld, Math. Ann. **194**, 180–96 (1971).
[15] S. G. Krein and Ju. I. Petunin, Usp. Mat. Nauk **21**, 89–168 (1966) [Russ. Math. Surv. **21**, 85–159 (1966)].
[16] R. S. Palais, *Seminar on the Atiyah–Singer Theorem*, Chap. VII, Ann. of Math. Studies #57 (Princeton U.P., Princeton, N.J., 1965).
[17] V. Bargmann, Commun. Pure Appl. Math. **20**, 1–101 (1967).
[18] A. C. Zaanen, *Integration*, 2nd ed. (North-Holland, Ansterdam, 1961) Chap. 15; W. A. J. Luxemburg and A. C. Zaanen, Math. Ann. **162**, 337–50 (1966).
[19] G. G. Gould, J. London Math. Soc. **34**, 161–72 (1959).
[20] M. A. Krasnoselskii, P. P. Zabreiko, E. I. Pustylnik, and P. E. Sbolevskii, *Integral Opetators in Spaces of Summable Functions* (Nauka, Moscow, 1966) [Engl. Transl. (Noordhoff, Leyden, 1976)].
[21] Z. Popowicz, Bull. Acad. Pol. Sci. Ser. Sci. Math. **23**, 1119–24 (1975).
[22] J. -P. Antoine, *Recent Developments in Relativistic Quantum Field Theory and its Applications*, Vol. II, Acta Univ. Wrat. #207 (Wroclaw, Poland, 1974), pp. 139–173.
[23] H. H. Schaefer, *Topological Vector Spaces* (Springer-Verlag, Berlin, 1971).

An integral transform related to quantization

I. Daubechies [a)]
Theoretische Natuurkunde, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel Belgium
A. Grossmann
CPT II, Centre de Physique Théorique, Centre de Luminy, Case 907, 70 Route Léon Lachamp, F-13288 Marseille Cedex 2 France

(Received 18 December 1978; accepted for publication 39 August 1979)

We study in some detail the correspondence between a function f on phase space and the matrix elements $(Q_f)(a, b)$ of its quantized Q_f between the coherent states $|a\rangle$ and $|b\rangle$. It is an integral transform: $Q_f(a, b) = \int \{a, b\,|v\} f(v)\, dv$ which resembles in many ways the integral transform of Bargmann. We obtain the matrix elements of Q_f between harmonic oscillator states as the Fourier coefficients of f with respect to an explicit orthonormal system.

I. INTRODUCTION

Quantization is a word which should be used with caution, since it means many things to many people. We understand it here in the sense first sketched by Weyl,[1] where it describes a "harmonic analysis" procedure. It consists in Fourier analyzing a (fairly arbitrary) function on phase space, and then replacing the "elementary building blocks" (i.e., exponentials on phase space) by appropriate operators (which have since been known as Weyl operators, and are exponentials of linear combinations of the operators X and P).

A satisfactory and intrinsic description of the procedure became possible when von Neumann[2] proved the uniqueness theorem (Steps towards the theorem can be found in Weyl's book.[1]) which states that for a given (finite) number of degrees of freedom there exists—up to unitary equivalence—essentially only one *irreducible* family of Weyl operators in Hilbert space. This theorem is a cornerstone of quantum mechanics for a finite number of degrees of freedom. It seems however to have appeared too late to be fully incorporated in the mainstream of texbooks on the subject.

The intrinsic and symplectic formulation of quantum mechanics, made possible by von Neumann's theorem, was developed by Segal[3,4] and Kastler,[5] largely as a by-product of work aimed at systems with infinitely many degrees of freedom (The MIT thesis of R. Lavine[6] is devoted to finite numbers of degrees of freedom). The ingredients are as follows:

(1) A phase space E which is defined (without any "*a priori*" decomposition into position and momentum) as an even-dimensional vector space ($\dim E = 2\nu$) with an antisymmetric nondegenerate bilinear form σ.

(2) A Weyl system W which is a family of unitary operators, labeled by points in phase space, acting irreducibly on a Hilbert space $\mathscr{H}$ and satisfying

$$W(v_1)W(v_2) = e^{i\sigma(v_1,v_2)}W(v_1 + v_2). \tag{1.1}$$

Given E and σ, von Neumann guarantees the existence and uniqueness (up to unitary equivalence) of W, but does not commit us to any concrete realization of W. The Weyl quantization procedure is then a two-step affair: (a) Fourier analysis: $f(v)$ is written as

$$f(v) = 2^{-\nu}\int e^{i\sigma(v,v')}\tilde f(v')dv'; \tag{1.2}$$

(b) substitution of $W(-v/2)$ for $e^{i\sigma(v,\cdot)}$, giving

$$Q(f) = 2^{-\nu}\int W\left(\frac{-v}{2}\right)\tilde f(v)\, dv \tag{1.3}$$

as the definition of the "quantized" of f.

It was shown in Ref. 3 that the correspondence $\tilde f \to Q(f)$ is inverted by

$$\tilde f(v) = 2^{-\nu}\mathrm{tr}\left(W\left(\frac{1}{2}v\right)Q(f)\right) = 2^{-\nu}\left(\left(W\left(\frac{-v}{2}\right), Q(f)\right)\right)_{\mathrm{HS}}, \tag{1.4}$$

where $((,))_{\mathrm{HS}}$ is the inner product in the Hilbert space $\mathscr{L}_{\mathrm{HS}}$ of Hilbert–Schmidt operators in $\mathscr{H}$, and that the map $f \to Q(f)$ is unitary from $L^2(E)$ onto $\mathscr{L}_{\mathrm{HS}}$. Consequently,

$$(f_1, f_2)_{L^2(E)} = ((Q(f_1), Q(f_2)))_{\mathrm{HS}}. \tag{1.5}$$

If e^a denotes the function $e^a(v) = e^{i\sigma(a,v)}$ we have $Q(e^a) = W(-a/2)$, and so, by extension of Eq. (1.5)

$$((Q(e^a), Q(e^b)))_{HS} = \int e^{-a}(v)e^b(v)dv = 2^{2\nu}\,\delta(a-b),$$

in a sense to be made precise (see, e.g., Ref. 7). This map is discussed in more detail by Pool.[8]

In Ref. 9, one of us made the remark that step (a) of the quantization procedure (Fourier analysis) can be avoided at the price of replacing the Weyl operators $W(v)$ by *Wigner operators* $\Pi(v)$ which are simply Weyl operators multiplied by parity, i.e., if $\Pi(0)$ is the parity operator (which can be defined intrinsically up to a sign in any Hilbert space that carries a Weyl system), and if $\Pi(v)$ is defined by

$$\Pi(v) = W(2v)\,\Pi(0) = W(v)\Pi W(-v),$$

then $Q(f)$ can be written directly as

$$Q(f) = 2^{\nu}\int f(v)\Pi(v)\, dv \tag{1.6}$$

[a)]Scientific collaborator at the Interuniversity Institute for Nuclear sciences (In the framework of research Project 21EN).

and we do not have to consider the Fourier transform $\tilde{f}$ of the function f. The reason for calling $\Pi(v)$ a Wigner operator is that the Wigner quasiprobability density $\rho_\psi(v)$ corresponding to a pure state ψ is just the expectation value of $\Pi(v)$:

$$\rho_\psi(v) = 2^\nu(\psi,\Pi(v)\psi)$$

(see Ref. 7).

Equation (1.6) expresses $Q(f)$ as a superposition of Wigner operators, which are in some ways simpler than Weyl operators, namely, (i) Every Wigner operator $\Pi(v)$, in addition to being unitary $[\Pi^*(v) = (\Pi(v))^{-1}]$ is also selfadjoint $[\Pi^*(v) = \Pi(v)]$. Consequently, $\Pi(v)$ is involutive $[(\Pi(v))^2 = 1]$ and its spectrum consists of the numbers $+1$ and -1. (ii) The relationship (1.1) for Weyl operators is replaced by

$$\Pi(v_1)\Pi(v_2)\Pi(v_3) = e^{i\varphi(v_1,v_2,v_3)}\Pi(v_1 - v_2 - v_3) \quad (1.7)$$

(see Huguenin[10]), where $\varphi(v_1,v_2,v_3) = 4[\sigma(v_1,v_3) + \sigma(v_3,v_2) + \sigma(v_2,v_1)]$ is the oriented area of the triangle spanned by v_1, v_2, v_3; thus, Eq. (1.7) is affine (i.e., independent of the choice of origin in phase space) while Eq. (1.1) is vectorial (i.e., dependent on the choice of origin).

We can again invert formula (1.6) to obtain an expression analogous to Eq. (1.4), but giving now a direct correspondence between f and Q_f:

$$f(v) = 2^\nu \mathrm{tr}\,(Q_f\Pi(v)). \quad (1.8)$$

Wigner operators (without the name) were already present in Ref. 11. They can also be found in Ref. 12, where a decomposition of operators with respect to Wigner operators is given, analogous to Eq. (1.6), and relation (1.8) is derived for the function used in this decomposition. The Wigner operators were however not discussed in Ref. 12 as a means to do Weyl quantization without having to pass through the Fourier analysis step. That the $\Pi(v)$ formulas (1.6) and (1.8) may be more convenient then the $W(v)$ formulas (1.3) and (1.4) was also implicitly recognized in Ref. 13, where indeed some nondiagonal matrix elements involving a parity operation were used rather than the diagonal ones to compute classical function, which amounts exactly to preferring Eq. (1.8) as a direct formula to the indirect version (1.4) containing still a Fourier transform.

In this paper we will exploit Eq. (1.6) to study directly the relationship between the function f and matrix elements of the corresponding operator Q_f. We are particularly interested in the matrix elements of Q_f between coherent states. So the coherent state formalism will be the second main input in this paper.

The coherent state formalism has "a long and proud history in quantum theory."[14] Coherent states can be considered as eigenstates of a displaced harmonic oscillator (it is in this form they historically made their first appearance; see Ref. 15), as wave packets satisfying the minimum-uncertainty conditions, or as the eigenfunctions of the annihilation operator associated to the harmonic oscillator. For more details concerning these different points of view, see Ref. 14 and the references quoted there. We will consider the coherent states as displacements of the harmonic oscillator vacuum Ω:

$$|a\rangle = \Omega^a = W(a)\,\Omega, \quad (1.9)$$

where the operator representing a shift in phase space by the vector a is the Weyl operator $W(a)$.

The coherent states have many interesting properties and have therefore been widely used in quantum physics. One property which has been frequently made use of is the resolution of the identity operator[16]:

$$1 = \int dv\, |\Omega^v)(\Omega^v|. \quad (1.10)$$

One can use this property to represent every vector ψ in $\mathscr{H}$ by the wavefunction ϕ_ψ defined as

$$\phi_\psi(v) = (\Omega^v,\psi). \quad (1.11)$$

The set of wavefunctions constructed in this way forms a Hilbert space when equipped with the L^2 norm, and one has

$$\forall\ \varphi,\ \psi\in\mathscr{H}: \quad (\varphi,\psi) = (\phi_\varphi,\phi_\psi) = \int dv\, \overline{\phi_\varphi(v)}\phi_\psi(v) \quad (1.12)$$

[this is again Eq. (1.10)] and

$$\phi_\psi(a) = (\phi_{\Omega^a},\phi_\psi). \quad (1.13)$$

In fact, the mapping $\psi \mapsto \phi_\psi$ is, up to a Gaussian factor, the unitary map from our abstract Hilbert space $\mathscr{H}$ to the Bargmann–Segal Hilbert space of analytic functions $\mathscr{H}_{BS}$[17,18] intertwining the operators $W(a)$ of the irreducible representation of the Weyl commutation relations on $\mathscr{H}$ with the usual Weyl operators on $\mathscr{H}_{BS}$. For the special choice $\mathscr{H} = L^2(R^\nu)$ with the Schrödinger realization of the commutation relations this unitary map was constructed explicitly in Ref. 17.

With the help of the unitary map $\psi \mapsto \phi_\psi$ one can easily transport the Weyl operators $W(a)$ on $\mathscr{H}$ to the Hilbert space $\{\phi_\psi;\ \psi\in\mathscr{H}\}$. The irreducible representation of the Weyl commutation relations obtained in this way is the coherent state representation of the canonical commutation relations. Note that to define the coherent state representation we have used the harmonic oscillator vacuum Ω. To define this Ω in an unambiguous way some additional structure on E, σ is needed. Therefore, one introduces a complex structure J on E, compatible with σ (see Sec. 2). This is analogous to but weaker than the usual decomposition of E into a direct sum of x space and p space.

In what follows we will define an integral transform which relates f to the matrix elements of Q_f between coherent states:

$$\langle a|Q_f|b\rangle = (\Omega^a,Q_f\Omega^b) = \int dv\, \{a,b\,|v\}f(v). \quad (1.14)$$

Such a direct relationship between f and the matrix elements of Q_f enables us to study some aspects of the quantization procedure by means of a correspondence between function or distribution spaces instead of as a map from functions on phase space to operators on a Hilbert space. Of course, we could achieve this by using directly the classical functions, equipped with the twisted product rather than the usual product—this is the point of view of deformation theory; see for instance Ref. 19— or the functions occurring in the so-called diagonal or P respresentation.[16,20-23] In both these approaches, however, functions corresponding to very nice operators may have quite singular features: the classical function corresponding to the parity operator is a delta function, and it is a well-known fact that the P-representation function of a trace class operator may have such big growth at infinity that it is not even a tempered distribution. Working with the correspondence $f \rightarrow \langle a|Q_f|b\rangle$ we gain in smoothness with respect to these two approaches; the price we pay for this is an increase of the numbers of variables used (4ν variables instead of 2ν).

Coherent states can be defined in any Hilbert space carrying an irreducible Weyl system, which means that the matrix elements (1.14) can be computed in any representation, and are representation independent. We will use this to choose one specific representation, namely, the coherent state representation (written in an intrinsic, i.e., coordinateless way), which is particularly well suited for calculations with coherent states; the matrix elements we compute will however be independent of this particular choice of representation. The kernel $\{a, b\,|v\}$ (which was briefly discussed in Ref. 24) is studied in Sec. 3; the notations are explained in Sec. 2. We consider in particular a bilinear expansion for the kernel with respect to a basis formed by the matrix elements $h_{mn}(v)$ of the Wigner operators $\Pi(v)$ between harmonic oscillator states. These functions h_{mn} are given explicitly by Eq. (3.28). The Fourier coefficients of an arbitrary function f with respect to the basis h_{mn} are the matrix elements of Q_f between harmonic oscillator states.

The integral transform (1.4) which is discussed in Sec. 4 is analogous in many ways to the transform of Bargmann.[17] This analogy and the differences are discussed in Sec. 6.

The discussion of the integral transform given here is not at all exhaustive: a deeper study will be carried out in a forthcoming paper; we will study in particular the correspondence between some classes of distributions and the corresponding operators. A first application of Eq. (1.14) can be found in the computation of the classical functions corresponding to linear canonical transformations in Ref. 25.

2. THE GEOMETRICAL SETTING

We find it convenient to work in phase space without coordinates whenever possible. We shall however also rewrite some of the main formulas in a notation with coordinates which may be more familiar to most readers.

A. Affine phase space (symplectic geometry)

We denote by E a set which has the structure of an affine space (i.e., which can be identified to a real vector space after the choice of an origin). Assume that E is even dimensional and that we have associated an "oriented area" $\varphi(a, b, c)$ to every triangle with vertices a, b, c (taken in a given order). We assume the following:

(i) φ does not change if all its arguments are shifted by the same vector.

(ii) If a point o is chosen as the origin, then $\sigma(a, b)$ defined by

$$\sigma(a,b) = \tfrac{1}{4}\varphi(a,o,b)$$

is symplectic (i.e., bilinear, antisymmetric, and nondegenerate).

The function φ can now be expressed in terms of σ:

$$\varphi(a, b, c) = 4(\sigma(b, a) + \sigma(a, c) + \sigma(c, b)).$$

We see that it is totally antisymmetric: it changes sign if any two arguments are interchanged.

B. Phase space with a symplectic and a Euclidean geometry

Consider in E a reference frame, i.e., a family of vectors $a_1, b_1, \dots, a_\nu, b_\nu$ that span E and such that $\sigma(a_i, a_j) = \sigma(b_i, b_j) = 0$, and $\sigma(a_i, b_j) = \delta_{ij}$. For our purposes (the building of a representation space for the Heisenberg commutation relations) all the relevant information is contained in the map J defined by

$$Ja_i = b_i, \quad J\,a_i = b_i, \quad J\,b_i = -a_i \qquad (i = 1, \dots, \nu).$$

Notice that J has the following properties:

$$J^2 = -\mathbb{1} \tag{2.1}$$

(which is expressed by saying that J is a complex structure),

$$\sigma(Ja, Jb) = \sigma(a, b), \tag{2.2}$$

and

$$\sigma(a, Ja) > 0, \text{ if } a \neq 0. \tag{2.3}$$

It follows that the bilinear form

$$s(a, b) = \sigma(a, Jb) \tag{2.4}$$

defines a Euclidean geometry on E. A Triangle a, b, c has now not only an oriented area $\varphi(a, b, c)$ but also side lengths $(s(a-b, a-b))^{1/2}, \dots$, which however depend on the choice of J.

We shall also use the complex combination

$$h(a, b) = s(a, b) + i\sigma(a, b), \tag{2.5}$$

which makes E into a ν-dimensional Hilbert space.

Examples: (1) Take $E = \mathbb{C}^\nu$ with

$$\sigma(a, b) = \mathrm{Im}(\bar{a}.b), \tag{2.6}$$

$$Ja = ia. \tag{2.7}$$

Then

$$s(a, b) = \mathrm{Re}(\bar{a}.b), \tag{2.8}$$

$$h(a, b) = \bar{a}.b, \tag{2.9}$$

and all the conditions above are satisfied. (2) Take $E = \mathbb{R}^\nu \oplus \mathbb{R}^\nu$. Any a in E is written as (x_a, p_a). Define

$$\sigma(a,b) = \tfrac{1}{2}(p_a.x_b - p_b.x_a), \tag{2.10}$$

$$J(x_a, p_a) = (p_a, -x_a). \tag{2.11}$$

Then

$$s(a,b) = \tfrac{1}{2}(x_a.x_b + p_a.p_b), \tag{2.12}$$

and all our conditions are fulfilled again. We can now use this structure to build a representation of the cannonical commutation relations.

C. A representation space for canonical commutation relations

On E, we consider the space of holomorphic functions

$$\mathscr{F}(E) = \{\varphi : E \to \mathbb{C} \mid \nabla^{Ja}\varphi = i\nabla^{a}\varphi \text{ for any } a \text{ in } E\}, \quad (2.13)$$

where

$$(\nabla^a \varphi)(v) = \lim_{\lambda \to 0} \frac{1}{\lambda} \varphi(v + \lambda a).$$

On the other hand, we define the Gaussian Ω by

$$\Omega(v) = \exp[-\tfrac{1}{4}s(v,v)]. \quad (2.14)$$

We shall say that a function ϕ on E is *modified holomorphic* if it is the product of a $\varphi \in \mathscr{F}$ and of the Gaussian: $\varphi(v) = \phi(v)\Omega(v)$.

This combination of both $\mathscr{F}$ and Ω gives us Z, the space of modified holomorphic functions:

$$Z(E) = \{\varphi\Omega \mid \varphi \in \mathscr{F}\}. \quad (2.15)$$

An alternative way of defining Z is

$$Z(E) = \{\phi : E \to \mathbb{C} \mid D^{Ja}\phi = iD^{a}\phi\}, \quad (2.15')$$

with $(D^a\phi)(v) = (\nabla^a\phi)(v) + s(a,v)\phi(v)$.

The Hilbert space we will use in the sequel whenever we want to consider a concrete representation space is now given by

$$\mathscr{L}_0 = Z(E) \cap L^2(E;dv), \quad (2.16)$$

where dv is the Lebesgue measure, translationally invariant, normalized by the requirement

$$\int \Omega^2(v)\,dv = 1. \quad (2.17)$$

On this Hilbert space, we define a set of unitary operators $W(a)$ by

$$(W(a)\phi)(v) = \exp[i\sigma(a,v)]\phi(v-a), \text{ for any } a \text{ in } E. \quad (2.18)$$

These $W(a)$ are called Weyl operators. It is easy to check that

$$W(a)W(b) = \exp[i\sigma(a,b)]W(a+b), \quad (2.19)$$

which implies we have a representation of the canonical commutation relations. One can easily see this in the example given above:

$$E = \mathbb{R}^\nu \oplus \mathbb{R}^\nu,$$

$$W((x_a,0))W(0,p_b))$$

$$= \exp\left(-\frac{i}{2}x_a \cdot p_b\right)W((x_a,p_b))$$

$$= \exp(-ix_a \cdot p_b)W((0,p_b))W((x_a,0)).$$

This is exactly what one would have expected from

$$W((x_a,p_b)) = \exp(i(x_a.X + p_b.P)),$$

with

$$[X_j,P_k] = i\delta_{jk}.$$

The representation given by Eq. (2.18) in the space (2.16) is irreducible.

Owing to von Neumann's uniqueness theorem for representations of the canonical commutation relations for a finite number of degrees of freedom, any result we will obtain in our particular representation on $\mathscr{L}_0$ can be transcribed to any irreducible representation.

Some particular functions in $\mathscr{L}_0$ will play a special role in the sequel: They are called the coherent states and are defined as

$$\Omega^a(v) = (W(a)\Omega)(v) = \exp(i\sigma(a,v))\Omega(v-a). \quad (2.20)$$

These coherent states have the following "reproducing" property[16,17]:

$$(\Omega^a,\phi) = \int \overline{\Omega^a(v)}\,\phi(v)\,dv = \phi(a) \text{ for any } \phi \in \mathscr{L}_0. \quad (2.21)$$

Writing this otherwise, we have

$$(\phi,\psi) = \int \overline{\phi(v)}\,\psi(v)\,dv$$

$$= \int (\phi,\Omega^v)(\Omega^v,\psi)dv;$$

hence

$$\int |\Omega^v\rangle\langle\Omega^v|\,dv = 1. \quad (2.22)$$

It is now easy to see that the Ω^a are normalized elements of $\mathscr{L}_0$:

$$(\Omega^a,\Omega^a) = \Omega^a(a) = \Omega(0) = 1. \quad (2.23)$$

As we already mentioned in the Introduction it is often useful to introduce Wigner operators, i.e., products of Weyl operators with the parity operator. We define

$$\pi : \psi \to \check{\psi}, \quad (2.24)$$

with $\check{\psi}(v) = \psi(-v)$. This operator conserves the modified holomorphy properties of and is thus an involutive unitary operator from $\mathscr{L}_0$ to itself. Moreover, one easily sees that

$$\Pi W(v) = W(-v)\Pi$$

or

$$\Pi W(v)\Pi = W(-v).$$

Hence, Π represents the parity $v \to -v$ on phase space. The Wigner operators $\Pi(a)$ are now defined as

$$\Pi(a) = W(2a)\Pi \quad (2.25)$$

i.e.,

$$(\Pi(a)\psi)(v) = e^{i\sigma(2a,v)}\psi(2a-v), \text{ for any } \psi \text{ in } \mathscr{L}_0. \quad (2.26)$$

It is easy to check that

$$\Pi(a)\Pi(b)\Pi(c) = \exp[i\varphi(a,b,c)]\Pi(a-b+c). \quad (2.27)$$

3. THE FUNCTIONS $\{a,b|v\} = \{\xi/v\}$

Definition: Let $\mathscr{H}$ be a Hilbert space carrying an irreducible representation of the Weyl commutation relations for ν degrees of freedom. Denote by Ω^a the coherent state centered at $a \in E$, and by $\Pi(v)$ the parity operator around v (Wigner operator). Given $a, b, v \in E$, we define

$$\{a,b|v\} = \kappa^{-1}(\Omega^a,\Pi(v)\Omega^b), \quad (3.1)$$

with

$$\kappa = 2^{-\nu}. \quad (3.1')$$

The numbers $\{a,b|v\}$ can be easily calculated and have simple properties.

A. Explicit form, symmetries, and special values

One has

$$\{a, b\,|v\} = 2^{\nu}\exp\left[i\varphi\left(\frac{a}{2}, v, \frac{b}{2}\right)\right] \Omega(a+b-2v), \tag{3.2}$$

i.e., the phase of $\{a, b\,|v\}$ is the oriented area of the triangle with vertices $\frac{1}{2}a$, v, $\frac{1}{2}b$. The number $\{a, b\,|v\}$ is real if and only if the three points $\frac{1}{2}a$, v, $\frac{1}{2}b$ are collinear.

The absolute value of $\kappa\{a, b\,|v\}$ is the exponential of the negative of half the squared Euclidean distance from v to the midpoint $(a+b)/2$ of the segment (a, b). It takes its maximal value (which is 1) when v is the midpoint of (a, b).

If we denote by ζ the pair $\{a, b\}$ we have

$$\{-\zeta\,|-v\} = \{-a, -b\,|-v\} = \{a, b\,|v\} = \{\zeta\,|v\}. \tag{3.3}$$

Denote by $\bar{\zeta}$ the pair $\{b, a\}$. (This will be justified below.) Then

$$\{\bar{\zeta}\,|v\} = \{b, a|v\} = \overline{\{a, b\,|v\}} = \overline{\{\zeta\,|v\}}. \tag{3.4}$$

If the arguments of $\{a, b\,|v\}$ are shifted, we have

$$\{a+c, b+c|v+c\} = \exp[i\sigma(c, a-b)]\{a, b\,|v\}. \tag{3.5}$$

One has

$$\{a, a|v\} = |\{a, a|v\}| = \kappa^{-1}\Omega(2v-2a) \tag{3.6}$$

and

$$\left\{a, b\,\Big|\,\frac{a+b}{2}\right\} = \kappa^{-1}. \tag{3.7}$$

B. Expression of $\{a, b|v\}$ in coordinates

(1) Identify E with $\mathbb{C}^{\nu}$. Then $\sigma(a, b) = \mathrm{Im}(\bar{a}.b)$, $Ja = ia$, and $s(a, b) = \mathrm{Re}(\bar{a}.b)$. So

$$\{a, b\,|v\} = 2^{\nu}\exp(-\tfrac{1}{2}|a|^2 - \tfrac{1}{2}|b|^2 - 2|v|^2 - a.\bar{b} + 2\bar{b}.v + 2\bar{v}.a). \tag{3.8}$$

(2) Identify E with $\mathbb{R}^{\nu} \oplus \mathbb{R}^{\nu}$. So a is the pair $a = \{x_a, p_a\}$: $\sigma(a, b) = \frac{1}{2}(p_a x_b - p_b x_a)$, $J(x_a, p_a) = (p_a, -x_a)$, and $s(a, b) = \frac{1}{2}(p_a p_b + x_a x_b)$.
Then

$$\{a, b\,|v\} = 2^{\nu}\exp[-\tfrac{1}{4}|x_a + x_b - 2x_v|^2 - \tfrac{1}{4}|p_a + p_b - 2p_v|^2 + i(x_v(p_b - p_a) - p_v(x_b - x_a) + \tfrac{1}{2}p_a x_b - \tfrac{1}{2}p_b x_a)]. \tag{3.9}$$

C. Analyticity and regularity properties

The expression (3.2) can be rewritten as

$$\{a, b\,|v\} = 2^{\nu}\exp[2h(b,v) + 2h(v,a) - h(b,a)] \times\Omega(a)\,\Omega(b)\,\Omega(2v), \tag{3.10}$$

where $h(a, v)$ is defined by Eq. (2.5). In coordinates, Eq. (3.10) is just Eq. (3.8).

Since Eq. (3.10) can be rewritten as

$$\begin{aligned}\{a, b\,|v\} &= 2^{\nu}\exp[2i\sigma(b,v)]\Omega^{2v-b}(a)\\ &= 2^{\nu}\exp[2i\sigma(v,a)]\,\overline{\Omega^{2v-b}(b)}\\ &= 2^{\nu}\exp[i\sigma(a,b)]\Omega^{-1}(a-b)\\ &\quad\times\Omega^{\sqrt{2}b}(\sqrt{2}v)\,\overline{\Omega^{\sqrt{2}a}(\sqrt{2}v)},\end{aligned}$$

we see that $\{a, b\,|v\}$ is modified holomorphic in a, modified antiholomorphic in b, and a product of a modified holomorphic function with a modified antiholomorphic one in v. In each of these variables it is infinitely differentiable and of Gaussian decrease at infinity.

D. Fourier transforms and integrals

The (symplectic) Fourier transform F can be defined by

$$(Fg)(v) = \kappa\int \exp[i\sigma(v,v')]g(v')\,dv'. \tag{3.11}$$

Then

$$F^2 = \mathbb{1}, \quad F\Omega = \Omega, \quad F\Omega^{a} = \Omega^{-a}. \tag{3.12}$$

If a function ϕ is modified holomorphic [see Eq. (2.15′)], it satisfies $(F\phi)(v) = \phi(-v)$. If it is modified antiholomorphic, it satisfies $(F\phi)(v) = \phi(v)$. So

$$\kappa\int \exp[i\sigma(a,a')]\{a', b\,|v\}\,da' = \{-a, b\,|v\}. \tag{3.13}$$

In particular,

$$\kappa\int \{a', b\,|v\}\,da' = \kappa^{-1}\exp[i\sigma(b, 2v)]\,\Omega(2v-b). \tag{3.14}$$

Similarly,

$$\kappa\int \exp[i\sigma(b,b')]\{a, b'|v\}\,db' = \{a, b\,|v\}. \tag{3.15}$$

In particular,

$$\kappa\int \{a, b'|v\}\,db' = \kappa^{-1}\exp[i\sigma(2v,a)]\,\Omega(2v-a). \tag{3.16}$$

The Fourier transform in the variable v can be computed directly. It is

$$\kappa\int \exp[i\sigma(v,v')]\{a, b\,|v'\}\,dv' = \kappa^2\{a, -b\,|\tfrac{1}{4}v\}. \tag{3.17}$$

In particular,

$$\kappa\int \{a, b\,|v'\}\,dv' = \kappa\,\Omega^{-b}(a). \tag{3.18}$$

One has also

$$\int \{a, a|v\}\,da = 1. \tag{3.19}$$

We now consider integrals that are bilinear in the symbols $\{a, b\,|v\}$. Particularly important is the relationship

$$\int\int \{a, b\,|v\}\{b, a|v'\}\,da\,db = \delta(v-v'), \tag{3.20}$$

which can also be written as

$$\int \{\zeta\,|v\}\{\bar{\zeta}\,|v'\}\,d\zeta = \int \{\zeta\,|v\}\overline{\{\zeta\,|v'\}}\,d\zeta = \delta(v-v'). \tag{3.21}$$

We shall derive it here, in order to show how simple the calculations are:

$$\int\int \{a, b\,|v\}\{b, a|v'\}\,da\,db$$

$$= \kappa^{-2}\int\int (\Omega^a,\Pi(v)\Omega^b)(\Omega^b,\Pi(v')\Omega^a)\,da\,db$$
$$= \kappa^{-2}\int (\Omega^a,\Pi(v)\Pi(v')\Omega^a)\,da$$
$$= \kappa^{-2}\int \exp(i\varphi(v,v',\frac{a}{2}))\Omega^{2(v-v')+a}(a)\,da$$
$$= \kappa^{-2}\exp(4i\sigma(v',v))\,\Omega(2v-2v')$$
$$\times\int \exp[4i\sigma(a,v-v')]\,da$$
$$= \exp[4i\sigma(v',v)]\,\Omega(2v'-2v)\kappa^{-4}\delta(4v'-4v)$$
$$= \delta(v'-v).$$

Another useful relation is

$$\int \{a,b\,|v\}\{c,d\,|v\}\,dv = \Omega^d(a)\Omega^b(c). \tag{3.22}$$

We give finally two integrals of triple products:

$$\int\int \{a,b\,|v\}\{b,c|v'\}\{c,d\,|v''\}\,db\,dc$$
$$= \kappa^{-2}\exp[i\varphi(v,v',v'')]\{a,d\,|v-v'+v''\} \tag{3.23}$$

and

$$\int\int\int \{a,b\,|v\}\{b,c|v'\}\{c,a|v''\}\,da\,db\,dc$$
$$= \kappa^{-2}\exp[i\varphi(v,v',v'')]. \tag{3.24}$$

E. Bilinear expansions of $\{\zeta|v\}$; The orthonormal family $h_{mn}(v)$

Since $\{\zeta\,|v\}$ will be used as the kernel of an integral transform, it is natural to expand it into a sum of products of functions of v and of functions of ζ. Such expansions can be found immediately since the $\{a,b\,|v\}$'s are matrix elements of irreducible families of operators.

1. Generalities

For any orthonormal base e_n in $\mathscr{H}$, we define the functions e_{mn} by

$$e_{mn}(v) = \kappa^{-1}(e_m,\Pi(v)e_n). \tag{3.25}$$

These e_{mn} form an orthonormal base in $L^2(E)$: orthogonality is a consequence of Eq. (3.22), and completeness follws from the fact that the family $\{\Pi(a)\}$ is an irreducible family of operators.

It is now obvious that

$$\{a,b\,|v\} = \sum_{m,n}(\Omega^a,e_m)(e_n,\Omega^b)e_{mn}(v). \tag{3.26}$$

2. In the representation on $\mathscr{L}_0$

Let $\{e_1,\ldots,e_\nu\}$ be a symplectic base on E,σ, i.e., a set of ν vectors satisfying

$$\sigma(e_j,e_k)=0,\qquad \sigma(e_j,Je_k)=\delta_{jk}.$$

We define the normalized monomials $h^{[n]}$ on E by

$$h^{[n]}(v) = \frac{1}{\sqrt{[n]!}}\prod_{j=1}^{\nu}(h(e_j,v))^{n_j}. \tag{3.27}$$

Here $[n]$ is a multi-index, with $[n]! = \Pi_{j=1}^{\nu}(n_j!)$.

The functions $h^{[n]}\Omega$ form an orthonormal base in $\mathscr{L}_0$. In fact, they are the eigenfunctions of the harmonic oscillator (see also Sec. 5). In the following we shall drop the square brackets in the notation $h^{[n]}$. In this particular case we have now

$$h_{mn}(v) = \kappa^{-1}(h^m\Omega,\Pi(v)h^n\Omega)$$
$$= \kappa^{-1}\Omega(2v)\sum_{s=0}^{\min(m,n)}(-1)^s\left(\binom{n}{s}\binom{m}{s}\right)^{1/2}$$
$$\times h^{n-s}(2v)\,\overline{h^{m-s}(2v)} \tag{3.28}$$

and

$$\{a,b\,|v\} = \Omega(a)\,\Omega(b)\sum_{m,n}h^m(a)\,\overline{h^n(b)}\,h_{mn}(v). \tag{3.29}$$

Formula (3.29) implies we can also write the h_{mn} as

$$h_{mn}(v) = \frac{1}{m!\,n!}\left[(D^a)^m(D^b)^n\{a,b\,|v\}\right]_{a=b=0}$$
$$= \frac{1}{m!\,n!}\left[(\nabla^a)^m(\nabla^b)^n(\{a,b\,|v\}\Omega^{-1}(a)\Omega^{-1}(b))\right]_{a=b=0},$$

where D^a and ∇^a are defined in Sec. 2c.

If one identifies E with $\mathbb{C}^\nu$, the h^n are simply

$$h^n(z) = \frac{1}{\sqrt{n!}}z^n,$$

and the h_{mn} are given by

$$h_{mn}(z) = 2^\nu 2^{m+n}e^{-2|z|^2}\sum_{s=0}^{\min(m,n)}\left[-\frac{1}{4}\right]^s$$
$$\times\frac{\sqrt{n!\,m!}}{s!\,(n-s)!\,(m-s)!}z^{n-s}\bar{z}^{m-s}.$$

Identifying E with $\mathbb{R}^{2\nu}$, one gets

$$h^n(x,p) = \frac{1}{\sqrt{n!}}2^{-n/2}(p+ix)^n$$

and

$$h_{mn}(x,p) = 2^\nu e^{-(x^2+p^2)}\sum_{s=0}^{\min(m,n)}(-2)^{-s}$$
$$\times\frac{\sqrt{n!\,m!}\,2^{(n+m)/2}}{s!\,(n-s)!\,(m-s)!}$$
$$\times(p+ix)^{n-s}(p-ix)^{m-s}. \tag{3.30}$$

In particular, we have

$$h_{nn}(x,p) = (-1)^n 2^\nu e^{-(x^2+p^2)}$$
$$\times\sum_{s=0}^{n}(-2)^s\frac{n!}{s!\,s!\,(m-s)!}(x^2+p^2)^s$$
$$= (-1)^n 2^\nu\exp(-x^2-p^2)L_n(2x^2+2p^2), \tag{3.31}$$

where L_n is the Laguerre polynomial with multi-index $[n]$ in the variables $x_j^2+p_j^2$ $(j=1,\cdots,\nu)$.

As a consequence of Eq. (3.17), we see that

$$(Fh_{mn})(v) = \kappa\left(h^m\Omega,\Pi\left(\frac{v}{4}\right)\Pi h^n\Omega\right)$$
$$= (-1)^n\kappa^2 h_{mn}\left(\frac{v}{4}\right). \tag{3.32}$$

4. THE INTEGRAL TRANSFORM: MAIN PROPERTIES

A. The map from $f(v)$ to $Q_f(\zeta)$

Let $f\in S'(E)$ be a tempered distribution on E. Let $\tilde{f}$ be its Fourier transform, which we write freely as

$$\tilde{f}(v) = \kappa \int \exp[i\sigma(v,v')]f(v')\,dv'. \tag{4.1}$$

Let

$$f(v) = \sum_{m,n} f_{mn} h_{mn}(v) \tag{4.2}$$

be the Fourier series expansion of f in the orthonormal system (3.28). By Eq. (3.30), the expansion of $\tilde{f}$ is

$$\tilde{f}(v) = \kappa^2 \sum (-1)^n f_{mn} h_{mn}\left(\frac{v}{4}\right).$$

Definition: The Q transform of f is the function Q_f on $E \times E$, defined by

$$Q_f(\zeta) = Q_f(a,b) = \int \{a,b\,|v\} f(v)\,dv, \tag{4.3}$$

to be interpreted, if necessary, as the evaluation of the functional f on the testing function $\{a, b\,|.\}$.

Remark: The above definition is more restrictive than necessary since the testing functions $\{a, b\,|.\}$ can handle more general distributions "of type S".[26,27] We shall not try here to study in detail the functional analysis associated with Eq. (4.3).

By Eq. (3.17), Q_f and $Q_{\tilde{f}}$ are related through

$$Q_{\tilde{f}}(a,b) = \kappa^{-2} Q_{\check{f}_4}(a, -b), \quad \text{with} \quad \check{f}_4(v) = f(-4v). \tag{4.4}$$

In fact, this relationship between the matrix elements of $Q_{\tilde{f}}$ and $Q_{\check{f}_4}$ is just a consequence of the equivalence of formulas (1.3) and (1.6). Indeed, we have from Eqs. (1.3) and (1.6)

$$\begin{aligned} Q_{\tilde{f}}(a,b) &= \kappa^{-1} \int dv \tilde{f}(v)(\Omega^a, \Pi(v)\Omega^b) \\ &= \kappa \int dv\, f(v)\left(\Omega^a, W\left(-\frac{v}{2}\right)\Omega^b\right) \\ &= \kappa \int dv\, f(-v)\left(\Omega^a, \Pi\left(\frac{v}{4}\right)\Pi\Omega^b\right) \\ &= \kappa^{-3} \int dv\, f(-4v)(\Omega^a, \Pi(v)\Omega^{-b}) \\ &= \kappa^{-2} Q_{\check{f}_4}(a, -b). \end{aligned} \tag{4.5}$$

Furthermore, by Eq. (3.29) and (4.3), the function $Q_f(a,b)$ can be expressed in terms of the Fourier coefficients f_{mn}:

$$Q_f(a,b) = \Omega(a)\Omega(b) \sum_{m,n} h^n(a) \overline{h^m(b)} f_{mn},$$

where the functions $h^n(a)$ and $h^m(b)$ are defined by Eq. (3.27).

An examination of either Eq. (4.3) or (4.5) shows that $Q_f(a,b)$ is modified holomorphic in a and modified antiholomorphic in b, i.e., it is holomorphic in ζ with respect to the complex structure $(J, -J)$.

B. Inverting the map $f \to Q_f$

Given $Q_f(\zeta)$, we can reconstruct f through

$$f(v) = \int \{\zeta\,|v\} Q_f(\zeta)\,d\zeta = \iint \{a,b\,|v\} Q_f(b,a)\,da\,db, \tag{4.6}$$

provided the integrals converge. This is an immediate consequence of (3.21).

In other words, the same kernel is used for (4.3) and its inverse just as in Fourier transforms and the integral transforms[17] of Bargmann.

C. Physical interpretation

So far, we have only defined some integral transform $f(v) \to Q_f(\zeta)$ by means of the kernel $\{\zeta\,|v\}$. Of course, we always had in mind the physical interpretation of all this when we defined our map from one function space to another. This physical interpretation follows immediately from formula (1.6) and definition (3.1) of the kernel $\{a, b\,|v\}$. One has

$$Q_f(a,b) = \kappa^{-1} \int dv\, f(v)(\Omega^a, \Pi(v)\Omega^b) = (\Omega^a, Q_f\,\Omega^b).$$

So for any a, b in E, $Q_f(a, b)$ is the matrix element between the coherent states Ω^a and Ω^b of the quantum mechanical operator Q_f corresponding to the "classical observable" f.

In an analogous way, the Fourier coefficient

$$f_{mn} = \int \overline{h_{mn}(v)} f(v)\,dv, \tag{4.7}$$

with

$$h_{mn}(v) = \kappa^{-1}(h^m\Omega, \Pi(v)h^n\Omega),$$

the h^s being normalized monomials [see Eq. (3.27)], is equal to the matrix element of Q_f between harmonic oscillator states:

$$f_{mn} = \langle n|Q_f|m\rangle.$$

D. Action of Q_f in $\mathscr{L}_0(E)$

The action of the operator Q_f on $\mathscr{L}_0(E)$ can be written explicitly with the help of the function $Q_f(\zeta)$ [see Eq. (2.21)]:

$$\forall \psi \in \mathscr{L}_0(E) : (Q_f\psi)(a) = \int db\, Q_f(a,b)\psi(b). \tag{4.8}$$

E. Unitarity of the correspondence $f \to Q_f$

The function $Q_f(\zeta)$ is an element of $Z(E \times E; (J, -J))$, i.e., modified holomorphic in its first variable and modified antiholomorphic in its second variable.

Define $\mathscr{L}(E \times E) = L^2(E \times E; dv \otimes dv) \cap Z(E \times E; (J, -J))$. Equipped with the L^2 norm, this is a Hilbert space. Suppose f is square integrable. Then

$$\begin{aligned} &\int d\zeta\, |Q_f(\zeta)|^2 \\ &= \iint da\,db\, \overline{Q_f(a,b)}\, Q_f(a,b) \\ &= \iiiint da\,db\,dv\,dv' f^*(v)\{b,a|v\}\, f(v')\{a,b\,|v'\} \\ &= \iint dv\,dv' f^*(v) f(v')\delta(v - v') \\ &= \int dv\, |f(v)|^2. \end{aligned}$$

Hence the map $f(.) \to Q_f(.)$ is a unitary map from $L^2(E)$ to $\mathscr{L}_0(E \times E)$; its inverse is defined by Eq. (4.6).

In fact, this unitarity is nothing else than the well-known unitarity of the correspondence between square integrable functions and Hilbert–Schmidt operators.[8] Indeed, one can check that the operator Q_f is Hilbert–Schmidt if the function $Q_f(.)$ is square integrable, and one has the equality

$$\|Q_f\|^2_{\text{HS}} = \text{Tr}(Q_f^* Q_f) = \int d\zeta\, |Q_f(\zeta)|^2.$$

F. Products

Let A and B be operators on $\mathscr{L}_0$. Then

$$(A, B)(a,b) = (\Omega^a, AB\Omega^b) = \int (\Omega^a, A\Omega^c)(\Omega^c, B\Omega^b)\,dc = \int A(a,c)B(c,b)\,dc.$$

Hence,

$$(Q_f \cdot Q_g)(a,b) = \int Q_f(a,c)Q_g(c,b)\,dc.$$

We define the twisted product $f \circ g$ by

$$Q_{f\circ g} = Q_f \cdot Q_g. \tag{4.9}$$

Hence,

$$\begin{aligned}(f\circ g)(v) &= \int \{\xi \,|v\} Q_{f\circ g}(\bar{\xi})\,d\xi \\ &= \int\int\int \{a,b\,|v\} Q_f(b,c)Q_g(c,a)da\,db\,dc \\ &= \int\int\int\int\int f(v')g(v'') \\ &\quad \times \{a,b\,|v\}\{b,c|v'\}\{c,a|v''\}da\,db\,dc\,dv'\,dv'' \\ &= \kappa^{-2}\int\int f(v'')g(v'')\exp[i\varphi(v,v',v'')]dv'\,dv'',\end{aligned} \tag{4.10}$$

which is a well-known expression.[10,28]

G. Bounds on $Q_f(a,b)$

Define the following two *regularized functions* associated to f:

$$f_R(v) = \int f(v')\Omega(2(v'-v))\,dv', \tag{4.11}$$

$$f_r(v) = \int f(v')\,\Omega\left(\frac{v'-v}{2}\right)dv'. \tag{4.12}$$

They can be obtained by choosing f as the the initial value of a diffusion (heat) equation and waiting the appropriate time.

Assume now that f is a (locally integrable) function so that $|f|$ (the absolute value of f) is well defined. Denote by $|f|_R$ the regularized Eq. (4.11) of $|f|$. Then Eq. (4.3) gives

$$|Q_f(a,b)| \leqslant \kappa^{-1}|f|_R\left(\frac{a+b}{2}\right). \tag{4.13}$$

On the other hand, if the Fourier transform $\tilde{f}$ of f is a function (here one should not think of Q_f as, say, a Hamiltonian but for example a resolvent),and if $|\tilde{f}|_r$ is the regularized Eq. (4.12) of $|\tilde{f}|$, we obtain from Eq. (4.4)

$$\begin{aligned}|Q_f(a,b)| &= \kappa^{-2}|Q_{\tilde{f}_4}(a,-b)| \\ &\leqslant \kappa^{-3}|\tilde{f}_4|_R\left(\frac{a-b}{2}\right) \\ &= \kappa|\tilde{f}|_r(2(a-b)).\end{aligned} \tag{4.14}$$

If both f and $\tilde{f}$ are functions we obtain

$$|Q_f(a,b)|^2 \leqslant |f|_R\left(\frac{a+b}{2}\right)|\tilde{f}|_r(2(a-b)). \tag{4.15}$$

For the diagonal matrix elements one obtains[24] an equality, which does not require any special assumption on f,

$$Q_f(a,a) = \kappa^{-1}f_R(a). \tag{4.16}$$

More generally, f can be assumed to be a measure.

H. Positivity of Q_f

Suppose that Q_f is positive, i.e.,

$$\forall \psi \in \mathscr{L}_0 : (\psi, Q_f\psi) = \int\int \overline{\psi(a)}\psi(b)Q_f(a,b)\,da\,db \geqslant 0. \tag{4.17}$$

Since

$$\{a,b\,|v\} = \kappa^{-1}\exp[2i\sigma(v,a)]\,\overline{\Omega^{2v-a}(b)}$$

(see Sec. 3c), this is equivalent to

$$\forall \psi\in\mathscr{L}_0 : \int\int \overline{\psi(a)}\psi(2v-a)\exp[2i\sigma(v,a)]\,f(v)\,da\,dv \geqslant 0, \tag{4.18}$$

provided we are allowed to change the order of the integrations, which is certainly true, for example, for $f \in L^2(E;dv)$.

We can rewrite condition (4.18) as

$$\forall\psi\in\mathscr{L}_0 : \int\int \overline{\psi(a)}\psi(b)\exp[i\sigma(b,a)]\,f\left(\frac{a+b}{2}\right)da\,db \geqslant 0. \tag{4.19}$$

For $f\in L^2(E;dv)$, Eq. (4.19) is a necessary and sufficient condition for Q_f to be positive.

If, moreover, we suppose f is essentially bounded ($f\in L_\infty$) and integrable ($f \in L^1$), then Eq. (4.19) is implied by

$$\forall n\in \mathbb{N},\ \forall\, a_1,\ldots,a_n \in E : \text{the matrix} \times\left[\exp[i\sigma(a_j,a_k)]\,f\left(\frac{1}{2}(a_j+a_k)\right)\right]_{j,k} \text{ is positive.} \tag{4.20}$$

So, for $f\in L_\infty \cap L^1$, Eq. (4.20) is a sufficient condition for Q_f to be positive.

A similar result, though in a different context, can be found in Ref. 29; the fact that the matrices

$$[\exp[i\sigma(a_j,a_k)]\,f(a_k-a_j)]_{jk}$$

are considered in Ref. 29 and not

$$\left[\exp[i\sigma(a_j,a_k)]\,f\left(\frac{a_j+a_k}{2}\right)\right]_{jk}$$

as here, is due to their studying the correspondence $f \rightarrow \int dv\, f(v)\, W(v)$ and not $f\rightarrow \int dv\, f(v)\Pi(v)$.

Examples: Using Eq. (4.20), one can easily check that the following functions yield positive operators:

$$f(v) = \Omega^{(1-\alpha)}(2v), \quad \text{with } \alpha < 1.$$

In particular,

$$f(v) = \Omega(2v) \text{ and } f(v) = \Omega(v),$$
$$f(v) = \exp[2s(c,v)]\,\Omega(2v),$$
$$f(v) = \exp[2\sigma(c,v)]\,\Omega(2v).$$

5. EXAMPLES

A. Operators corresponding to elementary functions

For some functions we shall use Eq. (4) to compute both Q_f and $Q_{\tilde{f}}$:

(1) $f(v) = 1$. Then $Q_1(a,b) = \Omega^b(a)$ or $Q_1 = \mathbb{1}$. $\tilde{1}(v) = \kappa^{-1}\delta(v)$. Hence, $Q_\delta(a,b) = \kappa^{-1}Q_1(a,-b) = \kappa^{-1}\Omega^{-b}(a)$. This implies $Q_\delta = \kappa^{-1}\Pi$, which is, of course, implicit in Eq. (4.3).

(2) $f(v) = \exp[i\sigma(c,v)]$. This gives $Q_f(a,b) = \exp[\frac{1}{2}i\sigma(c,b)]\,\Omega^{b+c/2}(a)$; hence $Q_f = \Pi(c/4)\Pi$. Applying again Eq. (4.4), we see that

$$[F e^{i\sigma(c,.)}](v) = \kappa^{-1}\delta(v-c) = \kappa^{-1}\delta_c(v)$$

and

$$Q_{\delta_c} = \kappa^{-1}\Pi(c).$$

(3) $f(v) = \Omega(\alpha v)$ for $\alpha \in \mathbb{R}$.

Then

$$Q_f(a,b) = \left[\frac{4}{4+\alpha^2}\right]^{\nu} \exp[i\sigma(a,b)]^{(\alpha^2-4)/(4+\alpha^2)} \times [\Omega(a+b)]^{(\alpha^2)/(4+\alpha^2)}[\Omega(a-b)]^{4/(4+\alpha^2)}.$$

As a consequence of the fact

$$[F\Omega(\alpha.)](v) = \alpha^{-2\nu}\Omega(\alpha^{-1}v),$$

we see that

$$Q_{\Omega(\alpha^{-1}.)}(a,b) = (2\alpha)^{2\nu} Q_{\Omega(4\alpha.)}(a,-b).$$

(4) As a special case of (3), we have

$$Q_{\Omega(2.)}(a,b) = \kappa\,\Omega(a)\,\Omega(b);$$

hence,

$$Q_{\Omega(2.)} = \kappa\,|\Omega)(\Omega|.$$

(5) $f(v) = s(v,v)$. This is the Hamiltonian of the harmonic oscillator. We have

$$Q_f(a,b) = \left[\frac{\nu}{2} + h(b,a)\right]\Omega^b(a),$$

or

$$Q_f = \frac{\nu}{2} + N,$$

with

$$N : \Omega^b \mapsto h(b,.)\Omega^b.$$

We see here the expected vacuum energy term $\nu/2$; moreover, one can easily check that for $u_n = h^n\Omega$, one has

$$N u_n = n\, u_n,$$

which is in accordance with the well-known fact that the u_n are the harmonic oscillator eigenfunctions.[17]

(6) $f(v) = \sigma(c, v)$. This gives $Q_f(a,b) = (i/2)\Omega^b(a) \times [h(b,c) - h(c,a)]$. Define $H_c : \psi \mapsto h(c,.)\psi$. Then $H_c^* : \psi \to D_c\psi$ with $D_c = \nabla_c + s(c,.)$ or $\varphi.\Omega \mapsto (\nabla_c\varphi)\Omega$ or $\Omega^b \to h(b,c)\Omega^b$, and $Q_{\sigma(c,.)} = (i/2)(H_c^* - H_c)$. Analogously, $Q_{s(c,.)} = \frac{1}{2}(H_c^* + H_c)$ and $Q_{h(c,.)} = H_c$.

(7) $f(v) = \sigma(c_1, v)\,\sigma(c_2, v)$. Then

$$Q_{\sigma(c_1,.)\sigma(c_2,.)} = \tfrac{1}{4}[s(c_1,c_2) + H_{c_1}H_{c_2}^* + H_{c_2}H_{c_1}^* - H_{c_1}H_{c_2} - H_{c_1}^*H_{c_2}^*].$$

Analogously,

$$Q_{h(c_1,.)h(c_2,.)} = H_{c_1}H_{c_2}.$$

B. Functions corresponding to elementary operators

1. Dyadics

Take A to be the dyadic $A = |\phi)(\psi|$. Then

$$f_A(v) = \int\int \{a,b\,|v\}\,\phi(b)\overline{\psi(a)}\,da\,db = (\psi, Q_{\delta_v}\phi).$$

In particular,

$$f_{|\Omega^d)(\Omega^c|}(v) = \{c,d\,|v\},$$

$$f_{|\Omega^c)(\Omega^c|}(v) = \kappa^{-1}\,\Omega(2v-2c),$$

$$f_{|\Omega)(\Omega|}(v) = \kappa^{-1}\,\Omega(2v),$$

$$f_{|u_n)(u_m|}(v) = h_{mn}(v).$$

From this last example we see that h_{mn} [given for instance by Eq. (3.31′)] is the classical function corresponding to the projection onto the subspace $\mathbb{C}u_n$. For the special case $\nu = 1$, this is the projection onto the nth eigenspace of the harmonic oscillator (a similar expression, obtained in a different way, can be found in Ref. 19).

2. Multiplication operators by holomorphic functions

Consider $A : \psi \to F.\psi$, where F is some (holomorphic) function such that $F.\Omega^a \in \mathscr{L}_0$ for any a.

Then

$$A(a,b) = F(a)\Omega^b(a),$$

and

$$f_A(v) = \kappa^{-1}\int \Omega(2v-2b)\,F(b)\,db;$$

hence,

$$f_A = \kappa^{-1}(\Omega_2 * F), \quad \Omega_2(v) = \Omega(2v).$$

3. Permutation operators

Suppose $E = E_1 \oplus \dots \oplus E_n$, with $JE_j \subset E_j, \forall j, \sigma(E_j,E_k) = 0$ for $j \neq k$: this is the phase space for the simultaneous description of n particles ($\dim E_j = 2\nu'$ for any j; $n\nu' = \nu$).

Let $\{e_1^j,\dots,e_{2\nu'}^j\}$ be a symplectic base in each E_j. For any $\pi \in P(1,\dots,n)$ [$P(1,\dots,n)$ is the set of all permutations of $(1,\dots,n)$], we define

$$P_\pi : E \to E,$$

$$\sum_{j=1}^{n}\sum_{k=1}^{2\nu'} v_k^j e_k^j \to \sum_{j=1}^{n}\sum_{k=1}^{2\nu'} v_k^{\pi j} e_k^j;$$

$$Q_\pi : \mathscr{L}_0 \mapsto \mathscr{L}_0,$$

$$\phi(v) \to \phi(P_\pi(v)).$$

Clearly,

$$Q_\pi(a,b) = \Omega^b(P_\pi(a)).$$

To compute the classical function corresponding to Q_π, we split up π into a product of independent cyclic permutation operators. The classical function splits up in a product of independent functions, corresponding to these cyclic permutations. For the cyclic permutation $\pi = (1,\dots,m)$ (this permutation maps 1 to 2, 2 to 3,..., $m-1$ to m, m to 1), we get

$$f_\pi(v_1,\dots,v_m) = 2^{\nu'(m-1)}\prod_{j=1}^{(1/2)(m-1)} \times \exp[i\varphi(V_{2j-1}, P_\pi^{-1}(v_{2j}), P_\pi^{-2}(v_{2j+1}))] \quad \text{for } m \text{ odd},$$

and

$$f_\pi(v_1,\dots,v_m) = 2^{\nu'(m-1)}\prod_{j=1}^{(m/2)-1} \times \exp[i\varphi(V_{2j-1}, P_\pi^{-1}(v_{2j}), P_\pi^{-2}(v_{2j+1}))]\delta(V_m),$$

for m even, with

$$V_k = \sum_{j=1}^{k} (-1)^{j-k}(P_\pi)^{j-k}(v_j).$$

In particular, if we describe two particles, and we want to consider the operator Q_π for $\pi 1 = 2$, $\pi 2 = 1$, we have

$$f_{(12)}(v_1,v_2) = \delta(v_1 - P_\pi(v_2)).$$

For three particles, we see that

$$f_{(123)} = 2^{2\nu}\exp[i\varphi(v_1,P_\pi^{-1}(v_2),P_\pi^{-2}(v_3))]$$

and

$$f_{(12)(3)} = \delta(v_1 - P_\pi(v_2)).$$

These different expressions can be considered as special cases of the classical functions corresponding to general symplectic transformations computed in Refs. 28 and 25.

6. A COMPARISON WITH BARGMANN'S INTEGRAL TRANSFORM REF. 17

In Ref. 17 some explicit expressions are given for the unitary operator intertwining the Schrödinger representation with the coherent state representation of the Weyl commutation relations. We rewrite here this result in our notations.

Identify E with $\mathbb{R}^{2\nu} = \mathbb{R}^{\nu} \oplus \mathbb{R}^{\nu} = x$ space $\oplus\, p$ space. Let us denote the x space by E_1. In what is usually called the Schrödinger representation the Hilbert space used is $L^2(E_1)$, i.e., the space of square integrable functions on E_1 with respect to a Lebesgue measure on E_1.

Bargmann's integral transform is a unitary map A from $L^2(E_1)$ to $\mathscr{L}_0(E, J)$ which can be represented by a kernel:

$$A : L^2(E_1) \to \mathscr{L}_0(E,J),$$

$$\forall \psi \in L^2(E_1) : (A\psi)(v) = \int dx\, A(v,x)\, \psi(x). \tag{6.1}$$

The kernel $A(v, x)$ has many interesting properties. For fixed x, it is an element of Z, and for fixed v it is square integrable on E_1. This is analogous with our kernel $\{\zeta\,|v\}$ which for fixed ζ is square integrable on E, and for fixed v an element of $Z\{(E\oplus E; (J, -J)\}$. Moreover, we know (see Sec. 4.E) that our integral transform Q is unitary from $L^2(E)$ on $\mathscr{L}_0(E\oplus E, (J, -J))$. So it would seem that our integral transform is just a double Bargmann transform:

$$A : L^2(E_1) \to \mathscr{L}_0(E, J),$$

$$Q: L^2(E) \simeq L^2(E_1) \otimes L^{2*}(E_1) \to \mathscr{L}_0(E, J) \otimes \mathscr{L}_0(E, J)^* \simeq \mathscr{L}_0(E\oplus E,(J, -J)).$$

We donote here by $\mathscr{H}^*$ the dual of $\mathscr{H}$; the isomorphy $\mathscr{L}_0(E, J) \otimes \mathscr{L}_0(E,J)^* \simeq \mathscr{L}_0(E\oplus E,(J, -J))$ follows from the fact that $\mathscr{L}_0(E\oplus E,(J, -J))$ is isomorphic to the Hilbert space of Hilbert–Schmidt operators on $\mathscr{L}_0(E, J)$ (see Sec. 4.E). It is however not altogether true that Q is just twice A. Indeed, on has

$$A(v,x) = \sum_{[m]} u_{[m]}(v)\phi_{[m]}(x), \tag{6.2}$$

where $u_{[m]} = h^{[m]}\Omega$ and $\phi_{[m]}$ are the eigenfunctions of the harmonic oscillator, respectively, in $\mathscr{L}_0(E, J)$ and $L^2(E_1)$[17]; on the other hand [see Eq. (3.29)],

$$\{a,b\,|(x_v,p_v)\} = \sum_{[m][n]} u_{[m]}(a)\, \overline{u_{[n]}(b)}\, h_{[m][n]}(x_v,p_v), \tag{6.3}$$

where $h_{[m][n]}(v)$ is given by Eq. (3.29) and is definitely different from $\phi_{[m]}(2^{1/2}x_v)\overline{\phi_{[n]}(2^{1/2}p_v)}$ (the factor $\sqrt{2}$ has to be introduced because of a difference in normalization in the measures on E_1 and E). This can readily be checked in an example. Take $\nu = 1$, $m = n = 1$. Then

$$\phi_1(2^{1/2}x_v)\overline{(\phi_1(\sqrt{2}p_v)} = x_v p_v e^{-(x_v^2+p_v^2)}$$

and

$$h_{11}(x_v,p_v) = 2\, e^{-(x_v^2+p_v^2)(-x_v^2-p_v^2+1)} \neq \phi_1(\sqrt{2}x_v)\overline{\phi_1(\sqrt{2}p_v)}.$$

Another way of seeing that the integral transform Q is not merely a double Bargmann transform is to look at the explicit expressions for the kernels $A(v, x)$ and $\{\zeta\,|v\}$. We have[17]

$$A(v,x) = \Pi^{-\nu/4} e^{(i/2)x_v p_v} e^{-ip_v x} e^{-(1/2)(x-x_v)^2} = \Omega^{(x_v,p_v)}(x), \tag{6.4}$$

where $\Omega^{(x_v,p_v)}(x)$ is the coherent state centered round (x_v,p_v), written in the x representation, while a direct calculation from Eq. (3.9) gives

$$\{a,b\,|(x_v,p_v)\} = \Omega^{(x_a+x_b)/\sqrt{2},(p_b-p_a)/\sqrt{2}}(\sqrt{2}x_v) \times \overline{\Omega^{(p_a+p_b)/\sqrt{2},(x_b-x_a)/\sqrt{2}}(\sqrt{2}p_v)}, \tag{6.5}$$

which is again very different from the expected

$$A(a,\sqrt{2}x_v)\overline{A(b,\sqrt{2}p_v)} = \Omega^{(x_a,p_a)}(\sqrt{2}x_v)\overline{\Omega^{(x_b,p_b)}(\sqrt{2}p_v)}. \tag{6.6}$$

In a certain sense these differences are due to the fact that the integral transform Q has to do with quantization, while A is just a unitary map from one quantum mechanical realization to another. Indeed, if we look at Eqs. (6.2) and (6.3), we see that on the $\mathscr{L}_0$ side everything is all right: Eq. (6.3) contains one straight copy $u_{[m]}(a)$ and one complex conjugate copy $\overline{u_{[n]}(b)}$ of the $\mathscr{L}_0$ function $u_{[m]}(v)$ in Eq. (6.2); but things go wrong with the L^2 function. This is precisely because $L^2(E)$, the initial space of Q, has to be considered as a space of classical functions, while the initial space of A is a quantum Hilbert space.

Another way of seeing this is the following: By taking a double Bargmann transform one treats x_v and p_v as two equivalent but independent ("commuting") variables: in Eq. (6.6) x_v is only linked with x_a, p_a, and p_v only with x_b, p_b. However, this is not what happens in a quantization procedure; there x_v and p_v are linked with the x_a, p_a as well as with x_b, p_b: some mixing has taken place.

ACKNOWLEDGMENT

One of the authors (I.D.) would like to thank the CPT, CNRS Marseille for hospitality.

[1] H. Weyl, *Theory of Groups and Quantum Mechanics* (English translation) (Dover, New York, 1950).
[2] J. von Neumann, Math. Ann. **104**, 570 (1931).
[3] I.E. Segal, Math. Scand. **13**, 31 (1963).

[4]I.E. Segal, "Representation of the canonical commutation relations," in *Applications of Mathematics to Problems in Theoretical Physics*, Cargèse Lecture Notes, edited by F. Lurçat, (Gordon Breach, New York, 1967).
[5]D. Kastler, Commun. Math. Phys. **1**, 14 (1965).
[6]R.B. Lavine, "The Weyl transform: Fourier analysis of operators in L^2-spaces, Ph.D. thesis, MIT (1965).
[7]A. Grossmann and P. Huguenin, Helv. Phys. Acta **51**, 252 (1978).
[8]J.C. Pool, J. Math. Phys. **7**, 66 (1966).
[9]A. Grossmann, Commun. Math. Phys. **48**, 191 (1976).
[10]P. Huguenin, Z. Naturforsch. **28**, 1090 (1968). *Mécanique de l'espace de phase*, Cours 3^e cycle (Neuchâtel, 1975).
[11]L.G. Suttorp and S.R.de Groot, Nuovo Cimento A **65**, 245 (1970); S.R.de Groot and L.G. Suttorp, *Foundations of Electrodynamics* (North-Holland, Amsterdam, 1972) Chaps. 6 and 7; S.R.de Groot, *La Transformation de Weyl et la Fonction de Wigner: une Forme Alternative de la Mécanique Quantique* (Université de Montréal, 1974).
[12]K.E. Cahill and R.J. Glauber, Phys. Rev. **177**, 1857 (1969).
[13]C.L. Mehta, Phys. Rev. Lett. **18**, 752 (1967).
[14]J.R. Klauder and E.C. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968), Chap. 7.
[15]E. Schrödinger, Naturwissenshaften **14**, 644 (1927).
[16]J.R. Klauder, Ann. Phys. (N.Y.) **11**, 123 (1960).
[17]V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961).
[18]I.E. Segal, Ill. Math. **6**, 500 (1962).
[19]F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, Ann. Phys. (Leipzig) **111**, 61, 111 (1978).
[20]J.R. Klauder, J. McKenna, and D.G. Currie, J. Math. Phys. **6**, 733 (1965).
[21]J.R. Klauder, Phys. Rev. Lett. **16**, 534 (1965).
[22]R. Gautrin, J. Math. Phys. **15**, 1686 (1974).
[23]N. Mukunda and E.C.G. Sudarshan, Prumana **10**, 227 (1978); E.C.G. Sudarshan and N. Mukunda, J. Math. Phys. **20**, 1801 (1979).
[24]A. Grossmann and R. Seiler, Commun. Math. Phys. **48**, 195 (1976).
[25]I. Daubechies, "Projective respresentation of the linear canonical transformations," VUB preprint, Brussels.
[26]I.M. Gel'fand and G.E. Shilov, *Generalized Functions* (Academic, New York, 1967), Vol. 2.
[27]A.Grossmann, J. Math. Phys. **6**, 54 (1965).
[28]P. Huguenin, Lett. Math. Phys. **2**, 321 (1978).
[29]G. Loupias, Sur le formalisme de la convolution gauche, thèse à l'Université d'Aix-Marseille (1966).

Nature of superspace[a)]

Krystyna Bugajska[b)]
Institute of Physics, Silesian University, Katowice, Poland

(Received 14 November 1978; accepted for publication 14 September 1979)

It is demonstrated that the superspace of the supersymmetry theory can be identified in a natural manner with a family of concrete spinor structures over space–time.

1. INTRODUCTION

The theories based on the Fermi–Bose supersymmetry invariance are recently a subject of considerable investigations. One introduces in these theories the concepts of superspace, and supersymmetry group. However, no deeper significance is given to the superspace itself, and the question about its interpretation is still open. The purpose of this paper is to propose an interpretation of superspace and superfields, and discuss it in some details. The interpretation is derived from the notion of spinor structure on a space–time manifold, hence we begin with a short discussion of this.

The starting point of our consideration is a condition of existence of a spinor structure on space–time E formulated by Crumeyrolle.[1] It should be stressed that the notion of spinor structure over E is to be understood here in a less common, but a very natural way as a possibility to attach half-spinor spaces $\Sigma(m)$ and $\Sigma^*(m)$ to every point m of E in a continuous manner.[2] The necessary and sufficient condition for this is the existence of a global field $f(m)$ of isotropic bivectors on E, generated by a family of real orthonormal tetrads.

This condition is equivalent to a reducibility of the bundle $\xi_{\mathscr{L}_0}$ of orthonormal tetrads to $\xi_{\mathscr{C}}$, where $\mathscr{C}$ (the Crumeyrolle group) is a two-parameter Abelian subgroup of the Lorentz group $\mathscr{L}_0$, topologically equivalent to $\mathbb{R}^2$. The reducibility of $\xi_{\mathscr{L}_0}$ to $\xi_{\mathscr{C}}$ assures the existence of a global cross section of $\xi_{\mathscr{L}_0}$, hence, according to the known[3] Geroch result, the existence of the SL(2,$\mathbb{C}$)-bundle $\xi_{\mathrm{SL}(2,\mathbb{C})}$ being the prolongation of $\xi_{\mathscr{L}_0}$. We see then that the Crumeyrolle condition is equivalent to the more known Milnor–Lichnerowicz one.[4]

On the other hand, the reduction of $\xi_{\mathrm{SL}(2,\mathbb{C})}$ to the bundle $\xi_{\mathscr{C}}$ defines a global field of isotropic bivectors $f(m)$, hence also[1] the spinor spaces $\Sigma_0(m)$ and $\Sigma_0^*(m)$ at each point m of E. These spinor spaces are odd and even parts of the invariant under SL(2,$\mathbb{C}$) decomposition of the minimal left ideal $C'f(m)$ of the complexified Clifford algebra $C'(m)$. The sum $\cup_{m\in E}\Sigma_0(m)$ will be denoted by $\mathscr{E}_0$ and called a concrete spinor structure over the space–time E.

We can pick up another field $f'(m)$ of isotropic bivectors as well another minimal left ideal $C'f'(m)$ of the Clifford algebra $C'(m)$. This means that we can attach to each point of E another spinor space $\Sigma_\theta(m)$ [$\Sigma_\theta^*(m)$] in a continuous way[5]. The resulting concrete spinor structure $\cup_{m\in E}\Sigma_\theta(m)$ will be denoted $\mathscr{E}_\theta$.

It is evident that one can define a lot of different, although equivalent, concrete spinor structures $\mathscr{E}_\theta$ over E, provided the space–time admits at least one such a structure. Hence some new degrees of freedom appear. To study them we observe that the above considerations lead to a one-to-one correspondence between elements of SL(2,$\mathbb{C}$)/$\widetilde{\mathscr{C}}$ and concrete spinor structures $\mathscr{E}_\theta$. On the other hand, it is known[6] that the homogeneous space SL(2,$\mathbb{C}$)/$\widetilde{\mathscr{C}}$ can be parametrized by elements

$$\theta=\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}, \theta^+(\theta_{\dot 1},\theta_{\dot 2}) \quad (\text{where } \theta_{\dot\alpha}=\theta_\alpha^*)$$

of a two-dimensional complex space, which transform as spinors under the Poincaré group, Thus these additional degrees of freedom lead to a richer than E structure parametrized by (x^μ,θ,θ^+). The main goal of this paper is to demonstrate that just this structure is the superspace as introduced[7] and considered by Wess and Zumino, Salam and Strathdee, and others.

Let us observe that our family of different spinor structures emerges also inside the Milnor–Lichnerowicz approach. Indeed, the structure group SL(2,$\mathbb{C}$) of the bundle $\xi_{\mathrm{SL}(2,\mathbb{C})}$ can be considered as a $\widetilde{\mathscr{C}}$-bundle over SL(2,$\mathbb{C}$)/$\widetilde{\mathscr{C}}$. The Lie algebra of the group $\widetilde{\mathscr{C}}$ defines a two-dimensional spinor space by the Cartan–Whittaker construction.[8] So we have at each point $m\in E$ a family of different spinor spaces, one space for one element of SL(2,$\mathbb{C}$)/$\widetilde{\mathscr{C}}$.

Coming back to the parameters θ, it is demonstrated in Sec. 2 that if we fix one spinor space, say $\Sigma_0(m)$, at point $m\in E$, then any other spinor space $\Sigma_\theta(m)$ is represented in a natural way by one and only one element u_θ of $\Sigma_0(m)$ (the so-call pure spinor, introduced by Cartan, and Chevalley[9]) with $u_\theta=u_0+\theta=s_\theta u_0$, where u_0 represents $\Sigma_0(m)$, and $s_\theta\in$SL(2,$\mathbb{C}$).

On the other hand, we can make use of the known[10] isomorphism between the Hermitian part of any $\Sigma_\theta(m)\otimes\Sigma_\theta^*(m)$ and the tangent space T_mE. If we pass now from $\Sigma_\theta(m)$ to $\Sigma_{\theta+\epsilon}(m)$ we see that it has to be accompanied by a translation in the affine tangent space A_mE. The translation vector z is determined by the condition:

$$x_\theta+x_\epsilon+z=x_{\theta+\epsilon}.$$

A simple calculation leads to

$$z=\theta\otimes\epsilon^*+\epsilon\otimes\theta^*. \qquad (1.1)$$

[a)]Supported by Polish Ministry of Science, Higher Education and Technology, Project MR I 7.

[b)]Postal address: K. Bugajska, Instytut Fizyki U. Sl., ul. Uniwersytecka 4, 40-007 Katowice, Poland.

This construction can be done at any point m of E, hence some global problem arises: The labelling spinor u_θ should be consistently defined for the whole spinor structure $\mathscr{E}_\theta$. To do this we fix a global cross section $r_0(m) = \{e_0,e_1,e_2,e_3\}_m$ of the bundle $\xi_{\mathscr{C}}$ (the existence of such a cross section is assured by topological properties of $\mathscr{C}$). This cross section defines a global field of canonical spinor frames $\{\rho,\sigma\}_m$ of $\Sigma_0(m)$. Now, because the pure spinor $u_0(m)$ representing $\Sigma_0(m)$ has the same coordinates in any canonical frame $\{\rho,\sigma\}_m$, we can identify spinors $u_\theta(m) = s_\theta u_0(m)$ at each point m of E by their coordinates in the frame $\{\rho,\sigma\}_m$ corresponding to m. Thus only one spinor u_θ is associated with the spinor structure $\mathscr{E}_\theta$.

A transition from $\mathscr{E}_\theta$ to $\mathscr{E}_{\theta+\epsilon}$ leads to the transformation

$$(x^\mu,\theta,\theta^+)\rightarrow(x^\mu+\theta^+\sigma^\mu\epsilon+\epsilon^+\sigma^\mu\theta,\theta+\epsilon,\theta^++\epsilon^+), \tag{1.2}$$

which is essentially a consequence of the translation introduced by (1.1), written down in the frames $r_0(m)$. Because any element (Λ,a) of the covering group $\widetilde{\mathscr{P}}_0$ of the Poincaré group acts in the following manner[6]:

$$(\Lambda,a)(x,\theta,\theta^+) = [\chi(\Lambda)x + a,\Lambda\theta,\theta^+\Lambda^+], \tag{1.3}$$

we obtain superspace $\mathscr{M}$ parametrized by (x^μ,θ,θ^+) with the action of the graded Lie algebra spanned by Poincaré generators $P_\mu, L_{\mu\nu}$ together with, anticommuting to the moments, operators generating transformations (1.2) defined on it.

The paper is organized as follows: In Sec. 2 we consider general properties of a spinor structure on the space–time E. Invetigations leading from the family of concrete spinor structures to the superspace are placed in Sec. 3. Section 4 is devoted to a study of superfields.

2. SPINOR STRUCTURE ON SPACE–TIME

The spinor structure on space–time is usually defined (following Milnor and Lichnerowicz[4]) as a prolongation of the Lorentzian structure $\xi_{\mathscr{L}_0}$ on space–time E to the spinor group SL(2,$\mathbb{C}$). If such a prolongation $\xi_{SL(2,\mathbb{C})}$ there exists, we can construct the associated bundle $\xi_{SL(2,\mathbb{C})}[\Sigma]$, where Σ is a two-dimensional complex vector space equipped with the skew bilinear form

$$\epsilon_{\alpha\beta} = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}.$$

Hence we can attach to each point m of E a spinor space $\Sigma(m)$ given by the fibre of this vector bundle.

However, we can reverse the question and look for conditions which allow us to attach to each point of the space–time a two-dimensional spinor space in a continuous way. It has been shown[1] that the necessary and sufficient condition for this is a reduction of the bundle of orthonormal frames $\xi_{\mathscr{L}_0}$ to the bundle $\xi_{\mathscr{C}}$, where the two-parameter Abelian group $\mathscr{C}$ is generated by

$$A_1 = L_{02} + L_{23}, \quad A_2 = L_{01} - L_{31},$$

(where $L_{\mu\nu}$ are the standard generators of the Lorentz group). The spinor space $\Sigma(m)$ at the point m is given by a minimal left ideal of the Clifford algebra of the tangent space T_mE.

It is well known that one can define a vector field (or a tensor field) on any manifold E, because the notion of tangent space at each point $m\in E$ does not require any additional assumptions about E. When the manifold E is additionally equipped with a nondegenerate metric tensor g of signature $(+,-,-,-)$, then a possibility of the existence of spinor fields appears. Namely we can construct the Clifford bundle $C(E,g)$, each fibre of which is the Clifford algebra $C(T_mE, Q_m)$ (here Q_m is the quadratic form over the tangent space T_mE defined by the metric tensor g_m at m). We shall denote the Clifford algebra $C(T_mE,Q_m)$ by C_m. If the Clifford algebra C_m is simple (it is true[10] for any even-dimensional manifold) then its finite dimensional irreducible representations are already given by its minimal left ideals. The elements of any minimal left ideal will be called spinors.

Thus the space–time E, as a 4-manifold which carries a smooth global Lorentzian tensor field g, allows the construction of the Clifford bundle $C(E,g)$, what enables us to define the spinor spaces S_m at each point m of E. To build up such a space S_m let us fix an orthonormal frame $r_0 = \{e_0,e_1,e_2,e_3\}$ of T_mE. This frame defines the Witt base ω_0 of the complexification $(T_mE)'$ of the vector space T_mE:

$$\omega_0 = \{x_1,x_2,y_1,y_2\}$$
$$= \{\tfrac{1}{2}(e_0+e_3),\tfrac{1}{2}(ie_1+e_2),\tfrac{1}{2}(e_0-e_3),\tfrac{1}{2}(ie_1-e_2)\}, \tag{2.1}$$

as well the isotropic bivector

$$f = y_1y_2 \tag{2.2}$$

(we will denote the Clifford product as above, although in this case it reduces to the exterior product). In this manner the spinor space S_m given by the tetrad r_0 is equal to

$$S_m = C'f. \tag{2.3}$$

(C' and Q' will denote the complexifications of $C = C_m$ and $Q = Q_m$ respectively. The obvious index m will be omitted.) The dimension of the spinor space S is equal to four, and we can represent it in one and only one way as a sum of two subspaces:

$$S = \Sigma \oplus \Sigma^* \tag{2.4}$$

of so called odd and even half-spinors respectively.

The spinor space $S = \Sigma\oplus\Sigma^*$ is the underlying vector space of the irreducible representation of the Clifford group G.[9] Hence the space S is an underlying space of corresponding representations of subgroups of G, such as Pin(1,3), Spin(1,3), and Spin$_+$(1,3)[1], which form the following chain:

$$G \xrightarrow[|N(s)|=1]{\text{Module norm condition}} \text{Pin}(1,3) \xrightarrow{\text{parity condition}} \text{Spin}(1,3) \xrightarrow[N(s)=1]{\text{norm condition}} \text{Spin}_+(1,3). \tag{2.5}$$

The representations of G and Pin(1,3) are irreducible, whereas the representation of Spin(1,3) is reducible to the sum of two inequivalent irreducible representations acting on the half-spinor spaces Σ and Σ^*. The same is true for the group Spin$_+$(1,3)$\simeq$SL(2,$\mathbb{C}$). The representation of Spin$_+$(1,3) on the half-spinor space Σ has an additional property: We can construct the bilinear invariant skew form $\epsilon_{\alpha\beta}$ on Σ. (The same holds for Σ^*.)

Every isotropic bivector f given by the orthonormal frame according to (2.2) defines a four-dimensional spinor space S with the Clifford group G as the symmetry group (it means that S is equipped with all G-equivalent spinor frames). This does not determine, however, the structure of S with a sufficient accuracy. Fortunately there are at our disposal the mentioned subgroups of the Clifford group (2.5), and we can define the spinor space as the vector space with one of these subgroups as the symmetry group. For physical reasons we prefer the symmetry group which preserves the decomposition of the space S onto half-spinor spaces Σ and Σ^*. This condition indicates the group Spin(1,3), which is the covering group of SO(1,3), as a possible candidate. But this group is still not good enough, as it does not preserve the (skew) scalar product in Σ (Σ^*). So we arrive at the spinor space as the complex vector space with $\mathrm{Spin}_+(1,3) \simeq \mathrm{SL}(2,\mathbb{C})$ as its symmetry group.

Two isotropic bivectors $f = y_1 y_2$ and $f' = y_1' y_2'$ define the same minimal left ideal of the algebra C' (that is $C'f = C'f'$) if and only if $f' = \lambda f$ where λ is a complex number[1]. Hence[1], if only f and f' define the same minimal left ideal of C', then they have to define the same maximal totally isotropic subspace of $(T_m E)'$.

Let us fix a maximal totally isotropic subspace (m.t.i.s.) of $(T_m E)'$, for example the subspace Z_0 spanned by y_1 and y_2 defined by (2.1). (We recall that y_1 and y_2 define $f = y_1 y_2$, hence the spinor space $S = C'f$.) It appears[11] that every totally isotropic two-dimensional subspace Z of $(T_m E)'$ can be represented by an element of S.

Indeed, let f_Z be the product of the elements of some base of Z. As we have told, f_Z is determined by Z up to a scalar factor different from zero. Then $f_Z C'$ is a minimal right ideal of C'. Because the intersection of any minimal left ideal of C' with any minimal right ideal is a one-dimensional vector subspace of S[9], we shall call any element of this subspace a representative spinor of Z (or pure spinor).

It is known[11] that a representative spinor of any totally isotropic subspace Z of $(T_m E)'$ is always a half-spinor. We are interested only in such m.t.i. subspaces Z of $(T_m E)'$ which are related to the Witt bases associated with real orthonormal tetrads by (2.1), because we are interested in the spinor (half-spinor) representation of the Minkowski space, but not of its complexification [from now on we shall restrict ourselves only to such subspaces Z of $(T_m E)'$]. So we take into account only such m.t.i. subspaces Z which are linked to Z_0 by an element of $\mathscr{L}_0$. All such Z's are represented by two even half-spinors belonging to Σ^* instead of the one-dimensional subspace $C'f_Z \cap f_{Z_0} C'$. Moreover, if u_Z is a representative spinor for Z, then su_Z is a representative spinor for $\chi(s)Z$, where $s \in \mathrm{SL}(2,\mathbb{C})$ and χ is the covering map, $\chi: \mathrm{SL}(2,\mathbb{C}) \to \mathscr{L}_0$.

Any orthonormal (real) frame r defines an isotropic bivector f by (2.2) and a m.t.i.s. Z spanned by y_1, y_2 [see formula (2.1)]. The same Z can be, however, defined in this manner by a whole class of orthonormal frames. If the symmetry group of the related spinor space S is the Clifford group G or the group Pin(1,3), then this class contains frames linked one to another by transformations of $\mathscr{L} = \mathrm{O}(1,3)$.[12] If we want the symmetry group of S to be Spin(1,3), then between frames of the class defining Z act transformations of SO(1,3). In this case the class into consideration is the image, under the covering map, of the subgroup of elements $s \in \mathrm{Spin}(1,3)$ fulfilling the condition:

$$sfs^{-1} = \pm f. \tag{2.6}$$

We limit our attention to the physically most important case of $\mathrm{Spin}_+(1,3) \simeq \mathrm{SL}(2,\mathbb{C})$. The considered class of frames determining the same m.t.i.s. Z is given now by the condition

$$sfs^{-1} = f, \tag{2.7}$$

where $s \in \mathrm{Spin}_+(1,3)$, instead of (2.6). This subgroup of $\mathrm{Spin}_+(1,3)$ will be denoted $\widetilde{\mathscr{C}}$ and called the Crumeyrolle group. Because the image of $\widetilde{\mathscr{C}}$ under the covering map: $\chi(\widetilde{\mathscr{C}}) = \mathscr{C} \simeq \widetilde{\mathscr{C}}$ leaves Z invariant, we can define a bijection from the family of all m.t.i subspaces into the homogeneous space $\mathscr{L}_0/\mathscr{C}$.

An orthonormal frame r can be used to construct, besides f and m.t.i.s. Z as above, half-spinor spaces Σ and Σ^* together with canonical spinor frames: $\{\rho,\sigma\} := \{x_1 f, x_2 f\}$ in Σ, and similar frame $\{\rho^*,\sigma^*\}$ in Σ^*. Any other frame from the class $r\mathscr{C}$ defines the same Σ (and Σ^*) with a $\mathscr{C}$-equivalent canonical spinor frame. Hence it is natural, and will be advantageous, to understand the spinor space as the linear space Σ together with the class of all $\mathscr{C}$-equivalent canonical frames. So, for example, we should distinguish between the two spinor spaces: Σ equipped with canonical frame $\{\rho,\sigma\}$, and Σ_- equipped with canonical frame $\{-\rho, -\sigma\}$, because these canonical frames are not $\mathscr{C}$-equivalent as canonical frames. We can consider these spinor spaces as defined by f and $-f$ respectively.

Let us take one arbitrary spinor space Σ_0^* (Σ_0) related to an orthonormal frame r_0 and a m.t.i.s. Z_0. Recall that any m.t.i.s. Z can be represented by two elements of Σ_0^*. If the nonzero spinors representing the m.t.i.s. Z_0 are denoted by $\pm u_0$, then any other m.t.i.s. $Z = \chi(x)Z_0$ is represented by spinors $\pm su_0$, $s \in \mathrm{SL}(2,\mathbb{C})$. The distinction we made between Σ and Σ_- allows us to represent Σ_0 by $u_0 \in \Sigma_0^*$, and Σ_{0-} by $-u_0$, and similarly every other Σ by su_0, and Σ_- by $-su_0$, with $s \in \mathrm{SL}(2,\mathbb{C})$. We can represent the spinor spaces by elements of Σ_0 instead of Σ_0^*, making use of the known antiisomorphism: $\Sigma^* \longleftrightarrow \Sigma$. Because the set of different m.t.i.s. $Z \subset (T_m E)'$ related to orthonormal frames in $T_m E$ is in the one-to-one correspondence with $\mathscr{L}_0/\mathscr{C}$, we obtain that the set of different spinor spaces Σ (Σ^*) at each point m of the space-time E is in a one-to-one correspondence with $\mathrm{SL}(2,\mathbb{C})/\mathscr{C}$.

Now let us summarize the arising picture: the quadratic form defined by g_m allows us to construct $\mathscr{L}_0/\mathscr{C}$ in number different m.t.i. subspaces Z of $(T_m E)'$ related to orthonormal real frames of $T_m E$, and $\mathrm{SL}(2,\mathbb{C})/\mathscr{C}$ in number different half-spinor spaces Σ (Σ^*). Choosing one frame r_0 we can construct an isotropic bivector f by (2.2), m.t.i.s. Z_0, and half-spinor spaces Σ_0 and Σ_0^* together with their canonical spinor frames $\{\rho,\sigma\}$ and $\{\rho^*,\sigma^*\}$. Any other half-spinor space Σ including Σ_0 itself is represented by some element of Σ_0. If Σ_0 is represented by u_0, then Σ_{0-} is represented by

$-u_0$, and any other Σ is represented by $su_0 \in \Sigma_0$ with $s \in SL(2,\mathbb{C})$, such that $r_0\chi(s^{-1}) = r$ and the frame r defines Σ. We want to point out that the group $\mathscr{C}$ which leaves invariant the m.t.i. subspace Z is also the stabilizer group of the homogeneous SL(2,ℂ),-space Σ.

The m.t.i. subspaces are a redundant element in the presented derivation of spinor spaces Σ and Σ^* from a given frame r of $T_m E$. However their vital importance manifests itself in our construction of the one-to-one correspondence between different spinor spaces at a given point of E and different elements of one of these spaces. This correspondence, in turn, will be essential in the next section. Besides, the considerations of m.t.i. subspaces could be of some interest in the general instanton problem, especially in the presence of the Ward observation, and result of Atiyah and others.

3. FROM SPINOR STRUCTURE TO SUPERSPACE

Section 2 demonstrated that, at each point of space–time, we can construct just as many half-spinor spaces $\Sigma(m)$ as there are elements of $SL(2,\mathbb{C})/\mathscr{C}$ and the same number of spaces $\Sigma^*(m)$. Moreover, if we fix one of them, say $\Sigma_0(m)$, then every one space $\Sigma(m)$ is represented by an element of $\Sigma_0(m)$. The space $\Sigma_0(m)$ itself is represented by spinor $u_0 \in \Sigma_0(m)$ which has components $\binom{0}{1}$ in the canonical base $\{\rho,\sigma\} = \{x_1 f, x_2 f\}$ associated with orthonormal frame $r_0 = \{e_0, e_1, e_2, e_3\}$. We shall denote by $\Sigma_\theta(m)$ the spinor space at the point $m \in E$ represented by $u_0 + \theta \in \Sigma_0(m)$. Making use of the canonical base we can define an element s_θ of SL(2,ℂ) such that

$$s_\theta u_0 = u_\theta := u_0 + \theta. \tag{3.1}$$

The spinor space $\Sigma_\theta(m)$ is related with the frame $r = r_0 \times \chi(s_\theta^{-1})$

As we know,[10] the Hermitian part of the tensor product $\Sigma_\theta(m) \otimes \Sigma^*_\theta(m)$ is isomorphic to the tangent space $T_m E$ for any choice of the spinor space $\Sigma_\theta(m)$ at the point $m \in E$. However, if we extend the Lorentz group $\mathscr{L}_0$ to the Poincaré group $\mathscr{P}_0$ as the symmetry group of the tangent space, then we should consider this tangent space as an affine space. Hence we should relate with any spinor space $\Sigma_\theta(m)$ a vector x_θ as an origin of the affine tangent space $A_m E$. It is natural to associate the vector 0 with the space Σ_0. It implies that the vector x_θ associated with Σ_θ has the form:

$$x_\theta = u_\theta \otimes u^*_\theta - u_0 \otimes u^*_0. \tag{3.2}$$

So we obtain a correctly defined isomorphism between the Hermitian part of $\Sigma_\theta(m) \otimes \Sigma^*_\theta(m)$ and $A_m E$.

Now observe, that any spinor ϵ of $\Sigma_0(m)$ defines a translation in $\Sigma_0(m)$

$$\epsilon(\theta) := \theta + \epsilon, \tag{3.3}$$

hence generates a transformation in the set of all spinor spaces given by:

$$\epsilon(\Sigma_\theta(m)) := \Sigma_{\theta+\epsilon}(m). \tag{3.4}$$

Moreover, the translation by ϵ in $\Sigma_0(m)$ generates a transformation of the corresponding points in $A_m E$,

$$\epsilon(x_\theta) := x_{\theta+\epsilon}. \tag{3.5}$$

But $x_{\theta+\epsilon}$ is not equal to $x_\theta + x_\epsilon$. If we want to preserve the previous correspondence (3.2) between the spinors which label the different spinor spaces and vectors associated with these spaces: $\epsilon \to x_\epsilon, \theta \to x_\theta, \theta + \epsilon \to x_{\theta+\epsilon}$, etc., then the translation (3.3) in $\Sigma_0(m)$ should be always accompanied by a translation of the affine tangent space. This translation is determined by a vector z calculated from

$$x_\theta + x_\epsilon + z = x_{\theta+\epsilon}. \tag{3.6}$$

We find easily that

$$z = \theta \otimes \epsilon^* + \epsilon \otimes \theta^*. \tag{3.7}$$

Hence we see that the transition from Σ_0 to Σ_ϵ does not cause any transformation in the affine tangent space, but if one goes from Σ_θ to $\Sigma_{\theta+\epsilon}$ with $\theta \neq 0$, one must accompany it with the appropriate translation in $A_m E$. These considerations can be visualized by associating with any spinor space Σ_θ the family $\{x_\theta, r_0 \mathscr{L}_0\}$ of affine frames in the affine tangent space $A_m E$.

Let us consider now some global aspects of the problem of existence of different concrete spinor structures $\mathscr{C}_\theta$ over E. The necessary and sufficient condition for the existence of spinor structure is the existence of a global field $f(m)$ of isotropic bivectors on E, related to a family of real orthonormal tetrads over E. This family consists of elements of $\xi_{\mathscr{C}}$ (the appropriate reduction of the bundle $\xi_{\mathscr{L}_0}$). The topological properties of the group $\mathscr{C}$, mentioned in Section 1, make it possible to pick up a global field $r_0(m)$ of orthonormal tetrads belonging to $\xi_{\mathscr{C}}$ (a global cross section of $\xi_{\mathscr{C}}$). We see now that we can attach the m.t.i.s. $Z_0(m)$ and half-spinor spaces $\Sigma_0(m)$ ($\Sigma^*_0(m)$) together with their canonical spinor frames $\{\rho,\sigma\}_m$ ($\{\rho^*,\sigma^*\}_m$) to every point m of E in a smooth manner. Then we obtain a concrete spinor structure $\mathscr{C}_0$ defined by the global field $f_0(m)$ of isotropic bivectors determined by $r_0(m)$, by (2.1), and (2.2). Any half-spinor space $\Sigma_0(m) \in \mathscr{C}_0$ is represented, according to Sec. 2, in itself by the spinor $u_0(m)$, which in the canonical frame $\{\rho,\sigma\}_m$ has coordinates independent of m. So we can attach to the whole $\mathscr{C}_0$ one element u_0 of the fibre Σ_0 of the vector fibre bundle $\mathscr{C}_0$. The concrete spinor structure $\mathscr{C}_{0-}$ defined by the field $-f(m)$ will be represented in this Σ_0 by spinor $-u_0$.

Let us take now $s \in SL(2,\mathbb{C})$, such that $s \notin \mathscr{C}$. The global cross section $r_0(m)\chi(\pm s^{-1})$ will define new concrete spinor structures $\mathscr{C}$ and $\mathscr{C}_-$ represented in Σ_0 by su_0 and $-su_0$ respectively. The concrete spinor structure $\mathscr{C}$ will be labelled by $\theta \in \Sigma_0$, with

$$su_0 = u_0 + \theta. \tag{3.8}$$

So the global fields of isotropic bivectors sfs^{-1} and $-sfs^{-1}$, generated by the global cross section $r_0(m) \times \chi(\pm s^{-1})$ gives concrete spinor structures $\mathscr{C}_\theta$ and $\mathscr{C}_{\theta-}$ respectively.

Proceeding in this way we attach to any element of $SL(2,\mathbb{C})/\mathscr{C}$ a concrete spinor structure $\mathscr{C}_\theta$ labelled by an element θ of the fibre Σ_0 of the vector fibre bundle $\mathscr{C}_0$. As a result we obtain an extension of the space–time E with additional degrees of freedom provided by elements of $SL(2,\mathbb{C})/\mathscr{C}$. This extension has virtually a structure of the assoicated bundle $\xi_{SL(2,\mathbb{C})}[SL(2,\mathbb{C})/\mathscr{C}]$. The transformations of the group $\widetilde{\mathscr{P}}_0$ act on this new space, parametrized by (x,θ,θ^+) in the following way:

$$(\Lambda,a)(x,\theta,\theta^+) = [\chi(\Lambda)x + a,\Lambda\theta,\theta^+\Lambda^+], \qquad (3.9)$$

what has been essentially shown by Bacry and Kihlberg.[6] Owing to the above discussion [and formulas (3.6) and (3.7)], we see that a transition from one concrete spinor structure $\mathscr{C}_\theta$ to another one, say $\mathscr{C}_{\theta+\epsilon}$, leads to the transformation:

$$(x^\mu,\theta,\theta^+)\rightarrow(x^\mu + \theta^+\sigma^\mu\epsilon + \epsilon^+\sigma_\mu\theta,\theta+\epsilon,\theta^+ + \epsilon^+), \qquad (3.10)$$

where

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma^1 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$
$$\sigma^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (3.11)$$

The obtained structure is manifestly the superspace $\mathscr{M}$, while the transformation above is the supersymmetry.

4. SUPERFIELDS

From the point of view assumed above we see that, roughly speaking, the superspace is just space–time equipped with all concrete spinor structures. The labels of these spinor structures are exactly the new parameters, widely exploited in supersymmetry theories[7,13].

Taking the superspace $\mathscr{M}$ as a background of a physical field theory, we will consider the simplest case of a scalar field $\phi(x,\theta,\theta^+)$ on $\mathscr{M}$. The field ϕ undergo the following transformation under a finite supersymmetry T_ϵ:

$$T_\epsilon\phi(x,\theta,\theta_+) = \phi(x^\mu + \epsilon^+\sigma^\mu\theta + \theta^+\sigma^\mu\epsilon,\theta+\epsilon,\theta^+ + \epsilon^+); \qquad (4.1)$$

or in the infinitesimal form:

$$\delta\phi = \left(\epsilon\frac{\partial}{\partial\theta} + \epsilon^+\frac{\partial}{\partial\theta^+} + (\epsilon^+\sigma^\mu\theta + \theta^+\sigma^\mu\epsilon)\frac{\partial}{\partial x^\mu}\right)\phi. \qquad (4.2)$$

Hence we obtain infinitesimal generators:

$$Q_\alpha = \frac{\partial}{\partial\theta^\alpha} + (\theta^+\sigma^\mu)_\alpha\frac{\partial}{\partial x^\mu}, \qquad (4.3)$$

$$\bar{Q}_\alpha = \frac{\partial}{\partial\theta^{+\alpha}} + (\sigma^\mu\theta)_\alpha\frac{\partial}{\partial x^\mu}. \qquad (4.4)$$

We can check[7,13] that:

$$\{Q_\alpha,Q_\beta\} = \{\bar{Q}_\alpha,\bar{Q}_\beta\} = 0, \qquad (4.5)$$

$$\{Q_\alpha,\bar{Q}_\beta\} = 2\sigma^\mu_{\alpha\beta}P_\mu, \qquad (4.6)$$

and

$$[Q_\alpha,P_\mu] = [P_\mu,P_\nu] = 0, \qquad (4.7)$$

where P_μ are generators of translations of the Poincaré group. Thus the graded Lie algebra (GLA) of operators P_μ, $L_{\mu\nu}$, Q_α, and $\bar{Q}_\alpha$ acts infinitesimally on our superspace.

Salam and Strathdee[7] have shown that irreducible representations (defining supermultiplets) of this GLA can be worked out by means of the Wigner method of induced representations. To every such representation corresponds a superfield which decomposes, according to this representation, into a sum of finite number of scalar terms describing particles of different spin. A similar decomposition appears if we regard the spinor parameters as Grassmannian ones in order to obtain the exponentiation of the anticommuting elements of the GLA.[14] The Grassmannian character of the spinor parameters can be also observed if we (according to Ref. 13) identify the Minkowski space coordinates as even elements of order two, and the spinor parameters as odd elements of order one of the Grassmann algebra of the spinor space.

Calculating $\delta\phi$ according to (4.2) we see that the supersymmetry turns boson fields into fermion ones, and vice versa. (Superfields, which we can regard as linked to supermultiplets, contain particles of different spin.)

Besides scalar superfields we also consider "vector" superfields $\phi_\mu(x,\theta,\theta^+)$ or "spinor" ones $\phi_\alpha(x,\theta,\theta^+)$ with the appropriate transformation properties under Poincaré group:

$$\phi'_\mu(x',\theta',\theta'^+) = \chi(\Lambda)^\nu_\mu\phi_\nu(x,\theta,\theta^+),$$
$$\phi'_\alpha(x',\theta',\theta'^+) = \Lambda^\beta_\alpha\phi_\beta(x,\theta,\theta^+),$$

and so on.

If we disregard the fact of existence of many different (nevertheless equivalent) concrete spinor structures on E, then superfields collapse into usual physical fields, as the parameter θ loses its meaning. Thus we see that the superfields are more fundamental then the conventional fields, because the later are results of neglecting the additional degrees of freedom.

If space–time does not admit a spinor structure (there is no prolongation of $\xi_{\mathscr{L}_0}$ to $\xi_{SL(2,C)}$, or there is no reduction of $\xi_{\mathscr{L}_0}$ to $\xi_{\mathscr{C}}$) then the superspace $\mathscr{M}$ cannot be globally constructed. Nevertheless a local trivialization of the bundle $\xi_{\mathscr{L}_0}$ allows us to make such a construction locally over any element U_i of an open covering $\{U_i\}_{i\in I}$ trivializing $\xi_{\mathscr{L}_0}$. Let $\mathscr{M}_i$ and $\mathscr{M}_j$ be two such "local" superspaces over U_i and U_j respectively. However the transition from a concrete "local" spinor structure defined over $U_i\cap U_j$ by $\mathscr{M}_i$ to a concrete "local" spinor structure defined over the same set by $\mathscr{M}_j$ is described generally by a transformation belonging to the Clifford group.

ACKNOWLEDGMENT

It is a pleasure to thank Professor A. Trautman for helpful discussions and constant encouragement.

[1]A. Crumeyrolle, Ann. Inst. H. Poincaré **11**, 19 (1969); A. Crumeyrolle, C.R. Acad. Sci. Paris **271**, 1070 (1970); see also: K. Bugajska, "Spinor structure on space–time," to appear in Int. Theor. Phys., and K. Bugajska, "On geometrical properties of spinor structure," J. Math. Phys. **21**, xxxx (1980).

[2]Here $\Sigma(m)$ is an odd half-spinor space, usually denoted by $S_o(m)$, whereas $\Sigma^*(m)$ is an even half-spinor space, usually denoted by $S_e(m)$, naturally anti-isomorphic to $\Sigma(m)$.

[3]R. Geroch, J. Math. Phys. **9**, 1739 (1968).

[4]J. Milnor, L'Enseignement mathématique **9**, 198 (1963); A. Lichnerowicz, *Topics on space–time*, in "Battelle Recontres, 1967 Lectures in Mathematics and Physics," edited by C.M. de Witt and J.A. Wheeler (Benjamin., New York, Amsterdam, 1968), Chapter V, p. 115.

[5]In the fibre-bundle language it is equivalent to the reduction of the structure group of the bundle $\xi_{SL(2,C)}$ to the group $s_0\mathscr{C}s_0^{-1}$ with $s_0\in SL(2,C)$. As the reduction of $\xi_{SL(2,C)}$ to $\xi_{\mathscr{C}}$ is given by a cross section σ of the associated

bundle $\xi_{SL(2,\mathbb{C})}$ [SL(2,C)/$\mathscr{C}$], then the reduction to $\xi_{s_\theta \tilde{\mathscr{C}} s_\theta^{-1}}$, equivalent to it, is given by the cross section $s_\theta \sigma$. $\tilde{\mathscr{C}}$ denotes the covering group of $\mathscr{C}$, $\tilde{\mathscr{C}} \simeq \mathscr{C}$.

[6] H. Bacry and A. Kihlberg, J. Math. Phys. **10**, 2132 (1969).

[7] J. Wess and B. Zumino, Nucl. Phys. B **78**, 1 (1974); A. Salam J. Strathdee, Nucl. Phys. B **76**, 477 (1974); A. Salam, J. Strathdee, Nucl. Phys. B. **80**, 499 (1974); S. Ferrara, J. Wess, B. Zumino, Phys. Lett. B **51**, 239 (1974); S. Ferrara and B. Zumino, Nucl. Phys. B **79**, 413 (1974).

[8] K. Bugajska, "On geometrical properties of spinor structure," J. Math. Phys. **21**, 2097 (1980).

[9] E. Cartan, *The Theory of Spinors* (Hermann, Paris, 1954); C. Chevalley, *The Algebraic Theory of Spinors* (Columbia University, N.Y., 1954).

[10] R. Penrose, *Structure of space–time*, in "Battelle Recontres, 1967 Lectures in Mathematics and Physics" edited by C.M. de Witt and J.A. Wheeler (Benjamin, New York - Amsterdam, 1968), p. 121; K. Bugajska, "Geometrical interpretation of spinors," to appear in Int. J. Theor. Phys.

[11] C. Chevalley, Ref. 9.

[12] Any element of the Clifford group G. defines an action on $T_m E$ belonging to O(1,3) = $\mathscr{L}$, see, e.g., Ref. 1.

[13] L. Corvin, Y. Ne'eman, and S. Sternberg, Rev. Mod. Phys. **47**, 573 (1975).

[14] F.A. Berezin and G.I. Katz, Matemat. Sb. USSR **82**, 343 (1970); A. Pais and V. Rittenberg, J. Math. Phys. **16**, 2062 (1975).

On geometrical properties of spinor structure[a)]

Krystyna Bugajska[b)]
Institute of Physics, Silesian University, Katowice, Poland

(Received 14 November 1978; accepted for publication 14 September 1979)

The Crumeyrolle group $\mathscr{C}$ for four-dimensional space-time E is explicitely calculated. It is shown that the complexification of the Lie algebra of the group $\mathscr{C}$ is a spinor space. In this manner the condition of the existence of a spinor structure over E, formulated as the reduction of the structure group of the bundle of orthonormal frames to $\mathscr{C}$, enables us to associated the spinor space to each point of space-time in a continuous way.

1. INTRODUCTION

Milnor, Lichnerowicz, Bichteler, and Penrose were the first to define the notion of a spinor structure on a Riemannian, and pseudo-Riemannian manifolds. They have also provided a necessary and sufficient condition for the existence of such a structure[1]: A space and time orientable space–time manifold E carries a spinor structure if and only if the second Stiefel–Whitney class of E vanishes. It has been demostrated by Geroch[2] that this condition, for a noncompact space–time, is equivalent to the existence of a global field of orthonormal tetrads on E.

There is, however, another approach to this problem, based on the Clifford algebras, developed by Chevalley[3] and Crumeyrolle.[4] This approach provides us with another (but equivalent to the previous one[5]) necessary and sufficient condition for the existence of a spinor structure on a space–time: A space and time orientable space–time manifold E carries a spinor structure if and only if the structure group $\mathscr{L}_0$ of the bundle $\xi_{\mathscr{L}_0}$ of orthonormal tetrads over E is reducible[6] to a group $\mathscr{C}$ which we will call the Crumeyrolle group.

One of the goals of the present paper is the explicit calculation of the Crumeyrolle group. This is done in Sec. 3. Before this we give a short exposition of the Clifford algebras approach to the spinor structure (Sec. 2).

In Sec. 4 we use the known relation[7] between generators of the proper Lorentz group $\mathscr{L}_0$ and antisymmetric tensors to demonstrate that the Lie algebra $\mathscr{h}$ of the Crumeyrolle group $\mathscr{C}$ is spanned by two Cartan–Whittaker[8] tensors. Applying the Cartan–Whittaker construction we obtain that $\mathscr{h}$ is spanned by two spinors, say u and v. This result leads in a natural manner to the following geometrical interpretation of the mentioned Crumeyrolle condition: The reducibility of $\xi_{\mathscr{L}_0}$ to the Crumeyrolle group $\mathscr{C}$ allows us to attach (in a continuous way) to each point of the space–time manifold a two-dimensional complex space spanned by the spinors u and v.

It is worth to note that the Crumeyrolle group $\mathscr{C}$ appeared also in quite different circumstances. Namely, it has been demonstrated by Finkelstein, Bacry, and others[9] that the spin degree of freedom is connected to certain coordinates in a homogeneous space of the Lorentz group. The only case which cannot be realized as a rigid system of space–time points is obtained when the corresponding homogeneous space has just the group $\mathscr{C}$ as its stabilizer group.[10] We comment on this topic in Sec. 5.

2. SPINOR FIELDS

In the general case of a pseudo-Riemannian space–time manifold E we cannot consistently define spinor fields.[11] To make it possible we have to attach to each point of E a two-dimensional spinor space (half-spinor space) in a continuous way.[5]

It is known[3,4] that the metric tensor field g on E allows us to construct the Clifford bundle $C(E,g)$ over E. We can define a quadratic form Q_m on the tangent space T_mE at a point m of E as $Q_m(x) = g_m(x,x)$, where $x \in T_mE$, and g_m denotes the metric tensor at m. The fibre of $C(E,g)$ over m is the Clifford algebra C_m of the tangent space T_mE. Let $\{e_0, e_1, e_2, e_3\}$ be a base of T_mE.

Definition 2.1: The Clifford algebra C_m of the given quadratic form Q_m has the underlying vector space spanned by the elements denoted by $e_{i_1}e_{i_2}\cdots e_{i_k}$ with an increasing sequence of integers $i_1, i_2, \cdots, i_k$ between 0 and 3, and by unit 1. The multiplication law (the Clifford multiplication) is determined by the form Q_m:

$$yz = y \wedge z + \tfrac{1}{2}B_m(y,z)\cdot 1\,, \qquad (2.1)$$

for any $y,z \in T_mE$, where $B_m(y,z)$ is a bilinear form associated with Q_m:

$$B_m(y,z) = Q_m(y+z) - Q_m(z) - Q_m(y)\,, \qquad (2.2)$$

and $\wedge$ denotes the exterior product.

It follows from (2.1) that

$$x^2 = Q_m(x)\cdot 1\,. \qquad (2.3)$$

It is easy to see that the exterior algebra $\wedge\, T_mE$ may be identified with the Clifford algebra of the zero form on T_mE.

Definition 2.2: By the spinor space S_m at a point m of E we understand the vector space of the unique, up to equivalence, irreducible representation of the Clifford algebra C_m.

The representation mentioned in Definition 2.2 is given by the minimal left ideal of the complexification C'_m of C_m.

Let $(T_mE)'$, Q'_m, C'_m denote the complexifications of T_mE, Q_m, and C_m respectively. Now every orthonormal frame $r_0 = \{e_0, e_1, e_2, e_3\}$ of T_mE allows us to introduce the

[a)]Supported by Polish Ministry of Science, Higher Education and Technology, Project MR I 7.

[b)]Postal address: K. Bugajska, Institut Fizykiu. Sl., Ul. Uniwersytechka 4, 40-007 Katowice, Poland.

Witt base $\omega_0 = \{x_1, x_2, y_1, y_2\}$ of $(T_m E)'$ given by:

$$x_1 = \tfrac{1}{2}(e_0 + e_3), \quad y_1 = \tfrac{1}{2}(e_0 - e_3),$$
$$x_2 = \tfrac{1}{2}(ie_1 + e_2), \quad y_2 = \tfrac{1}{2}(ie_1 - e_2). \tag{2.4}$$

Linear spaces: N, spanned by $\{x_1, x_2\}$, and P, spanned by $\{y_1, y_2\}$, are totally isotropic subspaces of $(T_m E)'$. They are supplimentary to each other, and

$$B'_m(x_i, x_j) = B'_m(y_i, y_j) = 0,$$
$$B'_m(x_i, y_j) = \delta_{ij}, \tag{2.5}$$

where B'_m is defined by (2.2) with Q'_m in the place of Q_m. We set $f = y_1 y_2$. Now $C'_m f$ is the minimal left ideal of C'_m. Let C^N and C^P be the subalgebras of C'_m generated by N and P respectively. Then we have: $C'_m f = C^N f$. We define a representation of C'_m on C^N by

$$(\rho(v)u)f = vuf, \text{ if } v \in C'_m, \quad u \in C^N. \tag{2.6}$$

Now, because $C'_m f$ is the minimal left ideal of C'_m, ρ is simple. The spinor space S_m appears to be identical with C^N (see Chevalley, Ref. 3, pp. 42 and 70),

$$S_m = C^N = C'_m f. \tag{2.7}$$

The spin space S_m is not only the underlying vector space of the simple representation of the Clifford algebra of the quadratic form $Q_m(x) = g_m(x,x)$, but also the space of representations of some groups contained in C'_m, and strictly connected with Q_m. Let $C^* \subset C'_m$, be the multiplicative group of invertible elements of C'_m. The Clifford group G is defined as the subgroup of C^* consisting of elements $s \in C^*$ such that

$$sxs^{-1} \in T_m E \quad \forall x \in T_m E. \tag{2.8}$$

Because $Q_m(sxs^{-1}) \cdot 1 = (sxs^{-1})^2 = Q_m(x) \cdot 1$, we conclude from (2.8) that there exists a natural mapping $\varphi: G \to O(1,3)$, with $\varphi(s)$ denoting the linear automorphism $x \to sxs^{-1}$ of $T_m E$. In fact we have: $\varphi(G) = O(1,3)$ with the kernel isomorphic to the group GL(1) of nonzero real numbers. There are some subgroups of G which are the covering groups of appropriate subgroups of the Lorentz group $\mathscr{L} = O(1,3)$ (see Fig. 1).

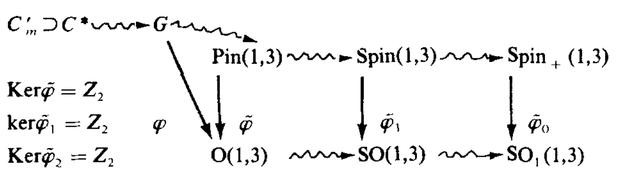


FIG. 1

In the middle row of the diagram of Fig. 1 there are the covering groups of the corresponding subgroups of $\mathscr{L}$ with covering mappings $\tilde{\varphi}, \tilde{\varphi}_1, \tilde{\varphi}_0$ (which are the appropriate restrictions of φ).[12] When we restrict ourselves to the group Spin(1,3) as a symmetry group of the spinor space S_m, then the spinor representation ρ is a sum of two inequivalent irreducible representations on the spaces of two-spinors (half-spinors), S_o and S_e, where S_o is spanned by $\{x_1 f, x_2 f\}$, or (because of 2.7) by $\{x_1, x_2\}$, and S_e is spanned by $\{f, x_1 x_2 f\}$, or by $\{1, x_1 x_2\}$. The spinor space $S_m = S_o \oplus S_e$ has the canonical base $\mathscr{S} = \{f, x_1 f, x_2 f, x_1 x_2 f\}$.

Now we wish to consider half-spinors as the basic quantities. This means that we should be able to obtain the space $T_m E$ from the tensor product $S_o S_e$ in such a way that for any transformation of S_m given by $s \in \text{Spin}(1,3)$, the space $T_m E$ will be transformed by $\varphi(s) \in SO(1,3)$. We have only one possibility: to assume that the $\text{Spin}_+(1,3)$-group is the symmetry group of our spinor space.[5]

So from the above considerations we see that every orthonormal frame $r_0 = \{e_0, e_1, e_2, e_3\}$ at m makes possible a construction of the minimal left ideal $C'_m f$ as well as the spaces S_o and S_e, which are the odd and the even parts of the invariant under SL(2,ℂ) decomposition of this minimal left ideal. More precisely, see Fig. 2.

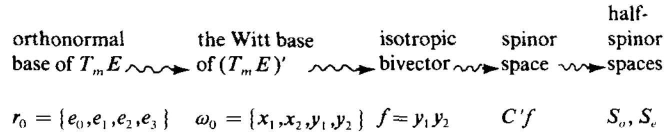


FIG. 2

Up to now we have defined the half-spinor spaces S_o and S_e at a-point m of E starting with a fixed orthonormal frame r_0 at m. Another basis r will define the same half-spinor spaces if and only if[5] there exists an element s of $\text{Spin}_+(1,3)$, such that $\varphi(s): r_0 \to r$, and

$$sfs^{-1} = f. \tag{2.9}$$

The last condition is equivalent to

$$sf = f. \tag{2.10}$$

The general isotropic bivector f' defining a minimal left ideal of C'_m is given by

$$sfs^{-1} = f' \neq \lambda f, \tag{2.11}$$

or

$$sfs^{-1} = f' = \lambda f, \tag{2.12}$$

where $\lambda \in \mathbb{R}$, $\lambda \neq 0$, and s belongs to the Clifford group G. In the case of (2.11) we obtain another spinor space S'_m, corresponding to the minimal left ideal $C'_m f' \neq C'_m f$, which obviously leads to another half-spinor spaces S'_o, S'_e. In the case of (2.12) we obtain the same spinor space S_m, but with the whole Clifford group as its symmetry group.

Hence two Lorentz observers r_0 and r at $m \in E$ can consistently introduce their own spinor spaces starting with r_0 and r respectively (according to Fig. 2), if and only if they are related by $\varphi(s)$, with s given by (2.9). In the opposite case the spinor bases: $\mathscr{S} = \{y_1 y_2, x_1 y_1 y_2, x_2 y_1 y_2, x_1 x_2 y_1 y_2\}$, and $\mathscr{S}' = \{y'_1 y'_2, x'_1 y'_1 y'_2, x'_2 y'_1 y'_2, x'_1 x'_2 y'_1 y'_2\}$ defined by r_0 and r respectively, will be bases of essentially different spinor spaces. On the other hand, the two observers can observe the same spinor space, say S_m, defined by r_0, and the only difference will be in different spinor bases introduced by them in S_m. So, if r_0 has introduced the base $\mathscr{S}$ and $r = r_0 \varphi(s) = r_0 \varphi(-s)$, then the basis $\mathscr{S}' = \pm s\mathscr{S}$ is connected to the observer r. (The Lorentz frame r defines $\mathscr{S}'$ only up to the sign.)

Definition 2.3: The elements s of the group $\text{Spin}_+(1,3)$ which satisfy the condition (2.9) form the subgroup of $\text{Spin}_+(1,3)$ which we shall denote by $\tilde{\mathscr{C}}$ and call the Crumeyrolle group. The group $\tilde{\varphi}_0(\tilde{\mathscr{C}})$ will be denoted by $\mathscr{C}$.

Now let us return to the problem of spinor fields in space–time. Because the bundle of orthonormal tetrads $\xi_{\mathscr{L}}$ (the Lorentz structure) is locally trivial, there is an open covering $\{U_i\}_{i\in I}$ of the space–time manifold with local fields of orthogonal tetrads h_i (here h_i is a local crossection over U_i given by the local trivialization). This enables us to construct spinor spaces over U_i in a continuous way. But we know that spinor spaces given by $h_i(m)$ and $h_j(m)$ for $m\in U_i\cap U_j\neq\emptyset$ will be the same only if the frame $h_i(m)$ can be obtained from $h_j(m)$ by means of an operation belonging to $\mathscr{C}$. It does not mean anything else but the fact that the transition functions of the principal bundle $\xi_{\mathscr{L}}$ must take their values from $\mathscr{C}$. However it is well known[6] that this is the necessary and sufficient condition for the reducibility of the structure group $\mathscr{L}$ of the Lorentz structure to $\mathscr{C}$.

To summarize: We can associate half-spinor spaces with any point m of E in a continuous way iff the space–time manifold is time and space orientable [that means that the Lorentz bundle $\xi_{\mathscr{L}}$ is reducible to $\mathscr{L}_0 = \mathrm{SO}_+(1,3)$], and the structure group of $\xi_{\mathscr{L}}$ is reducible to $\mathscr{C}$. When these conditions hold, we can consider spinor fields on E as basic quantities from which, by reduction of the tensor product of an appropriate number of spinor fields, we obtain every physical spinor, vector, or tensor fields (see, e.g., Penrose, Ref. 1, Geroch[2]).

3. THE CRUMEYROLLE GROUP $\mathscr{C}$

It is obvious that the Crumeyrolle group of Q_m will be the same for every point m of the space–time E (it depends only on the signature of the quadratic form Q_m). Therefore, we can and will now look for the explicit form of the group $\mathscr{C}$ for the abstract Minkowski space M. We know that for every Lie group G any subalgebra ℓ of its Lie algebra $\mathscr{g}$ is the Lie algebra of exactly one connected Lie subgroup G' of G.[13] For that reason we shall look for the Lie algebra $\mathscr{k}$ of $\mathscr{C}$.

A. Lie algebras of groups C*, G, and Spin+(1,3)

By C^* we have denoted the multiplicative group of invertible elements of the Clifford algebra C. Let us recall that on the Clifford algebra of the quadratic form $(+,-,-,-)$ we can define the natural vector-space topology. It allows us to treat the group C^* as a Lie group. Let us find its Lie algebra.

Every element $x\in C$ defines an endomorphism of the vector space of C given by the left Clifford product F_x: $u\to xu$ for every $u\in C$. Because the dimension of C is equal to $2^4 = 16$, F_x is represented by a matrix $2^4\times 2^4$. The exponential map is well defined for any matrix, so we have

$$\exp F_x = \sum_{k=0}^{\infty} \frac{1}{k!}(F_x)^k. \tag{3.1}$$

The map F of C into Hom (C,C) defined by: $x\to F_x$ for every $x\in C$ is a homeomorphism[3] between C and a subspace of the vector space of endomorphisms of the linear structure underlying C. Hence we can define the element $\exp x$ of the Clifford algebra C as such one for which we have

$$\exp F_x = F_{\exp x}\,. \tag{3.2}$$

Since for commuting elements x, y of C

$$\exp(x+y) = (\exp x)(\exp y)\,, \tag{3.3}$$

we see that $\exp x$ is an invertible element of C, and $(\exp x)^{-1} = \exp(-x)$. Now let us introduce the Lie product in C

$$[x,y] = xy - yx \quad \text{for each } x,y\in C \tag{3.4}$$

(xy is the Clifford product). From

$$F_{[x,y]} = [F_x, F_y] \tag{3.5}$$

we see that we may regard C as the Lie algebra of C^*.

Inner automorphisms of the group C^* define the adjoint representation of this group on its Lie algebra C. Indeed, an automorphism $\varphi_s: C^*\to C^*$, $s\in C^*$ given by

$$\varphi_s(s') = ss's^{-1} \quad \text{for each } s'\in C^*\,, \tag{3.6}$$

induces the transformation $(\varphi_s)_*: C\to C$ with

$$(\varphi_s)_* x = sxs^{-1} \quad \text{for each } x\in C\,, \tag{3.7}$$

and for the adjoint representation we have

$$(\varphi_{\exp})_* y = \exp(\mathrm{ad}x)y\,, \tag{3.8}$$

where $(\mathrm{ad}x)y = [x,y]$.

Let us recall that the Minkowski space M can be considered as a subspace of C, and the Clifford group G is the subgroup of C^* consisting of all such $s\in C^*$ that $(\varphi_s)_* M = M$. Thus elements X of the Lie algebra of the Clifford group G will satisfy

$$(\varphi_{\exp} X)_* y = \exp(\mathrm{ad}X)y\in M \quad \text{for each } y\in M\,. \tag{3.9}$$

It follows immediately that X belongs to the Lie algebra of G when $\mathrm{ad}X$ maps M into itself. It can be shown[3] that such a Lie algebra is spanned by elements e_ie_j, $i<j$, and by the unit 1. It can also be demonstrated that the Lie algebra of $\mathrm{Spin}_+(1,3)$ is spanned[3] by products e_ie_j, $i<j$ (e_i, e_j are vectors of M orthogonal to each other).

Now we demonstrate that the products e_ie_j are closely related to infinitesimal operators of one parameter subgroups of rotations and boosts. For $i,j = 1,2,3$ we have

$$2e_i\cdot e_j = B(e_i, e_j) = 2\delta_{ij} \tag{3.10}$$

(the dot denoting the scalar product). Now if we take (2.1) into account, then for every $x, y\in M$ we have

$$xy + yx = B(x,y)\cdot 1\,. \tag{3.11}$$

It implies the following result for matrix elements of e_ie_j, $i,j = 1,2,3,\quad i<j$:

$$\begin{aligned}(e_ie_j)_{kl} &= \tfrac{1}{2}B(e_k, (\mathrm{ad}e_ie_j)e_l)\\ &= \tfrac{1}{2}B(e_k, e_ie_je_l - e_le_ie_j)\\ &= 2(\delta_{ki}\delta_{jn} - \delta_{kj}\delta_{li})\,.\end{aligned} \tag{3.12}$$

Now for $i=0$ we have $B(e_0, e_j) = 2\delta_{0j}$, $j = 1,2,3$; and

$$\begin{aligned}(e_0e_j)_{kl} &= \tfrac{1}{2}B(e_k, (\mathrm{ad}e_0e_j)e_l)\\ &= 2(-\delta_{k0}\delta_{jl} + \delta_{kj}\delta_{0l})\,.\end{aligned} \tag{3.13}$$

Summarizing:

$$\tfrac{1}{2}(e_ie_j)_{kl} = -g_{ik}\delta_{jl} + g_{il}\delta_{jk}\,, \tag{3.14}$$

with $g_{ij} = \mathrm{diag}(+,-,-,-)$.

Let us define

$$L_{ij} = \tfrac{1}{2}ie_ie_j\,. \tag{3.15}$$

Commutation relations for L_{ij} are

$$[L_{ij}, L_{kl}] = ig_{jk}L_{il} - g_{jl}L_{ik} - g_{ik}L_{jl} + g_{il}L_{jk}\ , \tag{3.16}$$
$$L_{ij} = -L_{ji}\ , \quad i,j = 0,1,2,3\ .$$

We conclude that we have obtained the Lie algebra of the Lorentz group with L_{23}, L_{31}, L_{12} as infinitesimal operators of the one parameters subgroups of rotations, and L_{01}, L_{02}, L_{03} as infinitesimal operators of proper Lorentz transformations along e_i.

B. The Crumeyrolle group

Coming back to the group $\mathscr{C}$ we observe, that the condition (2.10) implies that if X belongs to the Lie algebra $\tilde{\mathscr{h}}$ of $\mathscr{C}$, then

$$Xf = 0\ , \tag{3.17}$$

with f introduced earlier, $f = y_1 y_2$. Because X belongs to the Lie algebra of Spin$_+$ (1,3), it must satisfy conditions

$$(\mathrm{ad}X)y \in M \quad \text{for each } y \in M\ , \tag{3.18}$$

and

$$\alpha(X) + X = 0\ , \tag{3.19}$$

where α is the so called main antiautomorphism[3] of C, and $\alpha(zy) = yz$ for $z,y \in M$. From the condition (3.18) we obtain that X has to be real, hence coefficients A_{ij} of the decomposition $X = \Sigma_{i<j} A_{ij} e_i e_j$ must be real,

$$A_{ij} = \bar{A}_{ij}\ . \tag{3.20}$$

On the other hand, the general element X of considered Lie algebras has the form

$$X = a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2 + by_1y_2\ , \tag{3.21}$$

where x_1, x_2, y_1, y_2 are the elements of the Witt base introduced earlier (see Fig. 2). From (3.20) we obtain

$$X = a_{11}x_1y_1 + \bar{b}x_2y_1 + by_1y_2\ , \tag{3.22}$$

with a_{11} real. But

$$a_{11}(\alpha(x_1y_1) + x_1y_1) = a_{11}\ , \tag{3.23}$$

so $a_{11} = 0$, and

$$\begin{aligned} X &= \bar{b}x_2y_1 + by_1y_2 \\ &= c_1(x_2y_1 + y_1y_2) + ic_2(x_2y_1 - y_1y_2) \\ &= \lambda_1(e_0e_2 + e_2e_3) + \lambda_2(e_0e_1 + e_1e_3), \quad \lambda_1, \lambda_2 \ \text{real}\ . \end{aligned} \tag{3.24}$$

We see that the Lie algebra of the Crumeyrolle group is spanned by

$$X_1 = e_0e_2 + e_2e_3 \quad \text{and} \quad X_2 = e_0e_1 + e_1e_3\ . \tag{3.25}$$

It can be easily checked that $[X_1, X_2] = 0$. From (3.15) we obtain that X_1 could be identified with $A_1 = L_{02} + L_{23}$, whereas X_2 could be identified with $A_2 = L_{01} - L_{31}$.

The Iwasawa decomposition of $\mathscr{L}_0$ has the form

$$\mathscr{L}_0 = \mathscr{K}\mathscr{A}\mathscr{N}\ , \tag{3.26}$$

where $\mathscr{K}$ is the maximal compact subgroup SO(3), $\mathscr{A}$ is the Abelian one-parameter subgroup generated by the acceleration L_{03}, and $\mathscr{N}$ is the nilpotent Abelian two-dimensional subgroup generated by A_1 and A_2 defined above. Thus the Crumeyrolle group is identical with $\mathscr{N}$.

4. SPINOR STRUCTURE

In the previous section we have determined the Lie algebra $\tilde{\mathscr{h}}$ of $\tilde{\mathscr{C}}$ which, of course, is isomorphic to the Lie algebra $\mathscr{h}$ of the subgroup $\mathscr{C}$ of the Lorentz group $\mathscr{L}_0$. Because for every subalgebra of a Lie algebra there exists exactly one connected Lie subgroup, we see that $\mathscr{C}$ is defined uniquely by $\mathscr{h}$.

It is known that every generator of the Lorentz group can be expressed as

$$X = F^{ik}L_{ik}\ , \tag{4.1}$$

where F^{ik} is an antisymmetric tensor (each tensor F^{ik} corresponding to a single element of the Lie algebra of $\mathscr{L}_0$. We can establish a connection between a complex 3-vector $\mathbf{F} = \mathbf{B} + i\mathbf{E}$ and a skew tensor F^{ik} by

$$\mathbf{F} = (F^{23} + iF^{01}, F^{31} + iF^{02}, F^{12} + iF^{03})\ , \tag{4.2}$$

which establishes a connection between the vector $\mathbf{B} + i\mathbf{E}$ and an element $F^{ik}L_{ik}$ of the Lie algebra of $\mathscr{L}_0$.

From (3.25) we see that the Lie algebra $\mathscr{h}$ is spanned by elements connected to vectors:

$$\mathbf{F}_1 = (1, i, 0)\ , \quad \mathbf{F}_2 = (i, -1, 0)\ . \tag{4.3}$$

These vectors obey the condition:

$$\mathbf{F}_i \cdot \mathbf{F}_i = 0\ , \quad i = 1, 2\ , \tag{4.4}$$

which implies that the two invariants of a skew-symmetric tensor F^{ik}:

$$F_{ik}F^{ik} = \mathbf{B}^2 - \mathbf{E}^2 \tag{4.5}$$

and

$$\mathscr{E}_{ijkl}F^{ij}F^{kl} = 2\mathbf{B}\cdot\mathbf{E} \tag{4.6}$$

vanish for the tensors connected to $\mathbf{F}_1$ and $\mathbf{F}_2$.

Now let us return to the problem of spinor fields on the space–time manifold E. We have already seen that such fields can be defined when the Lorentz group $\mathscr{L}_0$ of the bundle of orthonormal frames $\xi_{\mathscr{L}_0}$ over E is reducible to $\mathscr{C}$. On the other hand, this means that to each point $m \in E$ we attach the Lie algebra $\mathscr{h}$ spanned by two mutually orthogonal "null" complex 3-vectors $\mathbf{F}_1$ and $\mathbf{F}_2$. It is known,[8] however, that for every skew-tensor F^{ik} related to such "null" vectors by (4.2) (the Cantor–Whittaker tensor) there corresponds a complex 2-vector $w = \binom{w^0}{w^1}$ given by

$$\begin{aligned} (w^0)^2 &= \tfrac{1}{2}(F^{02} - F^{23} + i(F^{01} + F^{13}))\ , \\ (w^1)^2 &= \tfrac{1}{2}(F^{02} + F^{23} - i(F^{01} + F^{13}))\ , \\ w^0w^1 &= \tfrac{1}{2}(F^{12} - iF^{03})\ . \end{aligned} \tag{4.7}$$

Moreover,[8] if such a tensor F^{ik} with vanishing invariants of (4.5) and (4.6) was transformed by an element $\tilde{\varphi}_0(s) \in \mathscr{L}_0$, then the related spinor w would be transformed by $s \in \mathrm{SL}(2,\mathbb{C})$.

Now we should check the action of the Lorentz group $\mathscr{L}_0$ on its Lie algebra. To do this let us take into account that the Lorentz group is isomorphic to SO(3,$\mathbb{C}$) in the following way[7]: $\boldsymbol{\sigma}\cdot\mathbf{F}$ is a traceless matrix with determinant equal to $\mathbf{F}^2$ for any complex 3-vectors $\mathbf{F} = \mathbf{B} + i\mathbf{E}$, where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli 2×2 matrices. Any element s of SL (2,$\mathbb{C}$) acts linearly on such matrices:

$$s : \sigma\cdot\mathbf{F} \rightarrow s(\sigma\cdot\mathbf{F})s^{-1}, \qquad (4.8)$$

so we have the relation

$$s(\sigma\cdot\mathbf{F})s^{-1} = \sigma\cdot(R\,\mathbf{F}) \quad \text{with } R\in SO(3,\mathbb{C}). \qquad (4.9)$$

The last formula expresses the action of the Lorentz group on its Lie algebra through the convention of (4.1) and (4.2). Thus we see that F^{ik} is transformed by elements of the Lorentz group, hence the element w related to it by (4.7) is transformed by the elements of the group SL $(2,\mathbb{C})$.

In this manner we have obtained the following picture: The condition of the existence of a spinor structure over the space–time E, i.e., the reducibility of the Lorentz structure $\xi_{\mathscr{L}_0}$ to the group $\mathscr{C}$, can be regarded as a feasibility of setting up, at each point of E, two spinors u and v associated with generators A_1 and A_2 of the Lie algebra of $\mathscr{C}$ by means of (4.7). Hence we can construct at each point of E a two-dimensional complex space spanned by these u and v, with the group SL$(2,\mathbb{C})$ as a symmetry group. This provides a simple and visual interpretation of the existence condition for a spinor structure.

5. TWO-SPINOR SPACE AS A HOMOGENEOUS SPACE

It is surprising that when we wish to achieve the possibility of having continuous variables describing spin,[9] i.e., when spin degree of freedom is connected to certain coordinates in the homogeneous space of the Lorentz group, the smallest homogeneous space which admits half-integer spin wave functions and possesses an invariant measure, is the space with $\mathscr{C}$ as the stabilizer group. A homogeneous space of the Lorentz group (that is the space on which $\mathscr{L}_0$ acts transitively) may be realized as a coset space $\mathscr{L}_0/\mathscr{G}$ of $\mathscr{L}_0$ modulo some subgroup $\mathscr{G}$, the stabilizer group, and characterized by its invariant measure if it exists. The existence of such a measure is of importance for defining an interaction in a field theory based on a homogeneous space. Finkelstein[9] has shown that we can have Dirac-like states only when the stabilizer group is equal to : the unit element of $\mathscr{L}_0$, or the one-dimensional group generated by A_1, or the two-dimensional group generated by A_2 and A_2. However the first two cases can be realized by a rigid system of space–time points, hence are less appealing. So the most interesting is the third case, when the stabilizer group is the Crumeyrolle group. It is easy to see that in this case the corresponding homogeneous space is equivalent to the two-spinor space.

Indeed, let us take some base $\{\rho, \sigma\}$ of the abstract two-dimenstional spinor space Σ, $[\rho = \binom{1}{0}, \sigma = \binom{0}{1}]$. It may be easily verified that if we remove the point $\binom{0}{0}$ from Σ, the remaining space is a homogeneous space for SL$(2,\mathbb{C})$. We know that if a complex 3-vector $\mathbf{F}$ describes some element X of the Lie algebra of $\mathscr{L}_0$ by (4.2) and (4.1), then $\sigma\cdot\mathbf{F}$ defines an element in the Lie algebra of SL$(2,\mathbb{C})$ corresponding to X. Thus from (4.3) we see that the Lie algebra of $\tilde{\mathscr{C}}$ is spanned by matrices:

$$\begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad (5.1)$$

which means that any element of $\tilde{\mathscr{C}}$ has the form

$$s = \begin{pmatrix} 1 & \zeta \\ 0 & 0 \end{pmatrix} \quad \text{with } \zeta\in\mathbb{C}. \qquad (5.2)$$

So

$$\tilde{\mathscr{C}}\rho = \rho, \qquad (5.3)$$

which is sufficient for us to conclude that $\tilde{\mathscr{C}}$ is the stabilizer group of the spinor space Σ.

Thus spinor space itself is the smallest homogeneous space which can be used to describe half-integer spins by means of scalar wave functions in the sense of Finkelstein, Bacry, and others (all homogeneous space with the same or conjugate stabilizer groups are identified).

ACKNOWLEDGMENT

I would like to thank Professor A. Trautman for helpful discussions and encouragement.

[1]J. Milnor, L'Enseignement mathématique **9**, 198 (1963); A. Lichnerowicz, "Topics on space–time," in *Battelle Recontres* 1967, *Lectures in Mathematics and Physics*, edited by C.M. de Witt and J.A. Wheeler (Benjamin, New York, 1968), Chap. V, p. 115; R. Penrose, "Structure of space–time," *Ibid*, p. 121; K. Bichteler, J. Math. Phys. **9**, 813 (1968).
[2]R. Geroch, J. Math. Phys. **9**, 1739 (1968).
[3]C. Chevalley, *The Algebraic Theory of Spinors* (Columbia U.P., New York, 1954).
[4]A. Crumeyrolle, Ann. Inst. Henri Poincaré **11**, 19 (1969); C.R. Acad. Sci. Paris **271**, 1070 (1970).
[5]A. Crumeyrolle, Ref. 4; see also K. Bugajska, "Spinor structure on space–time," to appear in Int. J. Theor. Phys.
[6]See, e.g., S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. I (Interscience, New York, 1969), or D. Husemoller, *Fiber Bundles* (McGraw-Hill, New York, 1966).
[7]H. Bacry, P. Combe, and P. Sorba, Rep. Math. Phys. **5**, 361 (1974).
[8]E. Cartan, *The Theory of Spinors* (Hermann, Paris, 1954).
[9]D. Finkelstein, Phys. Rev. **100**, 924 (1955); H. Bacry and A. Kihlberg, J. Math. Phys. **10**, 2132 (1969); A. Bacry *et al.*, Ref. 7.
[10]D. Finkelstein, Ref. 9.
[11]See Refs. 1 and 2.
[12]See A. Crumeyrolle, Ref. 4; D. Husemoller, Ref. 6; and K. Bugajska, Ref. 5.
[13]P. Tondeur, *Introduction to Lie Groups and Transformation Groups*, Lecture Notes in Mathematics No. 7, (Springer, Berlin, 1969), 2nd ed.

Stochastic fields from stochastic mechanics [a)]

Steven M. Moore
Departmento de Física, Universidad de los Andes, Apartado Aéreo 4976, Bogotá, Colombia

(Received 19 July 1979; accepted for publication 24 December 1979)

Stochastic field theory for a real scalar field, considering both zero and positive temperatures, is developed from complements to Nelson's stochastic mechanics. These complements include path integral formulas for the moments of the stochastic process, a functional differential equation for the generating functional, and a virial theorem. Using these and Yasue's nonstandard analysis formulation of stochastic field theory, a rigorous meaning is given to the path integral formulas for the field moments and to the functional differential equation of the field's generating functional.

INTRODUCTION

Let us define stochastic physics as that part of physical theory which uses probabilistic methods as the fundamental mathematical tool and consistently employs a classical statistical interpretation in the description of atomic and subatomic phenomena. It is natural to divide this in two regimes, namely the nonrelativistic regime, or few-body theory, and the relativistic regime, or many-body theory.

There are two schools in nonrelativistic stochastic physics which will be called stochastic mechanics (SM) and classical stochastic electrodynamics (CSED).[1] The first school may be considered to be a phenomenological version of the second. One assumes that the mechanics is described by a diffusion process in the first school, the ideas being that no system is isolated and in fact interacts with a medium having infinitely many degrees of freedom, the physical cause of the diffusion.[2] CSED carries this latter idea even further, assuming that the cause of the stochastic fluctuations is a stochastic electric field that even exists when the temperature is zero.[3,4] In this case one no longer has a diffusion process, but in two recent papers, De la Peña and Cetto have shown that CSED has a Markov limit closely related to SM.[3]

Although it is admitted here that Nelson's theory[2] is phenomenological, it is useful for two reasons: The first is that it significantly overlaps with quantum mechanics (QM) without being equal to it.[2] In fact, it might be reinterpreted as an Euclidean formulation of QM, although this question needs much closer examination. The second reason is that SM plays an important role in stochastic field theory (SFT)[5-7]. Guerra and Ruggiero showed[5] that SM as formulated by Nelson[2] can be used to give a real time interpretation to Nelson's free Euclidean scalar field,[8] i.e., this field becomes a real zero-point field in SFT. It has also been shown[6,7] that the zero-point and positive-temperature electric fields used in CSED[3,4] can be obtained in the same way. Hence, SFT could be an alternative to quantum field theory (QFT) that limits to SM and CSED.

There exist obvious mathematical advantages in working with a stochastic description of high-energy phenomena if one thinks in terms of constructive QFT where many of the methods of statistical mechanics are employed. Moreover, SM and SFT provide a dynamical description of tunneling phenomena.[9,10] The purpose of this paper is to present some new mathematical techniques in SM that extend to SFT.

In the next section complements to Nelson's SM[2] are developed. These include path integral formulas for the moments of the stochastic process and a functional differential equation for the generating functional for both the zero and positive temperature theories. The virial theorem used in this section is derived in the Appendix. In the third section the results of the previous section and Yasue's nonstandard analysis formulation of SFT[11] are used in order to give a rigorous meaning to the path integral formulas for the moments of a real scalar stochastic field as well as to the functional differential equation for the field's generating functional.

COMPLEMENTS TO STOCHASTIC MECHANICS

Let us consider a system with N degrees of freedom. According to Nelson,[2] the stochastic process describing the state of the system satisfies

$$dq(t) = b\,(q(t), t\,)dt + dw(t)\,. \tag{1}$$

For zero temperature, it is postulated that dw is the Wiener process with each component independent and with an overall diffusion constant $\hbar/m$:

$$\langle dw_i(t)\rangle = 0\,, \tag{2}$$

$$\langle dw_i(t)dw_j(t')\rangle = \frac{\hbar}{m}\,\delta_{ij}\delta(t-t')dt\,dt'\,. \tag{3}$$

Here $\langle\ \rangle$ denotes the ensemble average (i.e., the average with respect to the underlying probability measure). For positive temperature, it is postulated that dw is the differential process with independent components satisfying Eq. (2) and[7]

$$\langle dw_i(t)dw_j(t')\rangle = \frac{\delta_{ij}}{\beta m}\sum_{n=-\infty}^{\infty} e^{i\omega_n(t-t')}dt\,dt'\,, \tag{3'}$$

where

$$\omega_n = 2\pi n/\beta\hbar\,. \tag{4}$$

It is easily seen that the limit $\beta\to\infty$ in Eq. (3′) gives Eq. (3). The form of Eq. (3′) is dictated by the KMS condition[12] if one thinks in terms of Euclidean quantum theory, but so far there exists no direct motivation for it in terms of stochastic

a)Research financed in part by Colciencias.

physics. Therefore, if one insists on a purely stochastic interpretation in the sense of Guerra and Ruggiero,[5] Eq. (3′) must be taken as a postulate. For this reason it is proposed that Eq. (1) through (4) be taken as the basic postulates of equilibrium SM. However, this proposal presents some serious interpretive problems (see the last section).

Thus, the point of view here differs from the usual one in SM, since Nelson considers a stochastic process to be associated with each quantum state.[2] However, then one would immediately confront the problem that the probability density can have nodes. As a solution to this problem, Albeverio and Høegh-Krohn showed[13] that any stationary state Ψ of the Schrödinger equation can be associated with a homogeneous Markov process of the heat equation with Dirichlet boundary conditions on the hypersurface $\Psi = 0$. Although one can consider such ideas important for systems that are not in equilibrium, equilibrium systems do not possess this problem of nodes since the process for positive temperatures mixes in "excited states" in the correct way and tends to the "ground state" process at zero temperature (later on this behavior will be seen explicitly for N noninteracting oscillators). This observation also should relax Nelson's concern about the problem of the superposition of stationary states.[14] It is proposed here that the only stationary states in stochastic physics are equilibrium states. Stationary states in the sense of nonrelativistic QM do not exist in CSED,[15] so there is no reason to suppose that they suddenly appear in SM, which can be considered to be the phenomenological limit of CSED.[3]

For practical calculations, the viewpoint of this paper presents a problem, however. At zero temperature, the most direct way to determine the forward derivative or drift $b(q(t),t)$ in Eq. (1) is by using the Schrödinger equation.[2] For positive temperature, $b(q(t),t)$ usually depends on the temperature (for further discussion of this point, see the Appendix), but there is nothing corresponding to the Schrödinger equation that will determine b.[16] It is fortunate that for the oscillator $b(q(t),t) = -\omega q(t)$, i.e., the temperature dependence is contained in $q(t)$ itself; this permitted applying the same direct methods of solution of Eq. (1) in the positive-temperature case[7] as the zero-temperature one.[2] However, in general, one must develop other methods for determining q since the form of b at positive temperature is not given.

The method that will be developed in this section is based on the assumption that the moments of q determine q itself. In the practical applications to SFT, this is all one will need. The final formulas (8) and (9) look very much like Euclidean quantum mechanics. This will be discussed in more detail later.

For zero temperature, Yasue has shown that the transition probabilities for q can be written as[9]

$$p(x,t\,|x_0,t_0) = \sqrt{\rho(x)/\rho(x_0)}\int d\mu_w(q(\cdot)) \times\exp\left[-\hbar^{-1}\int_{t_0}^{t}(V-E)dt'\right]. \tag{5}$$

Here μ_w represents the Wiener measure concentrated on paths q such that $q(t_0)=x_0$, $q(t)=x$, and E can be interpreted as the average energy[2]:

$$E = \int(\tfrac{1}{2}mb^2 + V)\rho\, dx\,. \tag{6}$$

It is important to note that Yasue's proof[9] does not depend on the Schrödinger equation but only on the Fokker–Planck equation. Moreover, accepting Eq. (3′) and the Markov nature of the process, one has the same equation with a different drift and periodic boundary conditions in time. Hence, the proof is valid for positive temperature if one uses a periodic Wiener measure. For applications, one will only need $p(x,\beta\,|x_0,0)$ due to this periodicity.[17]

Now let us note that the time-ordered moments of the stochastic process q are given by (for zero temperature)

$$\langle q_{i_1}(t_1)\cdots q'_{i_p}(t_p)\rangle = \lim_{\substack{T'\to\infty\\ T\to-\infty}} p(0,T'|0,T)^{-1} \times\int dx_1\cdots\int dx_p\, p(0,T'|x_p,t_p)x_{i_p} \times p(x_p,t_p|x_{p-1},t_{p-1})\times\cdots\times x_{i_2} \times p(x_2,t_2|x_1,t_1)x_{i_1}p(x_1,t_1|0,T)\,. \tag{7}$$

The notation is such that x_{i_j} is the i_jth component of the N-dimensional vector x_j and $t_1<t_2<\cdots<t_p$. Let us now introduce a generating functional $G\{J\}$ for these moments defined by

$$G\{J\} = \int d\mu_w(q(\cdot))\exp\left[-\hbar^{-1}\int_T^{T'}(V+J\cdot q)dt'\right]. \tag{8}$$

Thus, $G\{J\}$ is essentially the transition probability $p(0,T'|0,T)$ corresponding to the same potential V but with an external interaction $J\cdot q$ added. Using the usual definition of functional derivative, one sees that

$$\langle q_{i_1}(t_1)\cdots q_{i_p}(t_p)\rangle = \lim_{\substack{T'\to\infty\\ T\to-\infty}} \frac{(-\hbar)^p}{G\{J\}}\frac{\delta^p G\{J\}}{\delta J_{i_1}(t_1)\cdots\delta J_{i_p}(t_p)}\bigg|_{J=0}. \tag{9}$$

The same formulas (8) and (9) are valid for positive temperature, except $T=0$, $T'=\beta$, $0\leqslant t_1<\cdots<t_p\leqslant\beta$, and no limit is needed. In this case, the Wiener measure must be periodic, of course.

The generating functional (8) is closely related to the formal one used in Euclidean QM.[18] The Wiener integral contains a term that looks like

$$-\frac{m}{2}\int_T^{T'}\dot q(t')^2dt' \tag{10}$$

so one may formally write

$$G\{J\} = \int\mathscr{D}q(\cdot)\exp\left[\hbar^{-1}S_E\{q(\cdot)\} - \hbar^{-1}\int_T^{T'}J\cdot q\,dt'\right], \tag{11}$$

where $\int\mathscr{D}q(\cdot)$ represents a path integral,

$$S_E\{q(\cdot)\} = \int_T^{T'}L(i\dot q,q)dt', \tag{12}$$

and L is the classical Lagrangian. However, the subscript E has been written in (12) to emphasize that a real physical time interpretation has been given to Euclidean QM in the spirit of Guerra and Ruggiero.[5] Although Eq. (12) is a Wick-rotated[19] action, the real physics is in Eq. (9). QM would say that the right-hand side of (9) is the Wick-rotated ground state to ground state transition amplitude. SM states that it is the time-ordered moment of the stochastic process. Moreover, one has arrived at (9) without ever leaving the theoretical framework of SM.

Expressions (8) and (11) are in terms of functional integrals. It may be useful to have an expression for $G\{J\}$ in terms of a functional differential equation. It is to be observed that $\mathscr{D}q(\cdot)$ is translation invariant, i.e., $\mathscr{D}(q(\cdot)+q'(\cdot))=\mathscr{D}(q(\cdot))$ as long as $q'(T)=q'(T')=0$. Hence, one has the identity

$$\frac{\delta}{\delta q_i'(t)}\int \mathscr{D}(q(\cdot)+q'(\cdot))\exp\Big[\hbar^{-1}S_E\{q(\cdot)+q'(\cdot)\} - \hbar^{-1}\int_T^{T'} J\cdot(q+q')dt'\Big]\Big|_{q'=0}=0\,. \tag{13}$$

This gives the following functional differential equations for $G\{J\}$:

$$0=\left[J_i(t)+m\frac{d^2}{dt^2}\frac{\delta}{\delta J_i(t)}-\frac{\partial V}{\partial x_i}\left(\frac{\delta}{\delta J(t)}\right)\right]G\{J\}\,. \tag{14}$$

As an example of an application of formulas (8) and (9) and Eq. (14), let us consider the case of N noninteracting oscillators. This overworked example is an important one for SFT (see the next section).

For the zero-temperature case, the calculation of the moments has appeared before[18] (without being interpreted as SM!), so only the result is quoted. One has

$$\langle q_i(t)\rangle=0\,, \tag{15}$$

$$\langle q_i(t)q_j(t')\rangle=\frac{\hbar}{2m\omega}\delta_{ij}e^{-\omega(t-t')}\quad(t>t')\,. \tag{16}$$

Higher-order moments are determined by these (the process is Gaussian). Using the virial theorem that has been developed for SM (see the Appendix), one has

$$E=Nm\omega^2\langle q^2\rangle=\tfrac{1}{2}N\hbar\omega\,. \tag{17}$$

Thus, one may say that the average energy per oscillator is $\frac{1}{2}\hbar\omega$. Note that this is very different from saying that each oscillator has the ground state energy. This is an imporant difference in interpretation between quantum mechanics and stochastic mechanics that should not be overlooked.

For positive temperatures, one still finds that $\langle q_i(t)q_j(t')\rangle$ for $t>t'$ is the Green function D_E satisfying[18]

$$-\ddot{D}_E+\omega^2D_E=-\frac{1}{m}\delta(t-t')\,. \tag{18}$$

However, now one must have $D_E(\beta\hbar)=D_E(0)$ in order to guarantee the KMS condition. Thus,

$$\langle q_i(t)q_j(t')\rangle=\frac{\delta_{ij}}{\beta m}\sum_{n=-\infty}^{\infty}\frac{e^{i\omega_n(t-t')}}{\omega^2+\omega_n^2}\,, \tag{19}$$

in agreement with a previous result.[7] Thus, the average energy is

$$E=Nm\omega^2\langle q^2\rangle=\frac{N\omega^2}{\beta}\sum_{n=-\infty}^{\infty}(\omega^2+\omega_n^2)^{-1}\,. \tag{20}$$

Using

$$\sum_{n=-\infty}^{\infty}(n^2+a^2)^{-1}=\frac{\pi}{a}\coth(\pi a)\,, \tag{21}$$

one obtains

$$E=\tfrac{1}{2}N\hbar\omega\coth(\tfrac{1}{2}\beta\hbar\omega)\,. \tag{22}$$

This is exactly the quantum expression for the ensemble average of a system of noninteracting oscillators. It should be noted that these positive-temperature results have been confirmed in the recent article by De la Peña and Cetto which deals with the oscillator in CSED.[4]

In the case of the oscillator, Eq. (14) reduces to

$$0=\left[J_i(t)+\frac{d^2}{dt^2}\frac{\delta}{\delta J_i(t)}-m\omega^2\frac{\delta}{\delta J_i(t)}\right]G\{J\}\,. \tag{23}$$

One can easily verify that

$$G\{J\}=\prod_{i=1}^{N}\exp\left[-\hbar^{-1}\int_T^{T'}dt_1\int_T^{T'}dt_2J_i(t_1)D_E(t_1-t_2)J_i(t_2)\right], \tag{24}$$

is the solution to Eq. (23). This is just the functional integral (8).

STOCHASTIC FIELD THEORY

One should not consider SM to be a complete nonrelativistic theory considering the results of CSED mentioned earlier.[3,4] However, it does not seem to be as limited as one might have believed. As a phenomenological approximation to CSED, it serves a useful check on the more complete theory.

CSED, of course, cannot be a complete theory of atomic and subatomic phenomena either. SFT, to the extent that it has been developed, already limits to CSED since the zero-point and positive-temperature electric fields can be considered as free noninteracting fields in SFT.[6,7]

Let us consider a real stochastic scalar field satisfying

$$(\Box+m^2)\varphi(x)=j(x)\,. \tag{25}$$

Here $j(x)$ is the nonhomogeneous term representing all sources [this may contain self-interactions, e.g., $V'(\varphi(x))$, or terms that are fixed and external to the system, or terms that come from other fields].

If $j=0$, a simple procedure of randomization can be applied for turning the classical field into a stochastic field.[5-7] Let us take a bounded and smooth region G in $\mathbb{R}^3$. Let e_i be the characteristic functions of the negative Laplacian in G, i.e.,

$$-\Delta e_i=k_i^2e_i\,, \tag{26}$$

where the k_i^2 are the characteristic values. Then φ has an expansion

$$\varphi(\mathbf{x},t)=\sum a_i(t)e_i(\mathbf{x})\,. \tag{27}$$

One then associates one-dimensional stochastic oscillators with each $a_i(t)$ and takes the limit $G\to\mathbb{R}^3$. [One can equally well do the infinite-volume limit directly by taking the e_i to be a complete set of $\mathscr{L}(\mathbb{R}^3)$ in $\mathscr{S}(\mathbb{R}^3)$.) One gets Nelson's free Euclidean field[8] for $\beta=\infty$[5,6] and the positive-temperature free field[20] for $\beta<\infty$ [in the sense that the average energy, and hence $\log Z(\beta)$, are the same]. The moments are

$$\langle\varphi(\mathbf{x},t)\varphi(\mathbf{y},t')\rangle = \frac{1}{(2\pi)^3}\int d^3\mathbf{k}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}G^{(\beta)}(t-t') = \frac{1}{(2\pi)^4}\int d^4k\frac{e^{i(k,x-y)}}{m^2+(k,k)} \quad (\beta=\infty), \quad (28)$$

where $G^{(\beta)}(t-t')$ is the right-hand side of Eq. (16) for $\beta=\infty$ [and Eq. (19) for $\beta<\infty$] taking $\hbar=m=1$ and $i=j$, and (,) denotes the Euclidean inner product. In both cases we have the real time interpretation discovered originally by Guerra and Ruggiero.[5] Note that the moment is continuous for $\beta\to\infty$. This continuity relates to an important interpretive point which will be discussed in the last section.

Well known problems arises when $j\neq 0$. However, let us again suppose that the moments of φ determine φ. This is consistent with the interpretation of the moments as Wick rotations of the vacuum to vacuum transition amplitudes (vacuum expectation values of time-ordered products), but it may not be if one is to adhere to the strict stochastic interpretation, so this is simply taken as a hypothesis. In this case, by analogy with Eq. (9), (11), and (12), one might suppose that

$$\langle\varphi(x_1)\cdots\varphi(x_p)\rangle = \lim_{\substack{T'\to\infty\\ T\to-\infty}} \frac{(-1)^p}{G\{J\}}\frac{\delta^p G\{J\}}{\delta J(x_1)\cdots\delta J(x_p)}\bigg|_{J=0}, \quad (29)$$

where it is understood that $t_1<t_2<\cdots<t_p$ and

$$G\{J\} = \int \mathscr{D}\varphi(\cdot)\exp\left[\int_T^{T'}dt\int d^3\mathbf{x}(\mathscr{L}(\varphi,i\dot\varphi)-J\varphi)\right] \quad (30)$$

is a path integral over the classical φ.

For the free equilibrium field at zero temperature, the calculations can be carried out formally and Nelson's free Euclidean field[8] results [see Eq. (28)]. This is not surprising, since the formulas are the same ones used in Euclidean QFT.[18] Moreover, changing T' to β and T to 0 and integrating over cyclic paths gives the positive-temperature free equilibrium field[20] [also given by Eq. (28)]. However, again the left-hand side of (29) indicates that the interpretation is different in SFT since one is considering the moments of stochastic fields which satisfy the relativistic equation (25).

The advantage of Eqs. (29) and (30) is that they permit one to treat interacting fields and one can, at least formally, apply perturbation theory.[18,20] Nevertheless, the path integral formulas (29) and (30) are only formal expressions. The rest of this paper is dedicated to giving them a rigorous meaning using Yasue's nonstandard formulation of SFT.[11] In the process Yasue's construction will be clarified, especially in its relation with the simple randomization procedure outlined above. It should be noted that the rigorous path integral formulas may be given the standard Euclidean interpretation for the self-interacting scalar field, but for other fields this is not the case, as one will see upon reading the last section.

Let $\mathscr{F}$ be a regular ultrafilter in the natural numbers N. Then the ultraproduct

$$^*E = \prod_{N\in\mathbb{N}}\mathbb{R}^N/\mathscr{F} \quad (31)$$

is well defined by Robinson's theory of nonstandard analysis.[21] Yasue calls this ultra-Euclidean space,[11] but this name is misleading since $^*E$ is not a nonstandard version of any $\mathbb{R}^N$. However, as he notes, $^*E$ is a vector space over the nonstandard reals

$$^*\mathbb{R} = \prod_{N\in\mathbb{N}}\mathbb{R}/\mathscr{F} \quad (32)$$

and a * Euclidean space with respect to the inner product

$$[a^{(N)}]\cdot[b^{(N)}] = [a^{(N)}\cdot b^{(N)}] = \left[\sum_{i\leqslant N}a_i^{(N)}b_i^{(N)}\right]. \quad (33)$$

Here $[a^{(N)}]$ denotes the equivalence class of $(a^{(N)})$ modulo $\mathscr{F}$.

The right-hand side of Eq. (32) is, of course, only one of many non-Archimedean fields that contain the real numbers and infinitesimals. With respect to model theory and its relation to analysis,[21] one would like to be able to make compactness arguments in $^*\mathbb{R}$. (This is not needed in what follows, however.) This means that $\mathscr{F}$ must be a regular ultrafilter in the set of finite subsets of card $\mathscr{L}$, where $\mathscr{L}$ is the language of the model associated with $^*\mathbb{R}$.[21] Therefore, Yasue's construction necessarily implies that $\mathscr{L}$ is countable. This may seem strange from the point of view of type theory,[21] but from the mathematical physicist's point of view it is easy to justify: It is impossible in one's finite lifespan to write down more than a finite number of symbols and/or formulas from any language!

Yasue chooses a free ultrafilter,[11] but regularity of $\mathscr{F}$ guarantee that $\mathscr{F}$ is free since N in infinite. Moreover, this implies that each $A\in\mathscr{F}$ is infinite. This observation gives a direct way of associating $a\in\mathbb{R}$ with an element of $^*\mathbb{R}$: Take $a_i^{(N)}=a$ except possibly in a finite number of pairs (N,i).

Each $[a^{(N)}]\in{}^*E$ defines a field from $\mathbb{R}^3$ to $^*\mathbb{R}$ by Yasue's formula[11] [where now we take e_i to be a complete set of $\mathscr{L}(\mathbb{R}^3)$ in $\mathscr{S}(\mathbb{R}^3)$]:

$$\varphi(\mathbf{x}) = \left[\sum_{i\leqslant N}a_i^{(N)}e_i(\mathbf{x})\right]. \quad (34)$$

Let us denote the set of all such φ by $^*K$. Yasue calls this Kawabata space.[11] As he notes, $^*K$ is homeomorphic and isomorphic to $^*E$. Thus, $^*K$, like $^*E$, is not an extension of any standard set. However, it is possible to relate Eqs. (27) and (34) in such a way that one can embed standard fields in $^*K$. One simply takes $a_k^{(N)}=a_i\delta_{ki}$ for each N. Then the Fourier decomposition of Eq. (34) gives a Fourier coefficient $[\Sigma_{k\leqslant N}a_k^{(N)}\delta_{ki}]$, and this is a_i by the remarks in the preceding paragraph. This relation between standard and nonstandard fields is not obvious in Yasue's work[11] and will play an important role in later developments.

Following Yasue,[11] one can now construct stochastic fields via the homeomorphism between $^*E$ and $^*K$. Let $q^{(N)}$ be a diffusion process in $\mathbb{R}^N$ satisfying Eq. (1). Then $[q^{(N)}]$ is a stochastic process in $^*E$. Hence,

$$\varphi(\mathbf{x},t) = \left[\sum_{i\leqslant N}q_i^{(N)}(t)e_i(\mathbf{x})\right] \quad (35)$$

is a stochastic field in $^*K$.

If

$$j(\mathbf{x},t)=\sum j_i(t)e_i(\mathbf{x}), \tag{36}$$

it is possible to construct a solution to Eq. (25) using Eq. (35). One takes anharmonic oscillators $q^{(N)}$ satisfying[22]

$$\ddot{q}_i^{(N)}+(k_i^2+m^2)q_i^{(N)}=j_i^{(N)}, \tag{37}$$

where $j_i^{(N)}=\delta_{ij}j_j$. There is only one technical point to check: One must insure that $a^{(N)}$ does not depend on $q^{(M)}$ for $M>N$. For an external j, this is obvious. For polynomial or exponential self-interactions, each power of φ is a local product in the sense of Yasue,[11] so each Nth component of the power only contains $q^{(M)}$ with $M=N$. For interactions with other fields, one also has polynomials in φ and other fields (or derivatives of these), so again no Nth component contains a $q^{(M)}$ with $M>N$.

As an example, consider

$$j(x)=-4\lambda\varphi(x)^3. \tag{38}$$

Then

$$j_i^{(N)}=-4\lambda\sum_{i_1,i_2,i_3\leqslant N}q_{i_1}^{(N)}q_{i_2}^{(N)}q_{i_3}^{(N)}\times\delta_{ij}\int d^3\mathbf{x}e_je_{i_1}e_{i_2}e_{i_3}. \tag{39}$$

It should be noted that (37) cannot be solved by the method developed by Nelson for solving the anharmonic oscillator (see the second reference in Ref. 2), since, as in Eq. (39), the $a^{(N)}$ generally depend on $q^{(N)}$. However, assuming certain regularity conditions on the $a^{(N)}$, Eq. (37) can be solved in the sense that one can in principle determine all the moments by using Eq. (9). This is all that will be needed in what follows.

For $j=0$, the construction outlined above reduces to the randomization procedure used in the case of free fields. Thus, the free fields are standard. For interacting fields, one should expect that in general they will be nonstandard, since distributions can be considered as nonstandard elements. As an example, $\delta(\mathbf{x}-\mathbf{y})$ may be written as

$$\delta(\mathbf{x}-\mathbf{y})=\Big[\sum_{i\leqslant N}e_i(\mathbf{y})e_i(\mathbf{x})\Big], \tag{40}$$

which each $\mathbf{x}$ and $\mathbf{y}$ is a perfectly good nonstandard number. Therefore, one either has to cope with nonstandard elements of $^*K$ for the interacting case or one has to smear the $\varphi(x)$ in order to get a generalized stochastic process. This latter procedure is not excluded, but the meaning of Eqs. (29) and (30) is not obtained by considering it. Thus, such generalized stochastic processes are not considered in this paper.

Let us note here an amusing advantage of the nonstandard SFT. Consider the average energy E of the free stochastic field. As a nonstandard real number, it is

$$E=\Big[\sum_{i\leqslant N}\tfrac{1}{2}(k_i^2+m^2)^{1/2}\coth\big(\tfrac{1}{2}\beta(k_i^2+m^2)\big)\Big]. \tag{41}$$

This is infinite, of course, since it includes the infinite energy contributed by the zero-point field. However, it can easily be distiguished from the other nonstandard real numbers and used unambiguously in calculations. One can perform a simple renormalization of course, but this denies the real existence of the zero-point field. Actually, the renormalization is not necessary since only energy differences are important, i.e., E_1-E_2 should be near standard for any energies E_1 and E_2.[23] Thus, all energies can be referenced to the energy in Eq. (41) for $\beta=\infty$. Other renormalizations are not so obvious.

It is now an easy step to give meaning to formulas (29) and (30). Let us first define a transition probability that is an easy extension of Yasue's definition of the probability density.[11] Let φ and φ_0 be two elements of $^*K$ independent of t. Then the transition probability $p(\varphi,t\,|\varphi_0,t_0)$ that the stochastic field $\varphi(x)$ is φ at t if it is φ_0 at t_0 is defined by

$$p(\varphi,t\,|\varphi_0,t_0)=\big[p^{(N)}(x^{(N)},t\,|x_0^{(N)},t_0)\big], \tag{42}$$

where $p^{(N)}(x^{(N)},t\,|x_0^{(N)},t_0)$ is the transition probability for the component stochastic process $q^{(N)}$ in Eq. (37). Obviously, one has that $0\leqslant p(\varphi,t\,|\varphi_0,t_0)\leqslant 1$ by applying the definition of truth in the model $^*\mathbb{R}$,[21] and one easily sees that Eq. (42) satisfies the Chapman–Kolmogorov equation

$$\int p(\varphi,t\,|\varphi_1,t_1)p(\varphi_1,t_1|\varphi_0,t_0)\delta\varphi_1=p(\varphi,t\,|\varphi_0,t_0), \tag{43}$$

where one uses the definition[11]

$$\int F\{\varphi\}\delta\varphi=\Big[\int d^Nx^{(N)}F^{(N)}\{x^{(N)}\}\Big] \tag{44}$$

for $\varphi=[\Sigma_{i\leqslant N}x_i^{(N)}e_i]$ and for a functional F on $^*K$ of the type

$$F\{\varphi\}=[F^{(N)}\{x^{(N)}\}]. \tag{45}$$

From (42) and (5), one also has that

$$\begin{aligned}p(\varphi,t\,|\varphi_0,t_0)&=\Big[\sqrt{\rho(x^{(N)})/\rho(x_0^{(N)})}\int d\mu_w(q(\cdot))\\&\qquad\times\exp\Big\{-\int_{t_0}^{t}(V^{(N)}-E^{(N)})dt'\Big\}\\&\equiv\sqrt{\frac{\rho(\varphi)}{\rho((\varphi_0))}}\int d\mu_w(\varphi(\cdot))S\{\varphi(\cdot)\},\end{aligned} \tag{46}$$

where

$$S\{\varphi(\cdot)\}\equiv\Big[\exp\Big\{-\int_{t_0}^{t}(V^{(N)}-E^{(N)})dt'\Big\}\Big]. \tag{47}$$

Now let us consider explicitly the field equation

$$(\Box+m^2)\varphi=j+J, \tag{48}$$

where J is an arbitrary external interaction. When $J=0$, φ is the solution of (25) whose moments are needed. Define

$$\begin{aligned}G\{J\}&\equiv[G^{(N)}\{J^{(N)}\}]\\&\equiv\Big[\int d\mu_w(q^{(N)}(\cdot))\exp\Big\{-\int_T^{T'}(V^{(N)}+J^{(N)}\cdot q^{(N)})\Big\}\Big],\end{aligned} \tag{49}$$

where $J_i^{(N)}=\delta_{ij}J_i$ and

$$J(\mathbf{x},t)=\sum J_i(t)e_i(\mathbf{x}). \tag{50}$$

Also define

$$\frac{\delta G\{J\}}{\delta J(x)}=\Big[\sum_{i\leqslant N}\frac{\delta G^{(N)}\{J^{(N)}\}}{\delta J_i^{(N)}(t)}e_i(\mathbf{x})\Big], \tag{51}$$

with higher-order functional derivatives being defined in an analogous manner. Then

$$\langle\varphi(x_1)\cdots\varphi(x_p)\rangle = \lim_{\substack{T'\to\infty\\ T\to-\infty}} \frac{(-1)^p}{G\{J\}} \frac{\delta^p G\{J\}}{\delta J(x_1)\cdots\delta J(x_p)}\bigg|_{J=0} . \tag{52}$$

To prove Eq. (52), let us assume that $x_1,\dots,x_p$ are already time ordered. Each $J^{(N)}$ defines an external interaction for $q^{(N)}$, so $G^{(N)}\{J^{(N)}\}$ makes sense. Thus,

$$\begin{aligned}\langle\varphi(x_1)\cdots\varphi(x_p)\rangle &= \left\langle\left[\sum_{i_1,\dots,i_p\leq N} q_{i_1}^{(N)}(t_1)\cdots q_{i_p}^{(N)}(t_p) e_{i_1}(\mathbf{x}_1)\cdots e_{i_p}(\mathbf{x}_p)\right]\right\rangle \\ &= \left[\sum_{i_1,\dots,i_p\leq N} \langle q_{i_1}^{(N)}(t_1)\cdots q_{i_p}^{(N)}(t_p)\rangle e_{i_1}(\mathbf{x}_1)\cdots e_{i_p}(\mathbf{x}_p)\right] \\ &= \left[\lim_{\substack{T'\to\infty\\ T\to-\infty}} \sum_{i_1,\dots,i_p\leq N} \frac{(-1)^p}{G^{(N)}\{J^{(N)}\}} \right. \\ &\quad \left. \times \frac{\delta^p G^{(N)}\{J^{(N)}\}}{\delta J_{i_1}^{(N)}(t_1)\cdots\delta J_{2_p}^{(N)}(t_p)}\bigg|_{J^{(N)}=0} e_{i_1}(\mathbf{x}_1)\cdots e_{i_p}(\mathbf{x}_p)\right].\end{aligned} \tag{53}$$

Using the definition of multiplication in *ℝ, this expression reduces to the right-hand side of (52) as long as one can interchange limits and the brackets [ ]. This is always possible since one can choose T and T' large enough so that $T\leqslant t_1<\cdots<t_p\leqslant T'$ and $J(x)$ zero outside of $[T, T']$. For the positive temperature case, $T=0$, $T'=\beta$, and there is no problem, of course.

It is interesting to observe that Eq. (51) almost agrees with the usual definition of functional derivative, which satisfies

$$\int dx\, f(x)\frac{\delta G\{J\}}{\delta J(x)} = \lim_{\epsilon\to 0^+}\frac{G\{J+\epsilon f\}-G\{J\}}{\epsilon}. \tag{54}$$

To see this, write

$$f(\mathbf{x},t) = \left[\sum_{i\leq N} b_i^{(N)}(t) e_i(\mathbf{x})\right]. \tag{55}$$

Then

$$\int dx\, f(x)\frac{\delta G\{J\}}{\delta J(x)} = \left[\sum_{i\leq N}\lim_{\epsilon\to 0^+}\frac{G^{(N)}\{J^{(N)}+\epsilon b^{(N)}\}-G^{(N)}\{J^{(N)}\}}{\epsilon}\right]. \tag{56}$$

If one could interchange the limit and [ ], he would have the right-hand side of (54). It is not clear that this can always be done, however. Yasue's definition[11] of $\delta F/\delta\varphi(\mathbf{x})$ also suffers from this same problem.

In conclusion, (49) and (52) give a rigorous meaning to the path integral formulas (29) and (30). As far as the author knows, this is the first rigorous definition given in the mathematical physics literature, irrespective of whether one gives a SFT interpretation or an Euclidean field theory interpretation.

Let us now note that (54) also gives the means to derive the usual functional expansion for $G\{J\}$. Let us treat $J+zj$ as the external source, and consider $G_0\{J+zj\}$, which reduces to $G\{J\}$ for $z=1$. Then one has

$$G_0\{J+zj\} = \exp\left[z\int dx\, j(x)\frac{\delta}{\delta J(x)}\right]G_0\{J\}. \tag{57}$$

Hence,

$$G\{J\} = \exp\left[\int dx\, j(x)\frac{\delta}{\delta J(x)}\right]G_0\{J\}, \tag{58}$$

which gives the expansion of $G\{J\}$ in terms of the free field $G_0\{J\}$. Equation (58) can be used as a formal basis for perturbation theory.

Equations (37) and (39) can also be used to obtain a functional differential equation for $G\{J\}$. For $\lambda\varphi^4$ theory one has

$$0 = \left[J(x) + (\Delta_4 - m^2)\frac{\delta}{\delta J(x)} - 4\lambda\left(\frac{\delta}{\delta J(x)}\right)^3\right]G\{J\}, \tag{59}$$

where Δ_4 is the four-dimensional Laplacian. This is just the Wick-rotated version of the usual functional differential equation of QFT, but a rigorous meaning and interpretation has been given to it as an equation in SFT. It can be used to develop equations for the Green functions $D_E(x_1,\dots,x_p)$ defined by

$$\log G\{J\} = \sum_{p=0}^{\infty}\frac{(-1)^p}{p!}\int dx_1\cdots\int dx_p D_E\, J(x_1)\cdots J(x_p). \tag{60}$$

It is easy to see that

$$\frac{(-1)^p}{G\{J\}}\frac{\delta^p G\{J\}}{\delta J(x_1)\cdots\delta J(x_p)}\bigg|_{J=0} = \frac{1}{p!}D_E(x_1,\dots,x_p), \tag{61}$$

which is why such equations are of interest.

CONCLUDING REMARKS

Stochastic theories of microscopic phenomena have received general criticism on the grounds that they do not preclude the possibility of hidden variables. However, SFT is only a mathematical model used much in the same way that one uses ensembles in statistical mechanics. The fields are classical in the sense that they are solutions to classical field equations, and their stochastic nature is a model of all the random contributions from sources in the universe. If one accepts this argument, hidden variables are not needed.

More specific criticisms have been directed at Nelson's theory, so it may be interesting to dwell on this point a moment. Let us refer explicitly to the recent criticism of Grabert *et al.*[24] and treat several points. In their title, these authors ask: "Is quantum mechanics equivalent to a classical stochastic process?" The answer is obviously no; neither Nelson for anyone else has ever pretended that the theories are equivalent. Nelson has repeatedly indicated essential differences,[2] and the work in CSED by De la Peña and Cetto[3,4] shows clearly that CSED is not equivalent either.

It is interesting that Grabert *et al.*[24] completely ignore CSED as well as SFT. Criticisms about the form of the correlation [their Eq. (4.12), and Eq. (16) in the present paper] appeared earlier in CSED in a paper by Claverie and Diner,[25] for example. However, Nelson's process *is* a Markov process, and De la Peña and Cetto have also shown that CSED possesses a true Markov limit.[3]

Grabert *et al.* also observe that there exists a drift velocity for each solution of the Schrödinger equation.[24] Mathematically, this is true, and this could be a valid criticism for those who would consider that SM in its present form can directly describe all quantum states. Nevertheless, their comment is not relevant to the contents presented in the second section of this paper, since the author's proposal that SM is an equilibrium theory means that there exists only one drift velocity which is a function of temperature. In fact, in this author's modest opinion, none of the stochastic theories have been able to go beyond equilibrium states, so the entire study of nonequilibrium phenomena remains to be done.

The comments in Sec. V of the paper by Grabert *et al.*[24] in which they refer to the nonquantum nature of the stochastic correlations can hardly be taken as a criticism of SM when these same authors admit that "... various definitions of quantum-mechanical correlations have been introduced in different contexts." If QM cannot decide on which correlations are the right ones, no comparison can be made between it and SM.

Nelson shows how SM can be considered a logical extension of Newtonian mechanics.[2] Davidson takes a more abstract viewpoint.[16] As a result, he shows that there is an infinity of diffusion processes, each with a different diffusion constant, which lead to the Schrödinger equation. This ambiguity merely points out the phenomenological nature of the Markov approximation, of course, and hence the phenomenological nature of the Schrödinger equation.

One can also anticipate criticisms of SFT. The first is that the theory is not really Lorentz invariant. This can be seen most easily by considering the electromagnetic field. For the radiation gauge the free field moments are[26]

$$\langle A_i(\mathbf{x}, t)A_j(\mathbf{y}, t)\rangle = \frac{1}{2\pi^2}\int d^3\mathbf{k}G^{(\beta)}(t-t') \times e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}(\delta_{ij} - k_i k_j/|\mathbf{k}|^2). \quad (62)$$

For an arbitrary gauge and in 4-vector notation one has for $\beta = \infty$

$$\langle A_\mu(x)A_\nu(y)\rangle = \frac{1}{4\pi^3}\int d^4k\,(k,k)^{-1}e^{i(k,x-y)} \times[-g_{\mu\nu} - k_\mu k_\nu(1-f(|\mathbf{k}|))/((k\cdot\eta)^2 - k\cdot k)], \quad (63)$$

where f is the gauge function, (,) denotes the Euclidean product, and $\cdot$ the Minkowski product. Equation (63) may be compared with the free Proca field moments:

$$\langle A_\mu(x)A_\nu(y)\rangle = \frac{1}{(2\pi)^4}\int d^4k e^{i(k,x-y)} \times(m^2 + (k,k))^{-1}[-g_{\mu\nu} + k_\mu k_\nu/m^2]. \quad (64)$$

Note that neither (63) nor (64) possesses Lorentz invariance. The term in brackets are invariant because the fields satisfy their corresponding relativistic equations (they are just the polarization vector sums), but the rest of the integrands is not because it corresponds to the moments of the component oscillators. One can understand this phenomena by considering the fact that (63) and (64) are the limit $\beta \to \infty$ of the positive-temperature fields. The statistical ensemble present for positive temperatures is still present in the limit, so the moments should *not* be Lorentz invariant. However, the reference frame in which one calculates seems to be irrelevant for $\beta = \infty$ (even though the moments are not invariant) since the energy spectrum of the field *is* Lorentz invariant.

The second criticism that might be applied to SFT can also be applied to SM and CSED at positive temperatures. If one considers the periodicity of the positive-temperature paths in the second section of this paper, one finds that even at the low temperature of 1°K, the fundamental period is of the order of 10^{-11} sec, being smaller for higher temperatures. This seems very strange, and this author does not know how to interpret it in the real-time framework of SM. These remarks also hold for positive-temperature SFT and hence for positive-temperature CSED, since the positive-temperature fields are composed of positive-temperature oscillators (in general, anharmonic).

One way out of the dilemma mentioned above is to simply assume that SM and SFT are Euclidean theories. However, (63) and (64) are clearly not Euclidean correlations, and this will be true of all higher-spin fields. The Euclidean fields would have to be completely iSO(4) invariant, i.e., one also would need ($\mathbf{A}$, iA_0), the old-fashioned four vector, and not (A_0, $\mathbf{A}$). The essential point here is that the stochastic A_μ satisfies the relativistic equation.

These interpretative problems and criticisms aside, it should be mentioned that Yasue has already considered Yang–Mills fields in his study of tunneling phenomena.[10] Here the whole question of quantization of gauge fields should be re-examined in the context of SFT using the rigorous methods developed in this paper. Fermion fields can be studied using anticommuting *c*-numbers.[27] In this case the component stochastic processes of the field should give a method for studying spin in SM or CSED.

Various authors have indicated the usefulness of considering Euclidean fields as stochastic diffusion processes in infinite-dimensional space.[28] Nonstandard SFT clearly offers an alternative that is more closely related to our intuition. The author believes that it should be developed further in order to see if it can make more transparent some of the problems confronting high-energy physics.

ACKNOWLEDGMENTS

The author wishes to thank J. Challifour, L. De la Peña, E. Nelson, A.M. Rodriguez-Vargas, A. Rueda, E. Santos, K. Yasue, and the referee for useful references as well as P. Claverie, J. Giraldo, A.M. Rodriguez-Vargas, A. Rueda, E. Santos, and J. Vargas for interesting discussions.

APPENDIX

Here the virial theorem used in the second section of this paper is derived.

In classical mechanics, the virial theorem has a simple derivation.[29] The corresponding derivation in QM is slightly complicated by the formalism and by having to keep track of the commutators.[30] In SM, as will be seen, the derivation approaches the simplicity of the classical-mechanical derivation. Let us consider the case of zero temperature.

Let us start with Newton's law in the form[22]

$$-\frac{\partial V}{\partial x_i}(q_i) = \frac{m}{2}\ddot{q}_i. \tag{65}$$

Thus, the ensemble average of the virial is

$$\left\langle \sum q_i \frac{\partial V}{\partial x_i}(q_i)\right\rangle = \left\langle \sum m b_i^2 \right\rangle - \left\langle D\left(\sum q_i \cdot m b_i\right)\right\rangle. \tag{66}$$

The renormalized kinetic energy is[2]

$$\langle T \rangle = \left\langle \sum m b_i^2/2 \right\rangle. \tag{67}$$

Thus, the right-hand side of (66) reduces to

$$2\langle T \rangle - \left\langle D\left(\sum q_i \cdot m b_i\right)\right\rangle. \tag{68}$$

Hence, the virial theorem holds if the last term is zero.

Let us recall that[2]

$$Df(q(t),t) = \left(\frac{\partial}{\partial t} + b\cdot\nabla + \frac{\hbar}{2m}\Delta\right) f(q(t),t) \tag{69}$$

for any smooth function f, so that

$$\left\langle D\left(\sum q_i \cdot m b_i\right)\right\rangle = \int d^N x \rho(x)\left(\frac{\partial}{\partial t} + b\cdot\nabla + \frac{\hbar}{2m}\Delta\right)(x\cdot mb). \tag{70}$$

Using adjoints[2], one has that

$$\left\langle D\left(\sum q_i \cdot m b_i\right)\right\rangle = \int d^N x (x\cdot mb)\left(-\frac{\partial}{\partial t} - b\cdot\nabla - \nabla\cdot b + \frac{\hbar}{2m}\Delta\right)\rho(x). \tag{71}$$

However, ρ satisfies the Fokker–Planck equation[2]

$$\frac{\partial \rho}{\partial t} = -\nabla\cdot(b\rho) + \frac{\hbar}{2m}\Delta\rho. \tag{72}$$

Thus, the last term in (68) is indeed zero.

Thus, one has the virial theorem

$$2\langle T \rangle = \left\langle \sum q_i \cdot \frac{\partial V}{\partial x_i}(q_i)\right\rangle. \tag{73}$$

Note that the averages are ensemble averages instead of time averages as in the classical case.[29] However, since the process is assumed to be ergodic, (73) is equivalent to a virial theorem with time averages. The equivalence is not apparent in QM.[30] Indeed, the interpretation of the virial theorem in QM is not even clear.

One may also note that no boundedness assumptions were made on the sample paths of q, as is done in the classical case.[29] Here stationarity is the substitute for boundedness. In QM these considerations are not even treated,[30] although it is an interesting question when one considers that the quantum-mechanical propagator is also made up of classical trajectories.[18]

In the case where V is homogeneous of degree k, one may apply Euler's theorem as in the classical case.[29] One obtains

$$2\langle T \rangle = k\langle V \rangle. \tag{74}$$

This was used in the second section of this paper in order to express $\langle T \rangle$ in terms of $\langle V \rangle$.

For positive temperature, one only has to restrict his attention to periodic solutions $q(t)$. It is also important to note the b_i depend on temperature in general. This is easily seen in the example of the hydrogen atom: For zero temperature, one has

$$\mathbf{b} = -\frac{e^2}{\hbar}\frac{\mathbf{q}}{|\mathbf{q}|} \tag{75}$$

so the average energy per hydrogen atom is

$$\epsilon = \frac{1}{2} m\langle \mathbf{b}^2 \rangle = -\frac{me^4}{2\hbar^2}, \tag{76}$$

which agrees with the quantum result except for interpretation. However, for positive temperature, one sees that Eq. (75) cannot possibly be correct, since one would still get Eq. (76).

Thus, in general, Eqs. (73) and (74) are only useful in expressing $\langle T \rangle$ as a function of $\langle V \rangle$, since the temperature dependence of b is not *a priori* known.[16]

[1]For a more or less complete bibliography, see Refs. 3, 4, and 6 below.
[2]E. Nelson, Phys. Rev. **150**, 1079 (1966); *Dynamical Theories of Brownian Motion* (Princeton U. P., Princeton, New Jersey 1967); "Connection between Brownian motion and quantum mechanics," lecture presented at the Einstein Symposium, Berlin, March 1979.
[3]L. De la Peña and A.M. Cetto, J. Math. Phys. **18**, 1612 (1977); Found. Phys. **8**, 191 (1978).
[4]L. De la Peña and A.M. Cetto, J. Math. Phys. **20**, 469 (1979).
[5]F. Guerra and P. Ruggiero, Phys. Rev. Lett. **31**, 1022 (1973).
[6]S.M. Moore, Found. Phys. **9**, 237 (1979).
[7]S.M. Moore, Lett. Nuovo Cimento **24**, 284 (1979).
[8]E. Nelson, J. Funct. Anal. **12**, 211 (1973).
[9]K. Yasue, Phys. Rev. Lett. **40**, 665 (1978).
[10]K. Yasue, Phys. Rev. D **18**, 532 (1978).
[11]K. Yasue, J. Math. Phys. **19**, 1892 (1978).
[12]R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957); P.C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959).
[13]S. Albeverio and R. Høegh-Krohn, J. Math. Phys. **15**, 1745 (1974).
[14]E. Nelson, Bull. Am. Math. Soc. **84**, 121 (1978).
[15]E. Santos, Nuovo Cimento B **19**, 57 (1974).
[16]Recently, an essentially different and new physical alternative to Nelson's theory has been developed by M. Davidson [Physica (Utrecht) A **96**, 465 (1979)]. There a positive temperature drift velocity can be determined up to fourth order in $\hbar$. However, Davidson's basic equation is Eq. (1) of this paper with the zero temperature dw used even at positive temperature, so his drift velocity contains all the temperature dependence. Hence, it differs from the drift velocity in this paper, which can also be determined in principle by using the Markov limit of CSED and the positive-temperature electromagnetic field of Eq. (62). See the references in Ref. 3.
[17]The author has recently discovered that periodic paths were introduced some time ago by R. Høegh-Krohn [Commun. Math. Phys. **38**, 195 (1974)] in the Euclidean context. There is some resemblance between the results of his Sec. 2 and those of Ref. 7 and this paper, although the interpretation is entirely different.
[18]E.S. Abers and B.W. Lee, Phys. Lett. C **9**, 1 (1973).
[19]C.G. Wick, Phys. Rev. **96**, 1124 (1954).
[20]C.W. Bernard, Phys. Rev. D **9**, 3312 (1974).
[21]A. Robinson, *Non-Standard Analysis* (North-Holland, Amsterdam, 1966).
[22]Since we are only interested in equilibrium processes, the forward derivative D and backward derivative D_* are equal.
[23]These comments are particularly important when one applies other boundary conditions to situations where zero-point energy differences are measurable, e.g., in the Casimir effect. In this context see the interpretation of Lukosz' methods [W. Lukosz, Z. Phys. **258**, 99 (1973)] in Ref. 6.
[24]H. Grabert, P Hänggi, and P. Talkner, Phys. Rev. A **19**, 2440 (1979).
[25]P. Claverie and S. Diner, Ann. Fond. L. de Broglie **1**, 73 (1976).

[26]The moments in Eq. (62) are different than those previously calculated (Refs. 6 and 7) due to the added term $-k_i k_j/|\mathbf{k}|^2$. However, one can show that none of the basic conclusions developed earlier are changed.

[27]F.A. Berezin, *The Method of Second Quantization* (Academic, New York, 1966); C. Nash, *Relativistic Quantum Fields* (Academic, New York, 1978).

[28]S. Albeverio and R. Høegh-Krohn, "Quasi-invariant measures, symmetric diffusion processes and quantum fields," in Les Méthodes Mathématiques de la Théorie Quantique des Champs, Colloq. Int. C. N. R. S. **248**, 11 (1976); Z. Wahrscheinlichkeitstheorie verw. Gebiete **40**, 1 (1977); Z. Haba, Nuovo Cimento A **41**, 470 (1977); J. Math. Phys. **18**, 2133 (1977).

[29]L.D. Landau and E.M. Lifshitz, *Mechanics* (Pergamon, Oxford, 1976), third edition.

[30]E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970), second edition.

On a remarkable class of two-dimensional random walks

F. W. Wiegel [a)]
Department of Chemistry, Stanford University, Stanford, California 94305

(Received 12 March 1980; accepted for publication 14 April 1980)

This study describes some properties of random walks in a plane which differ from free random walks through an extra weightfactor (-1) for every crossing of some branchline T. The statistical distribution of these walks is derived, asymptotically for very long walks, in the case that T consists of a half-line. It is pointed out that these walks are relevant to (1) self-avoiding random walks in a plane; (2) the simple entanglement problem in polymer physics.

1. INTRODUCTION

In this paper we study the properties of random walks in a plane which have a weight which differs in some essential respects from the weight of the free random walks which have been studied extensively throughout the last sixty years.[1,2] The weight of these walks is defined as follows: Imagine some curve T in a plane, which may consist of several disconnected continuous parts. For an arbitrary random walk configuration (C) we determine the number of times $[n(C)]$ that C crosses T. The *a priori* weight $W(C)$ of this configuration is now defined by

$$W(C) = W_0(C)(-1)^{n(C)}, \tag{1}$$

where $W_0(C)$ equals the *a priori* weight of C in the standard case of free random walks. As $W(C)$ can be negative, the usual probabilistic interpretation does not apply. The most important new aspect of $W(C)$ is the fact that its sign depends on whether the total number of crossings between C and T is even or odd; this is a global, rather than a local, property of the configuration C.

We shall take for T the collection of points with Cartesian coordinates (x,y) with $-\infty < x < 0$ and $y = 0$; this will be called the "branchline" for reasons which will become obvious shortly. We shall calculate the function

$$p(x,y,N) \equiv \sum_C W(C), \tag{2}$$

where the summation extends over all those configurations which (1) start at some fixed point (x_0,y_0); (2) consist of N steps; (3) reach (x,y) at the end of the Nth step. This class of random walks is remarkable for two reasons. In the first place, a simple relation exists betwen them and the two-dimensional simple entanglement problem. The latter problem consists of calculating the configuration sum $Q_n(x,y,N)$ over all those configurations which have the properties (1), (2), and (3) as stated above and which, in addition, wind exactly n times around the origin of coordinates. It follows from these definitions that

$$p(x,y,N) = \sum_{n=-\infty}^{+\infty} (-1)^n Q_n(x,y,N). \tag{3}$$

The simple entanglement problem has been studied by Prager and Frisch,[3] Edwards,[4,5] Saito and Chen,[6] and Wiegel.[7,8,9]

The second reason these random walks are remarkable is their relation to the two-dimensional self-avoiding random walk problem. This relation has been discussed recently in detail by the author.[9,10] Actually, the relation with self-avoiding random walks become evident only after a further complex phase factor has been included in the weight. The resulting complex-weighted random walks also play a role in the combinatorial solution of the Ising model and the free-fermion case of the eight-vertex model.[11-17] For these reasons it is useful to study the random walks with weight (1) for their own sake.

2. GENERAL CONSIDERATIONS

We represent a random walk configuration by a set of N steps, each of the same length l. The functions $p(x,y,N)$ were defined in the introduction; they are connected by the recurrence relations

$$p(x,y,0) = \delta(x-x_0)\delta(y-y_0), \tag{4}$$

$$p(x,y,N) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} S(x,y|x-l\cos\alpha, y-l\sin\alpha) \times p(x-l\cos\alpha, y-l\sin\alpha, N-1)\, d\alpha, \tag{5}$$

where $S(x,y|x',y') = -1$ if the straight line which connects (x,y) and (x',y') intersects T; $S = +1$ if no intersection occurs.

Let B denote those points in the plane with a shortest distance to T which is smaller than l. If $(x,y)\notin B$ the integral relation (5) reads

$$p(x,y,N) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} p(x-l\cos\alpha, y-l\sin\alpha, N-1)\, d\alpha. \tag{6}$$

For $N \gg 1$, $p(x,y,N)$ will be a slowly varying function of x, y, and N and the integral relation can be replaced by the diffusion equation

$$\frac{\partial p}{\partial N} = \frac{l^2}{4}\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2}\right). \tag{7}$$

Now consider the original integral relation (5) in the case in which (x,y) is on the branchline T, i.e., $x<0, y=0$. In this case (5) gives

[a)]Permanent address: Department of Applied Physics, Twente Unviersity of Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands.

 0022-2488/80/082111-03$1.00

$$\lim_{y\downarrow 0} p(x,y,N) = -\lim_{y\uparrow 0} p(x,y,N) \quad (x<0). \tag{8}$$

Moreover, if the walks starts on the x axis, every walk from $(x_0,0)$ to an arbitrary point (x,y) gives, after reflection in the x axis, a mirror image which leads from $(x_0,0)$ to $(x,-y)$ and which has exactly the same number of intersections with T. For this case one finds

$$p(x,y,N) = p(x,-y,N), \quad (y_0 = 0). \tag{9}$$

Combination of the last two equations gives the boundary condition

$$p(x,0,N) = 0 \quad (x<0, y_0 = 0). \tag{10}$$

As a consequence of this boundary condition one can calculate the statistical distribution of the walks from the expansion

$$p(x,y,N) = \sum_s f_s(x,y) f_s^*(x_0,0) \exp(-\lambda_s N), \tag{11}$$

where the f_s denote the orthonormal eigenfunctions

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f_s + \frac{4\lambda_s}{l^2} f_s = 0, \tag{12}$$

under the boundary condition

$$f_s(x,0) = 0 \quad \text{for } x<0. \tag{13}$$

This calculation forms the subject of Sec. 3. The same formalism can be used if T consists, not of a single branchline but, for example, of two branchlines $x < -\frac{1}{2}r_0$ and $x > +\frac{1}{2}r_0$, as is the case for the self-avoiding random walk problem (compare the discussion in Refs. 9 and 10). It is also of interest to consider the case in which the walks are restricted to some domain in the plane. If D denotes the circumference of this domain then (11)–(13) need to hold inside D; on D one has to impose the additional boundary condition

$$f_s = 0, \quad (x,y) \in D, \tag{14}$$

in the case of an absorbing boundary, or

$$\frac{\partial f_s}{\partial n} = 0, \quad (x,y) \in D, \tag{15}$$

in the case of a hard boundary; in the last equation $\partial/\partial n$ denotes the derivative in the direction normal to D.

A peculiar consequence of the boundary condition (10) arises in the case $x = 0$. According to (10) one has

$$\lim_{x\downarrow 0} p(x,0,N) = 0. \tag{16}$$

But according to (5) one has also

$$\lim_{x\downarrow 0} p(x,0,N) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} p(x - l\cos\alpha, y - l\sin\alpha, N-1)\, d\alpha \neq 0. \tag{17}$$

Hence the endpoint (0,0) of the branchline T is a point in which the functions $p(x,y,N)$ have a finite discontinuity; this jump is a consequence of the geometric definition of the function S.

3. CALCULATIONAL DETAILS

If T consists of the negative x-axis it appears natural to use polar coordinates (r,θ) with $0<r<\infty$, $-\pi<\theta<+\pi$

$$x = r\cos\theta, \quad y = r\sin\theta. \tag{18}$$

Transforming (12) and (13) to these coordinates one finds

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\right) f + \frac{4\lambda}{l^2} f = 0, \tag{19}$$

$$f(r,\pm\pi) = 0 \quad (0<r<\infty). \tag{20}$$

The even eigenfunctions have the form

$$f(r,\theta) = A(r)\cos(n+\tfrac{1}{2})\theta \quad (n = 0,1,2,\cdots), \tag{21}$$

and the odd eigenfunctions have the form

$$f(r,\theta) = C(r)\sin n\theta \quad (n = 1,2,3,\cdots). \tag{22}$$

The odd eigenfunctions vanish for $\theta = 0$. They will, therefore, not contribute to the sum (11) in which the initial position now has polar coordinates (r_0,θ_0) with $\theta_0 = 0$.

The radial part of the eigenfunction, $A(r)$, has to be solved from the equation

$$\frac{d^2A}{dr^2} + \frac{1}{r}\frac{dA}{dr} + \left(\frac{4\lambda}{l^2} - \frac{(n+\frac{1}{2})^2}{r^2}\right) A = 0. \tag{23}$$

The solution, which has to be finite for $r\downarrow 0$, is given by

$$A(r) = BJ_{n+1/2}\left(\frac{r}{l}\sqrt{4\lambda}\right), \tag{24}$$

where $J_p(z)$ denotes the Bessel function of the first kind. In order to proceed with the calculation we impose the boundary condition (14) if r equals some very large value R; in a later stage of the calculation one takes the limit $R\to\infty$. If $z_{p,m}$ denotes the mth zero on the positive real z-axis of $J_p(z)$ this boundary condition gives the values of λ

$$\lambda_{n,m} = \frac{l^2}{4R^2} z^2_{n+1/2,m} \quad (n = 0,1,2,\ldots, m = 1,2,3,\cdots). \tag{25}$$

The normalization constants $B_{n,m}$ follow from the orthonormality condition

$$\pi B_{n,m} B_{n,m'} \int_0^R r J_{n+1/2}\left(\frac{r}{R} z_{n+1/2,m}\right) J_{n+1/2} \times \left(\frac{r}{R} z_{n+1/2,m'}\right) dr = \delta_{m,m'}. \tag{26}$$

The integral on the left vanishes for $m\neq m'$. If $m = m'$ Eq. 6.521.1 of Gradshteyn and Ryzhik[18] gives

$$B_{n,m} = (2/\pi)^{1/2} R^{-1} \left|J_{n+3/2}(z_{n+1/2,m})\right|^{-1} \quad (n = 0,1,2,\cdots, m = 1,2,3,\cdots). \tag{27}$$

Substituting these results into the eigenfunction expansion (11) one finds

$$p(r,\theta,N) = \frac{2}{\pi R^2} \sum_{n=0}^{\infty} \cos(n+\tfrac{1}{2})\theta \sum_{m=1}^{\infty} J_{n+1/2} \times \left(\frac{r}{R} z_{n+1/2,m}\right) J_{n+1/2}\left(\frac{r_0}{R} z_{n+1/2,m}\right) \cdot \times J^{-2}_{n+3/2}(z_{n+1/2,m}) \exp\left(-\frac{Nl^2}{4R^2} z^2_{n+1/2,m}\right). \tag{28}$$

At this point in the calculation it is convenient to take the limit $R \to \infty$. In this limit the summation over m will be dominated by those terms for which $z_{n+1/2,m} \gg 1$. This implies that one can use the asymptotic formula

$$J_p(z) \cong \sqrt{\frac{2}{\pi z}} \cos(z - \tfrac{1}{2}\pi p - \tfrac{1}{4}\pi) \quad (p > 0, z \gg 1). \tag{29}$$

This implies

$$z_{p,m} \cong \tfrac{1}{2}\pi p + \tfrac{3}{4}\pi + (m-1)\pi, \tag{30}$$

$$J^2_{n+3/2}(z_{n+1/2,m}) \cong 2(\pi z_{n+1/2,m})^{-1}. \tag{31}$$

For large values of R the variable

$$\zeta \equiv \frac{l}{R} z_{n+1/2,m} \tag{32}$$

behaves like a continuous variable. For the number $g(\zeta)\, d\zeta$ of ζ values in the interval $d\zeta$ one finds

$$g(\zeta) = R/l\pi. \tag{33}$$

When the last three equations are substituted into (28) one finds the expression

$$p(r,\theta,N) = (\pi l^2)^{-1} \sum_{n=0}^{\infty} \cos(n+\tfrac{1}{2})\theta \int_0^\infty \zeta J_{n+1/2} \times \left(\frac{r}{l}\zeta\right) J_{n+1/2}\left(\frac{r_0}{l}\zeta\right) \exp\left(-\frac{N}{4}\zeta^2\right) d\zeta. \tag{34}$$

The integral is given by Eq. 6.633.2 of Ref. 18,

$$p(r,\theta,N) = 2(\pi N l^2)^{-1} \exp\left\{-\frac{r_0^2 + r^2}{Nl^2}\right\} \times \sum_{n=0}^{\infty} \cos(n+\tfrac{1}{2})\theta I_{n+1/2}\left(\frac{2r_0 r}{Nl^2}\right), \tag{35}$$

where the $I_{n+1/2}$ denote the modified Bessel functions. The limiting value of p in the branchpoint follows by substituting of the last equation into (17); this gives for the size of the jump in the branchpoint

$$p(0,0,N) = 2\{\pi^2(N-1)l^2\}^{-1} \times \exp\left\{-\frac{1}{N-1} - \frac{r_0^2}{(N-1)l^2}\right\} \sum_{n=0}^{\infty} (-1)^n (n+\tfrac{1}{2})^{-1} I_{n+1/2}\left(\frac{2r_0}{(N-1)l}\right). \tag{36}$$

These results hold asymptotically for $N \gg 1$.

4. CONCLUDING REMARKS

The statistical distribution of the random walks under consideration is given by Eq. (35). Exactly the same result is found if one substitutes the solution of the two-dimensional simple entanglement problem[3-9] into our Eq. (3); this approach has been followed in Ref. 9.

It will be clear from the boundary condition (10) that the branchline functions as a perfect absorber. The total weight W_N of all walks of N steps, which for free random walks is given by

$$W_N^{(0)} = 1, \tag{37}$$

is now given by the expression

$$W_N = \frac{2}{\pi N l} \exp\left(-\frac{r_0^2}{Nl^2}\right) \sum_{n=0}^{\infty} \int_{-\pi}^{+\pi} \cos(n+\tfrac{1}{2})\theta\, d\theta \times \int_0^\infty r \exp\left(-\frac{r^2}{Nl^2}\right) I_{n+1/2}\left(\frac{2r_0 r}{Nl^2}\right) dr. \tag{38}$$

The integral over θ is trivial. The integral over r is first transformed by partial integration; the resulting integral can be calculated with the use of Eqs. 9.6.26 and 11.4.31 of Ref. 19. In this way one finds

$$W_N = (\pi N)^{-1/2} \frac{r_0}{l} \exp\left(\frac{-r_0^2}{2Nl^2}\right) \sum_{n=0}^{\infty} (-1)^n (n+\tfrac{1}{2})^{-1} \times \left[I_{n/2-1/4}\left(\frac{r_0^2}{2Nl^2}\right) + I_{n/2+3/4}\left(\frac{r_0^2}{2Nl^2}\right)\right]. \tag{39}$$

For $r_0^2 \gg Nl^2$ this expression reduces to (37); for $r_0^2 \ll Nl^2$ it has the asymptotic form

$$W_N \cong (8/\pi)^{1/2} \Gamma^{-1}(\tfrac{3}{4})(r_0^2/Nl^2)^{1/4} \cong 1.302226(r_0^2/Nl^2)^{1/4}. \tag{40}$$

This last formula shows the rapid "depletion" of the total weight $W_N \sim N^{-1/4}$ as a result of crossings of the branchline.

ACKNOWLEDGMENT

The author is indebted to Professor John Ross of Stanford University for critical reading of the manuscript.

[1]S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[2]M. N. Barber and B. W. Ninham, *Random and Restricted Walks* (Gordon and Breach, New York, 1970).
[3]S. Prager and H. L. Frisch, J. Chem. Phys. **46**, 1475 (1967).
[4]S. F. Edwards, Proc. Phys. Soc. London **91**, 513 (1967).
[5]S. F. Edwards in *Path Integrals*, edited by G. J. Papadopoulos and J. T. Devreese (Plenum, New York, 1978), p. 285.
[6]N. Saito and Y. Chen, J. Chem. Phys. **59**, 3701 (1973).
[7]F. W. Wiegel, J. Chem. Phys. **67**, 469 (1977).
[8]F. W. Wiegel, "Conformational Phase Transitions in a Macromolecule: Exactly Solvable Models," Sec. V, to appear in *Phase Transitions and Critical Phenomena*, Vol. 8, edited by C. Domb and M. S. Green (Academic, New York).
[9]F. W. Wiegel, "Fluid Flow through Porous Macromolecular Systems," *Lecture Notes in Physics* (Springer, New York, 1980, Vol. **121**).
[10]F. W. Wiegel, Physica A **98**, 345 (1979).
[11]M. Kac and J. C. Ward, Phys. Rev. **88**, 1332 (1952).
[12]S. Sherman, J. Math. Phys. **1**, 202 (1960).
[13]S. Sherman, J. Math. Phys. **4**, 1213 (1963).
[14]P. N. Burgoyne, J. Math. Phys. **4**, 1320 (1963).
[15]N. V. Vdovichenko, Zh. Eksp. Teor. Fiz. **47**, 715 (1964) [Sov. Phys. JETP **20**, 477 (1965)].
[16]F. W. Wiegel, Phys. Lett. A **41**, 225 (1972).
[17]F. W. Wiegel, Can. J. Phys. **53**, 1148 (1975).
[18]I. S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1965).
[19]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).

Nth-order multifrequency coherence functions: A functional path integral approach. II [a)]

Conrad M. Rose
Naval Surface Weapons Center, Dahlgren, Virginia 22448
Ioannis M. Besieris
Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

(Received 17 December 1979; accepted for publication 29 February 1980)

Nth-order multifrequency coherence functions arising in beam propagation through focusing media with random-axis misalignments and focusing media with additive statistical fluctuations are computed. The analysis is carried out by means of a simple formula which yields exact algorithmic solutions to a class of canonical path integrals.

1. INTRODUCTION

In a previous paper,[1] referred to in the sequel as Paper I, a functional (or path) integral applicable to a broad class of randomly perturbed media was constructed for the nth-order multifrequency coherence function, a quantity intimately linked to nth-order pulse statistics. This path integral was subsequently carried out explicitly in the case of a nondispersive, deterministically homogeneous medium characterized by a simplified (quadratic) Kolmogorov spectrum. It is our purpose in this paper to lift the restriction of a background flat medium. Specifically, we shall compute nth-order multifrequency coherence functions arising in beam propagation through focusing media with random-axis misalignments, and focusing media with additive statistical fluctuations. We shall carry out this task by means of a simple basic formula which allows exact algorithmic evaluations to a class of canonical path integrals.

Our work in Paper I was based on the stochastic Cauchy problem

$$\frac{i}{k}\frac{\partial}{\partial z}\psi(\underline{x},z,w;\alpha) = H_{\rm op}\Big(\underline{x}, -\frac{i}{k}\nabla_{\underline{x}}, z,w;\alpha\Big)\psi(\underline{x},z,w;\alpha), \tag{1.1a}$$

$\underline{x}\in R^2,\quad z>0,$

$$H_{\rm op}\Big(\underline{x}, -\frac{i}{k}\nabla_{\underline{x}},z,w;\alpha\Big) = -\frac{1}{2k^2}\nabla^2_{\underline{x}} + V(\underline{x},z,w;\alpha), \tag{1.1b}$$

$$\psi(\underline{x},0,w;\alpha) = \psi_0(\underline{x},w). \tag{1.1c}$$

The "Hamiltonian" $H_{\rm op}$ is a self-adjoint stochastic operator depending on a parameter $\alpha\in A$, (A,F,P) being an underlying probability measure space. In addition, w in (1.1) denotes the radian frequency, $k\equiv k(w)$ the wave number, $\psi(\underline{x},z,w;\alpha)$ the complex random wavefunction, and $V(\underline{x},z,w;\alpha)$, the "potential" field which is assumed to be a real random function. The initial condition $\psi_0(\underline{x},w)$ incorporates all the information concerning the temporal frequency spectrum and the spatial distribution of the source at the initial plane $z=0$.

In the course of this work we shall deal explicitly with the following two distinct categories of the potential field $V(\underline{x},z,w;\alpha)$ entering into (1.1b):

$$\text{(i)}\, V(\underline{x},z,w;\alpha) = \tfrac{1}{2}g^2[x - a\underline{H}(z;\alpha)]^2, \tag{1.2a}$$

$$\text{(ii)}\, V(\underline{x},z,w;\alpha) = \tfrac{1}{2}g^2 x^2 - \tfrac{1}{2}\epsilon_1(\underline{x},z;\alpha), \tag{1.2b}$$

where $x=|\underline{x}|$, a is a constant, and g is a spatial frequency (units: radians/meter). The first category corresponds to a parabolically focusing medium whose equilibrium axis is perturbed via the zero-mean, range-dependent, vector-valued, random function $\underline{H}(z;\alpha)$; on the other hand, the second category represents a medium whose parabolically graded deterministic profile is additively perturbed by the zero-mean random function $\epsilon_1(\underline{x},z;\alpha)$. The absence of the angular frequency w in the right-hand sides of (1.2) signifies that the media are assumed to be nondispersive.[2]

Besides their generic significance in quantum mechanics,[3] Schrödinger-like equations of the form (1.1) and (1.2) play a significant role in plane and beam electromagnetic and acoustic wave propagation. They are usually derived from a scalar Helmholtz equation within the framework of the parabolic (or small-angle) approximation. In this context, the complex stochastic parabolic equation (1.1) with potentials (i) and (ii) provides a good description of the forward propagation of low-order modes in a fiber lightguide having a randomly perturbed parabolically graded refractive index. It can also give some insight into the problem of forward propagation of low-order acoustic modes near an idealized, randomly perturbed, underwater, sound channel axis.

The problem under consideration in this paper, that is the study of nth-order pulse statistics associated with (1.1) and (1.2), can be made more specific as follows: Let $G(\underline{x},\underline{x}',z,w;\alpha)$ denote the fundamental solution (referred to alternatively as the propagator) of the stochastic complex parabolic equation (1.1). It follows, then, from the discussion in Paper I that the examination of pulse propagation in a random medium requires knowledge of the nth-order coherence functions $E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha\}$–at different frequencies and different transverse (with respect to z) coordinates. Here, the operator $E\{\cdot\}$ signifies ensemble averaging, the index n is assumed to be an even integer, $\underline{X}=(\underline{x}_1,\underline{x}_2,\cdots,\underline{x}_n)$

[a)] Research supported in part by the Office of Naval Research under contract No. N00014-76-C-0056.

$\in R^{2n}$, $\underline{X}' = (\underline{x}'_1, \underline{x}'_2, \cdots, \underline{x}'_n) \in R^{2n}$, and $\underline{w} = (w_1, w_2, \cdots, w_n) \in R^{n}$; finally, the n^{th}-order propagators $G^{(n)}$ are defined in terms of G as follows:

$$G^{(n)}(\underline{X}, \underline{X}', z, \underline{w}; \alpha) = \prod_{p=1}^{n/2} G^*(\underline{x}_{2p}, \underline{x}'_{2p}, z, w_{2p}; \alpha) \times G(\underline{x}_{2p-1}, \underline{x}'_{2p-1}, z, w_{2p-1}; \alpha). \tag{1.3}$$

The fundamental solution $G(\underline{x}, \underline{x}', z, w; \alpha)$ to (1.1) can be expressed as a continuous functional path integral. This can be used, in turn, as a basis for constructing a path integral representation for the n^{th}-order quantity $G^{(n)}(\underline{X}, \underline{X}', z, w; \alpha)$; specifically, $G^{(n)}(\underline{X}, \underline{X}', z, \underline{w}; \alpha)$

$$= \int d[\underline{X}(\zeta)] \exp\left\{\frac{i}{2} \sum_{p=1}^{n} \xi_p k_p \times \int_0^z d\zeta \left\{\dot{\underline{x}}_p^2(\zeta) - 2V[\underline{x}_p(\zeta), \zeta, w_p; \alpha]\right\}\right\}, \tag{1.4}$$

where $d[\underline{X}(\zeta)] = d[\underline{x}_1(\zeta)] d[\underline{x}_2(\zeta)] \cdots d[\underline{x}_n(\zeta)]$ is the usual Feymann path differential measure; $\xi_p = 1, p$ odd, $\xi_p = 1, p$ even; $k_p = k(w_p)$; and the integration is over all paths $\underline{x}_p(\zeta)$ $p = 1, 2, ..., n$, subject to the boundary conditions $\underline{x}_p(0) = \underline{x}'_p, x_p\ \underline{x}_p(z) = \underline{x}_p$.

The ensemble-averaged version of the path integral (1.4) for potentials (i) and (ii) [cf. Eq. (1.2)] will be evaluated in Secs. 3 and 4, respectively. These computations will be made on the basis of a simple formula which will be derived in the next section.

2. DERIVATION OF A BASIC FORMULA

A. Focusing Medium with Random-axis Misalignments: First-order Moment

Consider the stochastic parabolic equation (1.1) with a slightly modified potential of type (i) [cf. Eq. (1.2a)], viz.,

$$\frac{i}{k} \frac{\partial}{\partial z} \psi(x, z, w; \alpha) = -\frac{1}{2k^2} \nabla^2_{\underline{x}}\, \psi(\underline{x}\ z, w; \alpha) + \tfrac{1}{2} g^2\, \underline{x}^2 \psi(\underline{x}, z, w; \alpha) - g^2\, \bar{a} \underline{x} \cdot \underline{H}(z; \alpha) \psi(\underline{x}, z, w; \alpha) + \tfrac{1}{2}\lambda g^2 \bar{a}^2\, H^2(z; \alpha) \times \psi(\underline{x}, z, w; \alpha). \tag{2.1}$$

The extra multiplicative factor λ appearing in (2.1) should be set equal to unity in order to maintain a correspondence between (2.1) and (1.1), with potential (i) given in (1.2a). The importance of this factor will be made clear in the sequel. It should also be noted that the constant factor "a" incorporated in (1.2a) has been changed to "$\bar{a}$" in (2.1), again for reasons which will be explained later on.

We set as our immediate goal the derivation of an equation for the first moment of the wavefunction $\psi(\underline{x}, z, w; \alpha)$. In order to carry out this task, we ensemble-average both sides of (2.1) over the realizations $\alpha \in A$:

$$\frac{i}{k} \frac{\partial}{\partial z} E\{\psi(\underline{x}, z, w; \alpha)\} = \frac{1}{2k^2} \nabla^2_{\underline{x}}\, E\{\psi(\underline{x}, z, w; \alpha)\} + \tfrac{1}{2}g^2 x^2 E\{\psi(\underline{x}, z, w; \alpha)\} - g^2 \bar{a} \underline{x} \cdot E\{H(z; \alpha)\, \psi(\underline{x}, z, w; \alpha)\} + \tfrac{1}{2}\lambda g^2 \bar{a}^2\, E\{H^2(z; \alpha) \psi(\underline{x}, z, w; \alpha)\}. \tag{2.2}$$

To proceed further, we shall need expressions for the last two terms on the right-hand side of (2.2), viz.,

$$\underline{x} \cdot E\{\underline{H}(z; \alpha) \psi(\underline{x}, z, w; \alpha)\} = \sum_{j=1}^{2} x_j\, E\{H_j(z; \alpha) \psi(\underline{x}, z, w; \alpha)\}, \tag{2.3a}$$

$$E\{H^2(z; \alpha) \psi(\underline{x}, z, w; \alpha)\} = \sum_{j=1}^{2} E\{H_j^2(z; \alpha) \psi(\underline{x}, z, w; \alpha)\}, \tag{2.3b}$$

in terms of the first- and higher-order moments of $\psi(\underline{x}, z, w; \alpha)$. This "closure" problem will be examined next for a special class of random functions $\underline{H}(z; \alpha)$.

Let $\underline{H}(z; \alpha)$ be a zero-mean, wide-sense stationary, Gaussian random process with autocovariance tensor

$$E\{H_j(z; \alpha) H_k(z'; \alpha)\} = \Gamma_{jk}(z - z'),\ j, k = 1, 2. \tag{2.4}$$

It follows, then, from the Donsker–Furutsu–Novikov functional formalism[4-6] that

$$E\{H_j(z; \alpha) \psi(\underline{x}, z, w; \alpha)\} = \sum_{k=1}^{2} \int_0^z dz'\, \Gamma_{jk}(z - z') E\{\delta\psi(\underline{x}, z, w; \alpha)/\delta H_k(z'; \alpha)\}, \tag{2.5a}$$

$$E\{H_j^2(z; \alpha) \psi(\underline{x}, z, w; \alpha)\} = \Gamma_{jj}(0)\, E\{\psi(\underline{x}, z, w; \alpha)\} + \sum_{k=1}^{2} \int_0^z dz'\, \Gamma_{jk}(z - z')\, E\{H_j(z; \alpha) \times [\delta\psi(\underline{x}, z, w; \alpha)/\delta H_k(z'; \alpha)]\}, \tag{2.5b}$$

where $\delta(\cdot)/\delta(\cdot)$ denotes functional differentiation.

Unless further restrictions are imposed on the process $\underline{H}(z; \alpha)$, it turns out that the computation of the functional derivatives together with the performance of the ensemble averaging entering into (2.5) lead to an infinite hierarchy,[7] which, in turn, exhibits the impossibility of "closing" Eq. (2.2) for $E\{\psi(\underline{x}, z, w; \alpha)\}$. Closure may be effected by truncating this infinite hierarchy. Such a truncation leads to well-known statistical approximations (e.g., the first-order smoothing and the direct-interaction approximation).

In the following discussion we shall eliminate altogether the aforementioned closure difficulties by restricting the process $\underline{H}(z; \alpha)$ to be both isotropic and δ-correlated, viz.,

$$\Gamma_{jk}(z - z') = (\sigma^2/2) \delta_{jk} \delta(z - z'),$$

where σ is a constant. With this assumption, (2.5a) and (2.5b) simplify to[8,9]

$$E\{H_j(z; \alpha) \psi(\underline{x}, z, w; \alpha)\} = (\sigma^2/2) E\{\delta\psi(\underline{x}, z, w; \alpha)/\delta H_j(z; \alpha)\}, \tag{2.6a}$$

$$E\{H_j^2(z; \alpha) \psi(\underline{x}, z, w; \alpha)\} = \Gamma_{jj}(0)\, E\{\psi(\underline{x}, z, w; \alpha) + (\sigma^2/2)\, E\{H_j(z; \alpha)\, [\delta\psi(\underline{x}, z, w; \alpha)/\delta H_j(z; \alpha)]\}. \tag{2.6b}$$

The functional derivative $\delta\psi/\delta H_j$ required in (2.6) can be found from the original stochastic complex parabolic equation (2.1). Omitting intermediate steps, we present here the final result:

$$\frac{\delta\psi(\underline{x},z,w;\alpha)}{\delta H_j(z;\alpha)} = ikg^2\bar{a}x_j\psi(\underline{x},z,w;\alpha) - ik\lambda g^2\bar{a}^2H_j(z;\alpha)\psi(\underline{x},z,w;\alpha). \quad (2.7)$$

Equation (2.7) together with (2.6) result in closed form solutions for $E\{H_j\psi\}$ and $E\{H_j^2\psi\}$ in terms of $E\{\psi\}$. When these two expressions are used then in conjunction with (2.3) and (2.2), the desired equation for the first statistical moment of ψ is obtained; specifically,

$$\frac{i}{k}\frac{\partial}{\partial z}E\{\psi(\underline{x},z,w;\alpha)\} = -\frac{1}{2k^2}\nabla^2_{\underline{x}}E\{\psi(\underline{x},z,w;\alpha)\} + V_e(\underline{x})E\{\psi(\underline{x},z,w;\alpha)\},\ z>0, \quad (2.8a)$$

$$E\{\psi(\underline{x},0,w;\alpha)\} = \psi_0(\underline{x},w), \quad (2.8b)$$

with the effective potential $V_e(\underline{x})$ given as

$$V_e(\underline{x}) = B(\lambda;k;\bar{a}) + \tfrac{1}{2}\Omega^2(\lambda;k;\bar{a})x^2, \quad (2.8c)$$

where

$$B(\lambda;k;\bar{a}) = \tfrac{1}{4}\lambda g^2\bar{a}^2\sigma^2[1+ik\lambda g^2\bar{a}^2(\sigma^2/2)]^{-1}\sum_{j=1}^{2}\Gamma_{jj}(0), \quad (2.8d)$$

$$\Omega^2(\lambda;k;\bar{a}) = g^2\{1 - ikg^2\bar{a}^2\sigma^2[1+k\lambda g^2(\sigma^2/4)] \times[1+ik\lambda g^2\bar{a}^2(\sigma^2/2)]^{-2}\}. \quad (2.8e)$$

B. The Propagator of Eq. (2.8)

Let $G(\underline{x},\underline{x}',z,w;\alpha)$ denote the propagator of the stochastic parabolic equation (2.1). It is defined by means of the relationship

$$\psi(\underline{x},z,w;\alpha) = \int_{R^2}d\underline{x}'\,G(\underline{x},\underline{x}',z,w;\alpha)\,\psi_0(\underline{x}',w). \quad (2.9)$$

Averaging both sides of the last equation gives rise to the expression

$$E\{\psi(\underline{x},z,w;\alpha)\} = \int_{R^2}d\underline{x}'\,E\{G(\underline{x},\underline{x}',w;\alpha)\}\psi_0(\underline{x}',w). \quad (2.10)$$

The quantity $E\{G(\underline{x},\underline{x}',w;\alpha)\}$ in (2.10) is clearly the propagator of (2.8). It can be written as a continuous path integral, viz.,

$$E\{G(\underline{x},\underline{x}',w;\alpha)\} = \exp[-ikB(\lambda;k;\bar{a})z] \times\int d[\underline{x}(\zeta)]\exp\{ik\int_0^z d\zeta\,[\tfrac{1}{2}\dot{\underline{x}}^2(\zeta) - \tfrac{1}{2}\Omega^2(\lambda;k;\bar{a})\underline{x}^2(\zeta)]\}, \quad (2.11)$$

which can be carried out explicitly,[10] resulting in the expression

$$E\{G(\underline{x},\underline{x}',w;\alpha)\} = \exp[-ikB(\lambda;k;\bar{a})z] \times(k\Omega/2\pi i\sin\Omega z)\exp\{(ik\Omega/2\sin\Omega z) \times[(x^2+x'^2)\cos\Omega z - 2\underline{x}\cdot\underline{x}']\}. \quad (2.12)$$

C. The Propagator of Eq. (2.1)

The fundamental solution (or propagator) $G(\underline{x},\underline{x}',z,w;\alpha)$ of (2.1) [cf. also the definition in (2.9)] can be expressed as a continuous path integral; specifically,

$$G(\underline{x},\underline{x}',z,w;\alpha) = \exp\left\{-\frac{i}{2}kg^2\bar{a}^2\int_0^z d\zeta\,H^2(\zeta;\alpha)\right\} \times\int d[\underline{x}(\zeta)]\exp\left\{+\frac{i}{2}k\int_0^z d\zeta\left[\dot{\underline{x}}^2(\zeta) - g^2\underline{x}^2(\zeta) + 2g^2\bar{a}\underline{H}(\zeta;\alpha)\cdot\underline{x}(\zeta)\right]\right\}. \quad (2.13)$$

This path integral can be performed without difficulty[11]:

$$G(\underline{x},\underline{x}',z,w;\alpha) = (kg/2\pi i\sin gz) \times\exp\{(ikg/2\sin gz)[(x^2+x'^2)\cos gz - 2\underline{x}\cdot\underline{x}'] + (ik\bar{a}g^2/\sin gz)\int_0^z d\zeta\,[\underline{x}'\sin g(z-\zeta) + \underline{x}\sin g\zeta] \cdot\underline{H}(\zeta;\alpha) - (ik\bar{a}^2g^3/\sin gz)\int_0^z d\zeta\int_0^\zeta d\zeta' \times\sin g(z-\zeta)\sin g\zeta'\,\underline{H}(\zeta;\alpha)\cdot\underline{H}(\zeta';\alpha) - (ikg^2\bar{a}^2/2)\int_0^z d\zeta\,H^2(\zeta;\alpha)\}. \quad (2.14)$$

D. Basic Formula

Clearly, the quantity $E\{G(\underline{x},\underline{x}',z,w;\alpha)\}$ obtained by formally averaging both sides of (2.14) must be the same with the result derived in Sec. 2B [cf. Eq. (2.12)]. This observation leads to the following relationship:

$$E\left\{\exp\left\{(ik\bar{a}g^2/\sin gz)\int_0^z d\zeta\,[\underline{x}'\sin g(z-\zeta) + \underline{x}\sin g\zeta] \cdot\underline{H}(\zeta;\alpha) - (ik\bar{a}^2g^3/\sin gz)\int_0^z d\zeta \times\int_0^\zeta d\zeta'\sin g(z-\zeta)\sin g\zeta'\,\underline{H}(\zeta;\alpha)\cdot\underline{H}(\zeta';\alpha) - (ik\lambda g^2\bar{a}^2/2)\int_0^z d\zeta\,H^2(\zeta;\alpha)\right\}\right\}$$
$$= (\Omega/2\pi i\sin\Omega z)(g/2\pi i\sin gz)^{-1}\exp(-ikBz) \times\exp\{(kg/2i\sin gz)[(x^2+x'^2)\cos gz - 2\underline{x}\cdot\underline{x}']\} \times\exp\{(ik\Omega/2\sin\Omega z)[(x^2+x'^2)\cos\Omega z - 2\underline{x}\cdot\underline{x}']\}. \quad (2.15)$$

It should be noted in connection with formula (2.15) that the various constants, as well as $\underline{x}$ and $\underline{x}'$, need not be those associated with the original set in (2.14). Performing the averaging of (2.14) requires finding the $\underline{H}(z;\alpha)$ that makes the argument of the exponential an extremum. This $\underline{H}(z;\alpha)$, however, does not depend on the specific values of quantities such as $\underline{x}$ and $\underline{x}'$ which are defined at the end points only. Thus, $\underline{x}$ and $\underline{x}'$ may be any functions of the endpoints (within reason), and (2.15) will still hold.

The above important observation will be used in the following two sections in order to compute a series of nth-order multifrequency moments.

3. FOCUSING MEDIUM WITH RANDOM-AXIS MISALIGNMENTS: Nth-ORDER MULTIFREQUENCY MOMENTS

Substituting the potential field given in (1.2a) into (1.4) we obtain

$$G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)$$
$$= \int d[\underline{X}(\zeta)] \exp\Big\{\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p \int_0^z d\zeta\, [\dot{\underline{x}}_p^2(\zeta) - g^2 x_p^2(\zeta) + 2g^2 a \underline{x}_p(\zeta)\cdot\underline{H}(\zeta;\alpha) - \tfrac{1}{2}g^2 a^2 H^2(\zeta;\alpha)]\Big\}. \quad (3.1)$$

This path integral can be carried out explicitly [cf. definition (1.3); also, (2.13) and (2.14) with $\bar{a}\to a$, $\lambda\to 1$]:

$$G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)$$
$$= \Big(\prod_{i=1}^{n}\xi_i k_i\Big)(g/2\pi i \sin gz)^n \exp\Big\{(ig/2\sin gz) \times \sum_{p=1}^{n}\xi_p k_p\, [(x_p^2 + x_p'^2)\cos gz - 2\underline{x}_p\cdot\underline{x}_p']$$
$$+ \Big\{(iag^2/\sin gz)\int_0^z d\zeta \Big[\Big(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p'\Big)\sin g(z-\zeta) + \Big(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p\Big)\sin g\zeta\Big]\cdot\underline{H}(\zeta;\alpha)$$
$$- (ia^3 g^3/\sin gz)\Big(\sum_{p=1}^{n}\xi_p k_p\Big)\int_0^z d\zeta\int_0^{\zeta} d\zeta' \times \sin g(z-\zeta)\sin g\zeta'\, \underline{H}(\zeta;\alpha)\cdot\underline{H}(\zeta';\alpha)$$
$$- (ig^2 a^2/2)\Big(\sum_{p=1}^{n}\xi_p k_p\Big)\int_0^z d\zeta\, H^2(\zeta;\alpha)\Big\}\Big\}. \quad (3.2)$$

In order to evaluate the desired nth-order multifrequency coherence quantity $E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$ from (3.2), we must first compute the following statistical average:

$$E\Big\{\exp\Big\{(iag^2/\sin gz)\int_0^z d\zeta\Big[\Big(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p'\Big)\sin g(z-\zeta) + \Big(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p\Big)\sin g\zeta\Big]\cdot\underline{H}(\zeta;\alpha)$$
$$- (iag^2 \sin gz)\Big(\sum_{p=1}^{n}\xi_p k_p\Big)\int_0^z d\zeta\int_0^{\zeta} d\zeta' \times \sin g(z-\zeta)\sin g\zeta'\,\underline{H}(\zeta;\alpha)\cdot\underline{H}(\zeta';\alpha)$$
$$- (ig^2 a^2/2)\Big(\sum_{p=1}^{n}\xi_p k_p\Big)\int_0^z d\zeta\, H^2(\zeta;\alpha)\Big\}\Big\}. \quad (3.3)$$

This expression, however, can be brought into a one-to-one correspondence with the left-hand side of (2.15) provided that the following changes are made in the latter:

$$\lambda\to 1, \quad (3.4)$$

$$\underline{x}\to \sum_{p=1}^{n}\xi_p k_p \underline{x}_p, \quad (3.5a)$$

$$\underline{x}'\to \sum_{p=1}^{n}\xi_p k_p \underline{x}_p', \quad (3.5b)$$

$$\bar{a}\to a\Big(\sum_{p=1}^{n}\xi_p k_p\Big), \quad (3.6a)$$

$$k\to \Big(\sum_{p=1}^{n}\xi_p k_p\Big)^{-1}. \quad (3.6b)$$

It follows, therefore, that the statistical average (3.3) is equal to the right-hand side of the basic formula (2.15) if $\underline{x}$ and $\underline{x}'$ are those given in (3.5), and the quantities $kB(\lambda;k;\bar{a}) \equiv B'(\lambda;k;\bar{a})$ and $\Omega(\lambda;k;\bar{a})$ [cf. Eq. (2.8)] are evaluated at the arguments λ, $\bar{a}$, and k given in (3.4) and (3.6).

We present next the final form of the main result of this section:

$$E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$$
$$= \Big(\prod_{p=1}^{n}\xi_p k_p\Big)(g/2\pi i \sin gz)^{n-1}(\Omega/2\pi i \sin\Omega z) \times \exp(-iB'z)\exp\Big\{(i\Omega/2\sin\Omega z)\Big(\sum_{p=1}^{n}\xi_p k_p\Big)^{-1}$$
$$\times \Big\{\Big[\Big(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p\Big)^2\Big] + \Big(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p')\Big)^2\Big\}\cos\Omega z - 2\Big(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p\Big)\cdot\Big(\sum_{q=1}^{n}\xi_q k_q \underline{x}_q'\Big)\Big\}\Big\}$$
$$\times \exp(ig/2\sin gz)\Big(2\sum_{p=1}^{n}\xi_p k_p\Big)^{-1} \times \sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q\, [(\underline{x}_p - \underline{x}_q)^2 + (\underline{x}_p' - \underline{x}_q')^2] \times \cos gz - 2(\underline{x}_p - \underline{x}_q)\cdot(\underline{x}_p' - \underline{x}_q')\}. \quad (3.7)$$

4. FOCUSING MEDIUM WITH ADDITIVE STATISTICAL FLUCTUATIONS: Nth-ORDER MULTIFREQUENCY MOMENTS

Under the influence of the potential field (1.2b), the average of (1.4) over the realizations $\alpha\in A$ yields the expression

$$E\{G^{(n)}(\underline{X},\underline{X}',z,w;\alpha)\}$$
$$= \int d[\underline{X}(\zeta)] \exp\{\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p \times \int_0^z d\zeta\, [\dot{\underline{x}}_p^2(\zeta) - g^2 \underline{x}_p^2(\zeta)]\} \times E\{\exp\{\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p \int_0^z d\zeta\, \epsilon_1[\underline{x}_p(\zeta),\zeta;\alpha]\}\}. \quad (4.1)$$

To proceed further, we need to specify the structure of $\epsilon_1[\underline{x}_p(\zeta),\zeta;\alpha]$. If the latter is assumed to be a Gaussian random process, the statistical averaging appearing in (4.1) can be carried out explicitly, with the result

$$I_1 \equiv E\{\exp\{\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p \int_0^z d\zeta\,\epsilon_1[\underline{x}_p(\zeta),\zeta;\alpha)]\}\}$$
$$= \exp\{-\frac{1}{8}\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q \times \int_0^z d\zeta\int_0^{\zeta} d\zeta'\,\gamma[\underline{x}_p(\zeta),\underline{x}_q(\zeta'),\zeta,\zeta']\}, \quad (4.2)$$

where γ is the correlation function of the random process ϵ_1, viz.,

$$\gamma[\underline{x}_p(\zeta),\underline{x}_q(\zeta'),\zeta,\zeta'] = E\{\epsilon_1[\underline{x}_p(\zeta),\zeta;\alpha]\epsilon_1[\underline{x}_q(\zeta'),\zeta';\alpha]\}. \quad (4.3)$$

We resort, next, to the usual Markovian approximation (cf., also, Paper I, Sec. 3B), i.e., we assume that the process ϵ_1 is δ-correlated along the longitudinal direction of propagation. We have, then, in the place of (4.3)

$$\gamma[\underline{x}_p(\zeta),\underline{x}_p(\zeta'),\zeta,\zeta'] = A[\underline{x}_p(\zeta),\underline{x}_q(\zeta')]\delta(\zeta-\zeta'). \quad (4.4)$$

The corresponding expression for I_1 is given as follows:

$$I_1 = \exp\left\{ -\frac{1}{8} \sum_{p=1}^{n} \sum_{q=1}^{n} \xi_p \xi_q k_p k_q \times \int_0^z d\zeta A\,[\underline{x}_p(\zeta),\underline{x}_q(\zeta)]\right\}. \qquad (4.5)$$

In many physical problems, the transverse correlation $A\,[\underline{x}_p(\zeta),\underline{x}_q(\zeta)]$ is homogeneous, isotropic, and of a power-law type (cf. Paper I and references therein), viz.,

$$A\,[\underline{x}_p(\zeta),\underline{x}_q(\zeta)] = A\,(0)\left\{1 - \frac{1}{2}\left[\frac{1}{L_0}\,|\underline{x}_p(\zeta) - \underline{x}_q(\zeta)|\right]^{\beta}\right\}, \qquad (4.6)$$

where L_0 is a characteristic length, and the parameter β is usually within the range $1 < \beta < 4$.

Even under the restrictive assumptions made so far about the statistical characteristics of the random process ϵ_1, it is impossible to evaluate $E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$ exactly unless the parameter β introduced in (4.6) is equal to 2. For values of β different from 2, the most comprehensive contribution to the evaluation of n^{th}-order multifrequency coherence functions can be found in the recent work by Dashen,[12] whereby $E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$, with n even, is asymptotically expressible in terms of two-frequency mutual coherence functions $E\{G^{(2)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$. It should be pointed out, however, that, in contradistinction to single-frequency mutual coherence functions, the exact integration of the two-frequency quantities $E\{G^{(2)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$ still constitutes an open problem. The only recourse presently is to rely on approximate techniques. An excellent contribution along this direction was made recently by Furutsu[13] who examined second-order pulse statistics for an initially pulsed planar source distribution propagating in a channel devoid of a deterministic background profile.

In the following we shall restrict the discussion to the case $\beta = 2$ (simplified or quadratic Kolmogorov spectrum.)[14] Under this assumption, (4.1) assumes the form

$$\begin{aligned} E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\} &= \exp\{-\frac{1}{8}A\,(0)(\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q)z\} \\ &\times \int d\,[\underline{X}(\zeta)]\exp\left\{\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p \times \int_0^z d\zeta\left[\dot{\underline{x}}_p^2(\zeta) - g^2\underline{x}_p^2(\zeta) - \frac{i}{4}D\sum_{q=1}^{n}\xi_q k_q\,[\underline{x}_p(\zeta) - \underline{x}_q(\zeta)]^2\right]\right\}, \end{aligned} \qquad (4.7)$$

where $D = A\,(0)/2L_0^{\,2}$. With $g = 0$, this is precisely the path integral evaluated in Paper I. It is possible to modify the technique developed in that paper so that it can accommodate the presence of a focusing background profile ($g \neq 0$). Instead of following such a procedure in this paper, however, we shall recast (4.7) in a form which, when used in conjunction with the basic formula (2.15), will yield a solution for $E\{G^{(2)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$ in a straightforward manner.

We begin by expanding the quadratic form entering into (4.7) and recombining terms:

$$\begin{aligned} &-\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p\left\{\frac{i}{4}D\sum_{q=1}^{n}\xi_q k_q\,[\underline{x}_p(\zeta) - \underline{x}_q(\zeta)]^2\right\} \\ &= -\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p\left\{\frac{i}{2}D(\sum_{q=1}^{n}\xi_q k_q)\underline{x}_p^2(\zeta) - \frac{i}{2}D(\sum_{r=1}^{n}\xi_r k_r)^{-1}\left[\sum_{q=1}^{n}\xi_q k_q \underline{x}_q(\zeta)\right]^2\right\}. \end{aligned} \qquad (4.8)$$

With this change, (4.7) becomes

$$\begin{aligned} E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\} &= \exp\{-\frac{1}{8}A\,(0)(\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q z\} \\ &\times \int d\,[\underline{X}(\zeta)]\exp\left\{\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p \times \int_0^z d\zeta\left[\dot{\underline{x}}_p^2(\zeta) - \bar{g}^2\underline{x}_p^2(\zeta) + \frac{i}{2}D\left(\sum_{r=1}^{n}\xi_r k_r\right)^{-1}\left[\sum_{q=1}^{n}\xi_q k_q \underline{x}_q(\zeta)\right]^2\right]\right\}, \end{aligned} \qquad (4.9)$$

where the complex-valued spatial frequency $\bar{g}$ is defined as follows:

$$\bar{g} = \left[g^2 + i(D/2)(\sum_{q=1}^{n}\xi_q k_q)\right]^{1/2}. \qquad (4.10)$$

We introduce next a fictitious zero-mean, Gaussian, vector-valued, real process $\underline{F}(z;\alpha)$ with autocovariance tensor $E\{F_i(z;\alpha)F_j(z';\alpha)\} = (\frac{1}{2})\delta_{ij}\delta(z - z')$, $i,j = 1,2$. It is well-known in this case that

$$\int d\,[\underline{F}(\zeta;\alpha)]\exp\left\{i\int_0^z d\zeta\,\underline{y}(\zeta)\cdot\underline{F}(\zeta;\alpha) - \frac{1}{2}\int_0^z d\zeta\,\underline{F}^2(\zeta;\alpha)\right\} = \exp\left\{-\frac{1}{2}\int_0^z d\zeta\,\underline{y}^2(\zeta)\right\}. \qquad (4.11)$$

With the specific choice

$$\underline{y}(\zeta) = (D/2)^{1/2}\sum_{p=1}^{n}\xi_p k_p \underline{x}_p(\zeta), \qquad (4.12)$$

it follows from (4.9) and (4.11) that

$$\begin{aligned} E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\} &= \exp\left\{-\tfrac{1}{8}A\,(0)\left(\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q\right)z\right\} \\ &\times \int d\,[\underline{X}(\zeta)]\int d\,[\underline{F}(\zeta;\alpha)]\,P\,[\underline{F}(\zeta;\alpha)] \\ &\times \exp\left\{\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p\int_0^z d\zeta\,[\dot{\underline{x}}_p^2(\zeta) - \bar{g}^2\underline{x}_p^2(\zeta) + 2(D/2)^{1/2}\underline{x}_p(\zeta)\cdot\underline{F}(\zeta;\alpha)]\right\}, \end{aligned} \qquad (4.13)$$

where $P\,[\underline{F}(z;\alpha)]$ is the probability distribution functional of the process $\underline{F}(z;\alpha)$.

The path integral with respect to $\underline{X}(\zeta)$ in (4.13) is isomorphic to n uncoupled quantum mechanical harmonic oscillators; it can, therefore, be performed easily. The final result is

$$E\left\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\right. = \exp\left\{-\tfrac{1}{8}A\,(0)\left(\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q\right)z\right\}$$

$$\times \left(\prod_{p=1}^{n} \xi_p k_p\right) (\bar{g}/2\pi i \ \sin\bar{g}z)^n \exp\{(i\bar{g}/2 \sin\bar{g}z) \times \sum_{q=1}^{n} \xi_q k_q [(\underline{x}_q^2 + \underline{x}_q'^2)\cos\bar{g}z - 2\underline{x}_q\cdot\underline{x}_q']\} I_2, \qquad (4.14)$$

where

$$I_2 = E\left\{\exp\left\{[i(D/2)^{1/2}/\sin\bar{g}z \int_0^z d\zeta \left[\left(\sum_{p=1}^{n} \xi_p k_p \underline{x}_p'\right) \times \sin\bar{g}(z-\zeta) + \left(\sum_{p=1}^{n} \xi_p k_p \underline{x}_p\right)\sin\bar{g}\zeta\right] \cdot \underline{F}(\zeta;\alpha) - (iD/2\bar{g}\sin\bar{g}z)\left(\sum_{p=1}^{n}\xi_p k_p\right) \times \int_0^z d\zeta \int_0^\zeta d\zeta' \sin\bar{g}(z-\zeta)\sin\bar{g}\zeta' \underline{F}(\zeta;\alpha)\cdot\underline{F}(\zeta;\alpha)\right\}\right\}. \qquad (4.15)$$

We are now in a position to use the basic formula derived in Sec. 2D. We note that the expression for I_2 given (4.15) can be brought into a direct correspondence with the left-hand side of the basic formula (2.15) by means of the following changes:

$$\lambda \to 0, \qquad (4.16a)$$

$$\sigma \to 1, \qquad (4.16b)$$

$$g \to \bar{g}, \qquad (4.16c)$$

$$\underline{x} \to \sum_{p=1}^{n} \xi_p k_p \underline{x}_p, \qquad (4.16d)$$

$$\underline{x}' \to \sum_{p=1}^{n} \xi_p k_p \underline{x}_p', \qquad (4.16e)$$

$$\bar{a} \to (D/2)^{1/2}\bar{g}^{-2} \sum_{p=1}^{n} \xi_p k_p, \qquad (4.16f)$$

$$k \to \left(\sum_{p=1}^{n} \xi_p k_p\right)^{-1}. \qquad (4.16g)$$

As a consequence, the statistical average (4.15) is equal to the right-hand side of the basic formula (2.15). In the latter, g, $\underline{x}$ and $\underline{x}'$ are those given in (4.16c)–(4.16e); B must be set equal to zero by virtue of (4.16a); finally, Ω must be evaluated at the values of λ, σ, g, $\bar{a}$, and k specified above. (When this is done, one has the simple relationship $\Omega \to g$).

We present next the solution of the problem under consideration in this section:

$$E\{G^{(n)}(\underline{X}, X', z, \underline{w}; \alpha)\} = \exp\left\{-\tfrac{1}{8}A(0)\left(\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q\right)z\right\} \times \left(\prod_{p=1}^{n}\xi_p k_p\right)(2\pi i)^{-n}(\bar{g}/\sin\bar{g}z)^{n-1}(g/\sin gz) \times \exp\left\{(i\bar{g}/2\sin\bar{g}z)\left(2\sum_{r=1}^{n}\xi_r k_r\right)^{-1}\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q \times \{[(\underline{x}_p - \underline{x}_q)^2 + (\underline{x}_p' - \underline{x}_q')^2]\cos\bar{g}z - 2(\underline{x}_p - \underline{x}_q)\cdot(\underline{x}_p' - \underline{x}_q')\}\right\} \times \exp\left\{(ig/2\sin gz)\left(\sum_{r=1}^{n}\xi_r k_r\right)^{-1} \times \left\{\left[\left(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p\right)^2 + \left(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p'\right)^2\right]\cos gz - 2\left(\sum_{p=1}^{n}\xi_p k_p \underline{x}_p\right)\cdot\left(\sum_{q=1}^{n}\xi_q k_q \underline{x}_q'\right)\right\}\right\}. \qquad (4.17)$$

In the absence of a focusing background channel, i.e., $g = 0$, (4.17) coincides with the main result in Paper I [cf. Eq. (6.36)].

5. CONCLUDING REMARKS

Our main contribution in this paper is the computation of a set of nth-order multifrequency statistical moments $E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$ by means of a simple formula [cf. Eq. (2.15)] which provides straightforward solutions to a class of canonical path integrals. Alternative methods for obtaining these coherence functions, such as direct evaluation of the corresponding path integrals (cf. Paper I), or integration of the local transport equations satisfied by $E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$, are more difficult.

Once the quantities $E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\}$ are known, physically measurable pulse statistics are contained in the nth-order moments given in Eq. (2.11) of Paper I, viz.,

$$E\left\{\prod_{p=1}^{n} u_r(\underline{x}_p, z, t_p; \alpha)\right\} = \frac{1}{(2\pi)^n}\int_{R^n} d\underline{w} \int_{R^{2n}} d\underline{X}'\, E\{G^{(n)}(\underline{X},\underline{X}',z,\underline{w};\alpha)\} F_r^{(n)}(\underline{w}) \times \psi_0^{(n)}(\underline{X}',\underline{w}) \exp\left\{\sum_{p=1}^{n}(-i)\xi_p[w_p t_p - k(w_p)z]\right\}. \qquad (5.1)$$

Here, $u_r(\underline{x},z,t;\alpha)$ is a real field of radiation (acoustic or electromagnetic) whose time Fourier transform $U_r(\underline{x},z,w;\alpha)$ is linked to the wavefunction $\psi(\underline{x},z,w;\alpha)$ [cf. Eq. (1.1)] by the relation $U_r(\underline{x},z,w;\alpha) = \psi(\underline{x},z,w;\alpha)\exp[ik(w)z]$; furthermore, $F_r^{(n)}(\underline{w}) = F_r^*(w_2)F_r(w_1)\cdots F_r^*(w_n)F_r(w_{n-1})$, where $F_r(w)$ is the temporal spectrum—usually a bandpass function of w—characterizing the receiver at range z, and $\psi_0^{(n)}(\underline{X}',\underline{w}) = \psi_0^*(\underline{x}_2,w_2)\psi_0(\underline{x}_1,w_1)\cdots\psi_0^*(\underline{x}_n,w_n)\psi_0(\underline{x}_{n-1},w_{n-1})$, where $\psi_0(\underline{x},w)$ is the initial distribution associated with the stochastic parabolic equation (1.1).

It was demonstrated in Paper I that under very restrictive assumptions ($g = 0$, a broadband receiver, i.e., $F_r^{(2)}(w)\simeq 1$, and an impulsive planar source intensity), (5.1), with $n = 2$ and $\underline{x}_2 = \underline{x}_1$, can be evaluated exactly. If these conditions are relaxed, however, the nth-order moments shown in (5.1) can be computed only asymptotically (e.g., by the method of steepest descent), or numerically.

[1]C.M. Rose and I.M. Besieris, J. Math. Phys. 20, 1530 (1979).

[2]This statement is incorrect unless the wave number $k(w)$ is a linear function of the radian frequency w.

[3]I.M. Besieris, W.B. Stasiak, and F.D. Tappert, J. Math. Phys. 19, 359 (1978).

[4]M.D. Donsker, "On Function Space Integrals," in *Analysis in Function Space*, edited by W.T. Martin and I. Segal (M.I.T., Cambridge, Massachusetts, 1964).

[5]K. Furutsu, J. Res. Natl. Bur. Stand. D 67, 303 (1963).

[6]E.A. Novikov, Zh. Eksp. Teor. Fiz. 47, 191 (1964) [Sov. Phys. JETP 20, 1290 (1965)].

[7]V.I. Klyatskin and V.I. Tatarskii, Izv. Vyssh. Uchebn. Zaved. Radiofiz. 14, 1400 (1971).

[8]For practical considerations, it is only meaningful to specify a sufficiently wideband process $\underline{H}(z;\alpha)$ with finite power $\Sigma_{j=1}^{n}\Gamma_{jj}(0)$. This point is discussed further by Mazychuk (cf. Ref. 9).

[9]A.V. Mazychuk, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **21**, 217 (1978).
[10]R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw–Hill, New York, 1065). p. 63.
[11]Cf. Ref. 10, p. 64.
[12]R. Dashen, J. Math, Phys. **20**, 894 (1979).
[13]K. Furutsu, J. Math. Phys. **20**, 617 (1979).
[14]It should be noted that this model must be used with care since it behaves in a nonphysical manner for large spatial separations.

A simple approximate determination of stochastic transition for the standard mapping

Michael P. Mortell
The Registrar, University College, Cork, Ireland

Brian R. Seymour
Department of Mathematics and Institute of Applied Mathematics and Statistics University of British Columbia, Vancouver, B.C., Canada

(Received 6 December 1979; accepted for publication 20 March 1980)

The standard mapping results from a study of nonlinear forced oscillations of a gas in a closed tube. On interpreting stochastic transition as the breaking of all waves during a round trip in the tube, an estimate for the critical value of the relevant parameter is simply derived. The estimate agrees very well with the result of a more elaborate analysis of Greene.

1. INTRODUCTION

The study of a measure-preserving mapping of a plane onto itself arises in a number of problems in nonlinear dynamics.[1-3] While these mappings appear simple and *deterministic*, their solutions may be either ordered or *chaotic*, depending on the parameters of the system. The *standard mapping*.

$$F(\eta) = F(s) + A\sin(2\pi\eta), \quad \eta = s + F(s), \tag{1}$$

is a mapping which has received particular attention. It has been used to describe the motion of a particle in a "magnetic bottle," the oscillations of a pendulum with two degrees of freedom, and the motion of a particle constrained to move on the surface of nonsymmetrical bowl (see Greene[1] and Chirikov[2]). These papers derive the mapping (1) from appropriate Hamiltonian systems and suggest that many other dynamical systems can be approximated by the same mapping, hence the term *standard mapping*. One objective in studying this mapping has been to determine the maximum value of $A = A_c$ at which the transition from an ordered to a chaotic motion occurs.[1,2] The most recent estimate of A_c seems to be[1]

$$A_c = 0.971635(2\pi)^{-1}. \tag{2}$$

The standard mapping has also arisen in *nonlinear acoustics*, in the study of the nonlinear oscillatory motion of a gas in a closed tube which is driven by a reciprocating piston, see Seymour and Mortell,[4,5] and Mortell and Seymour.[6,7] It is the aim of this paper to give a brief derivation of the standard mapping in the context of nonlinear acoustics and to show that, from the vantage point of this physical model, we can very simply obtain the approximation

$$A_c = (2\pi)^{-1}. \tag{3}$$

This is close to the value of Eq. (2), which is found by a more elaborate analysis, and better than values in Ref. 2. The result (3) has been given as an aside in Mortell and Seymour,[7] but it is worthwhile to derive it here in the context of stochastic transition.

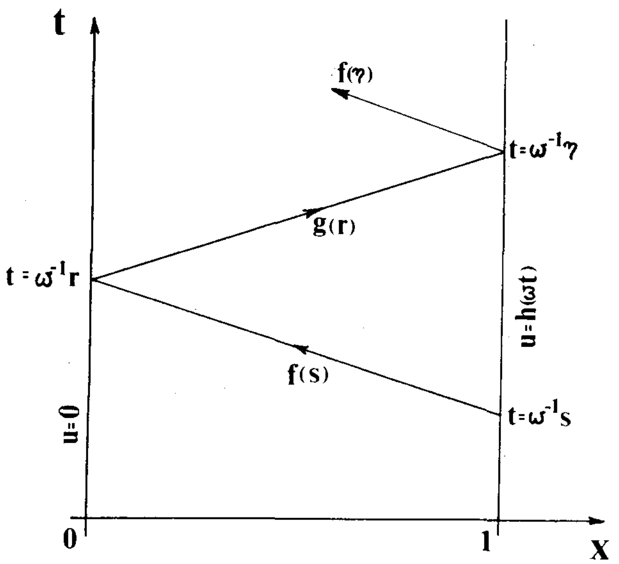


FIG. 1.

2. HEURISTIC DERIVATION OF THE STANDARD MAPPING

A careful derivation of the standard mapping as it arises in nonlinear acoustics has been given previously,[4,7] so here we sketch a derivation which is intuitively appealing. A column of gas is contained in a tube, one end of which is closed ($x = 0$), while at the other end ($x = 1$ in normalized Lagrangian coordinates) there is an oscillating piston. Figure 1 depicts the (t-x) space for the tube. The velocity of the piston at $x = 1$ is specified by $u = h(\omega t)$, where the period of h, in the variable ωt, has been normalized to unity, and $|h| \ll 1$. The dimensionless piston frequency is ω.

The Riemann representation for the particle velocity is

$$u = f(\beta) - g(\alpha), \tag{4}$$

where f, g are Riemann invariants; $\alpha(x,t)$ and $\beta(x,t)$ are the left-traveling and right-traveling nonlinear characteristics. We consider a wave with amplitude $f(s)$ which leaves the piston at $x = 1$ at time $t = \omega^{-1}s$. From the representation (4) and the condition $u(0,t) = 0$, this wave is reflected from the end $x = 0$ with amplitude $g(r) = f(s)$, at time $t = \omega^{-1}r$. It arrives back at $x = 1$ at time $t = \omega^{-1}\eta$. Having received the

 0022-2488/80/082121-03$1.00

input $h(\eta)$ from the piston at $x = 1$, it leaves with amplitude $f(\eta)$. Thus, from Eq. (4),

$$f(\eta) = f(s) + h(\eta). \quad (5)$$

To complete the determination of f we need an expression for the travel time, $\omega^{-1}(\eta - s)$. According to small amplitude nonlinear theory (i.e., $|f| \ll 1$ in the characteristic equations) the travel time is

$$\omega^{-1}(\eta - s) = 2 + 2Mf(s), \quad (6)$$

where $M = \frac{1}{4}(\gamma + 1)$ and γ is the ratio of specific heats for the gas. In Eq. (6), the term 2 is the linear contribution to the travel time, while the term $2Mf(s)$ is the first nonlinear correction. It is typical in such nonlinear problems that the travel time of a wave depends on the amplitude carried, which produces a distortion of the wave as it propagates. The results (5) and (6) may also be derived on the basis that the disturbance in the gas is the superposition of two small amplitude *simple waves*.[4]

The standard mapping (1) now results, on seeking f with unit period, by introducing

$$F(\eta) = 2\omega Mf(\eta) + \Delta, \quad H(\eta) = 2\omega Mh(\eta) = A\sin 2\pi\eta, \quad (7)$$

where $\Delta = 2(\omega - \omega_n)$ and $\omega_n = \frac{1}{2}n$, $n = 1,2,3,\cdots$, are the *linear resonant frequencies*. Δ is a measure of the detuning and, as f is required to have zero mean value if it is periodic, Eq. (7) implies

$$\int_0^1 F(s)\,ds = \Delta. \quad (8)$$

The acceleration and frequency parameters A and Δ, are determined from experimental conditions.

In the context of nonlinear acoustics Eq. (1) can be considered as the product, $T2T1$, of two mappings

$$T1: (s, F(s)) \to (\eta, \hat{F}(\eta)), \quad (9)$$

where $\hat{F}(\eta) = F(s(\eta))$, $\eta = s + F(s)$, can be regarded as a "simple-wave mapping". The function $\hat{F}(\eta)$ represents the distorted signal returning to the piston after reflection from $x = 0$, but before it has been reinforced by the piston motion.

$$T2: (\eta, \hat{F}(\eta)) \to (\eta, F(\eta)), \quad (10)$$

where $F(\eta) = \hat{F}(\eta) + H(\eta)$, then represents the action of the piston on $\hat{F}(\eta)$. To be an acceptable physical solution, F must not only map onto itself under $T2T1$, but must also satisfy the mean condition (8).

The mapping (1) has also been written in the *discrete* form[7]

$$F_{n+1} = F_n + H(x_{n+1}), \quad x_{n+1} = x_n + F_n, \quad (11)$$

which is equivalent to

$$F_{n+1} = F_n - H(x_n), \quad x_{n+1} = x_n + F_{n+1}, \quad (12)$$

the form used by Greene.[1]

3. STOCHASTIC TRANSITION

Rather than using Eq. (1) in the form (11) to map *discrete points*, it is more natural in nonlinear acoustics to use it to map *curves*. Thus if $F(s)$ is the signal which leaves the boundary $x = 1$, its image in the next cycle of the motion is $F(\eta)$ given by (1). If $F(s)$ is singlevalued, but its image $F(\eta)$ is multivalued, we say that the wave has "broken" during its round trip in the tube. In general, a wave will "break" at some point $\eta = \eta^*$ if

$$1 + F'(s^*) = 0, \quad \eta^* = s^* + F(s^*). \quad (13)$$

In the context of nonlinear acoustics, this corresponds to *shock waves* occurring in the flow. Such multivalued solutions of Eq. (1) are constructed in Ref. 5, where it is shown that each will contain an infinite number of multivalued loops. For a given value of A less than some value A_c, both continuous and multivalued solutions are possible depending on the specified value of Δ. The continuous solutions are known as *invariant curves* and correspond, in the context of periodic orbits of a nonlinear oscillator, to KAM (Kolmogorov, Arnol'd, and Moser) surfaces. For $A > A_c$ the solutions are multivalued for *all* values of Δ. Thus $A = A_c$ corresponds to the disappearance of KAM surfaces and hence to stochastic transition. The *chaotic* solutions of Greene[1] occur for those values of A and Δ at which the acoustic signal will break in a traversal of the tube. The *minimum* signal amplitude which will arise for a given $H(\eta)$ is when the frequency is not in the neighborhood of a linear resonant frequency ($\Delta = 0$), and hence the signal leaving the piston is never significantly reinforced. The condition that this signal, $H(\eta)$, breaks in one cycle is, by Eq. (13),

$$\max_{\eta} [-H'(\eta)] \geqslant 1. \quad (14)$$

When $H(\eta) = A\sin(2\pi\eta)$, condition (14) becomes the following: The wave will break in one cycle for *all* values of Δ if $A \geqslant A_c$, where

$$A_c = (2\pi)^{-1}. \quad (15)$$

This result was derived in Mortell and Seymour[7] in the context of nonlinear acoustics, where it was shown to be consistent with a numerical investigation of the transition curve in the (A-Δ) plane bounding the region of continuous periodic solutions. Several *exact* continuous solutions for a piecewise linear piston velocity $H(\eta)$ have also been constructed[6,8] in the vicinity of $\Delta = 0$ and $\Delta = \frac{1}{2}$. In Ref. 5 the concept of a discontinuous invariant is introduced when the "equal area rule" is added to Eqs. (1) and (8). These curves describe the discontinuous motions in the various resonance regions. Theoretically, resonance regions occur for Δ in the neighborhood of each rational (m/n) $(m < n)$, but the width of the region in the A-Δ plane decreases as m and n increase. Consequently, their detection depends on the sophistication of the numerical scheme used. A small amount of damping in the system will eliminate most of the higher order resonances in many practical situations, and indeed only the resonances near $\Delta = 0$ and $\Delta = \frac{1}{2}$ have been observed experimentally in the context of nonlinear acoustics.[7]

ACKNOWLEDGMENTS

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant No. A9117, and was completed while one of the authors (M.P.M.) was the 1979 Visitor in Applied Mathemat-

ics in the Department of Mathematics, University of Queensland, Brisbane, Australia.

[1]J. M. Greene, "A method for determining a stochastic transition," J. Math. Phys. **20**, 1183 (1979).
[2]B. V. Chirikov, "A universal instability of many-dimensional operators systems," Phys. Rep. **52**, 263 (1979).
[3]C. F. F. Karney, "Stochastic ion heating by a lower hybrid wave," Phys. Fluids **21**, 1584 (1978).
[4]B. R. Seymour and M. P. Mortell, "Resonant acoustic oscillations with damping: small rate theory," J. Fluid Mech. **58**, 353 (1973).
[5]B. R. Seymour, and M. P. Mortell, "A finite rate theory of resonance in a closed tube: discontinuous solutions of a functional equation," J. Fluid Mech. (to appear).
[6]M. P. Mortell, B. R. Seymour, "Exact solutions of a functional equation arising in nonlinear wave propagation," SIAM J. Appl. Math. **30**, 587 (1976).
[7]M. P. Mortell, and B. R. Seymour, "Nonlinear forced oscillations in a closed tube: continuous solutions of a functional equation," Proc. R. Soc. London Ser. A **367** (1979).
[8]B. R. Seymour, and M. P. Mortell, "Discontinuous solutions of a measure-preserving mapping," I.A.M.S. Tech. Report No. 79, 38 (1979).

Poincaré–Cartan integral invariant and canonical transformations for singular lagrangians

D. Dominici
Istituto Nazionale di Fisica Nucleare, Sezione di Firenze Istituto di Fisica Teorica dell'Università di Firenze, Firenze, Italia
J. Gomis
Department de Física Teòrica, Universitat de Barcelona, Barcelona, Spain

(Received 7 May 1979; accepted for publication 30 January 1980)

In this work we develop the canonical formalism for constrained systems with a finite number of degrees of freedom by making use of the Poincaré–Cartan integral invariant method. A set of variables suitable for the reduction to the physical ones can be obtained by means of a canonical transformation. From the invariance of the Poincaré–Cartan integral under canonical transformations we get the form of the equations of motion for the physical variables of the system.

1. INTRODUCTION

It is known that many interesting physical systems are described by singular Lagrangian. Some examples are provided by the electromagnetic, the gravitational, the Yang–Mills fields, and some relativistic models.[1] Features of all these theories are the invariance under certain transformations and the presence of relations (constraints) among the canonical variables, which restrict the motion to a hypersurface of the phase space.

A method for developing the canonical formalism and the quantization of constrained systems was proposed by Dirac.[2] The constraints are classified into two groups (first-class and second-class), depending on their algebraic properties with respect to Poisson brackets. The dynamics of the system is generated by an extended Hamiltonian, obtained by adding a linear combination of first-class constraints to the canonical one. One must take into account the presence of second-class constraints by working with generalized Poisson brackets (Dirac brackets). The problem of the quantization is complicated by the search for a set of variables independent and canonical with respect to Dirac brackets. Instead of following the Dirac technique, these variables can be directly obtained, as suggested by Shanmugadhasan,[3] as a subset of the variables of a canonical transformation, whose existence is based on some theorems on involutory systems.[4,5] and function groups.[6] We want to stress that this method, as well as the Dirac brackets technique, is a local one; in fact the existence of the canonical transformation is only locally guaranteed.[4]

In this work we pursue the study of the extension of the formalism of the Poincaré–Cartan integral invariant to constrained systems with a finite number of degrees of freedom, which one of us began in Ref. 7, and making use of the invariance of the Poincaré–Cartan integral under canonical transformations, the equations of motion for a set of variables free with respect to second-class constraints are easily obtained. Furthermore, working in this reduced space of the variables independent with respect to second-class constraints, a canonical transformation which isolates the gauge-independent variables from the gauge-dependent ones is performed. This is the great advantage of this technique with respect to the Dirac one. An interesting result is that, for Lagrangians homogeneous of first-degree in the velocities, this procedure corresponds to the Hamilton–Jacobi method.

In Sec. 2 we review and extend the Poincaré–Cartan integral formalism for constrained systems. Section 3 is devoted to the introduction of the concept of canonical transformation and to the proof of the invariance of the Poincaré–Cartan integral under canonical transformations. In Sec. 4 we perform the canonical transformation extended to the second-class constraints and the Hamilton equations for the new variables are obtained. In Sec. 5 the Hamilton equations for the set of variables free with respect to first- and second-class constraints are obtained.

2. POINCARÉ–CARTAN INTEGRAL INVARIANT FOR CONSTRAINED SYSTEMS

The Poincaré–Cartan integral invariant plays a fundamental role in standard classical mechanics since, from its invariance, it follows that the equations of motion of the dynamical system are Hamilton canonical equations.[8]

In Ref. 7 this result was generalized to systems described by singular Lagrangians.

Let us now review the essential points of this generalization. Let us consider a dynamical system described by a singular Lagrangian

$$L = L(q_s, \dot{q}_s, t), \quad (s = 1,\dots,n). \tag{2.1}$$

Due to the singularity of the Lagrangian, the motion of the system is restricted to a hypersurface of the phase space, determined by a set of constraints. Let

$$\Omega_\alpha(q_s, p_s) = 0, \quad (\alpha = 1,\dots,T - W), \tag{2.2}$$

be first-class constraints and

$$\Omega_\beta(q_s, p_s) = 0, \quad (\beta = T - W + 1,\dots,T) \tag{2.3}$$

be second-class.

Making a general variation of the action

 0022-2488/80/082124-07$1.00

$$W = \int_{t_0}^{t_1} dt\, L, \qquad (2.4)$$

it is possible to show that the integral

$$I = \oint_C (p_s \delta q_s - H_c \delta t), \qquad (2.5)$$

calculated along an arbitrarty closed contour lying on the hypersurface S of the extended phase space (q_s, p_s, t), defined by Eqs. (2.2) and (2.3), is invariant under an arbitrary displacement (with deformation) of the contour along any tube of dynamical trajectories. H_c is the canonical Hamiltonian of the system. I is called the Poincaré–Cartan integral invariant.

Let us now review the proof of the following theorem with some details.

Theorem 1: Let us suppose to have a dynamical system, constrained by Eqs. (2.2) and (2.3), whose trajectories satisfy a system of first order differential equations involving arbitrary functions l_α $(\alpha = 1,...,T - W)$

$$\frac{d}{dt} q_s \approx f_s(q_s, p_s, t., l_\alpha) \quad \frac{d}{dt} p_s \approx g_s(q_s, p_s, t, l_s), \qquad (2.6)$$

where the sign $\approx$ (weak equality) means equality on the hypersurface S [defined by Eqs. (2.2) and (2.3)]. Let H_c be a function with the property

$$\{\Omega_\alpha, H_c\} \approx 0. \qquad (2.7)$$

Then, the necessary and sufficient for Eqs. (2.6) be Hamilton equations is that the Poincaré–Cartan integral (2.5) be invariant.

Proof: Firstly, see that the invariance of the Poincaré–Cartan integral is a sufficient condition.

Following the book of Gantmacher, we introduce an auxiliary variable μ, supplementing Eq. (2.6) with one more equation

$$\frac{dq_1}{f_1} = \cdots = \frac{dq_n}{f_n} = \frac{dp_1}{g_1} = \cdots = \frac{dp_n}{g_n} = dt = \pi d\mu, \qquad (2.8)$$

π being an arbitrary function in the extended phase space. For each determination of the l_α's we find, integrating Eqs. (2.8),

$$\begin{cases} q_s = q_s(\mu; q_s^0, p_s^0, t_0) \\ p_s = p_s(\mu; q_s^0, p_s^0, t_0), \\ t = t(\mu; q_s^0, p_s^0, t_0) \end{cases} \qquad (2.9)$$

where q_s^0, p_s^0, t_0 are the initial values, corresponding to $\mu = 0$, which lie on the hypersurface S. In order to obtain a tube of dynamical trajectories (2.9), we choose the initial points on a closed curve, parametrized by means of α, contained in S.

The parametric equations for the dynamical paths that form the tube are

$$q_s = q_s(\mu,\alpha), \quad p_s = p_s(\mu,\alpha), \quad t = t(\mu,\alpha) \quad (0 \leqslant \alpha \leqslant l). \qquad (2.10)$$

The value of α isolates a generatrix of the tube while μ fixes a definite point on this generatrix. Assuming $\mu =$ const, Eqs. (2.10) define a closed curve embracing the tube; by calculating the integral along it, we get $I = I(\mu)$.

If we agree that d means differentiation with respect to μ and δ with respect to α, by invariance we have

$$dI = \oint [dp_s \delta q_s + p_s d\delta q_s - dH_c \delta t - H_c d\delta t] = 0. \qquad (2.11)$$

Integrating by parts, dividing by $d\mu = dt/\pi$ and using Eqs. (2.6) we get

$$\oint \left\{ \left(g_s + \frac{\partial H_c}{\partial q_s} \right) \delta q_s + \left(-f_s + \frac{\partial H_c}{\partial p_s} \right) \delta p_s + \left(-\frac{dH_c}{dt} + \frac{\partial H_c}{\partial t} \right) \delta t \right\} \pi = 0. \qquad (2.12)$$

Since π is an arbitrary factor we obtain

$$\left(g_s + \frac{\partial H_c}{\partial q_s} \right) \delta q_s + \left(-f_s + \frac{\partial H_c}{\partial p_s} \delta p_s \right) + \left(-\frac{dH_c}{dt} + \frac{\partial H_c}{\partial t} \right) \delta t \approx 0. \qquad (2.13)$$

The δq_s and the δp_s are not independent, since C must belong to S. So they must satisfy

$$\begin{aligned} \frac{\partial \Omega_\alpha}{\partial q_s} \delta q_s + \frac{\partial \Omega_\alpha}{\partial p_s} \delta p_s &= 0, \\ \frac{\partial \Omega_\beta}{\partial q_s} \delta q_s + \frac{\partial \Omega_\beta}{\partial p_s} \delta p_s &= 0. \end{aligned} \qquad (2.14)$$

Introducing a set of Lagrangian multipliers l_α, l_β $(\alpha = 1,...,T - W, \beta = T - W + 1,...,T)$, from Eqs. (2.13) and (2.14) we deduce

$$\begin{aligned} g_s &\approx -\frac{\partial H_c}{\partial q_s} - l_\alpha \frac{\partial \Omega_\alpha}{\partial q_s} - l_\beta \frac{\partial \Omega_\beta}{\partial q_s} \\ f_s &\approx \frac{\partial H_c}{\partial p_s} + l_\alpha \frac{\partial \Omega_\alpha}{\partial p_s} + l_\beta \frac{\partial \Omega_\beta}{\partial p_s}. \end{aligned} \qquad (2.15)$$

By requiring that the hypersurface be stationary, the l_β's can be determined

$$l_\beta \approx c_{\beta\beta'} \{\Omega_{\beta'}, H_c\}, \qquad (2.16)$$

where

$$c_{\beta\beta'} \{\Omega_{\beta'}, \Omega_{\beta''}\} \approx \delta_{\beta\beta''}. \qquad (2.17)$$

The Eqs. (2.6) can be written in the form

$$\begin{cases} \dot{q}_s \approx \dfrac{\partial H}{\partial p_s} = \{q_s, H\} \\ \dot{p}_s \approx -\dfrac{\partial H}{\partial q_s} = \{p_s, H\}, \end{cases} \qquad (2.18)$$

with

$$H = H_c + l_\alpha \Omega_\alpha - \Omega_\beta c_{\beta\beta'} \{\Omega_{\beta'}, H_c\}. \qquad (2.19)$$

By following an analogous reasoning and starting from Eqs. (2.18) and (2.19), it is possible to show the invariance of the Poincaré–Cartan integral (2.5). Thus the proof of the theorem is complete.

3. CANONICAL TRNASFORMATIONS AND POINCARÉ–CARTAN INTEGRAL INVARIANT

Let us now extend the concept of canonical transformation to constrained systems, by introducing, as in standard classical mechanics, the following

Definition: Given a dynamical system, constrained by Eqs. (2.2) and (2.3), whose equations of motion are given by Eq. (2.18), a transformation

$$Q_s = Q_s(q, p,t), \quad P_s = P_s(q, p,t) \quad (s = 1,...,n), \quad (3.1)$$

is called *canonical* if a function K_c exists so that Eqs. (2.18) become

$$\dot{Q}_s \approx \frac{\partial K}{\partial P_s} = \{Q_s,K\}, \quad \dot{P}_s \approx -\frac{\partial K}{\partial Q_s} = \{P_s,K\}, \quad (3.2)$$

with

$$K = K_c + l_\alpha \tilde{\Omega}_\alpha - \tilde{\Omega}_\beta \tilde{c}_{\beta\beta'} \left[\{\tilde{\Omega}_{\beta'},K_c\} + \frac{\partial \tilde{\Omega}_{\beta'}}{\partial t}\right], \quad (3.3)$$

and

$$\tilde{c}_{\beta\beta'} \{\tilde{\Omega}_{\beta'},\tilde{\Omega}_{\beta''}\} = \delta_{\beta\beta''}.$$

The $\tilde{\Omega}_\alpha(Q,P,t)$ and $\tilde{\Omega}_\beta(Q,P,t)$ appearing in Eq. (3.3) are obtained from Eqs. (2.2) and (2.3) by substitution of variables. The structure of K is suggested by Theorem 1 and guarantees the stationarity of the hypersurface of the constraints. The extra term $\partial\tilde{\Omega}_\beta/\partial t$ is due to the explicit dependence on t of the canonical transformation.

Following the usual procedure of standard classical mechanics we will prove the following theorem:

Theorem 2: Let Eq. (2.18) be the equations of motion of a dynamical system; a transformation

$$Q_s = Q_s(q,p,t), \quad P_s = P_s(q,p,t), \quad (3.4)$$

for which two functions K_c and F exist so that

$$p_s\delta q_s - H_c\delta t = P_s\delta Q_s - K_c\delta t - \delta F \quad (3.5)$$

is canonical.

Proof: From Eq. (3.5) we have

$$\oint_C [p_s\delta q_s - H_c\delta t - (P_s\delta Q_s - K_c\delta t)] = 0, \quad (3.6)$$

where C is an arbitrary closed contour in the extended phase space, that we will take lying on S. Let $\tilde{C}$ be the contour obtained from C by means of the transformation (3.4). Then the Poincaré–Cartan integral is invariant under the considered transformation. In fact, from Eq. (3.6) we get

$$\oint_C (p_s\delta q_s - H_c\delta t) = \oint_{\tilde{C}} (P_s\delta Q_s - K_c\delta t). \quad (3.7)$$

The left-hand side of Eq. (3.7) is invariant under displacement of the contour along the tube of the dynamical trajectories, solutions of Eq. (2.18) and lying on S. The right-hand side will be invariant under displacement of the contour $\tilde{C}$ along the tube obtained by means of the transformation (3.4) from the proceeding. On the other hand, the transformed trajectories obey a system of first order differential equations. Thus, by repeating the proof of Theorem 1 and by taking into account the explicit dependence of the constraints on the time, we get

$$\dot{Q}_s \approx \frac{\partial K}{\partial P_s}, \quad \dot{P}_s \approx -\frac{\partial K}{\partial Q_s}, \quad (3.8)$$

with K given by Eq. (3.3), and therefore the transformation is canonical.

4. A SET OF CANONICAL VARIABLES INDEPENDENT WITH RESPECT TO THE SECOND-CLASS CONSTRAINTS

In Ref. 7, as we have reviewed in Sec. 2, the Hamilton equations for a constrained system have been obtained [Eq. (2.18)]. The variables q_s and p_s are not independent, since they must satisfy Eqs. (2.2) and (2.3). A suitable method for isolating the true independent variables has been developed by Shanmugadhasan.[3] His theory is based on two theorems on function groups[6,9] and involutory systems[4,5] that we recall without giving proofs.

Theorem 3: A noncommutative function group G of rank r is a subgroup of a group of rank $2n$ whose basis $(\phi_1,...,\phi_n,\psi_1,...,\psi_n)$ can be chosen so that

$$\{\phi_i,\phi_j\} = \{\psi_i,\psi_j\} = 0, \quad \{\psi_j,\phi_i\} = \delta_{ij}, \quad (i,j = 1,...,n). \quad (4.1)$$

Theorem 4: A system of $2m + q$ independent equations (defining a surface S_D of dimension $D = 2n - 2m - q$)

$$\Omega_\tau = 0 \quad (\tau = 1,...,2m + q), \quad (4.2)$$

such that

$$\text{rank}\|\{\Omega_\sigma,\Omega_{\sigma'}\}\| = 2m, \quad (\sigma,\sigma' = 1,...,2m + q), \quad (4.3)$$

can be substituted by a *locally* equivalent system

$$\phi_\lambda = 0 \quad (\lambda = 1,...,m + q), \qquad \psi_\alpha = 0 \quad (\alpha = 1,...,m), \quad (4.4)$$

for which the relations

$$\{\phi_\lambda,\phi_\mu\} = \{\psi_\alpha,\psi_\beta\} = 0, \qquad \{\psi_\alpha,\phi_\lambda\} = \delta_{\alpha\lambda} \quad (4.5)$$

hold locally in the phase space.

First let us apply the last theorem to the set of second-class constraints[10] [Eqs. (2.3)]. Let

$$Q_f = 0, \quad P_f = 0, \quad (f = n_2 + 1,...,n), \quad (4.6)$$

$(n_2 = n - W/2)$ be the locally equivalent system such that

$$\{Q_f,P_{f'}\} = \delta_{ff'}, \quad \{Q_f,Q_{f'}\} = \{P_f,P_{f'}\} = 0. \quad (4.7)$$

The set $G = \{Q_f,P_f, f = n_2 + 1,...,n\}$ now forms a noncommutative function group. Theorem 3 enables us to find a $2n$-dimensional function group which contains G. Let

$$\{Q'_s,P'_s;s = 1,...,n\} \equiv \{Q'_j,P'_j,Q_f,P_f; j = 1,...,n_2, f = n_2 + 1,...,n\} \quad (4.8)$$

denote this function group. Due to the equations

$$\{Q'_s,P'_{s'}\} = \delta_{ss'}, \quad \{Q'_s,Q'_{s'}\} = \{P'_s,P'_{s'}\} = 0, \quad (4.9)$$

and denoting the new Hamitonian by K_c and the generating function by F, we will have, in the usual way,[8]

$$p_s\delta q_s - H_c\delta t = P'_s\delta Q'_s - K_c\delta t - \delta F. \quad (4.10)$$

We can now apply Theorem 2 and deduce that the transformation

$$q_s,p_s,t \to Q'_s,P'_s,t \tag{4.11}$$

is canonical. The equations of motion (3.2) become

$$\dot{Q}'_j \approx \{Q'_j,K\}, \quad \dot{P}'_j \approx \{P'_j,K\} \quad (j=1,\dots,n_2) \tag{4.12}$$

$$\dot{Q}_f \approx \{Q_f,K\}, \quad \dot{P}_f \approx \{P_f,K\} \quad (f=n_2+1,\dots,n), \tag{4.13}$$

with

$$K = K_c + l_\alpha \tilde{\Omega}_\alpha - \tilde{\Omega}_\beta \tilde{c}_{\beta\beta'} \left[\{\tilde{\Omega}_{\beta'},K_c\} + \frac{\partial \tilde{\Omega}_{\beta'}}{\partial t} \right], \tag{4.14}$$

where

$$\tilde{\Omega}_{\alpha,\beta}(Q'_s,P'_s,t) = \Omega_{\alpha,\beta}[q_s(Q'_sP'_st), p_s(Q'_sP'_st)] \tag{4.15}$$

are the expressions of the constraints in the new variables.

The term $\partial\tilde{\Omega}_\beta/\partial t$ in the Hamiltonian (4.14) can be removed as a consequence of the stationarity of the hypersurface.

In fact, since when $Q_f = P_f = 0$ we have $\tilde{\Omega}_\beta = 0$, we can develop $\tilde{\Omega}_\beta$ in a power series of Q_f and P_f, i.e.,

$$\tilde{\Omega}_\beta(Q'_s,P'_s,t) = a^f_\beta(Q'_j,P'_j,t)Q_f + b^f_\beta(Q'_j,P'_j,t)P_f + \text{higher orders}. \tag{4.16}$$

If we introduce, following Sudarshan and Mukunda,[2] the notation of "strong" equality ($\equiv$), we can rewrite Eq. (4.16) as

$$\tilde{\Omega}_\beta(Q'_s,P'_s,t) \equiv a^f_\beta(Q'_j,P'_j,t)Q_f + b^f_\beta(Q'_j,P'_j,t)P_f. \tag{4.17}$$

From Eq. (4.17) we can also locally get the inverse relations

$$\begin{cases} Q_f \equiv \tilde{c}^\beta_f(Q'_j,P'_j,t)\tilde{\Omega}_\beta \\ P_f \equiv \tilde{d}^\beta_f(Q'_j,P'_j,t)\tilde{\Omega}_\beta, \end{cases} \tag{4.18}$$

and in terms of the old variables

$$\begin{cases} Q_f(q_s,p_s,t) \equiv c^\beta_f(q_s,p_s,t)\Omega_\beta(q_s,p_s) \\ P_f(q_s,p_s,t) \equiv d^\beta_f(q_s,p_s,t)\Omega_\beta(q_s,p_s). \end{cases} \tag{4.19}$$

By taking the partial derivative with respect to t of Eqs. (4.19) we get

$$\frac{\partial Q_f}{\partial t} \approx 0. \quad \frac{\partial P_f}{\partial t} \approx 0. \tag{4.20}$$

On the other hand, if we take the total derivative with respect to t of Eq. (4.19) and use the stationarity of the Ω_β's we have

$$\frac{d}{dt}Q_f \approx 0, \quad \frac{d}{dt}P_f \approx 0, \tag{4.21}$$

and finally from Eq. (4.17),

$$\frac{\partial}{\partial t}\tilde{\Omega}_\beta \approx 0. \tag{4.22}$$

Thus the last term of the Hamiltonian can be dropped, because it is strongly equal to zero. Then Eqs. (4.12) and (4.13) become

$$\begin{cases} \dot{Q}'_j \approx \{Q'_j,K_c + l_\alpha\tilde{\Omega}_\alpha - \tilde{\Omega}_\beta\tilde{c}_{\beta\beta'}\{\tilde{\Omega}_{\beta'},K_c\}\}, \\ \dot{P}'_j \approx \{P'_j,K_c + l_\alpha\tilde{\Omega}_\alpha - \tilde{\Omega}_\beta\tilde{c}_{\beta\beta'}\{\tilde{\Omega}_{\beta'},K_c\}\}, \end{cases} \tag{4.23}$$

and

$$\begin{cases} \dot{Q}_f \approx \{Q_f,K_c + l_\alpha\tilde{\Omega}_\alpha - \tilde{\Omega}_\beta\tilde{c}_{\beta\beta'}\{\tilde{\Omega}_{\beta'},K_c\}\} \\ \dot{P}_f \approx \{P_f,K_c + l_\alpha\tilde{\Omega}_\alpha - \tilde{\Omega}_\beta\tilde{c}_{\beta\beta'}\{\tilde{\Omega}_{\beta'},K_c\}\}, \end{cases} \tag{4.24}$$

or, using Eqs. (4.17) and (4.18),

$$\begin{cases} \dot{Q}'_j \approx \{Q'_j,K_c + l_\alpha\tilde{\Omega}_\alpha\} = \{Q'_j,\tilde{K}\} \\ \dot{P}'_j \approx \{P'_j,K_c + l_\alpha\tilde{\Omega}_\alpha\} = \{P'_j,\tilde{K}\} \end{cases} \quad (j=1,\dots,n_2), \tag{4.25}$$

$$\begin{cases} \dot{Q}_f \approx 0 \\ \dot{P}_f \approx 0 \end{cases} \quad (f=n_2+1,\dots,n), \tag{4.26}$$

where the last equalities of Eq. (4.25) define $\tilde{K}$.

If we denote the set of variables which are independent with respect to second-class constraints by R

$$R \equiv \{Q'_j,P'_j; j=1,\dots,n_2\}, \tag{4.27}$$

the equations of motion, in this reduced phase space, can be rewritten as

$$\begin{cases} \dot{Q}'_j \approx \{Q'_j,\tilde{K}\}_R \approx \{Q'_j,\bar{K}\}_R, \\ \dot{P}'_j \approx \{P'_j,\tilde{K}\}_R \approx \{P'_j,\bar{K}\}_R, \end{cases} \tag{4.28}$$

where $\{\ ,\ \}_R$ are the Poisson brackets defined in the space R and

$$\bar{K} = \bar{K}_c + l_\alpha\bar{\Omega}_\alpha, \tag{4.29}$$

with $\bar{K}_c$ and $\bar{\Omega}_\alpha$ obtained by setting to zero the variables Q_f and P_f in K_c and $\tilde{\Omega}_\alpha$. The "weak" equalities of Eq. (4.28) are equalities on the surface determined by

$$\bar{\Omega}_\alpha(Q'_j,P'_j) = 0. \tag{4.30}$$

Let us observe that, since Ω_α are first-class, we have also

$$\{\tilde{\Omega}_\alpha,\tilde{\Omega}_{\alpha'}\} \approx 0, \tag{4.31}$$

$$\{\tilde{\Omega}_\alpha,\tilde{\Omega}_\beta\} \approx 0. \tag{4.32}$$

From Eqs. (4.18) and (4.32) we get

$$\frac{\partial\tilde{\Omega}_\alpha}{\partial P_f} = \{Q_f,\tilde{\Omega}_\alpha\} \approx 0, \quad \frac{\partial\tilde{\Omega}_\alpha}{\partial Q_f} = -\{P_f,\tilde{\Omega}_\alpha\} \approx 0. \tag{4.33}$$

Therefore, by defining

$$\{\tilde{\Omega}_\alpha,\tilde{\Omega}_{\alpha'}\}_R = \{\tilde{\Omega}_\alpha,\tilde{\Omega}_{\alpha'}\} - \left(\frac{\partial\tilde{\Omega}_\alpha}{\partial Q_f}\frac{\partial\tilde{\Omega}_{\alpha'}}{\partial P_f} - \frac{\partial\tilde{\Omega}_\alpha}{\partial P_f}\frac{\partial\tilde{\Omega}_{\alpha'}}{\partial Q_f} \right), \tag{4.34}$$

and using Eqs. (4.33) and (4.31), we have

$$\{\tilde{\Omega}_\alpha,\tilde{\Omega}_{\alpha'}\}_R \approx 0, \tag{4.35}$$

which also implies

$$\{\bar{\Omega}_\alpha,\bar{\Omega}_{\alpha'}\}_R \approx 0. \tag{4.36}$$

Finally let us prove that the hypersurface determined by Eq. (4.30) is stationary.

In fact,

$$\frac{d}{dt}\bar{\Omega}_\alpha \approx \{\bar{\Omega}_\alpha,\bar{K}\}_R \approx \{\bar{\Omega}_\alpha,\bar{K}_c\}_R \approx \{\tilde{\Omega}_\alpha,K_c\}_R \approx \{\tilde{\Omega}_\alpha,K_c\}, \tag{4.37}$$

where use was made of Eqs. (4.33) and (4.36). On the other hand, due to the canonical character of the transformation (4.11), Eqs. (2.7) imply

$$\{\tilde{\Omega}_\alpha,K_c\} \approx 0, \tag{4.38}$$

and thus

$$\frac{d}{dt}\bar{\Omega}_\alpha \approx \{\bar{\Omega}_\alpha,\bar{K}_c\}_R \approx 0. \quad (4.39)$$

Summing up, we have shown that it is possible, by making use of a canonical transformation, to write the equations of motion for a set of variables which are independent with respect to second-class constraints. As shown in Ref. 3, we have the following relation between Dirac brackets and Poisson brackets defined in the reduced space R:

$$\{\ ,\ \}^* = \{\ ,\ \}_R .$$

Therefore the variables which are canonical with respect to Dirac brackets are directly obtained by means of this canonical transformation.

5. EQUATIONS OF MOTION FOR A SET OF UNCONSTRAINED VARIABLES

A further step can be done by extending the transformation to include first-class constraints too.

In fact, Theorem 4 guarantees that it is always possible, *at least locally*, to replace the $\bar{\Omega}_\alpha$ by an equivalent set

$$P_e = 0, \quad (e = n_1+1,\dots,n_2), \quad (5.1)$$

$(n_1 = n - T + W/2)$, such that the equations

$$\{P_e,P_{e'}\} = 0 \quad (5.2)$$

are identically satisfied and not by virtue of Eqs. (5.1) themselves. The same theorem shows that the P_e can be obtained by solving Eq. (4.30) for $n_2 - n_1$ of the momenta P'_j in terms of the remaining momenta and of the coordinates Q'_j. We can, by renumbering the variables if necessary, assume that Eq. (4.30) can be solved for the last $n_2 - n_1$ P'_j in terms of the first $n_1 P'_j$ and all the Q'_j, i.e.,

$$P_e = P'_e - f_e(Q'_e,Q'_k,P'_k) \quad (k = 1,\dots,n_1) \quad (e = n_1+1,\dots,n_2). \quad (5.3)$$

We observe, from Eq. (5.3), the local character of this technique. Thus, in general, we will have to repeat the procedure we will develop in the following, for the different sheets of the hypersurface (4.30).

Let us notice that from Eqs. (4.9) and (5.3) we have

$$\{Q'_e,P_{e'}\}_R = \delta_{ee'} , \quad (5.4)$$

or, following the terminology of the function groups, Q'_e (that from now on we will call Q_e) and P_e form a noncommutative function group of dimension $2(n_2 - n_1)$. By applying again Theorem 3 we can construct a canonical transformation

$$Q'_j,P'_j,t \to Q_k,P_k,Q_e,P_e \quad (k = 1,\dots,n_1),(e = n_1+1,\dots,n_2), \quad (5.5)$$

with

$$\{Q_k,P_{k'}\} = \delta_{kk'}, \quad \{Q_e,P_{e'}\} = \delta_{ee'}, \quad (5.6)$$

and the other Poisson brackets vanishing.

If we denote the new canonical Hamiltonian by $\hat{K}_c$, and if we write the constraints (4.30) in terms of the new variables as

$$\hat{\Omega}(Q_k,P_k,Q_e,P_e) = \bar{\Omega}_\alpha(Q'_j(Q_k,P_k,Q_e,P_e),P'_j(Q_k,P_k,Q_e,P_e)) = 0, \quad (5.7)$$

the new equations of motion can be obtained by applying Theorem 2 in the reduced phase space R:

$$\begin{cases} \dot{Q}_k \approx \{Q_k,\hat{K}_c + l_\alpha\hat{\Omega}_\alpha\}_R , \\ \dot{P}_k \approx \{P_k,\hat{K}_c + l_\alpha\hat{\Omega}_\alpha\}_R , \end{cases} \quad (5.8)$$

and

$$\begin{cases} \dot{Q}_e \approx \{Q_e,\hat{K}_c + l_\alpha\hat{\Omega}_\alpha\}_R , \\ \dot{P}_e \approx \{P_e,\hat{K}_c + l_\alpha\hat{\Omega}_\alpha\}_R . \end{cases} \quad (5.9)$$

On the other hand, due to the fact that when $P_e = 0$, $\hat{\Omega}_\alpha = 0$, we can write, as for the second-class constraints [Eqs. (4.17)],

$$\hat{\Omega}_\alpha(Q_k,P_k,Q_e,P_e) \equiv g^e_\alpha(Q_k,P_k,Q_e)P_e . \quad (5.10)$$

Thus substituting Eq. (5.10) in Eq. (5.8) and (5.9), we get

$$\begin{cases} \dot{Q}_k \approx \{Q_k,\hat{K}_c + \lambda_{e'}P_{e'}\}_R \approx \{Q_k,\hat{K}_c\}_R , \\ \dot{P}_k \approx \{P_k,\hat{K}_c + \lambda_{e'}P_{e'}\}_R \approx \{P_k,\hat{K}_c\}_R , \end{cases} \quad (5.11)$$

$$\begin{cases} \dot{Q}_e \approx \{Q_e,\hat{K}_c + \lambda_{e'}P_{e'}\}_R \approx \{Q_e,\hat{K}_c\}_R + \lambda_e , \\ \dot{P}_e \approx \{P_e,\hat{K}_c + \lambda_{e'}P_{e'}\}_R \approx \{P_e,\hat{K}_c\}_R , \end{cases} \quad (5.12)$$

where $\lambda_e = l_\alpha g^e_\alpha$ are arbitrary functions of t.

Let us finally show that this sheet of hypersurface is stationary. By differentiating Eq. (4.30) with respect to any variable u (Q'_j or P'_j) and using Eq. (5.3) we get

$$\frac{\partial\bar{\Omega}_\alpha}{\partial u} = -\frac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\frac{\partial f_e}{\partial u} = \frac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\frac{\partial P_e}{\partial u}. \quad (5.13)$$

Therefore

$$\{\bar{\Omega}_\alpha,\bar{K}_c\}_R = \frac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\{P_e,\bar{K}_c\}_R , \quad (5.14)$$

and using Eq. (4.39) and the fact that we locally have

$$\det\left|\frac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\right| \neq 0, \quad (5.15)$$

we get

$$\{P_e,\bar{K}_c\}_R \approx 0, \quad (5.16)$$

and after the canonical transformation (5.5)

$$\frac{\partial\hat{K}_c}{\partial Q_e} = \{\hat{K}_c,P_e\}_R \approx 0, \quad (5.17)$$

which ensures the hypersurface be stationary. In addition Eq. (5.17) state that the variables Q_e are ignorable variables on the hypersurface.[3]

Thus we have the following equations of motion:

$$\dot{Q}_k \approx \{Q_k,\hat{K}_c\}_R \quad \dot{P}_k \approx \{P_k,\hat{K}_c\}_R ,$$
$$\dot{Q}_c \approx \{Q_e,\hat{K}_c\} + \lambda_e \quad \dot{P}_e \approx 0, \quad (5.18)$$

where λ_e are arbitrary functions.

We can finally consider a reduced space of unconstrained variables [Q_k, P_k and Q_e], which are the intrinisic coordinates of the hypersurface of the motion:

$$Q_f = 0, \quad P_f = 0, \quad P_c = 0. \quad (5.19)$$

Their evolution equations are

$$\dot{Q}_k = \frac{\partial \hat{K}_c}{\partial P_k}\bigg|_{P_e=0}, \quad \dot{P}_k = -\frac{\partial \hat{K}_c}{\partial Q_k}\bigg|_{P_e=0}, \tag{5.20}$$

$$\dot{Q}_e = \frac{\partial \hat{K}_c}{\partial P_e}\bigg|_{P_e=0} + \lambda_e . \tag{5.21}$$

If we put

$$\mathscr{K}_c(Q_k,P_k,t) = \hat{K}_c(Q_k,P_k,Q_e,P_e,t)\,|_{P_e=0}\,, \tag{5.22}$$

Eqs. (5.20) can be rewritten as

$$\dot{Q}_k = \frac{\partial \mathscr{K}_c}{\partial P_k}(Q_k,P_k,t),$$

$$\dot{P}_k = -\frac{\partial \mathscr{K}_c}{\partial Q_k}(Q_k,P_k,t) \quad (k = 1,...,n_1), \tag{5.23}$$

whereas Eqs. (5.21) are left unchanged since the two operations of setting $P_e = 0$ and differentiating with respect to P_e do not commute.

As shown by Eqs. (5.20) and (5.21) this method allows to isolate the gauge independent variables Q_k, P_k (physical variables) from the gauge dependent Q_e, whose evolution is determined only when the arbitrary functions are given.

We point out that when $H_c = 0$, that is when, if a lagrangian formulation exists, the action is parameter-invariant,[11] we have $\hat{K}_c = 0$ and Eqs. (5.20) and (5.21) become

$$\dot{Q}_k = 0, \quad \dot{P}_k = 0, \tag{5.24}$$

$$\dot{Q}_e = \lambda_e . \tag{5.25}$$

Thus, for what concerns the physical variables, the procedure is equivalent to the Hamilton–Jacobi method.

Let us finally observe that we can choose anyone of the coordinates Q_e,[12] for example $Q_{\bar{e}}$ ($n_1 < \bar{e} \leqslant n_2$), as evolution parameter and rewrite Eqs. (5.24) and (5.25) as

$$\frac{dQ_k}{dQ_{\bar{e}}} = 0, \quad \frac{dP_k}{dQ_{\bar{e}}} = 0, \tag{5.26}$$

$$\frac{dQ_r}{dQ_{\bar{e}}} = \lambda_r/\lambda_{\bar{e}} \quad (r = n_1 + 1,...\bar{e} - 1, \bar{e} + 1,...n_2) \tag{5.27}$$

$$\dot{Q}_{\bar{e}} = \lambda_{\bar{e}} . \tag{5.28}$$

Thus in order to get the relation between Q_r and $Q_{\bar{e}}$ we must give the ratio of the arbitrary functions $\lambda_r/\lambda_{\bar{e}}$ as a function of $Q_{\bar{e}}$ and if we are interested in the relation between $Q_{\bar{e}}$ and the unphysical parameter t we must give $\lambda_{\bar{e}}$ as a function of t.

Therefore, with this procedure we get a reduced class of gauge (we can only choose one of the coordinates Q_e as evolution parameter). This a consequence of the definition of P_e [Eqs. (5.3)]. On the other hand, different classes of gauges can be obtained by solving Eq. (4.30) to a different set of momenta.

CONCLUSIONS

Making use of the Poincaré–Cartan integral for constrained systems, we have shown that the invariance of this integral enables us to write the equations of motion for a dynamical system as Hamilton equations. We want to observe that with this procedure, all the first-class constraints appear in the Hamiltonian, because we cannot introduce any distinction between them. Recent papers, by Cawley[13] and Frenkel,[14] have shown, with some examples, that not all the first-class secondary constraints generate gauge transformations and therefore not all the first-class constraints appear in the Hamiltonian.[15] Thus we are investigating an algebraic procedure in order to take into account this result.

Furthermore, we have introduced a definition of canonical transformation, which is the trivial generalization of the usual one, and shown that the Poincaré–Cartan integral is invariant under this transformation. Then, following Shanmugadhasan,[3] we have performed a canonical transformation such that a subset of the new variables is equivalent to the second-class constraints. The reduced set of variables, independent with respect to second-class constraints, is nothing but the set of variables which are canonical with respect to Dirac brackets.

A further step is done by performing a new canonical transformation in the reduced phase space which isolates the variables corresponding to first-class constraints. This transformation is very useful because it isolates also the gauge independent variables from the gauge dependent ones. The evolution of these gauge dependent variables, contrary to the result of Shanmugadhasan, consistently depends on arbitrary functions.

When $H_c = 0$ this technique becomes equivalent to the Hamilton–Jacobi method. Explicit examples (the free relativistic point and a model of two interacting relativistic particles) have been already studied[16]; presently we are investigating the possibility of extending this technique to continuous systems, studying the relativistic string model (see Nambu and Scherk in Ref. 1).

ACKNOWLEDGMENTS

We wish to thank Dr. L. Lusanna of the "Istituto Nazionale di Fisica Nucleare" of Florence for many stimulating and interesting discussions.

[1]Y. Nambu, "Lectures for the Copenhagen Summer Symposium 1970" (unpublished); J. Scherk, Rev. Mod. Phys. **47**, 123 (1975); R. Casalbuoni and G. Longhi, Nuovo Cimento A **25**, 482 (1975); M. Kalb and P. Van Alstine, "Invariant Singular Actions for the Relativistic Two Body Problem: a Hamiltonian Formulation," Yale report C00–3075–146 (June 1976); T. Takabayasi and S. Kojima, Prog. Theor. Phys. **57**, 2127 (1977); D. Dominici, J. Gomis and G. Longhi, Nuovo Cimento B **48**, 152 (1978); Nuovo Cimento A **48**, 257 (1978).

[2]P. A. M. Dirac, "Lectures on Quantum Mechanics," Belfer Graduate School of Science, Yeshiva University (New York, 1964); for a comprehensive exposition of the Dirac theory see also E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: a Modern Perspective* (Wiley, New York, 1974).

[3]S. Shanmugadhasan, J. Math. Phys. **14**, 677 (1973).

[4]J. A. Shouten and W. v. der Kulk, *Pfaff's Problem and Its Generalizations* (Clarendon, Oxford, 1949).

[5]C. R. Chester, *Techniques in Partial Differential Equations*, Chap. XV (McGraw-Hill Kogakusha, Tokyo, 1971).

[6]L. P. Eisenhart, *Continuous Groups of Transformations*, (Dover, New York, 1961).

[7]F. Benavent and J. Gomis, "Poincaré–Cartan Integral Invariant for Constrained Systems," Ann. Phys. **118**, 476 (1979).

[8]F. Gantmacher, *Lectures in Analytical Mechanics* (Mir. Moscow, 1970).

[9]The set G of the functions of r independent functions $F_1(q_s, p_s)\cdots F_r(q_s, p_s)$, $(s = 1,...,n)$, such that $\{F_a, F_b\} = \phi(F_c)$, $(a,b,c, = 1,...,r)$ is said to be a *function group* of rank r. The function group is said commutative if $\{F_a, F_b\} = 0$ for all the values of a and b.

[10]L. Castellani, D. Domínici, and G. Longhi, Nuovo Cimento A **48**, 91 (1978).

[11]H. Rund, *The Hamilton–Jacobi Theory in the Calculus of Variations* (Van Nostrand, London, 1966).

[12]Also in Ref. 7 we have introduced as evolution parameter one of the coordinates q_s, but in that case the coordinates are not independent with respect to the constraints.

[13]R. Cawley, Phys. Rev. Lett. **42**, 413 (1979).

[14]A. Frenkel, "Comment on Cawley's Counter Example to a Conjecture of Dirac," preprint KFKI-1979-45, Central Research Institute for Physics, Budapest, Hungary.

[15]See also M. J. Gotay, "On the Validity of Dirac's Conjecture Regarding First-Class Secondary Constraints," University of Calgary preprint (November 1979) and references quoted therein.

[16]D. Dominici, J. Gomis and G. Longhi, in preparation.

Spectral broadening of waves propagating in a random medium

Alan R. Wenzel[a)]

Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, Virginia 23665

(Received 5 September 1978; accepted for publication 28 March 1980)

Formulas for the cross-correlation and spectral density functions of the scalar wave field radiated by a random point source in a weakly inhomogeneous three-dimensional time-dependent random medium are derived. The medium is assumed to be statistically homogeneous and isotropic and to be statistically independent of the source. The analysis is based on a modification of the smoothing method. An approximate expression for the power spectrum of the wave as a function of the source-field point distance (or propagation distance) is obtained for the case in which the characteristic frequency of the source is much greater than that of the medium. This expression shows that the wave spectrum approaches a limiting form, which is referred to here as the fully developed spectrum, with increasing propagation distance. It is also found that the total signal power is conserved as the spectrum evolves. Results obtained for the case of a narrow-band source indicate that the spectral bandwidth increases initially as the square root of the propagation distance, but that at larger distances it approaches a limiting value. Numerical results obtained for the narrow-band case show a progressive broadening of the wave spectrum with increasing propagation distance and/or with increasing strength of the randomness of the medium, in agreement with observations.

INTRODUCTION

Broadening of the frequency spectrum of an initially narrow-band wave field is a phenomenon which is characteristic of wave propagation in a time-dependent medium, and is a result of amplitude and frequency modulation of the spectral components of the wave by the time variations in the properties of the medium. Of particular practical interest is the effect of *random* fluctuations of the medium, and indeed spectral broadening due to propagation through turbulence has been observed in the case of both acoustic and electromagnetic waves.[1,2] The present investigation was undertaken with the purpose of studying this effect, i.e., spectral broadening arising from the presence of random, time-dependent fluctuations of the medium, from a rather general point of view.

Previous theoretical investigations of spectral broadening of waves in random media include those of Howe,[3] Fante,[4,5] and Woo *et al.*[6] Howe derived a kinetic equation and used it to study the effect of the random velocity field on the frequency spectrum of an acoustic wave propagating in a turbulent fluid. Fante used transport theory to study frequency spectra of beamed waves propagating in a turbulent atmosphere. The analysis of Woo *et al.* (see also Ref. 7, p. 422) is based on the parabolic approximation. Howe treated the case of an isotropic time-dependent turbulence field, whereas both Fante and Woo *et al.* assumed that the time variations of the medium were the result of a steady mean wind convecting a "frozen" turbulence field in a direction perpendicular to the direction of propagation.

The authors mentioned above based their analyses on different mathematical models and/or calculated different statistical properties of the wave field than those considered in the present investigation (e.g., Woo *et al.* considered the spectrum of the complex wave amplitude, whereas the present investigation deals with the entire wavefunction, which is a real quantity), and hence their results do not agree in all respects with those obtained here. The results of both Howe and Fante indicate that, over a suitably restricted propagation path and for high-frequency waves, the characteristic width of the wave spectrum increases as some power of the propagation distance. The results of Woo *et al.* are given in a more complicated form, but seem to show a similar effect. These results agree generally with those obtained here for small propagation distances. At large distances, however, the present results indicate that the spectral width approaches a limiting value, which is not predicted by any of the theories mentioned above.

The problem of spectral broadening has also been treated in a recent paper by Kuznetsova and Chernov.[8] Their analysis, like that of Woo *et al.*, is based on the parabolic approximation with a frozen turbulence model. Their results also indicate an increase of the spectral width as some power of the propagation distance for small propagation distances. However, since their expression for the wave spectrum is given in the form of a power series in the propagation distance, the behavior of the spectrum for large propagation distances, as predicted by their theory, is not clear.

Spectral broadening in a random medium has also been discussed from a theoretical viewpoint by Adomian[9]; however, that author did not obtain an explicit analytical expression for the wave spectrum. Related work, concerned mainly with spectra of scattered waves (in contrast to the present investigation, which deals with the spectrum of the total wave field) and with spectra of amplitude and phase fluctuations of waves propagating in random media, can be found in Refs. 10–16.

[a)]Present address: Naval Ocean Research and Development Activity, Code 340, NSTL Station, Mississippi 39529.

I. GENERAL ANALYSIS

The starting point of the analysis is the scalar wave equation

$$(c^{-2}\partial_t^2 - \nabla^2)u = f, \tag{1}$$

where u is the wavefunction, f is the source term, and c is the local propagation speed of small disturbances of the medium. All quantities are assumed to be real functions of t and $\mathbf{x}$, where t is time and $\mathbf{x}$ [$= (x_1,x_2,x_3)$] is a three-dimensional spatial coordinate.

The propagation speed c is assumed to be random; i.e., c is assumed to depend on a parameter a which is an element of a sample space A. The space A, together with a σ-algebra of subsets and a probability measure, forms a probability space. The source term f is also random; however f is assumed to be statistically independent of c. Thus, f may be regarded as being dependent on a parameter b ranging over a different sample space B which, together with its own σ-algebra of subsets and probability measure, also forms a probability space.

It is clear that the solution u of Eq. (1), as well as functions of it, will depend on both a and b. (The dependence on the parameters a and b of the various quantities appearing in the analysis will not, in general, be explicitly indicated.) It will be necessary, therefore, in what follows to distinguish between ensemble averages over the space A, which will be denoted by $\langle\ \rangle_A$, and averages over B, denoted by $\langle\ \rangle_B$. An average over both A and B (i.e., an ensemble average over the product sample space $A \times B$) will be denoted simply by $\langle\ \rangle$. We note that generally $\langle\ \rangle = \langle\langle\ \rangle_A\rangle_B = \langle\langle\ \rangle_B\rangle_A$.

In most cases involving wave propagation in real media such as the atmosphere or ocean the fluctuations in the medium properties can be regarded as small. Thus it is realistic, as well as mathematically convenient, to write c in the form

$$c(t,\mathbf{x}) = c_0[1 + \epsilon\mu(t,\mathbf{x})]. \tag{2}$$

Here ϵ is a small parameter which is a measure of the magnitude of the fluctuations of the medium, and μ is a random function with zero mean and unit variance; i.e., $\langle\mu\rangle_A = 0$, $\langle\mu^2\rangle_A = 1$. The quantity c_0, the average of c, is assumed to be a constant.

Writing c as in Eq. (2) allows the problem to be solved by a perturbation technique. To begin, we substitute the expression for c given by Eq. (2) into Eq. (1) and expand in powers of ϵ. This yields

$$(L_0 + \epsilon L_1 + \epsilon^2 L_2 + \cdots)u = f, \tag{3}$$

where the operators L_0, L_1, and L_2 are given by

$$L_0 = c_0^{-2}\partial_t^2 - \nabla^2, \tag{4}$$

$$L_1 = -2c_0^{-2}\mu\partial_t^2, \tag{5}$$

$$L_2 = 3c_0^{-2}\mu^2\partial_t^2. \tag{6}$$

From Eq. (3), approximate equations, valid when ϵ is small, can be obtained for $\bar{u}$ and $\tilde{u}$, where $\bar{u} \equiv \langle u\rangle_A$ and $\tilde{u} \equiv u - \langle u\rangle_A$. The procedure is entirely analogous to that described by Keller[17] (see also Ref. 18). It is only necessary to keep in mind that since f is independent of a it is unaffected by averaging over A. When $\langle L_1\rangle_A = 0$, which is usually the case in practice, these equations reduce to

$$M\bar{u} = f, \tag{7}$$

$$\tilde{u} = -\epsilon L_0^{-1}L_1\bar{u}, \tag{8}$$

where the operator M is defined by

$$M = L_0 + \epsilon^2(\langle L_2\rangle_A - \langle L_1 L_0^{-1} L_1\rangle_A). \tag{9}$$

Terms of order ϵ^3 have been dropped from Eq. (7); terms of order ϵ^2 have been dropped from Eq. (8).

In the special case in which f is determinate (i.e., nonrandom) the quantities $\bar{u}$ (which is then also determinate) and $\tilde{u}$ correspond respectively to the mean and fluctuating fields. In that context the type of approach leading to Eqs. (7) and (8), which involves obtaining separate equations for the mean and fluctuating fields, is referred to as the smoothing method by Frisch.[18]

We shall be concerned in the remainder of this paper only with random processes which are stationary in time. (By stationary we mean stationary in the wide sense, i.e., that correlation functions of the form given by Eq. (19) are independent of t.) In order to ensure that $u(t,\mathbf{x})$ is stationary in time we shall assume that both $\mu(t,\mathbf{x})$ and $f(t,\mathbf{x})$ are stationary in time. That these assumptions are sufficient for our purposes will become clear as the analysis proceeds. In addition, we shall assume for convenience that $\mu(t,\mathbf{x})$ is statistically homogeneous and isotropic in space, and that $\langle f\rangle_B = 0$.

We introduce next the Green's functions $G_0(t,\mathbf{x})$ and $G(t,\mathbf{x})$, which are solutions of the equations

$$L_0 G_0(t,\mathbf{x}) = \delta(t)\delta(\mathbf{x}), \tag{10}$$

$$MG(t,\mathbf{x}) = \delta(t)\delta(\mathbf{x}), \tag{11}$$

and which satisfy the initial conditions $G_0 = G = 0$ for $t<0$. (No boundary conditions need be imposed on G_0 or G since we are considering only free-space propagation.) Then L_0^{-1} can be written in the form

$$L_0^{-1}w(t,\mathbf{x}) = \int\int G_0(t-t',\mathbf{x}-\mathbf{x}')w(t',\mathbf{x}')\,dt'\,d\mathbf{x}', \tag{12}$$

where $w(t,\mathbf{x})$ is any function for which the integral exists. (Here, and henceforth, an integral sign without limits denotes an integral from $-\infty$ to $+\infty$.) Similarly, the solution of Eq. (7) can be expressed as

$$\bar{u}(t,\mathbf{x}) = \int\int G(t-t',\mathbf{x}-\mathbf{x}')f(t',\mathbf{x}')\,dt'\,d\mathbf{x}'. \tag{13}$$

By making a change of integration variable we can write Eqs. (12) and (13) in the form

$$L_0^{-1}w(t,\mathbf{x}) = \int\int G_0(t',\mathbf{x}')w(t-t',\mathbf{x}-\mathbf{x}')\,dt'\,d\mathbf{x}', \tag{14}$$

$$\bar{u}(t,\mathbf{x}) = \int\int G(t',\mathbf{x}')f(t-t',\mathbf{x}-\mathbf{x}')\,dt'\,d\mathbf{x}'. \tag{15}$$

It should be pointed out that, as a consequence of the assumption that $\mu(t,\mathbf{x})$ is stationary in t and $\mathbf{x}$, the operator M commutes with both time and space translations. This allows the Green's function G in Eq. (13) to be written as a function of the differences $t-t'$ and $\mathbf{x}-\mathbf{x}'$, instead of as a function of $t,t',\mathbf{x}$, and $\mathbf{x}'$ separately. Since the operator L_0 has constant coefficients, it also commutes with both time and space translations, and hence the Green's function G_0 in Eq. (12) can also be written as a function of $t - t'$ and $\mathbf{x} - \mathbf{x}'$. That G_0 and G can be written as functions of $t - t'$ in Eqs. (12) and (13) is necessary for the stationarity of u. Note also that both G_0 and G are determinate functions.

Operating with L_1 on $\bar{u}$, as given by Eq. (13), yields

$$L_1\bar{u}(t,\mathbf{x}) = -2c_0^{-2}\int\int \mu(t,\mathbf{x}) \times G_{tt}(t-t',\mathbf{x}-\mathbf{x}')f(t',\mathbf{x}')\,dt'\,d\mathbf{x}'. \quad (16)$$

(The subscripts on G denote derivatives.) By making a change of integration variable we can write Eq. (16) in the form

$$L_1\bar{u}(t,\mathbf{x}) = -2c_0^{-2}\int\int \mu(t,\mathbf{x})G_{tt}(t',\mathbf{x}') \times f(t-t',\mathbf{x}-\mathbf{x}')\,dt'\,d\mathbf{x}'. \quad (17)$$

Operating on Eq. (17) with L_0^{-1}, as given by Eq. (14), and substituting the result into Eq. (8) yields

$$\tilde{u}(t,\mathbf{x}) = 2\epsilon c_0^{-2}\int\cdots\int G_0(t',\mathbf{x}')G_{tt}(t'',\mathbf{x}'') \times \mu(t-t',\mathbf{x}-\mathbf{x}')f(t-t'-t'',\mathbf{x}-\mathbf{x}'-\mathbf{x}'') \times dt'\,d\mathbf{x}'\,dt''\,d\mathbf{x}''. \quad (18)$$

The cross-correlation function $R(\tau,\mathbf{x},\mathbf{y})$ is defined by

$$R(\tau,\mathbf{x},\mathbf{y}) = \langle u(t,\mathbf{x})u(t-\tau,\mathbf{y})\rangle. \quad (19)$$

Upon writing u as the sum $u = \bar{u} + \tilde{u}$ in Eq. (19) we obtain

$$R(\tau,\mathbf{x},\mathbf{y}) = \langle \bar{u}(t,\mathbf{x})\bar{u}(t-\tau,\mathbf{y})\rangle + \langle \bar{u}(t,\mathbf{x})\tilde{u}(t-\tau,\mathbf{y})\rangle + \langle \tilde{u}(t,\mathbf{x})\bar{u}(t-\tau,\mathbf{y})\rangle + \langle \tilde{u}(t,\mathbf{x})\tilde{u}(t-\tau,\mathbf{y})\rangle. \quad (20)$$

The two cross terms on the right-hand side of Eq. (20), i.e., the terms involving products of $\bar{u}$ and $\tilde{u}$, vanish. This follows from the fact that $\bar{u}$ is independent of a and that $\langle \tilde{u}\rangle_A = 0$. Thus, for the first cross term, we can write

$$\langle \bar{u}(t,\mathbf{x})\tilde{u}(t-\tau,\mathbf{y})\rangle = \langle \langle \bar{u}(t,\mathbf{x})\tilde{u}(t-\tau,\mathbf{y})\rangle_A\rangle_B = \langle \bar{u}(t,\mathbf{x})\langle \tilde{u}(t-\tau,\mathbf{y})\rangle_A\rangle_B = 0$$

(since $\langle \tilde{u}(t-\tau,\mathbf{y})\rangle_A = 0$), and similarly for the second cross term. Expressions for the remaining two terms on the right-hand side of Eq. (20) can be obtained with the aid of Eqs. (15) and (18), after which Eq. (19) can be written

$$R(\tau,\mathbf{x},\mathbf{y}) = \bar{R}(\tau,\mathbf{x},\mathbf{y}) + \tilde{R}(\tau,\mathbf{x},\mathbf{y}), \quad (21)$$

where

$$\bar{R}(\tau,\mathbf{x},\mathbf{y}) = \langle \bar{u}(t,\mathbf{x})\bar{u}(t-\tau,\mathbf{y})\rangle = \int\cdots\int G(t',\mathbf{x}')G(s',\mathbf{y}') \times R_0(\tau-t'+s',\mathbf{x}-\mathbf{x}',\mathbf{y}-\mathbf{y}')dt'\,d\mathbf{x}'\,ds'd\mathbf{y}', \quad (22)$$

$$\tilde{R}(\tau,\mathbf{x},\mathbf{y}) = \langle \tilde{u}(t,\mathbf{x})\tilde{u}(t-\tau,\mathbf{y})\rangle = 4\epsilon^2 c_0^{-4} \times \int\cdots\int G_0(t'\mathbf{x}')G_0(s',\mathbf{y}')G_{tt}(t'',\mathbf{x}'')G_{tt}(s'',\mathbf{y}'') \times \Gamma(\tau-t'+s',\mathbf{y}-\mathbf{x}+\mathbf{x}'-\mathbf{y}') \times R_0(\tau-t'+s'-t''+s'',\mathbf{x}-\mathbf{x}'-\mathbf{x}'',\mathbf{y}-\mathbf{y}'-\mathbf{y}'') \times dt'\,d\mathbf{x}'\,ds'\,d\mathbf{y}'\,dt''\,d\mathbf{x}''\,ds''\,d\mathbf{y}''. \quad (23)$$

The correlation functions Γ and R_0 are defined by

$$\Gamma(\tau,\boldsymbol{\xi}) = \langle \mu(t,\mathbf{x})\mu(t-\tau,\mathbf{x}+\boldsymbol{\xi})\rangle_A, \quad (24)$$

$$R_0(\tau,\mathbf{x},\mathbf{y}) = \langle f(t,\mathbf{x})f(t-\tau,\mathbf{y})\rangle_B. \quad (25)$$

In deriving Eqs. (22) and (23) use has been made of the fact that μ is independent of b and that f is independent of a.

Note that $u(t,\mathbf{x})$, as calculated here, is indeed stationary in time, as can be seen by referring to Eq. (19) and Eqs. (21)–(23).

The spectral density function $S(\omega,\mathbf{x},\mathbf{y})$ is defined by

$$S(\omega,\mathbf{x},\mathbf{y}) = \int R(\tau,\mathbf{x},\mathbf{y})e^{i\omega\tau}\,d\tau. \quad (26)$$

With the aid of Eq. (21) we can write

$$S(\omega,\mathbf{x},\mathbf{y}) = \bar{S}(\omega,\mathbf{x},\mathbf{y}) + \tilde{S}(\omega,\mathbf{x},\mathbf{y}), \quad (27)$$

where

$$\bar{S}(\omega,\mathbf{x},\mathbf{y}) = \int \bar{R}(\tau,\mathbf{x},\mathbf{y})e^{i\omega\tau}\,d\tau, \quad (28)$$

$$\tilde{S}(\omega,\mathbf{x},\mathbf{y}) = \int \tilde{R}(\tau,\mathbf{x},\mathbf{y})e^{i\omega\tau}\,d\tau. \quad (29)$$

To calculate $\bar{S}$ we insert into Eq. (28) the expression for $\bar{R}$ given by Eq. (22) and carry out the integration over τ, t', and s'. The result is

$$\bar{S}(\omega,\mathbf{x},\mathbf{y}) = \int\int H(\omega,\mathbf{x}')H^*(\omega,\mathbf{y}') \times S_0(\omega,\mathbf{x}-\mathbf{x}',\mathbf{y}-\mathbf{y}')\,d\mathbf{x}'\,d\mathbf{y}', \quad (30)$$

where we have defined

$$H(\omega,\mathbf{x}) = \int G(t,\mathbf{x})e^{i\omega t}\,dt, \quad (31)$$

$$S_0(\omega,\mathbf{x},\mathbf{y}) = \int R_0(\tau,\mathbf{x},\mathbf{y})e^{i\omega\tau}\,d\tau, \quad (32)$$

and the symbol $(\;)^*$ denotes a complex conjugate. Similarly, an expression for $\tilde{S}$ is obtained by substituting the formula for $\tilde{R}$ given by Eq. (23) into Eq. (29) and carrying out the integration over τ, t', s', t'', and s''. This procedure yields

$$\tilde{S}(\omega,\mathbf{x},\mathbf{y}) = (2\epsilon^2/\pi c_0^4)\int\cdots\int H_0(\omega,\mathbf{x}')H_0^*(\omega,\mathbf{y}') \times [Z(\omega,\mathbf{y}-\mathbf{x}+\mathbf{x}'-\mathbf{y}') * \omega^4 H(\omega,\mathbf{x}'')H^*(\omega,\mathbf{y}'') \times S_0(\omega,\mathbf{x}-\mathbf{x}'-\mathbf{x}'',\mathbf{y}-\mathbf{y}'-\mathbf{y}'')] \times d\mathbf{x}'\,d\mathbf{y}'\,d\mathbf{x}''\,d\mathbf{y}'', \quad (33)$$

where H_0 and Z are defined by

$$H_0(\omega,\mathbf{x}) = \int G_0(t,\mathbf{x})e^{i\omega t}\,dt, \quad (34)$$

$$Z(\omega,\boldsymbol{\xi}) = \int \Gamma(\tau,\boldsymbol{\xi})e^{i\omega\tau}\,d\tau. \quad (35)$$

The notation $(\ )*(\ )$ in Eq. (33) denotes a convolution with respect to ω; i.e.,

$$f*g(\omega) = \int f(\omega - \omega')g(\omega')\, d\omega'.$$

(Whenever the convolution symbol appears inside brackets, as in Eq. (33), it is to be understood that only the terms inside the brackets are involved in the convolution.) In deriving Eq. (33) we have made use of some known results relating the Fourier transform of a product of two functions to the convolution of the transformed functions.

The formulas for R and S given above are accurate to order ϵ^2; i.e., the error in them is of order ϵ^3. This is a consequence of the dropping of terms of order ϵ^3 in Eq. (7) and ϵ^2 in Eq. (8) (note that $\tilde{u}$ is of order ϵ), and the vanishing of the cross terms in Eq. (20).

For practical purposes it is usually convenient to assume that all processes under consideration are ergodic, as well as stationary, in time, in which case the average denoted by $\langle\ \rangle$ can be regarded as a time average.

The analysis given above can be generalized; i.e., instead of starting with Eq. (1) we can start with Eq. (3) and assume that the operator L_0 is determinate with a known inverse and that the operators L_1, L_2, etc., are random with known statistics. The operators L_0, L_1, L_2, etc., need not be otherwise specified. Formulas for the correlation and spectral density functions, analogous to Eqs. (21)–(23), (27), (30), and (33), can then be derived for various cases, depending on the additional assumptions made regarding the operators L_0, L_1, L_2, etc. A general analysis of this type has been carried out and is available in report form.[19]

In order to proceed further it is necessary to calculate the Green's functions G_0 and G and the transforms H_0 and H. The function G_0, which corresponds to a spherical pulsed wave propagating in a uniform medium, is obtained by solving Eq. (10), with L_0 given by Eq. (4), subject to the initial condition $G_0 = 0$ for $t < 0$. This yields the familiar waveform given by

$$G_0(t,\mathbf{x}) = (4\pi x)^{-1}\delta(t - c_0^{-1}x). \tag{36}$$

By inserting the expression for G_0 given by Eq. (36) into Eq. (14), carrying out the integration over t', and changing the spatial integration variable, we can express the operator L_0^{-1} in the form

$$L_0^{-1}w(t,\mathbf{x}) = (4\pi)^{-1}\int \xi^{-1}w(t - c_0^{-1}\xi,\mathbf{x} + \boldsymbol{\xi})\, d\boldsymbol{\xi}. \tag{37}$$

The function $H_0(\omega,\mathbf{x})$ is easily obtained by transforming Eq. (36) according to Eq. (34). The result is

$$H_0(\omega,\mathbf{x}) = (4\pi x)^{-1}e^{ikx}, \tag{38}$$

where $k = \omega/c_0$.

The function $G(t,\mathbf{x})$ is determined by Eq. (11), where the operator M is given by Eq. (9). With the aid of these equations, along with Eqs. (4), (5), (6), and (37), we can write the equation for G in the form

$$\begin{aligned}(c_0^{-2}\partial_t^2 - \nabla^2)G(t,\mathbf{x}) &+ \epsilon^2\{3c_0^{-2}G_{tt}(t,\mathbf{x}) - (\pi c_0^4)^{-1}\\ &\times\int \xi^{-1}[\Gamma(c_0^{-1}\xi,\boldsymbol{\xi})G_{tttt}(t - c_0^{-1}\xi,\mathbf{x} + \boldsymbol{\xi})\\ &- 2\Gamma_\tau(c_0^{-1}\xi,\boldsymbol{\xi})G_{ttt}(t - c_0^{-1}\xi,\mathbf{x} + \boldsymbol{\xi})\\ &+ \Gamma_{\tau\tau}(c_0^{-1}\xi,\boldsymbol{\xi})G_{tt}(t - c_0^{-1}\xi,\mathbf{x} + \boldsymbol{\xi})]\, d\boldsymbol{\xi}\}\\ &= \delta(t)\delta(\mathbf{x}),\end{aligned} \tag{39}$$

where the letter subscripts denote derivatives. The initial condition for G is that $G = 0$ for $t < 0$.

The procedure by which Eq. (39) is solved for $G(t,\mathbf{x})$ is described in Appendix A. Since we wish only to calculate the function S, we need only the transform $H(\omega,\mathbf{x})$ of $G(t,\mathbf{x})$, as defined by Eq. (31). For the case in which the medium is isotropic [i.e., when $\Gamma(\tau,\boldsymbol{\xi}) = \Gamma(\tau,\xi)$] this quantity is given by

$$H(\omega,\mathbf{x}) = C(k)(4\pi x)^{-1}e^{i\kappa x}, \tag{40}$$

where

$$\kappa = k\left[1 + \tfrac{1}{2}\epsilon^2\left(3 + 4k^{-1}\int_0^\infty e^{ik\xi}\chi(k,\xi)\sin k\xi\, d\xi\right)\right], \tag{41}$$

$$\chi(k,\xi) = k^2\Gamma(c_0^{-1}\xi,\xi) - 2ikc_0^{-1}\Gamma_\tau(c_0^{-1}\xi,\xi) - c_0^{-2}\Gamma_{\tau\tau}(c_0^{-1}\xi,\xi), \tag{42}$$

$$C(k) = 1 + 2\epsilon^2\psi(k), \tag{43}$$

and

$$\psi(k) = \int_0^\infty e^{ik\xi}\chi(k,\xi)\left(\cos k\xi - \frac{\sin k\xi}{k\xi}\right)\xi\, d\xi. \tag{44}$$

In deriving Eqs. (40)–(44) higher-order terms in ϵ have been dropped.

The source term f is assumed to represent a point source in space but one which is random in time. Accordingly we write

$$f(t,\mathbf{x}) = g(t)\delta(\mathbf{x}), \tag{45}$$

where $g(t)$ is a stationary random function with zero mean. Equation (25) then yields

$$R_0(\tau,\mathbf{x},\mathbf{y}) = P_0(\tau)\delta(\mathbf{x})\delta(\mathbf{y}), \tag{46}$$

where $P_0(\tau)$ is defined by

$$P_0(\tau) = \langle g(t)g(t - \tau)\rangle_B. \tag{47}$$

Upon transforming Eq. (46) according to Eq. (32) we obtain

$$S_0(\omega,\mathbf{x},\mathbf{y}) = Q_0(\omega)\delta(\mathbf{x})\delta(\mathbf{y}), \tag{48}$$

where Q_0 is the transform of P_0; i.e.,

$$Q_0(\omega) = \int P_0(\tau)e^{i\omega\tau}\, d\tau. \tag{49}$$

Expressions for $\bar{S}$ and $\tilde{S}$ can now be obtained by substituting the formula for S_0 given by Eq. (48) into Eqs. (30) and (33) and carrying out the integration over $\mathbf{x}'$ and $\mathbf{y}'$ in Eq. (30) and over $\mathbf{x}''$ and $\mathbf{y}''$ in Eq. (33). The result is

$$\bar{S}(\omega,\mathbf{x},\mathbf{y}) = Q_0(\omega)H(\omega,\mathbf{x})H^*(\omega,\mathbf{y}), \tag{50}$$

$$\begin{aligned}\tilde{S}(\omega,\mathbf{x},\mathbf{y}) = (2\epsilon^2/\pi c_0^4)\int\int H_0(\omega,\mathbf{x}')H_0^*(\omega,\mathbf{y}')\\ \times[Z(\omega,\mathbf{y} - \mathbf{x} + \mathbf{x}' - \mathbf{y}')*\omega^4 Q_0(\omega)H(\omega,\mathbf{x} - \mathbf{x}')\\ \times H^*(\omega,\mathbf{y} - \mathbf{y}')]\, d\mathbf{x}'\, d\mathbf{y}'.\end{aligned} \tag{51}$$

The spectral density function $S(\omega,\mathbf{x},\mathbf{y})$ can now be calculated in terms of known functions with the aid of Eqs. (27), (50), (51), (38), and (40).

II. HIGH-FREQUENCY WAVES

The expressions for $\overline{S}$ and $\widetilde{S}$ given by Eqs. (50) and (51) can be considerably simplified in the case of high-frequency waves; i.e., when the characteristic frequency of the source is much greater than that of the medium. In considering this case we shall restrict our attention to the power spectrum $Q(\omega,\mathbf{x})$, which is defined by

$$Q(\omega,\mathbf{x}) = S(\omega,\mathbf{x},\mathbf{x}). \tag{52}$$

From Eq. (27) we have

$$Q(\omega,\mathbf{x}) = \overline{Q}(\omega,\mathbf{x}) + \widetilde{Q}(\omega,\mathbf{x}), \tag{53}$$

where

$$\overline{Q}(\omega,\mathbf{x}) \equiv \overline{S}(\omega,\mathbf{x},\mathbf{x}), \tag{54}$$

$$\widetilde{Q}(\omega,\mathbf{x}) \equiv \widetilde{S}(\omega,\mathbf{x},\mathbf{x}). \tag{55}$$

Eqs. (50) and (51) yield

$$\overline{Q}(\omega,\mathbf{x}) = Q_0(\omega)|H(\omega,\mathbf{x})|^2, \tag{56}$$

$$\widetilde{Q}(\omega,\mathbf{x}) = (2\epsilon^2/\pi c_0^4)\int\int H_0(\omega,\mathbf{x}')H_0^*(\omega,\mathbf{y}') \times[Z(\omega,\mathbf{x}'-\mathbf{y}')*\omega^4 Q_0(\omega)H(\omega,\mathbf{x}-\mathbf{x}') \times H^*(\omega,\mathbf{x}-\mathbf{y}')]\,d\mathbf{x}'\,d\mathbf{y}'. \tag{57}$$

After changing the integration variables in Eq. (57) we can write

$$\widetilde{Q}(\omega,\mathbf{x}) = (2\epsilon^2/\pi c_0^4)\int\int H_0(\omega,\mathbf{x}-\mathbf{x}')H_0^*(\omega,\mathbf{x}-\mathbf{x}'') \times[Z(\omega,\mathbf{x}''-\mathbf{x}')*\omega^4 Q_0(\omega) \times H(\omega,\mathbf{x}')H^*(\omega,\mathbf{x}'')]\,d\mathbf{x}'\,d\mathbf{x}''. \tag{58}$$

The first step in the high-frequency analysis is to obtain an asymptotic expansion, valid for large k, for the quantity κ. This is easily accomplished by integrating by parts in Eq. (41), after substituting for χ from Eq. (42). This yields the approximation

$$\kappa \simeq k + i\alpha, \tag{59}$$

where

$$\alpha = \epsilon^2 k^2 l. \tag{60}$$

The quantity l is a characteristic length scale associated with the medium, and is defined by

$$l = \int_0^\infty \Gamma(c_0^{-1}\xi,\xi)\,d\xi. \tag{61}$$

With the aid of Eqs. (38), (40), and (59) we can write Eqs. (56) and (58) in the form

$$\overline{Q}(\omega,\mathbf{x}) = Q_0(\omega)|C(k)|^2(4\pi x)^{-2}e^{-2\alpha x}, \tag{62}$$

$$\widetilde{Q}(\omega,\mathbf{x}) = \frac{2\epsilon^2}{\pi(4\pi)^4}\int\int\frac{e^{ik(|\mathbf{x}-\mathbf{x}'|-|\mathbf{x}-\mathbf{x}''|)}}{|\mathbf{x}-\mathbf{x}'|x'\,|\mathbf{x}-\mathbf{x}''|x''} \times[Z(\omega,\mathbf{x}''-\mathbf{x}')*k^4|C(k)|^2Q_0(\omega) \times e^{ik(x'-x'')}e^{-\alpha(x'+x'')}]\,d\mathbf{x}'\,d\mathbf{x}'', \tag{63}$$

where, from Eq. (43),

$$|C(k)|^2 = 1 + 4\epsilon^2\,\mathrm{Re}\psi(k). \tag{64}$$

[In deriving Eq. (64) terms of order ϵ^4 were dropped.]

The integral over $\mathbf{x}'$ and $\mathbf{x}''$ in Eq. (63) has been evaluated using the forward-scatter approximation. The details of that calculation are given in Appendix B. The resulting approximate expression for $\widetilde{Q}(\omega,\mathbf{x})$ can be written

$$\widetilde{Q}(\omega,\mathbf{x}) = (4\pi x)^{-2}[W(\omega)*|C(k)|^2(1-e^{-2\alpha x})Q_0(\omega)], \tag{65}$$

where W is defined by

$$W(\omega) = (4\pi l)^{-1}\hat{Z}(\omega,k), \tag{66}$$

and

$$\hat{Z}(\omega,\nu) = 2\int_0^\infty Z(\omega,\xi)\cos\nu\xi\,d\xi. \tag{67}$$

An expression for $Q(\omega,\mathbf{x})$ can now be obtained by substituting the formulas for $\overline{Q}(\omega,\mathbf{x})$ and $\widetilde{Q}(\omega,\mathbf{x})$ given by Eqs. (62) and (65) into Eq. (53). In so doing we simplify matters slightly by making the approximation $|C(k)|^2 = 1$. After dividing through by the spherical-spreading term $(4\pi x)^{-2}$ we obtain finally

$$(4\pi x)^2 Q(\omega,\mathbf{x}) = e^{-2\alpha x}Q_0(\omega) + W(\omega)*(1-e^{-2\alpha x})Q_0(\omega). \tag{68}$$

It should be pointed out here that, although the error in the general formulas for R and S given by Eqs. (21), (22), (23), (27), (30), and (33) is of order ϵ^3, the error in Eq. (68) is of order ϵ^2. This is because some terms of order ϵ^2 were dropped in the derivation of this equation.

Equation (68) is the main result of the high-frequency analysis. It shows that, as $\alpha x \to 0$,

$$(4\pi x)^2 Q(\omega,\mathbf{x}) \to Q_0(\omega).$$

Thus, as ϵ and/or x (the source-field point distance) goes to zero, the wave spectrum (with the spherical-spreading term factored out) approaches the source spectrum, as we would expect. In the opposite limit, i.e., as $\alpha x \to \infty$, Eq. (68) shows that

$$(4\pi x)^2 Q(\omega,\mathbf{x}) \to W*Q_0(\omega).$$

We see therefore that the wave spectrum (again apart from the spherical-spreading term) tends to a limiting form as $x \to \infty$. This limiting form, which is given by the convolution of W with Q_0, is referred to here as the fully-developed spectrum.

It may be verified by direct integration of Eq. (68) that

$$(4\pi x)^2\int Q(\omega,\mathbf{x})\,d\omega = \int Q_0(\omega)\,d\omega. \tag{69}$$

In the derivation of Eq. (69) we have used the fact that

$$\int W(\omega)\,d\omega = 1.$$

Equation (69) shows that the total signal power; i.e., the area under the spectral curve, normalized by the spherical-spreading term, is conserved.

A. Narrow-band source

We can simplify Eq. (68) further by assuming a narrow-band source, i.e., by assuming that the characteristic width of the source spectrum $Q_0(\omega)$ is much less than that of the function $W(\omega)$. Then, insofar as the convolution integral is concerned, $Q_0(\omega)$ can be regarded as a delta function. Accordingly, we replace $Q_0(\omega)$ in the convolution term by

$$A_0[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$$

(since Q_0 must be an even function), where $\omega_0 > 0$ (ω_0 is called the carrier frequency) and $A_0 > 0$. We can also write, in this case,

$$e^{-2\alpha x} Q_0(\omega) \simeq e^{-2\alpha_0 x} Q_0(\omega),$$

where $\alpha_0 = \epsilon^2 k_0^2 l$ and $k_0 = \omega_0/c_0$. Then Eq. (68) becomes

$$(4\pi x)^2 Q(\omega,\mathbf{x}) = e^{-2\alpha_0 x} Q_0(\omega) + (1 - e^{-2\alpha_0 x}) Q_\infty(\omega), \tag{70}$$

where

$$Q_\infty(\omega) = A_0[W(\omega - \omega_0) + W(\omega + \omega_0)]. \tag{71}$$

Equation (70) shows that, in the narrow-band case, the wave spectrum broadens with increasing propagation distance, with increasing strength of the randomness of the medium, and with increasing carrier frequency.

By introducing a "broadening parameter" β, defined by

$$\beta = 1 - e^{-2\alpha_0 x}, \tag{72}$$

we can write Eq. (70) in the form

$$(4\pi x)^2 Q(\omega,\mathbf{x}) = (1 - \beta) Q_0(\omega) + \beta Q_\infty(\omega). \tag{73}$$

Thus we see that the wave spectrum (with the spherical-spreading term factored out) can be regarded in this case as a linear (in β) interpolation between the source spectrum $Q_0(\omega)$ and the fully-developed spectrum $Q_\infty(\omega)$.

We define the bandwidth Ω of the wave spectrum for the narrow-band case by writing

$$\Omega(x) = \left[\int_0^\infty (\omega - \omega_0)^2 Q(\omega,\mathbf{x})\, d\omega \Big/ \int_0^\infty Q(\omega,\mathbf{x})\, d\omega \right]^{1/2}. \tag{74}$$

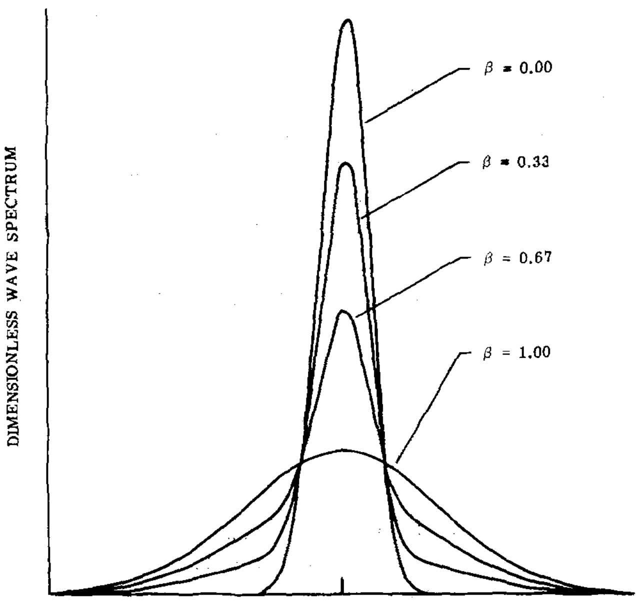


FIG. 1. Dimensionless wave spectrum (with the spherical-spreading term factored out) vs dimensionless frequency for various values of the broadening parameter β. The calculations are based on Eq. (73). The mark on the horizontal scale corresponds to the carrier frequency ω_0. The function $W(\omega)$ has a maximum at $\omega = 0$.

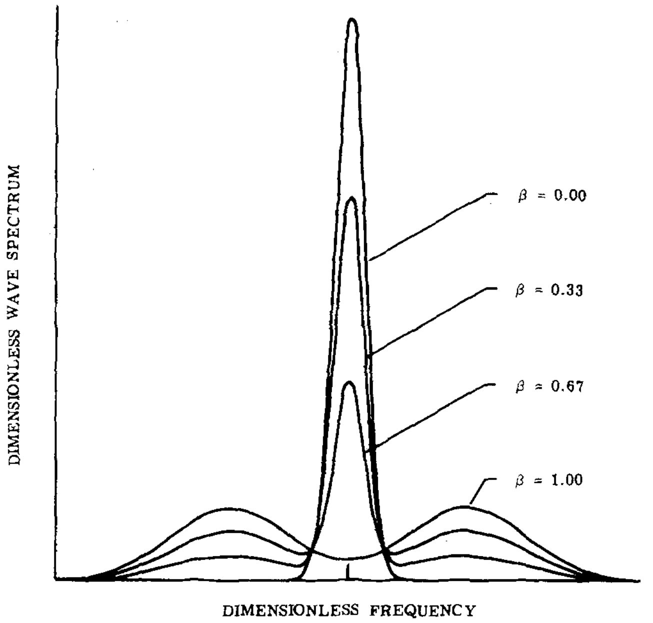


FIG. 2. Same as Fig. 1, except that the function $W(\omega)$ has a maximum at a nonzero value of ω.

By substituting the expression for Q given by Eq. (73) into Eq. (74) we obtain

$$\Omega(x) = [(1 - \beta)\Omega_0^2 + \beta \Omega_\infty^2]^{1/2}, \tag{75}$$

where Ω_0 is the bandwidth of the source spectrum and Ω_∞ is the bandwidth of the fully-developed spectrum; i.e.,

$$\Omega_0 = \left[\int_0^\infty (\omega - \omega_0)^2 Q_0(\omega)\, d\omega \Big/ \int_0^\infty Q_0(\omega)\, d\omega \right]^{1/2}, \tag{76}$$

$$\Omega_\infty = \left[\int_0^\infty (\omega - \omega_0)^2 Q_\infty(\omega)\, d\omega \Big/ \int_0^\infty Q_\infty(\omega)\, d\omega \right]^{1/2}. \tag{77}$$

Equation (75) shows that $\Omega(x)$ increases monotonically with β from the value Ω_0 at $\beta = 0$, and that it approaches Ω_∞ as $\beta \to 1$.

If we assume that $\Omega_0 \simeq 0$, which is consistent with the assumption of a narrow-band source, then Eq. (75) yields

$$\Omega(x) \simeq \beta^{1/2} \Omega_\infty. \tag{78}$$

When $\alpha_0 x \ll 1$ we have, from Eq. (72), $\beta \simeq 2\alpha_0 x = 2\epsilon^2 k_0^2 l x$, and hence, from Eq. (78),

$$\Omega(x) \simeq \epsilon k_0 (2lx)^{1/2} \Omega_\infty. \tag{79}$$

Equation (79) is valid when $\epsilon k_0 (2lx)^{1/2} \ll 1$, i.e., when $\Omega \ll \Omega_\infty$. This equation shows that, when the propagation distance is small, the spectral bandwidth increases as the square root of the propagation distance, and is also linear in the carrier frequency in this range.

In order to show the broadening phenomenon graphically, numerical calculations of the quantity $(4\pi x)^2 Q(\omega,\mathbf{x})$ as a function of ω have been made for various values of β using Eq. (73). For this purpose the source spectrum $Q_0(\omega)$ was chosen to be a narrow-band Gaussian function, centered at $\omega = \omega_0$ and reflected about the $\omega = 0$ axis (since Q_0 must be an even function). The function $W(\omega)$ was also chosen to be a reflected Gaussian, with a maximum in the first instance at

$\omega = 0$ and in the second at a frequency ω_1 for which $0 < \omega_1 \ll \omega_0$.

The results of these calculations are plotted (in dimensionless coordinates) in Figs. 1 and 2. All of the curves in each figure are plotted on the same scale. Note that in each figure the curve labeled $\beta = 0$ corresponds to the source spectrum, the curve labeled $\beta = 1$ corresponds to the fully-developed spectrum, and those labeled with values of β between zero and one correspond to intermediate stages in the broadening process. Both sets of curves show clearly the broadening of the wave spectrum with increasing β. The two sets differ, however, in one respect. The results shown in Fig. 2, for which the function $W(\omega)$ has a maximum at a nonzero value of ω, are marked by the appearance of side bands on the broadened spectrum. In Fig. 1, by contrast, for which the maximum of $W(\omega)$ occurs at $\omega = 0$, no such side bands appear.

The results obtained here appear generally to be in qualitative agreement with observations, as can be seen by, for example, comparing Fig. 1 with Fig. 11 of Ref. 1 or Fig. 3 of Ref. 2. Note, moreover, that the observations reported in Refs. 1 and 2 indicate conservation of total signal power, which is also consistent with the present results [cf. Eq. (69)].

ACKNOWLEDGMENT

This report was prepared as a result of work performed under NASA Contract Nos. NAS1-14101 and NAS1-14472 at ICASE, NASA Langley Research Center, Hampton, Virgina 23665.

APPENDIX A: CALCULATION OF $G(t,\mathbf{x})$ AND $H(\omega,\mathbf{x})$

The function $G(t,\mathbf{x})$ is determined by Eq. (39), together with the initial condition $G = 0$ for $t < 0$. An equation for the transform $H(\omega,\mathbf{x})$ of $G(t,\mathbf{x})$, as defined by Eq. (31), is obtained by transforming both sides of Eq. (39). The result is

$$[\nabla^2 + (1+3\epsilon^2)k^2]H(\omega,\mathbf{x}) + (\epsilon^2 k^2/\pi)\int \xi^{-1}e^{ik\xi} \times\chi(k,\boldsymbol{\xi})H(\omega,\mathbf{x}+\boldsymbol{\xi})\,d\boldsymbol{\xi} = -\delta(\mathbf{x}), \qquad \text{(A1)}$$

where the function $\chi(k,\boldsymbol{\xi})$ is defined by

$$\chi(k,\boldsymbol{\xi}) = k^2\Gamma(c_0^{-1}\xi,\boldsymbol{\xi}) - 2ikc_0^{-1}\Gamma_\tau(c_0^{-1}\xi,\boldsymbol{\xi}) - c_0^{-2}\Gamma_{\tau\tau}(c_0^{-1}\xi,\boldsymbol{\xi}). \qquad \text{(A2)}$$

In order to solve Eq. (A1) we introduce the spatial Fourier transform $\widehat{H}(\omega,\mathbf{m})$ of $H(\omega,\mathbf{x})$, defined by

$$\widehat{H}(\omega,\mathbf{m}) = \int H(\omega,\mathbf{x})e^{-i\mathbf{m}\cdot\mathbf{x}}\,d\mathbf{x}, \qquad \text{(A3)}$$

where $\mathbf{m}\cdot\mathbf{x} \equiv \Sigma_{i=1}^3 m_i x_i$. Transforming both sides of Eq. (A1) according to the prescription given by Eq. (A3) and solving for $\widehat{H}$ yields

$$\widehat{H}(\omega,\mathbf{m}) = [D(k,\mathbf{m})]^{-1}, \qquad \text{(A4)}$$

where we have defined

$$D(k,\mathbf{m}) = m^2 - (1+3\epsilon^2)k^2 - (\epsilon^2k^2/\pi)\int \xi^{-1}e^{ik\xi} \times\chi(k,\boldsymbol{\xi})e^{i\mathbf{m}\cdot\boldsymbol{\xi}}\,d\boldsymbol{\xi}. \qquad \text{(A5)}$$

With the aid of Eq. (A4) we can now express $H(\omega,\mathbf{x})$ as an inverse transform, i.e., we write

$$H(\omega,\mathbf{x}) = (8\pi^3)^{-1}\int [D(k,\mathbf{m})]^{-1}e^{i\mathbf{m}\cdot\mathbf{x}}\,d\mathbf{m}. \qquad \text{(A6)}$$

In order to proceed further we assume that the medium is statistically isotropic, so that we can write $\Gamma(\tau,\boldsymbol{\xi}) = \Gamma(\tau,\xi)$. Then, in view of Eq. (A2), we can also write $\chi(k,\boldsymbol{\xi}) = \chi(k,\xi)$. The angular integration in Eq. (A5) can now be carried out, yielding

$$D(k,\mathbf{m}) = m^2 - (1+3\epsilon^2)k^2 - 4\epsilon^2k^2m^{-1} \times\int_0^\infty e^{ik\xi}\chi(k,\xi)\sin m\xi\,d\xi. \qquad \text{(A7)}$$

Upon substituting the expression for D given by Eq. (A7) into Eq. (A6) and carrying out the angular integration we find that

$$H(\omega,\mathbf{x}) = (2\pi^2x)^{-1}\int_0^\infty [D(k,m)]^{-1}m\sin mx\,dm. \qquad \text{(A8)}$$

The integral in Eq. (A8) can be evaluated by means of contour integration, after which the expression for H can be written

$$H(\omega,\mathbf{x}) = (2\pi x)^{-1}[D_m(k,\kappa)]^{-1}\kappa e^{i\kappa x}. \qquad \text{(A9)}$$

Here D_m denotes the derivative of $D(k,m)$ with respect to m (regarded now as a complex variable), and κ is the root of the dispersion equation $D(k,\kappa) = 0$ which has the property that $\kappa \to k$ as $\epsilon \to 0$. This root is given, to lowest order in ϵ, by Eq. (41). Upon substituting the expression for κ given by Eq. (41) into Eq. (A9), after calculating D_m using Eq. (A7), we obtain the expression for H given by Eq. (40).

The function $G(t,\mathbf{x})$ can now be obtained by applying the inverse Fourier transform to Eq. (40). We shall not carry out that calculation here, however, since we need only the function $H(\omega,\mathbf{x})$. That calculation was carried out in Ref. 20 for the case of a time-independent medium.

APPENDIX B: CALCULATION OF $\widetilde{Q}$ USING THE FORWARD-SCATTER APPROXIMATION

By making explict the ω' integration in Eq. (63) and changing the order of integration we can write the expression for $\widetilde{Q}$ in the form

$$\widetilde{Q}(\omega,\mathbf{x}) = [2\epsilon^2/\pi(4\pi)^4] \times\int k'^4|C(k')|^2Q_0(\omega')I(\omega,\mathbf{x};\omega')\,d\omega', \qquad \text{(B1)}$$

where the integral I is defined by

$$I(\omega,\mathbf{x};\omega') = \int\int \frac{e^{ik'(|\mathbf{x}-\mathbf{x}'|-|\mathbf{x}-\mathbf{x}''|)}}{|\mathbf{x}-\mathbf{x}'|x'|\mathbf{x}-\mathbf{x}''|x''} Z(\omega-\omega',\mathbf{x}''-\mathbf{x}') \times e^{ik'(x'-x'')}e^{-\alpha'(x'+x'')}\,d\mathbf{x}'\,d\mathbf{x}''. \qquad \text{(B2)}$$

Here $k' = \omega'/c_0$ and $\alpha' = \epsilon^2k'^2l$. Eq. (B2) can be written in the alternate form

$$I(\omega,x;\omega') = \int\int \frac{e^{ik'(|\mathbf{x}-\mathbf{x}'|+x')}}{|\mathbf{x}-\mathbf{x}'|x'}\,\frac{e^{-ik'(|\mathbf{x}-\mathbf{x}''|+x'')}}{|\mathbf{x}-\mathbf{x}''|x''}$$

$$\times Z(\omega-\omega',\mathbf{x}''-\mathbf{x}')e^{i(k-k')|\mathbf{x}-\mathbf{x}'|}$$
$$\times e^{-i(k-k')|\mathbf{x}-\mathbf{x}''|}e^{-\alpha'(x'+x'')}\,d\mathbf{x}'\,d\mathbf{x}'', \qquad \text{(B3)}$$

which is more convenient for the application of the forward-scatter approximation.

We begin the analysis by substituting for Z in terms of Γ in Eq. (B3) with the aid of Eq. (35). By changing the order of integration in the resulting expression for I we get

$$I=\int e^{i(\omega-\omega')}\int\int\frac{e^{ik'(|\mathbf{x}-\mathbf{x}'|+x')}}{|\mathbf{x}-\mathbf{x}'|x'}\frac{e^{-ik'(|\mathbf{x}-\mathbf{x}''|+x'')}}{|\mathbf{x}-\mathbf{x}''|x''}$$
$$\times\Gamma(\tau,\mathbf{x}''-\mathbf{x}')e^{i(k-k')|\mathbf{x}-\mathbf{x}'|}e^{-i(k-k')|\mathbf{x}-\mathbf{x}''|}$$
$$\times e^{-\alpha'(x'+x'')}\,d\mathbf{x}'\,d\mathbf{x}''\,d\tau. \qquad \text{(B4)}$$

Next we use Eq. (24) to substitute for Γ in terms of μ in Eq. (B4). Upon reversing the order of the averaging and integration (over $\mathbf{x}'$ and $\mathbf{x}''$) processes, we note that the double spatial integral can be split into a product of two integrals. Equation (B4) can then be written

$$I=\int e^{i(\omega-\omega')\tau}\langle J\rangle_A\,d\tau, \qquad \text{(B5)}$$

where

$$J=J_+\,J_-, \qquad \text{(B6)}$$

$$J_+=\int\frac{e^{ik'(|\mathbf{x}-\mathbf{x}'|+x')}}{|\mathbf{x}-\mathbf{x}'|x'}\mu(t,\mathbf{x}')$$
$$\times e^{i(k-k')|\mathbf{x}-\mathbf{x}'|}e^{-\alpha'x'}\,d\mathbf{x}', \qquad \text{(B7)}$$

and

$$J_-=\int\frac{e^{-ik'(|\mathbf{x}-\mathbf{x}''|+x'')}}{|\mathbf{x}-\mathbf{x}''|x''}\mu(t-\tau,\mathbf{x}'')$$
$$\times e^{-i(k-k')|\mathbf{x}-\mathbf{x}''|}e^{-\alpha'x''}\,d\mathbf{x}''. \qquad \text{(B8)}$$

We can now apply the forward-scatter approximation, as discussed in Ref. 21, to the integrals J_+ and J_-. This yields

$$J_+=(2\pi i/k'x)e^{ik'x}\int_0^x\mu(t,0,0,x')$$
$$\times e^{i(k-k')(x-x')}e^{-\alpha'x'}\,dx'+O(k'^{-2}), \qquad \text{(B9)}$$

$$J_-=-(2\pi i/k'x)e^{-ik'x}\int_0^x\mu(t-\tau,0,0,x'')$$
$$\times e^{-i(k-k')(x-x'')}e^{-\alpha'x''}\,dx''+O(k'^{-2}). \qquad \text{(B10)}$$

In the derivation of Eqs. (B9) and (B10) we have set $\mathbf{x}=(0,0,x)$. This entails no loss of generality since the medium has been assumed statistically isotropic.

Conditions for the validity of the forward-scatter approximation are given in Ref. 21. In the present context these conditions take the form

$$k_1^{-1}\ll x\ll k_1 l^2, \qquad \text{(B11)}$$

where k_1 is a characteristic wave number associated with the wave field.

By substituting the expressions for J_+ and J_- given by Eqs. (B9) and (B10) into Eq. (B6), dropping terms of order k'^{-3}, and averaging, we obtain

$$\langle J\rangle_A=(4\pi^2/k'^2x^2)\int_0^x\int_0^x\Gamma(\tau,x''-x')$$
$$\times e^{i(k-k')(x''-x')}e^{-\alpha'(x''+x')}\,dx'\,dx''. \qquad \text{(B12)}$$

The double integral in Eq. (B12) can be partially evaluated with the aid of the coordinate transformation $\xi=x''-x'$, $\eta=x''+x'$. The result is

$$\langle J\rangle_A=(4\pi^2/\alpha'k'^2x^2)\int_0^x\Gamma(\tau,\xi)$$
$$\times(e^{-\alpha'\xi}-e^{-2\alpha'x}e^{\alpha'\xi})\cos[(k-k')\xi]\,d\xi. \qquad \text{(B13)}$$

In deriving Eq. (B13) we have made use of the fact that $\Gamma(\tau,\xi)$ is even in ξ.

We can now get a series expansion for $\langle J\rangle_A$ in powers of α' (which is equivalent to an expansion in powers of ϵ^2) by expanding the terms $\exp(\alpha'\xi)$ and $\exp(-\alpha'\xi)$ in Eq. (B13) and integrating term by term. This yields

$$\langle J\rangle_A=(4\pi^2/\alpha'k'^2x^2)\sum_{n=0}^{\infty}(-1)^n$$
$$\times[1-(-1)^n e^{-2\alpha'x}](\alpha'^n/n!)$$
$$\times\int_0^x\xi^n\Gamma(\tau,\xi)\cos[(k-k')\xi]\,d\xi. \qquad \text{(B14)}$$

When $x\gg l$ the integration in Eq. (B14) can be extended to $+\infty$ without introducing significant error into the integral. Upon dropping all but the first term of the resulting expansion we obtain

$$\langle J\rangle_A\simeq(4\pi^2/\alpha'k'^2x^2)(1-e^{-2\alpha'x})$$
$$\times\int_0^\infty\Gamma(\tau,\xi)\cos[(k-k')\xi]\,d\xi. \qquad \text{(B15)}$$

An approximate expression for the integral I can now be obtained by substituting the result for $\langle J\rangle_A$ given by Eq. (B15) into Eq. (B5) and carrying out the integration over τ. This yields

$$I=(2\pi^2/\alpha'k'^2x^2)(1-e^{-2\alpha'x})\hat{Z}(\omega-\omega',k-k'), \qquad \text{(B16)}$$

where $\hat{Z}$ is defined by Eq. (67). Upon combining Eqs. (B1) and (B16) we obtain the expression for $\tilde{Q}$ given by Eq. (65).

[1]D. C. Mathews, R. T. Nagel, and J. D. Kester, "Review of Theory and Methods for Turbine Noise Prediction," in *Progress in Astronautics and Aeronautics*, edited by M. Summerfield (MIT Press, Cambridge, Massachusetts, 1976), Vol. 46, pp. 155–176.

[2]R. M. Goldstein, "Suprior Conjunction of Pioneer 6," Science **166**, 598–601 (1969).

[3]M. S. Howe, "Multiple Scattering of Sound by Turbulence and Other Inhomogeneities," J. Sound Vib. **27**, 455–76 (1973).

[4]R. L. Fante, "Mutual Coherence Function and Frequency Spectrum of a Laser Beam Propagating Through Atmospheric Turbulence," J. Opt. Soc. Am. **64**, 592–98 (1974).

[5]R. L. Fante, "Electric Field Spectrum and Intensity Covariance of a Wave in a Random Medium," Radio Science **10**, 77–85 (1975).

[6]R. Woo, F.-C. Yang, and A. Ishimaru, "Structure of Density Fluctuations Near the Sun Deduced from Pioneer-6 Spectral Broadening Measurements," Astrophys. J. **210**, 593–602 (1976).

[7]A. Ishimaru, *Wave Propagation and Scattering in Random Media*, (Academic, New York, 1978). Vol. 2.

[8]N. G. Kuznetsova and L. A. Chernov, "Allowance for the Doppler Effect in the Theory of Multiple Wave Scattering in a Randomly Inhomogeneous Medium," Sov. Phys. Acoust. **23**, 93–94 (1977).

[9]G. Adomian, "Linear Random Operator Equations in Mathematical Physics III," J. Math. Phys. **12**, 1948–55 (1971).

[10]V. I. Tatarskii, *The Effects of the Turbulent Atmosphere on Wave Propagation* (Natl. Tech. Inf. Serv., Springfield, Virginia, 1971).
[11]A. Spizzichino, "Spectral Broadening of Acoustic and Radio Waves Scattered by Atmospheric Turbulence in the Case of Radar and Sodar Experiments," Ann. Geophys. **31**, 433–45 (1975).
[12]I. Kupiec, S. Rosenbaum, and L. B. Felsen, "Backscatter from Scattering Cells Located in a Random Medium," Radio Science **12**, 1–10 (1977).
[13]E. H. Brown and S. F. Clifford, "Spectral Broadening of an Acoustic Pulse Propagating Through Turbulence," J. Acoust. Soc. Am. **54**, 36–39 (1973).
[14]E. H. Brown, "Turbulent Spectral Broadening of Backscattered Acoustic Pulses," J. Acoust. Soc. Am. **56**, 1398–1406 (1974).
[15]A. Ishimaru, "Temporal Frequency Spectra of Multi-frequency Waves in Turbulent Atmosphere," IEEE Trans. Antennas Propag. **AP-20**, 10–19 (1972).
[16]H. T. Yura, "Temporal-Frequency Spectrum of an Optical Wave Propagating Under Saturation Conditions," J. Opt. Soc. Am. **64**, 357–59 (1974).
[17]J. B. Keller, "Stochastic Equations and Wave Propagation in Random Media," Proc. Symp. Appl. Math. **16**, 145–70 (1964).
[18]U. Frisch, "Wave Propagation in Random Media," in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid (Academic, New York, 1968), Vol. 1, p. 114.
[18]A. R. Wenzel, "Frequency Spectra of Solutions of Stochastic Equations, With Applications to Spectral Broadening of Waves Propagating in a Random Medium," ICASE Report No. 78-11 (1978).
[20]A. R. Wenzel, "Propagation of Transients in a Random Medium," J. Math. Phys. **16**, 44–49 (1975).
[21]A. R. Wenzel, "Note on the Forward-Scatter Approximation," J. Acoust. Soc. Am. **61**, 882–3 (1977).

Image methods for constructing Green's functions and eigenfunctions for domains with plane boundaries

Riho Terras
Department of Mathematics, University of California, San Diego, La Jolla, California 92093
Robert Swanson
Department of Physics, University of California, San Diego, La Jolla, California 92093

(Received 7 February 1980; accepted for publication 2 May 1980)

We consider the image problem for domains with plane boundaries. We list all three and two dimensional domains for which the image method yields solutions of the potential problem, and we describe the image arrays generated by these domains in familiar crystallographic terms. One obtains from the group–theoretic description of images two representations for the Dirichlet Green's functions for ∇^2. The first is obtained by summing the unrestricted Green's function over the crystal image structures, and the second is obtained in terms of an eigenfunction expansion using solutions of $\nabla^2\psi = \lambda\psi$ which vanish on the plane boundaries.

I. INTRODUCTION

If a point charge q is at some distance d from a grounded conducting plane, the boundary condition imposed by the plane on the resulting potential may be satisfied by replacing the plane with an "image charge" $-q$ located at a position which is the mirror image location of q. This type of solution was called the "method of images" by its inventor, Sir William Thomson,[1] and is illustrated in Fig. 1.

We have studied the general problem of a solution by images for a point charge in a domain bounded by several grounded conducting planes, with the unexpected result that we are able to list all such domains for which the image solution exists. The possible corners are limited to intersections of three planes and are well known in the theory of regular polyhedra. In each case, the set of image charges at such a corner forms a representation of a finite point group. Additional planes result in an infinite crystal structure of image charges in which the unit cell is the finite group of images at a corner. There are no domains whose boundary consists of more than six planes.

The existence of the group structure yields the surprise that one can determine complete systems of fundamental eigenfunctions of the Laplace operator ∇^2, i.e., solutions of $\nabla^2\psi = \lambda\psi$ which vanish on the boundaries. Moreover, the only classical cases are the eigenfunctions for the box, the square, and the three types of Lamé eigenfunctions.[2]

Given one of our domains there are two ways to represent a Green's function for it, i.e., the potential for a unit charge. One is a direct sum of multipoles determined by a unit cell of images, and the other is an eigenfunction expansion.

In Sec. II we deduce the allowed domains with plane boundaries. In Sec. III we summarize the group theory appropriate to the problem, and we show that reflections in a plane not containing the corner vertex generate an infinite crystal structure. In Sec. IV we show that the existence of the image solution for potentials and Green's functions follows directly from the group structure of the array of image charges. In Sec. V we calculate the image arrays for cylinders and prisms formed by terminating a cylinder with planes normal to the cylinder axis. In Sec. VI we consider the tetrahedral domains, which make essential demands on our group theoretic formalisms. In Sec. VII we consider the icosahedral Möbius corner. In Sec. VIII we display a general formula for eigenfunctions of image domains and we devise a completeness proof. In Sec. IX, X, and XI we display explicit eigenfunctions as well as the Green's function expansions for the more interesting image domains.

In an earlier publication[3] we illustrated the details of some constructions and announced some of our principal results.

II. IMAGE DOMAINS WITH PLANE BOUNDARIES

The image solution exists if the potential of the original charge and its images vanishes on each conducting boundary plane, and no proper image lies in the domain bounded by the conducting planes. We first consider the wedge formed by two intersecting planes, where it is well known that the necessary and sufficient condition for existence of an image solution is that the wedge angle is π/n, with n an integer greater than 1. For a domain bounded by more than two planes, it remains a necessary condition that each pair of intersecting planes meets at π/n. We proceed to find all domains bounded by planes which satisfy this necessary constraint. In succeeding sections we show that this necessary condition is sufficient by constructing the space group that generates the image charge array. This group determines the images completely. Thus existence of the domains we seek

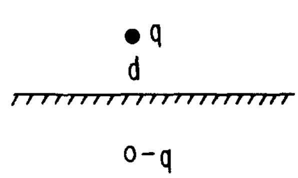


FIG. 1. *Single Plane.* Light and dark circles are charges of opposite sign.

 0022-2488/80/082140-14$1.00

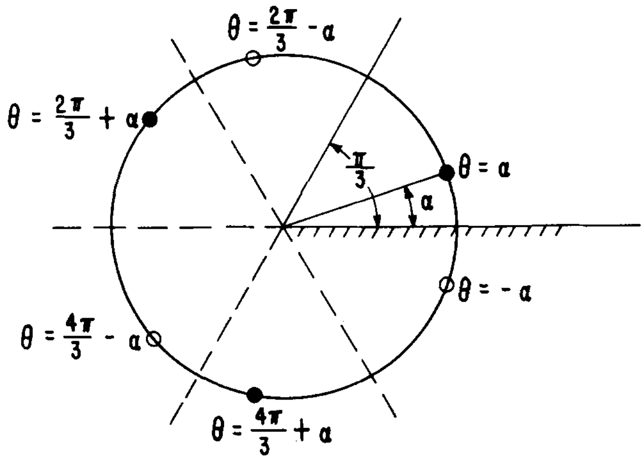


FIG. 2. *The* $\pi/3$*–Wedge*. The array generated is represented by dark and light circles (charges of opposite sign).

will follow from the group structure of the image array. A general domain for which the image method works is called an *image domain*.

A. Intersection of Two Planes

Many texts[4] apply the method of images to a point charge placed between a pair of intersecting planes, where it is easy to show that the image solution exists if and only if the angle between the planes in π/n, with n an integer greater than 1. If n is an integer, there are $2n-1$ image charges as shown in Fig. 2. If n is not an integer, the successive reflections needed to satisfy the boundary condition $V=0$ produce images which lie in the domain $0 \leqslant \theta \leqslant \pi/n$. We refer to the domain bounded by planes at an angle π/n as a π/n *wedge*. The case of two parallel planes may be considered as the limit $n \rightarrow \infty$, in which case the number of image charges is infinite.[5] A single plane may be considered as the special case $n=1$.

B. Open Cylinders

Since the interior angles of a polygon with n sides add up to $(n-2)\pi$, the only possible cylindrical cross sections are triangles with angles $(\pi/3|\pi/3|\pi/3)$, $(\pi/2|\pi/4|\pi/4)$, $(\pi/2|\pi/3|\pi/6)$, the rectangle $(\pi/2|\pi/2|\pi/2|\pi/2)$, and an open figure $(\pi/2|\pi/2|0)$, in which two sides meet at infinity. The cylinders are shown in Fig. 4.

C. Corners

If a corner of n planes has its apex at the center of a sphere, the intersection angles between planes are seen to be the interior angles at the corners of a spherical polygon. For a spherical polygon, the sum Σ of the interior angles satisfies $(n-2)\pi < \Sigma < (n+2)\pi$. Simple enumeration shows that the only possible corners have three planes with angles $(\pi/2|\pi/2|\pi/n)$, $(\pi/2|\pi/3|\pi/3)$, $(\pi/2|\pi/3|\pi/4)$, $(\pi/2|\pi/3|\pi/5)$. The spherical triangles associated with such corners are known in the theory of regular polyhedra as the *Möbius triangles*. We shall call such corners *Möbius corners*.

We also note that since the interior angles of an n–sided

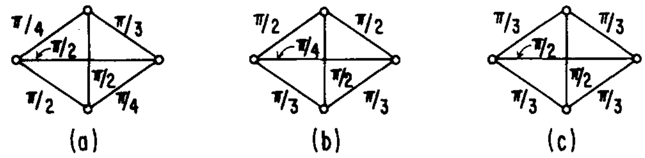


FIG. 3. *Angles*. Angles of intersection for the admissible bounded four sided image domains. With terminology from VI one has: (a) primitive octahedral domain, (b) centered octahedral domain, and (c) large tetrahedral domain.

plane polygon add up to $(n-2)\pi$, the faces of a closed domain with the above corners must be triangles or rectangles.

D. Other Open Domains

With the limitation to corners listed in (C), the only other open domains are cylinders from (B) terminated at one end by a plane perpendicular to the cylinder axis, and a wedge of two planes at an angle π/n intersected by two parallel planes having the π/n intersection as a normal. We find these by the enumeration described in part (E).

E. Closed Domains—4 Faces

Four planes intersect in six lines, each of which is shared by two faces. The only domains are thus tetrahedra with triangular faces. We find the allowed domains by a simple enumeration which consists of taking each corner from (C), intersecting the surfaces with a plane which makes one allowed corner, and then testing the remaining corners. Only three basic tetrahedra emerge. They are described in Fig. 3 by the topology of the corner angles. Fig. 5 shows their con-

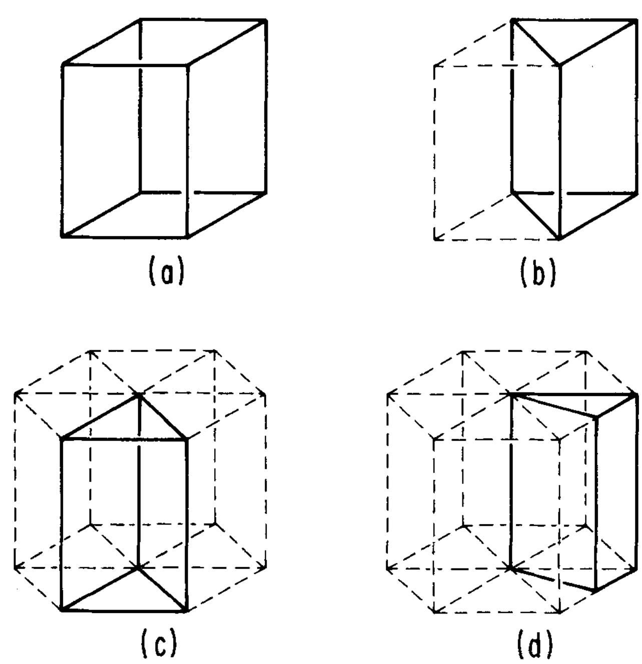


FIG. 4. *Geometric Types*. The bounded prismatic image domains are: (a) rectangular orthohombic domain, (b) $(\pi/2|\pi/4|\pi/4)$–triangular domain, (c) $(\pi/3|\pi/3|\pi/3)$–triangular domain, and (d) $(\pi/2|\pi/3|\pi/6)$–triangular domain.

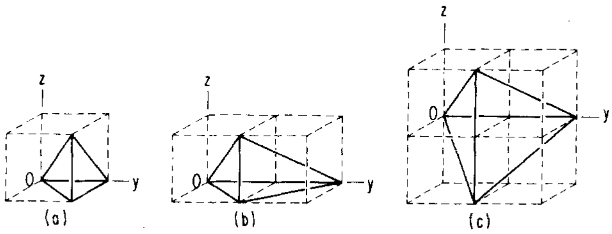


FIG. 5. *Geometric Types.* Cubic constructions show the relationship between the three distinct types of bounded four-sided image domains. Depicted are: (a) primitive octahedral domain, (b) centered octahedral domain, and (c) large tetrahedral domain.

struction by sectioning a cube. We can also arrive at these domains using Descartes' theorem in part (G).

F. Closed Domains—5 Faces

There are no five-sided domains in which each face intersects all the other faces. This would produce 10 lines of intersection; since the number of corners is integer and 2/3 the number of lines, this is impossible. If two planes do not intersect in the surface, this gives 9 lines of intersection which must be the edges of two triangles and three rectangles. The only such domains are the triangular cylinders of part (B) intersected by a pair of parallel planes normal to the cylinder axis. See Fig. 4.

G. Closed Domains—6 or More Faces

Since all planes meet at angles less than π, they form only convex polyhedra. If

$C =$ number of corners,

$E =$ number of edges,

$F =$ number of faces,

Euler's theorem states

$$C + F - E = 2.$$

Since all corners are formed from 3 edges, it follows that $2E = 3C$. Hence

$$C = 2F - 4.$$

Thus, a figure of 6 faces has 8 corners. The sum of the plane angles at a corner is less than 2π. The difference is called the *angular defect*; the corner $(\pi/2|\pi/2|\pi/2)$ has the smallest angular defect, which is $\pi/2$. A theorem of Decartes which can easily be derived from Euler's theorem,[3,6] states that the angular defects sum to 4π for the corners of a convex polyhedron. Hence, we have $C\leqslant 8$, or from Euler's theorem $F\leqslant 6$. For $F = 6$, all corners must be $(\pi/2|\pi/2|\pi/2)$ and the figure is a rectangular parallelipeped. Descartes' theorem, with the condition that an edge angle is shared by two adjacent corners, may be used to derive the results of parts (E) and (F).

The above domains were found by imposing the necessary condition that the angle between any two intersecting planes is π/n. We need to show that the image arrays created by our admissible domains do not contain additional image points which lie inside these domains. We also need to demonstrate that the potential and Green's functions that arise vanish on all the boundaries. As is shown in the following sections, each of the above domains satisfies these conditions. Our necessary condition for a region to be an image domain is thus also sufficient.

III. GROUP THEORY

In order to describe the crystal structures of images, we are fortunate in having at our disposal the language of groups as they occur in solid–state theory and crystallography. In this section we shall recall and elaborate on the necessary group-theoretical formalisms, but for a detailed elementary description of groups we shall refer the reader elsewhere.[7,8,9]

If $\mathbf{G}$ is a group and $\mathbf{H}$ and $\mathbf{K}$ are subgroups, then one defines $\mathbf{HK} = \{\mu\lambda \,|\mu\in\mathbf{H}, \lambda\in\mathbf{K}\}$. If $\mathbf{G} = \mathbf{HK}$ and the subgroups $\mathbf{H}$ and $\mathbf{K}$ have only the identity element in common the $\mathbf{G} = \mathbf{HK}$ will be called a *product decomposition*. If $\mathbf{G} = \mathbf{HK}$ is a product decomposition and $\mu\lambda = \lambda\mu$ for all $\mu\in\mathbf{H}$ and all $\lambda\in\mathbf{K}$, then $\mathbf{G} = \mathbf{HK}$ will be called a *direct product decomposition*.

In general, for arbitrary subroups $\mathbf{H}$ and $\mathbf{K}$ of $\mathbf{G}$, the set $\mathbf{S} = \mathbf{HK}$ will not be a subgroup of $\mathbf{G}$. One says that $\mathbf{H}$ is *normal* (or *invariant*) with respect to $\mathbf{K}$ if $x^{-1}\mathbf{H}x\subset\mathbf{H}$ for all $x\in\mathbf{K}$. If $\mathbf{H}$ and $\mathbf{K}$ are subgroups of $\mathbf{G}$ and if $\mathbf{H}$ is normal with respect to $\mathbf{K}$ then $\mathbf{S} = \mathbf{HK} = \mathbf{KH}$ is a subgroup of $\mathbf{G}$. The normality condition is useful for constructing groups, but it should be noted that $\mathbf{S} = \mathbf{HK}$ can be a group without the normality condition. A necessary and sufficient condition for $\mathbf{S} = \mathbf{HK}$ to be a group is that $\mathbf{HK} = \mathbf{KH}$. For a finite group $\mathbf{H}$, the number of elements of in $\mathbf{H}$, denoted by $n_{\mathbf{H}}$, is called the *order* of $\mathbf{H}$. If $\mathbf{S} = \mathbf{HK}$ is a product decomposition then $n_{\mathbf{S}} = n_{\mathbf{H}} n_{\mathbf{K}}$.

Let $\mathbf{E}$ be a vector space with inner product $\langle\mathbf{x}|\mathbf{y}\rangle$. A reflection with respect to a plane through the origin is given by the formula

$$\lambda\mathbf{x} = \mathbf{x} - 2\mathbf{a}\langle\mathbf{x}|\mathbf{a}\rangle/\langle\mathbf{a}|\mathbf{a}\rangle \quad (\mathbf{a}\in\mathbf{E},\mathbf{x}\in\mathbf{E}). \tag{1}$$

The vector $\mathbf{a}$ is said to be *normal* to the plane which determines λ.

In three dimensions one has the matrix representation

$$\lambda = \begin{bmatrix} 1-2a^2 & -2ab & -2ac \\ -2ab & 1-2b^2 & -2bc \\ -2ac & -2bc & 1-2c^2 \end{bmatrix} \tag{2}$$

when the normal (a,b,c) satisfies $a^2 + b^2 + c^2 = 1$.

For any reflection λ one has the determinant relation $\det(\lambda) = -1$. The *orthogonal group* $\mathrm{O}(n)$ for the vector space $\mathbf{E}$ is defined to be the set of all linear transformations σ on $\mathbf{E}$ which leave the inner product invariant, i.e., $\langle\sigma\mathbf{x}|\sigma\mathbf{y}\rangle = \langle\mathbf{x}|\mathbf{y}\rangle$ for all $\mathbf{x},\mathbf{y}\in\mathbf{E}$. It follows that every group generated by reflections in planes through the origin is a subgroup of $\mathrm{O}(n)$.

Two reflections $\lambda\mathbf{x} = \mathbf{x} - 2\mathbf{a}\langle\mathbf{x}|\mathbf{a}\rangle/\langle\mathbf{a}|\mathbf{a}\rangle$ and $\mu\mathbf{x} = \mathbf{x} - 2\mathbf{b}\langle\mathbf{x}|\mathbf{b}\rangle/\langle\mathbf{b}|\mathbf{b}\rangle$ commute if and only if $\langle\mathbf{a}|\mathbf{b}\rangle = 0$ or $\mathbf{a}\times\mathbf{b} = 0$. One infers this from the commutator relation

$$(\lambda\mu - \mu\lambda)\mathbf{x} = 4\langle\mathbf{a}|\mathbf{b}\rangle(\mathbf{x}\times(\mathbf{a}\times\mathbf{b})). \tag{3}$$

Two reflection λ and μ are *perpendicular* if $\langle\mathbf{a}|\mathbf{b}\rangle = 0$ holds for their normals. We have shown a reflection commutes with any perpendicular reflection and itself.

The Möbius corner $(\pi/2|\pi/2|\pi/n)$. For such a Möbius corner one reflection λ is perpendicular to the other two. If **D** is the order-two group generated by λ, and **H** is the group generated by the other two reflections, then perpendicularity implies that $\mathbf{S} = \mathbf{DH}$ is a group and a direct product. In the Schoenflies notation the group **S** is called a *dihedral* group and is denoted by the symbol $\mathbf{D}_{nh}$.

Generation of space groups. A set **L** in a vector space **E** is a *lattice* if **L** contains a set of linearly independent vectors such that every element of **L** can be expressed as an integral combination of these elements. The basis vectors will be called *primitive translation vectors*. The parallelepiped spanned by a set of such vectors is called a *primitive cell*.

The group of integers **Z** is a one-dimensional lattice in the space of real numbers. Henceforth, the space **E** will be a real three-dimensional space, even though this assumption is not formally necessary in this section.

Let S be a finite group generated by reflections with respect to planes containing the origin. If we adjoin to S a number of reflections with respect to planes not containing the origin, then the group thus generated will be called a *space group*. While the images of a point under **S** will be restricted to the surface of a sphere, the images under the space group can lie infinitely far from the origin.

Hypothesis I. We shall not add to **S** an arbitrary reflection. It is assumed that ξ is a reflection and that the related transformation $\lambda\mathbf{x} = \xi\mathbf{x} - \xi\,\mathbf{O}$, which defines a reflection λ with respect to a plane containing the origin, is an element of the group **S**.

Given this hypothesis about ξ, one writes $\mathbf{a} = \xi\,\mathbf{O}$ and thus ξ is expressible in the form $\xi\mathbf{x} = \lambda\mathbf{x} + \mathbf{a}$ with $\lambda\mathbf{a} = -\mathbf{a}$. Given a point **x** we are now interested in determing all images of **x** under the space group **G** generated by the finite group **S**.

Given $\sigma\in\mathbf{S}$ we now consider a transformation $\mu = \sigma\xi\lambda\sigma^{-1}$. One has $\mu\mathbf{x} = \mathbf{x} + \sigma\mathbf{a}$ and thus μ acts as a *raising operator*. Clearly μ is an element of the space group **G**. One has $\mu^{-1}\mathbf{x} = \mathbf{x} - \sigma\mathbf{a}$ and thus μ^{-1} acts as a *lowering operator*. It follows that $\mu^k\mathbf{x} = \mathbf{x} + k\sigma\mathbf{a}$ is an image of **x** for any integer $k\in\mathbf{Z}$. Let **L** be the integral span of $\{\sigma a|\sigma\in\mathbf{S}\}$.

The pair (**S**,**L**) can now be given a group structure by defining $(\sigma,\mathbf{n})\in(\mathbf{S},\mathbf{L})$ to be a transformation with action

$$(\sigma,\mathbf{n})\mathbf{x} = \sigma\mathbf{x} + \mathbf{n}. \tag{4}$$

The multiplication in (**S**,**L**) is defined by composition of transformations and has the formula

$$(\lambda,\mathbf{n})(\mu,\mathbf{m}) = (\lambda\mu,\lambda\mathbf{m} + \mathbf{n}), \tag{5}$$

and in the context of lattices is called *Seitz* multiplication

It may now be observed that **L** is invariant under **S** and ξ. It follows that all images of **x** under the space group **G** are of the form $(\sigma,\mathbf{n})\mathbf{x} = \sigma\mathbf{x} + \mathbf{n}$ with $\sigma\in\mathbf{S}$ and $\mathbf{n}\in\mathbf{L}$. Indeed, **G** and (**S**,**L**) can be isomorphically identified by assigning ξ to $(\lambda,\mathbf{a})$ and $\sigma\in\mathbf{S}$ to $(\sigma,\mathbf{O})$.

Our notation would suggest that **L** always has the structure of a lattice, but this need not be the case. For example, it is easy to see that the set of real numbers $\{n\sqrt{2} + m|n,m\in\mathbf{Z}\}$ can not be expressed in terms of integral multiples of a single generator.

Hypothesis II. The second hypothesis is that $\{\sigma\mathbf{a}|\sigma\in\mathbf{S}\}$ contains independent vectors such that every other vector in $\{\sigma\mathbf{a}|\sigma\in\mathbf{S}\}$ can be expressed as an integral combination of these vectors.

Under the second hypothesis the set **L** is a lattice.

In order to adjoin to **S** simulataneously two reflections ξ and ξ', it will be assumed that both satisfy Hypothesis I. One sets $\mathbf{a} = \xi\,\mathbf{O}$ and $\mathbf{a}' = \xi'\mathbf{O}$ and one defines **L** to be the integral span of the sets $\{\sigma\mathbf{a}|\sigma\in\mathbf{S}\}\cup\{\sigma\mathbf{a}'|\sigma\in\mathbf{S}\}$. If this larger integral span satisfies Hypothesis II then **L** is a lattice and (**S**,**L**) is the space group generated by $\{\mathbf{S},\xi,\xi'\}$. In this manner, any number of suitable reflections can be adjoined to **S**.

On the assumption that the space group (**S**,**L**) is determined by a lattice, our group (**S**,**L**) will be a space group in the crystalographic sense.[7] The only situation of interest in our theory and in electrostatics is the case where **L** is a Bravais lattice.

IV. IMAGE CRYSTAL STRUCTURES

The images of a point in suitable domains formed from planes form crystal structures in the abstract sense. We shall give an overview of this approach.

Space groups for image domains. One associates a space group to any domain **V** formed by plane surfaces by considering the group **G** generated by reflections with respect to the bounding planes.

If for every $g\in\mathbf{G}$ distinct from the identity and for every $\mathbf{x}\in\mathbf{V}$ the image $g\mathbf{x}$ lies outside of **V**, then **V** is an *image domain*. The group theory section shows that one should be able to represent the space group **G** in the form (**S**,**L**), where **S** is a finite group of reflections and **L** is the lattice which arises through the extension process. However, the fact that **L** is a lattice does not follow from general considerations and needs to be verified through explicit computation. Moreover, it turns out that the proper corner to choose for the generation of **S** is the sharpest corner of the domain. The resulting accounting of images will allow one to deduce that all the domains described in the first section are image domains.

Potentials for image domains. Let **V** be a corner domain or a wedge domain. Let **S** be the associated group of reflections. If a unit charge is placed at $\mathbf{x}\in\mathbf{V}$ then the potential is given by

$$\phi(\mathbf{u}) = \sum_{\sigma\in\mathbf{S}} \det(\sigma)\|\sigma\mathbf{x} - \mathbf{u}\|^{-1}. \tag{6}$$

Let **V** be a general image domain with space group (**S**,**L**). The potential for **V** is then given by

$$\Phi(\mathbf{u}) = \sum_{\mathbf{n}\in\mathbf{L}} \phi(\mathbf{n} + \mathbf{u}), \tag{7}$$

where $\phi(\mathbf{u})$ is the *corner potential* defined by (6). It is a consequence of Seitz multiplication (5) that the sign of the image charge at $(\sigma,\mathbf{n})\mathbf{x}$ is specified by the determinant of σ. When the monopole, dipole or quadrupole moments of the charge distribution $\{\det(\sigma),\sigma\mathbf{x}|\sigma\in\mathbf{S}\}$ vanish, then $\phi(\mathbf{u})$ tends to zero rapidly as **u** gets large. This circumstance allows one to assert that (7) converges absolutely. The explicit determinations of (S,L) will allow one to see this with complete rigor.

The potential satisfies the transformation property

$$\Phi\,((\sigma,\mathbf{n})\mathbf{u}) = \det(\sigma)\Phi\,(\mathbf{u}) \tag{8}$$

for an arbitrary $(\sigma,\mathbf{n})\in(\mathbf{S},\mathbf{L})$.

Vanishing of the potential. One has that $\Phi\,(\mathbf{u}) = 0$ when **u** lies on the boundary of **V**.

The assertion is easily proved. When **u** lies on the boundary of **V** then $(\sigma,\mathbf{n})\mathbf{u} = \mathbf{u}$, for $(\sigma,\mathbf{n})$ the reflection with respect to the boundary plane containing **u**. For a reflection one has $\det(\sigma) = -1$. One now computes that $\Phi\,(\mathbf{u}) = \Phi\,((\sigma,\mathbf{n})\mathbf{u}) = -\Phi\,(\mathbf{u})$. Thus $\Phi\,(\mathbf{u}) = 0$.

Interchange of summation and differentiation. One can prove mathematically that if **D** is any differential operator with respect to **u** then

$$\mathbf{D}\Phi\,(\mathbf{u}) = \sum_{\mathbf{n}\in\mathbf{L}} \mathbf{D}\phi\,(\mathbf{n}+\mathbf{u}) \tag{9}$$

and the derived series converges absolutely.

The differentiation interchange implies that $\Phi\,(\mathbf{u})$ satisfies the Laplace equation

$$\nabla^2\Phi\,(\mathbf{u}) = -4\pi\delta\,(\mathbf{x}-\mathbf{u}), \tag{10}$$

where δ is the Dirac delta function and ∇^2 is the Laplace operator with respect to **u**.

Interchange of summation and integration. This interchange is not always valid. It is worthwhile recalling the hypotheses for integration and summation interchange. If for a convergent series $\Sigma\phi_n$ of integrable functions the series $\Sigma \int |\phi_n|$ converges, then the interchange $\int \Sigma\, \phi_n = \Sigma \int \phi_n$ is valid. The example we mention is a rare instance where this interchange is not possible and where the problem arises in a natural context.

We shall consider the potential for parallel plates. The function $\|(x-u,y-v,z-w)\|^{-1}$ is the *unrestricted potential* for all space. The difference

$$\phi\,(u,v,w) = \|(x-u,y-v,z-w)\|^{-1} - \|(-x-u,y-v,z-w)\|^{-1}$$

is a special case of the group sum (7). The lattice sum

$$\Phi\,(u,v,w) = \sum_{n\in\mathbf{Z}} \phi\,(2n+u,v,w) \tag{11}$$

is an absolutely convergent series and represents the potential due to a unit charge at (x,y,z) between two conducting parallel plates at $x = 0$ and $x = 1$. The derived series

$$\frac{\partial\Phi}{\partial u} = \sum_{n\in\mathbf{Z}} \frac{\partial\phi}{\partial u}$$

is also seen to be absolutely convergent. However, if one now considers the surface integral over the whole infinite plane at $x = 0$, then term by term integration (up to factor of 2π) yields

$$2(1-x) = \cdots + (-1+1) + (1+1) + (1-1) + \cdots = 2, \tag{12}$$

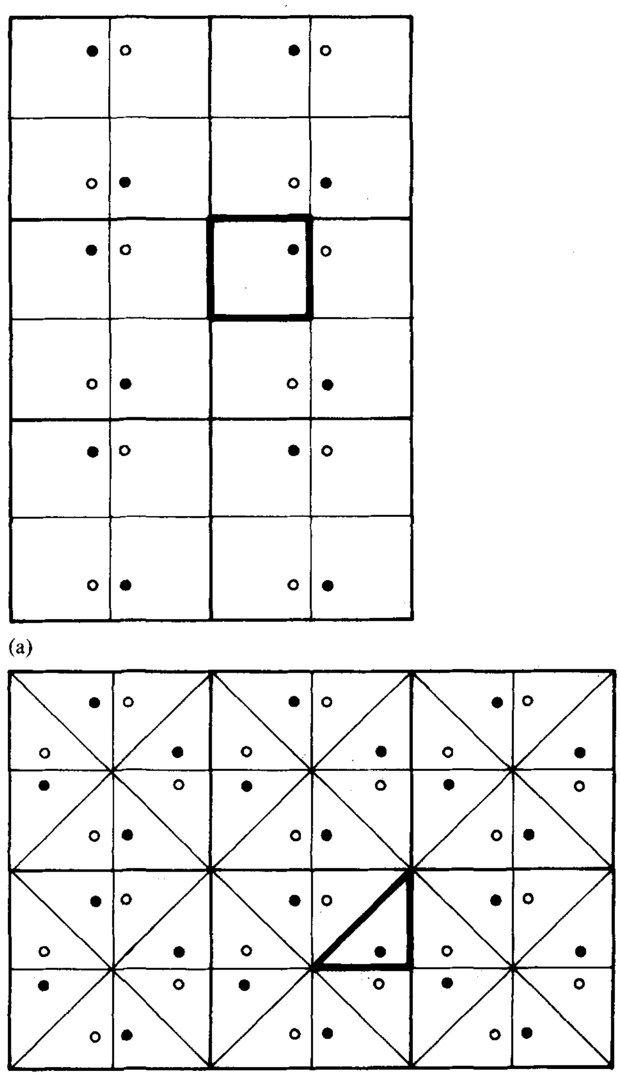
where the left-hand side is obtained from the correct total charge computation of Zahn,[5] Schockley,[10] and Kittel and Fong.[11] Despite assertions to the contrary by Pleines and Mahajan,[12] there exists no physically meaningful rearrangement of parentheses which will make (12) an identity. In this regard we also note the discussion of Epstein and Smith.[13] In Terras[14] the "method of theta functions" is applied to the parallel plate problem.

V. CYLINDERS AND PRISMS

In this section we will derive the space groups for the admissible cylinders and prisms. This computation will demonstrate that these figures are indeed image domains. If $\mathbf{S} = \mathbf{HK}$ is a product decomposition of a point group then

$$\sum_{\sigma\in S} \det(\sigma) f(\sigma\mathbf{x}) = \sum_{\lambda\in\mathbf{H}} \det(\lambda) \sum_{\mu\in\mathbf{K}} \det(\mu) f(\mu\lambda\mathbf{x}). \tag{13}$$

This decomposition of a group sum underlies our formulas for normal modes. Use of this method to construct efficient codings for potentials is illustrated by (31) and (36). It is for these reasons that in our listings we give point groups in product form, but we also identify these groups with the point groups of crystallography in Schoenflies notation.[7]



(a)

(b)

FIG. 6. *Crystal Structure.* (a) A charge in the dark rectangle generates a cluster which is replicated by a rectangular Bravais lattice; (b) a charge in a $(\pi/2|\pi/4|\pi/4)$–triangle generates a cluster which is also replicated by a rectangular Bravais lattice.

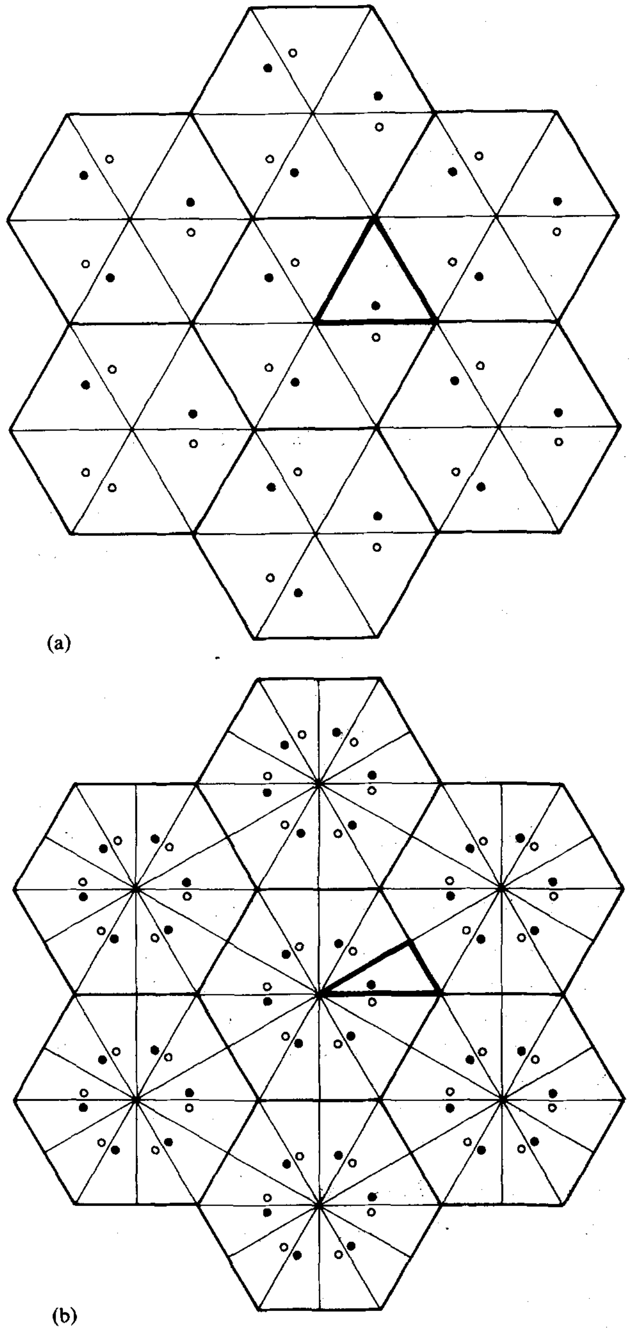


FIG. 7. *Crystal Structure.* (a) A charge in the dark $(\pi/3|\pi/3|\pi/3)$–triangle leads to a hexagonal Bravais lattice; (b) a charge in a $(\pi/2|\pi/3|\pi/6)$–triangle also leads to a hexagonal Bravais lattice.

A. Rectangular and $(\pi/4|\pi/4|\pi/2)$ triangular cross sections

The following notation for matrix generators of reflection groups will be used in this section:

$$\lambda = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mu = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$\eta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \nu = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{14}$$

1. *The open rectangular cylinder* $\mathbf{V} = \{(x,y,z)/0 \leqslant x \leqslant a, 0 \leqslant y \leqslant b\}$.

The bounding planes through the origin determine reflections λ and μ, and generate the point group

$$\mathbf{S} = \{1,\lambda\}\{1,\mu\} = \{1,\lambda,\mu,\lambda\mu\}, \text{(Schoenflies } C_{2v}) \tag{15}$$

which we term the *square corner group*. The cylinder is then seen to have space group (**S**,**L**), where

$$\mathbf{L} = \{(2am,2bn,0)|m,n\in\mathbf{Z}\}, \tag{16}$$

which is a *rectangular* Bravais lattice. The crystal structure generated by this space group is depicted in Fig. 6(a).

In terms of the unrestricted potential, formula (6) translates into the formulas

$$\begin{aligned}\phi(u,v,w) &= \sum_{\sigma\in\mathbf{S}} \det(\sigma)\|\sigma(x,y,z) - (u,v,w)\|^{-1} \\ &= \|(x,y,z) - (u,v,w)\|^{-1} - \|(-x,y,z) - (u,v,w)\|^{-1} \\ &\quad + \|(-x,-y,z) - (u,v,w)\|^{-1} \\ &\quad - \|(x,-y,z) - (u,v,w)\|^{-1}.\end{aligned} \tag{17}$$

It is easy to see that this is the potential due to a quadrupole and goes to zero as $\|(u,v,w)\|^{-3}$. It follows that

$$\Phi(u,v,w) = \sum_{m\in\mathbf{Z}} \sum_{n\in\mathbf{Z}} \phi(2am + u, 2bn + v, w)$$

converges absolutely. Such a verification would have to be made for every explicit realization of (6) in order to substantiate our claim regarding absolute convergence, but we will leave the remaining verifications for the reader. It is worth noting that we had good success in computer evaluation of this and other lattice sums due to additional cancellation which occurs when summation is carried out over blocks of indices that are invariant under the point group action.

We tabularize the remaining space groups with sparse detail. The data presented is needed to parametrize the normal modes discussed in Sec. VIII. The computations need to be performed in order to finish the formal proof that our admissible regions are indeed image domains as defined in Sec. I.

One should note that the domains are all oriented in such a manner that the sharpest corner is at the origin. The generators of the point groups are the reflections in the bounding planes through the origin. The lattice is generated by the adjoining to the point group the remaining reflections in the bounding planes, through the group extension process described in Sec. III. The geometric constructions of image arrays in Figs. 6 and 7 may be used to good advantage in checking our data.

2. *The closed cylinder* $V = \{(x,y,z)/0 \leqslant x \leqslant a, 0 \leqslant y \leqslant b, 0 \leqslant z\}$.

Generators: λ, μ, η;

$$\mathbf{S} = \{1,\lambda\}\{1,\mu\}\{1,\eta\} \quad \text{(Schoenflies } D_{2h}),$$
$$\mathbf{L} = \{(2am,2bn,0)|m,n\in\mathbf{Z}\} \quad \text{[see Fig. 6(a)].} \tag{18}$$

We shall refer to the frequently occurring point group **S** as the *cube corner group*.

3. The rectangular box $V = \{(x,y,z)/0<x<a,\ 0<y<b,\ 0\leqslant z\leqslant c\}$.

Generators: λ, μ, η;

$$\mathbf{S} = \{1,\lambda\}\{1,\mu\}\{1,\eta\},\quad \mathbf{L} = \{(2ak,2bm,2cn)|k,m,n\in\mathbf{Z}\}. \tag{19}$$

4. The $(\pi/2|\pi/4|\pi/4)$ *cylinder* $V = \{(x,y,z)/0\leqslant y\leqslant x\leqslant a\}$.

Generators: μ, ν; Relations: $\lambda = \nu\mu\nu$

$$\mathbf{S} = \{1,\lambda\}\{1,\mu\}\{1,\nu\}\quad (\text{Schoenflies } C_{4v}),\quad \mathbf{L} = \{(2am,2an,0)|m,n\in\mathbf{Z}\}\quad (\text{see Fig. 6(b)}. \tag{20}$$

5. The closed $(\pi/2|\pi/4|\pi/4)$ *cylinder* $V = \{(x,y,z)/0\leqslant y\leqslant x\leqslant a,\ 0\leqslant z\}$.

Generators: μ, ν, η; Relations: $\lambda = \nu\mu\nu$;

$$\mathbf{S} = \{1,\lambda\}\{1,\mu\}\{1,\nu\}\{1,\eta\}\quad (\text{Schoenflies } D_{4h}),\quad \mathbf{L} = \{2am,2an,0)|m,n\in\mathbf{Z}\}. \tag{21}$$

6. The $(\pi/2|\pi/4|\pi/4)$ *prism* $V = \{(x,y,z)/0\leqslant y\leqslant x\leqslant a,\ 0\leqslant z\leqslant c\}$.

Generators: μ, ν, η Relations: $\lambda = \nu\mu\nu$;

$$\mathbf{S} = \{1,\lambda\}\{1,\mu\}\{1,\nu\}\{1,\eta\},\quad \mathbf{L} = \{(2ak,2am,2cn)|k,m,n\in\mathbf{Z}\}. \tag{22}$$

B. Triangular $(\pi/3|\pi/3|\pi/3)$ and $(\pi/2|\pi/3|\pi/6)$ cylinders and prisms

We shall consider the additional matrix generators:

$$\kappa = \begin{bmatrix} 1/2 & \sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}\quad \gamma = \begin{bmatrix} -1/2 & \sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{23}$$

1. The open $\pi/3$ *cylinder* $V = \{(x,y,z)/0\leqslant y/\sqrt{3}\leqslant x,\ y\leqslant(a-x)\sqrt{3}\}$.

Generators: μ, γ;

$$\mathbf{S} = \{1,\mu\gamma,\gamma\mu\}\{1,\mu\}\quad (\text{Schoenflies } C_{3v}),\quad \mathbf{L} = \{(3a(n+m)/2,\sqrt{3}a(n-m)/2,0)|k,m,n\in\mathbf{Z}\}\quad [\text{see Fig. 7(a)}]. \tag{24}$$

2. The closed $\pi/3$ *cylinder* $V = \{(x,y,z)/0\leqslant y/\sqrt{3}\leqslant x,\ y\leqslant(a-x)\sqrt{3},\ 0\leqslant z\}$.

Generators: μ, γ, η;

$$\mathbf{S} = \{1,\mu\gamma,\gamma\mu\}\{1,\mu\}\{1,\eta\}\quad (\text{Schoenflies } D_{3h}),\quad \mathbf{L} = \{(3a(n+m)/2,\sqrt{3}a(n-m)/2,0)|m,n\in\mathbf{Z}\}. \tag{25}$$

3. The prism $V = \{(x,y,z)/0\leqslant y/\sqrt{3}\leqslant x,\ y\leqslant(a-x)\sqrt{3},\ 0\leqslant z\leqslant c\}$.

Generators: μ, γ, η;

$$\mathbf{S} = \{1,\mu\gamma,\gamma\mu\}\{1,\mu\}\{1,\eta\},\quad \mathbf{L} = \{(3a(n+m)/2,\sqrt{3}a(n-m)/2,2kc)|k,m,n\in\mathbf{Z}\}. \tag{26}$$

4. The open $\pi/6$ *cylinder* $V = \{(x,y,z)/0\leqslant\sqrt{3}y\leqslant x,\ y\leqslant(a-x)\sqrt{3}\}$.

Generators: κ, μ; Relations: $\gamma = \kappa\mu\kappa$, $\lambda = \kappa\mu\kappa\mu\kappa$;

$$\mathbf{S} = \{1,\mu\gamma,\gamma\mu\}\{1,\mu\}\{1,\lambda\}\quad (\text{Schoenflies } C_{6v}),\quad \mathbf{L} = \{(3a(n+m)/2,\sqrt{3}a(n-m)/2,0)|m,n\in\mathbf{Z}\}\quad [\text{see Fig. 7(b)}]. \tag{27}$$

5. The closed $\pi/6$ *cylinder* $V = \{(x,y,z)/0\leqslant\sqrt{3}y\leqslant x,\ y\leqslant(a-x)\sqrt{3},\ 0\leqslant z\}$.

Generators: κ, μ, η; Relations: $\gamma = \kappa\mu\kappa$, $\lambda = \kappa\mu\kappa\mu\kappa$;

$$\mathbf{S} = \{1,\mu\gamma,\gamma\mu\}\{1,\mu\}\{1,\lambda\}\{1,\eta\}\quad (\text{Scheonflies } D_{6h}),\quad \mathbf{L} = \{3a(n+m),\sqrt{3}a(n-m)/2,0)|m,n\in\mathbf{Z}\} \tag{28}$$

6. The prism $V = \{(x,y,z)/0\leqslant\sqrt{3}y\leqslant x,\ y\leqslant(a-x)\sqrt{3},\ 0\leqslant z\leqslant c\}$.

Generators: κ, μ, η; Relations: $\gamma = \kappa\mu\kappa$, $\lambda = \kappa\mu\kappa\mu\kappa$;

$$\mathbf{S} = \{1,\mu\gamma,\gamma\mu\}\{1,\mu\}\{1,\lambda\}\{1,\eta\},\quad \mathbf{L} = \{(3a(n+m)/2,\sqrt{3}a(n-m)/2,2kc)|k,m,n\in\mathbf{Z}\}. \tag{29}$$

VI. THE BOUNDED FOUR-SIDED DOMAINS

Analysis of these domanisn requires more use of group theory and less reliance on geometrical construction. These bounded domains are all tetrahedra, but will be distinguished according to the space groups associated with them. Our three tetrahedra will thus be termed the *large tetrahedral domain*, the *centered octahedral* domain, and the *primitive octahedral* domain.

A. The large tetrahedral domain

The representation considered is $\mathbf{V} = \{(x,y,z)|x\leqslant y,\ z\leqslant x,\ -z\leqslant x,\ x+y\leqslant 2a\}$, and is depicted in Fig. 5(c).

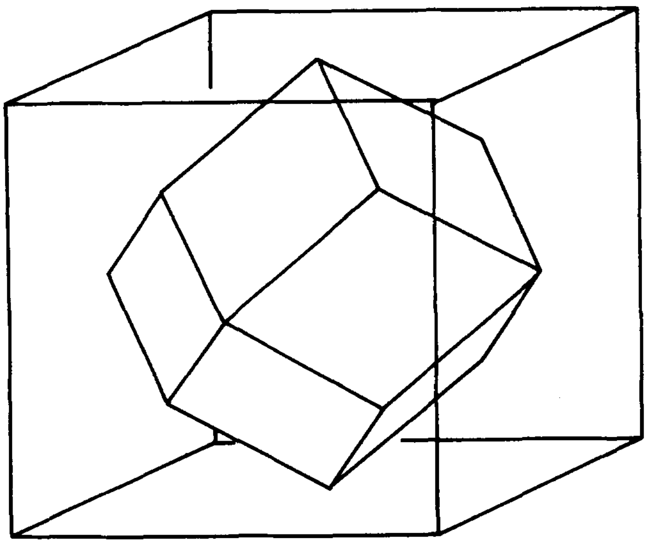
The Möbius corner $(\pi/2,\pi/3,\pi/3)$. This corner generates the *tetrahedral* group. The planes of **V** which pass through the origin generate the reflections

$$\lambda = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},\ \mu = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad \nu = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}. \tag{30}$$

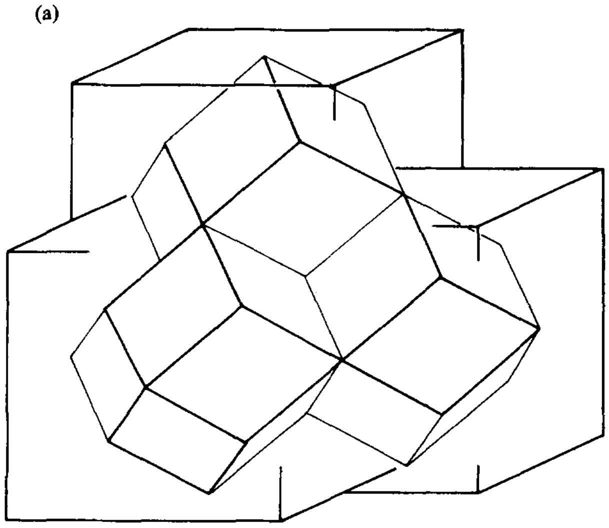
One finds that $\lambda\nu(x,y,z) = (-x,y,-z)$ and $\mu\lambda\nu\mu(x,y,z) = (x,-y,-z)$. The elements $\lambda\nu$ and $\mu\lambda\nu\mu$ generate the diagonal group **D** of order four, the matrices of determinant unity with diagonal entries plus or minus unity. The effect of **D** on (x,y,z) is to generate the images

$$\{(x,y,z),\ (x,-y,-z),\ (-x,-y,z),\ (-x,y,-z)\}.$$

On the other hand, λ and μ generate the permutation



(a)



(b)

FIG. 8. *Symmetric Cell.* (a) Under the action of the tetrahedral group T_h the large tetrahedral domain generates a Wigner–Seitz cell (a rhombic dodecahedron) for the associated cubic F–lattice; (b) the Wigner–Seitz cells stack to fill all space without gaps.

group **P** on three letters. The effect of **P** on (x,y,z) is to generate the images

$\{(x,y,z), (z,y,x), (y,x,z), (z,x,y), (x,z,y), (y,z,x)\}$

The group **D** is normal with respect to **P**. It follows that $\mathbf{S} = \mathbf{DP}$ is a group and has order 24. In Schoenflies notation the group **S** is denoted by $\mathbf{T}_h$.

The corner potential. Use of (13) with $f(x,y,z) = \|(x-u, y-v, z-w)\|^{-1}$ allows one to write the potential (6) with the help of a simple sequence of auxiliary functions. One has

$$g(u,v,w) = f(u,v,w) + f(-u,-v,w) + f(u,-v,-w) + f(-u,v,-w),$$
$$h(u,v,w) = g(u,v,w) + g(v,w,u) + g(w,u,v), \tag{31}$$
$$\phi(u,v,w) = h(u,v,w) - h(v,u,w).$$

In the manipulations of these formulas we made free use of the duality property

$$\sum_{\sigma\in S} \det(\sigma)\|\sigma\mathbf{x} - \mathbf{u}\|^{-1} = \sum_{\sigma\in S} \det(\sigma)\|\mathbf{x} - \sigma\mathbf{u}\|^{-1}.$$

The space group. Reflection across the plane $x + y = 2a$ has the formula $\xi(x,y,z) = (-y,-x,z) + 2a(1,1,0)$. It is this reflection which needs to be adjoined to **S**. One has $(\nu\lambda)\mu(\nu\lambda)(x,y,z) = (-y,-x,z)$ and thus Hypothesis I for space group generation is satisfied. One now needs to examine all the images of $\xi(0,0,0) = 2a(1,1,0)$ under the action of the tetrahedral group. A glance at our decomposition result $\mathbf{S} = \mathbf{DP}$ shows that the listing of the twelve elements $2a(\mp 1, \mp 1, 0)$, $2a(0, \mp 1, \mp 1)$, $2a(\mp 1, 0, \mp 1)$ yields all possible images. Thus Hypothesis II is also satisfied. Let

$$\mathbf{L} = \{2a(i+k, i+j, j+k) | i,j,k\in\mathbf{Z}\}. \tag{32}$$

This is a *face–centered cubic Bravais lattice.* The space group for **V** is thus seen to be (**S**,**L**).

The different images of the domain **V** produced by the action of the tetrahedral group **S** are disjoint and do not overlap except at the boundaries. The solid region represented by the union $\mathbf{V}_S = \cup\{\sigma\mathbf{V}|\sigma\in\mathbf{S}\}$ appears in Fig. 8(a), and is called a *Wigner–Seitz cell* for the cubic F–Lattice **L**. The effect of the lattice **L** is to fill all three dimensional space with copies of this cell is a non-overlapping manner as in Fig. 8(b). One thus deduces that the tetrahedral domain **V** is an image domain.

The potential function. If $\phi(u,v,w)$ is the potential (31) for the corner region, then the lattice sum

$$\Phi(u,v,w) = \sum_{i,j,k\in\mathbf{Z}} \phi((u,v,w) + 2a(i+k, i+j, j+k)) \tag{33}$$

is an absolutely convergent series for the Green's function for the image domain **V**.

B. Centered Octahedral Domain

A representation of this octahedral domain is $\mathbf{V} = \{(x,y,z)|x<y, 0<z<x, x+y<2a\}$. The orientation of this domain is depicted in Fig. 5(b).

The Möbius corner $(\pi/2, \pi/3, \pi/4)$. This corner generates the octahedral group. The planes passing through the origin generate reflections $\lambda(x,y,z) = (z,y,x)$, $\mu(x,y,z) = (y,x,z)$ as in (30), and the reflection $\eta(x,y,z) = (x,y,-z)$, which is not an element of the tetrahedral group. One constructs diagonal elements by noting that $\lambda\eta\lambda(x,y,z) = (-x,y,z)$ and $(\lambda\mu\lambda)\eta(\lambda\mu\lambda)(x,y,z) = (x,-y,z)$. The elements η, $\lambda\eta\lambda$, and $(\lambda\mu\lambda)\eta(\lambda\mu\lambda)$ are mutually orthogonal and generate the cube corner group **D**, which is a diagonal group of order eight. The elements λ and μ generate the same permutation group **P** as occurred in the tetrahedral case. Since **D** is normal with respect to **P** it follows that $\mathbf{S} = \mathbf{DP}$ is a group of order forty-eight.

The space group. One has to add the reflection $\xi(x,y,z) = (-y,-x,z) + 2a(1,1,0)$ to the octahedral group in order to generate the space group. The product decomposition $\mathbf{S} = \mathbf{DP}$ shows that hypothesis I is satisfied. The decomposition also shows that the images of

$\xi(0,0,0) = 2a(1,1,0)$ under the octahedral group are the same as under the tetrahedral group. Thus Hypothesis II is also satisfied and it follows that the space group is given by $(\mathbf{S},\mathbf{L})$, where $\mathbf{S}$ is the currently discussed octahedral group and $\mathbf{L}$ is the previously defined face-centered cubic lattice (32).

The Wigner–Seitz cell $\mathbf{V_S} = \{\sigma\mathbf{V}|\sigma\in\mathbf{S}\}$ is the same as the one that occurred in the tetrahedral case. It follows that $\mathbf{V}$ is an image domain.

C. Primitive Octahedral Domain

This region has the representation $\mathbf{V} = \{(x,y,z)|x\leqslant y, 0\leqslant z\leqslant x, y\leqslant a\}$. The orientation of the figure is as in Fig. 5(a). The Möbius corner at the origin generates the octahedral group. The reflection to adjoin is $\xi(x,y,z) = (x, -y,z) + 2a(0,1,0)$. One sees that Hypothesis I is satisfied. The images of $\xi(0,0,0) = 2a(0,1,0)$ under the octahedral group are just the eight vectors $\{\pm 2a(1,0,0), \pm 2a(0,1,0), \pm 2a(0,0,1)\}$. It follows that Hypothesis II is satisfied. We now define

$$\mathbf{L} = \{2a(i,j,k)|i,j,k\in\mathbf{Z}\}, \qquad (34)$$

which is a *primitive cubic lattice*. The space group for $\mathbf{V}$ is thus $(\mathbf{S},\mathbf{L})$, where $\mathbf{S}$ is the octahedral group, and $\mathbf{L}$ is the currently defined lattice (34).

The Wigner–Seitz cell $\mathbf{V_S} = \cup\{\sigma\mathbf{V}|\sigma\in\mathbf{S}\}$ is just the cube $\{(x,y,z)| -a\leqslant x,y,z\leqslant a\}$ and is formed as a union of domains which do not overlap. Moreover, under the action of the lattice $\mathbf{L}$ all space is filled in a non–overlapping manner with these cubes. It follows that $\mathbf{V}$ is an image domain.

VII. THE ICOSAHEDRAL POTENTIAL

The Möbius corner $(\pi/2,\pi/3,\pi/5)$. The reflection group generated by this corner is called the *icosahedral group* $\mathbf{Y}_h$. This group does not manifest itself as a crystallographic point group. Nevertheless, the unit cell of some intermetallic compounds, such as $MoAl_{12}$ contains an icosahedral structure.[15,16] The group is best known from the theory of regular polyhedra and was discussed by Möbius[17] and more recently by Coxeter.[6]

We shall obtain an explicit representation by exploiting the relation $\cos(\pi/5) = (\sqrt{5}+1)/4$. Set $\tau = (\sqrt{5}+1)/2$. One notes that $\tau^{-1} = (\sqrt{5}-1)/2$. One now considers the sector $\{(x,y,z)|z\leqslant\tau^{-1}x - \tau y, y\geqslant 0, z\geqslant 0\}$. The sector is depicted in Fig. 9. The reflections associated to the corner have matrix representations μ and η as in (14) and

$$\lambda = \frac{1}{2}\begin{bmatrix} \tau & 1 & \tau^{-1} \\ 1 & -\tau^{-1} & -\tau \\ \tau^{-1} & -\tau & 1 \end{bmatrix}.$$

The elements $\mathbf{K} = \{1,\mu\lambda,\lambda\mu,\lambda\mu\lambda\mu,\mu\lambda\mu\lambda\}$ form a cyclic group of order five and represent rotations through 0, $2\pi/5, 8\pi/5, 4\pi/5, 6\pi/5$ radians. The matrix elements for the non-identity rotations are

$$\mu\lambda = \frac{1}{2}\begin{bmatrix} \tau & 1 & \tau^{-1} \\ -1 & \tau^{-1} & \tau \\ \tau^{-1} & -\tau & 1 \end{bmatrix},$$

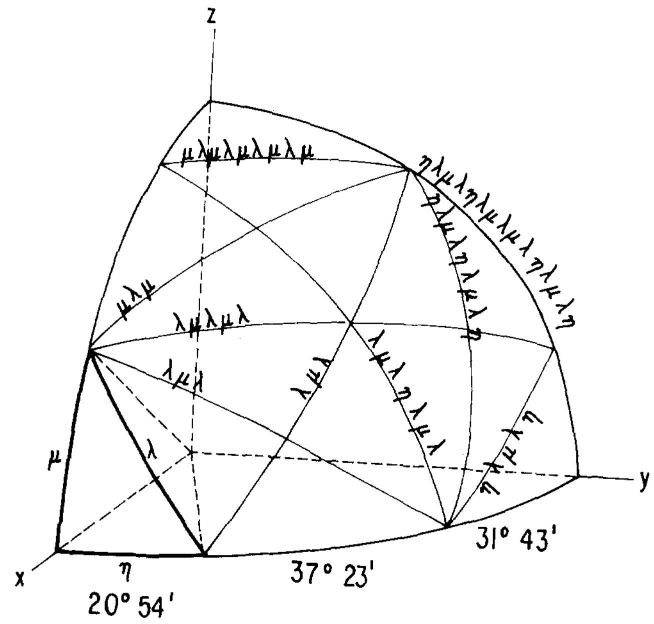


FIG. 9. *Reflection Planes.* Geodesics in the sphere are used to represent the reflection planes in the first octant generated by the icosahedral Möbius corner outlined in the lower left.

$$\lambda\mu = \frac{1}{2}\begin{bmatrix} \tau & -1 & \tau^{-1} \\ 1 & \tau^{-1} & -\tau \\ \tau^{-1} & \tau & 1 \end{bmatrix},$$

$$\mu\lambda\mu\lambda = \frac{1}{2}\begin{bmatrix} 1 & \tau^{-1} & \tau \\ -\tau^{-1} & -\tau & 1 \\ \tau & -1 & -\tau^{-1} \end{bmatrix}, \qquad (35)$$

$$\lambda\mu\lambda\mu = \frac{1}{2}\begin{bmatrix} 1 & -\tau^{-1} & \tau \\ \tau^{-1} & -\tau & -1 \\ \tau & 1 & -\tau^{-1} \end{bmatrix}.$$

A calculation yields $\rho\mu\eta\rho^2(x,y,z) = (-x,y,-z)$, where $\rho = (\lambda\mu\lambda\mu)\eta(\lambda\mu\lambda)(\eta\lambda\eta)$, for which one has $\rho(x,y,z) = (z,x,y)$. The element ρ generates the cyclic group of order three $\mathbf{P} = \{1,\rho,\rho^2\}$. The diagonal elements μ, $\mu\eta$, $\rho\mu\eta\rho^2$ generate the cube corner group $\mathbf{D}$, which has order eight. The product $\mathbf{S} = \mathbf{DPK}$ can be shown to be a group, and is the icosahedral group that we have been seeking to define.

The potential for the icosahedral corner is now easy to obtain by using (13). Let $f(u,v,w) = \|(x-u,y-v,z-w)\|^{-1}$. Define

$$\begin{aligned} g(u,v,w) &= f(u,v,w) + f(-u,-v,w) \\ &\quad + f(u,-v,-w) + f(-u,v,-w), \\ h(u,v,w) &= f(u,v,w) + g(v,w,u) + g(w,u,v), \\ k(u,v,w) &= h(u,v,w) - h(u,-v,w), \\ \phi(u,v,w) &= \sum_{\sigma\in\mathbf{K}} k(\sigma(u,v,w)). \end{aligned} \qquad (36)$$

The formidable problem of coding a matrix sum with 120 terms has thus been reduced to the problem of coding a matrix sum with 5 terms. The necessary elements of $\mathbf{K}$ were listed in (35).

VIII. EIGENFUNCTION EXPANSIONS

In this section we shall discuss normal modes for image domains. We also consider the abstract aspects of Green's function expansions in normal modes for the cylindrical image domains and the bounded image domains. We omit the famous parallel plate problem which was considered by Fong[11] and Jackson.[18]

The principal problem is to parametrize complete systems of normal modes. Because normal modes occur in the description of waveguides and cavity resonators, this aspect of the problem is useful for its own sake. The Dirichlet normal modes for the plane $(\pi/3|\pi/3|\pi/3)$ and $(\pi/2|\pi/3|\pi/6)$ triangles appeared already in the treatise of Lamé,[2] which, however, contained no proof of completeness. After years of controversy the completeness problem was settled by C. G. Nooney.[19] Our own completeness proof is based on group-theoretic methods and is short and novel and applies to the Lamé normal modes as well as our own varieties and is by far the most significant contribution of this section.

Normal Modes. We shall consider image domains in two or three dimensions, but most arguments are formal and apply quite generally. Let **V** be an image domain with space group (**S**,**L**). Let

$$\phi_{\mathbf{v}}(\mathbf{u}) = \sum_{\sigma\in\mathbf{S}} \det(\sigma) e^{-2\pi i\langle \mathbf{v}|\sigma\mathbf{u}\rangle}, \tag{37}$$

where **u** and **v** are elements of the same underlying vector space **E**.

We shall call $\phi_{\mathbf{v}}(\mathbf{u})$ a *Dirichlet normal mode* for **E**.

It is easy to see that

$$\nabla^2\phi_{\mathbf{v}}(\mathbf{u}) = -4\pi^2\langle \mathbf{v}|\mathbf{v}\rangle\phi_{\mathbf{v}}(\mathbf{u}), \tag{38}$$

where ∇^2 is the Laplace operator with respect to **u**. It follows that our terminology is appropriate.

We shall consider the symmetric cell $\mathbf{V}_S = \cup\{\sigma\mathbf{V}|\sigma\in\mathbf{S}\}$. One sees from the preceding sections that in every instance $\mathbf{E} = \{\mathbf{V}_S + \mathbf{n}|\mathbf{n}\in\mathbf{L}\}$ and that this union is formed in a nonoverlapping manner. It follows that an arbitrary function $f(\mathbf{u})$ on **V** can be extended to a function on all space by $f(\sigma\mathbf{u} + \mathbf{n}) = \det(\sigma)f(\mathbf{u})$ for $\mathbf{u}\in\mathbf{V}$. This extension of $f(\mathbf{u})$ to all space is called the *alternating extension*. If $f(\mathbf{u})$ is continuous and vanishes on the boundary of **V** then the alternating extension is a continuous function on all space.

The lattice **L** need not have the same dimension as **E**. We thus consider the space **F** spanned by **L** in **E**. The *dual* of **L** is defined to be a lattice in **F** defined by

$$\mathbf{L}' = \{\mathbf{m}\in\mathbf{F}|\langle\mathbf{m}|\mathbf{n}\rangle\in\mathbf{Z}, \text{ for all } \mathbf{n}\in\mathbf{L}\}. \tag{39}$$

When $\mathbf{n}\in\mathbf{L}$ and $\mathbf{m}\in\mathbf{L}'$ then one has

$$\phi_{(\sigma,\mathbf{n})\mathbf{m}}(\mathbf{u}) = \phi_{\mathbf{m}}((\sigma,\mathbf{n})\mathbf{u}) = \det(\sigma)\phi_{\mathbf{m}}(\mathbf{u}). \tag{40}$$

These relations show that $\phi_{\mathbf{m}}(\mathbf{u}) = 0$ when either **u** or **m** lies on the boundary of **V**.

Any **L**-periodic function $g(\mathbf{u})$ on **F** can be expanded as a Fourier series[20] through

$$g(\mathbf{u}) = \frac{1}{v(\mathbf{R})}\sum_{\mathbf{m}\in\mathbf{L}'} \hat{g}(\mathbf{m}) e^{2\pi i\langle\mathbf{m}|\mathbf{u}\rangle}, \tag{41}$$

where

$$\hat{g}(\mathbf{m}) = \int_{\mathbf{R}} g(\mathbf{u}) e^{-2\pi i\langle\mathbf{m}|\mathbf{u}\rangle}\,d\mathbf{u},$$

and where $v(\mathbf{R})$ is the volume of a primitive cell **R** in **L**, and where the integration **du** is carried out in **F**.

Bounded Image Domains. Let **V** be a bounded image domain with space group (**S**,**L**) and let $g(\mathbf{u})$ be a function which is continuous on **V** and vanishes on the boundary of **V**. If we now consider the alternating extension of $g(\mathbf{u})$ to **E** then (41) yields

$$g(\mathbf{u}) = \frac{1}{n_S v(\mathbf{R})}\sum_{\mathbf{m}\in\mathbf{L}'}\hat{g}(\mathbf{m})\sum_{\sigma\in S}\det(\sigma)e^{2\pi i\langle\mathbf{m}|\sigma\mathbf{u}\rangle},$$

which in terms of (37) becomes

$$g(\mathbf{u}) = \frac{1}{n_S v(\mathbf{R})}\sum_{\mathbf{m}\in\mathbf{L}'} g(\mathbf{m})\phi^*_{\mathbf{m}}(\mathbf{u}). \tag{42}$$

If $\mathbf{m}\in\mathbf{L}'$ then one has

$$\int_{\mathbf{V}} g(\mathbf{u})\phi_{\mathbf{m}}(\mathbf{u})d\mathbf{u} = \int_{\cup\{\sigma\mathbf{V}|\sigma\in\mathbf{S}\}} g(\mathbf{u})e^{-2\pi i\langle\mathbf{m}|\mathbf{u}\rangle}\,d\mathbf{u}$$
$$= \int_{\mathbf{R}} g(\mathbf{u})e^{-2\pi i\langle\mathbf{m}|\mathbf{u}\rangle}\,d\mathbf{u}, \tag{43}$$

where **R** is a primitive cell in **L**. The change in the domain of integration from $\cup\{\sigma\mathbf{V}|\sigma\in\mathbf{S}\}$ is permissible in every one of our explicit constructions.

Define $[\mathbf{n}] = \{\sigma\mathbf{n}|\sigma\in\mathbf{S}\}$. Two classes $[\mathbf{n}]$ and $[\mathbf{m}]$ have either no elements in common or are the same. Let **L(S)** denote a selection of parameter elements in the dual **L**′ obtained by choosing one element from each maximal equivalence class. In actual fact, **L(S)** may be obtained by choosing those elements of **L**′ which lie properly in the interior of the sector that defines **S**, as illustrated in Fig. 10.

For $\mathbf{n}, \mathbf{m}\in\mathbf{L(S)}$ one has that

$$\int_{\mathbf{V}}\phi_{\mathbf{m}}(\mathbf{u})\phi^*_{\mathbf{n}}(\mathbf{u})\,d\mathbf{u} = \begin{cases} 0 & [\mathbf{n}]\neq[\mathbf{m}] \\ v(\mathbf{R}) & [\mathbf{n}] = [\mathbf{m}]. \end{cases} \tag{44}$$

The normal mode expansion is thus summarized by

$$g(\mathbf{u}) = \frac{1}{v(\mathbf{R})}\sum_{\mathbf{m}\in\mathbf{L(S)}}\hat{\mathbf{g}}(\mathbf{m})\phi^*_{\mathbf{m}}(\mathbf{u}),$$

where

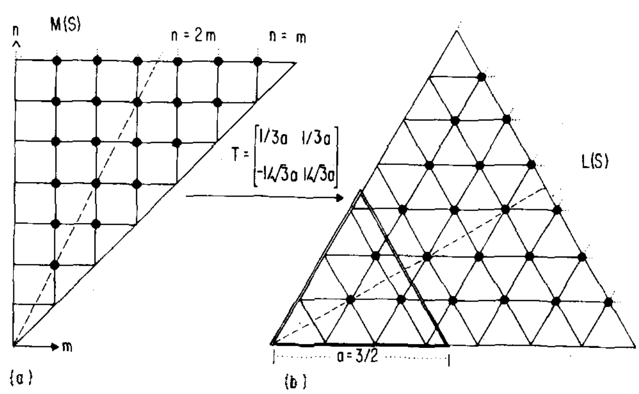


FIG. 10. *Index Sets.* (a) The standard index set, the sector **M(S)**, parametrizies the normal modes, for the $(\pi/3|\pi/3|\pi/3)$–triangle; (b) the natural parameter set **L(S)** is a sector with a simple relationship to the $(\pi/3|\pi/3|\pi/3)$–domain.

$$\hat{g}(\mathbf{m}) = \int_{\mathbf{V}} g(\mathbf{u})\phi_{\mathbf{m}}(\mathbf{u})\,d\mathbf{u}. \quad (45)$$

Since $g(\mathbf{u})$ was an arbitrary function on **V** with the homogeneous boundary condition $g(\mathbf{u}) = 0$, the above demonstrates completeness of the Dirichlet normal modes. Neumann normal modes can be defined by the symmetric sums $\psi_{\mathbf{v}}(\mathbf{u}) = \Sigma_{\sigma\in\mathbf{S}} e^{-2\pi i\langle\mathbf{v}|\sigma\mathbf{u}\rangle}$. A demonstration of completeness can be obtained through the use of symmetric extensions of a function on **V**.

Green's Functions for Bounded Image Domains. Let **V** be a bounded image domain with space group (**S**,**L**). Let $\Phi(\mathbf{u}) = G(\mathbf{x},\mathbf{u})$ be the Green's function for **V**, i.e., the potential for a unit charge at **x**. Computing as in Jackson[18] one obtains the eigenfunction expansion. One has

$$\Phi(\mathbf{u}) = [\pi v(\mathbf{R})]^{-1} \sum_{\mathbf{m}\in\mathbf{L}(\mathbf{S})} \langle\mathbf{m}|\mathbf{m}\rangle^{-1}\phi_{\mathbf{m}}(\mathbf{x})\,\overline{\phi_{\mathbf{m}}(\mathbf{u})}\,. \quad (46)$$

Set $\alpha_{\mathbf{m}}(\mathbf{u}) = [\phi_{\mathbf{m}}(\mathbf{u}) + \phi_{-\mathbf{m}}(\mathbf{u})]/2$ and $\beta_{\mathbf{m}}(\mathbf{u}) = [\phi_{\mathbf{m}}(\mathbf{u}) - \phi_{-\mathbf{m}}(\mathbf{u})]/2i$. One has $\nabla^2\alpha_{\mathbf{m}}(\mathbf{u}) = -4\pi^2\langle\mathbf{m}|\mathbf{m}\rangle\alpha_{\mathbf{m}}(\mathbf{u})$ and $\nabla^2\beta_{\mathbf{m}}(\mathbf{u}) = -4\pi^2\langle\mathbf{m}|\mathbf{m}\rangle\beta_{\mathbf{m}}(\mathbf{u})$. Thus $\alpha_{\mathbf{n}}(\mathbf{u})$ and $\beta_{\mathbf{n}}(\mathbf{u})$ are real eigenfunctions of the Laplace operator, which vanish on the boundary of **V**, and will be called the *real Dirichlet normal modes* for **V**. One sees that

$$\Phi(\mathbf{u}) = [\pi v(\mathbf{R})]^{-1} \sum_{\mathbf{m}\in\mathbf{L}(\mathbf{S})} \frac{[\alpha_{\mathbf{m}}(\mathbf{x})\alpha_{\mathbf{m}}(\mathbf{u}) + \beta_{\mathbf{m}}(\mathbf{x})\beta_{\mathbf{m}}(\mathbf{u})]}{\langle\mathbf{m}|\mathbf{m}\rangle}. \quad (47)$$

It follows that $\{\alpha_{\mathbf{m}}(\mathbf{u})|\mathbf{m}\in\mathbf{L}(\mathbf{S})\}\cup\{\beta_{\mathbf{m}}(\mathbf{u})|\mathbf{m}\in\mathbf{L}(\mathbf{S})\}$ forms a complete set, but it should be pointed out that this listing may contain duplications due to additional relations.

We will now explain why most three-dimensional Dirichlet normal modes are purely imaginary for three-dimensional image domains. One has the relations $\alpha_{\mathbf{m}}(\sigma\mathbf{u}) = \det(\sigma)\alpha_{\mathbf{m}}(\mathbf{u})$ and $\beta_{\mathbf{m}}(\sigma\mathbf{u}) = \det(\sigma)\beta_{\mathbf{m}}(\mathbf{u})$ for $\sigma\in\mathbf{S}$. If inversion is an element of **S**, i.e., if $-\mathbf{n}$ can be obtained from **n** by a group operation in **S**, then $\alpha_{-\mathbf{n}}(\mathbf{u}) = -\alpha_{\mathbf{m}}(\mathbf{u})$. However, the defining condition implies that $\alpha_{\mathbf{m}}(\mathbf{u}) = \alpha_{-\mathbf{m}}(\mathbf{u})$. It follows that $\alpha_{\mathbf{m}}(\mathbf{u}) = 0$ whenever inversion is an element of **S**.

Lattices of the form **Z**, **Z**×**Z**, **Z**×**Z**×**Z**,··· are called *standard lattices*. In practice it is desirable to index the eigenfunctions on a standard lattice. Let **M** be a standard lattice. Let **T** be an invertible linear map which transforms **M** into **L**′. One sets

$$\phi_{\mathbf{T}(\mathbf{m})}(\mathbf{u}) = \alpha_{\mathbf{m}}(\mathbf{u}) + i\alpha_{\mathbf{m}}(\mathbf{u}). \quad (48)$$

Our explicit expansions adhere to this notation. The parameter set for the real and imaginary parts $\alpha_{\mathbf{m}}(\mathbf{u})$ and $\beta_{\mathbf{m}}(\mathbf{u})$ then is $\mathbf{M}(\mathbf{S}) = \{\mathbf{m}\in\mathbf{M}|\mathbf{T}(\mathbf{m})\in\mathbf{L}(\mathbf{S})\}$. The relationship between **M**(**S**) and **L**(**S**) for the $(\pi/3|\pi/3|\pi/3)$ cylinder is illustrated in Fig. 10.

Green's Functions for Cylinder Domains. Let $\{\phi_{\mathbf{a}}(u,v)|\mathbf{a}\in\mathbf{I}\}$ be a complete set of orthonormal Dirichlet normal modes with respect to the Laplace operator $\partial^2/\partial u^2 + \partial^2/\partial v^2$ on a bounded plane domain **D**. With respect to such a system any continuous function $f(u,v)$ on **D** which vanishes on the boundary has the expansion

$$f(u,v) = \sum_{\mathbf{a}\in\mathbf{I}} \hat{f}(\mathbf{a})\phi_{\mathbf{a}}(u,v), \quad (49)$$

where

$$\hat{f}(\mathbf{a}) = \int_{\mathbf{D}} f(u,v)\phi_{\mathbf{a}}(u,v)\,du\,dv.$$

Let $\Phi(u,v,w) = G(x,y,z|u,v,w)$ be the Green's function for a perpendicular cylinder with cross section **D**. Computing as in Jackson[18] one obtains the expansion

$$\Phi(u,v,w) = 2\pi \sum_{\mathbf{a}\in\mathbf{I}} \frac{\exp(-\lambda_{\mathbf{a}}^{1/2}|z-w|)}{\lambda_{\mathbf{a}}^{1/2}}\phi_{\mathbf{a}}(x,y)\phi_{\mathbf{a}}^{*}(u,v), \quad (50)$$

where the eigenvalue $\lambda_{\mathbf{a}}$ corresponds to $\nabla^2\phi_{\mathbf{a}}(u,v) = -\lambda_{\mathbf{a}}\phi_{\mathbf{a}}(u,v)$.

IX. FORMULAS FOR CYLINDERS

The case of the rectangular cylinder is a well-known classical case,[21] but the remaining formulas are expressed in terms of our versions of the Lamé normal modes.

A. The Rectangular Cylinder $\mathbf{V} = \{(x,y,z)/0\leqslant x\leqslant a, 0\leqslant y\leqslant b\}$

$$\Phi(u,v,w) = \frac{1}{2ab}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \times\frac{\exp\{-\pi[(m/a)^2+(n/b)^2]^{1/2}|z-w|\}}{[(n/a)^2+(n/b)^2]^{1/2}} \times\alpha_{m,n}(x,y)\alpha_{m,n}(u,v), \quad (51)$$

where

$$\alpha_{m,n}(u,v) = 4\sin(\pi mu/a)\sin(\pi nv/b). \quad (52)$$

With the parameter set $\mathbf{M}(\mathbf{S}) = \{(m,n)|m\geqslant 1, n\geqslant 1\}$ one has a complete set of Dirichlet normal modes for the square $\{(x,y)|0\leqslant x\leqslant a, 0\leqslant y\leqslant b\}$.

B. The $\pi/4$–Triangular Cylinder $\mathbf{V} = \{(x,y,z)/0\leqslant y\leqslant x\leqslant a\}$

$$\Phi(u,v,w) = \frac{1}{2a}\sum_{n=2}^{\infty}\sum_{m=1}^{n-1} \times\frac{\exp[-\pi a^{-1}(m^2+n^2)^{1/2}|z-w|]}{(m^2+n^2)^{1/2}} \times\alpha_{m,n}(x,y)\alpha_{m,n}(u,v), \quad (53)$$

where

$$\alpha_{m,n}(u,v) = 4\begin{vmatrix}\sin(\pi mu/a) & \sin(\pi mv/a)\\ \sin(\pi nu/a) & \sin(\pi nv/a)\end{vmatrix}. \quad (54)$$

With the parameter set $\mathbf{M}(\mathbf{S}) = \{(m,n)|0<m<n\}$ these Dirichlet normal modes form a complete set.[21]

C. The $\pi/6$–Cylinder $\mathbf{V} = \{(x,y,z)|0\leqslant\sqrt{3}y\leqslant x, y\leqslant(a-x)\sqrt{3}\}$

$$\Phi(u,v,w) = \frac{1}{\sqrt{3}a}\sum_{n=3}^{\infty}\sum_{m=[n/2]+1}^{n-1} \times\frac{\exp[-(4\pi/3a)|z-w|(n^2-mn+m^2)^{1/2}]}{(n^2-mn+m^2)^{1/2}} \times\alpha_{m,n}(x,y)\alpha_{m,n}(u,v), \quad (55)$$

where $[x]$ denotes the greatest integer less than or equal to x, and where

$$\alpha_{m,n}(u,v) = 4[\sin(2\pi(n+m)u/3a)\sin(2\pi(n-m)v/\sqrt{3}a) - \sin(2\pi(2n-m)u/3a)\sin(2\pi mv/\sqrt{3}a) - \sin(2\pi(n-2m)u/3a)\sin(2\pi nv/\sqrt{3}a)]. \quad (56)$$

With the parameter set $\mathbf{M(S)} = \{(m,n)|0<m<n<2m\}$ one has a complete set of Dirichlet normal modes.

D. The $\pi/3$–Cylinder $V = \{(x,y,z)/0 \leqslant y \leqslant \sqrt{3}x, y \leqslant (a-x)\sqrt{3}\}$

$$\Phi(u,v,w) = \frac{1}{\sqrt{3}a}\sum_{n=2}^{\infty}\sum_{m=1}^{n-1} \times \frac{[(-4\pi/3a)|z-w|(n^2-mn+m^2)^{1/2}]}{(n^2-mn+m^2)^{1/2}} \times \phi_{(m,n)}(x,y)\phi^*_{(m,n)}(u,v), \quad (57)$$

where $\phi_{(m,n)} = \alpha_{m,n} + i\beta_{m,n}$, and where

$$\alpha_{m,n}(u,v) = 2[\sin(2\pi(n+m)u/3a)\sin(2\pi(n-m)v/\sqrt{3}a) - \sin(2\pi(2n-m)u/3a)\sin(2\pi mv/\sqrt{3}a) - \sin(2\pi(n-2m)u/3a)\sin(2\pi nv/\sqrt{3}a)], \quad (58)$$

$$\beta_{m,n}(u,v) = 2[\sin(2\pi(n+m)u/3a)\cos(2\pi(n-m)v/\sqrt{3}a) - \sin(2\pi(2n-m)u/3a)\cos(2\pi mv/\sqrt{3}a) - \sin(2\pi(n-2m)u/3a)\cos(2\pi nv/\sqrt{3}a)].$$

With the parameter set $\mathbf{M(S)} = \{(m,n)|0<m<n, m,n\in\mathbf{Z}\}$ the complex functions $\phi_{(m,n)}$ form a complete set of Dirichlet normal modes. The relationship between $\mathbf{L(S)}$ and $\mathbf{M(S)}$ is depicted in Fig. 10, which also allows one to infer the additional relations in the real modes. One may thus write

$$\Phi(u,v,w) = \frac{2}{3a}\sum_{n=1}^{\infty}\frac{1}{n}\exp\left(-\frac{4\pi n}{a\sqrt{3}}|z-w|\right) \times \beta_{n,2n}(x,y)\beta_{n,2n}(u,v) + \frac{2}{\sqrt{3}a}\sum_{n=3}^{\infty}\sum_{m=[n/2]+1}^{n-1} \times \exp\frac{[-(4\pi/3a)|z-w|(n^2-mn+m^2)^{1/2}]}{(n^2-mn+m^2)^{1/2}} \times\Big(\alpha_{m,n}(x,y)\alpha_{m,n}(u,v) + \beta_{m,n}(x,y)\beta_{m,n}(u,v)\Big). \quad (59)$$

X. FORMULAS FOR PRISMS

A. The Rectangular Box $V = \{(x,y,z)/0 \leqslant x \leqslant a, 0 \leqslant y \leqslant b, 0 \leqslant z \leqslant c\}$

$$\Phi(u,v,w) = \frac{1}{2\pi abc}\sum_{k=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \times \frac{\beta_{k,m,n}(x,y,z)\beta_{k,m,n}(u,v,w)}{(k/a)^2+(m/b)^2+(n/c)^2}, \quad (60)$$

where

$$\beta_{k,m,n}(u,v,w) = -8\sin(\pi ku/a)\sin(\pi mv/b)\sin(\pi nw/c). \quad (61)$$

With the parameter set $\mathbf{M(S)} = \{(k,m,n)|k\geqslant 1, m\geqslant 1, n\geqslant 1, k,m,n\in\mathbf{Z}\}$, one has a complete set of Dirichlet normal modes for the box. It should be noted that the expansion reported in Courant and Hilbert[22] is flawed, but the formula is very classical and appears in Jackson.[18]

B. The $\pi/4$ Prism $V = \{(x,y,z)/0 \leqslant y \leqslant x \leqslant a, 0 \leqslant z \leqslant b\}$

$$\Phi(u,v,w) = \frac{1}{2\pi a^2 b}\sum_{n=1}^{\infty}\sum_{k=2}^{\infty}\sum_{m=1}^{k-1} \times \frac{\beta_{k,m,n}(x,y,z)\beta_{k,m,n}(u,v,w)}{(k/a)^2+(m/a)^2+(n/b)^2}, \quad (62)$$

where

$$\beta_{k,m,n}(u,v,w) = -8\begin{vmatrix}\sin(\pi ku/a) & \sin(\pi kv/a)\\ \sin(\pi mu/a) & \sin(\pi mv/a)\end{vmatrix}\sin(\pi nw/b). \quad (63)$$

With the parameter set $\mathbf{M(S)} = \{(k,m,n)|1 \leqslant m \leqslant k-1, 1 \leqslant n\}$ one has a complete listing for these normal modes.

C. The $\pi/6$ Prism $V = \{(x,y,z)/0 \leqslant \sqrt{3}y \leqslant x, y \leqslant (a-x)\sqrt{3}, 0 \leqslant z \leqslant c\}$

$$\Phi(u,v,w) = \frac{1}{12\sqrt{3}\pi a^2 c}\sum_{k=1}^{\infty}\sum_{n=2}^{\infty}\sum_{m=[n/2]+1}^{n=1} \times \frac{\beta_{k,m,n}(x,y,z)\beta_{k,m,n}(u,v,w)}{[(m^2-mn+n^2)/(3a)^2]+[k^2/(4c)^2]}, \quad (64)$$

where $\beta_{k,m,n}(u,v,w) = 2\alpha_{m,n}(u,v)\sin(\pi kw/c)$, with $\alpha_{m,n}(u,v)$ defined by (56). One obtains a complete set of real Dirichlet normal modes with the parameter set $\mathbf{M(S)} = \{(k,m,n)|k\geqslant 1, 0<m<n<2m\}$.

D. The $\pi/3$ Prism $V = \{(x,y,z)/0 \leqslant y \leqslant \sqrt{3}x, y \leqslant (a-x)\sqrt{3}, 0 \leqslant z \leqslant c\}$

$$\Phi(u,v,w) = \frac{1}{12\sqrt{3}\pi a^2 c}\sum_{k=1}^{\infty}\sum_{n=2}^{\infty}\sum_{m=1}^{n-1}\frac{\phi_{(k,m,n)}(x,y,z)\,\phi^*_{(k,m,n)}(u,v,w)}{[n^2-mn+m^2)/(3a)^2]+[k^2/(4c)^2]}$$

$$= \frac{1}{12\sqrt{3}\pi a^2 c}\sum_{k=1}^{\infty}\sum_{m=1}^{\infty}\frac{\alpha_{k,m,2m}(x,y,z)\alpha_{k,m,2m}(u,v,w)}{[n^2/3a^2]+[k^2/(4c)^2]} + \frac{1}{6\sqrt{3}\pi a^2 c}\sum_{k=1}^{\infty}\sum_{n=2}^{\infty}\sum_{m=[n/2]+1}^{n-1} \left[\frac{\alpha_{k,m,n}(x,y,z)\alpha_{k,m,n}(u,v,w)+\beta_{k,m,n}(x,y,z)\beta_{k,m,n}(u,v,w)}{[(n^2-mn+m^2)/(3a)^2]+[k^2/(4c)^2]}\right], \quad (65)$$

where $\phi_{(k,m,n)} = \alpha_{k,m,n} + i\beta_{k,m,n}$ and where $\alpha_{k,m,n}(u,v,w) = 2\beta_{m,n}(u,v)\sin(\pi kw/c)$, and $\beta_{k,m,n}(u,v,w) = 2\alpha_{m,n}(u,v)\times\sin(\pi kw/c)$, with $\alpha_{m,n}$ and $\beta_{m,n}$ as in (58). The parameter set $\mathbf{M(S)} = \{(k,m,n)|k\geqslant 1, 0<m<n\}$ determines a complete orthogonal set of the complex modes.

XI. FORMULAS FOR THE TETRAHEDRA

Our most novel formulas appear in this section. The derivation of the normal modes for these domains makes essential use of our group–theoretic approach. In these cases it is even hard to depict and generate the image crystal structures through geometric constructions.

A. Primitive Octahedral Domain $\mathbf{V} = \{(x,y,z)/x\leqslant y,\ 0\leqslant z\leqslant x,\ y\leqslant a\}$

$$\Phi(u,v,w) = \frac{1}{2\pi a}\sum_{m=3}^{\infty}\sum_{k=2}^{m-1}\sum_{n=1}^{k-1}\frac{\beta_{k,m,n}(x,y,z)\,\beta_{k,m,n}(u,v,w)}{(k^2+m^2+n^2)}, \tag{66}$$

where

$$\beta_{k,m,n})(u,v,w) = -8\begin{vmatrix}\sin(\pi ku/a) & \sin(\pi kv/a) & \sin(\pi kw/a)\\ \sin(\pi mu/a) & \sin(\pi mv/a) & \sin(\pi mw/a)\\ \sin(\pi nu/a) & \sin(\pi nv/a) & \sin(\pi nw/a)\end{vmatrix}. \tag{67}$$

The above Dirichlet normal modes form a complete set with the parameter set $\mathbf{M(S)} = \{(k,m,n)|0<n<k<m\}$.

B. Centered Octahedral Domain $\mathbf{V} = \{(x,y,z)/x\leqslant y,\ 0\leqslant z\leqslant x,\ x+y\leqslant 2a\}$

$$\Phi(u,v,w) = \frac{1}{\pi a}\sum_{k=4}^{\infty}\sum_{n=[(k+3)/2]}^{k-1}\sum_{m=k-n+1}^{n-1}\frac{\beta_{k,m,n}(x,y,z)\,\beta_{k,m,n}(u,v,w)}{3(k^2+m^2+n^2)-2(km+kn+mn)} \tag{68}$$

where

$$\beta_{k,m,n})(u,v,w) = -8\begin{vmatrix}\sin(\pi u(k+m-n)/2a) & \sin(\pi v(k+m-n)/2a) & \sin(\pi w(k+m-n)/2a)\\ \sin(\pi u(k-m+n)/2a) & \sin(\pi v(k-m+n)/2a) & \sin(\pi w(k-m+n)/2a)\\ \sin(\pi u(-k+m+n)/2a) & \sin(\pi v(-k+m+n)/2a) & \sin(\pi w(-k+m+n)/2a)\end{vmatrix}. \tag{69}$$

One obtains a complete orthogonal set with $\mathbf{M(S)} = \{(k,m,n)|1\leqslant m<n<k<m+n\}$.

C. Large Tetrahedral Domain $\mathbf{V} = \{(x,y,z)/x\leqslant y,\ -x\leqslant z\leqslant x,\ x+y\leqslant 2a\}$

$$\begin{aligned}\Phi(u,v,w) &= \frac{1}{\pi a}\sum_{k=3}^{\infty}\sum_{n=2}^{k-1}\sum_{m=1}^{n-1}\frac{\phi_{(k,m,n)}(x,y,z)\,\phi^*_{(k,m,n)}(u,v,w)}{3(k^2+m^2+n^2)-2(km+kn+mn)},\\ &= \frac{1}{4\pi a}\sum_{n=2}^{\infty}\sum_{m=1}^{n-1}\frac{\alpha_{m+n,m,n}(x,y,z)\alpha_{m+n,m,n}(u,v,w)}{m^2+n^2} + \frac{2}{\pi a}\sum_{k=4}^{\infty}\sum_{n=[(k+3)/2]}^{k-1}\sum_{m=k-n+1}^{n-1}\\ &\quad\times\frac{\alpha_{k,m,n}(x,y,z)\alpha_{k,m,n}(u,v,w)+\beta_{k,m,n}(x,y,z)\,\beta_{k,m,n}(u,v,w)}{3(k^2+m^2+n^2)-2(km+kn+mn)},\end{aligned} \tag{70}$$

where

$$\alpha_{k,m,n}(u,v,w) = 4\begin{vmatrix}\cos(\pi u(k+m-n)/2a) & \cos(\pi v(k+m-n)/2a) & \cos(\pi w(k+m-n)/2a)\\ \cos(\pi u(k-m+n)/2a) & \cos(\pi v(k-m+n)/2a) & \cos(\pi w(k-m+n)/2a)\\ \cos(\pi u(-k+m+n)/2a) & \cos(\pi v(-k+m+n)/2a) & \cos(\pi w(-k+m+n)/2a)\end{vmatrix}, \tag{71}$$

and $\beta_{k,m,n}(u,v,w)$ is as in (71), except that cos is replaced by the sin. A complete set of orthogonal complex modes $\phi_{k,m,n} = \alpha_{k,m,n} + i\beta_{k,m,n}$ is obtained with $\mathbf{M(S)} = \{(k,m,n) = |1\leqslant m<n<k\}$.

XII. SUMMARY

We have determined all the domains bounded by linear planes for which the image method determines the solution of the potential problem. In the applications one considers such domains to have conducting walls. The domains are listed in Table I, together with the point group and lattice which determine the crystal structure of the images. The lattice parameters are given in the text, as are the Green's functions associated with each domain. Solutions for the wedge and rectangular prism are in the existing literature.[4,18] For the Green's functions we have numerically evaluated the Coulomb sums as well as the eigenfunction expansions, which required special summation methods, and we found numerical agreement. We found the partial Coulomb sums over invariant sets of indices to be rapidly convergent when three-dimensional lattices were involved, and to be moderately rapidly convergent for two-dimensional lattices, but we have refrained from including our mathematical esti-

TABLE I. Image domains

Domain	Schoenflies corner group	Lattice
Parallel Plates	C_{1v}	(1 dim)
π/n–Wedge	C_{nv}	None
$(\pi/2,\pi/2,\pi/n)$–Corner	D_{nh}	None
$(\pi/2,\pi/3,\pi/3)$–Corner	T_h	None
$(\pi/2,\pi/3,\pi/4)$–Corner	O_h	None
$(\pi/2,\pi/3,\pi/5)$–Corner	Y_h	None
Open Channel	C_{2v}	(1 dim)
Prismatic Wedge	D_{nh}	(1 dim)
Open Rectangular Cylinder	C_{2v}	Primitive rectangular (square if 2 sides equal)
Closed Rectangular Cylinder	D_{2h}	Primitive rectangular (square if 2 sides equal)
Rectangular Prism	D_{2h}	Primitive orthorhombic (tetrahgonal or cubic if equal sides)
Open $\pi/4$–Triangular Cylinder	C_{4v}	Square
Closed $\pi/4$–Triangular Cylinder	D_{4h}	Square
$\pi/4$–Triangular Prism	D_{4h}	Primitive tetragonal (cubic if equal sides)
Open $\pi/3$–Triangular Cylinder	C_{3v}	Hexagonal (2 dim)
Closed $\pi/3$–Triangular Cylinder	D_{3h}	Hexagonal (2 dim)
$\pi/3$–Triangular Prism	D_{3h}	Hexagonal (3 dim)
Open $\pi/6$–Triangular Cylinder	C_{6v}	Hexagonal (2 dim)
Closed $\pi/6$–Triangular Cylinder	D_{6h}	Hexagonal (2 dim)
$\pi/6$–Triangular Prism	D_{6h}	Hexagonal (3 dim)
Tetrahedral Domain	T_h	Face centered cubic
Primitive Octahedral Domain	O_h	Primitive cubic
Centered Octahedral Domain	O_h	Face centered cubic

mates on the rates of convergence of these sums in this paper.

The eigenfunctions that we have displayed also arise in a number of other boundary value problems, such as those connected with cavity resonators and waveguides, but these problems also require the Neumann normal modes, which, as we indicated, can be defined with symmetric sums.

ACKNOWLEDGMENT

We thank the UCSD Academic Senate and the National Science Foundation for partial support of this work.

[1]W. Thomson, Brit. Ass. Rep. **6** (1847).
[2]G. Lamé, *Lecons sur la Theorie Analytique de la Chaleur,* Paris (1861).
[3]R. Terras and R. Swanson, Am. J. Phys. **48**, 526 (1980).
[4]J. H. Jeans, *The Mathematical Theory of Electricity and Magnetism* (Cambridge U. P., Cambridge, 1925).
[5]M. Zahn, Am. J. Phys. **44**, 1132 (1976).
[6]H. S. M. Coxeter *Regular Complex Polytopes* (Cambridge U. P., London, 1974).
[7]G. Burns and A. M. Glazer, *Space Groups for Solid State Scientists* (Academic, New York 1978).
[8]M. Hamermesh, *Group Theory* (Addison–Wesley, Reading, 1964).
[9]N. F. M. Henry and K. Lonsdale, *International Tables for X-Ray Crystalography* I (Kynoch, Birmingham 1969).
[10]W. Schockley, J. Appl. Phys. **9**, 635 (1938).
[11]C. Kittel and C. Y. Fong, Am. J. Phys. **35**, 1091 (1967).
[12]J. Pleines and S. Mahajan, Am. J. Phys. **45**, 868 (1977).
[13]D. J. Epstein, A. C. Smith, Am. J. Phys. **47**, 99 (1979).
[14]R. Terras, J. Comput. Phys. (to appear 1981).
[15]J. Adams, J. B. Rich, Acta. Crystallogr. **7**, 813 (1954).
[16]L. Pauling, *The Nature of the Chemical Bond* (Cornell U. P. Ithaca 1960).
[17]A. F. Möbius, Gesamelte Werke (Hirzel, Leipzig, 1886).
[18]J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
[19]C. G. Nooney, "On the Vibrations of Triangular Membranes," Dissertation, Stanford University (1953).
[20]M. Born, K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon, Oxford, 1956).
[21]P. Morse and H. Feshbach, *Methods of Theoretical Physics* I, II (McGraw–Hill, New York, 1953).
[22]R. Courant, D. Hilbert, *Methoden der Mathematischen Physik* (Springer, Berlin, 1931).

Quantum friction in the c-number picture: The damped harmonic oscillator [a)]

J. R. Brinati
Instituto de Física, UFRJ, Ilha do Fundão, Bloco A, 21910, Rio de Janeiro, RJ., Brazil
S. S. Mizrahi
Instituto de Física Teórica, Rua Pamplona, 145, 01405, São Paulo, SP., Brazil

(Received 25 September 1979; accepted for publication 4 April 1980)

Considering the Lagrangian proposed by Havas, that describes the classical damped motion of a particle, new momentum and position are defined in order to write a Hamiltonian that is subsequently quantized and expressed in terms of non-Hermitian operators. Using the c-number formalism proposed by Lax and Yuen, we associate to the quantum Liouville equation a Fokker–Planck one in terms of c-numbers. From the properties of this equation we obtain the mean values of the position, momentum, and energy of a brownian particle and we also verify the uncertainty principle. We observe that when the system is considered under the Markov hypothesis, the stochastic force is intimately related to the uncertainty principle and to the zero point energy.

INTRODUCTION

The approach to the energy loss problem of a particle in quantum mechanics is divided, essentially, into two lines. The first one is concerned with the brownian motion of a particle embedded in a heat bath (reservoir), considering this bath constituted by N independent harmonic oscillators (HO). The brownian particle is also assumed as an HO which is coupled to the reservoir oscillators. In the limit of $N\rightarrow\infty$ the motion of the particle is described by the Langevin equation or governed by a Fokker–Planck (F–P) equation. In this line we emphasize the works of Ford *et al*,[1] Ullersma,[2] and Louisell,[3] in which the assumption of a Markovian process is essential to guarantee the irreversibility of the phenomenon.

The treatment in the second line consists in writing a one-body phenomenological Hamiltonian describing the quantum-mechanical dissipative process. Some Hamiltonians can be obtained from the Lagrangian proposed by Havas[4]; we can cite Kanai's Hamiltonian,[5] that depends explicitly on time, and the nonlinear Kostin[6] one (this last originally obtained[7] from Heisenberg equations of motion). Another method to construct a Hamiltonian is the one of Albrecht[8] and Hasse,[9] who proposed a new class of nonlinear friction operators under the condition that the correspondence principle be satisfied by requiring the Ehrenfest theorem to hold.

In all these second-line methods the way to treat the quantum friction is to solve the Schrödinger equation and to analyze the wavefunction behavior or to construct wave packets (see, for instance, Refs. 10 and 11, respectively). However, inherent difficulties are present and these are, essentially, the violation of the uncertainty principle in the case of Kanai's Hamiltonian and in the case of Kostin's, the superposition principle; moreover, the Schrödinger equation admits, also, stationary solutions. More details can be found in Ref. 9.

When one tries to write a one-body phenomenological Hamiltonian in order to describe quantum friction, the effects of the medium, responsible for the energy loss, can be simulated by a term due to a force proportional to the velocity of the particle plus a term responsible for the fluctuations (corresponding to a stochastic force), in accordance with the classical Langevin equation.[12] Originally, it was Senitzky[13] who pointed to the importance of the stochastic force in the quantum treatment.

As far as we know, Svin'in[14] was the first author to give a more convenient treatment to the quantum friction with a phenomenological Hamiltonian, Kanai's, in which a stochastic force was included. Then assuming a wave packet as a solution of the Schrödinger equation, he constructed distribution functions for the position and momentum separately, combining two averages, a quantum one and a statistical one; thereby being able to compute quantum statistical mean values. Messer[15] gave the same treatment, but with Kostin's Hamiltonian.

What we propose in this work is another phenomenological treatment, more suited to stochastic problems, making use of the quantum Liouville equation. Here the mean values of the quantities of interest are calculated in a more natural way through a distribution function which is the solution of a F–P equation. This treatment may be considered as the quantum analogue of the classical one developed by Wang and Uhlenbeck[16] for the F–P equation.

The c-number formalism proposed by Lax and Yuen[17] (L–Y) is exploited here, and for the Hamiltonian that we consider it permits us to associate to the quantum Liouville equation an F–P one in c-numbers. We treat the one-particle friction problem under the assumption of a Markov process and later we verify how the stochastic force affects the mean values of the relevant quantities.

In Sec. 1 we obtain the classical Hamiltonian for the motion of a particle, which is subsequently quantized. In Sec. 2 the L–Y method is used to derive a c-number equation associated to the quantum Liouville one, for a more general Hamiltonian. In Sec. 3 we write the F–P equation under the Markov hypothesis, and finally, in Sec. 4 we consider the

[a)]Work supported by FINEP.

relevant mean values such as position, momentum, and energy, and we also verify the uncertainty principle.

1. THE HAMILTONIAN OF THE SYSTEM

The Lagrangian proposed by Havas[4] to describe the motion of a classical particle subject to friction is written, in one dimension, as

$$L(x,\dot{x},t) = (\tfrac{1}{2}m\dot{x}^2 - V(x) + xF(t))e^{\lambda t}. \quad (1.1)$$

The first term corresponds to the kinetic energy, $V(x)$ is the potential energy due to a conservative force, and $F(t)$ is a time-dependent external force; λ is the friction constant. Considering $F(t)$ as a random force, the above Lagrangian yields the Langevin equation for a brownian particle under the condition that the average of $F(t)$ over a statistical ensemble (or over time) be zero

$$\langle F(t)\rangle = 0. \quad (1.2)$$

Defining a new position coordinate

$$X = xe^{\lambda t/2}, \quad (1.3)$$

we are able to write a new Lagrangian in terms of X and $\dot{X}$,

$$\mathscr{L}(X,\dot{X},t) = \tfrac{1}{2}m\dot{X}^2 + \tfrac{1}{8}m\lambda^2X^2 - \tfrac{1}{2}m\lambda X\dot{X} - e^{\lambda t}V(Xe^{-\lambda t/2}) + e^{\lambda t/2}XF(t). \quad (1.4)$$

The canonical momentum is readily obtained

$$P = \frac{\partial\mathscr{L}}{\partial\dot{X}} = m(\dot{X} - \tfrac{1}{2}\lambda X) = pe^{\lambda t/2}, \quad (1.5)$$

where $p = m\dot{x}$ is the mechanical momentum and x is the actual position, while we will refer to X as the virtual position. Therefore, the Hamiltonian of the system is written as

$$H(P,X,t) = P\dot{X} - \mathscr{L}(X,\dot{X},t) = \frac{1}{2m}P^2 + \tfrac{1}{2}\lambda XP + e^{\lambda t}V(Xe^{-\lambda t/2}) - e^{\lambda t/2}XF(t). \quad (1.6)$$

$H(P,X,t)$ does not represent the energy; it is the generator of the motion of an energy-dissipating open system. This dissipation is characterized by the linear term in P, which is also responsible for the violation of the time reversibility. The Hamiltonian can be quantized in the standard procedure under the requirement that it be Hermitian,

$$\hat{H} = \frac{1}{2m}\hat{P}^2 + \frac{\lambda}{4}(\hat{P}\hat{X} + \hat{X}\hat{P}) + e^{\lambda t}V(e^{-\lambda t/2}\hat{X}) - e^{\lambda t/2}\hat{X}F(t), \quad (1.7)$$

with $\hat{P} = -i\hbar(\partial/\partial X)$; furthermore, we consider the force $F(t)$ as a c-number. Assuming an HO potential $V(x) = \tfrac{1}{2}m\omega_0^2x^2$ and writing the Hamiltonian in terms of the non-Hermitian bosonic operators

$$a = (2\hbar\omega_0)^{-1/2}(\omega_0\hat{X} + i\hat{P}),$$
$$a^+ = (2\hbar\omega_0)^{-1/2}(\omega_0\hat{X} - i\hat{P}), \quad (1.8)$$

we obtain

$$\hat{H} = \hbar\omega_0(a^+a + \tfrac{1}{2}) + \tfrac{1}{4}i\hbar\lambda(a^{+2} - a^2) - (\hbar/2\omega_0)^{1/2}e^{\lambda t/2} \times F(t)(a^+ + a), \quad (1.9)$$

where ω_0 is the HO classical frequency. Here we have considered the particle mass $m = 1$.

2. THE c-NUMBER PICTURE

Lax and Yuen developed a formalism in which a c-number is associated to every operator and a c-number function is associated to a function of operators, for a previously chosen ordering of these operators (see also Ref. 3). In this section our purpose is to obtain a c-number equation associated to the quantum Liouville one,

$$i\hbar\frac{\partial\hat{\rho}(t)}{\partial t} = [\hat{H},\hat{\rho}], \quad (2.1)$$

$\hat{\rho}(t)$ being the density operator. We only assume Hamiltonians that can be written in the form of powers of a and a^+,

$$\hat{H} = \sum_{\{r,s\}} h_{rs}^{(n)}(t)a^{+r}a^s, \quad (2.2)$$

where the sum runs over the set of numbers compatible with the nature of $\hat{H}$, the $h_{rs}^{(n)}(t)$'s are the coefficients for a normal ordering, being eventually time-dependent.

First, we are going to associate a c-number $\alpha(\alpha^*)$ to the operator $a(a^+)$, and the distribution function $P(\alpha,\alpha^*,t)$ to the density operator, where by definition,

$$P(\alpha,\alpha^*,t) = \mathrm{Tr}[\hat{\rho}(t)\delta(\alpha^* - a^+)\delta(\alpha - a)], \quad (2.3)$$

with the condition

$$\int P(\alpha,\alpha^*,t)\,d^2\alpha = 1. \quad (2.4)$$

The function $P(\alpha,\alpha^*,t)$ may be interpreted as a classical probability function because we compute quantum statistical mean values as in the classical case, i.e.,

$$\langle a^{+r}a^s\rangle = \langle\alpha^{*r}\alpha^s\rangle = \int\alpha^{*r}\alpha^sP(\alpha,\alpha^*,t)\,d^2\alpha, \quad (2.5)$$

respecting the operator ordering adopted. The δ-functions in (2.3) must be written in the chosen normal order and they are represented by

$$\delta(\alpha^* - a^+)\delta(\alpha - a) = \frac{1}{\pi^2}\int d^2\xi\, e^{-i\xi^*(\alpha^* - a^+)}e^{-i\xi(\alpha - a)}. \quad (2.6)$$

Then we can derive the equation of motion for the distribution function; applying the time derivative operator $i\hbar(\partial/\partial t)$ to both sides of (2.3) and using Liouville equation (2.1) and the representation (2.6) for the δ-functions, we obtain

$$i\hbar\frac{\partial P(\alpha,\alpha^*,t)}{\partial t} = \frac{1}{\pi^2}\int d^2\xi\, e^{-i(\xi^*\alpha^* + \xi\alpha)}I(\xi,\xi^*,t), \quad (2.7)$$

where

$$I(\xi,\xi^*,t) = \mathrm{Tr}\{\hat{\rho}(t)[e^{i\xi^*a^+}e^{i\xi a},\hat{H}]\}. \quad (2.8)$$

In this last expression, substituting the Hamiltonian $\hat{H}$ from (2.2) in the comutator, the normal ordering happens to be destroyed, so the terms in the brackets must be rearranged in order to obtain a normal-ordered expression. Doing this and returning to (2.7), we obtain the following c-number equation associated to the quantum Liouville (2.1)

$$i\hbar \frac{\partial P(\alpha,\alpha^*,t)}{\partial t} = \sum_{\{r,s\}} h_{rs}^{(n)}(t)\Big\{\Big(\alpha^* - \frac{\partial}{\partial \alpha}\Big)^r(\alpha^s P) - \Big(\alpha - \frac{\partial}{\partial \alpha^*}\Big)^s(\alpha^{*r}P)\Big\}, \quad (2.9)$$

which is the equation of motion of $P(\alpha,\alpha^*,t)$ for the Hamiltonian (2.2).

3. THE FOKKER–PLANCK EQUATION

Now, with the c-number picture, we are able to obtain the equation associated to the quantum Liouville one in the case of the Hamiltonian (1.9). Let's, initially, consider this Hamiltonian without the last term (containing the stochatic force), so Eq. (2.9) reduces to

$$\frac{\partial P(\alpha,\alpha^*,t)}{\partial t} = -\frac{\partial}{\partial \alpha}(A_1 P) - \frac{\partial}{\partial \alpha^*}(A_2 P) + \frac{\lambda}{4}\Big(\frac{\partial^2 P}{\partial \alpha^2} + \frac{\partial^2 P}{\partial \alpha^{*2}}\Big), \quad (3.1)$$

with

$$A_1 = -i\omega_0\alpha + \tfrac{1}{2}\lambda\alpha^*, \quad A_2 = A_1^*. \quad (3.2)$$

Equation (3.1) is an (F–P) one, and is characteristic of irreversible processes; it can be written in the more familiar form

$$\frac{\partial P}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}(A_i P) + \sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}(D_{ij}P), \quad (3.3)$$

where $x_1 = \alpha$ and $x_2 = \alpha^*$, the A_i's are the drift coefficients and the D_{ij}'s are the elements of the diffusion matrix

$$\mathbb{D} = \frac{\lambda}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.4)$$

$P(\alpha,\alpha^*,t)$, the distribution function, is determined by solving the differential equation with given initial conditions. However, we can obtain the mean values of the position, momentum, and energy without having to solve the F–P equation, it being sufficient to use its properties, namely, the equations of motion for the first and second moments, $\langle x_i\rangle$ and $\langle x_i x_j\rangle$, respectively

$$\frac{d}{dt}\langle x_i\rangle = \langle A_i\rangle, \quad (3.5)$$

$$\frac{d}{dt}\langle x_i x_j\rangle = \langle x_i A_j\rangle + \langle x_j A_i\rangle + \langle D_{ij} + D_{ji}\rangle. \quad (3.6)$$

Considering the drift coefficients (3.2) and the diffusion matrix (3.4), we have

$$\frac{d}{dt}\langle \alpha\rangle = -i\omega_0\langle\alpha\rangle + \tfrac{1}{2}\lambda\langle\alpha^*\rangle, \quad (3.7)$$

$$\frac{d}{dt}\langle \alpha^2\rangle = \lambda\langle\alpha^*\alpha\rangle - 2i\omega_0\langle\alpha^2\rangle + \tfrac{1}{2}\lambda, \quad (3.8)$$

$$\frac{d}{dt}\langle \alpha^*\alpha\rangle = \tfrac{1}{2}\lambda(\langle\alpha^2\rangle + \langle\alpha^{*2}\rangle), \quad (3.9)$$

and the complex conjugate equations follow immediately.

Considering the complete Hamiltonian (1.9), the function $P(\alpha,\alpha^*,t)$ represents the distribution of α and α^* in a statistical ensemble in which $\langle F(t)\rangle = 0$. The F–P equation will be modified only in the drift coefficients, which will contain an extra term (compare with 3.2),

$$A_1^s = -i\omega_0\alpha + \tfrac{1}{2}\lambda\alpha^* + i(2\hbar\omega_0)^{-1/2}e^{\lambda t/2}F(t), \quad (3.10)$$

$$A_2^s = A_1^{s*}.$$

The inclusion of the stochastic term does not affect the equation of motion of the first moments (3.5) but it modifies the Eqs. (3.6) for the second moments. Now, using the hypothesis of a Markov process, we can introduce the effects of $F(t)$ in the diffusion matrix, modifying it and leaving the drift coefficients free of the last term in (3.10). To achieve this, let's consider the matrix whose elements are $A_{ij} = \langle x_i A_j^s + x_j A_i^s\rangle$ and in which the c-number formalism permits the substitution of $a(t)$ and $a^+(t)$ for α and α^*, respectively, in the mean values; now using Eqs. (A.1) and (A.6) from Appendix A, the matrix $\mathbb{A}$ can be written as

$$\mathbb{A} = \begin{pmatrix} \langle -i2\omega_0\alpha^2 + \lambda\alpha^*\alpha\rangle - \dfrac{d}{\hbar\omega_0}e^{\lambda t} & \dfrac{\lambda}{2}\langle\alpha^2 + \alpha^{*2}\rangle + \dfrac{d}{\hbar\omega_0}e^{\lambda t} \\ \dfrac{\lambda}{2}\langle\alpha^{*2} + \alpha^2\rangle + \dfrac{d}{\hbar\omega_0}e^{\lambda t} & \langle i2\omega_0\alpha^{*2} + \lambda\alpha^*\alpha\rangle - \dfrac{d}{\hbar\omega_0}e^{\lambda t} \end{pmatrix}. \quad (3.11)$$

The comparison between the elements of this matrix with Eqs. (3.8) and (3.9) enables us to write a F–P equation with the drift coefficients given by (3.2) and with a new diffusion matrix containing an additional term (the effects of the fluctuations)

$$\mathbb{D}^s = \frac{\lambda}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{d}{2\hbar\omega_0}e^{\lambda t}\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (3.12)$$

We observe that for the case in which the force is null, the constant d must be equal to zero and the diffusion matrix reduces to the original one (3.4). The constant d will be determined under the requirement that the system be in thermal equilibrium as $t\to\infty$.

4. MEAN VALUES AND UNCERTAINTY PRINCIPLE

The equations of motion (3.5) and (3.6) are more easily solved using the method proposed by Wang and Uhlenbeck[16] (see Appendix B). The solutions of the equations of motion are determined for the underdamped case ($\lambda/2 < \omega_0$); for the first moment we have

$$\langle\alpha\rangle_t = \alpha_0\Big(\cos\omega t - i\frac{\omega_0}{\omega}\sin\omega t\Big) + \frac{\lambda}{2\omega_0}\sin\omega t\,\alpha_0^* = \langle a(t)\rangle, \quad (4.1)$$

and $\langle\alpha^*\rangle_t$ is its complex conjugate; ω is the shifted frequency, $\alpha_0 = \langle\alpha\rangle_{t=0}$, and $\alpha_0^* = \langle\alpha^*\rangle_{t=0}$ are the initial values. For the second moments we obtain

$\langle \alpha^*\alpha\rangle_t$

$$= \langle \alpha^*\alpha\rangle_0\Big[1+\frac{\lambda^2}{4\omega^2}(1-\cos 2\omega t)\Big]+(\langle\alpha^2\rangle_0+\langle\alpha^{*2}\rangle_0)\times\frac{\lambda}{4\omega}\sin 2\omega t - i(\langle\alpha^2\rangle_0-\langle\alpha^{*2}\rangle_0)\frac{\lambda\omega_0}{4\omega^2}(1-\cos 2\omega t)+\frac{d}{\hbar\lambda\omega_0}(e^{\lambda t}-1)+\frac{\lambda^2}{4\omega^2}\times\Big(\frac{1}{2}-\frac{d}{\hbar\omega_0\lambda}\Big)(1-\cos 2\omega t) \tag{4.2}$$

and

$$\langle\alpha^2\rangle_t = \langle\alpha^2\rangle_0\Big[1+i\frac{\omega_0}{\omega}\sin 2\omega t-\frac{1}{2}\Big(1+\frac{\omega_0^2}{\omega^2}\Big)\times(1-\cos 2\omega t)\Big]-\langle\alpha^{*2}\rangle_0\frac{\lambda^2}{8\omega^2}(1-\cos 2\omega t)+\langle\alpha^*\alpha\rangle_0\frac{\lambda}{2\omega}\Big[\sin 2\omega t - i\frac{\omega_0}{\omega}(1-\cos 2\omega t)\Big]-i\frac{\lambda\omega_0}{2\omega^2}\Big(\frac{1}{2}-\frac{d}{\hbar\omega_0\lambda}\Big)(1-\cos 2\omega t)+\frac{\lambda}{2\omega}\Big(\frac{1}{2}-\frac{d}{\hbar\omega_0\lambda}\Big)\sin 2\omega t, \tag{4.3}$$

and $\langle\alpha^{*2}\rangle_t$ is the complex conjugate of (4.3); $\langle\alpha^2\rangle_0$, $\langle\alpha^{*2}\rangle_0$, and $\langle\alpha^*\alpha\rangle_0$ correspond to the initial values. Now, with the expressions for the first and second moments we are able to evaluate the mean values of quantities of physical interest. The mean values for the position and momentum of the brownian particle are obtained from the first moments; from Sec. 1, the actual position and mechanical momentum are respectively $\hat{x}=\hat{X}e^{-\lambda t/2}$ and $\hat{p}=\hat{P}e^{-\lambda t/2}$. Therefore

$$\langle\hat{x}\rangle = \langle\hat{X}\rangle e^{-\lambda t/2} = \langle\alpha+\alpha^*\rangle\Big(\frac{\hbar}{2\omega_0}\Big)^{1/2}e^{-\lambda t/2} = [x_0\cos\omega t+\omega^{-1}(\tfrac{1}{2}\lambda x_0+p_0)\sin\omega t]e^{-\lambda t/2} \tag{4.4}$$

and

$$\langle\hat{p}\rangle = \langle\hat{P}\rangle e^{-\lambda t/2} = \frac{1}{i}\Big(\frac{\hbar\omega_0}{2}\Big)^{1/2}\langle\alpha-\alpha^*\rangle e^{-\lambda t/2} = [p_0\cos\omega t-\omega^{-1}(\tfrac{1}{2}\lambda p_0+\omega_0^2x_0)\sin\omega t]e^{-\lambda t/2}, \tag{4.5}$$

with x_0 and p_0 the initial values of the mean position and momentum. We notice that these mean values evolve in time as the position and momentum of a classical damped harmonic oscillator.

Now we going to compute the mean energy and in order to make the physics involved more transparent, we consider the case of small damping, $\lambda/\omega_0 \ll 1$. The energy operator is given by

$$\hat{E} = \tfrac{1}{2}\hat{p}^2+\tfrac{1}{2}\omega_0^2\hat{x}^2-\hat{x}F(t) = (\tfrac{1}{2}\hat{P}^2+\tfrac{1}{2}\omega_0^2\hat{X}^2)e^{-\lambda t}-\hat{X}F(t)e^{-\lambda t/2} = \hbar\omega_0(a^+a+\tfrac{1}{2})e^{-\lambda t}-\Big(\frac{\hbar}{2\omega_0}\Big)^{1/2}(a^++a)\times F(t)e^{-\lambda t/2}. \tag{4.6}$$

Taking the mean value we can see that the last terms vanish [see expression (A6) in Appendix A], thus

$$\langle\hat{E}(t)\rangle = \hbar\omega_0(\langle\alpha^*\alpha\rangle_t+\tfrac{1}{2})e^{-\lambda t} = E_0e^{-\lambda t}+\frac{d}{\lambda}(1-e^{-\lambda t}), \tag{4.7}$$

where $E_0=\hbar\omega_0(\langle\alpha^*\alpha\rangle_0+\tfrac{1}{2})$ is the initial energy value. Requiring that the particle be in thermal equilibrium with the bath when $t\to\infty$, we must have[12]

$$\lim_{t\to\infty}\langle\hat{E}(t)\rangle = \frac{d}{\lambda} = \frac{\hbar\omega_0}{2}\coth\frac{\hbar\omega_0}{2kT}, \tag{4.8}$$

corresponding to the energy of an HO in thermal equilibrium at temperature T (k is the Boltzmann constant). Substituting the value of d in (4.7) we have

$$\langle\hat{E}(t)\rangle = E_0e^{-\lambda t}+\frac{\hbar\omega_0}{2}\coth\frac{\hbar\omega_0}{2kT}(1-e^{-\lambda t}). \tag{4.9}$$

Furthermore, the determination of the constant d verifies the fluctuation-dissipation theorem,

$$\lambda = (\hbar\omega_0)^{-1}\tanh\Big(\frac{\hbar\omega_0}{2kT}\Big)\int_{-\infty}^{\infty}\langle F(0)F(t)\rangle\,dt, \tag{4.10}$$

where relation (A4) has been used.

The Heisenberg uncertainty principle can also be verified, and this is achieved using the mean values of $\hat{X}^2$ and $\hat{P}^2$,

$$\langle\hat{X}^2(t)\rangle = \frac{\hbar}{\omega_0}[(\langle\alpha^*\alpha\rangle_t+\tfrac{1}{2})+\tfrac{1}{2}(\langle\alpha^2\rangle_t+\langle\alpha^{*2}\rangle_t)], \tag{4.11}$$

$$\langle\hat{P}^2(t)\rangle = \hbar\omega_0[(\langle\alpha^*\alpha\rangle_t+\tfrac{1}{2})-\tfrac{1}{2}(\langle\alpha^2\rangle_t+\langle\alpha^{*2}\rangle_t)], \tag{4.12}$$

together with the expressions (4.4) and (4.5); then we immediately evaluate the product $\Delta x\Delta p$

$$\Delta x\Delta p = [(\langle\hat{P}^2(t)\rangle-\langle\hat{P}(t)\rangle^2)(\langle\hat{X}^2(t)\rangle-\langle\hat{X}(t)\rangle^2)]^{1/2}\times e^{-\lambda t}, \tag{4.13}$$

which in the limit $t\to\infty$ reduces to

$$\Delta x\Delta p \underset{t\to\infty}{\longrightarrow} \frac{\hbar}{2}\coth\frac{\hbar\omega_0}{2kT}, \tag{4.14}$$

and we have $\Delta x\Delta p \geqslant \hbar/2$ at any time and any temperature of the reservoir.

Therefore our treatment describes the behavior of a quantum particle subject to friction, using a convenient formalism in which we obtain an F–P equation in c-number similar to the classical one of Wang and Uhlenbeck in their treatment of the brownian particle. Furthermore, it can clearly be seen that the stochastic force is responsible for the preservation of the quantum effects, namely, the uncertainty principle and the zero point energy. Finally, we would like to mention that the equation of motion (2.9) enables us to consider systems with potentials other than the HO one here considered, that can be written in powers of the operators a and a^+ [see (2.2)].

ACKNOLWEDGMENT

We would like to thank Professor A. H. Zimerman, of Instituto de Física Teórica, for reading the manuscript.

APPENDIX A

The Markov process is characterized by the conditions

$$\langle a(t)F(t')\rangle = 0,$$
$$\langle a^+(t)F(t')\rangle = 0 \qquad \text{(A1)}$$

for $t' > t$, with $a(t)$ expressed in the Heisenberg picture and satisfying the equation of motion

$$\frac{da(t)}{dt} = -i\omega_0 a(t) + \frac{\lambda}{2}a^+(t) + i(2\hbar\omega_0)^{-1/2}e^{\lambda t/2}F(t), \qquad \text{(A2)}$$

obtained from

$$i\hbar\frac{da(t)}{dt} = [a,\hat{H}], \qquad \text{(A3)}$$

with the Hamiltonian (1.9). The equation of motion for $a^+(t)$ is the self-adjoint of (A2). Another condition that charcterizes the Markov process is

$$\langle F(t)F(t')\rangle = 2d\delta(t-t'). \qquad \text{(A4)}$$

Conditions (A1) imply the nonexistence of correlations between the stochastic force and the system operators for $t' > t$, meaning that $F(t)$ does not depend on the behavior of the system for intervals of time $(t'-t) > \mathcal{T}_c$.[18] Equation (A4) is the autocorrelation function of $F(t)$ for a Markov process; the constant d is determined from the equilibrium condition of the system when $t\to\infty$. Now we can estimate the mean value $\langle a(t)F(t)\rangle$, writing

$$\langle a(t)F(t)\rangle \equiv \left\langle\left(a(t_0) + \int_{t_0}^{t}\frac{da(t')}{dt'}dt'\right)F(t)\right\rangle, \qquad \text{(A5)}$$

with the condition that $\mathcal{T}_c \ll t - t_0 \ll \lambda^{-1}$. In these intervals of time the operators remain practically constant while the force undergoes strong variations. From relation (A1) we have $\langle a(t_0)F(t)\rangle = 0$; substituting (A2) in (A5) and making use of relation (A4), we obtain

$$\langle a(t)F(t)\rangle = i(2\hbar\omega_0)^{-1/2}de^{\lambda t/2}; \qquad \text{(A6a)}$$

similarly we have

$$\langle a^+(t)F(t)\rangle = -i(2\hbar\omega_0)^{-1/2}de^{\lambda t/2}. \qquad \text{(A6b)}$$

APPENDIX B

Wang and Uhlenbeck presented a method to evaluate the first and second moments from the F–P equation

$$\frac{\partial P(x_1,x_2,\ldots,t)}{\partial t} = -\sum_i\frac{\partial}{\partial x_i}(A_iP) + \sum_{i,j}\frac{\partial^2}{\partial x_i\partial x_j}(D_{ij}P), \qquad \text{(B1)}$$

when the drift coefficients are of the form

$$A_i = \sum_k c_{ik}x_k. \qquad \text{(B2)}$$

In matrix form it is written as

$$\mathbb{A} = \mathbb{C}\mathbb{X}. \qquad \text{(B3)}$$

Making a change of variables from the x_i's to y_i's

$$\mathbb{Y} = \mathbb{B}\mathbb{X}, \qquad \text{(B4)}$$

the matrix $\mathbb{B}$ being chosen so as to diagonalize $\mathbb{C}$. Under the above transformation the F–P equation takes the form

$$\frac{\partial P(y_1,y_2,\ldots,t)}{\partial t} = -\sum_k \Omega_k\frac{\partial}{\partial y_k}(y_kP) + \sum_{k,l}\frac{\partial^2}{\partial y_k\partial y_l}(\sigma_{kl}P), \qquad \text{(B5)}$$

where the Ω_k's are the eigenvalues of $\mathbb{C}$ and

$$\sigma = \mathbb{B}\mathbb{D}\mathbb{B}^t \qquad \text{(B6)}$$

(the index t means transpose). The equations of motion for the first and second moments are therefore

$$\frac{d}{dt}\langle y_i\rangle = \Omega_i, \qquad \text{(B7)}$$

$$\frac{d}{dt}\langle y_i y_j\rangle = (\Omega_i + \Omega_j)\langle y_i y_j\rangle + \langle\sigma_{ij} + \sigma_{ji}\rangle. \qquad \text{(B8)}$$

Considering the F–P equation for our specific problem, the matrices $\mathbb{C}$ and $\mathbb{B}$ are

$$\mathbb{C} = \begin{pmatrix} -i\omega_0 & \lambda/2 \\ \lambda/2 & i\omega_0 \end{pmatrix}, \qquad \text{(B9)}$$

$$\mathbb{B} = \begin{pmatrix} -\dfrac{\omega_0-\omega}{\lambda\omega} & -\dfrac{i}{2\omega} \\ \dfrac{\omega_0+\omega}{\lambda\omega} & \dfrac{i}{2\omega} \end{pmatrix}, \qquad \text{(B10)}$$

and $\Omega_1 = i\omega$, $(\Omega_2 = \Omega_1^*)$. $\omega = (\omega_0^2 - \frac{1}{4}\lambda^2)^{1/2}$ is the shifted frequency for the underdamped case, $\lambda/2 < \omega_0$. Once the mean values $\langle y_i\rangle_t$ and $\langle y_i y_j\rangle_t$ have been determined by solving Eqs. (B7) and (B8), the mean values $\langle\alpha\rangle_t$, $\langle\alpha^2\rangle_t$ and $\langle\alpha^*\alpha\rangle_t$ are obtained by the inverse transformation of (B4) (with $x_1 = \alpha$ and $x_2 = \alpha^*$).

[1]G. W. Ford, M. Kac, and P. Mazur, J. Math. Phys. 6, 504 (1965).
[2]P. Ullersma, Physica, 32, 27 (1966).
[3]W. H. Louisell, *Quantum Statistical Properties of Radiation* (Wiley, New York, 1973).
[4]P. Havas, Nuovo Cimento Suppl. 5, 363 (1957).
[5]E. Kanai, Prog. Theor. Phys. 3, 440 (1948).
[6]M. Razavy, Can. J. Chem. Phys. 56, 311 (1978).
[7]M. D. Kostin, J. Chem. Phys. 57, 3589 (1972).
[8]K. Albrecht, Phys. Lett. B 56, 127 (1975).
[9]R. Hasse, J. Math. Phys. 16, 2005 (1975).
[10]E. H. Kerner, Can. J. Phys. 36, 371 (1958).
[11]K. K. Kan and J. J. Griffin, Phys. Lett. B 50, 241 (1974).
[12]F. Reif, *Fundamental of Statistical and Thermal Physics* (McGraw–Hill, New York, 1965).
[13]I. R. Senitzky, Phys. Rev. 124, 642 (1961).
[14]I. R. Svin'in, Theo. Mat. Fiz. 27, 270 (1976).
[15]J. Messer, Acta Phys. Austriaca 50, 75 (1979).
[16]M. C. Wang and G. E. Uhlenbeck, Phys. Rev. 36, 323 (1945); also *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York, 1954).
[17]M. Lax and H. Yuen, Phys. Rev. 172, 362 (1968).

[18]$\mathcal{T}_c$ is the correlation time; it is of the order of the mean period of the fluctuations of $F(t)$ and is very small in macroscopic scale, $\mathcal{T}_c \ll \lambda^{-1}$. Here, the δ-function of (A4) gives the drastic situation $\mathcal{T}_c = 0$.

Transition amplitudes for time dependent harmonic oscillators

L. F. Landovitz
Graduate Center of the City University of New York, New York, New York 10036
A. M. Levine and W. M. Schreiber
The College of Staten Island of the City University of New York, New York, New York 10301

(Received 16 October 1979; accepted for publication 29 January 1980)

Utilizing the Green's function for a time dependent harmonic oscillator, we calculate the corresponding transition amplitudes. Particular examples of damped and runaway oscillators are discussed.

I. INTRODUCTION

In a previous paper,[1] the exact quantum mechanical Green's function for an arbitrary time dependent harmonic oscillator was derived. Equivalent results have been obtained by other authors using different techniques.[2] In this paper, we use the Green's function to compute transition amplitudes. The derivation makes no assumption about the details of the amplitude or on the frequency spectrum of the time-dependent parameters. Therefore, the solution can be used for a wide variety of different problems: exponential coefficients, periodic coefficients, and/or stochastically varying coefficients. As an example of the method, we calculate transition amplitudes for exponential coefficients. The simplest case is the "damped" harmonic oscillator as originally described by Kanai.[3] There are some papers that claim that the Kanai model is inappropriate[4] and suggest using an approximate solution to a more complicated model describing the coupling between a lossless oscillator and the loss mechanisms. It is our view that there is validity in obtaining an exact solution to a simple model rather than an approximate solution to a more complicated model. The resolution of this question requires computation of physically interesting parameters such as transition amplitudes. This paper is one step in that direction.

II. CALCULATION OF THE TRANSITION AMPLITUDE

The Hamiltonian H is expressed as[1]

$$H = f(t)\frac{p^2}{2M} + g(t)\tfrac{1}{2}M\omega_0^2x^2. \tag{2.1}$$

We set

$$f(0) = g(0) = 1. \tag{2.2}$$

The corresponding solutions for the operators $x_+(t)$ and $p_+(t)$ are expressible as

$$x_+ = a(t)x + b(t)p, \tag{2.3}$$

$$p_+ = c(t)x + d(t)p, \tag{2.4}$$

where

$$a(0) = d(0) = 1, \tag{2.5}$$

$$b(0) = c(0) = 0. \tag{2.6}$$

The Hamilton equations imply

$$\dot{a} = \frac{f}{M}c, \tag{2.7}$$

$$\dot{b} = \frac{f}{M}d, \tag{2.8}$$

$$\dot{c} = -gM\omega_0^2a, \tag{2.9}$$

$$\dot{d} = -gM\omega_0^2b. \tag{2.10}$$

The condition

$$[x_+, p_+] = [x, p], \tag{2.11}$$

requires

$$ad - bc = 1. \tag{2.12}$$

From Eqs. (2.5)–(2.6), Eq. (2.12) is satisfied at $t = 0$. Also, from Eqs. (2.7)–(2.10), one obtains

$$\frac{d}{dt}(ad - bc) = 0, \tag{2.13}$$

and hence Eq. (2.12) is satisfied for all times.

The corresponding Green's function obtained is

$$G(x,x';t) = \left(\frac{\beta}{i\pi}\right)^{1/2}\exp[i\beta(dx^2 + ax'^2 - 2xx')], \tag{2.14}$$

where

$$\beta = \frac{1}{2\hbar b}. \tag{2.15}$$

The wavefunction $\psi(x,t)$ is obtainable from $\psi(x',0)$ by the expression

$$\psi(x,t) = \int_{-\infty}^{\infty} G(x,x';t)\psi(x',0)\,dx'. \tag{2.16}$$

Utilizing the above Green's function, we compute the transition amplitude a_{nm} for an oscillator from a state $|m\rangle$ to a state $|n\rangle$, where $|m\rangle$ and $|n\rangle$ are the usual harmonic oscillator eigenstates. Thus,

$$a_{nm} = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx'\, u_n(x)G(x,x';t)u_m(x'), \tag{2.17}$$

$$u_n(x) = \left(\frac{\alpha}{\pi^{1/2}2^n n!}\right)^{1/2} H_n(\alpha x)e^{-(1/2)\alpha^2x^2}, \tag{2.18}$$

$$\alpha = \sqrt{M\omega_0/\hbar}. \tag{2.19}$$

In Eq. (2.18), $H_n(\alpha x)$ is the nth order Hermite polynomial.

We define the generating function $A_{ss'}$ as

$$A_{ss'} = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx'\, S(\alpha x,s)e^{-(1/2)\alpha^2x^2}$$

$$\times G(x,x';t) S(\alpha x',s') e^{-(1/2)\alpha^2 x'^2}$$

$$= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \sum_{n=0}^{\infty} \frac{H_n(\alpha x)}{n!} s^n$$

$$\times e^{-(1/2)\alpha^2 x^2} G(x,x';t) \sum_{m=0}^{\infty} \frac{H_m(\alpha x')}{m!} s'^m e^{-(1/2)\alpha^2 x'^2}, \tag{2.20}$$

where

$$S(z,s) = e^{-s^2+2sz} = \sum_{n=0}^{\infty} \frac{H_n(z)}{n!} s^n. \tag{2.21}$$

Hence,

$$a_{nm} = \frac{\alpha}{\sqrt{\pi}} (2^n 2^m n! m!)^{-1/2} \left[\frac{\partial^{n+m}}{\partial s^n \partial s'^m} A_{ss'} \right]_{\substack{s'=0\\ s=0}}. \tag{2.22}$$

Substituting the expression (2.14) for $G(x,x';t)$,

$$A_{ss'} = \left(\frac{\beta}{i\pi} \right)^{1/2} e^{-(s^2+s'^2)} B_{ss'} \tag{2.23}$$

$$B_{ss'} = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' e^{2\alpha(sx+s'x')}$$

$$\times \exp(-1) \left[\left(\frac{\alpha^2}{2} - i\beta d \right) x^2 + 2i\beta x x' + \left(\frac{\alpha^2}{2} - i\beta a \right) x'^2 \right]. \tag{2.24}$$

Let

$$x = qy + ry', \tag{2.25}$$

$$x' = -ry + qy', \tag{2.26}$$

and set

$$q = \sin\theta, \tag{2.27}$$

$$r = \cos\theta, \tag{2.28}$$

$$\frac{1}{d-a} = -\tfrac{1}{2} \tan 2\theta. \tag{2.29}$$

$B_{ss'}$ is then expressed by

$$B_{ss'} = e^{(\alpha^2/\gamma_1)(qs-rs')^2} e^{(\alpha^2/\gamma_2)(rs+qs')^2}$$

$$\times \int_{-\infty}^{\infty} dy\, e^{-\gamma_1 [y-(\alpha/\gamma_1)(qs-rs')]^2}$$

$$\times \int_{-\infty}^{\infty} dy'\, e^{-\gamma_2 [y'-(\alpha/\gamma_2)(rs+qs')]^2}$$

$$= \frac{\pi}{\sqrt{\gamma_1\gamma_2}} \exp\left\{ \frac{\alpha^2}{\gamma_1\gamma_2} [(q^2\gamma_2 + r^2\gamma_1)s^2 - 2qr(\gamma_2-\gamma_1)ss' + (r^2\gamma_2 + q^2\gamma_1)s'^2] \right\}, \tag{2.30}$$

where

$$\gamma_1 = \alpha^2/2 - i\beta(dq^2 + ar^2 + 2qr), \tag{2.31}$$

$$\gamma_2 = \alpha^2/2 - i\beta(dr^2 + aq^2 - 2qr). \tag{2.32}$$

Using the above definitions for the respective parameters and substituting in Eq. (2.23), one obtains

$$A_{ss'} = \frac{1}{\alpha} \left(\frac{2\sigma\pi}{i\lambda} \right)^{1/2} T_{ss'}. \tag{2.33}$$

In the above expression,

$$T_{ss'} = \exp[(1/\lambda)(\mu^* s^2 - 4i\sigma ss' + \mu s'^2)], \tag{2.34}$$

$$\sigma = 1/M\omega_0 b, \tag{2.35}$$

$$\mu = 1 + i\sigma(a-d) - \sigma^2(1-ad), \tag{2.36}$$

$$\lambda = 1 - i\sigma(a+d) + \sigma^2(1-ad). \tag{2.37}$$

From Eq. (2.22),

$$a_{nm} = \left(\frac{2\sigma n! m!}{i\lambda\, 2^n 2^m} \right)^{1/2} t_{nm}, \tag{2.38}$$

where

$$t_{nm} = \frac{1}{n! m!} \left[\frac{\partial^{n+m}}{\partial s^n \partial s'^m} T_{ss'} \right]_{\substack{s'=0\\ s=0}}. \tag{2.39}$$

Using Eq. (2.21), we express $T_{ss'}$ as

$$T_{ss'} = \sum_{l=0}^{\infty} \frac{1}{l!} \left(-i\sqrt{\frac{\mu}{\lambda}} \right)^l s'^l H_l\left(\frac{2\sigma}{\sqrt{\lambda\mu}} s \right) e^{(\mu^*/\lambda)s^2}. \tag{2.40}$$

Defining

$$T_{ms} = \frac{1}{m!} \left[\frac{\partial^m}{\partial s'^m} T_{ss'} \right]_{s'=0}, \tag{2.41}$$

one obtains from Eq. (2.40)

$$T_{ms} = \frac{1}{m!} \left(-i\sqrt{\frac{\mu}{\lambda}} \right)^m H_m\left(\frac{2\sigma}{\sqrt{\lambda\mu}} s \right) e^{(\mu^*/\lambda)s^2}. \tag{2.42}$$

Equation (2.39) implies

$$t_{nm} = \frac{1}{n!} \left[\frac{\partial^n}{\partial s^n} T_{ms} \right]_{s=0}. \tag{2.43}$$

Hence,

$$t_{nm} = \frac{1}{n! m!} \left(-i\sqrt{\frac{\mu}{\lambda}} \right)^m \frac{\partial^n}{\partial s^n} \times \left[H_m\left(\frac{2\sigma}{\sqrt{\lambda\mu}} s \right) e^{(\mu^*/\lambda)s^2} \right]_{s=0}. \tag{2.44}$$

The differentiation in Eq. (2.44) is facilitated by the identity

$$\frac{\partial^n}{\partial s^n}(FG) = \sum_{l=0}^{n} C_l^n \frac{\partial^l F}{\partial s^l} \frac{\partial^{n-l} G}{\partial s^{n-l}}, \tag{2.45}$$

where C_l^n is the binomial coefficient, i.e.,

$$C_l^n = \frac{1}{l!} \frac{n!}{(n-l)!}. \tag{2.46}$$

Also,

$$\left[\frac{d^p}{dx^p} e^{\zeta x^2} \right]_{x=0} = \begin{cases} (\zeta)^{p/2} \dfrac{p!}{(p/2)!} & \text{even } p, \\ 0 & \text{odd } p. \end{cases} \tag{2.47}$$

The Hermite polynomials are expressible as[5]

$$H_n(x) = (-1)^{n/2} \frac{n!}{(n/2)!} v_1(x) \quad \text{even } n,$$

$$= (-1)^{(n-1)/2} \frac{2(n!)}{[(n-1)/2]!} v_2(x) \quad \text{odd } n, \tag{2.48}$$

where

$$v_1(x) = 1 + \sum_{p=1}^{n/2} (-1)^p \frac{(2)^p}{(2p)!} \frac{n!!}{(n-2p)!!} x^{2p}, \tag{2.49}$$

$$v_2(x) = x + \sum_{p=1}^{(n-1)/2} (-1)^p \frac{(2)^p}{(2p+1)!} \times \frac{(n-1)!!}{(n-1-2p)!!} x^{2p+1}. \tag{2.50}$$

From Eqs. (2.40)–(2.50), one obtains for even n

$$\left[\frac{d^q}{dx^q}H_n(x)\right]_{x=0} = (-1)^{(n+q)/2}(2)^{q/2}\frac{n!}{(n/2)!}\frac{n!!}{(n-q)!!} \quad q\leqslant n \text{ and even},$$

$$= 0 \text{ otherwise,} \tag{2.51}$$

and for odd n

$$\left[\frac{d^q}{dx^q}H_n(x)\right]_{x=0} = -(-1)^{(n+q)/2}(2)^{(q+1)/2}\frac{n!}{[(n-1)/2]!} \times \frac{(n-1)!!}{(n-q)!!} \quad q\leqslant n \text{ and odd},$$

$$= 0 \text{ otherwise.} \tag{2.52}$$

After performing the indicated differentiation in Eq. (2.44) and substituting the resulting t_{nm} in Eq. (2.38), one obtains

$$a_{nm} = \begin{cases} (-1)^m\left(\dfrac{2\sigma}{i\lambda}\right)^{1/2}\dfrac{(m!n!)^{1/2}}{(2)^{(m+n)/2}}\left(\dfrac{\mu}{\lambda}\right)^{m/2}\left(\dfrac{\mu^*}{\lambda}\right)^{n/2} \\ \times\frac{1}{2}\displaystyle\sum_{l=0}^{\min(m,n)}[1+(-1)^{m+l}](-1)^{l/2}(2)^{2l}\frac{1}{l![(m-l)/2]![(n-l)/2]!}\left(\frac{\sigma}{|\mu|}\right)^l & \text{even } |m-n|, \\ 0 \quad \text{odd } |m-n|. \end{cases} \tag{2.53}$$

The transition probability P_{nm} is expressed by

$$P_{nm} = |a_{nm}|^2. \tag{2.54}$$

Hence from Eq. (2.53),

$$P_{nm} = \begin{cases} (-1)^m\dfrac{m!n!}{(2)^{m+n}}\left|\dfrac{2\sigma}{\lambda}\right|\left|\dfrac{\mu}{\lambda}\right|^{m+n}\left(\dfrac{1}{4}\right)\displaystyle\sum_{l,l'=0}^{\min(m,n)}[1+(-1)^{m+l}][1+(-1)^{m+l'}] \\ \times(-1)^{(l+l')/2}(2)^{2(l+l')}\dfrac{1}{l!l'![(m-l)/2]![(n-l)/2]![(m-l')/2]![(n-l')/2]!}\left(\dfrac{\sigma}{|\mu|}\right)^{l+l'} & \text{even } |m-n|, \\ 0 \quad \text{odd } |m-n|. \end{cases} \tag{2.55}$$

Since $u_n(x)$ is of even parity for even n and of odd parity for odd n, the fact that the transition amplitudes are nonzero only for even $|m-n|$ corresponds to conservation of parity. Also, Eqs. (2.53) and (2.55) imply that only even (l,l')'s in the sum contribute to the (m,n) even case and odd (l,l')'s to the $(m.n)$ odd case.

If one substitutes the identity[5]

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2} \tag{2.56}$$

into Eq. (2.44), one can obtain the following equivalent expressions for a_{nm} and P_{nm}:

$$a_{nm} = \begin{cases} (-1)^n(i)^{m-1/2}\left(\dfrac{n!}{m!}\right)^{1/2}\left(\dfrac{2\sigma}{\lambda}\right)^{n+1/2}\left|\dfrac{\mu}{\lambda}\right|^{(m-n)/2} \\ \times z^{-n}(1-z^2)^{(n-m)/4}P_{(m+n)/2}^{(m-n)/2}(z) \quad \text{even } |m-n|, \\ 0 \quad \text{odd } |m-n|. \end{cases} \tag{2.57}$$

$$P_{nm} = \begin{cases} \dfrac{n!}{m!}(2)^{2n+1}\left|\dfrac{\sigma}{\lambda}\right|^{2n+1}\left|\dfrac{\mu}{\lambda}\right|^{m-n}z^{-2n}(1-z^2)^{(n-m)/2}[P_{(m+n)/2}^{(m-n)/2}(z)]^2 \quad \text{even } |m-n|, \\ 0 \quad \text{odd } |m-n|. \end{cases} \tag{2.58}$$

In the above equations,

$$z = \left(1+\frac{1}{4}\left|\frac{\mu}{\sigma}\right|^2\right)^{-1/2}, \tag{2.59}$$

and $P_{(m+n)/2}^{(m-n)/2}(z)$ is an associated legendre function of the first kind.[6] $P_p^q(z)$ has the series representation

$$P_p^q(z) = (-1)^q\frac{(1-z^2)}{(2)^p}\sum_{k=0}^{[(p-q)/2]}\frac{(-1)^k(2p-2k)!}{k!(p-k)!(p-q-2k)!}z^{p-q-2k}. \tag{2.60}$$

III. PARTICULAR EXAMPLES

We consider particular exponential choices of $f(t)$ and $g(t)$ in the Hamiltonian (2.1). The corresponding expressions for $a(t)$, $b(t)$, $c(t)$, and $d(t)$ have been obtained previously.[1]

The first case corresponds to

$$f(t) = g(t) = e^{-t/\tau}. \tag{3.1}$$

This corresponds to a decaying oscillator whose ω is also decaying with time according to the prescription

$$\omega = e^{-t/\tau}\omega_0. \tag{3.2}$$

The results obtained were $\bar{a}(\theta)$, $\bar{b}(\theta)$, $\bar{c}(\theta)$, and $\bar{d}(\theta)$. The bar denotes the corresponding time-independent harmonic oscillator expressions and t is replaced by θ, where

$$\theta = \tau(1 - e^{-t/\tau}). \tag{3.3}$$

Substitution of these expressions into Eq. (2.55) yields

$$P_{nm} = \delta_{nm}, \tag{3.4}$$

similar to the corresponding time-independent harmonic oscillator expression. The runaway oscillator expressions correspond to positive exponents in Eqs. (3.1)–(3.3) and yield identical results for P_{nm}.

The second case corresponds to

$$f(t) = e^{-t/\tau}, \tag{3.5}$$

$$g(t) = e^{t/\tau}. \tag{3.6}$$

and is the Kanai Hamiltonian.[3] The corresponding expressions for $a(t)$, $b(t)$, $c(t)$, and $d(t)$ are

$$a(t) = e^{-t/2\tau}[\cos\omega t + (1/2\omega\tau)\sin\omega t], \tag{3.7}$$

$$b(t) = e^{-t/2\tau}(1/M\omega)\sin\omega t, \tag{3.8}$$

$$c(t) = -e^{t/2\tau}M(\omega_0^2/\omega)\sin\omega t, \tag{3.9}$$

$$d(t) = e^{t/2\tau}[\cos\omega t - (1/2\omega\tau)\sin\omega t]. \tag{3.10}$$

In the above expressions,

$$\omega = [\omega_0^2 - (1/4\tau^2)]^{1/2}. \tag{3.11}$$

Referring to the above equations and Eqs. (2.35)–(2.37), one obtains the following expressions for the parameters to be substituted in Eqs. (2.53) and (2.55):

$$\frac{\sigma}{\lambda} = \frac{1}{2}\left\{\frac{\omega_0}{\omega}\cosh\left(\frac{t}{2\tau}\right)\sin\omega t - i\left[\cosh\left(\frac{t}{2\tau}\right)\cos\omega t - \frac{1}{2\omega\tau}\sinh\left(\frac{t}{2\tau}\right)\sin\omega t\right]\right\}^{-1}, \tag{3.12}$$

$$\frac{\mu}{\lambda} = -\frac{\{(\omega_0/\omega)\sinh(t/2\tau)\sin\omega t + i[\sinh(t/2\tau)\cos\omega t - (1/2\omega\tau)\cosh(t/2\tau)\sin\omega t]\}}{\{(\omega_0/\omega)\cosh(t/2\tau)\sin\omega t + i[\sinh(t/2\tau)\cos\omega t - (1/2\omega\tau)\cosh(t/2\tau)\sin\omega t]\}}, \tag{3.13}$$

$$\frac{\sigma}{|\mu|} = \frac{1}{2}\left\{\frac{\omega_0^2}{\omega^2}\sinh^2\left(\frac{t}{2\tau}\right)\sin^2\omega t + \left[\sinh\left(\frac{t}{2\tau}\right)\cos\omega t - \frac{1}{2\omega\tau}\cosh\left(\frac{t}{2\tau}\right)\sin\omega t\right]^2\right\}^{-1/2}. \tag{3.14}$$

Substituting the above expressions into Eq. (2.55), one obtains

$$\begin{aligned} P_{nm} = {} & (-1)^m \frac{m!n!}{(2)^{m+n}}\left\{\frac{\omega_0^2}{\omega^2}\cosh^2\left(\frac{t}{2\tau}\right)\sin^2\omega t + \left[\cosh\left(\frac{t}{2\tau}\right)\cos\omega t - \frac{1}{2\omega\tau}\sinh\left(\frac{t}{2\tau}\right)\sin\omega t\right]^2\right\}^{-1/2} \\ & \times\left\{\frac{\omega_0^2}{\omega^2}\cosh^2\left(\frac{t}{2\tau}\right)\sin^2\omega t + \left[\sinh\left(\frac{t}{2\tau}\right)\cos\omega t - \frac{1}{2\omega\tau}\cosh\left(\frac{t}{2\tau}\right)\sin\omega t\right]^2\right\}^{-(m+n)/2} \\ & \times \sum_{\substack{\text{even}(l,l')=0 \\ [\text{odd}(l,l')=1]}}^{\min(m,n)} (-1)^{(l+l')/2}(2)^{l+l'} \frac{1}{l!l'![(m-l)/2]![(n-l)/2]![(m-l')/2]![(n-l')/2]!} \\ & \times\left\{\frac{\omega_0^2}{\omega^2}\sinh^2\left(\frac{t}{2\tau}\right)\sin^2\omega t + \left[\sinh\left(\frac{t}{2\tau}\right)\cos\omega t - \frac{1}{2\omega\tau}\cosh\left(\frac{t}{2\tau}\right)\sin\omega t\right]^2\right\}^{[m+n-(l+l')]/2}, \\ & \text{even }(m,n) \quad [\text{odd }(m,n)], \end{aligned} \tag{3.15}$$

$$P_{nm} = 0 \quad \text{odd}\,|m-n|. \tag{3.16}$$

The runaway oscillator is obtainable by the substitution $\tau \to -\tau$ in the above expressions and yields the same expression for P_{nm}.

We consider P_{nm} asymptotically as t approaches infinity. Then

$$\sinh\left(\frac{t}{2\tau}\right) \to \cosh\left(\frac{t}{2\tau}\right) \to \frac{e^{t/2\tau}}{2}. \tag{3.17}$$

Referring to Eq. (3.15),

$$\begin{aligned} P_{nm} \to {} & (-1)^m \frac{m!n!}{(2)^{m+n-1}} e^{-t/2\tau} \sum_{\substack{\text{even}(l,l')=0 \\ [\text{odd}(l,l')=1]}}^{\min(m,n)} (-1)^{(l+l')/2}(2)^{2(l+l')} \frac{1}{l!l'![(m-l)/2]![(n-l)/2]![(m-l')/2]![(n-l')/2]!} \\ & \times e^{-(l+l')t/2\tau}\left\{\frac{\omega_0^2}{\omega^2}\sin^2\omega t + \left[\cos\omega t - \frac{1}{2\omega\tau}\sin\omega t\right]^2\right\}^{[(m+n)/2-(l-l')/2-1]} \quad \text{even}(m,n) \quad [\text{odd}(m,n)]. \end{aligned} \tag{3.18}$$

Because of the factor $e^{-(l+l')t/2\tau}$, the dominant term corresponds to $l = l' = 0$ for even(m,n) and $l = l' = 1$ for odd(m,n). Hence,

$$P_{nm}\to e^{-t/2\tau}\frac{2(m-1)!!(n-1)!!}{m!!n!!}\left\{\frac{\omega_0^2}{\omega^2}\sin^2\omega t+\left[\cos\omega t-\frac{1}{2\omega\tau}\sin\omega t\right]^2\right\}^{[(m+n)/2-1]}\quad \text{even}(m,n), \tag{3.19}$$

$$P_{nm}\to e^{-3t/2\tau}(2)^3\frac{mn(m-2)!!(n-2)!!}{(m-1)!!(n-1)!!}\left\{\frac{\omega_0^2}{\omega^2}\sin^2\omega t+\left[\cos\omega t-\frac{1}{2\omega\tau}\sin\omega t\right]^2\right\}^{[(m+n)/2-2]}\quad \text{odd }(m,n). \tag{3.20}$$

Thus, for $t\to\infty$

$$P_{nm}\to 0\sim e^{-t/2\tau}\quad \text{even }(m,n), \tag{3.21}$$

$$P_{nm}\to 0\sim e^{-3t/2\tau}\quad \text{odd }(m,n). \tag{3.22}$$

Expressions (3.21) and (3.22) imply that the asymptotic transition probabilities for even (m,n) states and odd (m,n) states behave quantitatively different.

IV. DISCUSSION

In the above sections, the general expression for transition amplitudes was presented as well as the specific expressions for exponential coefficients. The general expressions have been calculated with respect to harmonic oscillator states. Since these states are complete, corresponding formulas for other states are readily obtainable. The results for the Kanai dampled oscillator are particularly interesting. If there is no dissipative mechanism, the transition probabilities are obviously zero. In the presence of a dissipating environment, the transitions are allowed. But the selection rules are that even (odd) states can decay into states of the same parity. Even more surprising is the different asymptotic behavior of the even and odd states. The dispersion into odd states from an odd state occurs faster than the corresponding dispersion for the even states.

We could imagine preparing two separate systems each linked to the same dissipative environment, one system in an even state and the second in an odd state. Subsequent examination of the system would find faster dispersion for the odd system than for the even system. The authors are currently investigating whether the same behavior would occur in experimentally interesting atomic systems.

A final point, adding linear terms to the Hamiltonian (2.1), i.e., terms proportional to x and p, would allow transitions from even to odd states. This case will be considered in a future paper.

[1]L. F. Landovitz, A. M. Levine, and W. M. Schreiber, Phys. Rev. A **20**, 1162 (1979).

[2]I. A. Malkin and V. I. Man'ko, Int. J. Theor. Phys. **14**, 37 (1975).

[3]E. Kanai, Prog. Theor. Phys. (Kyoto) 3, 440 (1948).

[4]I. R. Senitzky, Phys. Rev. **119**, 670 (1960).

[5]See, e.g., E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970), Chap. 5.

[6]See, e.g., I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals Series and Products* (Academic, New York, 1965), see 8.81.

Number operators for composite particles in nonrelativistic many-body theory

Wesley E. Brittin
Department of Physics, University of Colorado, Boulder, Colorado 80309
Arthur Y. Sakakura
Department of Physics, Colorado School of Mines, Golden, Colorado 80401

(Received 21 January 1980; accepted for publication 4 April 1980)

Commuting physical occupation number operators for composite particles are constructed using projection operator techniques. The composite particle occupation number operators are constructed from creation and annihilation operators of the elementary particles which make up the many-body system. They appear as positive operators in any given second quantized theory and represent observables within the framework of that theory. Bose-type composites have number operators with eigenvalues 0,1,2,..., and Fermi-type composites have number operators with eigenvalues 0,1. There does not arise here any problem having to do with exchange symmetry—exchange symmetry is exact, since the number operators act in the Fock space of the elementary particles. The composite particle number operators may be used in the construction of theories of composite particle reactions or equilibrium from a first principles standpoint. The construction used here not only establishes the existence of composite particle number operators but also provides some computational machinery which hopefully will aid in more practical applications.

I. INTRODUCTION

We consider a many-body system made up of large numbers of elementary particles which interact via a given nonrelativistic Hamiltonian. All observables can then be constructed from complete sets of annihilation and creation operators corresponding to elementary particles. The physical states of the system range over enormous numbers of possibilities. In certain physical circumstances the system may appear as a collection of nuclei, electrons, atoms, and/or molecules, while under other circumstances it may appear to be nearly fully ionized, etc. If the physical conditions are such that the system is readily describable in terms of a collection of interacting composite particles, the question arises as to how to introduce many-body observables corresponding to composites from a first principles standpoint. For example, if it is physically meaningful to talk about the number of composites of a certain type, then we should be able to construct for that composite a number operator whose expectation values correspond to the measured values of these quantities.

The first problem to be solved is that of finding a suitable description of a single composite particle in the many-body context. A hydrogen *atom* in a dense plasma may be quite distinct from a free hydrogen atom. This question has been studied by a number of investigators[1-3] and does not yet have a solution which is entirely satisfactory. We assume nevertheless that this part of the problem has been solved and that we have in our possession, for example, the wavefunctions of the individual composites which we are going to use in our description. The precise nature of these single composite particle wavefunctions or states is not important as far as our construction is concerned. They are arbitrary, subject to very few conditions, so the theory is quite general.

The construction of physical occupation numbers presented here may be regarded as more in the nature of an existence theorem than as a practical tool for computation. However, it is hoped that the mathematical aids introduced in our construction will also provide leads for methods of computation.

II. PHYSICAL PRELIMINARIES

We consider a many-body system made up of elementary particles of types $a,b,c,\cdots$. Let $a(i)$, $a(i)^*$, $b(j)$, $b(j)^*$, $c(k)$, $c(k)^*$, $\cdots$ be complete sets of one-particle annihilation and creation operators for the elementary particles. The physical many-body state space $\mathscr{F}$ is a separable Hilbert space built up from the normalized vacuum state $|0\rangle \in \mathscr{F}$, $\langle 0|0\rangle = 1$, by applications to $|0\rangle$ of the elementary particle creation operators. Let

$$|I\rangle \equiv (1/\sqrt{N_a!N_b!N_c!\cdots})a(i_1)^*a(i_2)^*\cdots b(j_1)^*b(j_2)^*\cdots|0\rangle, \tag{2.1}$$

where the index I includes the numbers N_a, N_b, N_c,$\cdots$ of elementary particles, there being precisely $N_a a(i)^*$'s, $N_b b(j)^*$'s, $N_c c(k)^*$'s,$\cdots$. Then the vectors $|I\rangle$ span $\mathscr{F}$ and the unit operator $1_{\mathscr{F}}$ is represented by the sum (strong limit)

$$\sum |I\rangle\langle I| = 1_{\mathscr{F}}. \tag{2.2}$$

The operators $a(i)$, $a(i)^*$,$\cdots$ satisfy elementary fermion anticommutation relations or elementary boson commutation relations, e.g.,

$$[a(i),a(i)^*]_\pm \equiv a(i)'a(i')^* \pm a(i')^*a(i)$$

 0022-2488/80/082164-06$1.00

$$= \delta(i,i') = \begin{cases} 1, & i = i', \\ 0 & i \neq i', \end{cases} \quad (2.3)$$

when applied to the vectors $|I\rangle$. The vectors $|I\rangle$ may be considered to be linearly independent in that if

$$\sum_I C_I \, |I\rangle = 0, \quad (2.4)$$

and if the C_I are completely symmetric in boson indices and completely antisymmetric in fermion indices, then $C_I \equiv 0$. Any vector $|\psi\rangle \in \mathscr{F}$ may be expanded in terms of the $|I\rangle$, since by Eq. (2.2),

$$|\psi\rangle = \sum_I |I\rangle \langle I|\psi\rangle \equiv \sum_I |I\rangle \psi(I), \quad (2.5)$$

and

$$\langle\psi|\psi\rangle \equiv \|\psi\|^2 = \sum_I |\psi(I)|^2. \quad (2.6)$$

We call $\psi(I)$ the wavefunction corresponding to $|\psi\rangle$.

Individual composite particles of types $A,B,C,\cdots$ are introduced by specifying their wavefunctions $\psi_\alpha(I)$, $\psi_\beta(I)$, $\psi_\gamma(I),\cdots$. We assume that the wavefunctions are orthonormal but not complete in the individual composite particle subspace of $\mathscr{F}$. That is, if $|\alpha\rangle$ represents the composite particle of type A having the wavefunction $\psi_\alpha(I)$,

$$|\alpha\rangle = \sum_\alpha |I\rangle \psi_\alpha(I), \quad (2.7)$$

then we assume that

$$\langle\alpha|\alpha'\rangle = \delta(\alpha,\alpha'), \quad (2.8)$$

but that the orthogonal projection operator ($\equiv$projector) represented by

$$\hat{P}_A = \sum_\alpha |\alpha\rangle \langle\alpha|, \quad (2.9)$$

is smaller than the projector corresponding to the complete subspace of $\mathscr{F}$ spanned by states of the elementary particles comprising $|\alpha\rangle$.

From the definition of $|I\rangle$, we may express $|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle,\cdots$ corresponding to the wavefunctions $\psi_\alpha(I)$, $\psi_\beta(I)$, $\psi_\gamma(I)$, as

$$\begin{aligned} |\alpha\rangle &= A(\alpha)^*|0\rangle \equiv \sum_I |I\rangle \psi_\alpha(I), \\ |\beta\rangle &= B(\beta)^*|0\rangle \equiv \sum_I |I\rangle \psi_\beta(I), \\ |\gamma\rangle &= C(\gamma)^*|0\rangle \equiv \sum_I |I\rangle \psi_\gamma(I). \end{aligned} \quad (2.10)$$

We observe that although the $\psi(I)$ are defined for all I, ψ will be zero for all sets of indices not corresponding to those indices corresponding to the numbers of elementary particles making up the given composite. The A^*, B^*, will usually be simple linear expressions in products of elementary particle creation operators. For example if $\psi_\alpha(ij)$ represents *bound* states of *hydrogen*

$$A^*_\alpha \equiv \sum_{i,j} a(i)^* b(j)^* \psi_\alpha(i,j), \quad (2.11)$$

where $a(i)^*$ creates a proton in the single elementary particle state $|i\rangle$ and $b(j)^*$ creates an electron in the single elementary particle state $|j\rangle$. The label α includes center of mass as well as internal degrees of freedom, and, as pointed out earlier, the ψ_α do not span all possible proton–electron states. We take ψ_α, ψ_β, $\psi_\gamma\cdots$ to represent bound composites. Ions are not excluded from the theory, but they will be represented as distinct composite particles. As far as our theory is concerned, the states $|\alpha\rangle$, $|\beta\rangle,\cdots$ may be more general than those representing individual composite particles. For some theories other types of excitations may be used.

The composite particle annihilation and creation operators $A(\alpha) \equiv [A(\alpha)^*]^*$, $A(\alpha)^*$, $B(\beta)$, $B(\beta)^*,\cdots$ do not have simple commutation or anticommutation relations.[2] If overlap between the single composite particle states *could* be neglected, then they would satisfy simple boson commutation or fermion anticommutation relations depending upon whether a given composite is a boson composite or a fermion composite. [A composite is called a boson composite if the number of elementary fermions making up the composite is even (including zero), and is called a fermion composite if the number of elementary fermions making up the composite is odd.] If the composite particle operators did satisfy elementary type commutation relations, then the vectors

$$|N\rangle = |\{N(\alpha)\},\{N(\beta)\},\cdots\rangle$$
$$\equiv \prod_{\alpha,\beta,\gamma,\cdots} \frac{[A(\alpha)^*]^{N(\alpha)}}{\sqrt{N(\alpha)!}} \frac{[B(\beta)^*]^{N(\beta)}}{\sqrt{N(\beta)!}} \cdots|0\rangle, \quad (2.12)$$

would represent a physical state of the many-body system in which there are $N(\alpha)$ composites of type A in the single composite particle state $|\alpha\rangle$, $N(\beta)$ composites of type B in the single composite particle state $|\beta\rangle,\cdots\alpha,\beta,\gamma, = 1,2,3,\cdots$. The operators in the product of Eq. (2.12) are ordered according to $\alpha = 1,2,3,\ldots$, $\beta = 1,2,3,\ldots$, and N represents the sets $\{N(\alpha)\}$ $\{N(\beta)\},\cdots\alpha,\beta,\cdots = 1,2,3$. The sums $\Sigma_\alpha(\alpha) < \infty$, $\Sigma_\beta N(\beta) < \infty,\cdots$ are finite but are otherwise unrestricted. We will use vectors of a more general type, but similar to $|N\rangle$ for our construction of number operators for composites. Let us define the vectors $|N,I\rangle$:

$$|N,I\rangle \equiv \prod_{\alpha,\beta,\cdots} \frac{[A(\alpha)^*]^{N(\alpha)}}{\sqrt{N(\alpha)!}} \frac{[B(\beta)^*]^{N(\beta)}}{\sqrt{N(\beta)!}} \cdots|I\rangle. \quad (2.13)$$

The vectors given by Eq. (2.13) above are highly overcomplete. Indeed, if $N = 0$, $|0,I\rangle = |I\rangle$, and $\{|I\rangle\}$ already forms a complete set. If the product of composite particle operators in $|N,I\rangle$ is not zero, then, as $|I\rangle$ varies over all I, the collection of $|N,I\rangle$ for fixed N spans a subspace of $\mathscr{F}$ which we describe as *having* $N(\alpha)$ or more composites of type A in the single composite particle state $|\alpha\rangle$, $N(\beta)$ or more composites of type B in the single composite particle state $|\beta\rangle,\ldots$, $\alpha = 1,2,3,\ldots,\beta = 1,2,3,\ldots$. The vectors $|N,I\rangle$ will be used to construct orthogonal subspaces of $\mathscr{F}$, each subspace of which corresponds to having precisely $N(\alpha)$ composites of type A in the single composite particle state $|\alpha\rangle$, $N(\beta)$ in $|\beta\rangle,\cdots,\alpha,\beta,\cdots = 1,2,3,\cdots$. Actually we will construct orthogonal projection operators (projectors) which act on $\mathscr{F}$ to yield the desired subspaces. These projectors will then be used to construct a commuting set of composite particle occupation number operators $\hat{N}(\alpha)$, $\hat{N}(\beta),\ldots$. First, however, we develop some mathematical techniques for dealing with noncommuting (in general) collections of projectors.

III. MATHEMATICAL PRELIMINARIES—PROJECTORS

A projector $\hat{P}$ on a Hilbert space $\mathscr{F}$ is a self-adjoint operator for which

$$\hat{P}^2 = \hat{P}. \tag{3.1}$$

It is clear that $\hat{P}$ is a positive (nonnegative) operator,

$$\langle\psi|\hat{P}\psi\rangle = \langle\psi|\hat{P}^*\hat{P}\psi\rangle = \langle\hat{P}\psi|\hat{P}\psi\rangle = \|\hat{P}\psi\|^2 \geqslant 0,$$

with operator norm 1 if $\hat{P} \neq 0$. The result of applying $\hat{P}$ to $\mathscr{F}$ yields a closed subspace $P = \hat{P}\mathscr{F}$, and, conversely, for every closed subspace $P \subset \mathscr{F}$ there exists a unique projector $\hat{P}$ such that $P = \hat{P}\mathscr{F}$. There is therefore a one-to-one correspondence between closed subspaces of $\mathscr{F}$ and projectors on $\mathscr{F}$. The collection of all projectors in $\mathscr{F}$ is endowed with a partial ordering, $<$, $\hat{P}_1 < \hat{P}_2$ if $P_1 \subset P_2$. If $\hat{P}_1 < \hat{P}_2$, then $\hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1 = \hat{P}_1$ and conversely. For every projector $\hat{P}$, $0 < \hat{P} < 1$. (Note $\hat{P} < \hat{P}$, so our partial ordering symbol $<$ could be written $\leqslant$.)

Two projectors $\hat{P}_1$, $\hat{P}_2$, are *orthogonal* to each other if

$$\hat{P}_1\hat{P}_2 = 0, \tag{3.2}$$

and the vectors in $P_1 = \hat{P}_2\mathscr{F}$ are orthogonal to those in $P_2 = \hat{P}_2\mathscr{F}$.

If $\hat{P}_2 > \hat{P}_1$, then $\hat{P}_2 - \hat{P}_1 \equiv \hat{P}$ is a projector and $\hat{P}$ is orthogonal to $\hat{P}_1$, for

$$\begin{aligned}\hat{P}^2 &= \hat{P}_2^2 + \hat{P}_1^2 - \hat{P}_2\hat{P}_1 - \hat{P}_1\hat{P}_2\\ &= \hat{P}_2 + \hat{P}_1 - \hat{P}_1 - \hat{P}_1 = \hat{P}_2 - \hat{P}_1 = \hat{P},\end{aligned}$$

and

$$\hat{P}\hat{P}_1 = (\hat{P}_2 - \hat{P}_1)\hat{P}_1 = \hat{P}_2\hat{P}_1 - \hat{P}_1 = \hat{P}_1 - \hat{P}_1 = 0.$$

We assume these and other elementary propeties of projectors to be known.

Given any two projectors $\hat{P}_1$, $\hat{P}_2$ then there exists a greater lower bound for $\hat{P}_1$, $\hat{P}_2$ designated by $\hat{P}_1 \wedge \hat{P}_2$ and defined to be the supremum of all projectors $\hat{P}$ for which $\hat{P} < \hat{P}_1$ and $\hat{P} < \hat{P}_2$,

$$\hat{P}_1 \wedge \hat{P}_2 \equiv \sup\{\hat{P}; \hat{P} < \hat{P}_1, \hat{P} < \hat{P}_2\}. \tag{3.3}$$

The corresponding closed subspace P is the largest closed subspace which is contained in P_1 and also contained in P_2. Hence

$$\hat{P}_1 \wedge \hat{P}_2 \leftrightarrow P = P_1 \cap P_2. \tag{3.4}$$

That is, P is the intersection of the closed subspaces P_1 and P_2. The classical construction of $\hat{P} = \hat{P}_1 \wedge \hat{P}_2$ in terms of $\hat{P}_1$ and $\hat{P}_2$ is given by von Neumann.[4] The result is expressed as a strong operator limit:

$$\hat{P}_1 \wedge \hat{P}_2 = \text{s-lim}(\hat{P}_1\hat{P}_2\hat{P}_1)^n, \tag{3.5}$$

which means that on any $|\psi\rangle \in \mathscr{F}$,

$$\hat{P}_1 \wedge \hat{P}_2|\psi\rangle = \lim_{n\to\infty} (\hat{P}_1\hat{P}_2\hat{P}_1)^n|\psi\rangle, \tag{3.6}$$

with the right-hand side coverging in vector norm. The proof rests upon the fact that $(\hat{P}_1\hat{P}_2\hat{P}_1)^n$ is a bounded monotone sequence of bounded positive operators, and therefore has a unique strong limit which is again a bounded positive operator $\hat{P}$. That $(\hat{P}_1\hat{P}_2\hat{P}_1)^n$ is a monotone decreasing sequence of positive operators means that

$$0 \leqslant \langle\psi|(\hat{P}_1\hat{P}_2\hat{P}_1)^n\psi\rangle \leqslant \langle\psi|(\hat{P}_1\hat{P}_2\hat{P}_1)^m\psi\rangle \leqslant \langle\psi|\psi\rangle, \quad \text{for } n \geqslant m.$$

It is easy to show that the unique $\text{s-lim}_{n\to\infty}(\hat{P}_1\hat{P}_2\hat{P}_1)^n \equiv \hat{P}$ has the property $\hat{P}^2 = \hat{P}$ and $\hat{P}|\psi\rangle = |\psi\rangle$ iff $\hat{P}_1|\psi\rangle = \hat{P}_2|\psi\rangle = |\psi\rangle$. There are equivalent ways of expressing $\hat{P}$:

$$\begin{aligned}\hat{P} \equiv \hat{P}_1 \wedge \hat{P}_2 &= \underset{n\to\infty}{\text{s-lim}}\,(\hat{P}_2\hat{P}_1\hat{P}_2)^n\\ &= \underset{n\to\infty}{\text{s lim}}\,(\hat{P}_1\hat{P}_2)^n = \underset{n\to\infty}{\text{s lim}}\,(\hat{P}_2\hat{P}_1)^n.\end{aligned}$$

Clearly if $\hat{P}_1$ and $\hat{P}_2$ commute $\hat{P}_1 \wedge \hat{P}_2 = \hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1$.

The concept of greatest lower bound can be extended to any collection $\{\hat{P}\alpha\}_{\alpha\in A}$, A an index set, of projectors $\hat{P}\alpha$:

$$\wedge_\alpha \hat{P}_\alpha = \sup\{\hat{P}, \hat{P} \text{ a projector}, \hat{P} < \hat{P}\alpha, \forall\alpha\in A\}, \tag{3.7}$$

with the corresponding closed subspace $\cap_{\alpha\in A} P\alpha$. We have need only for $\wedge_n \hat{P}_n$ where n goes over an enumerable set. The point of introducing a sketchy outline of von Neumanns construction of $\hat{P}_1 \wedge \hat{P}_2$ was to show that it would be desirable to have a much simpler construction if we wish to use operators such as $\wedge_{n=1}^{\infty} \hat{P}_n$ in computations.

The other lattice operator $\vee$ on projectors is complementary to $\wedge$. Given two projectors $\hat{P}_1$, $\hat{P}_2$, $\hat{P}_1 \vee \hat{P}_2$, the least upper bound of $\hat{P}_1$ and $\hat{P}_2$ is the smallest projector $\hat{P}$ for which $\hat{P}_1 < \hat{P}$ and $\hat{P}_2 < \hat{P}$. The closed subspace P corresponding to $\hat{P}_1 \vee \hat{P}_2$ is just the closure of the set of all linear combinations of vectors $|x_1\rangle$, $|x_2\rangle$, $|x_1\rangle \in P_1$, $|x_2\rangle \in P_2$. We have

$$\hat{P}_1 \vee \hat{P}_2 = (\hat{P}_1^\perp \wedge \hat{P}_2^\perp)^\perp \equiv 1_{\mathscr{F}} - (1_{\mathscr{F}} - \hat{P}_1) \wedge (1_{\mathscr{F}} - \hat{P}_2), \tag{3.8}$$

so that $\vee$ can be expressed in terms of $\wedge$ and vice versa. More generally,

$$\vee_n \hat{P}_n = (\wedge_n \hat{P}_n^\perp)^\perp, \quad \hat{P}^\perp \equiv 1_{\mathscr{F}} - \hat{P}. \tag{3.9}$$

We note that $\hat{P}_1 \vee \hat{P}_2 = \hat{P}_1 + \hat{P}_2$ iff $\hat{P}_1 \cdot \hat{P}_2 = 0$, and if $\hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1$, $\hat{P}_1 \vee \hat{P}_2 = \hat{P}_1 + \hat{P}_2 - \hat{P}_1\hat{P}_2$.

Let $\{\hat{P}_n\}_{n=1}^{\infty}$ be an enumerable set of projectors from which we wish to construct $\hat{P} \equiv \vee_n \hat{P}_n$ (or $\wedge_n \hat{P}_n$). We define the positive operator $\hat{g}$, by

$$\hat{g} \equiv \sum_{n=1}^{\infty} \hat{P}_n / 2^n. \tag{3.10}$$

The sum in Eq. (3.10) converges in operator norm (uniformly), since

$$\|\hat{g}\| \leqslant \sum_{n=1}^{\infty} \frac{\|\hat{P}_n\|}{2^n} \leqslant \sum_{n=1}^{\infty} \frac{1}{2^n} = 1, \tag{3.11}$$

thus $\hat{g}$ is a bounded positive operator with norm less than or equal to unity. We observe that $\hat{P}$ is precisely the *support* of $\hat{g}$. The *support* $\hat{S}$ of a bounded self-adjoint operator $\hat{g}$ is the smallest projector $\hat{S}$ for which $\hat{g}\hat{S} = \hat{g}$, so $\hat{S}$ is the orthogonal complement of $\hat{N}$, the projector onto the null space of $\hat{g}$. To prove the above assertion let $|\psi\rangle \in N$, then

$$g|\psi\rangle = 0, \tag{3.12}$$

or

$$\sum_{n=0}^{\infty} \frac{\hat{P}_n|\psi\rangle}{2^n} = 0.$$

If we take the inner product of the above relation with $|\psi\rangle$, we find

$$0=\sum_{n=1}^{\infty}\frac{\langle\psi|\hat{P}_n\psi\rangle}{2^n}=\sum_{n=1}^{\infty}\frac{\langle\hat{P}_n\psi|\hat{P}_n\psi\rangle}{2^n}$$
$$=\sum_{n=1}^{\infty}\frac{\|\hat{P}_n\psi\|^2}{2^n}. \tag{3.13}$$

Hence

$$\|\hat{P}_n\psi\|=0 \text{ and } \hat{P}_n|\psi\rangle=0, \quad n=1,2,3,\cdots,$$

which proves that $|\psi\rangle\in P^{\perp}$. Conversely if $|\psi\rangle\in P^{\perp}$, then $\hat{g}|\psi\rangle=0$. Therefore $N^{\perp}=\hat{P}=\vee_n\hat{P}_n$. If $|\psi\rangle\in P=\widehat{S\mathscr{F}}$, and $|\psi\rangle\neq 0$, then $\hat{g}|\psi\rangle\neq 0$.

Since $\hat{g}$ is positive, $\hat{g}_\epsilon\equiv\hat{g}+\epsilon 1_{\mathscr{F}}$, $\epsilon>0$, is strictly positive so[5] $\hat{g}_\epsilon^{-1}$ exists and is a bounded positive operator. If $|\psi\rangle\in P$, then

$$\hat{g}_\epsilon^{-1}\hat{g}|\psi\rangle=\hat{g}_\epsilon^{-1}(\hat{g}_\epsilon-\epsilon 1_{\mathscr{F}})|\psi\rangle$$
$$=|\psi\rangle-\epsilon\hat{g}_\epsilon^{-1}|\psi\rangle,$$

and

$$\lim_{\epsilon\to 0}\hat{g}_\epsilon^{-1}g|\psi\rangle=|\psi\rangle, \tag{3.14}$$

since $\hat{g}|\psi\rangle\neq 0, 0\neq|\psi\rangle\in P$. If $|\psi\rangle\in N=P^{\perp}$, then $\hat{g}_\epsilon^{-1}\hat{g}|\psi\rangle=0$, and

$$\lim_{\epsilon\downarrow 0}\hat{g}_\epsilon^{-1}\hat{g}|\psi\rangle=0. \tag{3.15}$$

Therefore

$$\hat{P}=\vee_n\hat{P}_n=\operatorname*{s-lim}_{\epsilon\downarrow 0}(\hat{g}_\epsilon^{-1}\hat{g}). \tag{3.16}$$

We have constructed $\vee_n\hat{P}_n$ using techniques that are no more cumbersome than those used in many areas of mathematical physics, e.g., scattering theory.

For some applications other forms for $\hat{P}$ may be more useful. For example, since $\hat{g}$ is a bounded positive operator, $e^{-t\hat{g}}$ is a bounded positive operator for $t\geqslant 0$. It is not difficult to establish that

$$\hat{P}=\vee_n\hat{P}_n=\left(\operatorname*{s-lim}_{t\to\infty}e^{-t\hat{g}}\right)^{\perp}, \tag{3.17}$$

which is another expression for $\vee_n\hat{P}_n$.

We have need for one more construction. Given an arbitrary collection $\{|\psi_n\rangle\}_{n=0}^{\infty}$ of vectors $|\psi_n\rangle$, we construct the projector $\widehat{P}$ corresponding to the closed subspace P spanned by $\{|\psi_n\rangle\}_{n=1}^{\infty}$. Let $C_n>0$ for $|\psi_n\rangle\neq 0$, be such that the sum $\Sigma_{n=1}^{\infty}C_n|\psi_n\rangle\langle\psi_n|\equiv\hat{g}$ converges uniformly. This can always be done, for example, by setting $C_n=0$, if $|\psi_n\rangle=0$ and $C_n=(1/2^n)(1/\|\Psi_n\|^2)$ if $|\psi_n\rangle\neq 0$. Then

$$\left|\left|\sum_n C_n|\psi_n\rangle\langle\psi_n|\right|\right|\leqslant\sum_n C_n||\Psi_n||^2\leqslant 1. \tag{3.18}$$

Then the projector $\widehat{P}$ on the closed span of $\{|\psi_n\rangle\}_{n=1}^{\infty}$ is given by

$$\widehat{P}=\operatorname*{s-lim}_{\epsilon\downarrow 0}(\hat{g}_\epsilon^{-1}\hat{g}). \tag{3.19}$$

IV. CONSTRUCTION OF COMPOSITE PARTICLE OCCUPATION NUMBER OPERATORS

We now classify states of a given many-body system in terms of occupation numbers for composite particles. This will enable us to construct corresponding occupation numbers for composite particles. The problem of such a classification cannot be expected to have a unique solution. Our solution is one which we have tried to make as free from arbitrariness as possible, and at the same time one which is sufficiently general to include most anticipated applications.

The vectors

$$|N,I\rangle\equiv|\{N(\alpha)\},\{N(\beta)\},\cdots,I\rangle, \tag{4.1}$$

given by Eq. (2.13), where $\alpha=1,2,3,\cdots$, $\beta=1,2,3,\cdots,\cdots$, and $\Sigma_\alpha N(\alpha)$, $\Sigma_\beta N(\beta),\cdots$ finite, if not zero, for fixed N span, as I varies, a closed subspace $P(N)$ of $\mathscr{F}$. If $|\psi\rangle$ is a normalized vector and if $|\psi\rangle\in P(N)$, we say that the state $|\psi\rangle$ *has* $N(\alpha)$ or more composites of type A in the single composite particle state $|\alpha\rangle$, $\alpha=1,2,3,\cdots$, $N(\beta)$ or more composites of type B in the single composite particle state $|\beta\rangle$, $\beta=1,2,3,\cdots,\cdots$. For fermion composites $N(\alpha)=0,1$, and for boson composites $N(\alpha)=0,1,2,3,\cdots$. The projector $\widehat{P}(N)$ on the subspace $P(N)$ may be constructed using the techniques of Sec. II. Let

$$\hat{g}(N)=\Sigma_I|N,I\rangle\langle N,I|$$
$$=\sum_I\prod_{\alpha,\beta,\cdots}\frac{[A(\alpha)^*]^{N(\alpha)}}{\sqrt{N(\alpha)!}}\frac{[B(\beta)^*]^{N(\beta)}}{\sqrt{N(\beta)!}}\cdots|I\rangle\langle I|$$
$$\cdots\frac{[B(\beta)]^{N(\beta)}}{\sqrt{N(\beta)!}}\frac{[A(\alpha)]^{N(\alpha)}}{\sqrt{N(\alpha)!}}. \tag{4.2}$$

But $\Sigma_I|I\rangle\langle I|=1_{\mathscr{F}}$, therefore

$$\hat{g}(N)=\prod_{\alpha,\beta,\cdots}\frac{[A(\alpha)^*]^{N(\alpha)}}{\sqrt{N(\alpha)!}}\frac{[B(\beta)^*]^{N(\beta)}}{\sqrt{N(\beta)!}}\cdots$$
$$\cdots\frac{B(\beta)^{N(\beta)}}{\sqrt{N(\beta)!}}\frac{A(\alpha)^{N(\alpha)}}{\sqrt{N(\alpha)!}}, \tag{4.3}$$

or

$$\hat{g}(N)=:\prod_{\alpha,\beta,\cdots}\frac{[A^*(\alpha)A(\alpha)]^{N(\alpha)}}{N(\alpha)!}\frac{[B^*(\beta)B(\beta)]^{N(\beta)}}{N(\beta)!}\cdots:, \tag{4.4}$$

where : : is the Wick symbol for the normal ordering of operators. The projector $\widehat{P}(N)$ is then given by the strong limit as $\epsilon\downarrow 0$ of

$$\widehat{P}_\epsilon(N)\equiv\hat{g}_\epsilon(N)^{-1}\hat{g}(N). \tag{4.5}$$

If all $N(\alpha)$, $N(\beta)$, are zero $\widehat{P}(0)=1_{\mathscr{F}}$. Let all $N(\alpha')$, $N(\beta),\cdots$ be fixed except for a given $N(\alpha)$, then, because of the incompleteness assumption concerning the single composite particle states $|\alpha\rangle$,

$$\widehat{P}(\cdots,N(\alpha)+1,\cdots)<\widehat{P}(\cdots N(\alpha)\cdots). \tag{4.6}$$

This means that

$$\widehat{P}(N,N(\alpha))\equiv\widehat{P}(\cdots N(\alpha)\cdots)-\widehat{P}(\cdots,N(\alpha)+1,\cdots), \tag{4.7}$$

is a projector, for $N(\alpha)=0,1,2,\cdots$. Further

$$\widehat{P}(N,N(\alpha))\widehat{P}(N,N'(\alpha))=\delta(N(\alpha),N'(\alpha))\widehat{P}(N,N(\alpha)), \tag{4.8}$$

thus $\{\widehat{P}(N,N(\alpha))\}_{N(\alpha)=0,1,2,\ldots}$ is an orthogonal set of projectors. Many-body states lying in the subspace $\widehat{P}(N,N(\alpha))\mathscr{F}=P(N,N(\alpha))$ may be considered to have precisely $N(\alpha)$ composites of type A in the single composite particle state $|\alpha\rangle$, $N(\alpha')(\alpha'\neq\alpha)$ or more composites of type A in single composite particle state $|\alpha'\rangle$, $N(\beta)$ or more composites of type B in the single composite particle state $|\beta\rangle$,$\cdots$. The subspaces $P(N,N(\alpha))$, $P(N,N'(\alpha))$, $N'(\alpha)\neq N(\alpha)$, are orthogonal because of Eq. (4.8).

We wish to construct closed subspaces P_c of $\mathscr{F}$ which correspond to precisely $N(\alpha)$, $N(\beta)$,$\cdots\alpha = 1,2$, $\beta = 1,2,\cdots$ particles in the various single composite particle states. A logical candidate for the projector on P_c is

$$\hat{Q}(N)\equiv \wedge_{\alpha}\hat{P}(N,N(\alpha)) \wedge \wedge_{\beta}\hat{P}(N,N(\beta))\cdots.$$

The corresponding subspace $Q(N)$ would then correspond to states having precisely $N(\alpha)_{\alpha=1,2,\ldots}$ A composites in the state $|\alpha\rangle$, $N(\beta)_{\beta=1,2,\ldots}$ B composites in the state $|\beta\rangle$, $\cdots$. Unfortunately quantum logic is not the same as everyday logic. The subspaces $Q(N)$, $Q(N')$ are not in general orthogonal. If we fix all $N(\alpha')$, $N(\beta)$, $N(\gamma)$, except $N(\delta)$, say, then $Q(\cdots N(\delta)\cdots)$ and $Q(\cdots N(\delta)'\cdots)$ are orthogonal for $N(\delta)\neq N'(\delta)$, but in general $Q(\cdots N(\alpha)\cdots N(\beta)\cdots)$ and $Q(\cdots N(\alpha)'\cdots N(\beta)'\cdots)$, $N(\alpha)\neq N'(\alpha)$, $N(\beta)\neq N'(\beta)$, overlap. The subspaces $Q(N)$ are too large and must be peeled down. This state of affairs was anticipated since we placed very little restrictions on the various single composite particle states. There may be linear relations among the vectors built up from the creation operators of the single composites, etc. The authors have constructed a theory[5] which avoids this difficulty by selecting carefully certain linearly independent collections of vectors. The theory presented here is very general and a certain lack of uniqueness is to be expected. Nevertheless, a satisfactory theory does emerge. We must construct a family of orthogonal subspaces starting with the $Q(N)$, and this must be done in as unarbitrarily a way as possible. The $\hat{Q}(N)$ are labeled by a finite number of nonzero integers $\{N(\alpha), N(\beta)\cdots\}$ and may therefore be linearly ordered

$$O{:}\hat{Q}_1,\hat{Q}_2,\cdots\hat{Q}_{N^-},\hat{Q}(N), \tag{4.9}$$

where Q_{N^-} is the immediate predecessor of $Q(N)$ in the ordering. The ordering (4.9) may be done using physical arguments, or it may be done for example by means of a lexographic scheme. There does not seem to be any way out of selecting a certain order. (This appears to be a generalization of the Schmidt process for vectors.)

The generalized Schmidt process for projectors may now be carried out. Define the projectors

$$\hat{R}_M \equiv \vee_{j=1}^{M} \hat{Q}_j, \tag{4.10}$$

then

$$\hat{R}_M < \hat{R}_{M'}, \quad M<M', \tag{4.11}$$

and $\hat{R}_M - \hat{R}_{M-1}$ is a projector for $M = 1,2,3,\cdots$ ($\hat{R}_0\equiv 0$). Let

$$\hat{P}_c(N)\equiv\hat{R}_N - \hat{R}_{N^-}, \tag{4.12}$$

where $N = \{N(\alpha), N(\beta), N(\gamma),\cdots\}$ and N^- is the immediate predecessor of N in the ordering $\hat{Q}_1,\ldots,\hat{Q}(N)$. The $\hat{P}_C(N)$ forms a orthogonal set of projectors:

$$\hat{P}_c(N)\hat{P}_c(N') = \delta(N,N')\hat{P}_c(N), \quad \hat{P}_c(N)\lessdot\hat{P}_c(N'), \tag{4.13}$$

and, in obvious notation,

$$\sum_{k=1}^{N} \hat{P}_c(k) = \vee_{k=1}^{N} \hat{Q}_k, \tag{4.14}$$

where $\hat{P}_c(k)\leftrightarrow\hat{Q}_k$ are ordered in the same fashion as the $\hat{Q}_k$.

We take $\hat{P}_c\{N(\alpha)\}\ \{N(\beta)\}$ to be the projector on the subspace P_c of $\mathscr{F}$ which corresponds to precisely $N(\alpha)$, $N(\beta)\cdots\alpha = 1,2,\cdots,\beta = 1,2,3,\cdots$ composites in the states $|\alpha\rangle$, $|\beta\rangle$, $\cdots$. For certain N, $\hat{P}_c$ may be zero. However, this is quite satisfactory and just corresponds to the fact that we are then dealing with an overcomplete description.

Commuting operators $\hat{N}(\alpha)$, $\hat{N}(\beta)$, $\cdots$ may now be introduced for composite particle occupation numbers, by means of the definitions

$$\hat{N}(\alpha) = \sum_N N(\alpha)\hat{P}_c(N),$$

$$\hat{N}(\beta) = \sum_N N(\beta)\hat{P}_c(N),\cdots, \tag{4.15}$$

where the sum extends over all $N(\alpha)$, $N(\beta)$, $\alpha = 1,2,3,\ldots$, $\beta = 1,2,3,\cdots$, $N(\alpha)$, $N(\beta)$,$\cdots = 0,1,2,\cdots$.

The number operators commute with each other and with the total number operators $\Sigma_i a(i)^*a(i)$, $\Sigma_j b(j)^*b(j)$,$\cdots$ of the elementary particles. The composite number operators do not in general commute with the Hamiltonian, so they may be used as a basis for a first principles approach to nonequilibrium problems.

The projectors $\hat{P}_c(N)$ may be used to classify states of the many-body system. Let

$$\hat{P}_0 + \sum_{N\neq 0} \hat{P}_c(N) = 1_{\mathscr{F}}, \tag{4.16}$$

where $N\neq 0$ means that at least one member of the set $\{N(\alpha)$, $N(\beta)$, $N(\gamma)\}$ is not zero. We then interpret $\hat{P}_0$ to be the projector on the subspace P_0 corresponding to the completely ionized state. This does not mean that there is no ionization in other states. For example, if $\hat{P}_c(N)|\psi\rangle = |\psi\rangle$, and $|\psi\rangle$ corresponds to more than the numbers of elementary particles accounted for in $\{N(\alpha), N(\beta), \cdots\}$, the excess elementary particles may be regarded as ionized. For any $|\psi\rangle\in\mathscr{F}$, according to Eq. (4.16), we have the orthogonal decomposition

$$|\psi\rangle = |\psi_0\rangle + \sum_{N\neq 0} |\psi(N)\rangle, \tag{4.17}$$

and

$$\||\psi\rangle\|^2 = \||\psi_0\rangle\|^2 + \sum_N \||\Psi(N)\rangle\|^2, \tag{4.18}$$

with

$$|\psi(N)\rangle\equiv\hat{P}_c(N)|\psi\rangle.$$

We may extend the above considerations to situations where it is desirable to consider number operators for the elementary particles which commute with the composite particle number operators. Such operators would then be used to provide a description of the many-body system in terms of *free* elementary particles and bound composite particles.

V. CONCLUSION

We have classified states of nonrelativistic many-body systems through the use of occupation numbers for composite particles. The occupation numbers specify how many composites are in a given single composite particle state. This description is kinematic and does not say anything concerning the dynamics of the many body system. Any state may be decomposed, $|\psi\rangle = |\psi_0\rangle + \Sigma_{N\neq 0}|\psi(N)\rangle$, where $|\psi_0\rangle$ corresponds to the completely ionized state and $|\psi(N)\rangle$ to precise numbers of composites. Commuting occupation number operators were constructed using projection operator techniques. These number operators are central to fundamental theories having to do with composite particle kinetics and equilibria.

[1]S. Golden, *Quantum Statistical Foundations of Chemical Kinetics* (Clarendon, Oxford 1969).
[2]M. D. Girardeau, J. Math. Phys. **19**, 2605 (1978).
[3]J. D. Gilbert, J. Math. Phys. **18**, 791 (1977).
[4]John von Neumann, in *Functional Operators*, Vol. II: *The Geometry of Orthogonal Spaces*, Princeton U. P., Princeton, N.J., 1950), pp. 51–59.
[5]W. E. Brittin and A. Y. Sakakura, Phys. Rev. A **21**, 2050 (1980).

The isotropic harmonic oscillator in an angular momentum basis: An algebraic formulation

A. J. Bracken and H. I. Leemon
Department of Mathematics, University of Queensland, St. Lucia, Queensland 4067 Australia

(Received 18 July 1979; accepted for publication 30 August 1979)

A completely algebraic and representation-independent solution is presented of the simultaneous eigenvalue problem for H, $\mathbf{L}^2$, and $\mathbf{L}_3$, where $\mathbf{H}$ is the Hamiltonian operator for the three-dimensional, isotropic harmonic oscillator, and $\mathbf{L}$ is its angular momentum vector. It is shown that H can be written in the form $\hbar\omega(2\nu^\dagger\nu + \boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda} + 3/2)$, where $\nu^\dagger$ and ν are raising and lowering (boson) operators for $\nu^\dagger\nu$, which has nonnegative integer eigenvalues k; and $\boldsymbol{\lambda}^\dagger$ and $\boldsymbol{\lambda}$ are raising and lowering operators for $\boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda}$, which has nonnegative integer eigenvalues l, the total angular momentum quantum number. Thus the eigenvalues of H appear in the familiar form $\hbar\omega(2k + l + 3/2)$, previously obtained only by working in the coordinate or momentum representation. The common eigenvectors are constructed by applying the operators $\nu^\dagger$ and $\boldsymbol{\lambda}^\dagger$ to a "vacuum" vector on which ν and $\boldsymbol{\lambda}$ vanish. The Lie algebra so(2,1) ⊕ so(3,2) is shown to be a spectrum-generating algebra for this problem. It is suggested that coherent angular momentum states can be defined for the oscillator, as the eigenvectors of the lowering operators ν and $\boldsymbol{\lambda}$. A brief discussion is given of the classical counterparts of ν, $\nu^\dagger$, $\boldsymbol{\lambda}$, and $\boldsymbol{\lambda}^\dagger$, in order to clarify their physical interpretation.

1. INTRODUCTION

The eigenvalue problem for the three-dimensional, isotropic harmonic oscillator Hamiltonian operator,

$$H = \frac{\mathbf{p}^2}{2M} + \tfrac{1}{2}M\omega^2\mathbf{x}^2 , \qquad (1)$$

is often solved algebraically (see for example Stehle,[1] Sec. 8). One introduces the boson creation and annihilation operators

$$\mathbf{a}^\dagger = (2M\hbar\omega)^{-1/2}(-i\mathbf{p} + M\omega\mathbf{x}) ,$$
$$\mathbf{a} = (2M\hbar\omega)^{-1/2}(i\mathbf{p} + M\omega\mathbf{x}) , \qquad (2)$$

which are Hermitian conjugate to each other, and which satisfy the commutation relations

$$[a_i, a_j] = 0 = [a_i^\dagger, a_j^\dagger] ,$$
$$[a_i, a_j^\dagger] = \delta_{ij} , \quad i,j = 1,2,3 . \qquad (3)$$

Then one has

$$H = \hbar\omega(N + 3/2) , \qquad (4)$$

where

$$N = \mathbf{a}^\dagger\cdot\mathbf{a} = N_1 + N_2 + N_3 , \qquad (5)$$

with, for example, $N_1 = a_1^\dagger a_1$. The usual boson calculus leads to the conclusion that the commuting operators N_1, N_2, and N_3, have simultaneous eigenvalues n_1, n_2, and n_3, running over all nonnegative integers independently, so that the eigenvalues of N appear in the form $n_1 + n_2 + n_3$. The corresponding normalized eigenvector may be denoted $|n_1, n_2, n_3\rangle$, and is nondegenerate. It may be obtained from a normalized "vacuum vector" $|0\rangle$ as

$$|n_1, n_2, n_3\rangle = (n_1!n_2!n_3!)^{-1/2}(a_1^\dagger)^{n_1}(a_2^\dagger)^{n_2}(a_3^\dagger)^{n_3}|0\rangle , \qquad (6)$$

where

$$a_i|0\rangle = 0 , \quad i = 1,2,3 , \qquad (7)$$

so that

$$N|0\rangle = 0 = N_i|0\rangle , \quad i = 1,2,3 . \qquad (8)$$

The eigenvalue problem for H (equivalently, for N) may be solved also in an "angular momentum basis." (See for example Davydov,[2] Sec. 37.) One works in either the coordinate or the momentum representation and looks for the common eigenfunctions of N, $\mathbf{L}^2$, and L_3, where

$$L_i = \tfrac{1}{2}\hbar\epsilon_{ijk}l_{jk} ,$$
$$l_{jk} = (x_jp_k - x_kp_j)/\hbar$$
$$= i(a_ja_k^\dagger - a_ka_j^\dagger) , \qquad (9)$$

so that

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k ,$$
$$i[l_{ij}, l_{km}] = \delta_{jk}l_{im} + \delta_{im}l_{jk} - \delta_{ik}l_{jm} - \delta_{jm}l_{ik} . \qquad (10)$$

The simultaneous eigenvalues are found to be

$$N: 2k + l , \quad \mathbf{L}^2: l(l+1)\hbar^2 , \quad L_3: m\hbar , \qquad (11)$$

where k and l run over the nonnegative integers independently, and for a given l, m runs over $l, l-1, \cdots, -l$. The corresponding normalized eigenfunction may be denoted ϕ_{klm} and is nondegenerate.

In this paper we show how the simultaneous eigenvalue problem for N, $\mathbf{L}^2$, and L_3 can be solved in a purely algebraic way, with the introduction of operators which raise and lower the values of k, l, and m, rather than n_1, n_2, and n_3. More precisely we find that N can be written in the form [contrast with Eq. (5)]

$$N = 2\nu^\dagger\nu + \boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda} , \qquad (12)$$

where $\nu^\dagger$ and ν are raising and lowering operators for $\nu^\dagger\nu$, which has eigenvalues k; and $\boldsymbol{\lambda}^\dagger$ and $\boldsymbol{\lambda}$ are raising and lowering operators for $\boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda}$, which has eigenvalues l, the total angular momentum quantum number. The normalized common eigenvectors, denoted $|klm\rangle$, are obtained by applying suitable combinations of the raising operators to a normalized "vacuum vector" $|0\rangle$, which satisfies

$$N|0\rangle = \mathbf{L}^2|0\rangle = L_3|0\rangle = 0. \quad (13)$$

[It is readily seen that this vector can be identified with the vector defined by Eqs. (7) or (8), hence the common notation.] In this approach, the fundamental dynamical variables in the problem are $\nu^\dagger$, ν, $\boldsymbol{\lambda}^\dagger$, and $\boldsymbol{\lambda}$ rather than $\mathbf{a}^\dagger$ and $\mathbf{a}$.

Of all the many investigations of the harmonic oscillator and related problems (for reviews and many references, see Kramer and Moshinsky[3] and McIntosh[4]), the closest in spirit to ours is that by Rose,[5] who examined the algebraic structure of operators χ_{kl}^m satisfying

$$|k\,l\,m\rangle = \chi_{kl}^m|0\rangle .$$

However, Rose did not identify the elementary operators ν, $\nu^\dagger$, $\boldsymbol{\lambda}$, and $\boldsymbol{\lambda}^\dagger$ in terms of which the Hamiltonian operator and all such χ_{kl}^m can be expressed [see our Eqs. (12) and (53)], and in terms of which the eigenvalue problem can be formulated and solved completely.

The algebraic solution of this problem is of some intrinsic interest, being independent of the choice of a particular representation space. Although one knows that any problem in quantum mechanics can be formulated in a variety of equivalent representations, and that the eigenvalues of any particular operator are determined by the structure of the relevant algebra of operators, rather than by the choice of representation space, few problems have been analyzed completely in a representation-independent way. (For examples, see the book of Green.[6] Of course, our constructions necessarily also define in the coordinate representation, for example, shift-operators associated with the differential operators N, $\mathbf{L}^2$, and L_3. There is a point of contact here with the so-called "factorization method."[7] We note however that the operator L which we introduce in the next section and which plays a central role in our analysis, is an integral operator, not a differential operator, in both the coordinate and the momentum representation.)

Having an algebraic formulation, we readily identify a hitherto unrecognized spectrum-generating algebra for this problem, namely the Lie algebra so(2,1) $\oplus$ so(3,2). However, our motivation for this work is primarily to set up an algebraic framework within which we can construct "coherent angular momentum states" for the oscillator. The investigation of such states will be the subject of a subsequent publication. They will be defined as common eigenvectors of the lowering operators ν and $\boldsymbol{\lambda}$, just as the usual coherent states can be defined as common eigenvectors of the lowering operators $\mathbf{a}$. They have many interesting properties in common with the usual coherent states, leading us to hope that they also will prove useful. Further motivation for the study of such states may be found in the work of Atkins and Dobson,[8] and of Delbourgo,[9] where the idea of superposing eigenvectors corresponding to all the possible values of the total angular momentum quantum number of a system, to form "coherent angular momentum states," has been proposed in a more general context.

In Sec. 2 we derive expressions for the operators $\nu^\dagger$, ν, $\boldsymbol{\lambda}^\dagger$ and $\boldsymbol{\lambda}$, and investigate some of their properties. Some proofs are relegated to Appendix A. The method used to determine $\boldsymbol{\lambda}^\dagger$ and $\boldsymbol{\lambda}$, in particular, depends heavily on techniques developed by Bracken and Green[10] for the analysis of vector operators. Indeed, the idea of constructing from the vector operators $\mathbf{a}^\dagger$ and $\mathbf{a}$ other vector operators which form "creation and annihilation operators for angular momentum" was partly developed some years ago by them.[11]

In Sec. 3, we present with the help of these operators the solution of the common eigenvalue problem for N, $\mathbf{L}^2$, and L_3, relegating some proofs to Appendix B. Then in Sec. 4, we discuss the time-dependence of these operators (in the Heisenberg picture) and the meaning of their classical counterparts.

It is known[12] that the Lie algebra sp(6, R) is a relevant spectrum-generating algebra for the oscillator Hamiltonian when N_1, N_2, and N_3 are to be diagonalized. In Sec. 5, we show that the Lie algebra so(2, 1) $\oplus$ so (3, 2) [$\approx$sp(2, R) $\oplus$ sp(4, R)] is a more appropriate spectrum-generating algebra when N, $\mathbf{L}^2$, and L_3 are to be diagonalized.

2. THE APPROPRIATE DYNAMICAL VARIABLES

In order to introduce the operators $\nu^\dagger$, ν, $\boldsymbol{\lambda}^\dagger$, and $\boldsymbol{\lambda}$ with the desirable properties described above, it is necessary in the first place to define the operator $L+\frac{1}{2}$, as the positive, scalar, Hermitian square-root of the positive operator $\frac{1}{2}l_{ij}l_{ij}+\frac{1}{4}(=\hbar^{-2}\mathbf{L}^2+\frac{1}{4})$, so that

$$\mathbf{L}^2 = L(L+1)\hbar^2 . \quad (14)$$

It follows from the nonnegativity of $\mathbf{L}^2$ that any of its eigenvalues can be written in the form $l(l+1)\hbar^2$, with l nonnegative. On the same eigenvector, the eigenvalue of L will then be l. Of course, it will turn out that l runs over all the nonnegative integers—but we deduce this, not assume it.

We define also the Hermitian operator

$$K = \tfrac{1}{2}(N-L), \quad (15)$$

so that $N = 2K + L$. Like all scalar operators, N (and hence K) commutes with all l_{ij}, and therefore with L.

However, the vector operator $\mathbf{a}$ (and likewise $\mathbf{a}^\dagger$) can be resolved into the sum of a vector operator which shifts the eigenvalue of L up by one unit, and a vector operator which shifts it down by one unit. This may be seen with the help of the techniques developed by Bracken and Green[10] as follows: From Eqs. (3) and the definition (9) of l_{ij} we have

$$\epsilon_{ijk}a_i l_{jk} = 0, \quad (16)$$

or, equivalently,

$$a_i l_{jk} + a_k l_{ij} + a_j l_{ki} = 0. \quad (17)$$

Contracting on the right with $\frac{1}{2}l_{ij}$, and using the commutation relations (10) and the definition of L, we find

$$a_i l_{ij} l_{jk} + i a_i l_{ik} + a_k L(L+1) = 0, \quad (18)$$

that is

$$\begin{aligned} 0 &= a_i[l_{ij} + i(L+1)\delta_{ij}][l_{jk} - iL\delta_{jk}] \\ &= a_i[l_{ij} - iL\delta_{ij}][l_{jk} + i(L+1)\delta_{jk}] . \end{aligned} \quad (19)$$

We define the operators $\mathbf{a}^{(\pm)}$ by

$$a_j^{(\pm)} = a_i[(L+\tfrac{1}{2})\delta_{ij} \pm \tfrac{1}{2}\delta_{ij} \mp i l_{ij}][2L+1]^{-1} \quad (20)$$

(noting that $[2L+1]$ has a well-defined inverse, since L is nonnegative). Then $\mathbf{a}^{(\pm)}$ is evidently a vector-operator so that

$$i[a_i^{(\pm)}, l_{jk}] = \delta_{ij} a_k^{(\pm)} - \delta_{ik} a_j^{(\pm)}, \tag{21}$$

and hence

$$[a_i^{(\pm)}, \tfrac{1}{2} l_{jk} l_{jk}] = 2 i a_k^{(\pm)} l_{ki} - 2 a_i^{(\pm)}. \tag{22}$$

But, according to Eqs. (19) and (20),

$$i a_k^{(\pm)} l_{ki} = a_i^{(\pm)} [\tfrac{1}{2} \mp (L + \tfrac{1}{2})]. \tag{23}$$

Combining Eqs. (22) and (23), and using again the definition of L, we have

$$[a_i^{(\pm)}, L(L+1)] = -a_i^{(\pm)} [1 \pm (2L+1)], \tag{24}$$

or, equivalently,

$$L(L+1)\mathbf{a}^{(\pm)} = \mathbf{a}^{(\pm)} (L \pm 1)(L \pm 1 + 1). \tag{25}$$

From the nonnegativty of L, it then follows that

$$L\mathbf{a}^{(\pm)} = \mathbf{a}^{(\pm)} (L \pm 1), \tag{26}$$

so that $\mathbf{a}^{(\pm)}$ is a vector shift-operator for L. We have from Eq. (20) that

$$\mathbf{a} = \mathbf{a}^{(+)} + \mathbf{a}^{(-)}, \tag{27}$$

which is the required resolution of $\mathbf{a}$.

In the same way we find

$$\mathbf{a}^\dagger = \mathbf{a}^{\dagger(+)} + \mathbf{a}^{\dagger(-)}, \tag{28}$$

with

$$\begin{aligned} a_j^{\dagger(\pm)} &= a_i^\dagger [(L + \tfrac{1}{2})\delta_{ij} \pm \tfrac{1}{2}\delta_{ij} \mp i l_{ij}][2L+1]^{-1}, \\ i a_j^{\dagger(\pm)} l_{ji} &= a_i^{\dagger(\pm)} [\tfrac{1}{2} \mp (L + \tfrac{1}{2})], \end{aligned} \tag{29}$$

and

$$L\mathbf{a}^{\dagger(\pm)} = \mathbf{a}^{\dagger(\pm)} (L \pm 1). \tag{30}$$

In Appendix A we prove that

$$[\mathbf{a}^{(\pm)}]^\dagger = \mathbf{a}^{\dagger(\mp)}. \tag{31}$$

Now

$$N\mathbf{a} = \mathbf{a}(N-1), \tag{32}$$

and since N commutes with l_{ij} and L, it follows from the definition (20) that

$$N\mathbf{a}^{(\pm)} = \mathbf{a}^{(\pm)}(N-1). \tag{33}$$

It then follows from Eq. (26) and the definition (15) of K, that

$$\begin{aligned} [K, \mathbf{a}^{(-)}] &= 0, \\ K\mathbf{a}^{(+)} &= \mathbf{a}^{(+)}(K-1). \end{aligned} \tag{34}$$

In a similar way [or by conjugation of Eqs. (34)] we deduce that

$$\begin{aligned} [K, \mathbf{a}^{\dagger(+)}] &= 0, \\ K\mathbf{a}^{\dagger(-)} &= \mathbf{a}^{\dagger(-)}(K+1). \end{aligned} \tag{35}$$

It is easily seen from Eqs. (3) that

$$N(\mathbf{a}\cdot\mathbf{a}) = (\mathbf{a}\cdot\mathbf{a})(N-2), \tag{36}$$

and that $(\mathbf{a}\cdot\mathbf{a})$, being a scalar, commutes with L. Hence, using Eq. (15) we have

$$K(\mathbf{a}\cdot\mathbf{a}) = (\mathbf{a}\cdot\mathbf{a})(K-1), \tag{37a}$$

and, by a similar argument

$$K(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) = (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(K+1). \tag{37b}$$

It now follows that $\mathbf{a}^{(+)}$ and $\mathbf{a}^{\dagger(+)}(\mathbf{a}\cdot\mathbf{a})$ have the same shifting properties for N, K, and L, so it is not surprising to find that (see Appendix A for proofs)

$$\mathbf{a}^{(+)} = \mathbf{a}^{\dagger(+)}(\mathbf{a}\cdot\mathbf{a})(2K+2L+1)^{-1}, \tag{38a}$$

and similarly

$$\mathbf{a}^{\dagger(-)} = \mathbf{a}^{(-)}(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(2K+2L+3)^{-1}. \tag{38b}$$

We therefore isolate as fundamental the operator $\mathbf{a}^{\dagger(+)}$ and its conjugate $\mathbf{a}^{(-)}$, which are raising and lowering operators for L, but which commute with K; and the operators $(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)$ and $(\mathbf{a}\cdot\mathbf{a})$, which are raising and lowering operators for K but which commute with L.
[The operators $\mathbf{a}^{(+)}$ and $\mathbf{a}^{\dagger(-)}$ are relegated to a secondary position, and they may be regarded as defined by Eqs. (38).] However, the operators $\boldsymbol{\lambda}$, ν, and their conjugates $\boldsymbol{\lambda}^\dagger$, $\nu^\dagger$, defined by

$$\begin{aligned} \boldsymbol{\lambda} &= \mathbf{a}^{(-)} f(K,L) = f(K,L+1)\mathbf{a}^{(-)}, \\ \boldsymbol{\lambda}^\dagger &= f(K,L)\mathbf{a}^{\dagger(+)} = \mathbf{a}^{\dagger(+)} f(K,L+1), \\ \nu &= (\mathbf{a}\cdot\mathbf{a}) g(K,L) = g(K+1,L)(\mathbf{a}\cdot\mathbf{a}), \\ \nu^\dagger &= g(K,L)(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) = (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) g(K+1,L), \end{aligned} \tag{39}$$

may equally well be regarded as fundamental, for any reasonable Hermitian operator functions f and g. They evidently have the same shifting properties for K and L as have $\mathbf{a}^{(-)}$, $\mathbf{a}^{\dagger(+)}$, $(\mathbf{a}\cdot\mathbf{a})$, and $(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)$, respectively, viz

$$\begin{aligned} &L\boldsymbol{\lambda} = \boldsymbol{\lambda}(L-1), \quad L\boldsymbol{\lambda}^\dagger = \boldsymbol{\lambda}^\dagger(L+1), \\ &[L,\nu] = 0 = [L,\nu^\dagger], \\ &K\nu = \nu(K-1), \quad K\nu^\dagger = \nu^\dagger(K+1), \\ &[K,\boldsymbol{\lambda}] = 0 = [K,\boldsymbol{\lambda}^\dagger]. \end{aligned} \tag{40}$$

Furthermore, for any f we have (see Appendix A)

$$\begin{aligned} &[\lambda_i, \lambda_j] = 0 = [\lambda_i^\dagger, \lambda_j^\dagger], \\ &\boldsymbol{\lambda}\cdot\boldsymbol{\lambda} = 0 = \boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda}^\dagger, \end{aligned} \tag{41}$$

$$\begin{aligned} i\lambda_k l_{ki} &= \lambda_i(L+1), \\ -i\lambda_k^\dagger l_{ki} &= \lambda_i^\dagger L, \end{aligned} \tag{42}$$

and also

$$\begin{aligned} i[\lambda_i, l_{jk}] &= \delta_{ij}\lambda_k - \delta_{ik}\lambda_j, \\ i[\lambda_i^\dagger, l_{jk}] &= \delta_{ij}\lambda_k^\dagger - \delta_{ik}\lambda_j^\dagger. \end{aligned} \tag{43}$$

We choose the functions f and g so that, in addition to Eqs. (40), (41), (42), and (43), the operators $\boldsymbol{\lambda}$, $\boldsymbol{\lambda}^\dagger$, ν, and $\nu^\dagger$ have other simple algebraic properties, which make them most useful for the solution of the problem at hand (and for the construction of coherent states–see the comments at the end of Sec. 5). Noting that $2K+2L+1(=2N+L+1)$ is positive definite, and so has well-defined negative powers, we take

$$f = [(2L+1)/(2K+2L+1)]^{1/2}, \qquad g = (4K+4L+2)^{-1/2}, \tag{44}$$

and find (see Appendix A)

$$\begin{aligned} &[\nu, \nu^\dagger] = 1, \\ &[\lambda_i, \nu] = 0 = [\lambda_i^\dagger, \nu^\dagger], \\ &[\lambda_i, \nu^\dagger] = 0 = [\lambda_i^\dagger, \nu], \\ &(2\boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda}+1)[\lambda_i, \lambda_j^\dagger] = (2\boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda}+1)\delta_{ij} - 2\lambda_i^\dagger\lambda_j, \end{aligned} \tag{45}$$

and also

$$K = \nu^\dagger\nu, \quad L = \boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda}, \qquad il_{ij} = \lambda_i^\dagger\lambda_j - \lambda_j^\dagger\lambda_i. \tag{46}$$

The definition of the operators $\boldsymbol{\lambda}$, $\boldsymbol{\lambda}^\dagger$, ν, and $\nu^\dagger$ in terms of $\mathbf{a}$ and $\mathbf{a}^\dagger$, as presented above, is rather complicated. However one may choose to regard them, rather than $\mathbf{a}$ and $\mathbf{a}^\dagger$, as the basic variables. Then Eqs. (15) and (46) become definitions of N, K, L, and l_{ij}, and it can be shown that all relations in the algebra, such as those in Eqs. (40), (42), and (43), follow from Eqs. (41) and (45). In particular, $\mathbf{a}$ and $\mathbf{a}^\dagger$, which from this point of view have the complicated *definitions*

$$\begin{aligned} \mathbf{a} &= \boldsymbol{\lambda}[(2K+2L+1)/(2L+1)]^{1/2} + \boldsymbol{\lambda}^\dagger\nu[2/(2L+3)]^{1/2}, \\ \mathbf{a}^\dagger &= \boldsymbol{\lambda}^\dagger[(2K+2L+3)/(2L+3)]^{1/2} + \boldsymbol{\lambda}\nu^\dagger[2/(2L+1)]^{1/2}, \end{aligned} \tag{47}$$

can be shown to satisfy the boson commutation relations (3).

The commutation relations satisfied by $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$ as given in Eqs. (41) and (45) make these operators more difficult to manipulate than the boson operators ν and $\nu^\dagger$. However, the last of Eqs. (45), although complicated in appearance, has an important property in common with boson commutation relations: It does permit an annihilation operator λ_i to be shuffled through a product of creation operators $\lambda_j^\dagger$ acting on a "vacuum" vector, with the accumulation of terms which are free of annihilation operators. Using these operators we are able to solve the common eigenvalue problem for K, L,an L_3 in a manner quite similar to that usually adopted for N_1, N_2, and N_3.

The algebraic relations satisfied by the operators $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$ as listed above, are the same as those satisfied by the "modified boson operators" introduced in a quite different context by Lohe and Hurst.[13] Accordingly the algebraic structure of the eigenvectors $|k\,l\,m\rangle$ defined in the next section, in so far as it involves the variables $\boldsymbol{\lambda}^\dagger$, is essentially the same as the structure of the vectors $|^l_m\rangle$ of Ref. 13.

However, there is an important difference between the two sets of operators (apart from the fact that no analogs of ν and $\nu^\dagger$ appear in the work of Lohe and Hurst). The operators $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$ have been defined in terms of boson operators $\mathbf{a}$ and $\mathbf{a}^\dagger$ and act in the same space as those operators. While this space can be taken to be that of the usual coordinate representation of quantum mechanics, $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$ have been defined in a representation-independent way, and are perhaps best thought of as acting in an abstract Hilbert space, not tied to any particular representation. In contrast, the operators of Lohe and Hurst are defined by modifying not only a set of boson operators $\mathbf{a}$ and $\mathbf{a}^\dagger$, but also the particular space in which they are taken to act. As a result, their modified boson operators are only defined in a space of harmonic functions of three variables. The reason that they satisfy the same algebraic relations as $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$ may be traced to the fact that equivalent representations of the Lie algebra so(3,2) underly the two structures. In our case this so(3,2) is a subalgebra of a spectrum-generating algebra for the oscillator (Sec. 5), whereas in the case of Lohe and Hurst, though not metioned by them, it arises as a well-known invariance algebra of Laplace's equation in three dimensions.

3. SOLUTION OF THE EIGENVALUE PROBLEM

Since K and L cannot have negative eigenvalues, we see at once that there must be a vector on which the lowering operators $\boldsymbol{\lambda}$ and ν vanish. Thus we assert the existence of a normalized vector $|0\rangle$ such that

$$\nu|0\rangle = 0 = \lambda_i|0\rangle, \quad i = 1,2,3. \tag{48}$$

Since $K = \nu^\dagger\nu$, $L = \boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda}$, and $il_{ij} = \lambda_i^\dagger\lambda_j - \lambda_j^\dagger\lambda_i$, we have

$$K\,|0\rangle = L\,|0\rangle = N\,|0\rangle = 0, \qquad l_{ij}|0\rangle = 0 = L_i|0\rangle, \quad i,j = 1,2,3. \tag{49}$$

The other common eigenvectors of K, L, and L_3 can now be built up by applying the raising operators $\nu^\dagger$ and $\boldsymbol{\lambda}^\dagger$ to this "vacuum" vector. We define

$$\lambda_\pm = (\lambda_1 \pm i\lambda_2), \quad \lambda_\pm^\dagger = (\lambda_1^\dagger \pm i\lambda_2^\dagger), \tag{50}$$

so that

$$L_3\lambda_\pm = \lambda_\pm(L_3 \pm \hbar), \qquad L_3\lambda_\pm^\dagger = \lambda_\pm^\dagger(L_3 \pm \hbar). \tag{51}$$

Now let k, r, and s run over the nonnegative integers independently and let ϵ denote either $+$ or $-$. Then it is evident that on the vector

$$(\nu^\dagger)^k(\lambda_\epsilon^\dagger)^r(\lambda_3^\dagger)^s|0\rangle, \tag{52}$$

K, L, and L_3 have the eigenvalues k, $r+s$, and $\epsilon r\hbar$, respectively. Setting l equal to $r+s$, and m equal to ϵr, we write

$$|k\,l\,m\rangle = c_{klm}(\nu^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle \tag{53}$$

as the normalized common eigenvector of these operators, corresponding to the eigenvalues k, l, and $m\hbar$. (We postpone for the moment discussion of the values of the normalization constants c_{klm}.) Here k and l run over the nonnegative integers independently, while for a given value of l, m runs over l, $l-1$,..., $-l$, and ϵ is the sign of m. It is easily shown that for fixed k and l, the $2l+1$ vectors $|k\,l\,m\rangle$ form the basis for an irreducible representation of the Lie algebra so(3) spanned by the operators l_{ij} of Eqs. (46) (cf Ref. 13). [Alternatively one can consider for any fixed k, the vectors

$$|k;\alpha,\beta,...,\tau\rangle = (\nu^\dagger)^k\lambda_\alpha^\dagger\lambda_\beta^\dagger...\lambda_\tau^\dagger|0\rangle, \tag{54}$$

where the subscripts $\alpha, \beta, ..., \tau$ are l in number, and run over 1, 2, 3 independently. These vectors form a rank-l tensor basis for this representation of so(3). Note that in view of Eqs. (41), this tensor is automatically symmetric and traceless.]

For completeness, it is necessary to show that the vectors $|k\, l\, m\rangle$ are, up to multiplication by constants, the only common eigenvectors of K, L, L_3 which one can construct by applying to $|0\rangle$ any operator in the algebra generated by ν, $\nu^\dagger$, $\boldsymbol{\lambda}$, and $\boldsymbol{\lambda}^\dagger$. To do that, it suffices to show that the subspace of all finite linear combinations of the vectors $|k\, l\, m\rangle$ is invariant under the action of ν, $\nu^\dagger$, λ_3, $\lambda_3^\dagger$, $\lambda_\pm$, and $\lambda_\pm^\dagger$; and this is true, for we see in Eqs. (59) below that when any of these operators is applied to any $|k\, l\, m\rangle$, a constant multiple of another such eigenvector is produced.

Turning to the calculation of the normalization constant c_{klm} in Eqs. (53), we see at once that

$$\langle <k\, l\, m|k\, l\, m\rangle = k!|c_{klm}|^2\langle 0|(\lambda_3)^{l-|m|} \times(\lambda_{-\epsilon})^{|m|}(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle\,. \quad (55)$$

Using the last of the relations (45), it is straightforward to show by induction that

$$\begin{aligned}\lambda_i\lambda_\alpha^\dagger\lambda_\beta^\dagger\lambda_\gamma^\dagger\cdots\lambda_\sigma^\dagger\lambda_\tau^\dagger|0\rangle &= \{(\delta_{i\alpha}\lambda_\beta^\dagger\lambda_\gamma^\dagger\cdots\lambda_\sigma^\dagger\lambda_\tau^\dagger + \delta_{i\beta}\lambda_\alpha^\dagger\lambda_\gamma^\dagger\cdots\lambda_\sigma^\dagger\lambda_\tau^\dagger + \cdots \\ &+ \delta_{i\tau}\lambda_\alpha^\dagger\lambda_\beta^\dagger\lambda_\gamma^\dagger\cdots\lambda_\sigma^\dagger) - \frac{2}{(2l-1)}\lambda_i^\dagger(\delta_{\alpha\beta}\lambda_\gamma^\dagger\cdots\lambda_\sigma^\dagger\lambda_\tau^\dagger \\ &+ \delta_{\alpha\gamma}\lambda_\beta^\dagger\cdots\lambda_\sigma^\dagger\lambda_\tau^\dagger + \cdots + \delta_{\alpha\tau}\lambda_\beta^\dagger\lambda_\gamma^\dagger\cdots\lambda_\sigma^\dagger \\ &+ \delta_{\beta\gamma}\lambda_\alpha^\dagger\cdots\lambda_\sigma^\dagger\lambda_\tau^\dagger + \cdots + \delta_{\beta\tau}\lambda_\alpha^\dagger\lambda_\gamma^\dagger\cdots\lambda_\sigma^\dagger \\ &+ \cdots + \delta_{\sigma\tau}\lambda_\alpha^\dagger\lambda_\beta^\dagger\lambda_\gamma^\dagger\cdots)\}|0\rangle\,, \end{aligned} \quad (56)$$

where l is the number of creation operators $\lambda_\alpha^\dagger, \lambda_\beta^\dagger, \ldots, \lambda_\tau^\dagger$. With the help of this result, we are able to show (see Appendix B) that

$$\langle k\, l\, m|k\, l\, m\rangle = \frac{k!|c_{klm}|^2 2^l l!(l-m)!(l+m)!}{(2l)!}\,, \quad (57)$$

so that $|k\, l\, m\rangle$ as defined in Eq. (53) is normalized if we take (with a convenient choice of phases)

$$c_{klm} = (-\epsilon)^m\left(\frac{(2l)!}{k!l!(l-m)!(l+m)!2^l}\right)^{1/2}. \quad (58)$$

It is then found that (see Appendix B)

$$\begin{aligned}\nu|k\, l\, m\rangle &= (k)^{1/2}|k-1\, l\, m\rangle\,, \\ \nu^\dagger|k\, l\, m\rangle &= (k+1)^{1/2}|k+1\, l\, m\rangle\,, \\ \lambda_3^\dagger|k\, l\, m\rangle &= \left(\frac{(l+1-m)\,(l+1+m)}{(2l+1)}\right)^{1/2}|k\, l+1\, m\rangle\,, \\ \lambda_\pm^\dagger|k\, l\, m\rangle &= \mp\left(\frac{(l\pm m+2)\,(l\pm m+1)}{(2l+1)}\right)^{1/2} \\ &\quad\times|k\, l+1\, m\pm 1\rangle\,, \\ \lambda_3|k\, l\, m\rangle &= \left(\frac{(l-m)\,(l+m)}{(2l-1)}\right)^{1/2}|k\, l-1\, m\rangle\,, \\ \lambda_\pm|k\, l\, m\rangle &= \pm\left(\frac{(l\mp m)\,(l\mp m-1)}{(2l-1)}\right)^{1/2} \\ &\quad\times|k\, l-1\, m\pm 1\rangle\,. \end{aligned} \quad (59)$$

We close this section by remarking that we have chosen phases in Eq. (58) in such a way that the vector $|k\, l\, m\rangle$ appears in the coordinate representation as

$$\phi_{klm} = (-1)^k\left(\frac{2a^3 k!}{\Gamma(k+l+3/2)}\right)^{1/2}\xi^l e^{-(1/2)\xi^2} \times L_k^{(l+1/2)}(\xi^2)Y_{lm}(\theta,\phi)\,,$$

where $a = (M\omega/\hbar)^{1/2}$, $\xi = ar$ (r,θ and ϕ are the usual spherical polar coordinates), $L_k^{(l+1/2)}$ is the generalized Laguerre polynomial, defined as in Ref. 14, and the spherical harmonic Y_{lm} is defined as in Ref. 15.

4. TIME-DEPENDENCE AND INTERPRETATION

In the Heisenberg picture, the time-dependence of an observable A (or of any complex linear combination A of observables) is determined by

$$i\hbar\frac{dA}{dt} = [A, H]\,.$$

Now the operators ν, $\nu^\dagger$, $\boldsymbol{\lambda}$, and $\boldsymbol{\lambda}^\dagger$ are shift-operators for H, allowing us to deduce at once that

$$\begin{aligned}\frac{d\nu}{dt} &= -2i\omega\nu\,, \quad \frac{d\nu^\dagger}{dt} = 2i\omega\nu^\dagger\,, \\ \frac{d\boldsymbol{\lambda}}{dt} &= -i\omega\boldsymbol{\lambda}\,, \quad \frac{d\boldsymbol{\lambda}^\dagger}{dt} = i\omega\boldsymbol{\lambda}^\dagger\,. \end{aligned} \quad (60)$$

Thus

$$\begin{aligned}\nu &= \nu_0 e^{-2i\omega t}\,, \quad \nu^\dagger = \nu_0^\dagger e^{2i\omega t}\,, \\ \boldsymbol{\lambda} &= \boldsymbol{\lambda}_0 e^{-i\omega t}\,, \quad \boldsymbol{\lambda}^\dagger = \boldsymbol{\lambda}_0^\dagger e^{i\omega t}\,, \end{aligned} \quad (61)$$

where the (constant) operators ν_0, $\nu_0^\dagger$, $\boldsymbol{\lambda}_0$, and $\boldsymbol{\lambda}_0^\dagger$ satisfy the same algebraic relations as ν, $\nu^\dagger$, $\boldsymbol{\lambda}$, and $\boldsymbol{\lambda}^\dagger$.

We gain some insight into the physical interpretation of these variables by considering their classical counterparts. Denoting the classical coordinate and momentum vectors by $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ respectively and the classical Hamiltonian by $\hat{H}$, we define

$$\hat{\mathbf{a}} = (2M\omega)^{-1/2}(i\hat{\mathbf{p}} + M\omega\hat{\mathbf{x}})\,, \quad (62)$$

and its *complex* conjugate $\hat{\mathbf{a}}^*$. [Note the extra factor of $(\hbar)^{1/2}$ in comparison with Eq. (2).] Then

$$\hat{H} = \omega\hat{\mathbf{a}}^*\cdot\hat{\mathbf{a}}\,. \quad (63)$$

Introducing the classical angular momentum vector

$$\hat{\mathbf{L}} = \hat{\mathbf{x}}\times\hat{\mathbf{p}}\,, \quad (64)$$

with length $\hat{L}$, we define

$$\hat{K} = \frac{1}{2\omega}(\hat{H} - \omega\hat{L})\,. \quad (65)$$

In the definitions of ν, $\nu^\dagger$, $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$ above, we let $\hbar\to 0$, with $H\to\hat{H}$ and

$$\begin{aligned}&(\hbar)^{1/2}\mathbf{a}\to\hat{\mathbf{a}}\,, \quad (\hbar)^{1/2}\mathbf{a}^\dagger\to\hat{\mathbf{a}}^*\,, \\ &\hbar l_{ij}\to\epsilon_{ijk}\hat{L}_k\,, \\ &\hbar L\to\hat{L}\,, \quad \hbar K\to\hat{K}\,, \end{aligned} \quad (66)$$

in order to obtain the classical variables corresponding to ν and $\boldsymbol{\lambda}$,

$$\hat{\nu} = \tfrac{1}{2}(\hat{\mathbf{a}}\cdot\hat{\mathbf{a}})\,(\hat{K} + \hat{L})^{-1/2}$$

$$= \frac{(\hat{K}+\hat{L})^{-1/2}}{4M\omega}[(M^2\omega^2\hat{\mathbf{x}}^2-\hat{\mathbf{p}}^2)+i(2M\omega\hat{\mathbf{x}}\cdot\hat{\mathbf{p}})]\,, \tag{67}$$

$$\begin{aligned}\hat{\boldsymbol{\lambda}} &= \tfrac{1}{2}[\hat{L}(\hat{K}+\hat{L})]^{-1/2}(\hat{L}\hat{\mathbf{a}}+i\hat{\mathbf{L}}\times\hat{\mathbf{a}})\\ &= [8M\omega\hat{L}(\hat{K}+\hat{L})]^{-1/2}[(M\omega\hat{L}\hat{\mathbf{x}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}})\\ &\quad + i(\hat{\mathbf{L}}\hat{\mathbf{p}}+M\omega\hat{\mathbf{L}}\times\hat{\mathbf{x}})]\,,\end{aligned}$$

and their complex conjugates $\hat{v}^*$ and $\hat{\boldsymbol{\lambda}}^*$, which correspond to $v^\dagger$ and $\boldsymbol{\lambda}^\dagger$. Apart from the overall factors involving the constants of the motion $\hat{K}$ and $\hat{L}$, these expressions are reasonably simple. It is straightforward to verify in particular that

$$\begin{aligned}&\hat{v}^*\hat{v}=\hat{K}\,,\quad \hat{\boldsymbol{\lambda}}^*\cdot\hat{\boldsymbol{\lambda}}=\hat{L}\,,\\ &\hat{\boldsymbol{\lambda}}\cdot\hat{\boldsymbol{\lambda}}=0=\hat{\boldsymbol{\lambda}}^*\cdot\hat{\boldsymbol{\lambda}}^*\,,\\ &\hat{H}=\omega(2\hat{v}^*\hat{v}+\hat{\boldsymbol{\lambda}}^*\cdot\hat{\boldsymbol{\lambda}})\,,\\ &\hat{\boldsymbol{\lambda}}\times\hat{\mathbf{L}}=i\hat{L}\hat{\boldsymbol{\lambda}}\,,\quad \hat{\boldsymbol{\lambda}}^*\times\hat{\mathbf{L}}=-i\hat{L}\hat{\boldsymbol{\lambda}}^*\,,\\ &\hat{\boldsymbol{\lambda}}^*\times\hat{\boldsymbol{\lambda}}=i\hat{\mathbf{L}}\,,\end{aligned} \tag{68}$$

and also that the time-dependence of the classical variables is the same as that of their counterparts, as in Eqs. (61).

From the relations (68), it can be seen that if $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ denote the real and imaginary parts of $(\sqrt{2})\hat{\boldsymbol{\lambda}}$, then $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ are orthogonal, and of the same length $(\hat{L})^{1/2}$. From the last three of the relations (68), we see then that $\mathbf{e}=\hat{\boldsymbol{\alpha}}(\hat{L})^{-1/2}$, $\mathbf{f}=\boldsymbol{\beta}(\hat{L})^{-1/2}$, and $\mathbf{g}=\hat{\mathbf{L}}(\hat{L})^{-1}$ form a right-handed system of orthogonal unit vectors, of which the first two are time-dependent.

Classically, the motion is elliptical, in the plane perpendicular to $\hat{\mathbf{L}}$, i.e., in the plane determined by $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$. For any particular motion we can choose time-origin and space-axes such that the motion is anticlockwise in the XY-plane, with

$$\hat{\mathbf{x}}=(A\cos\omega t, B\sin\omega t, 0),\quad A\geqslant B\geqslant 0\,.$$

Then

$$\begin{aligned}&\hat{\mathbf{p}}=M\omega(-A\sin\omega t, B\cos\omega t, 0)\,,\\ &\hat{\mathbf{L}}=M\omega(0,0,AB)\,,\quad \hat{L}=M\omega AB\,,\\ &\hat{H}=\tfrac{1}{2}M\omega^2(A^2+B^2)\,,\quad \hat{K}=\tfrac{1}{4}M\omega(A-B)^2\,,\\ &\hat{v}=\tfrac{1}{2}\sqrt{M\omega}(A-B)e^{-2i\omega t}=\sqrt{\hat{K}}\,e^{-2i\omega t}\,,\\ &\hat{\boldsymbol{\lambda}}=\sqrt{\tfrac{1}{2}M\omega AB}\,e^{-i\omega t}(1,i,0)=\sqrt{\tfrac{1}{2}\hat{L}}\,e^{-i\omega t}(1,i,0)\,,\\ &\mathbf{e}=(\cos\omega t,\sin\omega t,0)\,,\quad \mathbf{f}=(-\sin\omega t,\cos\omega t,0)\,,\\ &\mathbf{g}=(0,0,1)\,.\end{aligned} \tag{69}$$

The periodic variables $\hat{\boldsymbol{\lambda}}$ and $\hat{\boldsymbol{\lambda}}^*$ have angular frequency ω, the natural frequency of the oscillator, but $\hat{v}$ and $\hat{v}^*$ have angular frequency 2ω. This is at first glance rather puzzling, but we can understand it as follows, and perhaps at the same time appreciate the geometrical significance of all these variables.

The elliptical motion can be regarded as arising from the superposition of two uniform circular motions, with angular frequencies 2ω and ω, respectively. In the particular coordinate system adopted above,

$$\begin{aligned}\hat{\mathbf{x}} &= (A\cos\omega t,\ B\sin\omega t,\ 0)\\ &= \tfrac{1}{2}(A+B)\mathbf{e}+\tfrac{1}{2}(A-B)(\cos 2\omega t\,\mathbf{e}-\sin 2\omega t\,\mathbf{f})\,.\end{aligned} \tag{70}$$

Thus the particle can be regarded as moving uniformly *clockwise*, with angular frequency 2ω, in a circle which is fixed relative to the vectors $\mathbf{e}$ and $\mathbf{f}$ and has radius $\tfrac{1}{2}(A-B)$. This circle (epicycle) itself rotates *anticlockwise* (along with $\mathbf{e}$ and $\mathbf{f}$) about an exterior point 0, so that its center moves uniformly, with angular frequency ω, around the circumference of a larger circle (deferent) of radius $\tfrac{1}{2}(A+B)$, centered at 0. The point 0 is the center of the resultant anticlockwise elliptical motion. (See Fig. 1.)

The relevance of the variables $\hat{v}$, $\hat{v}^*$, $\hat{\boldsymbol{\lambda}}$, and $\hat{\boldsymbol{\lambda}}^*$ to this decomposition of the motion can be appreciated when one notes that Eq. (70) is [in the particular coordinate system of Eqs. (69)] just the real part of the formula

$$\hat{\mathbf{a}}=[(\hat{K}+\hat{L})/\hat{L}]^{1/2}\hat{\boldsymbol{\lambda}}+[\hat{L}]^{-1/2}\hat{v}\hat{\boldsymbol{\lambda}}^*\,, \tag{71}$$

which is the classical equivalent of the first of Eqs. (47).

The resolution of the harmonic motion into two circular motions can also be seen and understood in the following way. The equation of motion for the oscillator is

$$m\frac{d^2\hat{\mathbf{x}}}{dt^2}=-m\omega^2\hat{\mathbf{x}}\,. \tag{72}$$

Since the force on the particle is central, the motion is in a fixed plane perpendicular to the angular momentum $\hat{\mathbf{L}}$. We make a change of reference frame, to the frame rotating anticlockwise, with angular frequency ω, about a unit vector $\mathbf{n}$ which passes through the origin and which is parallel to $\hat{\mathbf{L}}$. In the rotating frame, the equation of motion for the particle at $\mathbf{r}$ is

$$m\frac{d^2\mathbf{r}}{dt^2}=-m\omega^2\mathbf{r}-2m\boldsymbol{\omega}\times\frac{d\mathbf{r}}{dt}-m\,\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})\,, \tag{73}$$

with $\boldsymbol{\omega}=\omega\mathbf{n}$. Here the second term is the Coriolis "force," and the third is the centrifugal "force" on the particle

Now $\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})=-\omega^2\mathbf{r}$,

since $\boldsymbol{\omega}$ is orthogonal to $\hat{\mathbf{x}}$, and hence to $\mathbf{r}$. Thus in this frame the centrifugal force exactly cancels the true force, and the particle moves under the Coriolis force alone, with

$$\frac{d^2\mathbf{r}}{dt^2}=-2\boldsymbol{\omega}\times\frac{d\mathbf{r}}{dt}. \tag{74}$$

Then

$$\begin{aligned}\frac{d^3\mathbf{r}}{dt^3} &= -2\boldsymbol{\omega}\times\frac{d^2\mathbf{r}}{dt^2}\\ &= -2\boldsymbol{\omega}\times\left(-2\boldsymbol{\omega}\times\frac{d\mathbf{r}}{dt}\right)\\ &= -4\omega^2\frac{d\mathbf{r}}{dt}.\end{aligned} \tag{75}$$

Integrating once we have

$$\frac{d^2\mathbf{R}}{dt^2}=-4\omega^2\mathbf{R}\,, \tag{76}$$

where $\mathbf{R}=\mathbf{r}-\mathbf{r}_0$, with $\mathbf{r}_0$ an arbitrary constant vector, which must be orthogonal to $\boldsymbol{\omega}$ in view of Eq. (74).

We see that in this rotating frame the motion is harmonic with angular frequency 2ω, about an arbitrary fixed point $\mathbf{r}_0$ in the plane of the motion. That this motion must actually be circular (it is the motion around the smaller circle in Fig. 1.) follows from the fact that we also have, from Eq. (74)

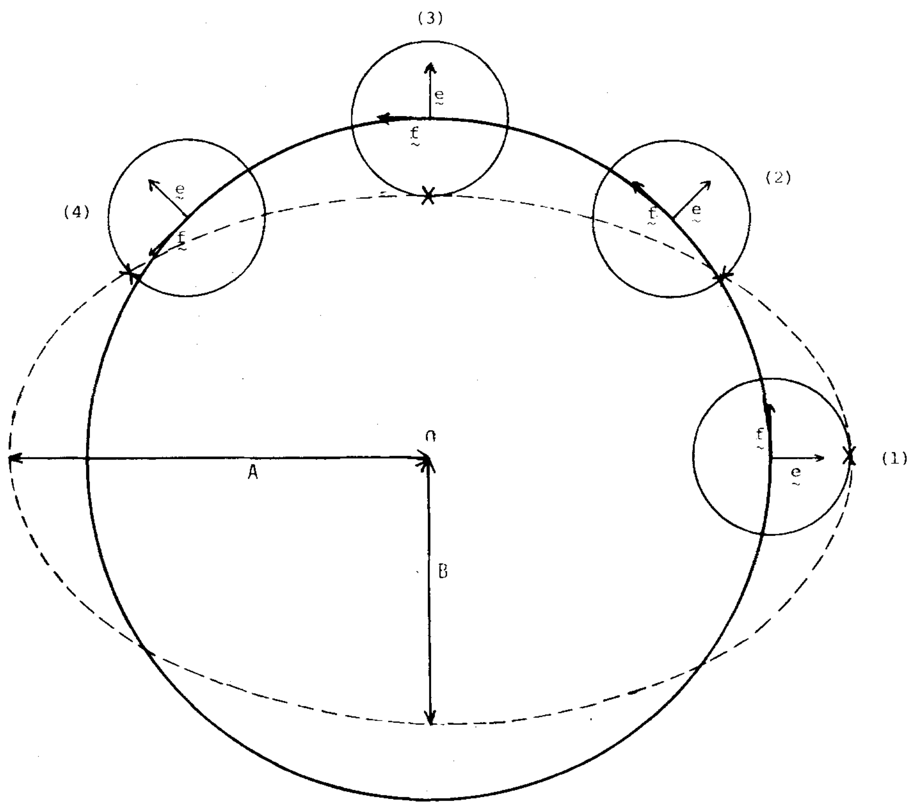


FIG. 1. Resolution of elliptical motion into two circular motions. The particle, whose position is marked X, rotates clockwise with angular frequency 2ω around the smaller circle, whose center moves anticlockwise with angular frequency ω around the larger circle. Positions are shown at (1) $\omega t = 0$; (2) $\omega t = \pi/4$; (3) $\omega t = \pi/2$; (4) $\omega t = 3\pi/4$.

$$\frac{d^2\mathbf{R}}{dt^2} = -2\boldsymbol{\omega}\times\frac{d\mathbf{R}}{dt}, \tag{77}$$

so that

$$-4\omega^2\mathbf{R} = -2\boldsymbol{\omega}\times\frac{d\mathbf{R}}{dt}, \tag{78}$$

implying that $\mathbf{R}\cdot(d\mathbf{R}/dt) = 0$. and hence that $\mathbf{R}^2$ is constant. We also note from Eq. (78) that

$$\begin{aligned}\mathbf{R}\times\frac{d\mathbf{R}}{dt} &= \frac{1}{2\omega^2}\left(\boldsymbol{\omega}\times\frac{d\mathbf{R}}{dt}\right)\times\frac{d\mathbf{R}}{dt} \\ &= -\frac{1}{2\omega^2}\left(\frac{d\mathbf{R}}{dt}\right)^2\boldsymbol{\omega},\end{aligned} \tag{79}$$

so that the circular motion is in the opposite sense to $\boldsymbol{\omega}$, i.e., it is clockwise about an axis which passes through $\mathbf{R} = \mathbf{0}$ and which is parallel to $\mathbf{n}$.

5. A NEW SPECTRUM-GENERATING ALGEBRA FOR THE OSCILLATOR

For the treatment of the common eigenvalue problem for the operators N_1, N_2, and N_3, a spectrum-generating (Lie) algebra is the 21-dimensional Lie algebra sp(6,R), with Hermitian basis

$$\begin{aligned}&(a_i a_j + a_i^\dagger a_j^\dagger),\quad i(a_i a_j - a_i^\dagger a_j^\dagger),\\ &(a_i^\dagger a_j + a_i a_j^\dagger),\quad i(a_i^\dagger a_j - a_j^\dagger a_i).\end{aligned} \tag{80}$$

The vectors $|n_1, n_2, n_3\rangle$ with odd $(n_1 + n_2 + n_3)$ span one irreducible representation of this algebra, and those with even $(n_1 + n_2 + n_3)$ span another.[12]

For the common eigenvalue problem for N, $\mathbf{L}^2$, and L_3 another Lie algebra is more relevant. Define the Hermitian operators

$$\begin{aligned}\boldsymbol{\Lambda} &= (2L+1)^{1/2}\boldsymbol{\lambda} = \boldsymbol{\lambda}(2L-1)^{1/2},\\ \boldsymbol{\Lambda}^\dagger &= \boldsymbol{\lambda}^\dagger(2L+1)^{1/2} = (2L-1)^{1/2}\boldsymbol{\lambda}^\dagger,\end{aligned} \tag{81}$$

and note that, as well as commuting with ν and $\nu^\dagger$, and having the same shifting properties for L as $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$, they satisfy

$$\begin{aligned}&[\Lambda_i, \Lambda_j] = 0 = [\Lambda_i^\dagger, \Lambda_j^\dagger],\\ &[\Lambda_i, \Lambda_j^\dagger] = (2L+1)\delta_{ij} - 2il_{ij},\\ &\boldsymbol{\Lambda}\cdot\boldsymbol{\Lambda} = 0 = \boldsymbol{\Lambda}^\dagger\cdot\boldsymbol{\Lambda}^\dagger,\\ &\boldsymbol{\Lambda}^\dagger\cdot\boldsymbol{\Lambda} = L(2L-1),\\ &\Lambda_i^\dagger\Lambda_j - \Lambda_j^\dagger\Lambda_i = il_{ij}(2L-1).\end{aligned} \tag{82}$$

The proof of the results (82) is elementary, with the use of Eqs. (41), (45), and (46).

Now define

$$\begin{aligned}&A_1 = \tfrac{1}{4}(\nu\nu + \nu^\dagger\nu^\dagger),\quad A_2 = \tfrac{1}{4}i(\nu\nu - \nu^\dagger\nu^\dagger),\\ &A_3 = \tfrac{1}{2}(\nu^\dagger\nu + \tfrac{1}{2}),\\ &B_{4i} = \tfrac{1}{2}(\Lambda_i + \Lambda_i^\dagger) = -B_{i4},\\ &B_{5i} = \tfrac{1}{2}i(\Lambda_i - \Lambda_i^\dagger) = -B_{i5},\\ &B_{ij} = l_{ij},\quad B_{54} = (L + \tfrac{1}{2}) = -B_{45}.\end{aligned} \tag{83}$$

It is easily checked that these operators span an Hermitian representation of the Lie algebra so(2,1) $\oplus$ so(3,2) [$\simeq$sp(2,R) $\oplus$ sp(4,R)], with the only nontrivial commutation relations being

$$[A_1, A_2] = -iA_3,\quad [A_2, A_3] = iA_1,\quad [A_3, A_1] = iA_2, \tag{84}$$

$$i[B_{\mu\nu}, B_{\rho\sigma}] = g_{\nu\rho}B_{\mu\sigma} + g_{\mu\sigma}B_{\nu\rho} - g_{\mu\rho}B_{\nu\sigma} - g_{\nu\sigma}B_{\mu\rho} ,$$

where μ, ν, ρ, and σ run over 1,2,3,4,5, and the metric tensor $g_{\mu\nu}$ is diagonal, with $g_{11} = g_{22} = g_{33} = -g_{44} = -g_{55} = 1$.

The quadratic invariant of the so(2,1) algebra has the value

$$(A_1)^2 + (A_2)^2 - (A_3)^2 = \tfrac{3}{16} . \tag{85}$$

There are two irreducible Hermitian representations of so(2,1), labelled $\mathscr{D}^{(+)}(-1/4)$ and $\mathscr{D}^{(+)}(-3/4)$ by Barut and Fronsdal,[16] for which the invariant has this value, and in which the spectrum of A_3 is bounded below (as it evidently is in the present situation). In the representation $\mathscr{D}^{(+)}(-1/4)$, A_3 has eigenvalues 1/4, 5/4, 9/4, ⋯; and in the representation $\mathscr{D}^{(+)}(-3/4)$ it has eigenvalues 3/4, 7/4, 11/4, ⋯. It can be seen that representations of both types are involved in the problem under discussion—the former associated with even-integral eigenvalues of $K\ (= \nu^\dagger \nu)$, the latter with odd-integral eigenvalues.

A simple calculation shows that the quadratic invariant of the so(3,2) algebra has the value

$$\tfrac{1}{2}B_{\mu\nu}B^{\mu\nu} = -\tfrac{5}{4} . \tag{86}$$

Moreover, the two invariants of the so(3,1) subalgebra spanned by the B_{ij} and B_{4i}, have the values

$$\tfrac{1}{2}B_{ij}B_{ij} - B_{4i}B_{4i} = -\tfrac{3}{4} ,$$
$$\tfrac{1}{2}\epsilon_{ijk}B_{ij}B_{4k} = 0 , \tag{87}$$

indicating that any irreducible representation of so(3,2) which appears here, remains irreducible when restricted to the so(3,1) subalgebra. In the commonly used[17] $[k_0, c]$ labelling of the irreducible representation of so(3,1), these two invariants have values $(k_0^2 + c^2 - 1)$ and ik_0c, respectively. Thus the irreducible representations of so(3,1) appearing here can only be $[\tfrac{1}{2}, 0]$ or $[0, \tfrac{1}{2}]$; and since the eigenvalues of B_{12} are integral, only the representation $[0, \tfrac{1}{2}]$ can be involved. It is known (see for example Böhm,[18]) that this representation of so(3,1) extends to either of two irreducible Hermitian representations (two of the four Majorana representations) of so(3,2), each consistent with Eq. (86). But in only one of these—let us call it $\mathscr{T}$—is the spectrum of B_{54} bounded below, as it evidently is in the present situation. In this representation $\mathscr{T}$, B_{54} has eigenvalues 1/2, 3/2, 5/2⋯.

The representation of so(2,1) ⊕ so(3,2) associated with the harmonic oscillator in an angular momentum basis can now be identified, in view of the nondegeneracy of the eigenvectors $|k\,l\,m\rangle$, as simply

$$(\mathscr{D}^{(+)}(-1/4), \mathscr{T}) \oplus (\mathscr{D}^{(+)}(-3/4), \mathscr{T}) . \tag{88}$$

The Hamiltonian operator appears in the form

$$H = \hbar\omega(4A_3 + B_{54}) , \tag{89}$$

and its eigenvalues are immediately deducible from the known spectra of A_3 in the representations $\mathscr{D}^{(+)}(-1/4)$, $\mathscr{D}^{(+)}(-3/4)$, and of B_{54} in $\mathscr{T}$.

The reader may wonder why we did not, in Sec. 2, choose to work with the operators $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^\dagger$ rather than $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$. A simple change of the function f in Eqs. (44) {to $f(K,L) = [(4L^2-1)/(2K+2L+1)]^{1/2}$} would have accomplished such a substitution. The commutation relations satisfied by $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^\dagger$ are simpler than those satisfied by $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$, and the connection with the spectrum-generating algebra is more immediate. For these reasons it may be argued that the operators $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^\dagger$ are more suitable for the algebraic treatment of the eigenvalue problem.

Our preference for the operators $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$ is mainly determined by our intention to define in a subsequent publication, "coherent angular momentum states" for the oscillator as eigenvectors of the lowering operators. The expectation values of the important operators H, K, L, and l_{ij} will be very simple in such states, if we diagonalize the operators $\boldsymbol{\lambda}$ and ν, because

$$K = \nu^\dagger\nu , \quad L = \boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda} , \quad il_{ij} = \lambda_i^\dagger\lambda_j - \lambda_j^\dagger\lambda_i . \tag{90}$$

On the other hand, if we diagonalize the operators $\mathbf{\Lambda}$, we shall need to work with the expressions

$$L = \tfrac{1}{4} + \tfrac{1}{4}(1 + 8\underline{\Lambda}^\dagger\cdot\underline{\Lambda})^{1/2} ,$$
$$il_{ij} = (2L-1)^{-1}(\Lambda_i^\dagger\Lambda_j - \Lambda_j^\dagger\Lambda_i) , \tag{91}$$

whose expectation values will not be simple.

ACKNOWLEDGMENTS

We thank Professor L. Bass for very helpful discussions, and for a critical reading of the manuscript. A.J.B. wishes to thank Professor H.S. Green for impressing on him the possibility of a purely algebraic approach to this and other problems in quantum mechanics.

APPENDIX A

Here and in Appendix B we present the derivations of some of the results stated above. The ends of proofs are indicated thus: □

From the definition of $\mathbf{a}^{\dagger(+)}$ in the first of Eqs. (29), it follows that

$$(a_j^{\dagger(+)})^\dagger = (2L+1)^{-1}[(L+1)\delta_{ij} + il_{ij}]a_i . \tag{A1}$$

Using Eqs. (21) and then Eqs. (23) we see that

$$\begin{aligned} il_{ij}a_i^{(\pm)} &= ia_i^{(\pm)}l_{ij} - 2a_j^{(\pm)} \\ &= -a_j^{(\pm)}[3/2 \pm (L + \tfrac{1}{2})] . \end{aligned} \tag{A2}$$

Now using Eq. (27) in Eq. (A1), we have

$$\begin{aligned} (a_j^{\dagger(+)})^\dagger &= (2L+1)^{-1}(L+1)(a_j^{(+)} + a_j^{(-)}) \\ &\quad + (2L+1)^{-1}(il_{ij}a_i^{(+)} + il_{ij}a_i^{(-)}) \\ &= (2L+1)^{-1}[a_j^{(+)}(L+2) + a_j^{(-)}L \\ &\quad - a_j^{(+)}(L+2) + a_j^{(-)}(L-1)] \quad \text{[using Eq. (A2)]} \\ &= (2L+1)^{-1}a_j^{(-)}(2L-1) \\ &= a_j^{(-)} \quad \text{[using Eq. (26)]}. \end{aligned}$$

In a similar way we show that $(\mathbf{a}^{\dagger(-)})^\dagger = \mathbf{a}^{(+)}$, so completing the verification of Eqs. (31). □

From the definition in Eqs. (20) we have that

$$\begin{aligned} a_j^{(-)}(2L+1) &= a_jL + ia_il_{ij} \\ &= a_jL - a_i(a_ia_j^\dagger - a_ja_i^\dagger) \quad \text{[using Eq. (9)]} , \end{aligned}$$

$$= -a_j^\dagger(\mathbf{a}\cdot\mathbf{a}) + a_j(N+L+1) \quad \text{[using Eqs. (3)]}. \qquad (A3)$$

Similarly, we find

$$\begin{aligned} a_j^{(+)}(2L+1) &= a_j^\dagger(\mathbf{a}\cdot\mathbf{a}) - a_j(N-L), \\ a_j^{\dagger(-)}(2L+1) &= a_j(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) - a_j^\dagger(N-L+2), \\ a_j^{\dagger(+)}(2L+1) &= -a_j(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) + a_j^\dagger(N+L+3). \end{aligned} \qquad (A4)$$

It is easily deduced from the definitions (9) and the relations (3) that

$$\begin{aligned} \tfrac{1}{2}l_{ij}l_{ij} &= N^2 + N - (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(\mathbf{a}\cdot\mathbf{a}) \\ &= N^2 + 5N + 6 - (\mathbf{a}\cdot\mathbf{a})(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger). \end{aligned}$$

Since $\frac{1}{2}l_{ij}l_{ij} = L(L+1)$, it follows that

$$\begin{aligned} (\mathbf{a}\cdot\mathbf{a})(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) &= (N-L+2)(N+L+3), \\ (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(\mathbf{a}\cdot\mathbf{a}) &= (N-L)(N+L+1). \end{aligned} \qquad (A5)$$

Multiplying on the right by $(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)$ in Eq. (A3), we get

$$\begin{aligned} a_j^{(-)}&(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(2L+1) \\ &= -a_j^\dagger(\mathbf{a}\cdot\mathbf{a})(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) + a_j(N+L+1)(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) \\ &= -a_j^\dagger(N-L+2)(N+L+3) + a_j(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) \\ &\quad \times(N+L+3) \text{ [using Eqs. (A5) and (3)]} \\ &= a_j^{\dagger(-)}(2L+1)(N+L+3) \text{ [using Eqs. (A4)].} \end{aligned}$$

Thus

$$\mathbf{a}^{(-)}(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) = \mathbf{a}^{\dagger(-)}(N+L+3),$$

and in a similar way we show that

$$\mathbf{a}^{\dagger(+)}(\mathbf{a}\cdot\mathbf{a}) = \mathbf{a}^{(+)}(N+L+1),$$

establishing Eqs. (38). □

Consider the product $\lambda_i\lambda_j$, with $\boldsymbol{\lambda}$ having the general form given in Eqs. (39):

$$\begin{aligned} \lambda_i\lambda_j &= a_i^{(-)}f(K,L)a_j^{(-)}f(K,L) \\ &= f(K,L+1)f(K,L+2)a_i^{(-)}a_j^{(-)}. \end{aligned} \qquad (A6)$$

From Eq. (A3) we have (recalling that $N = 2K + L$)

$$\begin{aligned} a_i^{(-)}&(2L+1)a_j^{(-)}(2L+1) \\ &= [-a_i^\dagger(\mathbf{a}\cdot\mathbf{a}) + a_i(2K+2L+1)]a_j^{(-)}(2L+1) \\ &= -a_i^\dagger(\mathbf{a}\cdot\mathbf{a})a_j^{(-)}(2L+1) + a_ia_j^{(-)}(2L+1) \\ &\quad \times(2K+2L-1) \quad \text{[using Eqs. (26) and (34)]} \\ &= -a_i^\dagger(\mathbf{a}\cdot\mathbf{a})[-a_j^\dagger(\mathbf{a}\cdot\mathbf{a}) + a_j(2K+2L+1)] \\ &\quad + a_i[-a_j^\dagger(\mathbf{a}\cdot\mathbf{a}) + a_j(2K+2L+1)] \\ &\quad \times(2K+2L-1) \quad \text{[using Eq. (A3) again]} \\ &= a_i^\dagger a_j^\dagger(\mathbf{a}\cdot\mathbf{a})^2 + a_ia_j(2K+2L+1)(2K+2L-1) \\ &\quad - (a_i^\dagger a_j + a_ia_j^\dagger)(\mathbf{a}\cdot\mathbf{a})(2K+2L-1) \\ &\qquad \text{[using Eqs. (3)].} \end{aligned} \qquad (A7)$$

The right-hand side of this equation is symmetric in i and j. Thus

$$a_i^{(-)}(2L+1)a_j^{(-)}(2L+1) = a_j^{(-)}(2L+1)a_i^{(-)}(2L+1),$$

that is,

$$(2L+3)(2L+5)[a_i^{(-)}, a_j^{(-)}] = 0,$$

which implies that $[a_i^{(-)}, a_j^{(-)}] = 0$. It follows at once from Eq. (A6) that $[\lambda_i, \lambda_j] = 0$; and in a similar way we deduce that $[\lambda_i^\dagger, \lambda_j^\dagger] = 0$.

We see also from Eq. (A7) that

$$\begin{aligned} (2L+3)&(2L+5)\,\mathbf{a}^{(-)}\cdot\mathbf{a}^{(-)} \\ &= (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(\mathbf{a}\cdot\mathbf{a})^2 + (\mathbf{a}\cdot\mathbf{a})(N+L+1)(N+L-1) \\ &\quad - (2N+3)(\mathbf{a}\cdot\mathbf{a})(N+L-1) \\ &= [(N-L)(N+L+1) + (N+L+3) \\ &\quad \times(N+L+1) - (2N+3)(N+L+1)](\mathbf{a}\cdot\mathbf{a}) \\ &= 0. \end{aligned}$$

Thus $\mathbf{a}^{(-)}\cdot\mathbf{a}^{(-)} = 0$, and it follows from Eq. (A6) that $\boldsymbol{\lambda}\cdot\boldsymbol{\lambda} = 0$. In a similar way, we deduce that $\boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda}^\dagger = 0$.

Equations (41) have now been confirmed. Their validity can be seen also from more general arguments. Since $\boldsymbol{\lambda}$ shifts the value of L down by one unit, the vector operator

$$\theta_i = \epsilon_{ijk}[\lambda_j, \lambda_k]$$

shifts the value of L down by two units. But a vector operator can only have components which commute with L, or shift its value up or down by one unit. Thus $\boldsymbol{\theta}$, and hence $[\lambda_i, \lambda_j]$ must vanish. Similarly, the scalar $\boldsymbol{\lambda}\cdot\boldsymbol{\lambda}$ shifts the value of L down by two units. But a scalar operator commutes with L; and therefore $\boldsymbol{\lambda}\cdot\boldsymbol{\lambda} = 0$. □

Equations (42) follow trivially from Eqs. (23) and (29), since $f(K,L)$ is a scalar operator, commuting with the l_{ki}; and Eqs. (43) follow at once from the fact that $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\dagger$ are vector operators by the manner of their construction. □

Consider now

$$\begin{aligned} \nu\nu^\dagger &= (\mathbf{a}\cdot\mathbf{a})g^2(K,L)(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) \\ &= (\mathbf{a}\cdot\mathbf{a})(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)g^2(K+1,L) \\ &= 2(K+1)(2K+2L+3)\,g^2(K+1,L) \\ &\qquad \text{[using Eqs. (15) and (A5)].} \end{aligned}$$

Similarly we find

$$\begin{aligned} \nu^\dagger\nu &= (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(\mathbf{a}\cdot\mathbf{a})g^2(K,L) \\ &= 2K(2K+2L+1)\,g^2(K,L). \end{aligned}$$

With g as in Eqs. (44), these two equations reduce to

$$\nu\nu^\dagger = K+1, \quad \nu^\dagger\nu = K,$$

establishing the first of Eqs. (45) and the first of Eqs. (46). □

Next consider the products

$$\begin{aligned} \lambda_i\nu &= a_i^{(-)}f(K,L)(\mathbf{a}\cdot\mathbf{a})g(K,L) \\ &= a_i^{(-)}(\mathbf{a}\cdot\mathbf{a})f(K-1,L)g(K,L), \\ \nu\lambda_i &= (\mathbf{a}\cdot\mathbf{a})g(K,L)a_i^{(-)}f(K,L) \\ &= (\mathbf{a}\cdot\mathbf{a})a_i^{(-)}g(K,L-1)f(K,L) \\ &= a_i^{(-)}(\mathbf{a}\cdot\mathbf{a})g(K,L-1)f(K,L) \\ &\qquad \text{[using Eqs. (3) and (20)].} \end{aligned}$$

With f and g as in Eqs. (44) we have

$$g(K,L-1)f(K,L) = f(K-1,L)g(K,L)$$

and it then follows that $[\lambda_i, \nu] = 0$. Taking the Hermitian conjugate of this equation, we deduce that $[\lambda_i^\dagger, \nu^\dagger] = 0$, and the second set of Eqs. (45) is verified. □

Consider next the products

$$\lambda_i \nu^\dagger = a_i^{(-)} f(K,L)(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) g(K+1,L)$$
$$= a_i^{(-)}(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) f(K+1,L) g(K+1,L),$$
$$\nu^\dagger\lambda_i = (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) a_i^{(-)} g(K+1,L-1) f(K,L)$$
$$= [a_i^{(-)}(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) - 2a_i^{\dagger(-)}] g(K+1,L-1) \times f(K,L) \quad \text{[using Eqs. (3) and (20)]},$$
$$= [a_i^{(-)}(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) - 2a_i^{(-)}(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(2K+2L+3)^{-1}] \times g(K+1,L-1) f(K,L) \quad \text{[using Eq. (38b)]}.$$

Thus

$$[\lambda_i, \nu^\dagger] = a_i^{(-)}(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(2K+2L+3)^{-1} \times [(2K+2L+3) g(K+1,L) f(K+1,L) - (2K+2L+1) g(K+1,L-1) f(K,L)] = 0 \tag{A8}$$

because of the form of f and g in Eqs. (44). Taking the Hermitian conjugate of Eq. (A8) we deduce also that $[\lambda_i^\dagger, \nu] = 0$, so that the third set of Eqs. (45) is confirmed. □

Now consider the product

$$\lambda_i^\dagger\lambda_j = f(K,L) a_i^{\dagger(+)} a_j^{(-)} f(K,L)$$
$$= a_i^{\dagger(+)}(2L+1) a_j^{(-)} f^2(K,L)(2L-1)^{-1}$$
$$= [-a_i(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) + a_i^\dagger(N+L+3)] a_j^{(-)} \times f^2(K,L)(2L-1)^{-1} \quad \text{[using Eqs. (A4)]}$$
$$= [-a_i(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) a_j^{(-)}(2L+1) + a_i^\dagger a_j^{(-)}(2L+1) \times (N+L+1)] f^2(K,L)(4L^2-1)^{-1}$$
$$= \{-a_i(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)[-a_j^\dagger(\mathbf{a}\cdot\mathbf{a}) + a_j(N+L+1)] + a_i^\dagger[-a_j^\dagger(\mathbf{a}\cdot\mathbf{a}) + a_j(N+L+1)](N+L+1)\} \times f^2(K,L)(4L^2-1)^{-1} \quad \text{[using Eq. (A3)]}$$
$$= \{a_i a_j^\dagger(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(\mathbf{a}\cdot\mathbf{a}) - a_i a_j(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)(N+L+1) + 2a_i a_j^\dagger(N+L+1) - a_i^\dagger a_j^\dagger(\mathbf{a}\cdot\mathbf{a})(N+L+1) + a_i^\dagger a_j(N+L+1)^2\} f^2(K,L)(4L^2-1)^{-1}$$
$$= \{2a_i a_j^\dagger(K+1) + a_i^\dagger a_j(2K+2L+1) - a_i a_j(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) - a_i^\dagger a_j^\dagger(\mathbf{a}\cdot\mathbf{a})\} \times f^2(K,L)(2K+2L+1)(4L^2-1)^{-1} \tag{A9}$$

[using Eqs. (A5) and (15)].

In a similar way, we show that

$$\lambda_i\lambda_j^\dagger = \{2a_i^\dagger a_j K + a_i a_j^\dagger(2K+2L+3) - a_i a_j(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) - a_i^\dagger a_j^\dagger(\mathbf{a}\cdot\mathbf{a})\} f^2(K,L+1)(2K+2L+3) \times (2L+1)^{-1}(2L+3)^{-1}. \tag{A10}$$

It follows from Eqs. (A9) and (A5) that

$$\boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda} = \{2(2K+L+3)(K+1) + (2K+L) \times (2K+2L+1) - 2(K+1)(2K+2L+3) - 2K(2K+2L+1)\} \times f^2(K,L)(2K+2L+1)(4L^2-1)^{-1}$$
$$= L f^2(K,L)(2K+2L+1)(2L+1)^{-1}.$$

Thus $\boldsymbol{\lambda}^\dagger\cdot\boldsymbol{\lambda} = L$ for the choice of f in Eqs. (44), verifying the second of Eqs. (46). □

It can also be seen from Eqs. (A9) that

$$\lambda_i^\dagger\lambda_j - \lambda_j^\dagger\lambda_i = \{2(a_i a_j^\dagger - a_j a_i^\dagger)(K+1) + (a_i^\dagger a_j - a_j^\dagger a_i)(2K+2L+1)\} \times f^2(K,L)(2K+2L+1) \times (4L^2-1)^{-1}$$
$$= i l_{ij} f^2(K,L)(2K+2L+1)(2L+1)^{-1},$$

so that, again with f as in Eqs. (44), we confirm the last of Eqs. (46). □

From Eqs. (A9) and (A10), we see that with this choice of f,

$$(2L+1)[\lambda_i, \lambda_j^\dagger] + 2\lambda_i^\dagger\lambda_j$$
$$= (2L+1)\lambda_i\lambda_j^\dagger - (2L-1)\lambda_j^\dagger\lambda_i + 2(\lambda_i^\dagger\lambda_j - \lambda_j^\dagger\lambda_i)$$
$$= 2a_i^\dagger a_j K + a_i a_j^\dagger(2K+2L+3) - a_i a_j(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) - a_i^\dagger a_j^\dagger(\mathbf{a}\cdot\mathbf{a}) - 2a_j a_i^\dagger(K+1) - a_j^\dagger a_i(2K+2L+1) + a_j a_i(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger) + a_j^\dagger a_i^\dagger(\mathbf{a}\cdot\mathbf{a}) + 2il_{ij}$$
$$= -2a_j a_i^\dagger - 2K\delta_{ij} + 2a_i a_j^\dagger + \delta_{ij}(2K+2L+1) + 2i\, l_{ij}$$
$$= (2L+1)\delta_{ij},$$

thus confirming the last of Eqs. (45). □

APPENDIX B

In order to derive Eq. (57) from Eq. (55), it is necessary to calculate the effect of $\lambda_{-\epsilon}$ on the vector

$$(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle.$$

We note that

$$\lambda_{-\epsilon}\lambda_\epsilon^\dagger = (\lambda_1 - i\epsilon\lambda_2)(\lambda_1^\dagger + i\epsilon\lambda_2^\dagger)$$
$$= \boldsymbol{\lambda}\cdot\boldsymbol{\lambda}^\dagger - \lambda_3\lambda_3^\dagger + i\epsilon(\lambda_1\lambda_2^\dagger - \lambda_2\lambda_1^\dagger).$$

It follows from Eqs. (45) and (46) that

$$\boldsymbol{\lambda}\cdot\boldsymbol{\lambda}^\dagger = (2L+3)(L+1)(2L+1)^{-1},$$
$$\lambda_1\lambda_2^\dagger - \lambda_2\lambda_1^\dagger = -i(2L+3)(2L+1)^{-1} l_{12},$$

so that

$$\lambda_{-\epsilon}\lambda_\epsilon^\dagger = (L+1+\epsilon l_{12})(2L+3)(2L+1)^{-1} - \lambda_3\lambda_3^\dagger. \tag{B1}$$

We now see that if $|m| \geqslant 1$ (which requires $l \geqslant 1$),

$$\lambda_{-\epsilon}(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= (l+|m|-1)(2l+1)(2l-1)^{-1} \times (\lambda_\epsilon^\dagger)^{|m|-1}(\lambda_\epsilon^\dagger)^{l-|m|}|0\rangle - \lambda_3(\lambda_\epsilon^\dagger)^{|m|-1} \times (\lambda_3^\dagger)^{l-|m|+1}|0\rangle. \tag{B2}$$

Working from Eq. (56), we deduce that if r, s, and t are nonnegative integers, then

$$\lambda_3(\lambda_1^\dagger)^r(\lambda_2^\dagger)^s(\lambda_3^\dagger)^t|0\rangle$$
$$= [2(r+s+t)-1]^{-1}\{t(t+2r+2s)(\lambda_1^\dagger)^r(\lambda_2^\dagger)^s \times (\lambda_3^\dagger)^{t-1} - r(r-1)(\lambda_1^\dagger)^{r-2}(\lambda_2^\dagger)^s(\lambda_3^\dagger)^{t+1} - s(s-1)(\lambda_1)^r(\lambda_2)^{s-2}(\lambda_3^\dagger)^{t+1}\}|0\rangle. \tag{B3}$$

Now

$$(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle = \sum_{r=0}^{|m|} \frac{|m|!}{r!(|m|-r)!} \times (\lambda_1^\dagger)^r(i\epsilon\lambda_2^\dagger)^{|m|-r}(\lambda_3^\dagger)^{l-|m|}|0\rangle, \tag{B4}$$

and so

$$\lambda_3(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= (2l-1)^{-1}\sum_{r=0}^{|m|}\frac{|m|!}{r!(|m|-r)!}\{(l^2-m^2)(\lambda_1^\dagger)^r \times(i\epsilon\lambda_2^\dagger)^{|m|-r}(\lambda_3^\dagger)^{l-|m|-1} - r(r-1)(\lambda_1^\dagger)^{r-2} \times(i\epsilon\lambda_2^\dagger)^{|m|-r}(\lambda_3^\dagger)^{l-|m|+1} + (|m|-r) \times(|m|-r-1)(\lambda_1^\dagger)^r(i\epsilon\lambda_2^\dagger)^{|m|-r-2} \times(\lambda_3^\dagger)^{l-|m|+1}\}|0\rangle$$
$$= \{(l^2-m^2)(2l-1)^{-1}(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|-1} - (2l-1)^{-1}\sum_{r=2}^{|m|}\frac{|m|!}{(r-2)!(|m|-r)!}(\lambda_1^\dagger)^{r-2} \times(i\epsilon\lambda_2^\dagger)^{|m|-r}(\lambda_3^\dagger)^{l-|m|+1} + (2l-1)^{-1}\sum_{r=0}^{|m|-2}\frac{|m|!}{r!(|m|-r-2)!} \times(\lambda_1^\dagger)^r(i\epsilon\lambda_2^\dagger)^{|m|-r-2}(\lambda_3^\dagger)^{l-|m|+1}\}|0\rangle$$
$$= (l^2-m^2)(2l-1)^{-1}(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|-1}|0\rangle, \quad \text{(B5)}$$

since the two sums cancel.

Combining this result with Eq. (B2), we see that if $|m|\geqslant 1$,

$$\lambda_{-\epsilon}(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= [(l+|m|-1)(2l+1)(2l-1)^{-1} - (l-|m|+1) \times(l+|m|+1)(2l-1)^{-1}](\lambda_\epsilon^\dagger)^{|m|-1}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= (l+|m|)(l+|m|-1)(2l-1)^{-1} \times(\lambda_\epsilon^\dagger)^{|m|-1}(\lambda_3^\dagger)^{l-|m|}|0\rangle. \quad \text{(B6)}$$

It follows that, if $|m|\geqslant 1$,

$$\langle 0|(\lambda_3)^{l-|m|}(\lambda_{-\epsilon})^{|m|}(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= (l+|m|)(l+|m|-1)(2l-1)^{-1}\langle 0|(\lambda_3)^{l-|m|} \times(\lambda_{-\epsilon})^{|m|-1}(\lambda_\epsilon^\dagger)^{|m|-1}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= \left[\prod_{i=0}^{|m|-1}\left(\frac{(l-2i+|m|)(l-2i+|m|-1)}{(2l-2i-1)}\right)\right] \times\langle 0|(\lambda_3)^{l-|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle. \quad \text{(B7)}$$

We see from Eq. (B3) that

$$\lambda_3(\lambda_3^\dagger)^{l-|m|}|0\rangle = [2(l-|m|)-1]^{-1}(l-|m|)^2 \times(\lambda_3^\dagger)^{l-|m|-1}|0\rangle,$$

so that

$$\langle 0|(\lambda_3)^{l-|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= [2(l-|m|-1)]^{-1}(l-|m|)^2 \times\langle 0|(\lambda_3)^{l-|m|-1}(\lambda_3^\dagger)^{l-|m|-1}|0\rangle$$
$$= \left[\prod_{j=0}^{l-|m|-1}\left(\frac{(l-|m|-j)^2}{(2l-2|m|-2j-1)}\right)\right]\langle 0|0\rangle. \quad \text{(B8)}$$

We now combine Eqs. (B7) and (B8) to obtain (for $|m|\geqslant 1$)

$$\langle 0|(\lambda_3)^{l-|m|}(\lambda_{-\epsilon})^{|m|}(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle = \frac{2^l l!(l+|m|)!(l-|m|)!}{(2l)!}. \quad \text{(B9)}$$

It is seen from Eq. (B8) that this result is valid also if $m=0$. Combining Eqs. (B9) and (55), we obtain Eq. (57). □

The first two of Eqs. (59) are well known in the boson calculus, and require no derivation here. Consider

$$\lambda_3^\dagger|k\,l\,m\rangle = \lambda_3^\dagger c_{klm}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle = \frac{c_{klm}}{c_{kl+1m}}|k\,l+1\,m\rangle = \left(\frac{(l+1-m)(l+1+m)}{(2l+1)}\right)^{1/2}|k\,l+1\,m\rangle,$$

verifying the third of Eqs. (59). □

Next consider, for $m\neq 0$,

$$\lambda_\epsilon^\dagger|k\,l\,m\rangle = \lambda_\epsilon^\dagger c_{klm}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle = \frac{c_{klm}}{c_{kl+1m+\epsilon}}|k\;l+1\;m+\epsilon\rangle = (-\epsilon)^\epsilon\left(\frac{(l+|m|+2)(l+|m|+1)}{(2l+1)}\right)^{1/2} \times|k\;l+1\;m+\epsilon\rangle.$$

From this equation we have

$$\lambda_+^\dagger|k\,l\,m\rangle = -\left(\frac{(l+m+2)(l+m+1)}{(2l+1)}\right)^{1/2} \times|k\;l+1\;m+1\rangle, \text{ for } m>0,$$
$$\lambda_-^\dagger|k\,l\,m\rangle = +\left[\frac{(l-m+2)(l-m+1)}{(2l+1)}\right]^{1/2} \times|k\;l+1\;m-1\rangle, \text{ for } m<0. \quad \text{(B10)}$$

Now consider, also for $m\neq 0$,

$$\lambda_{-\epsilon}^\dagger|k\,l\,m\rangle = \lambda_{-\epsilon}^\dagger c_{klm}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= (\lambda_{-\epsilon}^\dagger\lambda_\epsilon^\dagger)c_{klm}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|-1}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$
$$= -(\lambda_3^\dagger)^2 c_{klm}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|-1}(\lambda_3^\dagger)^{l-|m|}|0\rangle$$

[using the last of Eqs. (41)]

$$= -\frac{c_{klm}}{c_{kl+1m-\epsilon}}|k\;l+1\;m-\epsilon\rangle = -(-\epsilon)^\epsilon\left(\frac{(l-|m|+2)(l-|m|+1)}{(2l+1)}\right)^{1/2} \times|k\;l+1\;m-\epsilon\rangle.$$

From this equation we have

$$\lambda_+^\dagger|k\,l\,m\rangle = -\left(\frac{(l+m+2)(l+m+1)}{(2l+1)}\right)^{1/2} \times|k\;l+1\;m+1\rangle, \text{ for } m<0,$$
$$\lambda_-^\dagger|k\,l\,m\rangle = \left(\frac{(l-m+2)(l-m+1)}{(2l+1)}\right)^{1/2} \times|k\;l+1\;m-1\rangle, \text{ for } m>0. \quad \text{(B11)}$$

Next consider (with $\epsilon=\pm 1$)

$$\lambda_\epsilon^\dagger|k\,l\,0\rangle = \lambda_\epsilon^\dagger c_{kl0}(v^\dagger)^k(\lambda_3^\dagger)^l|0\rangle = \frac{c_{kl0}}{c_{kl+1\epsilon}}|k\;l+1\;\epsilon\rangle = (-\epsilon)^\epsilon\left(\frac{(l+2)(l+1)}{(2l+1)}\right)^{1/2}|k\;l+1\;\epsilon\rangle. \quad \text{(B12)}$$

Combining Eqs. (B10), (B11) and (B12), we arrive at the fourth set of Eqs. (59). □

From Eq. (B5) we have

$$\begin{aligned}\lambda_3|k\,l\,m\rangle &= c_{klm}\frac{(l^2-m^2)}{(2l-1)}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|-1}|0\rangle\\ &= \frac{(l^2-m^2)}{(2l-1)}\frac{c_{klm}}{c_{kl-1m}}|k\ l-1\ m\rangle\\ &= \left(\frac{(l-m)\,(l+m)}{(2l-1)}\right)^{1/2}|k\ l-1\ m\rangle\,,\end{aligned}$$

as in the fifth of Eqs. (59). □

Using Eq. (B6), we see that, for $|m|\geqslant 1$,

$$\begin{aligned}\lambda_{-\epsilon}|k\,l\,m\rangle &= \lambda_{-\epsilon}c_{klm}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle\\ &= \frac{(l+|m|)\,(l+|m|-1)}{(2l-1)}\frac{c_{klm}}{c_{kl-1m-\epsilon}}\\ &\quad\times|k\ l-1\ m-\epsilon\rangle\\ &= (-\epsilon)^\epsilon\left(\frac{(l+|m|)\,(l+|m|-1)}{(2l-1)}\right)^{1/2}\\ &\quad\times|k\ l-1\ m-\epsilon\rangle\,.\end{aligned}$$

From this equation we have

$$\begin{aligned}\lambda_-|k\,l\,m\rangle &= -\left(\frac{(l+m)\,(l+m-1)}{(2l-1)}\right)^{1/2}\\ &\quad\times|k\ l-1\ m-1\rangle\,,\ \text{for}\ m>0\,,\\ \lambda_+|k\,l\,m\rangle &= \left(\frac{(l-m)\,(l-m-1)}{(2l-1)}\right)^{1/2}\\ &\quad\times|k\ l-1\ m+1\rangle\,,\ \text{for}\ m<0\,.\end{aligned}\tag{B13}$$

Now consider the vector

$$\lambda_\epsilon|k\,l\,m\rangle = \lambda_\epsilon c_{klm}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|}|0\rangle\,,$$

on which L has the value $l-1$, and L_3 the value $\hbar(m+\epsilon)$. Since $\hbar^2L(L+1)=(L_1)^2+(L_2)^2+(L_3)^2$, the value of L_3 cannot be greater in modulus than that of $\hbar L$. Therefore, this vector vanishes unless $l-1\geqslant|m+\epsilon|=|m|+1$; i.e., unless $l\geqslant|m|+2$. Supposing this inequality is satisfied, we write

$$\begin{aligned}\lambda_\epsilon|k\,l\,m\rangle &= c_{klm}\lambda_\epsilon(\lambda_3^\dagger)^2(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|}(\lambda_3^\dagger)^{l-|m|-2}|0\rangle\\ &= -c_{klm}\lambda_\epsilon\lambda_{-\epsilon}^\dagger(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|+1}(\lambda_3^\dagger)^{l-|m|-2}|0\rangle\\ &\qquad\text{[using the last of Eqs. (41)]}\,.\end{aligned}$$

Now using Eq. (B1) (with ϵ replaced by $-\epsilon$), we have

$$\begin{aligned}\lambda_\epsilon|k\,l\,m\rangle &= -c_{klm}\{(L+1-\epsilon l_{12})\,(2L+3)\\ &\quad\times(2L+1)^{-1}(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|+1}(\lambda_3^\dagger)^{l-|m|-2}\\ &\quad-\lambda_3(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|+1}(\lambda_3^\dagger)^{l-|m|-1}\}|0\rangle\\ &= -c_{klm}\{(l-|m|-1)\,(2l+1)\,(2l-1)^{-1}\\ &\quad-[l^2-(|m|+1)^2]\,(2l-1)^{-1}\}\\ &\quad\times(v^\dagger)^k(\lambda_\epsilon^\dagger)^{|m|+1}(\lambda_3^\dagger)^{l-|m|-2}|0\rangle\\ &\qquad\text{[using Eq. (B5)]}\\ &= -\frac{(l-|m|)\,(l-|m|-1)}{(2l-1)}\frac{c_{klm}}{c_{kl-1m+\epsilon}}\\ &\quad\times|k\ l-1\ m+\epsilon\rangle\\ &= -(-\epsilon)^\epsilon\left(\frac{(l-|m|)\,(l-|m|-1)}{(2l-1)}\right)^{1/2}\\ &\quad\times|k\ l-1\ m+\epsilon\rangle\,.\end{aligned}$$

From this equation we see that

$$\begin{aligned}\lambda_+|k\,l\,m\rangle &= \left(\frac{(l-m)\,(l-m-1)}{(2l-1)}\right)^{1/2}\\ &\quad\times|k\ l-1\ m+1\rangle\,,\ \text{for}\ m>0\,,\\ \lambda_-|k\,l\,m\rangle &= -\left(\frac{(l+m)\,(l+m-1)}{(2l-1)}\right)^{1/2}\\ &\quad\times|k\ l-1\ m-1\rangle\,,\ \text{for}\ m<0\,.\end{aligned}\tag{B14}$$

It is easily seen from Eq. (56) that

$$\lambda_\pm(\lambda_3^\dagger)^l|0\rangle = -\frac{l(l-1)}{(2l-1)}\lambda_\pm^\dagger(\lambda_3^\dagger)^{l-2}|0\rangle\,,$$

so that

$$\begin{aligned}\lambda_\pm|k\,l\,0\rangle &= \lambda_\pm c_{kl0}(v^\dagger)^k(\lambda_3^\dagger)^l|0\rangle\\ &= -\frac{l(l-l)}{(2l-1)}\frac{c_{kl0}}{c_{kl-1\pm1}}|k\ l-1\pm1\rangle\,.\end{aligned}$$

From this equation we see that

$$\lambda_\pm|k\,l\,0\rangle = \pm\left(\frac{l(l-1)}{(2l-1)}\right)^{1/2}|k\ l-1\pm1\rangle\,,\tag{B15}$$

and combining Eqs. (B13), (B14), and (B15) we obtain the last of Eqs. (59). □

[1]Philip Stehle, *Quantum Mechanics* (Holden-Day, San Francisco, 1966).
[2]A.S. Davydov, *Quantum Mechanics* (Pergamon, Oxford, 1965).
[3]P. Kramer and M. Moshinsky, *Group Theory and its Applications*, edited by E. M. Loebl, Vol. I (Academic, New York, 1968).
[4]Harold V. McIntosh, *ibid.*, Vol. II (Academic, New York, 1971).
[5]M.E. Rose in *Spectroscopic and Group Theoretical Methods in Physics* edited by F. Bloch *et al.* (North-Holland, Amsterdam, 1968).
[6]H.S. Green, *Matrix Mechanics* (Noordhoff, Groningen, 1965).
[7]L. Infeld and T.E. Hull, Rev. Mod. Phys. **23**, 21 (1951); Willard Miller, Jr., *Lie Theory and Special Functions* (Academic, New York, 1968).
[8]P.W. Atkins and J.C. Dobson, Proc. R. Soc. London Ser. A **321**, 321 (1971).
[9]R. Delbourgo, J. Phys. A: Math. Gen. **10**, 1837 (1977).
[10]A.J. Bracken and H.S. Green, J. Math. Phys. **12**, 2099 (1971).
[11]A.J. Bracken, Ph.D. thesis, The University of Adelaide, 1970 (unpublished).
[12]R.C. Hwa and J. Nuyts, Phys. Rev. **145**, 1188 (1966).
[13]M.A. Lohe and C.A. Hurst, J. Math. Phys. **12**, 1882 (1971).
[14]Milton Abramowitz and Irene A. Stegun (Editors), *Handbook of Mathematical Functions* (Dover, New York, 1965).
[15]A.R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, N.J., 1960).
[16]A.O. Barut and C. Fronsdal, Proc. R. Soc. London Ser. A **287**, 532 (1965).
[17]M.A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, Oxford, 1964).
[18]A. Böhm, Phys. Rev. **175**, 1767 (1968).

Tight lower bounds to eigenvalues of the Schrödinger equation

Robert Nyden Hill
Physics Department, University of Delaware, Newark, Delaware

(Received 7 June 1979; accepted for publication 8 February 1980)

A new tight lower bound to eigenvalues of the Schrödinger equation, which tends to be better than the well-known Temple lower bound, is presented. The new bound works (as does the Temple bound) for both ground and excited states. Optimization of both the Temple bound and the new tight bound with respect to a variational trial function is discussed. Numerical results are given for the anharmonic oscillator.

I. INTRODUCTION

Traditional variational (Rayleigh–Ritz) calculations of eigenvalues of the Schrodinger equation yield upper bounds to those eigenvalues. Complementary lower bounds are needed to complete the picture and locate the eigenvalues with certainty. Tight lower bounds comparable in accuracy with the Rayleigh–Ritz upper bounds are normally wanted for this purpose.[1]

The best-known tight lower bound is the Temple lower bound[2]

$$E_{\rm T} = \langle H\rangle - (\widetilde{E}_{n+1} - \langle H\rangle)^{-1}(\langle H^2\rangle - \langle H\rangle^2) \qquad (1)$$

to the energy E_n of the nth state of a Hamiltonian H. Here $\langle H\rangle = \langle\Phi|H|\Phi\rangle/\langle\Phi|\Phi\rangle$ and $\langle H^2\rangle = \langle\Phi|H^2|\Phi\rangle/\langle\Phi|\Phi\rangle$, where $|\Phi\rangle$ is an approximation to the nth state eigenfunction. $\widetilde{E}_{n+1}$, which must be greater than $\langle H\rangle$, is a rough prior lower bound to the energy of the $(n+1)$th state. In practice, $|\Phi\rangle$ is usually the approximation obtained by minimizing $\langle H\rangle$ with respect to the parameters in a suitably chosen variational trial function. Typically the Rayleigh–Ritz upper bound to E_n is much closer to E_n than the Templer lower bound to E_n. For example, in Bazley's well-known calculation[3] on the ground state of helium, values of $\langle H\rangle$ and $\langle H^2\rangle - \langle H\rangle^2$ from an 80-term variational computation by Kinoshita yielded $-2.9037474 = E_T \leqslant E_1$ for the Temple bound and $E_1 \leqslant \langle H\rangle = -2.9037237$ for the upper bound. Comparison with the value $E_1 \cong -2.9037244$ taken from a variational computation by Frankowski and Pekeris,[4] which is generally believed to be exact to at least one part in 10^9, shows that the error in the Temple lower bound is greater than the error in the Rayleigh–Ritz upper bound by a factor of about 33.

The fact that the Temple lower bound tends to be much further from the true value than the Rayleigh–Ritz upper bound motivated a search for alternative methods of obtaining tight lower bounds. Section II reviews relevant existing work and presents a new tight lower bound which can be expected to be more accurate than the Temple bound, and which works for both ground states and excited states. Section III presents numerical comparison of the Rayleigh–Ritz upper bound, the Temple lower bound, and the new tight lower bound for the anharmonic oscillator. Section IV discusses optimization of both the Temple lower bound and the new tight lower bound with respect to a variational trial function $|\Phi\rangle$.

II. A NEW TIGHT LOWER BOUND

Construction of the Temple lower bound to E_n requires values for $\langle H\rangle$, for $\langle H^2\rangle$, and for $\widetilde{E}_{n+1}$. The methods introduced by Bazley[3] and by Bazley and Fox[5–8] to obtain the rough lower bound $\widetilde{E}_{n+1}$ to the $(n+1)$th eigenvalue E_{n+1} can be used to obtain rough lower bounds to other excited states; with this extra information available, it is possible to obtain improved tight lower bounds which are closer to the true value than the Temple bound for both the ground state and the excited states. This section begins by stating the basic theorem used to establish the lower bounds. Relevant existing lower bound methods are then briefly reviewed. Introduction of the new tight lower bound follows. The section concludes with a comparison of the errors expected for the Rayleigh–Ritz upper bound, the Temple lower bound, and the new tight lower bound.

A. The basic theorem

The basic theorem used to establish the lower bounds is a comparison theorem[9] well known among mathematicans who work on eigenvalue problems:

Let $H^{(1)}$ and $H^{(2)}$ be two essentially self-adjoint (Hermitian) Hamiltonians whose discrete eigenvalues below the bottom of the essential spectrum can be characterized by the familiar variational principle $E = \min\langle\psi|H|\psi\rangle/\langle\psi|\psi\rangle$, with the minimization for excited states carried out subject to the constraint that $|\psi\rangle$ be orthogonal to preceding eigenvectors. Denote the ordered eigenvalues of $H^{(i)}$ by $E_1^{(i)} \leqslant E_2^{(i)} \leqslant \cdots \leqslant E_n^{(i)} \leqslant \cdots \leqslant E_{\rm ess}^{(i)}$, where $E_{\rm ess}^{(i)}$ is the energy at which the essential spectrum (if any) begins. Assume $\langle\psi|H^{(1)}|\psi\rangle$ is defined for all vectors $|\psi\rangle$ for which $\langle\psi|H^{(2)}|\psi\rangle$ is defined. Then if $\langle\psi|H^{(1)}|\psi\rangle \leqslant \langle\psi|H^{(2)}|\psi\rangle$ holds for all admissible state vectors $|\psi\rangle$, $E_n^{(1)} \leqslant E_n^{(2)}$ holds for all n, and $E_{\rm ess}^{(1)} \leqslant E_{\rm ess}^{(2)}$.

The result $E_1^{(1)} \leqslant E_1^{(2)}$ for the ground state energy follows immediately from the "familiar variational principle." Proofs of the results for the excited states are usually based on one of the minimax characterizations[10] of eigenvalues. In practical applications of the basic theorem to the computation of lower bounds, $H^{(2)}$ is the original Hamiltonian, while $H^{(1)}$ is something more tractable.

B. Review of relevant existing methods

Several methods of obtaining a tractable lower bound-

 0022-2488/80/082182-11$1.00

ing Hamiltonian $H^{(1)}$ have been introduced by Bazley,[3] and by Bazley and Fox.[5-8] Assume that the Hamiltonian of interest has the form

$$H = H_0 + V, \tag{2}$$

where the eigenvalue problem for H_0 is exactly solvable, and where V is a nonnegative operator. Then V will have a unique positive square root, which will be denoted by $V^{1/2}$ Suppose that P is a projection operator. The fact that a projection operator such as P cannot increase the length of a vector such as $V^{1/2}|\psi\rangle$ implies that $\langle\psi|V^{1/2}PV^{1/2}|\psi\rangle \leqslant \psi|V|\psi\rangle$ for all admissible $|\psi\rangle$. Aronszajn-type projection, also known as inner projection, replaces V in H by a modified interaction $V^{1/2}PV^{1/2}$ and leads to a new Hamiltonian H' defined by

$$H' = H_0 + V^{1/2}PV^{1/2}. \tag{3}$$

Then the hypotheses of the basic theorem are satisfied with $H^{(1)} = H'$ and $H^{(2)} = H$, so that the eigenvalues of H' are lower bounds to the eigenvalues of the original Hamiltonian H. It remains to choose P.

Bazley's method of special choice[1,3] for constructing P starts with the first M eigenvectors $|\xi_i\rangle$ of H_0 and uses Schmidt orthogonalization to construct vectors $|\xi_i'\rangle$ such that

$$\langle\xi_i'|V^{-1}|\xi_j'\rangle = \delta_{i,j}. \tag{4}$$

The projection operator P, which will for this case be denoted by P_{Bazley}, is then chosen such that the modified interaction is

$$V^{1/2}P_{\text{Bazley}}V^{1/2} = \sum_{i=1}^{M} |\xi_i'\rangle\langle\xi_i'|. \tag{5}$$

It follows immediately from (4) and (5) that $P^{\dagger}_{\text{Bazley}} = P_{\text{Bazley}}$ and $P^2_{\text{Bazley}} = P_{\text{Bazley}}$, so that P_{Bazley} is a projection operator as is required. Because $|\xi_i'\rangle$ is a linear combination of the first $M\,|\xi_i\rangle$, $V^{1/2}P_{\text{Bazley}}V^{1/2}$ couples only the first M states of H_0, and the eigenvalue problem for $H_0 + V^{1/2}P_{\text{Bazley}}V^{1/2}$ reduces to an eigenvalue problem for an $M\times M$ matrix. The above described method of special choice was used by Bazley to obtain the lower bound $\widetilde{E}_2$ to the first excited state needed for evaluation of the Temple lower bound to the ground state energy of helium. A generalization of Bazley's method of special choice was used by the present author to prove that the H^- ion has only one bound state.[11]

The method of special choice is inconvenient if some of the matrix elements $\langle\xi_l|V^{-1}|\xi_m\rangle$ are hard to calculate, and fails if they do not exist. (This happens for the anharmonic oscillator example of Sec. III, where $V=\lambda x^4$.) An alternative, which works if $\langle\xi_l|V^{1/2}|\xi_m\rangle$ is a band matrix which couples only states for which $|l-m|\leqslant r$, was introduced by Bazley and Fox and called the method of generalized special choice.[8] This alternative is the choice $P = P_k$ where

$$P_k \equiv \sum_{l=1}^{k} |\xi_l\rangle\langle\xi_l|. \tag{6}$$

The modified interaction $V^{1/2}P_kV^{1/2}$ then couples only the states $|\xi_l\rangle$ for which $l\leqslant k+r$, so that the eigenvalue problem for $H_0 + V^{1/2}P_kV^{1/2}$ reduces to an eigenvalue problem for a $(k+r)\times(k+r)$ matrix. A slightly different version of the method of generalized special choice, which works when $\langle\xi_l|V|\xi_m\rangle$ is a band matrix, has also been given by Bazley and Fox.[5,7]

In some cases more accurate bounds can be achieved by starting with N vectors $|\eta_l\rangle$ which are *not* eigenvectors of H_0 and proceeding as outlined above. A modified interaction $V^{1/2}PV^{1/2}$ which couples a finite number of vectors $|\eta_l\rangle$ then results; in order to obtain a tractable problem, H_0 must then also be modified. The necessary modification, suggested by Bazley and Fox,[6] is called truncation. Denote the eigenvalues of H_0 by $\epsilon_1\leqslant\epsilon_2\leqslant\cdots\leqslant\epsilon_l\leqslant\cdots$. Define $H_0^{(s)}$ by

$$H_0^{(s)} = \sum_{l=1}^{s-1} (\epsilon_l - \epsilon_n)|\xi_l\rangle\langle\xi_l| + \epsilon_n I, \tag{7}$$

where I is the identity operator. Then $\langle\psi|H_0^{(s)}|\psi\rangle \leqslant \langle\psi|H_0|\psi\rangle$ for all admissable $|\psi\rangle$, and the eigenvalues of $H_0^{(s)} + V^{1/2}P_kV^{1/2}$, which can be obtained by solving a finite dimensional matrix problem, are lower bounds to the eigenvalues of the original Hamiltonian $H = H_0 + V$.

A generalization of these methods to three or more electron atomic systems, where the continuous spectrum of H_0 can overlap the ground state of the atom, was given by Fox,[12] who was later joined by Sigillito.[13] A similar but mathematically less complete analysis was given by Reid,[14] who later tried the method on lithium[15] but was unable to raise the bottom of the continuum for his lower-bounding problems above the ground state of the atom.

C. The new tight lower bound

Suppose that a projection operator P has been chosen in such a way that $V^{1/2}PV^{1/2}$ couples only the first M eigenstates of H_0. Let the subspace spanned by these first M eigenstates be denoted by $S_{\parallel}$, and let $P_{\parallel}$ be the projection operator which projects onto $S_{\parallel}$. Let $S_{\perp}$ be the orthogonal complement of $S_{\parallel}$, and let $P_{\perp}$ be the projection operator which projects on $S_{\perp}$. Let E_{lb} be a lower bound to the spectrum of H_0 restricted to $S_{\perp}$, and define H_1, H_2 by

$$H_1 = P_{\parallel}(H_0 + V^{1/2}PV^{1/2})P_{\parallel} + E_{\text{lb}}P_{\perp}, \tag{8}$$

$$H_2 = H - H_1 = V - V^{1/2}PV^{1/2} + P_{\perp}(H_0 - E_{\text{lb}})P_{\perp}. \tag{9}$$

Thus H_1 is obtained from H by truncating H_0 and performing an Aronszajn-type projection (inner projection) on V. The difference H_2 is then an operator whose expectation value is nonnegative. H_2 will now be modified via an Aronszajn-type projection. Suppose now that $|\Phi\rangle$ is an approximation to the nth eigenvector of H, obtained by any method, for which $H_2|\Phi\rangle$ is not the null vector. Then

$$P' = H_2^{1/2}|\Phi\rangle(\langle\Phi|H_2|\Phi\rangle)^{-1}\langle\Phi|H_2^{1/2} \tag{10}$$

is a projection operator, and $H_2' \equiv H_2^{1/2}P'H_2^{1/2} = H_2|\Phi\rangle(\langle\Phi|H_2|\Phi\rangle)^{-1}\langle\Phi|H_2$ is an operator whose expectation value is less than or equal to the expectation value of H_2. The eigenvalues of

$$H'' \equiv H_1 + H_2' = H_1 + H_2|\Phi\rangle(\langle\Phi|H_2|\Phi\rangle)^{-1}\langle\Phi|H_2 \tag{11}$$

are then lower bounds to the eigenvalues of H. If solving the eigenvalue problem for H_1 requires diagonalizing an $M\times M$ matrix, the eigenvalue problem for H'' can be solved by diagonalizing an $(M+1)\times(M+1)$ matrix *if* $P_{\perp}H_2|\Phi\rangle$ is not the null vector; this follows immediately from using

$((\Phi|H_2P_\perp H_2|\Phi))^{-1/2}\Phi\rangle P_\perp H_2|\Phi\rangle$ as one of the basis vectors in $S_\perp$. If $P_\perp H_2|\Phi\rangle$ is the null vector, the eigenvalue problem for H'' can be solved by diagonalizing an $M\times M$ matrix.

The above construction of H'' is motivated by the observation that $H''|\Phi\rangle = H_1|\Phi\rangle + H_2|\Phi\rangle = H|\Phi\rangle$; thus if $|\Phi\rangle$ were an exact eigenstate of H with eigenvalue E_n, it would also be an exact eigenstate of H'' with eigenvalue E_n. If $|\Phi\rangle$ is a good approximation to the nth eigenstate of H, and if $E_{\rm lb}$ is above E_n and not too close to E_n ($E_{\rm lb}$ must be above E_n because $E_{\rm lb}$ is an infinitely degenerate eigenvalue of H''), then the nth eigenvalue of H'' will be a tight lower bound to E_n.

D. Comparison with the Rayleigh–Ritz and Temple bounds

The lower bound produced by H'' can be compared with the Rayleigh–Ritz upper bound and the Temple lower bound produced by a given approximate $|\Phi\rangle$ by supposing that $|\Phi\rangle = |\Phi_{\rm ex}\rangle + \epsilon|\delta\Phi\rangle$, where $|\Phi_{\rm ex}\rangle$ is the exact wave function and $\epsilon|\delta\Phi\rangle$ is a small error assumed perpendicular to $|\Phi_{\rm ex}\rangle$: $\langle\Phi_{\rm ex}|\delta\Phi\rangle = 0$. Denote the exact energy by $E_{\rm ex}$, the Rayleigh–Ritz upper bound by $E_{\rm RR}$, the Temple lower bound by E_T, and the new tight lower bound by E''. Expanding formally to second order in ϵ then produces the following:

$$E_{\rm RR} = E_{\rm ex} + \epsilon^2\langle\delta\Phi|(H-E_{\rm ex})|\delta\Phi\rangle + O(\epsilon^3), \quad (12)$$

$$E_T = E_{\rm ex} + \epsilon^2[\langle\delta\Phi|(H-E_{\rm ex})|\delta\Phi\rangle - (\widetilde{E}_{n+1}-E_{\rm ex})^{-1}\langle\delta\Phi|(H-E_{\rm ex})^2|\delta\Phi\rangle] + O(\epsilon^3), \quad (13)$$

and

$$E'' = E_{\rm ex} - \epsilon^2[\langle\delta\Phi|(H_2-H_2')|\delta\Phi\rangle + \langle\delta\Phi|(H_2-H_2')(H_1+H_2'-E_{\rm ex})^{-1}\times(H_2-H_2')|\delta\Phi\rangle] + O(\epsilon^3). \quad (14)$$

The following observations can be made: In all three cases, the error is of order ϵ^2. The errors in the Temple bound and in the new tight lower bound both look something like second order perturbation theory, but in the Temple formula, the energy denominator is $(\widetilde{E}_{n+1} - E_{\rm ex})$, whereas in the new tight lower bound the energy denominator is the improved approximation $(H_1 + H_2' - E_{\rm ex})$ which incorporates not just the lower bound $\widetilde{E}_{n+1}$ to the $(n+1)$th excited state, but *all* of the lower bounds contained in $H'' = H_1 + H_2'$.

III. NUMERICAL RESULTS FOR THE ANHARMONIC OSCILLATOR

The anharmonic oscillator, with Hamiltonian $H = H_0 + V$, where

$$H_0 = -\frac{d^2}{dx^2} + x^2 \quad (15)$$

and

$$V = \lambda x^4, \quad (16)$$

provides a well-understood one-dimensional model problem which can be easily solved to high accuracy.[16] Thus it provides a useful vehicle for the preliminary exploration of the quality of lower-bound methods. The above bounds have been tested on this model as outlined below.

The eigenstates of H_0 restricted to the even parity sector [symmetric wave function $\psi(x) = \psi(-x)$] are

$$\xi_l(x) = C_{2l-2}\exp(-x^2/2)h_{2l-2}(x), \quad (17)$$

where h_r is the Hermite polynomial defined by the Rodrigues' formula

$$h_r(x) = (-1)^r\exp(x^2)\frac{d^r}{dx^r}\exp(-x^2) \quad (18)$$

and $C_r = (\pi^{1/2}2^r r!)^{-1/2}$. The Bazley method of special choice does not work here, because the matrix elements $\langle\xi_l|x^{-4}|\xi_m\rangle$ are divergent. However, $\langle\xi_l|V^{1/2}|\xi_m\rangle = \lambda^{1/2}\langle\xi_l|x^2|\xi_m\rangle$ is a band matrix which couples only states with $|l-m|\leqslant 1$; thus the alternative of Eq. (6) works very nicely.[17] The modified interaction $V^{1/2}P_k V^{1/2} = \lambda x^2 P_k x^2$ then couples only the states $|\xi_1\rangle, |\xi_2\rangle, \ldots, |\xi_{k+1}\rangle$, which span $S_\parallel$. The needed matrix elements are easily obtained with the aid of the differential equation, the recursion relations, and the orthogonality relation for the Hermite polynomials. The matrix elements of H_0 and x^4 are

$$\langle\xi_l|H_0|\xi_m\rangle = (4l-3)\delta_{l,m} \quad (19)$$

and

$$\langle\xi_l|x^4|\xi_m\rangle = \tfrac{1}{4}[(2l+2)(2l+1)(2l)(2l-1)]^{1/2}\delta_{l+2,m} + \tfrac{1}{2}(4l-1)[2l(2l-1)]^{1/2}\delta_{l+1,m} + (8l^2-12l+5)\delta_{l,m} + \tfrac{1}{2}(4l-5)\times[(2l-2)(2l-3)]^{1/2}\delta_{l-1,m} + \tfrac{1}{4}[(2l-2)(2l-3)\times(2l-4)(2l-5)]^{1/2}\delta_{l-2,m}. \quad (20)$$

The matrix element $\langle\xi_l|x^2P_kx^2|\xi_m\rangle$ equals the matrix element $\langle\xi_l|x^4|\xi_m\rangle$ for either $l<k$ or $m<k$. All other matrix elements $\langle\xi_l|x^2P_kx^2|\xi_m\rangle$ are zero except for

$$\langle\xi_k|x^2P_kx^2|\xi_k\rangle = \tfrac{1}{4}(20k^2-34k+15), \quad (21a)$$

$$\langle\xi_{k+1}|x^2P_kx^2|\xi_{k+1}\rangle = \tfrac{1}{2}k(2k-1), \quad (21b)$$

and

$$\langle\xi_{k+1}|x^2P_kx^2|\xi_k\rangle = \langle\xi_k|x^2P_kx^2|\xi_{k+1}\rangle = \tfrac{1}{4}(4k-3)[2k(2k-1)]^{1/2}. \quad (21c)$$

It follows that $E_{\rm lb} = 4k+5$, and that the null space of H_2 is the space spanned by the $k-1$ vectors $|\xi_1\rangle, |\xi_2\rangle, \cdots, |\xi_{k-1}\rangle$.

Table I shows results obtained for the $\lambda = 1$ ground state eigenvalue E_1 when $|\Phi\rangle$ is the N-term trial function

$$|\Phi\rangle = \sum_{l=1}^{N} a_l|\xi_l\rangle, \quad (22)$$

with the a_l chosen to minimize $\langle H\rangle = \langle\Phi|H|\Phi\rangle/\langle\Phi|\Phi\rangle$. Table II shows corresponding results for the second excited state eigenvalue E_3. All computations were done with a 78 bit mantissa (about 22 significant figures) on a Burroughs B7700 computer. The exact results used for comparison, and the lower bounds $\widetilde{E}_2$ and $\widetilde{E}_4$ to the first and third excited states needed for the Temple lower bound, were obtained by using a 52-term Rayleigh–Ritz trial function to obtain upper bounds and by solving the eigenvalue problem for H_1 with $k = 51$ to obtain lower bounds. The upper and lower bounds to E_1 agree to 21 significant figures and give

$$E_1 = 1.39235164153029185566.$$

The upper and lower bounds to E_3 agree to 17 significant figures and give

$$E_3 = 18.057557436303253.$$

The values obtained for $\widetilde{E}_2$ and $\widetilde{E}_4$ are

$$\widetilde{E}_2 = 8.65504995775930968779$$

and

$$\widetilde{E}_4 = 28.8353384595042467469.$$

Comparison with the corresponding Rayleigh–Ritz upper bounds, which are 8.65504995775930968804 and 28.8353384595042542932, shows that $\widetilde{E}_2$ and $\widetilde{E}_4$ are exact to 19 and 15 significant figures, respectively. The values of $E_1 - E''$ given for $N = 1$ in Table I and of $E_3 - E''$ given for $N = 3$ in Table II are the same as the differences between E_{ex} and the lower bounds obtained from H_1. Inspection of Table I shows that the E'' bounds are better than the Temple bounds to E_1 in every case shown. As N gets large, the ratio $(E_1 - E_{\text{T}})/(E_1 - E'')$, which is a quantitative measure of the superiority of the E'' bounds, does not vary much with N; it is roughly 1.6, 2.7, 3.8, 5.1 and 6.0 for $k = 2,4,6,8$, and 10, respectively. Inspection of Table II shows that the Temple bounds are better than the E'' bounds to E_3 for $k = 4$, but the E'' bounds are better for $k = 6,8,10$, and 12. As N gets large, the ratio $(E_3 - E_{\text{T}})/(E_3 - E'')$ does not vary much with N; it is roughly 0.27, 1.0, 1.8, 2.5 and 5.3 for $k = 4, 6, 8, 10$, and 12, respectively. For $k = 2$ (not shown in Table II), the error $E_3 - E''$ is $5.06 = E_3 - E_{\text{lb}}$ for all N because E_{lb}, which is 13 for $k = 2$, is an infinitely degenerate eigenvalue of the lower bound Hamiltonian H''.

IV. OPTIMIZATION OF THE LOWER BOUNDS

Inspection of Tables I and II shows that the errors in the Temple lower bound and in the new tight lower bound do not always get smaller when the number of terms N in the variational trial function $|\Phi\rangle$ increases. This is a clear indication that the choice of the coefficients a_l which gives the best Rayleigh–Ritz upper bound for a given N is not the same as the choice of the a_l which gives the best lower bound for

TABLE I. The differences between E_1 and the various bounds to E_1 are listed for different numbers N of terms in Φ when $\lambda = 1.0$.

N	$E_{\text{RR}} - E_1$	$E_1 - E_{\text{T}}$	$E_1 - E''$ $k=2$	$E_1 - E''$ $k=4$	$E_1 - E''$ $k=6$	$E_1 - E''$ $k=8$	$E_1 - E''$ $k=10$
1	3.58×10^{-1}	5.11×10^{-1}	8.66×10^{-2}	1.95×10^{-3}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
2	2.03×10^{-2}	1.23×10^{-1}	1.88×10^{-2}	1.95×10^{-3}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
3	2.72×10^{-3}	7.59×10^{-3}	3.29×10^{-3}	1.95×10^{-3}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
4	2.56×10^{-3}	1.38×10^{-2}	6.63×10^{-3}	1.54×10^{-4}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
5	1.02×10^{-3}	1.31×10^{-2}	7.00×10^{-3}	8.74×10^{-4}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
6	2.14×10^{-4}	5.14×10^{-3}	3.01×10^{-3}	8.20×10^{-4}	8.43×10^{-5}	2.75×10^{-5}	5.82×10^{-7}
7	2.33×10^{-5}	9.69×10^{-4}	5.93×10^{-4}	2.87×10^{-4}	7.74×10^{-5}	2.75×10^{-5}	5.82×10^{-7}
8	3.85×10^{-6}	6.23×10^{-5}	3.75×10^{-5}	2.03×10^{-5}	1.17×10^{-5}	5.31×10^{-6}	5.82×10^{-7}
9	3.85×10^{-6}	5.68×10^{-5}	3.41×10^{-5}	1.83×10^{-5}	1.05×10^{-5}	5.02×10^{-6}	5.82×10^{-7}
10	2.28×10^{-6}	7.40×10^{-5}	4.55×10^{-5}	2.56×10^{-5}	1.53×10^{-5}	7.94×10^{-6}	1.68×10^{-7}
11	8.02×10^{-7}	4.26×10^{-5}	2.63×10^{-5}	1.51×10^{-5}	9.75×10^{-6}	5.93×10^{-6}	3.59×10^{-7}
12	1.77×10^{-7}	1.43×10^{-5}	8.89×10^{-6}	5.18×10^{-6}	3.54×10^{-6}	2.49×10^{-6}	4.00×10^{-7}
13	2.34×10^{-8}	2.68×10^{-6}	1.67×10^{-6}	9.78×10^{-7}	6.85×10^{-7}	5.17×10^{-7}	2.35×10^{-7}
14	5.54×10^{-9}	1.91×10^{-7}	1.17×10^{-7}	6.71×10^{-8}	4.60×10^{-8}	3.44×10^{-8}	2.54×10^{-8}
15	5.52×10^{-9}	1.56×10^{-7}	9.58×10^{-8}	5.44×10^{-8}	3.71×10^{-8}	2.75×10^{-8}	2.03×10^{-8}
16	3.74×10^{-9}	2.36×10^{-7}	1.46×10^{-7}	8.50×10^{-8}	5.93×10^{-8}	4.50×10^{-8}	3.36×10^{-8}
17	1.64×10^{-9}	1.62×10^{-7}	1.01×10^{-7}	5.91×10^{-8}	4.15×10^{-8}	3.17×10^{-8}	2.44×10^{-8}
18	4.89×10^{-10}	6.96×10^{-8}	4.33×10^{-8}	2.55×10^{-8}	1.79×10^{-8}	1.38×10^{-8}	1.10×10^{-8}
19	9.64×10^{-11}	1.93×10^{-8}	1.20×10^{-8}	7.08×10^{-9}	5.00×10^{-9}	3.85×10^{-9}	3.11×10^{-9}
20	1.45×10^{-11}	2.85×10^{-9}	1.78×10^{-9}	1.05×10^{-9}	7.39×10^{-10}	5.69×10^{-10}	4.62×10^{-10}
21	8.15×10^{-12}	2.23×10^{-10}	1.36×10^{-10}	7.74×10^{-11}	5.26×10^{-11}	3.90×10^{-11}	3.03×10^{-11}
22	7.74×10^{-12}	5.01×10^{-10}	3.11×10^{-10}	1.81×10^{-10}	1.26×10^{-10}	9.60×10^{-11}	7.70×10^{-11}
23	4.90×10^{-12}	6.02×10^{-10}	3.75×10^{-10}	2.20×10^{-10}	1.55×10^{-10}	1.19×10^{-10}	9.61×10^{-11}
24	2.17×10^{-12}	3.94×10^{-10}	2.46×10^{-10}	1.44×10^{-10}	1.02×10^{-10}	7.86×10^{-11}	6.38×10^{-11}
25	6.96×10^{-13}	1.73×10^{-10}	1.08×10^{-10}	6.38×10^{-11}	4.51×10^{-11}	3.48×10^{-11}	2.83×10^{-11}
26	1.55×10^{-13}	5.23×10^{-11}	3.26×10^{-11}	1.93×10^{-11}	1.36×10^{-11}	1.05×10^{-11}	8.58×10^{-12}
27	2.65×10^{-14}	9.24×10^{-12}	5.77×10^{-12}	3.41×10^{-12}	2.41×10^{-12}	1.86×10^{-12}	1.52×10^{-12}
28	1.24×10^{-14}	6.84×10^{-13}	4.23×10^{-13}	2.46×10^{-13}	1.71×10^{-13}	1.30×10^{-13}	1.04×10^{-13}
29	1.24×10^{-14}	9.64×10^{-13}	5.98×10^{-13}	3.49×10^{-13}	2.44×10^{-13}	1.87×10^{-13}	1.50×10^{-13}
30	8.94×10^{-15}	1.48×10^{-12}	9.22×10^{-13}	5.42×10^{-13}	3.83×10^{-13}	2.95×10^{-13}	2.39×10^{-13}
31	4.64×10^{-15}	1.16×10^{-12}	7.24×10^{-13}	4.27×10^{-13}	3.02×10^{-13}	2.33×10^{-13}	1.89×10^{-13}
32	1.80×10^{-15}	6.14×10^{-13}	3.83×10^{-13}	2.26×10^{-13}	1.60×10^{-13}	1.24×10^{-13}	1.01×10^{-13}
33	5.19×10^{-16}	2.34×10^{-13}	1.46×10^{-13}	8.64×10^{-14}	6.12×10^{-14}	4.74×10^{-14}	3.86×10^{-14}
34	1.07×10^{-16}	6.10×10^{-14}	3.81×10^{-14}	2.25×10^{-14}	1.60×10^{-14}	1.24×10^{-14}	1.01×10^{-14}
35	2.47×10^{-17}	8.40×10^{-15}	5.25×10^{-15}	3.10×10^{-15}	2.19×10^{-15}	1.69×10^{-15}	1.38×10^{-15}
36	1.97×10^{-17}	1.00×10^{-15}	6.18×10^{-16}	3.58×10^{-16}	2.48×10^{-16}	1.88×10^{-16}	1.50×10^{-16}
37	1.82×10^{-17}	2.71×10^{-15}	1.69×10^{-15}	9.94×10^{-16}	7.00×10^{-16}	5.39×10^{-16}	4.37×10^{-16}
38	1.23×10^{-17}	3.31×10^{-15}	2.07×10^{-15}	1.22×10^{-15}	8.63×10^{-16}	6.66×10^{-16}	5.42×10^{-16}
39	6.20×10^{-18}	2.40×10^{-15}	1.50×10^{-15}	8.85×10^{-16}	6.27×10^{-16}	4.85×10^{-16}	3.95×10^{-16}
40	2.42×10^{-18}	1.24×10^{-15}	7.78×10^{-16}	4.59×10^{-16}	3.26×10^{-16}	2.52×10^{-16}	2.05×10^{-16}

TABLE II. The differences between E_3 and the various bounds to E_3 are listed for different numbers N of terms in Φ when $\lambda = 1.0$.

N	$E_{RR} - E_3$	$E_3 - E_T$	$E_3 - E''$ $k=4$	$E_3 - E''$ $k=6$	$E_3 - E''$ $k=8$	$E_3 - E''$ $k=10$	$E_3 - E''$ $k=12$
3	2.67×10		2.33	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
4	9.72	4.41×10^{2}	6.65×10^{-1}	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
5	3.68	3.39×10	1.90	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
6	1.25	1.35×10	1.99	2.67×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
7	3.38×10^{-1}	5.49	1.92	2.08×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
8	6.28×10^{-2}	1.66	1.53	2.13×10^{-1}	7.90×10^{-2}	2.38×10^{-2}	6.67×10^{-4}
9	9.68×10^{-3}	2.54×10^{-1}	6.18×10^{-1}	1.24×10^{-1}	7.36×10^{-2}	2.38×10^{-2}	6.67×10^{-4}
10	6.64×10^{-3}	2.65×10^{-2}	1.07×10^{-1}	2.25×10^{-2}	1.05×10^{-2}	4.71×10^{-3}	6.67×10^{-4}
11	5.89×10^{-3}	7.88×10^{-2}	2.56×10^{-1}	5.78×10^{-2}	3.23×10^{-2}	9.94×10^{-3}	6.67×10^{-4}
12	3.31×10^{-3}	8.82×10^{-2}	2.76×10^{-1}	6.41×10^{-2}	3.71×10^{-2}	1.23×10^{-2}	3.35×10^{-4}
13	1.26×10^{-3}	5.34×10^{-2}	1.77×10^{-1}	4.35×10^{-2}	2.52×10^{-2}	1.04×10^{-2}	4.69×10^{-4}
14	3.31×10^{-4}	2.08×10^{-2}	7.36×10^{-2}	1.90×10^{-2}	1.09×10^{-2}	5.94×10^{-3}	5.35×10^{-4}
15	5.70×10^{-5}	5.04×10^{-3}	1.84×10^{-2}	4.87×10^{-3}	2.80×10^{-3}	1.82×10^{-3}	4.39×10^{-4}
16	1.21×10^{-5}	5.47×10^{-4}	2.03×10^{-3}	5.37×10^{-4}	3.05×10^{-4}	2.10×10^{-4}	1.25×10^{-4}
17	1.08×10^{-5}	1.36×10^{-4}	5.25×10^{-4}	1.33×10^{-4}	7.25×10^{-5}	4.76×10^{-5}	3.15×10^{-5}
18	8.96×10^{-6}	3.25×10^{-4}	1.21×10^{-3}	3.20×10^{-4}	1.81×10^{-4}	1.24×10^{-4}	8.05×10^{-5}
19	4.99×10^{-6}	3.13×10^{-4}	1.16×10^{-3}	3.08×10^{-4}	1.76×10^{-4}	1.21×10^{-4}	8.05×10^{-5}
20	2.00×10^{-6}	1.82×10^{-4}	6.73×10^{-4}	1.80×10^{-4}	1.03×10^{-4}	7.16×10^{-5}	5.04×10^{-5}
21	5.76×10^{-7}	7.33×10^{-5}	2.70×10^{-4}	7.22×10^{-5}	4.15×10^{-5}	2.90×10^{-5}	2.15×10^{-5}
22	1.15×10^{-7}	1.97×10^{-5}	7.24×10^{-5}	1.94×10^{-5}	1.11×10^{-5}	7.80×10^{-6}	5.93×10^{-6}
23	2.27×10^{-8}	2.77×10^{-6}	1.02×10^{-5}	2.73×10^{-6}	1.57×10^{-6}	1.09×10^{-6}	8.37×10^{-7}
24	1.68×10^{-8}	2.92×10^{-7}	1.11×10^{-6}	2.87×10^{-7}	1.59×10^{-7}	1.07×10^{-7}	7.84×10^{-8}
25	1.55×10^{-8}	7.85×10^{-7}	2.92×10^{-6}	7.73×10^{-7}	4.40×10^{-7}	3.05×10^{-7}	2.31×10^{-7}
26	1.01×10^{-8}	9.48×10^{-7}	3.50×10^{-6}	9.33×10^{-7}	5.35×10^{-7}	3.73×10^{-7}	2.85×10^{-7}
27	4.80×10^{-9}	6.61×10^{-7}	2.43×10^{-6}	6.51×10^{-7}	3.74×10^{-7}	2.61×10^{-7}	2.00×10^{-7}
28	1.72×10^{-9}	3.22×10^{-7}	1.18×10^{-6}	3.17×10^{-7}	1.83×10^{-7}	1.28×10^{-7}	9.81×10^{-8}
29	4.52×10^{-10}	1.13×10^{-7}	4.15×10^{-7}	1.11×10^{-7}	6.41×10^{-8}	4.49×10^{-8}	3.45×10^{-8}
30	8.79×10^{-11}	2.59×10^{-8}	9.51×10^{-8}	2.55×10^{-8}	1.47×10^{-8}	1.03×10^{-8}	7.93×10^{-9}
31	2.80×10^{-11}	2.74×10^{-9}	1.01×10^{-8}	2.70×10^{-9}	1.55×10^{-9}	1.08×10^{-9}	8.26×10^{-10}
32	2.67×10^{-11}	9.12×10^{-10}	3.41×10^{-9}	8.98×10^{-10}	5.08×10^{-10}	3.49×10^{-10}	2.63×10^{-10}
33	2.27×10^{-11}	2.18×10^{-9}	8.05×10^{-9}	2.15×10^{-9}	1.23×10^{-9}	8.59×10^{-10}	6.57×10^{-10}
34	1.41×10^{-11}	2.25×10^{-9}	8.28×10^{-9}	2.22×10^{-9}	1.27×10^{-9}	8.92×10^{-10}	6.84×10^{-10}
35	6.64×10^{-12}	1.49×10^{-9}	5.47×10^{-9}	1.46×10^{-9}	8.43×10^{-10}	5.91×10^{-10}	4.54×10^{-10}
36	2.42×10^{-12}	7.19×10^{-10}	2.64×10^{-9}	7.08×10^{-10}	4.08×10^{-10}	2.86×10^{-10}	2.20×10^{-10}
37	6.66×10^{-13}	2.57×10^{-10}	9.43×10^{-10}	2.53×10^{-10}	1.46×10^{-10}	1.02×10^{-10}	7.88×10^{-11}
38	1.40×10^{-13}	6.20×10^{-11}	2.27×10^{-10}	6.11×10^{-11}	3.52×10^{-11}	2.47×10^{-11}	1.90×10^{-11}
39	4.57×10^{-14}	7.31×10^{-12}	2.69×10^{-11}	7.20×10^{-12}	4.14×10^{-12}	2.90×10^{-12}	2.22×10^{-12}
40	4.26×10^{-14}	1.79×10^{-12}	6.66×10^{-12}	1.76×10^{-12}	9.99×10^{-13}	6.90×10^{-13}	5.22×10^{-13}

that value of N. Subsection A below shows how to choose the a_l to get the best Temple lower bound for a given N. Subsection B shows how to choose the a_l to optimize the new tight lower bound for a given N.

A. Optimization of the Temple lower bound

We begin with the observation that the usual formula (1) for the Temple lower bound E_T can be rearranged to take the form

$$E_T = \widetilde{E}_{n+1} + \frac{\langle \Phi |(H - \widetilde{E}_{n+1} I)^2 | \Phi \rangle}{\langle \Phi |(H - \widetilde{E}_{n+1} I) | \Phi \rangle}. \tag{23}$$

It follows that the best Temple bound $\widetilde{E}_T$ can be characterized by the minimum principle

$$(\widetilde{E}_T - \widetilde{E}_{n+1})^{-1} = \min \frac{\langle \Phi |(H - \widetilde{E}_{n+1} I) | \Phi \rangle}{\langle \Phi |(H - \widetilde{E}_{n+1} I)^2 | \Phi \rangle}. \tag{24}$$

The minimum principle (24) for the lower bound can also be obtained via the procedure of Lehman[18] and Maehly,[18] which amounts to the calculation of Rayleigh–Ritz bounds on the operator $(H - \widetilde{E}_{n+1} I)^{-1}$ with trial functions of the form $(H - \widetilde{E}_{n+1} I)|\Phi\rangle$. Assume now that $|\Phi\rangle$ is constrained to have the form

$$|\Phi\rangle = \sum_{l=1}^{N} a_l |\zeta_l\rangle, \tag{25}$$

where $|\zeta_l\rangle$ are members of a suitably chosen orthonormal set and the a_l are (linear) variational parameters to be determined. Varying (24) with $|\Phi\rangle$ given the form (25) leads, via a standard argument, to the matrix eigenvalue problem

$$\sum_{l=1}^{N} \langle \zeta_k |(H - \widetilde{E}_{n+1} I)^2 | \zeta_l \rangle a_l = (\widetilde{E}_T - \widetilde{E}_{n+1}) \sum_{l=1}^{N} \langle \zeta_k |(H - \widetilde{E}_{n+1} I) | \zeta_l \rangle a_l. \tag{26}$$

If $\widetilde{E}_{n+1}$ is not an eigenvalue of H so that $(H - \widetilde{E}_{n+1} I)^2$ is positive definite, the existence of solutions to (26) follows from the minimum principle (24) by invoking the theorem of Weierstrass that a continuous function of several variables which are restricted to a finite closed domain assumes a minimum on that domain. The finite, closed domain can, for example, then be chosen to be $\Sigma_{k,l} \langle \zeta_k |(H - \widetilde{E}_{n+1} I)^2 | \zeta_l \rangle \bar{a}_k a_l = 1$. Actually, it is sufficient that $\widetilde{E}_{n+1}$ not be an eigenvalue of the $N \times N$ matrix $\langle \zeta_k | H | \zeta_l \rangle$.

The numerical solution of the matrix eigenvalue problem (26) is hampered by the fact that the matrices

$\langle \zeta_k | (H - \widetilde{E}_{n+1} I)^2 | \zeta_l \rangle$ and $\langle \zeta_k | (H - \widetilde{E}_{n+1} I) | \zeta_l \rangle$ are almost singular if $\widetilde{E}_{n+1}$ is close to the $(n+1)$th eigenvalue of $\langle \zeta_k | H | \zeta_l \rangle$. An effective iterative procedure, which circumvents this problem and exploits the fact that the vector which gives the best Rayleigh–Ritz upper bound is close to the vector which optimizes the Temple bound, follows.

The iterative procedure starts with the nth and $(n+1)$th eigenvectors $\mathbf{a}^{(n)}$, $\mathbf{a}^{(n+1)}$ of $\langle \zeta_k | H | \zeta_l \rangle$, whose components will be denoted by $a_l^{(n)}$ and $a_l^{(n+1)}$. The vector $\mathbf{a}^{(n)}$ is presumably a good starting approximation to the vector which optimizes the Temple bound; the vector $\mathbf{a}^{(n+1)}$ is the vector whose presence causes trouble when $\widetilde{E}_{n+1}$ is close to the $(n+1)$th eigenvalue of $\langle \zeta_k | H | \zeta_l \rangle$. The Temple bound $E_{\rm T}^{(1)}$ which results from using (25) with $a_l = a_l^{(n)}$ is calculated. The matrix A with elements

$$A_{k,l} = \langle \zeta_k | (H - \widetilde{E}_{n+1} I)^2 | \zeta_l \rangle - (E_T^{(1)} - \widetilde{E}_{n+1}) \times \langle \zeta_k | (H - \widetilde{E}_{n+1} I) | \zeta_l \rangle \tag{27}$$

and the vector $A\mathbf{a}^{(n)}$ are then computed. Let $\|\mathbf{a}\|$ denote the usual Euclidean norm

$$\|\mathbf{a}\| = \left(\sum_{l=1}^{N} |a_l|^2 \right)^{1/2} \tag{28}$$

of the vector $\mathbf{a}$. The iteration stops when $\|A\mathbf{a}^{(n)}\|^2 / \|\mathbf{a}^{(n)}\|^2$ is less than some prescribed small parameter ϵ. If $\|A\mathbf{a}^{(n)}\|^2 / \|\mathbf{a}^{(1)}\|^2$ is not sufficiently small, the matrix B, given by

$$B = A^2 + I_N \delta \tag{29}$$

is computed. Here I_N is the $N \times N$ identity matrix and δ is a small parameter; the term $I_N \delta$ is inserted to prevent the matrix B from becoming too singular. The linear equations

$$B\mathbf{b}^{(i)} = \mathbf{a}^{(n+i-1)}, \quad i = 1,2 \tag{30}$$

are then solved to obtain new vectors $\mathbf{b}^{(i)}$.[19] The ratio $\|A\mathbf{b}^{(i)}\| / \|\mathbf{b}^{(i)}\|$ will in general be smaller than the ratio $\|A\mathbf{a}^{(n+i-1)}\| / \|\mathbf{a}^{(n+i-1)}\|$. The matrix eigenvalue problem (26) is solved exactly (to machine accuracy) in the two dimensional subspace spanned by the vectors $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$ to obtain new N-dimensional vectors $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ which are the eigenvectors of (26) in the subspace spanned by $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$. If the eigenvalues of (26) in the subspace spanned by $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$ are degenerate or almost degenerate, the eigenvectors are not computed and $\mathbf{c}^{(1)}$, $\mathbf{c}^{(2)}$ are taken equal to $\mathbf{b}^{(1)}$, $\mathbf{b}^{(2)}$. The procedure is then repeated with $\mathbf{a}^{(n)}$ replaced by the member of the pair $\mathbf{c}^{(1)}$, $\mathbf{c}^{(2)}$ which gives the best Temple bound, and $\mathbf{a}^{(n+1)}$ replaced by a linear combination of the $\mathbf{c}^{(i)}$ which is orthogonal to the new $\mathbf{a}^{(n)}$. This orthogonalization is needed to prevent the $\mathbf{b}^{(i)}$ from approaching linear dependence as the iteration proceeds. It was found essential to carry along the approximations $\mathbf{a}^{(n+1)}$, $\mathbf{b}^{(2)}$, $\mathbf{c}^{(2)}$ to the "troublemaking" vector and solve exactly in the subspace spanned by $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$; without this refinement the iteration did not converge.

Table III shows optimized Temple lower bounds to E_1 obtained for six different values of the lower bound $\widetilde{E}_2$ to the first excited state. Comparison with the unoptimized Temple bounds of Table I shows that the improvement obtained fluctuates with N almost in periodic fashion: the ratio of the error $E_1 - E_{\rm T}$ for the unoptimized bound to the error $E_1 - E_{\rm T}$ for the optimized bound, which is a quanitative measure of the improvement, has maxima for $N = 2$, 4 or 5, 10 or 11, 17, 23 or 24, 31, and 39, and minima for $N = 3$, 8, 14, 21, 28, and 36. The various values of $\widetilde{E}_2$ used are obtained from $\widetilde{E}_2 = \alpha E_{1\,\rm ub} + (1-\alpha) E_{2\,\rm lb}$ where $E_{1\,\rm ub} = 1.39235164153029185566$ is the 52-term Rayleigh–Ritz upper bound, which is exact to the 21 significant figures shown, and $E_{2\,\rm lb} = 8.65504995775930968779$ is the $k = 51$ lower bound used for the unoptimized Temple bounds, which is exact to 19 significant figures. It is straightforward to prove that the Temple bound (optimized or unoptimized) improves as $\widetilde{E}_{n+1}$ moves closer to E_{n+1}. Comparing the results for different values of $\widetilde{E}_2$ shows when the error is significant: The error $E_1 - E_{\rm T}$ in the optimized Temple bound is sensitive to the value of $\widetilde{E}_2$ only when the error $E_2 - \widetilde{E}_2$ in the $\widetilde{E}_2$ lower bound is within an order of magnitude of $E_1 - E_{\rm T}$. The unoptimized Temple bound showed no such sensitivity to the value of $\widetilde{E}_2$. Similar results (not shown) have been obtained for optimized Temple lower bounds to excited states.

The parameters δ and ϵ in the iterative optimization scheme were given the values $\delta = 10^{-5}$ and $\epsilon = 10^{-28}$. With these values of δ and ϵ, computation of the optimized Temple lower bound to E_1 required at most three iterations, except for $\alpha = 0.1$ and $N = 3$, where five iterations were needed. For N greater than 25, one iteration was sufficient. The choices of ϵ and δ were not critical.

The approach of Lehman[18] and Maehly[18] to the optimized Temple bound shows that a lower bound to E_{n+p+1} is given by the pth eigenvalue of (26). Some lower bounds to E_1 of this type are shown in Table IV for $n = p = 2$ and for $n = p = 3$. The various values of $\widetilde{E}_{n+1}$ used are obtained from $\widetilde{E}_{n+1} = \alpha E_{n\rm ub} + (1-\alpha) E_{(n+1)\rm lb}$, where the upper bounds $E_{2\,\rm ub}, E_{3\,\rm ub}$, and the lower bounds $E_{3\rm lb}, E_{4\rm lb}$ agree with the exact values of E_2, E_3, and E_4 to at least 15 significant figures. The bounds for $n = p = 3$ are better than those for $n = p = 1$ and $n = p = 2$ for all N greater than 8. The bounds for $n = p = 2$ are better than those for $n = p = 1$ (for a given α) except for $N = 10$, 16, and 17 when $\alpha = 10^{-8}$ and $N = 16$, 17, 23, 25, 30, 31, and 32 when $\alpha = 10^{-12}$. Again significant improvement in the lower bound is obtained when the difference $E_{n+1} - \widetilde{E}_{n+1}$ is within an order of magnitude of $E_1 - E_{\rm T}$. The lower bounds are more accurate than the Rayleigh–Ritz upper bounds, for the same N, for $N = 3$–6 when $\alpha = 10^{-2}$ and $n = p = 2$, for $N = 3$–7 when $\alpha = 10^{-8}$ or 10^{-12} and $n = p = 2$, for $N = 9$–11 when $\alpha = 10^{-2}$ and $n = p = 3$, and for $n = 9$–12 and 14–33 when $\alpha = 10^{-12}$ and $n = p = 3$.

The improvement obtained by making $E_{n+1} - \widetilde{E}_{n+1}$ of the order of the accuracy desired suggests an iterative procedure which starts with rough lower bounds $\widetilde{E}_{n+1}$ and uses the Lehman–Maehly optimized Temple bounds obtained at a given step as improved $\widetilde{E}_{n+1}$ for use in the computation of better Lehman–Maehly optimized Temple lower bounds at the next step.

B. Optimization of the new tight lower bound

The best new tight lower bound to the nth eigenvalue of the original Hamiltonian H will be obtained when the nth

TABLE III. The differences $E_{RR} - E_1$ between E_1 and the Rayleigh–Ritz upper bound and the differences $E_1 - E_T$ between E_1 and the optimized Temple lower bound to E_1 for different lower bounds $\widetilde{E}_2$ to the first excited state are listed for different numbers N of terms in Φ when $\lambda = 1.0$. $\widetilde{E}_2$ is given by $\widetilde{E}_2 = \alpha E_{1\,ub} + (1-\alpha)E_{2\,lb}$, where $E_{1\,ub}$ is a Rayleigh–Ritz upper bound to E_1 which is exact to 21 significant figures and $E_{2\,lb}$ is a lower bound to E_2 which is exact to 19 significant figures.

N	$E_{RR} - E_1$	$E_1 - E_T$ $\alpha = 10^{-1}$	$E_1 - E_T$ $\alpha = 10^{-2}$	$E_1 - E_T$ $\alpha = 10^{-4}$	$E_1 - E_T$ $\alpha = 10^{-8}$	$E_1 - E_T$ $\alpha = 10^{-12}$	$E_1 - E_T$ $\alpha = 0$
1	3.58×10^{-1}	6.13×10^{-1}	5.21×10^{-1}	5.11×10^{-1}	5.11×10^{-1}	5.11×10^{-1}	5.11×10^{-1}
2	2.03×10^{-2}	5.75×10^{-2}	4.78×10^{-2}	4.68×10^{-2}	4.68×10^{-2}	4.68×10^{-2}	4.68×10^{-2}
3	2.72×10^{-3}	8.36×10^{-3}	7.32×10^{-3}	7.21×10^{-3}	7.21×10^{-3}	7.21×10^{-3}	7.21×10^{-3}
4	2.56×10^{-3}	3.52×10^{-3}	2.10×10^{-3}	1.93×10^{-3}	1.92×10^{-3}	1.92×10^{-3}	1.92×10^{-3}
5	1.02×10^{-3}	3.50×10^{-3}	1.57×10^{-3}	1.08×10^{-3}	1.07×10^{-3}	1.07×10^{-3}	1.07×10^{-3}
6	2.14×10^{-4}	1.81×10^{-3}	1.39×10^{-3}	1.07×10^{-3}	1.07×10^{-3}	1.07×10^{-3}	1.07×10^{-3}
7	2.33×10^{-5}	4.23×10^{-4}	3.74×10^{-4}	3.58×10^{-4}	3.58×10^{-4}	3.58×10^{-4}	3.58×10^{-4}
8	3.85×10^{-6}	5.75×10^{-5}	5.17×10^{-5}	5.07×10^{-5}	5.07×10^{-5}	5.07×10^{-5}	5.07×10^{-5}
9	3.85×10^{-6}	2.54×10^{-5}	2.08×10^{-5}	1.14×10^{-5}	1.09×10^{-5}	1.09×10^{-5}	1.09×10^{-5}
10	2.28×10^{-6}	2.42×10^{-5}	2.04×10^{-5}	5.84×10^{-6}	3.87×10^{-6}	3.87×10^{-6}	3.87×10^{-6}
11	8.02×10^{-7}	1.37×10^{-5}	1.22×10^{-5}	5.72×10^{-6}	2.35×10^{-6}	2.35×10^{-6}	2.35×10^{-6}
12	1.77×10^{-7}	4.91×10^{-6}	4.41×10^{-6}	3.89×10^{-6}	2.31×10^{-6}	2.31×10^{-6}	2.31×10^{-6}
13	2.34×10^{-8}	1.11×10^{-6}	1.01×10^{-6}	9.96×10^{-7}	9.94×10^{-7}	9.94×10^{-7}	9.94×10^{-7}
14	5.54×10^{-9}	1.69×10^{-7}	1.53×10^{-7}	1.51×10^{-7}	1.50×10^{-7}	1.50×10^{-7}	1.50×10^{-7}
15	5.52×10^{-9}	6.86×10^{-8}	6.18×10^{-8}	5.90×10^{-8}	3.22×10^{-8}	3.21×10^{-8}	3.21×10^{-8}
16	3.74×10^{-9}	6.79×10^{-8}	6.12×10^{-8}	5.88×10^{-8}	1.06×10^{-8}	1.04×10^{-8}	1.04×10^{-8}
17	1.64×10^{-9}	4.51×10^{-8}	4.07×10^{-8}	3.98×10^{-8}	5.43×10^{-9}	5.13×10^{-9}	5.13×10^{-9}
18	4.89×10^{-10}	2.01×10^{-8}	1.81×10^{-8}	1.79×10^{-8}	4.46×10^{-9}	3.94×10^{-9}	3.94×10^{-9}
19	9.64×10^{-11}	6.22×10^{-9}	5.63×10^{-9}	5.57×10^{-9}	4.16×10^{-9}	3.86×10^{-9}	3.86×10^{-9}
20	1.45×10^{-11}	1.30×10^{-9}	1.18×10^{-9}	1.16×10^{-9}	1.16×10^{-9}	1.16×10^{-9}	1.16×10^{-9}
21	8.15×10^{-12}	2.36×10^{-10}	2.14×10^{-10}	2.12×10^{-10}	2.05×10^{-10}	2.03×10^{-10}	2.03×10^{-10}
22	7.74×10^{-12}	1.56×10^{-10}	1.41×10^{-10}	1.40×10^{-10}	8.32×10^{-11}	5.06×10^{-11}	5.06×10^{-11}
23	4.90×10^{-12}	1.47×10^{-10}	1.33×10^{-10}	1.32×10^{-10}	8.00×10^{-11}	1.81×10^{-11}	1.81×10^{-11}
24	2.17×10^{-12}	9.53×10^{-11}	8.63×10^{-11}	8.54×10^{-11}	6.81×10^{-11}	9.21×10^{-12}	9.18×10^{-12}
25	6.96×10^{-13}	4.41×10^{-11}	3.99×10^{-11}	3.95×10^{-11}	3.72×10^{-11}	6.83×10^{-12}	6.78×10^{-12}
26	1.55×10^{-13}	1.49×10^{-11}	1.35×10^{-11}	1.33×10^{-11}	1.33×10^{-11}	6.82×10^{-12}	6.78×10^{-12}
27	2.65×10^{-14}	3.57×10^{-12}	3.24×10^{-12}	3.20×10^{-12}	3.20×10^{-12}	3.17×10^{-12}	3.17×10^{-12}
28	1.24×10^{-14}	6.77×10^{-13}	6.14×10^{-13}	6.08×10^{-13}	6.08×10^{-13}	6.01×10^{-13}	6.01×10^{-13}
29	1.24×10^{-14}	3.27×10^{-13}	2.96×10^{-13}	2.93×10^{-13}	2.92×10^{-13}	1.46×10^{-13}	1.43×10^{-13}
30	8.94×10^{-15}	3.27×10^{-13}	2.96×10^{-13}	2.92×10^{-13}	2.92×10^{-13}	5.65×10^{-14}	4.73×10^{-14}
31	4.64×10^{-15}	2.46×10^{-13}	2.23×10^{-13}	2.21×10^{-13}	2.20×10^{-13}	3.78×10^{-14}	2.13×10^{-14}
32	1.80×10^{-15}	1.35×10^{-13}	1.22×10^{-13}	1.21×10^{-13}	1.21×10^{-13}	3.67×10^{-14}	1.31×10^{-14}
33	5.19×10^{-16}	5.58×10^{-14}	5.06×10^{-14}	5.01×10^{-14}	5.01×10^{-14}	3.23×10^{-14}	1.13×10^{-14}
34	1.07×10^{-16}	1.73×10^{-14}	1.57×10^{-14}	1.56×10^{-14}	1.56×10^{-14}	1.48×10^{-14}	1.07×10^{-14}
35	2.47×10^{-17}	3.94×10^{-15}	3.57×10^{-15}	3.54×10^{-15}	3.54×10^{-15}	3.54×10^{-15}	3.52×10^{-15}
36	1.97×10^{-17}	9.16×10^{-16}	8.31×10^{-16}	8.23×10^{-16}	8.23×10^{-16}	8.13×10^{-16}	7.31×10^{-16}
37	1.82×10^{-17}	6.61×10^{-16}	5.99×10^{-16}	5.93×10^{-16}	5.92×10^{-16}	5.62×10^{-16}	1.99×10^{-16}
38	1.23×10^{-17}	6.37×10^{-16}	5.77×10^{-16}	5.71×10^{-16}	5.71×10^{-16}	5.48×10^{-16}	7.27×10^{-17}
39	6.20×10^{-18}	4.57×10^{-16}	4.14×10^{-16}	4.10×10^{-16}	4.10×10^{-16}	4.02×10^{-16}	3.49×10^{-17}
40	2.42×10^{-18}	2.48×10^{-16}	2.25×10^{-16}	2.23×10^{-16}	2.23×10^{-16}	2.21×10^{-16}	2.22×10^{-17}

eigenvalue E''_n of H'' is a maximum as a function of the parameters in the trial function $|\Phi\rangle$ (provided, of course, that $E''_n \leqslant E_{lb}$). Assume again that $|\Phi\rangle$ is constrained to have the form (25). Necessary conditions for E''_n to be a maximum are that

$$\partial E''_n/\partial a_l = \partial E''_n/\partial \bar{a}_l = 0, \quad l = 1,2,\ldots,N \tag{31}$$

(sufficiency will be discussed below). The derivatives in (31) can be computed from the Hellmann–Feynman theorem (which holds for finite dimensional matrices when the eigenvalues are nondegenerate):

$$\partial E''_n/\partial a_l = (\langle\psi_n|\psi_n\rangle)^{-1}\langle\psi_n|\partial H''/\partial a_l|\psi_n\rangle, \quad l = 1,2,\ldots,N. \tag{32}$$

Here $|\psi_n\rangle$ is the eigenfunction of H'' belonging to E''_n. The conditions (31) then imply that either

$$\langle\Phi|H_2|\psi_n\rangle = 0 \tag{33}$$

or

$$\langle\zeta_l|H_2|\Phi\rangle = (\langle\Phi|H_2|\psi_n\rangle)^{-1}\langle\Phi|H_2|\Phi\rangle\,\langle\zeta_l|H_2|\psi_n\rangle, \quad l = 1,2,\ldots,N. \tag{34}$$

If (33) is satisfied, E''_n is the same as the nth eigenvalue of H_1 and the term in H'' which contains $|\Phi\rangle$ does not improve the lower bound. On the other hand, (25) and (34) imply that

$$a_k = C\sum_{l=1}^{N} G_{k,l}\langle\zeta_l|H_2|\psi_n\rangle + b_k, \tag{35}$$

where $C = (\langle\Phi|H_2|\psi_n\rangle)^{-1}\langle\Phi|H_2|\Phi\rangle$ is an undetermined nonzero constant, $G_{k,l}$ is the generalized inverse of the $N\times N$ matrix $\langle\zeta_k|H_2|\zeta_l\rangle$, and the b_k are the components of an arbitrary vector $\mathbf{b}$ in the null space of $\langle\zeta_k|H_2|\zeta_l\rangle$. The generalized inverse $G_{k,l}$ is defined in the usual way by

$$\sum_{l=1}^{N} G_{k,l}\langle\zeta_k|H_2|\zeta_m\rangle = Q_{k,m} \tag{36}$$

TABLE VI. The difference between E_3 and the Rayleigh–Ritz upper bound E_{RR}, the difference between E_3 and the optimized Temple bound E_T for $\widetilde{E}_4 = E_{4\,\mathrm{lb}}$, and the differerences between E_3 and the optimized new tight lower bounds E'' are listed for different numbers N of terms in Φ when $\lambda = 1.0$.

N	$E_{RR} - E_3$	$E_3 - E_T$ $n = 3, p = 1$	$E_3 - E''$ $k = 4$	$E_3 - E''$ $k = 6$	$E_3 - E''$ $k = 8$	$E_3 - E''$ $k = 10$	$E_3 - E''$ $k = 12$
3	2.67×10		2.33	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
4	9.72	9.12	6.65×10^{-1}	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
5	3.68	8.94	3.81×10^{-1}	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
6	1.25	2.52	3.69×10^{-1}	2.67×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
7	3.38×10^{-1}	6.39×10^{-1}	3.54×10^{-1}	1.87×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
8	6.28×10^{-2}	1.79×10^{-1}	2.89×10^{-1}	9.83×10^{-2}	7.90×10^{-2}	2.38×10^{-2}	6.67×10^{-4}
9	9.68×10^{-3}	5.53×10^{-2}	1.71×10^{-1}	4.22×10^{-2}	2.75×10^{-2}	2.38×10^{-2}	6.67×10^{-4}
10	6.64×10^{-3}	2.00×10^{-2}	7.23×10^{-2}	1.72×10^{-2}	7.96×10^{-3}	4.71×10^{-3}	6.67×10^{-4}
11	5.89×10^{-3}	9.64×10^{-3}	5.75×10^{-2}	1.15×10^{-2}	2.86×10^{-3}	8.68×10^{-4}	6.67×10^{-4}
12	3.31×10^{-3}	6.89×10^{-3}	5.44×10^{-2}	1.15×10^{-2}	2.28×10^{-3}	4.75×10^{-4}	3.35×10^{-4}
13	1.26×10^{-3}	6.86×10^{-3}	3.66×10^{-2}	8.16×10^{-3}	2.21×10^{-3}	4.73×10^{-4}	3.22×10^{-4}
14	3.31×10^{-4}	3.94×10^{-3}	1.69×10^{-2}	3.82×10^{-3}	1.48×10^{-3}	3.69×10^{-4}	2.14×10^{-4}
15	5.70×10^{-5}	8.97×10^{-4}	5.24×10^{-3}	1.23×10^{-3}	5.84×10^{-4}	2.03×10^{-4}	1.04×10^{-4}
16	1.21×10^{-5}	2.19×10^{-4}	1.09×10^{-3}	2.70×10^{-4}	1.41×10^{-4}	7.10×10^{-5}	3.96×10^{-5}
17	1.08×10^{-5}	7.05×10^{-5}	3.00×10^{-4}	7.23×10^{-5}	3.63×10^{-5}	2.10×10^{-5}	1.35×10^{-5}
18	8.96×10^{-6}	3.07×10^{-5}	2.76×10^{-4}	6.39×10^{-5}	3.00×10^{-5}	1.63×10^{-5}	8.68×10^{-6}
19	4.99×10^{-6}	1.89×10^{-5}	2.35×10^{-4}	5.59×10^{-5}	2.72×10^{-5}	1.53×10^{-5}	8.64×10^{-6}
20	2.00×10^{-6}	1.69×10^{-5}	1.40×10^{-4}	3.41×10^{-5}	1.73×10^{-5}	1.02×10^{-5}	6.32×10^{-6}
21	5.76×10^{-7}	1.47×10^{-5}	6.05×10^{-5}	1.50×10^{-5}	7.90×10^{-6}	4.85×10^{-6}	3.18×10^{-6}
22	1.15×10^{-7}	4.08×10^{-6}	1.91×10^{-5}	4.84×10^{-6}	2.61×10^{-6}	1.67×10^{-6}	1.14×10^{-6}
23	2.27×10^{-8}	8.75×10^{-7}	4.28×10^{-6}	1.11×10^{-6}	6.11×10^{-7}	4.06×10^{-7}	2.91×10^{-7}
24	1.68×10^{-8}	2.48×10^{-7}	9.55×10^{-7}	2.44×10^{-7}	1.34×10^{-7}	8.87×10^{-8}	6.42×10^{-8}
25	1.55×10^{-8}	9.36×10^{-8}	6.92×10^{-7}	1.71×10^{-7}	8.94×10^{-8}	5.63×10^{-8}	3.81×10^{-8}
26	1.01×10^{-8}	4.74×10^{-8}	6.58×10^{-7}	1.64×10^{-7}	8.66×10^{-8}	5.51×10^{-8}	3.77×10^{-8}
27	4.80×10^{-9}	3.31×10^{-8}	4.55×10^{-7}	1.15×10^{-7}	6.16×10^{-8}	3.99×10^{-8}	2.79×10^{-8}
28	1.72×10^{-9}	3.18×10^{-8}	2.33×10^{-7}	5.94×10^{-8}	3.23×10^{-8}	2.12×10^{-8}	1.52×10^{-8}
29	4.52×10^{-10}	2.35×10^{-8}	9.05×10^{-8}	2.33×10^{-8}	1.28×10^{-8}	8.56×10^{-9}	6.21×10^{-9}
30	8.79×10^{-11}	5.82×10^{-9}	2.62×10^{-8}	6.83×10^{-9}	3.81×10^{-9}	2.58×10^{-9}	1.90×10^{-9}
31	2.80×10^{-11}	1.34×10^{-9}	5.68×10^{-9}	1.49×10^{-9}	8.41×10^{-10}	5.76×10^{-10}	4.31×10^{-10}
32	2.67×10^{-11}	4.11×10^{-10}	1.72×10^{-9}	4.42×10^{-10}	2.43×10^{-10}	1.63×10^{-10}	1.19×10^{-10}
33	2.27×10^{-11}	1.64×10^{-10}	1.55×10^{-9}	3.92×10^{-10}	2.12×10^{-10}	1.39×10^{-10}	9.93×10^{-11}
34	1.41×10^{-11}	8.54×10^{-11}	1.40×10^{-9}	3.56×10^{-10}	1.94×10^{-10}	1.28×10^{-10}	9.25×10^{-11}
35	6.64×10^{-12}	5.94×10^{-11}	9.33×10^{-10}	2.40×10^{-10}	1.32×10^{-10}	8.81×10^{-11}	6.42×10^{-11}
36	2.42×10^{-12}	5.56×10^{-11}	4.77×10^{-10}	1.24×10^{-10}	6.85×10^{-11}	4.61×10^{-11}	3.39×10^{-11}
37	6.66×10^{-13}	4.70×10^{-11}	1.90×10^{-10}	4.96×10^{-11}	2.77×10^{-11}	1.88×10^{-11}	1.40×10^{-11}
38	1.40×10^{-13}	1.40×10^{-11}	5.80×10^{-11}	1.52×10^{-11}	8.58×10^{-12}	5.87×10^{-12}	4.40×10^{-12}
39	4.57×10^{-14}	3.25×10^{-12}	1.35×10^{-11}	3.56×10^{-12}	2.02×10^{-12}	1.39×10^{-12}	1.05×10^{-12}
40	4.26×10^{-14}	9.70×10^{-13}	3.88×10^{-12}	1.01×10^{-12}	5.65×10^{-13}	3.84×10^{-13}	2.86×10^{-13}

The optimized E'' lower bounds can also be compared with the optimized Temple bounds E_T. $E_1 - E''$ tends to be smaller than $E_1 - E_T$ for $n = p = 1$, but there are a number of exceptions, which occur at and near $N = 10$, 17, 24, 31, and 39, for cases in which $E_2 - \widetilde{E}_2 \lesssim E_1 - E_T$. $E_3 - E''$ tends to be larger than the $n = 3, p = 1$ $E_3 - E_T$ for $k = 4$, except for $N < 8$. $E_3 - E''$ also tends to be larger for $k = 6$, except for $N < 11$ and $N = 14$, 24, or 29. $E_3 - E''$ tends to be smaller than $E_3 - E_T$ for $k = 8$,10, and 12, but there are a number of exceptions, which occur at and near $N = 11$, 19, 27, and 34, for cases in which $E_4 - \widetilde{E}_4 \lesssim E_3 - E_T$. Comparison of the $k = 10$ $E_1 - E''$ with the optimized Temple $E_1 - E_T$ for $n = p = 2$ and $n = p = 3$ shows that the optimized Temple bound is more accurate for $n = p = 2, N = 19$ when $\alpha = 10^{-8}$ or 10^{-12}, for $n = p = 3$, $N = 15$–31, when $\alpha = 10^{-8}$, and for $n = p = 3$, $N = 11$, 12, and 15–40 when $\alpha = 10^{-12}$. Thus it appears that the optimized Temple bounds tend to be better than the optimized E'' bounds when a lower bound $\widetilde{E}_{n+1}$ is known whose accuracy is comparable to the accuracy desired in the optimized Temple bound, but poorer than the E'' bounds when such a highly accurate $\widetilde{E}_{n+1}$ is not known. It would be interesting to see if these tendencies persist for other model problems.

Note added in proof: After the present paper had been submitted, the author became aware of a paper of M. Cohen and T. Feldman, J. Phys. B **12**, 2771–9 (1979), entitled "A generalization of Temple's lower bound to eigenvalues." The Cohen–Feldman lower bound appears to coincide with the Lehman-Maehly optimized Temple bound; in particular, their Eq. (29) is just the secular equation for the eigenvalue problem of Eq. (26) of the present paper.

ACKNOWLEDGMENTS

The author would like to thank Dr. David W. Fox for a helpful conversation about the Lehman–Maehly approach to the optimized Temple bound. This work was supported in part by the National Science Foundation under Grant No. PHY-7906936.

[1]The standard references on bounds to eigenvalues are: S.H. Gould, *Variational Methods for Eigenvalue Problems: An Introduction to the Weinstein*

TABLE V. The difference between E_1 and the Rayleigh–Ritz upper bound E_{RR}, the difference between E_1 and the optimized Temple bound E_T for $\widetilde{E}_2 = E_{2\,lb}$, and the differences between E_1 and the optimized new tight lower bounds E'' are listed for different numbers N of terms in Φ when $\lambda = 1.0$.

N	$E_{RR} - E_1$	$E_1 - E_T$ $n = p = 1$	$E_1 - E''$ $k = 2$	$E_1 - E''$ $k = 4$	$E_1 - E''$ $k = 6$	$E_1 - E''$ $k = 8$	$E_1 - E''$ $k = 10$
1	3.58×10^{-1}	5.11×10^{-1}	8.66×10^{-2}	1.95×10^{-3}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
2	2.03×10^{-2}	4.68×10^{-2}	1.88×10^{-2}	1.95×10^{-3}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
3	2.72×10^{-3}	7.21×10^{-3}	3.08×10^{-3}	1.95×10^{-3}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
4	2.56×10^{-3}	1.92×10^{-3}	7.87×10^{-4}	1.54×10^{-4}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
5	1.02×10^{-3}	1.07×10^{-3}	7.25×10^{-4}	1.41×10^{-4}	1.41×10^{-4}	2.75×10^{-5}	5.82×10^{-7}
6	2.14×10^{-4}	1.07×10^{-3}	5.56×10^{-4}	9.63×10^{-5}	8.43×10^{-5}	2.75×10^{-5}	5.82×10^{-7}
7	2.33×10^{-5}	3.58×10^{-4}	1.87×10^{-4}	4.14×10^{-5}	2.97×10^{-5}	2.75×10^{-5}	5.82×10^{-7}
8	3.85×10^{-6}	5.07×10^{-5}	2.94×10^{-5}	1.16×10^{-5}	7.28×10^{-6}	5.31×10^{-6}	5.82×10^{-7}
9	3.85×10^{-6}	1.09×10^{-5}	1.18×10^{-5}	4.89×10^{-6}	1.87×10^{-6}	7.37×10^{-7}	5.82×10^{-7}
10	2.28×10^{-6}	3.87×10^{-6}	1.15×10^{-5}	4.89×10^{-6}	1.40×10^{-6}	2.69×10^{-7}	1.68×10^{-7}
11	8.02×10^{-7}	2.35×10^{-6}	6.86×10^{-6}	3.14×10^{-6}	1.28×10^{-6}	2.69×10^{-7}	1.65×10^{-7}
12	1.77×10^{-7}	2.31×10^{-6}	2.54×10^{-6}	1.25×10^{-6}	6.65×10^{-7}	1.94×10^{-7}	1.01×10^{-7}
13	2.34×10^{-8}	9.94×10^{-7}	5.95×10^{-7}	3.13×10^{-7}	1.91×10^{-7}	8.52×10^{-8}	4.23×10^{-8}
14	5.54×10^{-9}	1.50×10^{-7}	9.17×10^{-8}	5.08×10^{-8}	3.33×10^{-8}	2.08×10^{-8}	1.28×10^{-8}
15	5.52×10^{-9}	3.21×10^{-8}	3.57×10^{-8}	1.83×10^{-8}	1.08×10^{-8}	6.63×10^{-9}	4.14×10^{-9}
16	3.74×10^{-9}	1.04×10^{-8}	3.55×10^{-8}	1.83×10^{-8}	1.08×10^{-8}	6.49×10^{-9}	3.58×10^{-9}
17	1.64×10^{-9}	5.13×10^{-9}	2.39×10^{-8}	1.27×10^{-8}	7.81×10^{-9}	4.95×10^{-9}	3.10×10^{-9}
18	4.89×10^{-10}	3.94×10^{-9}	1.08×10^{-8}	5.85×10^{-9}	3.74×10^{-9}	2.49×10^{-9}	1.69×10^{-9}
19	9.64×10^{-11}	3.86×10^{-9}	3.38×10^{-9}	1.87×10^{-9}	1.23×10^{-9}	8.61×10^{-10}	6.15×10^{-10}
20	1.45×10^{-11}	1.16×10^{-9}	7.13×10^{-10}	4.05×10^{-10}	2.74×10^{-10}	1.99×10^{-10}	1.50×10^{-10}
21	8.15×10^{-12}	2.03×10^{-10}	1.29×10^{-10}	7.30×10^{-11}	4.93×10^{-11}	3.63×10^{-11}	2.79×10^{-11}
22	7.74×10^{-12}	5.06×10^{-11}	8.36×10^{-11}	4.52×10^{-11}	2.90×10^{-11}	2.00×10^{-11}	1.42×10^{-11}
23	4.90×10^{-12}	1.81×10^{-11}	7.91×10^{-11}	4.32×10^{-11}	2.80×10^{-11}	1.96×10^{-11}	1.41×10^{-11}
24	2.17×10^{-12}	9.18×10^{-12}	5.17×10^{-11}	2.86×10^{-11}	1.89×10^{-11}	1.34×10^{-11}	9.93×10^{-12}
25	6.96×10^{-13}	6.78×10^{-12}	2.40×10^{-11}	1.35×10^{-11}	9.02×10^{-12}	6.53×10^{-12}	4.93×10^{-12}
26	1.55×10^{-13}	6.78×10^{-12}	8.17×10^{-12}	4.64×10^{-12}	3.14×10^{-12}	2.31×10^{-12}	1.78×10^{-12}
27	2.65×10^{-14}	3.17×10^{-12}	1.97×10^{-12}	1.13×10^{-12}	7.80×10^{-13}	5.84×10^{-13}	4.58×10^{-13}
28	1.24×10^{-14}	6.01×10^{-13}	3.75×10^{-13}	2.16×10^{-13}	1.49×10^{-13}	1.12×10^{-13}	8.88×10^{-14}
29	1.24×10^{-14}	1.43×10^{-13}	1.78×10^{-13}	9.90×10^{-14}	6.59×10^{-14}	4.76×10^{-14}	3.60×10^{-14}
30	8.94×10^{-15}	4.73×10^{-14}	1.77×10^{-13}	9.89×10^{-14}	6.59×10^{-14}	4.76×10^{-14}	3.59×10^{-14}
31	4.64×10^{-15}	2.13×10^{-14}	1.34×10^{-13}	7.57×10^{-14}	5.09×10^{-14}	3.72×10^{-14}	2.84×10^{-14}
32	1.80×10^{-15}	1.31×10^{-14}	7.41×10^{-14}	4.20×10^{-14}	2.85×10^{-14}	2.10×10^{-14}	1.63×10^{-14}
33	5.19×10^{-16}	1.13×10^{-14}	3.08×10^{-14}	1.76×10^{-14}	1.20×10^{-14}	8.97×10^{-15}	7.01×10^{-15}
34	1.07×10^{-16}	1.07×10^{-14}	9.61×10^{-15}	5.54×10^{-15}	3.82×10^{-15}	2.87×10^{-15}	2.27×10^{-15}
35	2.47×10^{-17}	3.52×10^{-15}	2.19×10^{-15}	1.27×10^{-15}	8.86×10^{-16}	6.73×10^{-16}	5.37×10^{-16}
36	1.97×10^{-17}	7.31×10^{-16}	5.07×10^{-16}	2.92×10^{-16}	2.02×10^{-16}	1.52×10^{-16}	1.20×10^{-16}
37	1.82×10^{-17}	1.99×10^{-16}	3.62×10^{-16}	2.04×10^{-16}	1.38×10^{-16}	1.01×10^{-16}	7.79×10^{-17}
38	1.23×10^{-17}	7.27×10^{-17}	3.49×10^{-16}	1.98×10^{-16}	1.34×10^{-16}	9.89×10^{-17}	7.65×10^{-17}
39	6.20×10^{-18}	3.49×10^{-17}	2.52×10^{-16}	1.44×10^{-16}	9.80×10^{-17}	7.28×10^{-17}	5.68×10^{-17}
40	2.42×10^{-18}	2.22×10^{-17}	1.37×10^{-16}	7.86×10^{-17}	5.40×10^{-17}	4.04×10^{-17}	3.18×10^{-17}

eigenvalues of H'' for every $|\Phi\rangle$ of the form (25). Thus we have indeed optimized the new tight lower bound.

Tables V and VI show optimized E'' bounds obtained for the ground state energy E_1 and the second excited state energy E_3. Comparison with the unoptimized E'' bounds of Tables I and II shows that the improvement obtained fluctuates with N almost in periodic fashion; the ratio of the error $E_1 - E_1''$ for the unoptimized bound to the error $E_1 - E_1''$ for the optimized bound has maxima for $N = 10$, 16 or 17, 23 or 24, 30 or 31, and 38 or 39 and minima for $N = 8$, 14, 21, 28, and 36, which is almost the same as the location of the maxima and minima for the Temple bound. The ratio of the errors $E_3 - E_3''$ for the unoptimized and optimized bounds behaves similarly, with maxima for $N = 12$, 19, 26 or 27, and 34, and minima for $N = 17$, 24, 31, and 40. These ratios also increase with k for fixed N, except for a few cases in which N is comparable to k. The greatest improvement for the ground state (by a factor of 29.5) occurs for $N = 10$ and $k = 8$; the greatest improvement for the second excited state (by a factor of 25.9) occurs for $N = 12$ and $k = 10$. Improvement by factors of 2–5 are more typical.

The optimized lower bounds can also be compared with Rayleigh–Ritz upper bounds for the same N. For $N < k$, the error $E_{ex} - E''$ in the optimized lower bound is smaller than the error $E_{RR} - E_{ex}$ in the Rayleigh–Ritz upper bound in all but two cases ($N = 7$, $k = 8$ for the ground state and $N = 9$, $k = 10$ for the second excited state). For $N \geqslant k$, the error $E_{ex} - E''$ tends to be larger, with a few exceptions, all of which occur for $N \leqslant k + 6$. The ratio $(E_{ex} - E'')/(E_{RR} - E'')$ fluctuates with N, taking on minima at $N = 9$ or 10, 15, 22, 29, and 37 for the ground state, and at 11 or 12, 17 or 18, 25, and 32 or 33 for the second excited state. The values of this ratio tend to increase at successive minima. For example, for the ground state with $k = 10$, values of this ratio at $N = 10$, 15, 22, 29, and 37 are 0.074, 0.75, 1.84, 2.91, and 4.28 respectively.

TABLE VI. The difference between E_3 and the Rayleigh–Ritz upper bound E_{RR}, the difference between E_3 and the optimized Temple bound E_T for $\widetilde{E}_4 = E_{4\,lb}$, and the differerences between E_3 and the optimized new tight lower bounds E'' are listed for different numbers N of terms in Φ when $\lambda = 1.0$.

N	$E_{RR} - E_3$	$E_3 - E_T$ $n = 3, p = 1$	$E_3 - E''$ $k = 4$	$E_3 - E''$ $k = 6$	$E_3 - E''$ $k = 8$	$E_3 - E''$ $k = 10$	$E_3 - E''$ $k = 12$
3	2.67×10		2.33	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
4	9.72	9.12	6.65×10^{-1}	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
5	3.68	8.94	3.81×10^{-1}	2.68×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
6	1.25	2.52	3.69×10^{-1}	2.67×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
7	3.38×10^{-1}	6.39×10^{-1}	3.54×10^{-1}	1.87×10^{-1}	1.78×10^{-1}	2.38×10^{-2}	6.67×10^{-4}
8	6.28×10^{-2}	1.79×10^{-1}	2.89×10^{-1}	9.83×10^{-2}	7.90×10^{-2}	2.38×10^{-2}	6.67×10^{-4}
9	9.68×10^{-3}	5.53×10^{-2}	1.71×10^{-1}	4.22×10^{-2}	2.75×10^{-2}	2.38×10^{-2}	6.67×10^{-4}
10	6.64×10^{-3}	2.00×10^{-2}	7.23×10^{-2}	1.72×10^{-2}	7.96×10^{-3}	4.71×10^{-3}	6.67×10^{-4}
11	5.89×10^{-3}	9.64×10^{-3}	5.75×10^{-2}	1.15×10^{-2}	2.86×10^{-3}	8.68×10^{-4}	6.67×10^{-4}
12	3.31×10^{-3}	6.89×10^{-3}	5.44×10^{-2}	1.15×10^{-2}	2.28×10^{-3}	4.75×10^{-4}	3.35×10^{-4}
13	1.26×10^{-3}	6.86×10^{-3}	3.66×10^{-2}	8.16×10^{-3}	2.21×10^{-3}	4.73×10^{-4}	3.22×10^{-4}
14	3.31×10^{-4}	3.94×10^{-3}	1.69×10^{-2}	3.82×10^{-3}	1.48×10^{-3}	3.69×10^{-4}	2.14×10^{-4}
15	5.70×10^{-5}	8.97×10^{-4}	5.24×10^{-3}	1.23×10^{-3}	5.84×10^{-4}	2.03×10^{-4}	1.04×10^{-4}
16	1.21×10^{-5}	2.19×10^{-4}	1.09×10^{-3}	2.70×10^{-4}	1.41×10^{-4}	7.10×10^{-5}	3.96×10^{-5}
17	1.08×10^{-5}	7.05×10^{-5}	3.00×10^{-4}	7.23×10^{-5}	3.63×10^{-5}	2.10×10^{-5}	1.35×10^{-5}
18	8.96×10^{-6}	3.07×10^{-5}	2.76×10^{-4}	6.39×10^{-5}	3.00×10^{-5}	1.63×10^{-5}	8.68×10^{-6}
19	4.99×10^{-6}	1.89×10^{-5}	2.35×10^{-4}	5.59×10^{-5}	2.72×10^{-5}	1.53×10^{-5}	8.64×10^{-6}
20	2.00×10^{-6}	1.69×10^{-5}	1.40×10^{-4}	3.41×10^{-5}	1.73×10^{-5}	1.02×10^{-5}	6.32×10^{-6}
21	5.76×10^{-7}	1.47×10^{-5}	6.05×10^{-5}	1.50×10^{-5}	7.90×10^{-6}	4.85×10^{-6}	3.18×10^{-6}
22	1.15×10^{-7}	4.08×10^{-6}	1.91×10^{-5}	4.84×10^{-6}	2.61×10^{-6}	1.67×10^{-6}	1.14×10^{-6}
23	2.27×10^{-8}	8.75×10^{-7}	4.28×10^{-6}	1.11×10^{-6}	6.11×10^{-7}	4.06×10^{-7}	2.91×10^{-7}
24	1.68×10^{-8}	2.48×10^{-7}	9.55×10^{-7}	2.44×10^{-7}	1.34×10^{-7}	8.87×10^{-8}	6.42×10^{-8}
25	1.55×10^{-8}	9.36×10^{-8}	6.92×10^{-7}	1.71×10^{-7}	8.94×10^{-8}	5.63×10^{-8}	3.81×10^{-8}
26	1.01×10^{-8}	4.74×10^{-8}	6.58×10^{-7}	1.64×10^{-7}	8.66×10^{-8}	5.51×10^{-8}	3.77×10^{-8}
27	4.80×10^{-9}	3.31×10^{-8}	4.55×10^{-7}	1.15×10^{-7}	6.16×10^{-8}	3.99×10^{-8}	2.79×10^{-8}
28	1.72×10^{-9}	3.18×10^{-8}	2.33×10^{-7}	5.94×10^{-8}	3.23×10^{-8}	2.12×10^{-8}	1.52×10^{-8}
29	4.52×10^{-10}	2.35×10^{-8}	9.05×10^{-8}	2.33×10^{-8}	1.28×10^{-8}	8.56×10^{-9}	6.21×10^{-9}
30	8.79×10^{-11}	5.82×10^{-9}	2.62×10^{-8}	6.83×10^{-9}	3.81×10^{-9}	2.58×10^{-9}	1.90×10^{-9}
31	2.80×10^{-11}	1.34×10^{-9}	5.68×10^{-9}	1.49×10^{-9}	8.41×10^{-10}	5.76×10^{-10}	4.31×10^{-10}
32	2.67×10^{-11}	4.11×10^{-10}	1.72×10^{-9}	4.42×10^{-10}	2.43×10^{-10}	1.63×10^{-10}	1.19×10^{-10}
33	2.27×10^{-11}	1.64×10^{-10}	1.55×10^{-9}	3.92×10^{-10}	2.12×10^{-10}	1.39×10^{-10}	9.93×10^{-11}
34	1.41×10^{-11}	8.54×10^{-11}	1.40×10^{-9}	3.56×10^{-10}	1.94×10^{-10}	1.28×10^{-10}	9.25×10^{-11}
35	6.64×10^{-12}	5.94×10^{-11}	9.33×10^{-10}	2.40×10^{-10}	1.32×10^{-10}	8.81×10^{-11}	6.42×10^{-11}
36	2.42×10^{-12}	5.56×10^{-11}	4.77×10^{-10}	1.24×10^{-10}	6.85×10^{-11}	4.61×10^{-11}	3.39×10^{-11}
37	6.66×10^{-13}	4.70×10^{-11}	1.90×10^{-10}	4.96×10^{-11}	2.77×10^{-11}	1.88×10^{-11}	1.40×10^{-11}
38	1.40×10^{-13}	1.40×10^{-11}	5.80×10^{-11}	1.52×10^{-11}	8.58×10^{-12}	5.87×10^{-12}	4.40×10^{-12}
39	4.57×10^{-14}	3.25×10^{-12}	1.35×10^{-11}	3.56×10^{-12}	2.02×10^{-12}	1.39×10^{-12}	1.05×10^{-12}
40	4.26×10^{-14}	9.70×10^{-13}	3.88×10^{-12}	1.01×10^{-12}	5.65×10^{-13}	3.84×10^{-13}	2.86×10^{-13}

The optimized E'' lower bounds can also be compared with the optimized Temple bounds E_T. $E_1 - E''$ tends to be smaller than $E_1 - E_T$ for $n = p = 1$, but there are a number of exceptions, which occur at and near $N = 10$, 17, 24, 31, and 39, for cases in which $E_2 - \widetilde{E}_2 \lesssim E_1 - E_T$. $E_3 - E''$ tends to be larger than the $n = 3, p = 1$ $E_3 - E_T$ for $k = 4$, except for $N < 8$. $E_3 - E''$ also tends to be larger for $k = 6$, except for $N < 11$ and $N = 14$, 24, or 29. $E_3 - E''$ tends to be smaller than $E_3 - E_T$ for $k = 8$,10, and 12, but there are a number of exceptions, which occur at and near $N = 11$, 19, 27, and 34, for cases in which $E_4 - \widetilde{E}_4 \lesssim E_3 - E_T$. Comparison of the $k = 10$ $E_1 - E''$ with the optimized Temple $E_1 - E_T$ for $n = p = 2$ and $n = p = 3$ shows that the optimized Temple bound is more accurate for $n = p = 2, N = 19$ when $\alpha = 10^{-8}$ or 10^{-12}, for $n = p = 3$, $N = 15$–31, when $\alpha = 10^{-8}$, and for $n = p = 3$, $N = 11$, 12, and 15–40 when $\alpha = 10^{-12}$. Thus it appears that the optimized Temple bounds tend to be better than the optimized E'' bounds when a lower bound $\widetilde{E}_{n+1}$ is known whose accuracy is comparable to the accuracy desired in the optimized Temple bound, but poorer than the E'' bounds when such a highly accurate $\widetilde{E}_{n+1}$ is not known. It would be interesting to see if these tendencies persist for other model problems.

Note added in proof: After the present paper had been submitted, the author became aware of a paper of M. Cohen and T. Feldman, J. Phys. B **12**, 2771–9 (1979), entitled "A generalization of Temple's lower bound to eigenvalues." The Cohen–Feldman lower bound appears to coincide with the Lehman–Maehly optimized Temple bound; in particular, their Eq. (29) is just the secular equation for the eigenvalue problem of Eq. (26) of the present paper.

ACKNOWLEDGMENTS

The author would like to thank Dr. David W. Fox for a helpful conversation about the Lehman–Maehly approach to the optimized Temple bound. This work was supported in part by the National Science Foundation under Grant No. PHY-7906936.

[1]The standard references on bounds to eigenvalues are: S.H. Gould, *Variational Methods for Eigenvalue Problems: An Introduction to the Weinstein*

Method of Intermediate Problems, 2nd ed. (University of Toronto Press, Toronto, 1966); A. Weinstein and W. Stenger, *Methods of Intermediate Problems for Eigenvalues* (Academic, New York, 1972), and H.F. Weinberger, *Variational Methods for Eigenvalue Approximation* (SIAM Regional Conference Series in Applied Mathematics, 1974). Methods of computing lower bounds to eigenvalues of the Schrödinger equation have been reviewed more recently by Charles E. Reid on pp. 315–347 of Jean Louis Calais, Osvald Goscinski, Jan Linderberg, and Yngve Ohrn, editors, *Quantum Science Methods and Structure a Tribute to Per-Olov Löwdin* (Plenum, New York, 1976) and by William Stenger in Chapter 14, pp. 277–291, of William C. Price and Seymour S. Chissick, editors, *The Uncertainty Principle and Foundations of Quantum Mechanics A Fifty Years Survey* (Wiley, New York, 1977).

[2]G. Temple, Proc. R. Soc. London Ser. A **119**, 276 (1928). A derivation may be found in Weinstein and Stenger, Ref. 1, pp. 103–104. Most derivations in the literature, including these, derive only the Temple bound to the ground state. However, it can be established easily for an eigenvalue E_n from the inequality $\langle \Phi | (H - E_n)(H - \widetilde{E}_{n+1}) | \Phi \rangle > 0$, which itself follows directly from a spectral representation of H.

[3]N.W. Bazley, Proc. Natl. Acad. Sci. U.S.A. **45**, 850–853 (1959); Phys. Rev. **120**, 144–149 (1960).

[4]K. Frankowski and C.L. Pekeris, Phys. Rev. **146**, 46 (1966).

[5]N.W. Bazley and D.W. Fox, Phys. Rev. **124**, 483–492 (1961).

[6]N.W. Bazley and D.W. Fox, J. Res. Nat. Bur. Stand. Sect. **65B**, 105–111 (1961).

[7]N.W. Bazley and D.W. Fox, J. Reine Angew. Math. **223**, 142–149 (1966).

[8]N.W. Bazley and D.W. Fox, Arch. Ration. Mech. Anal. **10**, 352–360 (1962).

[9]Essentially this thoerem appears as Theorem 1 on p. 21 of Weinstein and Stenger, Ref. 1, as Theorem 9 on p. 62 of Weinberger, Ref. 1, and is implicit in the discussion of intermediate problems of the second type in Gould, Ref. 1.

[10]There are actually two distinct minimax principles, one going back to H. Weyl and R. Courant, the other due to Poincaré. See Weinstein and Stenger, Ref. 1, Chap. 3, for a careful discussion of the distinction and of the relation between the two principles. See also Weinberger, Ref. 1, p. 55–57.

[11]R.N. Hill, Phys. Rev. Lett. **38**, 643–646 (1977); J. Math. Phys. **18**, 2316–2330 (1977).

[12]D.W. Fox, SIAM J. Math. Anal. **3**, 617–24 (1972).

[13]D.W. Fox and V.G. Sigillito, Chem. Phys. Lett. **13**, 85–87 (1972); **14**, 583–585 (1972); D.W. Fox and V.G. Sigillito, J. Appl. Math. Phys. **23**, 392–411 (1972).

[14]C.E. Reid, Int. J. Quantum. Chem. (USA) **6**, 793–795 (1972).

[15]C.E. Reid, Chem. Phys. Lett. **26**, 243–245 (1974).

[16]The anharmonic oscillator has been studied by a variety of methods. Recent papers include S.K. Bose and D.N. Tripathy, J. Math. Phys. **19**, 2555–2557 (1978); K. Banerjee and S.P. Bhatnagar, Phys. Rev. D. **18**, 4767–4769 (1978); J.P. Boyd, J. Math. Phys. **19**, 1445–1456 (1978); V. Singh, S.N. Biswas, and K. Datta, Phys. Rev. D **18**, 1901 (1978).

[17]The lower bound analysis given here for the harmonic oscillator, which exploits the fact that $\langle \xi_k | V^{1/2} | \xi_m \rangle$ is a band matrix, differs from that given in Refs. 1 and 5. The method of Refs. 1 and 5, which exploits the fact that $\langle \xi_l | V | \xi_m \rangle$ is a band matrix, requires a matrix inversion which is not needed here, and leads to a slightly different lower–bounding Hamiltonian.

[18]N.J. Lehman, Z. Angew. Math. Mech. **29**, 342–356 (1949); and **30**, 1–16 (1950). H.J. Maehly, Helv. Phys. Acta **25**, 547–568 (1952). A nice review can be found in Sec. 3 of N.W. Bazley and D.W. Fox, J. Res. Natl. Bur. Stand. **68B**, 173–183 (1964).

[19]This part of the iterative improvement method is closely related to the power method for the iterative calculation of matrix eigenvalues and eigenvectors. See J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Oxford U.P., London, 1965), and, for a recent review, G. Peters and J.S. Wilkinson, SIAM Rev. **21**, 339–360 (1979).

Asymptotic distribution of eigenvalues for the multidimensional Schrödinger equation

G. F. Newell[a)]

Division of Transportation Engineering, University of California, Berkeley, California 94720

(Received 26 June 1979; accepted for publication 26 September 1979)

Let $N(\lambda)$ be the number of eigenvalues less than λ of the k-dimensional Schrödinger equation for which the potential energy becomes infinite as $x\to\infty$ in all directions of the k-dimensional space. Titchmarsh has derived a first order asymptotic approximation to $N(\lambda)$ for $\lambda\to\infty$. His formula is rederived here by a different method and then extended so as to give at least one more term of the asymptotic series. The method is based upon approximating $q(x)$ locally by a quadratic function and evaluating the short time properties of the propagator for a hypothetical time-dependent diffusion equation having the same eigenvalues.

1. INTRODUCTION

One of the standard problems in quantum mechanics is to determine the eigenfunctions $\psi_n(x)$ and eigenvalues λ_n of the Schrödinger equation, a dimensionless form of which is

$$H\psi_n = -\nabla^2\psi_n(x) + q(x)\psi_n(x) = \lambda_n\psi_n(x), \qquad (1.1)$$
$$x = (x_1, x_2 \ldots, x_k), \quad \lambda_n \leqslant \lambda_{n+1},$$

in the space $\mathbb{R}^k$. This in turn is obtained from the time-dependent equation

$$H\psi(x,t) = \frac{i\partial\psi(x,t)}{\partial t}, \qquad (1.2)$$

the complete solution of which is

$$\psi(x,t) = \sum_{n=0}^{\infty} a_n\psi_n(x)\exp(-i\lambda_n t),$$

for suitable coefficients a_n.

It is assumed here that $q(x)\to+\infty$ for $x\to\infty$ in all directions in $\mathbb{R}^k$ so that the spectrum of H is, in fact, discrete. We also assume that $q(x)$ is bounded from below. By suitable translation of the λ_n we can assume that $q(x)\geqslant 0$.

One special aspect of this problem is to determine just the spectrum of H, i.e., a function

$$N(\lambda) = \text{number of eigenvalues with } \lambda_n < \lambda, \qquad (1.3)$$

or at least the asymptotic properties of $N(\lambda)$ for $\lambda\to\infty$.

In one dimension, $k=1$, the asymptotic properties of $N(\lambda)$ can be determined very accurately. The standard method for evaluating $N(\lambda)$ or λ_n is to exploit the fact that $\psi_n(x)$ has n zeros. One can obtain asymptotic approximations for $\psi(x,\lambda)$ using WKB or analogous schemes and then simply count the number of zeros of the approximate $\psi(x,\lambda)$ as λ increases. It can be shown that[1,2]

$$n+\tfrac{1}{2} = \frac{1}{\pi}\int[\lambda_n - q(x)]_+^{1/2}dx + O(1/n), \qquad (1.4)$$

$$[y]_+ \equiv \begin{cases} y & \text{if } y\geqslant 0,\\ 0 & \text{if } y\leqslant 0.\end{cases}$$

This means that the function $N(\lambda)$, which is a step function of integer steps, jumps from

$$N(\lambda_{n-}) = -\tfrac{1}{2} + \frac{1}{\pi}\int[\lambda_n - q(x)]_+^{1/2}dx,$$

to

$$N(\lambda_{n+}) = +\tfrac{1}{2} + \frac{1}{\pi}\int[\lambda_n - q(x)]_+^{1/2}dx,$$

with a mean value at λ_n of

$$\tfrac{1}{2}N(\lambda_{n+}) + \tfrac{1}{2}N(\lambda_{n-}) = \frac{1}{\pi}\int[\lambda_n - q(x)]_+^{1/2}dx + O(1/n). \qquad (1.5)$$

Note that the convention for counting is to start with $n=0$, thus $N(\lambda)$ jumps from 0 to 1 at λ_0.

Titchmarsh[3] shows that the generalization of (1.5) in k-dimensions, $k>1$, is

$$N(\lambda) = \frac{[1+o(1)]}{2^k\pi^{k/2}\Gamma(k/2+1)}\int[\lambda - q(x)]_+^{k/2}dx^{(k)}. \qquad (1.6)$$

The derivation is quite different than for $k=1$, however, since one cannot easily count nodal planes or cells of the $\psi_n(x)$ in $\mathbb{R}^k$ for $k>1$. He gives two derivations of (1.6), but neither produces a quantitative estimate of the error term $o(1)$. One could infer some bounds on the error from his derivation but the methods are such that the bounds would not be very tight. In fact, (1.6) is much more accurate than one might expect. Even for $k=1$, the error term $O(1/n)$ in (1.5) is extremely small compared with the leading term, $O(n)$; it is even small compared with the jump of 1 at $\lambda=\lambda_n$.

The purpose of the following analysis is to obtain some reasonable estimate of the error term in (1.6), or at least to understand why, in comparison with accurately known special solutions, it gives such good results.

2. FORMULATION

For various reasons, it is more convenient to work with the equation

$$H\psi(x,t) = \frac{-\partial\psi(x,t)}{\partial t} \qquad (2.1)$$

than with (1.2). They differ only in that t is replaced by it, but

[a)]This research was done while the author was on sabbatical leave visiting the Institute for Fundamental Studies, Physics Department, University of Rochester, Rochester, New York.

 0022-2488/80/082193-09$1.00

the $\psi(x,t)$ in (2.1) can be interpreted as the (positive) probability density of a diffusing particle in an absorbing medium. For $q(x)=0$, (2.1) is the classic diffusion or heat conduction equation in $\mathbb{R}^k$. For $q(x)\geqslant 0$, $q(x)$ can be interpreted as the rate at which the probability density decreases at x due to absorption. We can also imagine a hypothetic motion of a Brownian particle which upon collision with particles of the medium has a certain probability of being captured.

Suppose that one could evaluate the propagator solution of (2.1)

$$G(x,t\,|x_0) = \text{probability density at } x, t \text{ for } t>0, \text{ given that the particle was at } x_0 \text{ at } t=0, \quad (2.2)$$

i.e., $G(x,t\,|x_0)$ is the solution of (2.1) with initial condition $G(x,0|x_0)=\delta(x-x_0)$. Since it is assumed that $q(x)\rightarrow\infty$ for $x\rightarrow\infty$ in all directions, the particle is certain to be absorbed eventually.

The general solution of (2.1) is

$$\psi(x,t) = G(t)\psi(x,0)$$

in which $G(t)$ is an integral operator, for each t, defined by

$$\psi(x,t) = \int G(x,t\,|x)\psi(x_0,0)\,dx_0{}^{(k)}. \quad (2.3)$$

Since the general solution of (2.1) can also be written in terms of the eigenfunctions of H

$$\psi(x,t) = \sum_{n=0}^{\infty} a_n\psi_n(x)\exp(-\lambda_n t),$$

it follows that $G(t)$ has the same eigenfunctions as H

$$G(t)\psi_n(x) = \exp(-\lambda_n t)\psi_n(x),$$

with eigenvalues $\exp(-\lambda_n t)$, for each t.

If $G_{nm}(t)$ is a matrix associated with the $G(t)$ relative to any orthonormal base in the Hilbert space of $\mathbb{R}^k$, then

$$\operatorname{trace}G(t) = \sum_n G_{nn}(t) = \sum_n \exp(-\lambda_n t).$$

In the orthonormal base corresponding to $\delta(x-x_0)$, $G_{nn}(t)$ becomes $G(x_0,t\,|x_0)$ and trace G can be written as

$$\int G(x_0,t\,|x_0)dx_0{}^{(k)} = \sum_n \exp(-\lambda_n t) = \int_0^{\infty} e^{-\lambda t}\,dN(\lambda). \quad (2.4)$$

The function $N(\lambda)$ describes a measure on $(0,\infty)$, but it is not bounded, $N(\infty)=\infty$. We assume, however, that (2.4) is finite for all real $t>0$ (but infinite for $t=0$). It is monotone decreasing in t on the real line $t>0$ and analytic in the complex t plane for $\mathrm{Re}\,t>0$. Since (2.4) represents the Laplace transform of the measure $N(\lambda)$, $N(\lambda)$ can be determined from G by inversion of the transform

$$N(\lambda) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\left[\int G(x_0,t\,|x_0)\,dx_0{}^{(k)}\right]\frac{e^{\lambda t}}{t}\,dt, \quad (2.5)$$

for any $a>0$.

If we reverse the order of integration, we can write (2.5) in the form

$$N(\lambda) = \int n(\lambda,x_0)\,dx_0{}^{(k)}, \quad (2.6)$$

with

$$n(\lambda,x_0) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} G(x_0,t\,|x_0)\frac{e^{\lambda t}}{t}\,dt. \quad (2.7)$$

We can think of $n(\lambda,x_0)$ as the contribution per unit volume to $N(\lambda)$ coming from the propagator G evaluated at x_0.

Thus, if we can determine an approximate $G(x,t\,|x)$, we can obtain a corresponding approximate $N(\lambda)$. Much of the subsequent analysis will be concerned with approximate forms for the G and the errors which they induce in the $N(\lambda)$. In anticipation that these approximations will be valid only for "short" time, we note here that the short time properties of G determine the large scale properties of $N(\lambda)$. If, for example, we replace λt by t' in (2.7) we have

$$n(\lambda,x_0) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} G\left(x_0,\frac{t'}{\lambda}\middle|x_0\right)\frac{e^{t'}}{t'}\,dt', \quad (2.7a)$$

involving the G evaluated in t'/λ.

Since the exact $N(\lambda)$ and $n(\lambda,x_0)$ will be step functions, we also note that one could smooth these functions on a scale of order σ by defining an approximate spectral density of $n(\lambda,x_0)$

$$\frac{dn}{d\lambda_0}(\lambda_0,x_0,\sigma) = \int_{-\infty}^{+\infty}\frac{1}{(2\pi)^{1/2}\sigma}\exp\left(-\frac{(\lambda-\lambda_0)^2}{2\sigma^2}\right)dn(\lambda,x_0). \quad (2.8)$$

from (2.7) we can evaluate the $n(\lambda_0,x_0,\sigma)$ as

$$n(\lambda_0,x_0,\sigma) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} G(x_0,t\,|x_0)\frac{\exp(\lambda_0 t)}{t}\exp\left(\frac{\sigma^2t^2}{2}\right)dt. \quad (2.9)$$

As one integrates (2.9) along or near the imaginary t-axis the factor $\exp(\sigma^2t^2/2)$ will further help to truncate the contributions to the integral from large $\mathrm{Im}\,t$. Roughly speaking, a truncation of the time coordinate on a scale of order $1/\sigma$ gives a smearing of the spectrum on a scale of order σ.

3. FIRST APPROXIMATION

If a particle started at x_0 and $q(x)=q$ were a constant, the particle would survive by time t with probability e^{-qt}. If it survived it would have a probability density

$$\frac{1}{2^k\pi^{k/2}t^{k/2}}\exp\left(-\frac{|x-x_0|^2}{4t}\right), \quad (3.1)$$

$$|x-x_0|^2 = \sum_{j=1}^{k}(x_i-x_{0i})^2.$$

Thus

$$G(x_0,t\,|x_0) = \frac{e^{-qt}}{2^k\pi^{k/2}t^{k/2}} \quad \text{for} \quad q(x)=q. \quad (3.2)$$

One might now argue that the particle does not travel very far (a distance of order $t^{1/2}$) in a short time. If $q(x)$ is a continuous function of x, a particle starting at x_0 temporarily sees a nearly constant q, namely the $q(x_0)$ at its origin. Thus for sufficiently small t and variable $q(x)$

$$G(x_0,t\,|x_0) \simeq \frac{\exp[-q(x_0)t]}{2^k\pi^{k/2}t^{k/2}}. \quad (3.3)$$

If we substitute (3.3) into (2.7) we have

$$n(\lambda,x_0)\simeq \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{\exp\{[\lambda-q(x_0)]t\}}{2^k\pi^{k/2}t^{k/2+1}}\,dt.$$

For $\lambda<q(x_0)$ the integrand is analytic in the right half plane (rhp) and vanishes for $\mathrm{Re}\,t\to+\infty$. The path of integration can be closed by a semicircle in the rhp enclosing no singularities. The integral vanishes. For $\lambda>q(x_0)$ the contour can be closed in the lhp

$$n(\lambda,x_0)=\frac{[\lambda-q(x_0)]^{k/2}}{2^k\pi^{k/2}}\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{e^u}{u^{k/2+1}}\,du$$
$$=\frac{[\lambda-q(x_0)]^{k/2}}{2^k\pi^{k/2}\Gamma(k/2+1)}. \quad (3.4)$$

Thus we obtain the Titchmarsh formula

$$N(\lambda)=\frac{1}{2^k\pi^{k/2}\Gamma(k/2+1)}\int[\lambda-q(x_0)]_+^{k/2}\,dx_0^{(k)}.$$

Suffice it to say that one can use various continuity theorems for transforms to prove that this formula is correct for $\lambda\to\infty$ in the sense of (1.6). Our goals are more ambitious than this, however. The above derivation gives a somewhat more intuitive reasoning than others [starting from (2.5)], from which we can make further corrections.

4. SECOND APPROXIMATION

The motivation for the first approximation to $G(x_0,t\,|x_0)$ was that the large λ behavior of $N(\lambda)$ would be determined mainly by the short time behavior of $G(x_0,t\,|x_0)$, but during a short time a particle does not travel very far. It sees the $q(x)$ as if it were nearly constant with a value $q(x_0)$.

As a second approximation (for short time) it seems logical to assume that, in the vicinity of x_0, the $q(x)$ can be expanded in a Taylor series in $x-x_0$;

$$q(x)=q(x_0)+\sum_{j=1}^{k}\frac{\partial q(x_0)}{\partial x_{0j}}(x_j-x_{0j})$$
$$+\tfrac{1}{2}\sum_{j,l=1}^{k}\frac{\partial^2 q(x_0)}{\partial x_{0j}\partial x_{0l}}(x_j-x_{0j})(x_l-x_{0l})+\cdots.$$

It is possible, however, for each value of x_0, to choose a coordinate system such that the matrix of quadratic terms $\partial^2 q(x_0)/\partial x_j\partial x_k$ is diagonal. Thus we can write $q(x)$ as

$$q(x)=q(x_0)+\sum_{j=1}^{k}q_j(x_0)(x_j-x_{0j})$$
$$+\tfrac{1}{2}\sum q_{jj}(x_0)(x_j-x_{0j})^2+\cdots, \quad (4.1)$$

in which

$$q_{jj}(x_0)=\frac{\partial^2 q(x_0)}{\partial x_{0j}^{\;2}},\quad q_j(x_0)=\frac{\partial q(x_0)}{\partial x_{0j}}, \quad (4.2)$$

are respectively the second derivatives in the coordinate system for which the matrix is diagonal, and the corresponding first derivatives in the same coordinates. To simplify notation we will hereafter write these coefficients as simply q, q_j, q_{jj}, etc., the dependence upon x_0 being understood.

If we neglect all terms in (4.1) of third or higher degree in (x_j-x_{0j}), we can still evaluate $G(x,t\,|x_0)$ exactly, and consequently also $G(x_0,t\,|x_0)$. Equation (2.1) is separable in the spatial coordinates and has a solution of the form

$$G(x,t\,|x_0)=e^{-qt}\prod_{j=1}^{k}\exp[h_{j0}(t)+h_{j1}(t)(x_j-x_{0j})$$
$$+h_{j2}(t)(x_j-x_{0j})^2], \quad (4.3)$$

i.e., $G(x,t\,|x_0)$ is an exponential of a quadratic form in the (x_j-x_{0j}) with suitable time-dependent coefficients, $h_{j0}(t)$, $h_{j1}(t)$, $h_{j2}(t)$.

If we substitute (4.3) into (2.1) and collect the coefficients of $(x_j-x_{0j})^m$, $m=0,1,2$, we see that (4.3) is indeed a solution of (2.1) provided the h_j's satisfy the ordinary differential equations

$$2h_{j2}(t)+h_{j1}^2(t)-dh_{j0}(t)/dt=0, \quad (4.4a)$$
$$4h_{j1}(t)h_{j2}(t)-dh_{j1}(t)/dt=q_j, \quad (4.4b)$$
$$4h_{j2}^2(t)-dh_{j2}(t)/dt=\tfrac{1}{2}q_{jj}. \quad (4.4c)$$

Equation (4.4c) involves only $h_{j2}(t)$ and can be solved by standard methods to give

$$h_{j2}(t)=-\frac{\beta_j}{4}\coth(\beta_j t),\quad \beta_j=(2q_{jj})^{1/2}, \quad (4.5a)$$

with an initial condition that for $t\to0$, $h_{j2}(t)\to-1/(4t)$ as in (3.1). Equation (4.4b) involves only h_{j1} and h_{j2}. With $h_{j2}(t)$ known from (4.5a), this becomes a linear equation in $h_{j1}(t)$. This is solved subject to the initial condition, $h_{j1}(t)\to0$ for $t\to0$, giving

$$h_{j1}(t)=-\frac{q_j}{\beta_j}\frac{[\cosh(\beta_j t)-1]}{\sinh(\beta_j t)}. \quad (4.5b)$$

Finally, with $h_{j2}(t)$ and $h_{j1}(t)$ known, (4.4a) can be solved by direct integration, subject to the initial condition that it agree with (3.1) for $t\to0$. This gives

$$h_{j0}(t)=-\tfrac{1}{2}\ln\left(\frac{4\pi}{\beta_j}\sinh(\beta_j t)\right)+\frac{2q_j^2}{\beta_j^3}\left[\frac{\beta_j t}{2}-\tanh\left(\frac{\beta_j t}{2}\right)\right]. \quad (4.5c)$$

The function $G(x_0,t\,|x_0)$ involves only the $h_{j0}(t)$;

$$G(x_0,t\,|x_0)=\frac{e^{-qt}}{2^n\pi^{n/2}}\prod_{j=1}^{k}\left(\beta_j^{1/2}\exp\left\{\frac{2q_j^2}{\beta_j^3}\right.\right.$$
$$\left.\left.\times\left[\frac{\beta_j t}{2}-\tanh\left(\frac{\beta_j t}{2}\right)\right]\right\}\Big/[\sinh(\beta_j t)]^{1/2}\right)\Big/ \quad (4.6)$$

In order to understand the meaning of this it is convenient to look at the complete $G(x,t\,|x_0)$ for small and large t, specifically for $|\beta_j t|\ll1$ or $|\beta_j t|\gg1$. For $|\beta_j t|\ll1$ we can expand the h's in powers of $\beta_j t$ to give

$$h_{j2}(t)=-\frac{1}{4t}-\frac{q_{jj}t}{6}+O(q_{jj}^2t^3), \quad (4.7a)$$
$$h_{j1}(t)=-q_jt/2+O(q_jq_{jj}t^3), \quad (4.7b)$$
$$h_{j0}(t)=-\tfrac{1}{2}\ln(4\pi t)-q_{jj}\frac{t^2}{6}$$
$$+q_j^2\frac{t^3}{3\cdot4}+O(q_{jj}^2t^4)+O(q_j^2q_{jj}t^5). \quad (4.7c)$$

The first terms of $h_{j2}(t)$ and $h_{j0}(t)$ give the approximation (3.1). The former describes the diffusion of the particle, the latter the "renormalization". The second term of h_{j2} and the first term of $h_{j1}(t)$ describe the fact that, for $x_j - x_{j0} \neq 0$, the particle feels the decay associated with $q(x)$ along some average (linear) path from x_0 to x. For $q_{jj} > 0$, the decay is greater the greater the displacement $|x_j - x_{j0}|^2$. Indeed, a particle traveling at constant speed from x_0 to x will see an average quadratic term of $\frac{1}{3}(q_{jj}/2)(x_j - x_{j0})^2$. Similarly, for $q_j > 0$ the decay is greater if the particle moves in the direction of increasing $q(x)$, i.e., $x_j - x_{j0} > 0$.

These help to explain the meaning of the other terms of $h_{j0}(t)$ which are important to the evaluation of $G(x_0,t|x_0)$. The second term of (4.7c) is associated with the fact that a particle starting at x_0 and returning to x_0 at time t will have wandered a distance of order $\sqrt{t}$ before returning. Thus it would have been in a range of $(x_j - x_{j0})$ where the quadratic term of $q(x)$ is higher than at x_0 by the order of $q_{jj}t$, and will have decayed by an extra amount of order $q_{jj}t^2$.

The third term of $h_{j0}(t)$ is proportional to q_j^2 and positive. If a particle wanders in the direction of increasing q. it decays faster, but it is equally likely (for short times) to wander in the direction of decreasing q and decay slower. The first order effect of this is that the two cancel. The second order effect, however, is that the particle can survive if it can stay on the side of lower $q(x)$. If it wanders a distance of order $\sqrt{t}$ it will survive or decay at an extra rate of order $q_j t^{1/2}$ and have a total extra decay proportional to $q_j t^{3/2}$. The net effect of the positive and negative contributions, however, is proportional to the square of this, namely $q_j^2 t^3$.

For $\beta_j t \gg 1$, $q_{jj} > 0$,

$$h_{j2}(t) \to -\frac{\beta_j}{4}[1 + O(e^{-2\beta_j t})], \tag{4.8a}$$

$$h_{j1}(t) \to -\frac{q_j}{\beta_j}[1 + O(e^{-\beta_j t})], \tag{4.8b}$$

$$h_{j0}(t) \to \left[-\frac{\beta_j}{2} + \frac{q_j^2}{2q_{jj}}\right]t - \frac{1}{2}\ln\left(\frac{2\pi}{\beta_j}\right) - \frac{2q_j^2}{\beta_j^3} + O\left(\frac{q_j^2}{\beta_j^3}e^{-\beta_j t}\right). \tag{4.8c}$$

To see the significance of this, one should first note that, if $q_{jj} > 0$ for all j, then $q(x) - q(x_0)$ has a minimum with respect to the x_j at

$$-(x_j - x_{0j}) = q_j/q_{jj}$$

[possibly outside the range of validity of the expansion (4.1)], where

$$[q(x) - q(x_0)]_{\min} = \sum_{j=1}^{k} q_j^2/2q_{jj}.$$

The complete $G(x,t|x_0)$ becomes

$$G(x,t|x_0) \to \exp\left[-\sum_{j=1}^{k}\frac{\beta_j}{4}\left(x_j - x_{0j} + \frac{q_j}{q_{jj}}\right)^2\right]\exp\left(-\sum_{j=1}^{k}\frac{q_j^2}{\beta_j^3}\right) \times\left[\prod_{j=1}^{k}\left(\frac{\beta_j}{2\pi}\right)^{1/2}\right]\exp\left[-qt + \sum_{j=1}^{k}\left(-\frac{\beta_j}{2} + \frac{q_j^2}{2q_{jj}}\right)t\right], \tag{4.9}$$

and

$$G(x_0,t|x_0) \to \left[\prod_{j=1}^{k}\left(\frac{\beta_j}{2\pi}\right)^{1/2}\right] \times\exp\left[-qt + \sum_{j=1}^{k}\left(-\frac{\beta_j}{2} + \frac{q_j^2}{2q_{jj}}\right)t\right] \times\exp\left(-\sum_{j=1}^{k}\frac{2q_j^2}{\beta_j^3}\right). \tag{4.10}$$

The first factor of (4.9) is a normal distribution around the minimum of $q(x)$. It is the eigenfunction of H corresponding to the lowest eigenvalue. The last factor is the exponential decay associated with the lowest eigenvalue of H, namely

$$-q + \sum_{j=1}^{k}\left(-\frac{\beta_j}{2} + \frac{q_j^2}{2q_{jj}}\right). \tag{4.11}$$

The factor $\exp(-\Sigma_{j=1}^{k} q_j^2/\beta_j^3)$ can be interpreted as the decay suffered by the particle before it reaches the minimum of $q(x)$. Correspondingly, in (4.10), the factor $\exp(-2 \times\Sigma_{j=1}^{k} q_j^2/\beta_j^3)$ can be interpreted as the decay suffered in reaching the minimum of $q(x)$ and then returning to the starting point x_0.

5. INVERSION OF THE TRANSFORM

To determine $n(\lambda,x_0)$ we must substitute (4.6) into (2.7) and evaluate (or approximate) the integral with respect to t along a path parallel with the imaginary axis or an equivalent path. The integrand (4.6), however, has very unpleasant singularities along the imaginary axis due to the $\tanh(\beta_j t/2)$ in the exponential; also some branch point singularities from $[\sinh(\beta_j t)]^{1/2}$. The former occur at $t = i(2n_j + 1)\pi/\beta_j$, the latter at $in_j\pi/\beta_j$ for every integer n_j and each $\beta_j, j = 1,...,k$.

If one wished to invert the transform exactly, one could expand $G(x_0,t|x_0)$ in powers of $\exp(-\beta_j t), j = 1,...,k$ and integrate term by term. The fact that $G(x_0,t|x_0)$ has such an expansion means that, for fixed x_0, the spectrum of G is discrete with jumps in $n(\lambda,x_0)$ at each λ of the form

$$\lambda = q - \sum_{j=1}^{k} q_j^2/2q_{jj} + \sum_{j=1}^{k}\beta_j(n_j + \tfrac{1}{2}) \tag{5.1}$$

for every integer value of n_j with $n_j \geqslant 0$. For $n_1 = n_2 \cdots n_k = 0$, this λ corresponds to the lowest eigenvalue of H as in (4.11).

To carry out such an evaluation is quite tedious and not very instructive because, in the next step, one must integrate the $n(\lambda,x_0)$ with respect to x_0. If the $q(x)$ were exactly a quadratic function of x, i.e., the coefficients β_j were independent of x_0, then the jumps in $n(\lambda,x_0)$ would occur at the same values of λ for all x_0. Indeed one can easily show, by reversing the order of integration with respect to t and x_0, that $N(\lambda)$ has, in this case, a unit step at each value of λ satisfying (5.1). If, however, the β_j and the $q_j^2/2q_{jj}$ vary with x_0 and are continuous functions of x_0, an integration of the $n(\lambda,x_0)$ with respect to x_0 will yield an $N(\lambda)$ which is a continuous function of λ.

Although in some limit in which the β_j approach constants, our approximate $N(\lambda)$ will approach a step function, it is obvious that it would require only a very small deviation from an exactly quadratic $q(x)$ to smear out completely any ripples in the $N(\lambda)$, for large λ.

The spectrum of H is certainly discrete, i.e., the exact $N(\lambda)$ has integer steps, but an exact evaluation of the eigenvalues λ_n requires that the $\psi_n(x)$ in (1.1) remain bounded for $x \to \infty$. Even for the one dimensional case, $k = 1$, any solution $\psi_n(x,\lambda)$ which is well behaved near one value of x where $q(x) \sim \lambda$ must also be well behaved on the other side of the "potential well" where $q(x) \sim \lambda$. We certainly did not intend that the locally quadratic approximation to $q(x)$ near some x_0 should be accurate all the way to the other side of the surface where $q(x)$ is equal to $q(x_0)$, except possibly for x_0 near the minimum of $q(x)$. The best we can hope for is that our approximate $N(\lambda)$ gives a good smoothed approximation to the correct $N(\lambda)$, hopefully one which passes nearly through the midpoints of the steps in the exact $N(\lambda)$, at least for $k = 1$; as it should according to (1.5).

Since our approximate $G(x_0,t\,|x_0)$ is accurate only for $|\beta_j t| \ll 1$, we will use the expansion (4.7c)

$$G(x_0,t\,|x_0) \simeq \frac{\exp[-qt - \gamma_1 t^2 + \gamma_2 t^3 + \cdots]}{2^k \pi^{k/2} t^{k/2}}, \quad (5.2)$$

with

$$\gamma_1 = \tfrac{1}{6} \sum_{j=1}^{k} q_{jj} = \tfrac{1}{6}\nabla^2 q > 0, \quad (5.2a)$$

$$\gamma_2 = \tfrac{1}{12} \sum_{j=1}^{k} q_j^2 = \tfrac{1}{12}|\mathrm{grad} q|^2 > 0. \quad (5.2b)$$

In the evaluation of $n(\lambda,x_0)$ for large λ, the most important terms in this expansion are the qt and $\gamma_2 t^3$ since the coefficients q and γ_2 will both be large (in some sense) for most values of x_0. The term $\gamma_1 t^2$ (of order $|\beta_j t|^2$) is accurate only for $|\gamma_1 t^2| \ll 1$. It is not clear whether one should leave this in the exponent as in (5.2) or write it as

$$\exp(-\gamma_1 t^2) \simeq 1 - \gamma_1 t^2 + \cdots. \quad (5.3)$$

As it stands, this factor is an annoyance because it increases along the imaginary t axis. We could remove it by smearing the spectrum as in (2.9) with $\sigma^2 = 2\gamma_1$. This would smear the spectrum on a λ-scale of the order σ, which is of the order of the one-dimensional spacing β_j of the steps in (5.1). We shall, however, keep this term in the form (5.3). Thus we will approximate $n(\lambda,x_0)$ by

$$n(\lambda,x_0) \simeq \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{(1-\gamma_1 t^2)\exp[(\lambda - q)t + \gamma_2 t^3]\,dt}{2^k \pi^{k/2} t^{k/2+1}}. \quad (5.4)$$

If we rescale the t variable

$$u = \gamma_2^{1/3} t,$$

we can also express (5.4) as

$$n(\lambda,x_0) = \frac{\gamma_2^{k/6}}{2^k \pi^{k/2}} \left(F_k(\xi) - \frac{\gamma_1}{\gamma_2^{2/3}} F_{k-4}(\xi) \right), \quad (5.4a)$$

with

$$\xi = (\lambda - q)\gamma_2^{-1/3}, \quad (5.4b)$$

and

$$F_l(\xi) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{\exp(\xi u + u^3)\,du}{u^{l/2+1}}. \quad (5.4c)$$

It can be shown, see Appendix A, that $F_l(\xi)$ is an entire function of ξ with a power series expansion

$$F_l(\xi) = \tfrac{1}{3} \sum_{n=0}^{\infty} \frac{\xi^n}{n!\,\Gamma((l-2n)/6+1)}. \quad (5.5a)$$

For $\xi < 0$ and $-\xi \gg 1$ it has an asymptotic form

$$F_l(\xi) \simeq \frac{\exp(-2|\xi/3|^{3/2})}{2(3\pi)^{1/2}|\xi/3|^{(l+3)/4}}, \quad (5.5b)$$

a very rapidly decreasing function of $|\xi|$. For $\xi \gg 1$ the asymptotic behavior is the sum of two series

$$F_l(\xi) \simeq \xi^{l/2} \sum_{n=0}^{\infty} \frac{\xi^{-3n}}{n!\Gamma(l/2+1-3n)} + \frac{\cos[2(\xi/3)^{3/2} - (\pi/4)(l+3)]}{(3\pi)^{1/2}(\xi/3)^{(l+3)/4}} \times [1 + O(\xi^{-3/2})]. \quad (5.5c)$$

The first series decreases in powers of ξ^{-3}; the second is highly oscillatory with an amplitude proportional to $\xi^{-(l+3)/4}$. The first term of (5.5c), for $n = 0$, leads to the Titchmarsh formula.

6. EVALUATION OF $N(\lambda)$

The functions $F_l(\xi)$ can be tabulated for all relevant values of l and can, therefore, be considered as "known functions." Except for very special analytic forms for $q(x)$, it would be necessary to do the final integration (2.6) with respect to x_0 numerically, regardless of the analytic form of $n(\lambda,x_0)$. We could, therefore, consider the evaluation of our approximate $N(\lambda)$ to be complete. It is possible, however, to obtain a somewhat simpler form for $N(\lambda)$.

Let

$$N'(\lambda) \equiv N(\lambda) - \int \frac{[\lambda - q]_+^{k/2}\,dx^{(k)}}{2^k \pi^{k/2} \Gamma(k/2+1)} = \int \frac{\gamma_2^{k/6} F_k^{(1)}(\xi)\,dx^{(k)}}{2^k \pi^{k/2}} - \int \frac{\gamma_1 \gamma_2^{(k-4)/6} F_{k-4}(\xi)\,dx^{(k)}}{2^k \pi^{k/2}}, \quad (6.1)$$

represent the correction to the Titchmarsh formula, with

$$F_l^{(1)}(\xi) \equiv \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{e^{\xi u}(\exp u^3 - 1)}{u^{l/2+1}}\,du = F_l(\xi) - \frac{[\xi]_+^{l/2}}{\Gamma(l/2+1)}. \quad (6.2)$$

The two correction terms in (6.1) are actually of comparable size. To see this, consider the following identity

$$\sum_{l=1}^{k} \frac{\partial}{\partial x_l} \left[\left(\frac{1}{12} \frac{\partial q}{\partial x} \right) \gamma_2^{(k-4)/6} F_{k+2}^{(1)}(\xi) \right] = \sum_{l=1}^{k} \frac{\partial}{\partial x_l} \left[\left(\frac{1}{12\gamma_2} \frac{\partial q}{\partial x} \right) \times \int_{a-i\infty}^{a+i\infty} \frac{e^{(\lambda - q)t}[\exp(\gamma_2 t^3) - 1]\,dt}{t^{k/2+2}} \right]$$

$$= \frac{(\gamma_1 - \gamma_3)}{2} \gamma_2^{(k-4)/6} F_{k+2}^{(1)}(\xi) - \gamma_2^{k/6} F_k^{(1)}(\xi)$$

$$+ \frac{\gamma_3}{2} \gamma_2^{(k-4)/6} F_{k-4}(\xi), \quad (6.3)$$

in which

$$3\gamma_3 = \sum_{l,m=1}^{k} \frac{\partial q}{\partial x_l} \frac{\partial^2 q}{\partial x_l \partial x_m} \frac{\partial q}{\partial x_m} \Big/ |\mathrm{grad} q|^2 \quad (6.4)$$

can be interpreted as the second derivative of q in the direction of grad q.

The function $F_{k+2}^{(1)}(\xi)$ has continuous derivatives for all ξ including $\xi = 0$. If we integrate both sides of (6.3) over any region of $\mathbb{R}^k$, the left hand side can be integrated to a surface integral (by the divergence theorem). The integrand of the surface integral is bounded. Furthermore, if we let the surface go to infinity, then $q \to \infty$, $\xi \to -\infty$ and the surface term vanishes because $F_{k+2}^{(1)}(\xi)$ goes to zero very rapidly for $\xi \to -\infty$.

The integral of the second term on the right-hand side of (6.3) is proportional to the first term of (6.1). Thus (6.3) can be used to transform (6.1) into the form

$$N'(\lambda) = \int \frac{(\gamma_1 - \gamma_3)\gamma_2^{(k-4)/6}}{2^{k+1}\pi^{k/2}} F_{k+2}^{(1)}(\xi)\, dx^{(k)}$$

$$- \int \frac{(2\gamma_1 - \gamma_3)\gamma_2^{(k-4)/6}}{2^{k+1}\pi^{k/2}} F_{k-4}(\xi)\, dx^{(k)}. \quad (6.5)$$

Suppose we now define $F_l^{(l)}(\xi)$ so that

$$F_l(\xi) - F_l^{(l)}(\xi) = \begin{cases} \xi^{l/2} \displaystyle\sum_{n=0}^{(l+1)/6} \frac{\xi^{-3n}}{n!\Gamma(l/2 + 1 - 3n)}, & \text{for } \xi > 0 \\ 0, & \text{for } \xi < 0 \end{cases}, \quad (6.6)$$

in which the sum over integer n extends only to the largest integer less than or equal to $(l+1)/6$. For even integer l, the series (6.6) actually terminates by itself because $\Gamma(x) = \infty$ for $x = 0, -1, -2, \cdots$, i.e., $F_l(\xi) - F^{(l)}(\xi)$ is a polynomial. Generally, however, the lowest power of ξ in (6.6) is periodic in l with period 6 and is always at least $-\frac{1}{2}$ for $l > -2$, lowest power of $\xi = -1/2 + 1/2\,[(l+1)\mathrm{mod}6]$. For $l \leqslant -2$, $F_l(\xi) = F_l^{(l)}(\xi)$.

The function $F_l^{(l)}(\xi)$, which is the difference between $F_l(\xi)$ and the terms (6.6) of the asymptotic expansion for $\xi > 0$, decreases very rapidly for $-\xi \gg 1$ as in (5.5b). For l an even integer, $F_l^{(l)}(\xi)$ is finite at $\xi = 0$ and behaves like just the second series of (5.5c) for $\xi \gg 1$. For l an odd integer, $F_l^{(l)}(\xi)$ will be infinite for $\xi \to 0_+$ like $\xi^{-1/2}$ for l mod $6 = 3$, $l > 0$, but otherwise is finite. For $\xi \gg 1$ it behaves like the second series of (5.5c) plus those terms from the first series which decreases at least as fast as $\xi^{-3/2}$.

Substitution of (6.6) into (6.5) now gives

$$N' = -\int \frac{[\lambda - q]^{k/2-2}}{2^{k+1}\pi^{k/2}} \times \sum_{n=0}^{(k-3)/6} \frac{[\gamma_1(2n+1) - n\gamma_3]\gamma_1^n [\lambda - q]^{-3n}}{(n+1)!\Gamma(k/2 - 1 - 3n)} dx^{(k)}$$

$$+ \int \frac{(\gamma_1 - \gamma_3)\gamma_2^{(k-4)/6}}{2^{k+1}\pi^{k/2}} F_{k+2}^{(k+2)}(\xi)\, dx^{(k)}$$

$$- \int \frac{(2\gamma_1 - \gamma_3)\gamma_2^{(k-4)/6}}{2^{k+1}\pi^{k/2}} F_{k-4}^{(k-4)}(\xi)\, dx^{(k)}. \quad (6.7)$$

The first series in (6.7) is in decreasing powers of $[\lambda - q]$ starting with power $k/2 - 2$ (two less than the Titchmarsh formula) but decreasing in steps of three. There is no term for $k < 3$ and only one term for $3 \leqslant k < 9$. Furthermore for $n = 0$ there is no contribution from the $n\gamma_3$. For $k \geqslant 9$, there will also be terms from $n = 1$. In view of the various approximations that have been made to obtain the $n(\lambda, x_0)$ in (5.4), it is not clear, however, whether or not these higher order terms ($n \geqslant 1$) are reliable.

The last two terms of (6.7) should be smaller than any nonzero terms of the first series, but for $k < 3$ they are the only terms. Consequently we should try to estimate these terms also.

We know that $F_l^{(l)}(\xi)$ decays very rapidly for $\xi \to -\infty$ and at least as fast as $\xi^{-3/2}$ for $\xi \to +\infty$. It can be integrated over $\xi = 0$ where, at worst, it behaves as $\xi^{-1/2}$ for $\xi \to 0_+$. We expect, therefore, that most of the contribution to the integrals of (6.7) will come from points near the surface S_λ where $q(x) = \lambda$, in particular for ξ comparable with 1. This also suggests that we introduce a new coordinate system in $\mathbb{R}^k$ such that one of the coordinates is $\xi = (\lambda - q)\gamma_2^{-1/3}$ itself, which vanishes on S_λ. The other coordinates can describe positions in the surface ξ = constant. We can then integrate first with respect to ξ.

From (5.4c) it follows that

$$F_{l+2}(\xi) = \int_{-\infty}^{\xi} F_l(\xi')\, d\xi',$$

and, consequently, from (6.6) we obtain

$$\int_{-\infty}^{\xi} F_l^{(l)}(\xi')\, d\xi' = F_{l+2}^{(l+2)}(\xi) + \begin{cases} \xi^{-1/2} \Big/ \left(\frac{l+3}{6}\right)! \Gamma\left(\frac{1}{2}\right) & \text{for } \xi > 0 \text{ and } l \bmod 6 = 3, \\ 1 \Big/ \left(\frac{l+2}{6}\right)! & \text{for } \xi > 0 \text{ and } l \bmod 6 = 4, \\ 0 & \text{for } \xi < 0 \text{ or } l \bmod 6 \neq 3 \text{ or } 4 \end{cases} \quad (6.8)$$

In particular for $\xi \to \infty$ we have, since $F_{l+2}^{(l+2)}(\xi) \to 0$ for $\xi \to \infty$,

$$\int_{-\infty}^{+\infty} F_l^{(l)}(\xi')\,d\xi' = \begin{cases} 1\Big/\left(\dfrac{l+2}{6}\right)! & \text{for } l \bmod 6 = 4 \\ 0 & \text{otherwise} \end{cases} \tag{6.9}$$

If, in (6.7), we consider γ_1, γ_2, and γ_3 as essentially constant over the range $\xi \sim 1$, these integrals nearly vanish due to cancellation of positive and negative contributions of the $F_l^{(l)}(\xi)$ over the range $\xi \sim 1$, except for $k \bmod 6 = 2$.

In the special case $k \bmod 6 = 2$ we can write

$$dx^{(k)} = dS_\lambda^{(k-1)} d\xi / |\text{grad}\xi|$$

with $dS_\lambda^{(k-1)}$ a surface element of the surface $q(x) = \lambda$. Since

$$|\text{grad}\xi| = \left|\text{grad}\,\frac{(\lambda - q)}{\gamma_2^{1/3}}\right| \simeq \left|\frac{(\text{grad}q)}{\gamma_2^{1/3}}\right| = (12)^{1/2}\gamma_2^{1/6}, \tag{6.10}$$

the last two integrals in (6.7) give approximately

$$-\int \frac{[\gamma_1 + [(k-2)/6](2\gamma_1 - \gamma_3)]\gamma_2^{(k-5)/6}\,dS_\lambda}{2^{k+2}\pi^{k/2}3^{1/2}[(k+4)/6]!}, \quad \text{for } k \bmod 6 = 2,$$

$$0, \quad \text{for } k \bmod 6 \neq 2. \tag{6.11}$$

We could go on to obtain a second approximation to (6.11) correcting for the deviation of the γ's from constants or that ξ does not go to $+\infty$, but the former corrections would involve either third derivatives of q or squares of second derivatives. It does not seem reasonable to keep such terms in view of the previous approximations to $n(\lambda, x_0)$. The contributions from large ξ would be comparable with a hypothetical next term in the sum over n of the series in (6.7). Our final result is (6.7) along with (6.11).

7. SPECIAL CASES

Since one is usually interested in the $N(\lambda)$ for small values of k, it is instructive to consider explicitly the values of $k = 1, 2, \cdots$.

For $k = 1$, (6.7) gives nothing meaningful, thus

$$N(\lambda) \simeq \frac{1}{\pi}\int [\lambda - q]_+^{1/2}\,dx, \quad \text{for } k = 1, \tag{7.1}$$

with an error term with is small compared with the integer steps of the exact $N(\lambda)$. Given that the present approximations yield a smooth $N(\lambda)$, this is the best one can do. It does properly show that the term of order 1 is missing in (1.5). If there had been such a term it would have appeared as a positive contribution from the last terms of (6.7).

For $k = 2$, the only correction comes from the surface integral (6.11)

$$N(\lambda) \simeq \frac{1}{4\pi}\int [\lambda - q]_+\,dx^{(2)} - \frac{1}{48\pi}\int \frac{\nabla^2 q\,dS_\lambda^{(1)}}{|\text{grad}q|}, \quad \text{for } k = 2. \tag{7.2}$$

For a symmetric quadratic $q(x)$, $q(x) = x_1^2 + x_2^2$, the first term of (7.2) is $\lambda^2/8$; the second term is $-1/12$. The exact $N(\lambda)$ for this $q(x)$ is highly degenerate with steps at $\lambda = 2n + 2$, $n = 0,1,2$, but the steps are of magnitude $n + 1$. For an asymmetric quadratic $q(x)$ with irrational ratios $(q_{11}/q_{22})^{1/2}$, the steps of $N(\lambda)$ will occur singly. The second term of (7.2), however, is still small compared with these integer steps and only describes some appropriate smoothing of the exact $N(\lambda)$.

For $k = 3$, we obtain, for the first time, a correction from the leading term of (6.7), giving

$$N(\lambda) \simeq \frac{1}{6\pi^2}\int [\lambda - q]_+^{3/2}\,dx^{(3)} - \frac{1}{3\cdot 2^5\pi^2}\int [\lambda - q]_+^{-1/2}\nabla^2 q\,dx^{(3)}, \quad \text{for } k = 3. \tag{7.3}$$

For a symmetric quadratic $q(x)$, the first term now gives $\frac{1}{6}(\lambda/2)^3$; the second term $-\lambda/32$.

For $k = 4$ to 7, we continue to obtain just a single correction.

$$N(\lambda) \simeq \frac{1}{2^5\pi^2}\int [\lambda - q]_+^2\,dx^{(4)} - \frac{1}{3\cdot 2^6\pi^2}\int_{q<\lambda} \nabla^2 q\,dx^{(4)}, \quad \text{for } k = 4, \tag{7.4}$$

(Note that one can also write $\int \nabla^2 q\,dx^{(4)}$ as $\int |\text{grad}q| \times dS_\lambda^{(3)}$)

$$N(\lambda) \simeq \frac{1}{60\pi^3}\int [\lambda - q]_+^{5/2}\,dx^{(5)} - \frac{1}{3\cdot 2^5\pi^3}\int [\lambda - q]_+^{1/2}\nabla^2 q\,dx^{(5)}, \quad \text{for } k = 5, \tag{7.5}$$

$$N(\lambda) \simeq \frac{1}{3\cdot 2^6\pi^3}\int [\lambda - q]_+^3\,dx^{(6)} - \frac{1}{3\cdot 2^8\pi^3}\int [\lambda - q]_+\nabla^2 q\,dx^{(6)}, \quad \text{for } k = 6, \tag{7.6}$$

$$N(\lambda) \simeq \frac{1}{2^3\cdot 7\cdot 5\cdot 3\pi^4}\int [\lambda - q]_+^{7/2}\,dx^{(7)} - \frac{1}{9\cdot 2^7\pi^4}\int [\lambda - q]_+^{3/2}\nabla^2 q\,dx^{(7)}, \quad \text{for } k = 7. \tag{7.7}$$

For $k = 8$ we start a sequence of formulas with three terms, the third term coming from (6.11) for $k = 8$ but from the series in (6.7), $n = 1$, for $k = 9$–13.

At least the first correction term in the above formulas should be valid for "almost all" $q(x)$. The case of the completely symmetric quadratic $q(x)$, $q(x) = q_0 + (\beta/4)\Sigma x_j^2$ is exceptional, however, in that the λ's are linear functions of $n + k/2$, n integer. But the multiplicity of the eigenvalues increase like n^{k-1} making the steps of $N(\lambda)$ of larger order than the correction $N'(\lambda)$ for all k. For any other rotationally symmetric $q(x)$ one can expect the steps in $N(\lambda)$ to be of an order comparable with the $N'(\lambda)$.

Except in such cases in which the multiplicity of the eigenvalues is of order greater than 1 because of some obvious symmetry of H, one does not expect the steps of $N(\lambda)$ to be larger than 1. The exact location of the steps, however, is sensitive to detailed irregularities in the $q(x)$. It seems likely that, for "most" $q(x)$, the difference between the exact $N(\lambda)$ and some smoothed approximation will look like a stochastic process (noise) probably with an amplitude of order somewhat less than $N^{1/2}(\lambda)$. In particular, for $k = 3$ one might expect this noise to be of a magnitude comparable with $N'(\lambda)$ but to be of order smaller than $N'(\lambda)$ for $k > 3$. A study of the magnitude of this noise is a possible subject for future research.

8. RELATED PROBLEMS

Correction terms for the Titchmarsh formula have been derived previously[5] for the special case $q(x) = 0$ (Laplace's equation) over some finite region, with the $\psi_n(x)$ satisfying certain boundary conditions [for example $\psi_n(x) = 0$] on the surface. They were obtained also by Brownian motion type arguments analogous to those described here. Indeed such methods have been used very extensively in the analysis of partial differential equations.

If one considers a sequence of smooth functions $q^{(j)}(x)$ for which

$$\lim_{j\to\infty} q^{(j)}(x) = \begin{cases} 0, & x\in\Omega, \\ \infty, & x\notin\Omega, \end{cases}$$

the eigenfunctions $\psi_n^{(j)}(x)$ would automatically satisfy the condition $\lim \psi_n^{(j)}(x) \to 0$ on the surface of Ω. It is tempting, therefore, to ask if the results for $q(x) = 0$ can be derived from appropriate limits of the formulas of Secs. 6 and 7.

The latter formulas were obtained from an assumption that $q(x)$ can be approximated locally by a quadratic function. This assumption is clearly violated in the limit $j\to\infty$ at or near the boundary so one could not justify taking any such limit. In fact the present correction terms to the Titchmarsh formula are meaningless in this limit even though the Titchmarsh formula itself gives the correct first approximation for $q(x) = 0$ in Ω.

The key arguments in the above analysis were that a particle cannot travel very far in a short time and consequently sees a nearly quadratic potential, and that the propagator for the harmonic potential is known exactly. It is interesting to note that if we were interested in the thermodynamic properties of the system and t were interpreted as $1/KT$, T = temperature, K = Boltzmann's constant, then (2.4) would be the partition function.

For sufficiently large T (small t) we can use (5.2), (5.3) directly [without inverting the transform to obtain $N(\lambda)$], and conclude that

$$Z = \sum_n \exp(-\lambda_n/KT) \simeq \left(\frac{KT}{4\pi}\right)^{k/2} \int [1 - \gamma_1/(KT)^2] \times \exp[-q/KT + \gamma_2/(KT)^3]\, dx^{(k)}. \tag{8.1}$$

The lowest approximation corresponding to $\gamma_1 = \gamma_2 = 0$ is, of course, the classical partition function. The "quantum mechanical corrections" for $\gamma_1, \gamma_2 \neq 0$ are disguised somewhat because we have taken a dimensionless form of (1.1).

Possible uses of this formula have not been explored yet. Neither is it obvious from thermodynamics what it means that a particle cannot travel very far in a "time" $1/KT$.

ACKNOWLEDGMENT

The author would like to thank Professor Elliott Montroll for suggesting this problem and for his very generous hospitality at the University of Rochester.

APPENDIX

For even integer l, most of the properties of the $F_l(\xi)$ can be derived from known properties of the Airy function[3,4] since

$$F_{-2}(\xi) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \exp[\xi u + u^3]\, du = 3^{-1/3}\mathrm{Ai}(-\xi/3^{1/3}),$$

and

$$F_{l+2}(\xi) = \int_{-\infty}^{\xi} F_l(\xi')d\xi'. \tag{A1}$$

Equations (5.5) can, however, be derived quite easily for arbitrary l.

The path of integration parallel with the imaginary u-axis can be deformed to the path C of Fig. 1a at angles of $\pm\pi/3$. If, now, we let $z = u^3$, then

$$F_l(\xi) = \frac{1}{2\pi i}\frac{1}{3}\int_{C'} \frac{\exp[\xi z^{1/3} + z]\, dz}{z^{l/6+1}},$$

where C' a path along the negative z line as in Fig. 1a. Equation (5.5a) is obtained by expanding $\exp(\xi z^{1/3})$,

$$\exp(\xi z^{1/3}) = \sum_{n=0}^{\infty} \frac{(\xi z^{1/3})^n}{n!},$$

and integrating term by term.

For $-\xi \gg 1$, the factor $\exp(\xi u + u^3)$ has a saddle point at u_0

$$u_0 = |\xi/3|^{1/2},$$

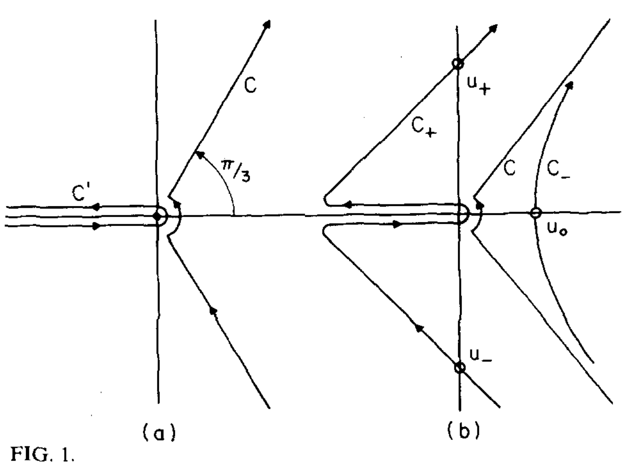


FIG. 1.

on the real line, $u_0 > 0$. Equation (5.5b) is obtained by deforming the contour C_- passing through this saddle point in the imaginary direction as in Fig. 1b, and doing the usual saddle point integration.

For $\xi > 0$, the saddle points of the exponential are on the imaginary axis at

$$u_\pm = \pm i|\xi/3|^{1/2}.$$

Since the integrand is rapidly decreasing along the negative line, $u < 0$, one would like to integrate along a path like C', but the path must start and end at $u = \pm\infty$ in the rhp. The simplest way to reach the rhp from $u = -\infty$ is to deform the contour to a path C_+ passing through the saddle point $u_\pm$ as in Fig. 1b. The path should go through $u_\pm$ at angles of $\pm\pi/4$.

The second term of (5.5c) comes from saddle point integrations through $u_\pm$. The first term of (5.5c) comes from the integration along the negative line, C'. If we let $v = \xi u$, this part of $F_l(\xi)$ becomes

$$\frac{1}{2\pi i}\int_{C'}\frac{\exp[v + (v/\xi)^3]\,dv}{\xi\,(v/\xi)^{k/2+1}}.$$

This time we expand $\exp[(v/\xi)^3]$ in a power series and integrate term by term.

The properties of $F_l(\xi)$ for l an odd integer can be obtained also by multiple integrations or differentiations of $F_1(\xi)$.

[1]E.C. Titchmarsh, *Eigenfunction Expansions Associated with Second Order Differential Equations* (Oxford, London, 1946).

[2]F.W.J. Olver, *Asymptotics and Special Functions* (Academic, New York, 1974).

[3]E.C. Titchmarsh, *Eigenfunction Expansions* (Oxford, London, 1958), Vol. II.

[4]*Handbook of Mathematical Functions*, edited by M. Abramowitz and I.A. Stegun (U.S. Goverment Printing Office, Washington, D.C., 1964).

[5]M. Kac, Amer. Math. Monthly 73, 1–23 (1966).

Coulomb scattering as the limit of scattering off smoothly screened Coulomb potentials

David M. Goodmanson
Department of Physics, Whitman College, Walla Walla, Washington 99362
John R. Taylor
Department of Physics, University of Colorado, Boulder, Colorado 80309

(Received 21 February 1980; accepted for publication 4 April 1980)

We examine the amplitude $f^{\rho}(\cos\theta)$ for scattering off a screened Coulomb potential $V^{\rho}(r) = V^{C}(r)\alpha^{\rho}(r)$, where $V^{C}(r)$ is the Coulomb potential γ/r. We prove that, for a wide class of smooth screening functions $\alpha^{\rho}(r)$, the screened amplitude $f^{\rho}(\cos\theta)$ approaches the pure Coulomb amplitude $f^{C}(\cos\theta)$, times a phase factor, as the screening radius $\rho\to\infty$. This limit is pointwise and uniform in angle and energy (excluding arbitrarily small neighborhoods of $\theta = 0$ and $E = 0$), and therefore implies that the corresponding cross sections become equal as $\rho\to\infty$.

1. INTRODUCTION

This is one of a series of papers on Coulomb scattering, the nonrelativistic scattering of two charged particles.[1-4] Our aim in these papers has been to establish the extent to which Coulomb scattering can be treated as the limit of scattering by a screened Coulomb potential, in the limit that the screening radius ρ tends to infinity.

We consider a screened potential $V^{\rho}(r)$ with the form

$$V^{\rho}(r) = V^{C}(r)\alpha^{\rho}(r), \tag{1.1}$$

where $V^{C}(r)$ denotes the pure Coulomb potential

$$V^{C}(r) = \gamma/r \tag{1.2}$$

(or, more generally, the Coulomb potential plus some short-range potential) and $\alpha^{\rho}(r)$ is a screening function that goes to zero as $r\to\infty$ (with ρ fixed) but approaches 1 as the screening radius $\rho\to\infty$ (with r fixed). Examples of the kinds of screening function one might consider are the exponential function

$$\alpha^{\rho}(r) = \exp(-r/\rho) \tag{1.3}$$

and sharp cutoff

$$\alpha^{\rho}(r) = \theta(\rho - r), \tag{1.4}$$

where $\theta(x)$ is the step function, which is zero for x negative, but is one for x positive.

We denote the amplitude and differential cross section for scattering off the screened potential V^{ρ} by f^{ρ} and $(d\sigma/d\Omega)^{\rho}$. The corresponding pure Coulomb quantities are f^{C} and $(d\sigma/d\Omega)^{C}$; in particular, the Coulomb amplitude f^{C} is well known to be[5]

$$f^{C}(\cos\theta) = -\gamma\exp(2i\sigma_0)\frac{\exp[-i\gamma\ln\sin^2(\theta/2)]}{2\sin^2(\theta/2)} \tag{1.5}$$

$$= \frac{\gamma'}{(1-\cos\theta)^{1+i\gamma}}, \tag{1.6}$$

where θ is the scattering angle, σ_0 is the s-wave Coulomb phase shift, $\arg\Gamma(1+i\gamma)$, and

$$\gamma' = -\gamma 2^{i\gamma}e^{2i\sigma_0}. \tag{1.7}$$

In Refs. 1 and 2 we considered the amplitudes f^{ρ} and f^{C} as distributions, and showed that, in this sense, f^{ρ} approaches f^{C} times a certain phase factor as $\rho\to\infty$:

$$[f^{\rho}(\cos\theta) - e^{2i\zeta(\rho)}f^{C}(\cos\theta)] \underset{\rho\to\infty}{\to} 0 \quad \text{(as a distribution)}, \tag{1.8}$$

where the phase $\zeta(\rho)$ is

$$\zeta(\rho) = -\int_{1/2}^{\infty} dr\, V^{\rho}(r). \tag{1.9}$$

The result (1.8) means that if the left-hand side is multiplied by a smooth function $\phi(\cos\theta)$ (which vanishes at $\theta = 0$)[6] and is integrated over all angles, then the resulting integral goes to zero as $\rho\to\infty$. The physical significance of the result is clear if one recalls that, for a given incident wave packet $\phi_{\text{in}}(\mathbf{p})$, the scattered wave is the integral over angles of $\phi_{\text{in}}(\mathbf{p})$ times the amplitude. Thus Eq. (1.8) means that the wave scattered off the screened potential V^{ρ} approaches that for the Coulomb potential (within a phase factor) as $\rho\to\infty$. Thus for a given incident packet, the probability for scattering by the screened potential V^{ρ} approaches that for the Coulomb potential as $\rho\to\infty$. In Ref. 1 this result was proved for the case that the potential that is screened is a pure Coulomb potential [as in Eq. (1.1)]. In Ref. 2 it was extended to the case of a Coulomb plus a short-range potential. In either case, the result was proved for a wide class of screening functions $\alpha^{\rho}(r)$ including both smooth functions like the exponential (1.3) and sharp ones like the cutoff (1.4).

The results of Refs. 1 and 2 establish equality of the relevant scattering probabilities *for a given incident packet* $\phi_{\text{in}}(\mathbf{p})$. This does not, by itself, guarantee equality of the corresponding cross sections, since the definition of a cross section involves a large number of different incident packets with randomly distributed impact parameters. In Refs. 3 and 4 we took up the question of cross sections. We showed that the screened cross section $(d\sigma/d\Omega)^{\rho}$ does approach the Coulomb cross section.

$$\left(\frac{d\sigma}{d\Omega}\right)^{\rho} \underset{\rho\to\infty}{\to} \left(\frac{d\sigma}{d\Omega}\right)^{C}, \tag{1.10}$$

provided the limit (1.8) holds pointwise and uniformly for all relevant momenta [see Ref. 3, paragraph below Eq. (2.3).]

 0022-2488/80/082202-06$1.00

We then examined the amplitudes f^{ρ} and f^{C} computed in Born and eikonal approximations and reached the following two conclusions: First, based on examination of the approximate amplitudes, it appears that for *smooth* screening functions like the exponential (1.3) the limit (1.8) does hold pointwise and uniformly. Thus it appears that the cross section for a smoothly screened Coulomb potential does approach the Coulomb cross section as $\rho \to \infty$; that is, the limit (1.10) is true. On the other hand, for the sharp cutoff (1.4), it appears that the limit (1.10) does *not* hold (again based on examination of the Born and eikonal approximations). Instead, we found that, in Born and eikonal approximations,

$$\left(\frac{d\sigma}{d\Omega}\right)^{\rho} \underset{\rho \to \infty}{\to} \frac{3}{2}\left(\frac{d\sigma}{d\Omega}\right)^{C}, \tag{1.11}$$

and were able to show that the extra $(1/2)(d\sigma/d\Omega)^{C}$ is contributed by particles that are reflected off the discontinuity in the potential at $r = \rho$ (no matter how large we make ρ).

Both of these conjectures were based on the examination of approximate amplitudes. Although one can argue that there are circumstances where the approximations are almost certainly reliable, it is obviously desirable to find proper proofs based on the exact amplitudes. In this paper we give an exact proof of the first result. That is, we prove that, for screening functions that are sufficiently smooth, the amplitude f^{ρ} does approach the Coulomb amplitude f^{C} within the expected phase factor,

$$[f^{\rho}(\cos\theta) - e^{2i\zeta(\rho)} f^{C}(\cos\theta)] \underset{\rho \to \infty}{\to} 0, \tag{1.12}$$

pointwise and uniformly for $\cos\theta$ in any closed interval excluding $\cos\theta = 1$ and for energy in any compact interval excluding 0. This ensures that the desired limit (1.10) for the cross sections holds.

Our smoothness conditions on the screening function are described in Sec. 3. The conditions are sufficient, but quite likely not necessary, and could probably be weakened. However, they are already sufficiently weak to admit a wide class of smooth functions, including the exponential $\alpha^{\rho}(r) = \exp(-r/\rho)$, the Gaussian $\exp(-r^2/\rho^2)$, and an inverse power like $\rho^n/(\rho + r)^n$. On the other hand, they exclude the sharp cutoff (1.4) as one might expect. [We have not yet managed to show what does happen with the sharp cutoff, but we conjecture that the limit (1.11) holds.]

Our proof of the limit (1.12) uses the partial-wave expansion for the amplitudes concerned. For this reason we begin our proof, in Sec. 2, with some useful results relating to the partial-wave series for the Coulomb amplitude. In Sec. 3 we give our conditions on the screening function. In Sec. 4 we outline the main proof, and in Secs. 5 and 6 we fill in some of the details; the remainder can be found in the thesis of the first named author.[7]

In the last few years there have appeared many papers on the scattering of charged particles. In particular, considerable progress has been made in understanding of scattering involving three or more bodies,[8-10] some of it based on the use of screening functions. (See especially Ref. 8 and 9.) Nevertheless, as far as we know there has appeared no proof of the simple two-body result (1.12), and this is therefore what we now prove.

2. PROPERTIES OF THE COULOMB PARTIAL-WAVE SERIES

The standard definition of the Coulomb phase shift σ_l is

$$\sigma_l = \arg\Gamma(l + 1 + i\gamma). \tag{2.1}$$

In terms of these one can define a Coulomb partial-wave S matrix $s_l^C = \exp(2i\sigma_l)$ and a partial-wave amplitude $f_l^C = (s_l^C - 1)/2i$. It is well known that the corresponding partial-wave series

$$\sum (2l+1) f_l^C P_l(\cos\theta)$$

is pointwise divergent.[1] On the other hand, it was shown in Ref. 1 that this same series, when considered as a distribution, is convergent and converges to the Coulomb amplitude

$$f^{C}(x) \equiv \gamma'/(1-x)^{1+i\gamma} = \sum (2l+1) f_l^C P_l(x) \quad \text{(as a distribution)}, \tag{2.2}$$

where we have introduced the abbreviation $x = \cos\theta$.

The reason that the series (2.2) is pointwise divergent is that $f^{C}(x)$ has an infinite singularity at $x = 1$ (the forward direction) due to the factor $(1 - x)$ in its denominator. This suggests that one would obtain a better series if one multiplied Eq. (2.2) by $(1 - x)$. If we do this and use the recurrence relation for the Legendre polynomials, we can rewrite $(1 - x)P_l(x)$ in terms of P_{l-1}, P_l, and P_{l+1}, and after some simple algebra obtain

$$(1-x) f^{C}(x) = \gamma'(1-x)^{-i\gamma} = \sum (2l+1) \frac{\gamma^2 s_{l-1}^C}{i(l+1-i\gamma)(l-i\gamma)} P_l(x). \tag{2.3}$$

From (2.2) it follows that the series (2.3) converges as a distribution. However, it follows from Laplace's formula,[11]

$$P_l(\cos\theta) = \left(\frac{2}{\pi l \sin\theta}\right)^{1/2} \cos[(l+\tfrac{1}{2})\theta - \tfrac{1}{4}\pi] + O(l^{-3/2}), \tag{2.4}$$

uniformly for $-1 + \epsilon \leqslant \cos\theta \leqslant 1 - \epsilon$ (any $\epsilon > 0$), that $P_l(x)$ is $O(l^{-1/2})$ and hence that the series (2.3) is actually convergent pointwise for $-1 < x < 1$ and uniformly for $-1 + \epsilon \leqslant x \leqslant 1 - \epsilon$. Furthermore, it is easily seen that the sum of the series is precisely the function on the left, $(1 - x)f^{C}(x)$. Thus even though the partial-wave series (2.2) is itself of no use in studying the pointwise properties of the Coulomb amplitude, the series (2.3) obtained by multiplying (2.2) by $(1 - x)$ *is* pointwise convergent and can be used to study the pointwise properties of $f^{C}(x)$.

The series (2.3) is unfortunately still divergent at $x = \pm 1$. This divergence reflects that the function $(1 - x)f^{C}(x)$ is discontinuous at $x = 1$. We shall not be interested in the point $x = 1$, since we know that the Coulomb cross section is undefined in the forward direction. On the other hand we *shall* be interested in the point $x = -1$ (the backward direction). To obtain a series that converges at

$x = -1$ (and, incidentally, at $x = 1$) we can multiply (2.3) by a second factor of $(1-x)$. The function $(1-x)^2 f^C(x)$ is continuous for $-1 \leqslant x \leqslant 1$ and has a Legendre expansion that can be found either by multiplying (2.3) by $(1-x)$ and rearranging, or directly, using standard techniques. It is[12]

$$(1-x)^2 f^C(x) = \sum_l g_l^C P_l(x), \tag{2.5}$$

where

$$g_l^C = \frac{2i\gamma^2(1-i\gamma)^2(2l+1)s_{l-2}^C}{(l+2-i\gamma)(l+1-i\gamma)(l-i\gamma)(l-1-i\gamma)}. \tag{2.6}$$

Obviously the coefficient g_l^C is $O(l^{-3})$ and, since $|P_l(x)| \leqslant 1$ for all x, the series (2.5) is uniformly and absolutely convergent for all x in $[-1,1]$.

In Sec. 4 we shall use the series (2.5) and the corresponding series for $(1-x)^2 f^\rho(x)$ to establish our main result that, as $\rho \to \infty$, $f^\rho(x)$ converges pointwise to $f^C(x)$ times the expected phase factor.

3. CONDITIONS ON THE SCREENING FUNCTION

The proof of our main result requires various assumptions concerning the screening function $\alpha^\rho(r)$. We assume first that $\alpha^\rho(r)$ satisfies the same conditions as in Ref. 1; that is we suppose that

(1) $\alpha^\rho(0) = 1$;

(2) $\alpha^\rho(r) \to 0$ monotonically like $O(r^{-2-\epsilon})$ as $r \to \infty$ (ρ fixed);

(3) $\alpha^\rho(r) \to 1$ as $\rho \to \infty$ (r fixed).

The main point of these assumptions is that $V^\rho(r) = \alpha^\rho(r)\gamma/r$ is a "well behaved" short-range potential [in the sense that it is $O(r^{-3-\epsilon})$ as $r \to \infty$] and that $V^\rho(r) \to V^C(r)$ as $\rho \to \infty$. Concerning the smoothness of $\alpha^\rho(r)$, we suppose that it has at least five continuous derivatives satisfying

$$\left|\left(\frac{d}{dr}\right)^n \alpha^\rho(r)\right| \leqslant \frac{K}{r^m}, \quad 0 < r < \infty, \tag{3.1}$$

for $m = 1,\ldots,5$, where K is independent of r and ρ.

Although the assumption (3.1) could perhaps be somewhat relaxed, it already admits a wide class of functions including the exponential $\alpha^\rho(r) = \exp(-r/\rho)$, the Gaussian $\exp(-r^2/\rho^2)$, and a power like $\rho^n/(\rho + r)^n$ $(n > 2)$.

For future reference we note that our assumptions imply that the screened potential $V^\rho(r) = \alpha^\rho(r)\gamma/r$ satisfies

$$\left|\left(\frac{d}{dr}\right)^m V^\rho(r)\right| \leqslant \frac{M}{r^{m+1}}, \quad 0 < r < \infty, \tag{3.2}$$

for $m = 0,\ldots,5$, where M is a constant independent of r and ρ.

4. PROOF OF CONVERGENCE OF THE AMPLITUDES

We have to show, subject to the conditions of Sec. 3, that

$$[f^\rho(x) - e^{2i\xi(\rho)} f^C(x)] \underset{\rho\to\infty}{\to} 0, \tag{4.1}$$

uniformly for $-1 \leqslant x \leqslant 1 - \epsilon$ (any $\epsilon > 0$). To this end it is obviously sufficient to show that

$$(1-x)^2[f^\rho(x) - e^{2i\xi(\rho)} f^C(x)] \underset{\rho\to\infty}{\to} 0, \tag{4.2}$$

uniformly for all x. We shall prove this using the partial-wave expansions of the two amplitudes concerned.

The screened amplitude $f^\rho(x)$ has a convergent partial-wave series, since the potential $V^\rho(r)$ is of short range. However, as $\rho \to \infty$ and $V^\rho(r)$ approaches the Coulomb potential, the series for $f^\rho(x)$ converges more and more slowly, reflecting the divergence of the pure Coulomb partial-wave series. Now, we have seen that multiplication of $f^C(x)$ by $(1-x)^2$ leads to a convergent Legendre expansion. Thus one might guess that multiplication of $f^\rho(x)$ by $(1-x)^2$ would produce a series that converges uniformly in ρ, and this proves to be the case.

Multiplying the partial-wave series for f^ρ by $(1-x)$ and using the recurrence relation for Legendre polynomials to rewrite $(1-x)P_l(x)$ in terms of P_{l-1}, P_l, and P_{l+1}, we obtain

$$(1-x)f^\rho(x) = \frac{i}{2}\sum [(l+1)s_{l+1}^\rho - (2l+1)s_l^\rho + l s_{l-1}^\rho] P_l(x), \tag{4.3}$$

where s_l^ρ is the partial-wave S matrix for the screened potential $V^\rho(r)$. This can be rewritten in terms of the first difference,

$$\Delta s_l \equiv s_{l+1} - s_l, \tag{4.4}$$

to read

$$(1-x)f^\rho(x) = \frac{i}{2}\sum \{(l+1)\Delta s_l^\rho - l\Delta s_{l-1}^\rho\} P_l(x). \tag{4.5}$$

If we multiply once more by $(1-x)$ and use the recurrence relation again, we obtain a series

$$(1-x)^2 f^\rho(x) = \sum g_l^\rho P_l(x), \tag{4.6}$$

where the coefficient g_l^ρ is easily written down in terms of the differences $\Delta s_{l+1}^\rho,\ldots,\Delta s_{l-2}^\rho$. The expression for g_l^ρ simplifies if we introduce the higher differences

$$\Delta^{m+1} s_l = \Delta^m s_{l+1} - \Delta^m s_l. \tag{4.7}$$

With this definition it is easily seen that[14]

$$g_l^\rho = \frac{1}{2i(2l+3)(2l-1)}[(2l^3 + l^2 - 3l)\Delta^4 s_{l-2}^\rho + (4l^2 + 6l)\Delta^3 s_{l-1}^\rho - (2l+4)\Delta^2 s_l^\rho + 2\Delta s_{l+1}^\rho]. \tag{4.8}$$

Now, if one examines the differences $\Delta^m s_l^C$ for the Coulomb S matrix, it is found that as $l \to \infty$[15]

$$\Delta^m s_l^C = O(l^{-m}); \tag{4.9}$$

that is, each successive difference goes to zero more rapidly than its predecessor by one power of l. This suggests that the same might be true of the screened S matrix, and that this would be true *uniformly for all screening radii* ρ. In Secs. 5 and 6 we prove that this is so. Specifically, we prove, subject to the conditions of Sec. 3 on $\alpha^{\rho}(r)$, that as $\rho \to \infty$

$$\Delta^{m} s_l^{\rho} = O(l^{-m}), \tag{4.10}$$

uniformly in ρ for $m = 1,2,3,4$.

Substituting the bound (4.10) into the expression (4.8) for the coefficient g_l^{ρ} we see immediately that as $l \to \infty$

$$g_l^{\rho} = O(l^{-3}), \tag{4.11}$$

uniformly in ρ. This means that the series (4.6) for $(1-x)^2 f^{\rho}(x)$ converges uniformly for all angles and all ρ.

We are now ready to prove the essential result (4.2). Both amplitudes in (4.2) can be expanded in Legendre series so that

$$(1-x)^2[f^{\rho}(x) - e^{2i\zeta(\rho)} f^{C}(x)] = \sum [g_l^{\rho} - e^{2i\zeta(\rho)} g_l^{C}] P_l(x). \tag{4.12}$$

This series converges uniformly for all ρ. Now, it follows from the results of Ref. 1 (p. 328) that for any fixed l,

$$[g_l^{\rho} - e^{2i\zeta(\rho)} g_l^{C}] \underset{\rho\to\infty}{\to} 0. \tag{4.13}$$

Since the series (4.12) converges uniformly in ρ we can interchange the summation with the limit $\rho \to \infty$. It immediately follows that (4.12) goes to zero as $\rho \to \infty$, and our proof is complete.

5. BOUND ON THE FIRST DIFFERENCE Δs_l

It remains to prove that $\Delta^{m} s_l^{\rho} = O(l^{-m})$ as $l \to \infty$, uniformly in ρ for $m = 1,2,3,4$. In this section we treat the case $m = 1$ in detail; that is, we prove that

$$\Delta s_l^{\rho} = O(l^{-1}), \tag{5.1}$$

uniformly in ρ as $l \to \infty$. In the next section we sketch the proof for the higher differences, referring the reader to Ref. 7 for further details. Throughout these two sections we shall make repeated use of the "large O" symbol, as in Eq. (5.1). In all cases the bound will apply as $l \to \infty$ and will be uniform in ρ.

The partial-wave S matrix s_l^{ρ} (which we abbreviate to s_l from now on) can be found from the normalized radial wave function $\psi_l(r)$. This function satisfies the radial equation[16]

$$\psi_l'' = \left(2V^{\rho}(r) + \frac{l(l+1)}{r^2} - 1\right)\psi_l, \tag{5.2}$$

and the boundary conditions

$$\psi_l \to \text{const} \times r^{l+1} \quad (\text{as } r \to 0), \tag{5.3}$$

and

$$\psi_l \to \tfrac{1}{2} i\,[e^{-i(r - l\pi/2)} - s_l e^{i(r - l\pi/2)}] \quad (\text{as } r \to \infty). \tag{5.4}$$

The first difference Δs_l can be expressed as an integral of wave functions using an integral identity first published by Calogero.[17] We let $\psi_1(r)$ and $\psi_2(r)$ denote two functions satisfying equations of the form

$$\psi_i''(r) = U_i(r)\psi_i(r) \quad (i = 1,2), \tag{5.5}$$

and let $F(r)$ be any function (such as $F = 1$ or $F = r^{-n}$) for which the integrals and limits in Eqs. (5.6) and (5.7) below exist. Finally we denote by $G(r)$ the integral

$$G(r) = \int_{\infty}^{r} dr' (U_1 - U_2) F. \tag{5.6}$$

It is an elementary (though tedious) exercise to verify, using successive integrations by parts, that

$$\int_0^{\infty} dr\, \psi_1\psi_2\, [F''' - 2(U_1 + U_2)F' - (U_1 + U_2)'F + (U_1 - U_2)G] = [\psi_1\psi_2\{F'' - (U_1 + U_2)F\} - (\psi_1\psi_2)'F' + 2\psi_1'\psi_2' F + (\psi_1'\psi_2 - \psi_1\psi_2')G]_0^{\infty}. \tag{5.7}$$

Calogero calls this useful, if inelegant, identity the generalized Wronskian relation, since it implies, as a special case, the standard Wronskian relation.[18] By making suitable choices for the functions ψ_1, ψ_2, and F we can use (5.7) to prove the desired bound, (5.1).

We first write the identity (5.7) for the case that $\psi_1 = \psi_{l+1}$, $\psi_2 = \psi_l$, and $F = 1$. With these choices most of the terms in (5.7) are zero, and it is easily checked [using the asymptotic forms (5.3) and (5.4)] that (5.7) reduces to

$$\Delta s_l = -4i \int_0^{\infty} dr\, \psi_{l+1}\psi_l V' \tag{5.8}$$

(where we omit the superscript ρ from V^{ρ} and V' denotes dV/dr as usual). This is the starting point for our bound on Δs_l.

Applying the Schwartz inequality to (5.8) we see that[19]

$$|\Delta s_l| \leqslant 4\left(\int |\psi_{l+1}|^2 |V'| \int |\psi_l|^2 |V'|\right)^{1/2}. \tag{5.9}$$

Now, we know from (3.2) that $|V'| \leqslant M/r^2$ (where M is independent of r and ρ). Thus (5.9) will imply the desired result if we can prove that

$$\int |\psi_l|^2/r^2 = O(l^{-1}). \tag{5.10}$$

In order to prove this we shall prove more generally that

$$I_n = I_n(l) \equiv \int |\psi_l|^2/r^n = O(l^{1-n}) \quad (2 \leqslant n \leqslant 2l+2), \tag{5.11}$$

where, as usual, the bound applies as $l \to \infty$ and is uniform in ρ.

To prove (5.11) we use the generalized Wronskian relation (5.7) with $\psi_1 = \psi_l$ and $\psi_2 = \psi_l^*$. If we first take $F = 1$, then (5.7) reduces to

$$1 = 2\int |\psi_l|^2 \left(\frac{l(l+1)}{r^3} - V'\right). \tag{5.12}$$

Since, by (3.2), $|V'| \leqslant M/r^2$, this equation implies that

$$1 = O(1) I_2 + \bar{O}(l^2) I_3, \tag{5.13}$$

where as usual $O(1)$ denotes a number that is bounded (uniformly in ρ, as $l \to \infty$) and we have introduced the convenient notation[20]

$$f = \bar{O}(g), \tag{5.14}$$

to signify that

$$f = O(g) \quad \text{and} \quad g = O(f). \tag{5.15}$$

If we again take $\psi_1 = \psi_l$ and $\psi_2 = \psi_l^*$ in (5.7), but choose $F = r^{-n}$ $(n \geqslant 2)$, then we find[21]

$$\int \frac{|\psi_l|^2}{r^n} = \int \frac{|\psi_l|^2}{r^n}\left(2V - \frac{V'r}{n-1} + \frac{n}{n-1}\frac{(2l+1)^2 - n^2}{4r^2}\right). \tag{5.16}$$

Again using the bound (3.2) on V and V' we obtain the relation

$$I_n = O(1)\, I_{n+1} + \bar{O}(l^2) I_{n+2}, \tag{5.17}$$

for any $n \geqslant 2$.

The two relations (5.13) and (5.17) are sufficient to prove the required bound (5.11). It is convenient to define a quantity J_n such that

$$I_n = l^{1-n} J_n, \tag{5.18}$$

in terms of which (5.13) reads

$$1 = O(l^{-1})\, J_2 + \bar{O}(1)\, J_3, \tag{5.19}$$

while (5.17), with $n = 2$ and 3, gives the two relations

$$J_2 = O(l^{-1})\, J_3 + \bar{O}(1)\, J_4 \tag{5.20}$$

and

$$J_3 = O(l^{-1})\, J_4 + \bar{O}(1)\, J_5. \tag{5.21}$$

From the Schwartz inequality it is clear that

$$J_4^2 \leqslant J_3 J_5. \tag{5.22}$$

Inserting this into (5.21) we obtain a quadratic equation for $\surd J_5$ in terms of $\surd J_3$. Solving this we find that

$$J_5 = \bar{O}(1)\, J_3, \tag{5.23}$$

and hence from (5.22) that

$$J_4 = O(1)\, J_3. \tag{5.24}$$

Next (5.20) implies that

$$J_2 = O(1)\, J_3, \tag{5.25}$$

and then (5.19) that $J_3 = \bar{O}(1)$. Inserting this in (5.25), we find

$$J_2 = O(1),$$

and thence immediately

$$J_n = O(1), \tag{5.26}$$

for all n for which the integrals exist (i.e., $2 \leqslant n \leqslant 2l+2$).

From (5.18) it now follows that $I_n = O(l^{1-n})$. In particular, this implies the bound (5.10) which, as discussed in connection with (5.9), implies that $\Delta s_l = O(l^{-1})$ as required.

6. BOUNDS ON HIGHER DIFFERENCES

The proof that $\Delta^m s_l = O(l^{-m})$ for $m = 2,3,4$ is analogous to that for the case $m = 1$, though significantly more complicated. We define a sequence of auxiliary functions $\psi_l^m(r)$ such that

$$\psi_l^0 \equiv \psi_l,$$

while

$$\psi_l^{m+1} = (\psi_{l+1}^m)' - \psi_l^m + (l+1)\psi_{l+1}^m / r.$$

From the asymptotic form (5.4) of ψ_l it follows at once by induction on m that

$$\psi_l^m(r) \underset{r\to\infty}{\longrightarrow} \Delta^m s_l e^{i(r - l\pi/2)}/2i \quad (m \geqslant 1).$$

In other words, the asymptotic form of the auxiliary function ψ_l^m tells us the mth difference $\Delta^m s_l$ in much the same way that that of ψ_l tells us s_l itself.

If the potential is zero, then all of the auxiliary functions ψ_l^m $(m \geqslant 1)$ are zero. (This follows from the recurrence relation for Riccati–Bessel function.) In general ψ_l^m satisfies an equation like the radial equation (5.2) but with an inhomogeneous term that is a linear combination of functions $\psi_{l'}^{m'}$ with $m' < m$. From this equation one can derive an identity analogous to the generalized Wronskian relation (5.7). From this, one can obtain integrals for the $\Delta^m s_l$ and thence bounds analogous to the bounds (5.19)–(5.21) (although somewhat more complicated). And from these, one can derive the required bound $\Delta^m s_l = O(l^{-m})$ as described in detail in ref. 7. With this, our proof is complete.

7. CONCLUSION

We have shown that for smooth screening functions $\alpha^\rho(r)$ (satisfying the conditions of Sec. 3) the screened Coulomb amplitude f^ρ approaches the pure Coulomb amplitude, times the expected phase factor,

$$[f^\rho(\cos\theta) - e^{2i\xi(\rho)} f^C(\cos\theta)] \to 0 \tag{7.1}$$

as $\rho \to \infty$. This limit is pointwise and uniform for $-1 \leqslant \cos\theta \leqslant 1 - \epsilon$ (any $\epsilon > 0$) and for energies in an finite closed interval excluding zero.[22]

Our proof can easily be extended to the case of a Coulomb potential plus a smooth short–range potential. However, it should be emphasized that we do have to require that the short-range potential is smooth [in the sense that the complete screened potential satisfies conditions (3.2)].

Finally, we should mention that for the case of a repulsive Coulomb potential our method yields an estimate of the rate at which the limit (7.1) is approached. For this we must, of course, specify how the screening function $\alpha^\rho(r)$ depends on ρ. A very natural way to do this is to suppose that $\alpha^\rho(r)$ is in fact a function of r/ρ,

$$\alpha^\rho(r) = \alpha(r/\rho).$$

With this assumption, our method shows that the difference in (7.1) is $O(\rho^{-1/2})$ as $\rho \to \infty$. From this it is easily seen that the differential cross section satisfies the same bound

$$\left(\frac{d\sigma}{d\Omega}\right)^\rho - \left(\frac{d\sigma}{d\Omega}\right)^C = O(\rho^{-1/2}), \tag{7.2}$$

as $\rho \to \infty$. If the potential is not repulsive, then the left side of (7.2) goes to zero, of course; but we, as yet, have no estimate of its rate of convergence.

[1] J. R. Taylor, Nuovo Cimento B **23**, 313 (1974).
[2] M. D. Semon and J. R. Taylor, Nuovo Cimento A **26**, 48 (1975).
[3] M. D. Semon and J. R. Taylor, J. Math. Phys. **17**, 1366 (1976).
[4] M. D. Semon and J. R. Taylor, Phys. Rev. A **16**, 33 (1977).

[5]We follow the notation of Ref. 1 and use units with $\hbar = m = p = 1$ (where m is the mass and p the momentum of the scattered particle).

[6]The requirement that $\phi = 0$ when $\theta = 0$ simply excludes discussion of forward scattering, which is well known to be infinite for Coulomb scattering. We required $\phi(\cos\theta)$ to be "smooth" in the sense that it has a continuous second derivative.

[7]D. M. Goodmanson, "Coulomb Scattering as the Limit of Screened Coulomb Scattering," thesis submitted to the University of Colorado in partial fulfillment of requirements for the Ph.D. degree, 1978.

[8]A. M. Veselova, Theor. Mat. Fiz. **3**, 326 (1970).

[9]E. O. Alt, W. Sandhas, and H. Ziegelman, Phys. Rev. C **17**, 1981 (1978).

[10]G. Bencze, Nucl. Phys. A **196**, 135 (1972); S. P. Merkuriev, Yad. Fiz. **24**, 289 (1976).

[11]See, for example, *Higher Transcendental Functions*, edited by A. Erdelyi (MaGraw-Hill, New York, 1953), Vol. II, p. 198.

[12]The simplest way to *prove* the two pointwise limits (2.3) and (2.5) is perhaps the following: The function $(1 - x)^2 f^C(x)$ in (2.5) is well behaved and certainly has a normal Legendre expansion. The coefficients in this expansion are given by the usual formula, $1/2(2l + 1)$ times the integral of $(1 - x)^2 f^C(x)$ times $P_l(x)$. This integral is given in Ref. 13, Eq. (7.127) and yields the coefficient (2.6). We can then *divide* Eq. (2.5) by $(1 - x)$ and, provided $-1 < x < 1$, we can rearrange terms to give (2.3).

[13]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1965).

[14]It is easily checked that the corresponding Coulomb coefficient g_l^C defined in (2.6) is given by a corresponding expression involving the differences $\Delta^m s_l^C$.

[15]The notation $a_l = O(b_l)$ as $l \to \infty$ means that there is a constant K such that $|a_l| < K |b_l|$ for l sufficiently large.

[16]We use units with $\hbar = m = p = 1$. Note that what we call V here was denoted by $V/2$ in Ref. 7.

[17]F. Calogero, "Generalized Wronskian Relations" in *Studies in Mathematical Physics*, edited by E. Lieb, B. Simon, and A. S. Wightman (Princeton U.P., Princeton, N.J., 1976).

[18]$\int_0^\infty dr\, \psi_1\psi_2(U_1 - U_2) = [\psi_1'\psi_2 - \psi_1\psi_2']_0^\infty$.

[19]From now on all integrals are over r, running from 0 to ∞.

[20]We have been unable to find an accepted notation for this convenient idea. Order relations like $f = O(g)$ can be multiplied together, but not divided. One can also *divide* by the relation $f = \bar{O}(g)$. In particular, $[\bar{O}(l^n)]^{-1} = \bar{O}(l^{\ n})$. In Eq. (5.13) we use the obvious result that $2l(l + 1) = \bar{O}(l^2)$.

[21]The condition $n \geqslant 2$ is needed to ensure convergence of the integrals at the upper limit. The integrals diverge at their lower limit if $2l < n$, but this does not matter since we are interested in bounds as $l \to \infty$.

[22]By using units with $p = 1$ we have obscured the fact that our bounds are uniform in p (on any compact interval excluding zero). However, it is easy to check that this is so.

Resonances, scattering theory, and rigged Hilbert spaces [a)]

Guido Parravicini and Vittorio Gorini
Department of Physics, Center for Particle Theory, The University of Texas at Austin, Austin, Texas 78712 [b)]

E. C. G. Sudarshan
Department of Physics, Center for Particle Theory, The University of Texas at Austin, Austin, Texas 78712 and Centre for Theoretical Studies, Indian Institute of Sciences, Bangalore 560012, India [b)]

(Received 2 February 1979; accepted for publication 5 December 1979)

The problem of decaying states and resonances is examined within the framework of scattering theory in a rigged Hilbert space formalism. The stationary free,"in," and "out" eigenvectors of formal scattering theory, which have a rigorous setting in rigged Hilbert space, are considered to be analytic functions of the energy eigenvalue. The value of these analytic functions at any point of regularity, real or complex, is an eigenvector with eigenvalue equal to the position of the point. The poles of the eigenvector families give origin to other eigenvectors of the Hamiltonian: the singularities of the "out" eigenvector family are the same as those of the continued S matrix, so that resonances are seen as eigenvectors of the Hamiltonian with eigenvalue equal to their location in the complex energy plane. Cauchy theorem then provides for expansions in terms of "complete" sets of eigenvectors with complex eigenvalues of the Hamiltonian. Applying such expansions to the survival amplitude of a decaying state, one finds that resonances give discrete contributions with purely exponential time behavior; the background is of course present, but explicitly separated. The resolvent of the Hamiltonian, restricted to the nuclear space appearing in the rigged Hilbert space, can be continued across the absolutely continuous spectrum; the singularities of the continuation are the same as those of the "out" eigenvectors. The free, "in" and "out" eigenvectors with complex eigenvalues and those corresponding to resonances can be approximated by physical vectors in the Hilbert space, as plane waves can. The need for having some further physical information in addition to the specification of the total Hamiltonian is apparent in the proposed framework. The formalism is applied to the Lee–Friedrichs model and to the scattering of a spinless particle by a local central potential.

1. INTRODUCTION AND SUMMARY

Recently, there has been a revival of interest in the study of the dynamical behavior of unstable states and resonances in quantum mechanics, based on technically an conceptually different approaches (see, e.g., Ref. 1 and references cited therein; for a different approach, in the context of nonequilibrium statistical mechanics, see Ref. 2). One of the purposes of these investigations is the study of the time behavior of the survival probability of an unstable system, in order to compare the classical and quantum predictions and reconcile their apparent contradictions. As is well known,[1] a purely exponential decay law for the survival probability of an unstable system is forbidden in quantum mechanics within the Hamiltonian formalism, since deviations from exponential behavior are necessarily bound to occur at short as well as at long times compared to the mean life.[3] A line of research has been developed in which unstable states have been considered from the point of view of the theory of open systems, in so far as their interaction with the environment (e.g., the measuring devices) is assumed to play a significant role in their evolution. Within this scheme, one can justify the derivation of a semigroup law for the evolution of the density matrix representing the unstable state, thereby obtaining for the survival probability an exponential behavior at all times, in agreement with the classical prediction.[1] This derivation is based on physical arguments concerning the reduction of the density matrix of the system due to the interaction with the environment. In this context, the Markovian approximation leading to a semigroup law for the density matrix and consequently to an exponential decay law, may be justified by the usual argument based on the separation of two typical time scales for the system and the surroundings.[4]

In S matrix theory, the treatment of unstable states and resonances has been based on the association of such states to second sheet poles of the analytic continuation of the S matrix in the lower half-plane across the cut determined by the absolutely continuous spectrum of the Hamiltonian. In this paper we conform to this point of view as a working hypothesis, even though there is no logical reason for its general validity. For example, it is possible to construct models in which a resonance appears, whereas the analytically continued S matrix has no poles in the unphysical sheet.[5] Conversely, the Friedrichs model[6] with a nonanalytic potential provides an example of a resonance model in which the S matrix is not even analytically continuable.[7] Resonances appear also as singuarities of the continuation of an appropriate family of matrix elements of the resolvent or of operators connected to the resolvent.[8–15] Such treatments apply equal-

[a)]Supported in part by DOE Contract No. EY-76-S-05-3992, by NATO Research Grant No. 1380, by CNR Contratto di Ricerca No 77.01543.63, and by INFN, Sezione di Milano.

[b)]Permanent address: Istituto di Fisica dell'Università, via Celoria 16 20133 Milano, Italy, and INFN, Sezione di Milano, Italy.

ly well in two different classes of models in both of which the Hamiltonian is envisaged as a sum $H = H^0 + V$ of a "free" part plus an "interaction" V. For the sake of simplicity, consider the case of a two-body resonance. On the one hand, one can consider the system from the point of view of a typical scattering problem. Then H_0 is the Hamiltonian one would ascribe to the two particles if they were noninteracting and V represents their interaction potential which may give rise to peaks in the scattering cross section, which are interpreted as quasibound states. The mean life of these resonant states is given by $1/\Gamma_0$, if Γ_0 denotes the full width at half maximum. This behavior of the cross section is ascribed to the presence of a pole of the second sheet continuation of the S matrix at the point $z_0 = E_0 - i(\Gamma_0/2)$. In an alternative approach, which suggests itself more naturally when one wishes to describe long-lived unstable states, one considers H_0 as the hypothetical Hamiltonian of the system plus its decay products in the absence of the interaction V which is responsible for the decay. In these models, H^0 has one or more eigenvalues m_0 embedded in the continuum. As V is switched on, the corresponding stationary states may become unstable and the eigenvalues are absorbed in the continuum. Granted suitable analyticity properties, the poles of the resolvent reduced to the proper eigenstates of H^0 move off the real axis to points $z_0 = E_0 - i(\Gamma_0/2)$ into the second sheet, close to the unperturbed energies m_0. There are of course physical situations in which both features of the above description appear, for example in the case of scattering of electrons off He^+ atoms, where the cross section has peaks occurring at the energies of the autoionizing states. A simple treatment which unifies the two models outlined above has been proposed some time ago by Horwitz and Marchand,[14] based on the notion of decay-scattering system. The problem of eigenvalues embedded in the continuum which may dissolve under the action of a perturbation was first considered by Weisskopf and Wigner[16] in connection with the study of the width of spectral lines. Its first mathematically rigorous treatment was given by Friedrichs in a particular model[6] and subsequently developed in a general context, in connection with the theory of spectral concentration. [12a,b,17] The Friedrichs model is equivalent to the Lee model[18] in the first nontrivial sector, with an unstable V particle. The theory of dilation analytic potentials,[10,11,19] has allowed to interpret resonances as isolated eigenvalues of a suitable non-self-adjoint analytic continuation of the Hamiltonian, thereby providing also a natural dense set of vectors in the Hilbert space (the analytic vectors of the group of dilations) for which the expectation values of the resolvent can be continued to the second sheet and exhibit poles exactly at the location of the resonances. As we shall see, a physically motivated selection of the matrix elements of the resolvent which are to be continued is of crucial importance.

In the analysis of the behavior of an unstable system described by a normalized state vector φ prepared at the time $t = 0$, the main object of study is the survival amplitude

$$A(t) = (\varphi, \exp(-iHt)\varphi), \quad t \geqslant 0, \tag{1.1}$$

where H is the total Hamiltonian of the system. In two preceding papers,[20,21] we have carried out a detailed study of the structure of (1.1) in a simple analytic resonance model. Assuming the existence at the point $z_0 = E_0 - i(\Gamma_0/2)$ of a simple pole of the second sheet continuation in the lower right quadrant of the matrix element $(\varphi, R(z)\varphi)$ of the resolvent $(H - z)^{-1}$ and deforming the integration contour, one expresses the survival amplitude as a sum of two contributions:

$$A(t) = A_0(t) + A_\Gamma(t). \tag{1.2}$$

They have respectively the following forms:

$$A_0(t) = b_0 \exp(-iE_0 t)\exp[-(\Gamma_0/2)t], \tag{1.3}$$

and

$$A_\Gamma(t) = \int_\Gamma \exp(-izt) b(z)\, dz. \tag{1.4}$$

The purely exponential term (1.3) arises from the residue of the S matrix at the pole z_0 and, if the resonance is sufficiently narrow, it dominates the survival amplitude over a long range of intermediate times $\tau_1 < t < \tau_2$, with τ_1 and τ_2 being respectively very short and very long as compared to the mean life $1/\Gamma_0$ of the unstable state. On the other hand, the background integral (1.3), which is performed along a complex contour Γ running below the location of the pole, is responsible for the deviations of the decay law from the pure exponential at short ($t < \tau_1$) and long ($t > \tau_2$) times (estimates of the time parameters τ_1 and τ_2 have been calculated in various models: se Refs. 3a, 20, and 1, and references quoted therein). In the special case of the Friedrichs model, using a procedure of analytic continuation which formally corresponds to a deformation of the spectrum of the Hamiltonian in the complex plane, we have shown in Ref. 21 that the representations (1.2)–(1.4) of the survival amplitude can be interpreted in a formal way as an expansion of the unstable state over a "complete set" of generalized eigenvectors of the Hamiltonian, which corresponds to a continuum of *complex* eigenvalues z along the contour Γ plus a "discrete" and likewise complex eigenvalue z_0 at which the second sheet continuation of the S matrix exhibits a pole. In this way, it is natural to associate the unstable particle to the pole $z_0 = E_0 - i(\Gamma_0/2)$, with mass E_0 and width Γ_0, and ascribe to it the corresponding generalized eigenvector $F(z_0)$ of the Hamiltonian. It is the purpose of this paper to give a rigorous mathematical foundation to this interpretaton in terms of complex eigenvalues and eigenvectors of the Hamiltonian in the framework of the theory of self-adjoint operators in rigged Hilbert spaces.[22–25] We do this in a way which provides an immediate technique for the analytic continuation of the resolvent across the absolutely continuous spectrum, so that we are able to associate to each other isolated singularities of the continued resolvent, "pure exponential" contributions to the survival amplitude of decaying states and complex eigenvalues of the Hamiltonian H which are characterized by the fact that they are isolated singularities of analytic families of eigenvectors of H. In order to get rid of the arbitrariness in the location of such singularities, we tie our formalism to S matrix theory, thus providing also a "deformalization" of the formal theory of scattering.

Even though physicists have liked to think intuitively of resonances as states associates to "complex eigenvalues" of the Hamiltonian, it was not until Grossman's paper appeared[9] that such an idea was given rigorous ground. Later, Baumgärtel showed explicitly how, whithin the rigged Hilbert space formalism, one can associate resonances to eigenvectors of the Hamiltonian, with complex eigenvalues, for a large class of perturbations.[13] Our approach is close to his. The main difference is that for us the solution of the eigenvalue equation outside the position of the resonances is of relevant importance.

The general framework is laid down in Sec. 2. There, we show that by extending a self-adjoint operator A to the dual of a suitable nuclear space, it is possible to associate to A various "complete sets" of generalized eigenvectors corresponding to complex eigenvalues. Even if the spectrum of A is absolutely continuous, these complete sets will in general contain both a "continuum" as well as "discrete" contribution, the latter arising from singularities of the analytic family of eigenvectors.

In Sec. 3, we apply this formalism to nonrelativistic scattering theory. We consider a Hamiltonian of the form $H = H^0 + V$ on a Hilbert space $\mathscr{H}$ and a suitable nuclear space Φ densely and continuously embedded into $\mathscr{H}$. We assume Φ to be in the domains of H^0 and H and that both H^0 and H map Φ into itself continuously. Then, under suitable assumptions regarding the analyticity properties of the eigenvectors of the absolutely continuous parts of the operators H^0 and H extended to the dual space Φ', we show that the second sheet singularities, in the lower half-plane, of the S matrix and of the resolvent of H, as a continuous operator from Φ to Φ', are the same as the singularities of the second sheet continuation of the stationary scattering states. Whenever these singularities are poles z_0, which can be interpreted as resonances of the system, certain coefficients of the Laurent expansion of the eigenvectors of H about z_0 are themselves eigenvectors (and possibly associated vectors) of H with z_0 as eigenvalue. They form the "discrete" part of a suitable complete set of eigenvectors of the restriction of H to its subspace of absolute continuity, with complex eigenvalues. These "discrete" eigenvectors are the stationary states of H which are responsible for the leading exponential decay of the resonant wave packet. Therefore, they can be interpreted as describing the decaying state, much in the same way as the stationary scattering states can be viewed as representing an incoming or outgoing particle in the idealized limiting situation of a well-defined value of the magnitude of the momentum.

In Sec. 4 ad 5 we implement the above theory in the cases of the Friedrichs model[6] and of the scattering of a spinless particle by a local central potential, with suitable assumptions on the corresponding potentials, by providing explicitly nuclear spaces Φ which fit all the required purposes. We show in another paper[26] that the same formalism goes through for the class of degenerate potentials, which have applications in nuclear physics. The theory can also be applied to other interesting models.[27]

Section 6 is dedicated to some concluding remarks and particularly to the physical consequences of the choice of the rigged Hilbert space.

A short summary of our results is contained in Ref. 28.

2. ANALYTIC FAMILIES OF EIGENVECTORS OF SELF-ADJOINT OPERATORS

In this section we deal with the topic of complex eigenvalues for self-adjoint operators in rigged Hilbert spaces. We show that under certain conditions the completeness expansion in terms of generalized eigenvectors of a self-adjoint operator may be rewritten in terms of eigenvectors relative to complex eigenvalues. This will be the starting point for our treatment of resonances and/or unstable states. Concerning rigged Hilbert spaces (or Gel'fand triples) and generalized eigenvectors see, e.g., Refs. 22–25.[29] In order to implement our formalism, we need to recall and reformulate a few concepts relative to these topics.

Let $\mathscr{H}$ be a separable Hilbert space with inner product (h,k), linear in k and let $J: h \rightarrow \bar{h}$ be an antilinear unitary involution (conjugation) on $\mathscr{H}$. A bilinear form is defined on $\mathscr{H}$ by

$$\langle h | k \rangle = :(\bar{h},k). \tag{2.1}$$

All nuclear spaces we consider in this paper are barreled and complete,[30] a fact that will be henceforth understood without further specifications. Let Φ be a nuclear space continuously and densely embedded into $\mathscr{H}$ via the linear map $\mathscr{J}$. We assume Φ to be invariant under J, and J to be continuous with respect to (wrt) the topology of Φ. We denote by Φ' the (topological) dual of Φ endowed with the weak topology, and if $\varphi \in \Phi$ and $F \in \Phi'$ we write $\langle F|\varphi\rangle$ or $\langle \varphi|F\rangle$ for the image of φ under F. $\mathscr{H}$ is continuously, densely, and *linearly* embedded into Φ' by the map $\mathscr{J}'$ defined by:

$$\langle \mathscr{J}'h | \varphi \rangle = :\langle h | \mathscr{J}\varphi \rangle = (\bar{h}, \mathscr{J}\varphi). \tag{2.2}$$

Therefore, we have a rigged Hilbert space $\Phi \subset \mathscr{H} \subset \Phi'$, where both embeddings $\mathscr{J}$ and $\mathscr{J}'$ are linear. As J maps Φ onto itself continuously, we can extend its action to the whole of Φ' by $\hat{J}: F \rightarrow \bar{F}$, where

$$\langle \bar{F} | \varphi \rangle = :\overline{\langle F | \bar{\varphi} \rangle}, \quad \forall \varphi \in \Phi. \tag{2.3}$$

The bar over complex numbers denotes ordinary complex conjugation. The above formula actually defines the extension of J, because $\overline{\mathscr{J}'h} = \mathscr{J}'\bar{h}$. $\hat{J}$ is an antilinear continuous involution on Φ'.

Next, let A be a densely defined linear oprator in $\mathscr{H}$ with domain $D(A)$ and adjoint A^*, such that $D(A) \supset \Phi$, $D(A^*) \supset \Phi$, $A\Phi \subset \Phi$, $A^*\Phi \subset \Phi$, and both A and A^* are continuous wrt the topology of Φ. The *transpose* A' of A is defined on Φ' by:

$$\langle A'F | \varphi \rangle = :\langle F | A\varphi \rangle \quad \forall \varphi \in \Phi, \quad \forall F \in \Phi'. \tag{2.4}$$

When $\mathscr{H} = \mathbb{C}^n$, A' is actually represented by the transpose of the matrix A, when J is the componentwise complex conjugation.

The *extension* $\hat{A}$ of A is defined on Φ' according to:

$$\langle \hat{A}F | \varphi \rangle = :\langle F | \overline{A^*\bar{\varphi}} \rangle \quad \forall \varphi \in \Phi, \quad \forall F \in \Phi'. \tag{2.5}$$

This definition is justified by noting that $\hat{A}\mathscr{J}'h = \mathscr{J}'Ah$, $\forall h \in D(A)$. It has to be remarked that, if A is symmetric and $h \in D(A^*)$, then $\hat{A}\mathscr{J}'h = \mathscr{J}'A^*h$. Indeed, $\forall h \in D(A^*)$ and

$\forall\varphi\in\Phi$, the following identities hold: $\langle \hat{A}\mathscr{F}'h|\varphi\rangle = \langle \mathscr{F}'h|\overline{A^*\overline{\varphi}}\rangle = \langle \mathscr{F}'h|\overline{A\overline{\varphi}}\rangle = (\overline{h, A\overline{\varphi}}) = (A\overline{\varphi}, h) = (\overline{\varphi}, A^*h) = (A^*h, \varphi) = \langle \mathscr{F}'\hat{A}^*h|\varphi\rangle$. Therefore, if A is symmetric, it follows that $\widehat{A^*} = \hat{A}$ and, similarly, $A^{*\prime} = A'$.

Definition 2.1: Let $\mathscr{H}$ and Φ be as above and A be an operator in $\mathscr{H}$ satisfying the assumptions following (2.3). We say that the complex number λ (respectively, μ) is a *right* (respectively, *left*) eigenvalue of A corresponding to the right (respectively, *left*) eigenvector $F(\lambda)\in\Phi'$ (respectively, $G(\mu)\in\Phi'$ if $\hat{A}F(\lambda) = \lambda F(\lambda)$ [respectively, $A'G(\mu) = \mu G(\mu)$].

This definition is justified by the usual convention regarding matrix multiplication in the finite-dimensional case. The relation between right and left eigenvalues and eigenvectors is given by the following proposition.

Proposition 2.1: Let λ be a right (respectively, left) eigenvalue of A corresponding to the right (respectively, left) eigenvector $F(\lambda)$. Then $\overline{F}(\lambda)$ is a left (respectively, right) eigenvector of A^* with corresponding left (respectively, right) eigenvalue $\bar{\lambda}$. The same statement holds with A and A^* interchanged.

Proof: Suppose that $\langle F(\lambda)|\overline{A^*\overline{\varphi}}\rangle = \lambda\langle F(\lambda)|\varphi\rangle$ for all $\varphi\in\Phi$. Then,

$$\langle A^{*\prime}\overline{F}(\lambda)|\varphi\rangle = \langle \overline{F}(\lambda)|A^*\varphi\rangle = \overline{[\langle F(\lambda)|\overline{A^*\overline{\varphi}}\rangle]} = \overline{\lambda\langle F(\lambda)|\overline{\varphi}\rangle} = \bar{\lambda}\langle \overline{F}(\lambda)|\varphi\rangle.$$

The remaining statements are proved analogously.

Corollary: Let A be as above and symmetric. If λ is a right (left) eigenvalue corresponding to the right (left) eigenvector $F(\lambda)$, then $\overline{F}(\lambda)$ is a left (right) eigenvector with corresponding left (right) eigenvalue $\bar{\lambda}$.

Next, we recall the definition of Φ'-valued analytic functions[22,30] and draw a few easy consequences from it.

Definition 2.2: A function $F(\lambda)$ defined in a region $\Omega\subset\mathbb{C}$ and with values in Φ' is said to be analytic in Ω if, for any given $\varphi\in\Phi$, the ordinary function of λ, $\langle F(\lambda)|\varphi\rangle$, is analytic in Ω.

The following proposition follows immediately from the definition and from the principle of identity of analytic functions.

Proposition 2.2: Let $F(\lambda)$ be a function with values in Φ', analytic in the region Ω. Let $\lambda_i = 1, 2, 3,\cdots$, be a sequence of points in Ω with proper accumulation point within Ω. Assume that $\forall i, \lambda_i$ is a left (right) eigenvalue of an operator A as above corresponding to the left (right) eigenvector $F(\lambda_i)$. Then $\forall\lambda\in\Omega$, λ is a left (right) eigenvalue of A corresponding to the left (right) eigenvector $F(\lambda)$.

To make life simpler later on, we shall say that a function $F(\lambda)$ with values in Φ', analytic in a region Ω, is a family of right (left) eigenvectors of the operator A, analytic in Ω, whenever for any $\lambda\in\Omega$, $F(\lambda)$ is a right (left) eigenvector of A with λ as right (left) eigenvalue.

From Definition 2.1, Proposition 2.1, and well-known facts about analytic functions, Proposition 2.3 follows.

Proposition 2.3: Let A, Φ, $\mathscr{H}$ be as above, and such that there exists a family $F(\lambda)$ of right (left) eigenvectors, analytic in some region Ω. Then, the family $G(\lambda) = \overline{F}(\bar{\lambda})$ is analytic in $\bar{\Omega}$, the complex conjugate of Ω, and is a family of left (right) eigenvectors for A^*, each $G(\lambda)$ with corresponding eigenvalue λ.

Of particular relevance to us are the isolated singularities of analytic functions with values in Φ'. To this purpose we recall the following definition.

Definition 2.3:[22,30] We say that λ_0 is an isolated singularity for the function $F(\lambda)$, $F(\lambda)\in\Phi'$, if $\exists\rho>0$ such that, for any given $\varphi\in\Phi$, the function $\langle F(\lambda)|\varphi\rangle$ is analytic in the annulus $0<|\lambda-\lambda_0|<\rho$, and it is singular at λ_0 for some $\varphi\in\Phi$. Classification of singularities is obvious.

When C is a simple closed rectifiable positively oriented curve (scroc) inside the region Ω of analyticity of the function $F(z)$ and $z_0\notin C$, the integral $(1/2\pi i)\int_C [F(z)/(z-z_0)^n]\,dz$, n an integer, defines, in the weak sense, a vector in Φ', because of our assumptions on Φ.[30] Therefore, if λ_0 is an isolated singularity for the function $F(z)$, analytic in the annulus $0<|\lambda-\lambda_0|<\rho$, from the ordinary Laurent expansion $\langle\varphi|F(\lambda)\rangle = \Sigma_{n=-\infty}^{\infty} C_n(\varphi)(\lambda-\lambda_0)^n$, $\varphi\in\Phi$, one has that each $C_n(\varphi)$ is given by $\langle\varphi|C_n\rangle$ with $C_n\in\Phi'$, and that[22,30]:

$$F(\lambda) = \sum_{n=-\infty}^{\infty} C_n(\lambda-\lambda_0)^n, \tag{2.6}$$

$$C_n = \frac{1}{2\pi i}\int_C \frac{F(\lambda)}{(\lambda-\lambda_0)^{n+1}}\,d\lambda, \tag{2.7}$$

whenever C is a scroc encircling λ_0 and inside the given annulus. The series is weakly convergent.[30] When $F(\lambda)$ is a family of left or right eigenvectors for an operator A as above, one expects the coefficients of the Laurent expansion about the isolated singularity λ_0 to have properties similar to those of the coefficients of the expansion of the resolvent about its singularities. Indeed, we have[31] Proposition 2.4.

Proposition 2.4: Let $\mathscr{H}$, Φ, A be as above, and let $F(\lambda)$ be a family of right (left) eigenvectors for A. Assume $F(\lambda)$ is analytic in some region Ω and let λ_0 be an isolated singularity. Then, the coefficients of the Laurent expansion (2.6) of $F(\lambda)$ about λ_0 satisfy equations

$$\hat{A}C_n = \lambda_0 C_n + C_{n-1}, \tag{2.8}$$

or

$$A'C_n = \lambda_0 C_n + C_{n-1}, \tag{2.9}$$

according to whether the eigenvectors $F(\lambda)$ are right or left.

Proof: The operators $\hat{A}$ and A' are continuous on Φ' wrt the weak topology. Then, from (2.7) we have

$$\hat{A}C_n = \frac{1}{2\pi i}\int_C \frac{\hat{A}F(\lambda)}{(\lambda-\lambda_0)^{n+1}}\,d\lambda,$$

so that, if the $F(\lambda)$ are right eigenvectors,

$$\begin{aligned}\hat{A}C_n &= \frac{1}{2\pi i}\int_C \frac{\lambda F(\lambda)}{(\lambda-\lambda_0)^{n+1}}\,d\lambda \\ &= \frac{1}{2\pi i}\int_C \frac{F(\lambda)}{(\lambda-\lambda_0)^{n}}\,d\lambda + \frac{\lambda_0}{2\pi i}\int_C \frac{F(\lambda)}{(\lambda-\lambda_0)^{n+1}}\,d\lambda \\ &= C_{n-1} + \lambda_0 C_n.\end{aligned}$$

Equation (2.9) is proved analogously.

Remark: When $C_{n-1} = 0$, it follows that C_n is a right (left) eigenvector and λ_0 an eigenvalue. This happens in particular for the first nontrivial coefficient, when λ_0 is either regular or a pole for the family $F(\lambda)$. In general, we state[31] definition 2.4.

Definition 2.4: Let $\mathscr{H}$, Φ, A be as above and let C_n, n integer, be a sequence in Φ' such that (2.8) or, respectively, (2.9) holds for every n for some $\lambda_0 \in \mathbb{C}$. Then, if C_n and C_{n-1} are nonzero, we say that C_n is a right (respectively, left) *associated vector* of A.

Remark: Let C_n, n integer, be a sequence in Φ' as in the preceding definition and assume that the series $F(\lambda) = \Sigma_{n=-\infty}^{\infty} C_n(\lambda - \lambda_0)^n$ is weakly convergent in an annulus $\Omega = \{\lambda \,|0 < \lambda - \lambda_0| < \rho\}$. Then, according to whether (2.8) or (2.9) holds for all n, $F(\lambda)$ is an analytic family in Ω of right or left eigenvectors of A.

The Gel'fand–Maurin theorem[22,23] can be reformulated in terms of left and right eigenvectors. This requires the existence of a suitable conjugation, a fact which we shall assume once for all in the following.[32]

Theorem 2.1 (*Gel'fand–Maurin*): Let $\mathscr{H}$ be a complex separable Hilbert space, A a self-adjoint operator on $\mathscr{H}$ with domain $D(A)$ and let $\mathscr{F}$ be the generalized Fourier transform associated to the direct integral decomposition of $\mathscr{H}$ wrt A [$\sigma(A)$ denotes the spectrum of A]:

$$\mathscr{H} \to \widehat{\mathscr{H}} = \int_{\oplus \sigma(A)} \mathscr{H}(\lambda)\, d\mu(\lambda), \tag{2.10}$$

$$\mathscr{F}: h \to h(\lambda) = \{h^i(\lambda)\}, \quad h(\lambda) \in \mathscr{H}(\lambda), \quad i = 1,\ldots,\dim \mathscr{H}(\lambda), \tag{2.11}$$

$$(h(\lambda), k(\lambda))_\lambda = \sum_i \overline{h^i(\lambda)} k^i(\lambda), \tag{2.12}$$

$$[\mathscr{F}(Ah)]^i = \lambda h^i(\lambda), \quad \text{if } h \in D(A). \tag{2.13}$$

Then there exists a (not necessarily unique) nuclear space $\Phi \subset D(A)$, densely and continuously embedded in $\mathscr{H}$, which is mapped by A continuously into itself and such that $\mathscr{F}$ can be implemented by a system of left eifenvectors $G^i(\lambda)$ of A in Φ' in such a way that

$$\langle G^i(\lambda)|\varphi\rangle = [\mathscr{F}(\varphi)]^i(\lambda), \tag{2.14}$$

for all $\varphi \in \Phi$ and μ-almost everywhere (a.e.). Since $\mathscr{F}$ is unitary, this system is complete:

$$(\varphi,\psi) = \int_{\sigma(A)} d\mu(\lambda) \sum_i \overline{\langle G^i(\lambda)|\varphi\rangle} \langle G^i(\lambda)|\psi\rangle \quad \forall \varphi, \; \psi \in \Phi. \tag{2.15}$$

The vectors $F^i(\lambda) = \overline{G}^i(\lambda)$ are right eigenvectors and the completeness relation (2.15) can be rewritten in the equivalent form:

$$\langle \varphi|\psi\rangle = \int_{\sigma(A)} d\mu(\lambda) \sum_i \langle \varphi|F^i(\lambda)\rangle \langle G^i(\lambda)|\psi\rangle \quad \forall \varphi, \; \psi \in \Phi. \tag{2.16}$$

Remarks: (i) In the following, we shall restrict ourselves to the case when the operator A has no singular continuous spectrum. Then (2.15) writes

$$(\varphi,\psi) = \sum_{j,l} \overline{\langle G_j^l|\varphi\rangle} \langle G_j^l|\psi\rangle + \int_{\sigma_{ac}(A)} d\mu_{ac}(\lambda) \sum_i \overline{\langle G_{ac}^i(\lambda)|\psi\rangle} \langle G_{ac}^i(\lambda)|\psi\rangle \quad \forall \varphi, \; \psi \in \Phi, \tag{2.16'}$$

where the vectors G_j^l are proper eigenvectors of A, viz., $AG_j^l = \lambda_j G_j^l$, and where in $\sigma_{ac}(A)$ we have $G_{ac}^i(\lambda) = G^i(\lambda)\,\mu_{ac}$-a.e. In particular, it follows tht if P is the orthogonal projection onto the subspace $\mathscr{H}_{ac}$ of absolute continuity of A then for all $\varphi \in \Phi$

$$\langle G_{ac}^i(\lambda)|\varphi\rangle = [\mathscr{F}(P\varphi)]^i(\lambda) \quad \mu_{ac}\text{-a.e.} \tag{2.17}$$

It has to be noted that (2.17) holds even though it might happen that $P\Phi \not\subset \Phi$. (ii) Given a nuclear space Φ densely and continuously embedded in $\mathscr{H}$, which is mapped into itself continuously by a self-adjoint operator A, in general there might be solutions to the eigenvalue equation $A'G(\lambda) = \lambda G(\lambda)$ in Φ' which correspond (a) to eigenvalues not in $\sigma(A)$ and (b) to eigenvalues in $\sigma(A)$ whose corresponding eigenvectors play no role in the above theorem [an example of (a) is provided by $\mathscr{H} = L^2(\mathbb{R})$, $A = -id/dx$, Φ the space $\mathscr{D}$[22,23] of Schwarz test functions; an example of (b) is given in Sec. 5].

For our purposes, we need a sufficient condition which ensures that given $\mathscr{H}$, A, Φ as in the remark above, there are solutions in Φ' to the (left) eigenvalue problem for A yielding eigenvalues in $\sigma(A)$ with corresponding eigenvectors which implement the generalized Fourier transform as in Theorem 2.1. An answer to our needs is provided by the following proposition.

Proposition 2.5: Let $\mathscr{H}$ be complex separable Hilbert space, A an essentially self-adjoint operator in $\mathscr{H}$ with domain D, and $\Phi \subset D$ a nuclear space densely and continuously embedded into $\mathscr{H}$, such that the restriction $A \upharpoonright \Phi$ maps Φ into itself continuously. Let $G(\lambda)$ be a function on some closed set $\sigma \subset \mathbb{R}$ with values in Φ' and let μ be a finite regular Borel measure concentrated on σ such that

(a) $A'G(\lambda) = \lambda G(\lambda)$ μ-a.e.,

(b) $(\varphi,\psi) = \int_\sigma \overline{\langle G(\lambda)|\varphi\rangle} \langle G(\lambda)|\psi\rangle\, d\mu(\lambda) \quad \forall \varphi, \quad \psi \in \Phi.$

Let $\widetilde{\mathscr{F}}: \mathscr{H} \to L^2(\sigma,\mu)$ be the extension to $\mathscr{H}$ of the operator $\mathscr{F}: \Phi \to L^2(\sigma,\mu)$ which associates to $\varphi \in \Phi$ the function $\varphi(\lambda) \in L^2(\sigma,\mu)$ by the rule $\varphi(\lambda) = \langle G(\lambda)|\varphi\rangle\ \mu$ – a.e., and assume that

(c) $\mathscr{F}(\Phi)$ is dense in $L^2(\sigma,\mu)$;

(d) $[\widetilde{\mathscr{F}}(Ah)](\lambda) = \lambda[\widetilde{\mathscr{F}}(h)](\lambda)$ μ-a.e., $\forall h \in D$.

Then, σ is the spectrum of A, and $\widetilde{\mathscr{F}}$, Φ and the family $G(\lambda)$ are a possible choice for the implementation of the generalized Fourier transform, nuclear space and left eigenvectors of Theorem 2.1.

Proof: It follows from (b) that $\mathscr{F}$ is an isometry of Φ, endowed with the norm inherited from $\mathscr{H}$, into $L^2(\sigma,\mu)$. Its unique extension $\widetilde{\mathscr{F}}$ is unitary by (c). The operator $Q = \widetilde{\mathscr{F}} A \widetilde{\mathscr{F}}^{-1}$ in $L^2(\sigma,\mu)$ with domain $\widetilde{\mathscr{F}}(D)$ is essentially self-adjoint, and acts as a multiplication operator, by (d). The operator of multiplication by λ in $L^2(\sigma,\mu)$ with domain $D = \{h(\lambda)| \int_\sigma \lambda^2 |h(\lambda)|^2\, d\mu(\lambda) < \infty\} \supset \widetilde{\mathscr{F}}(D)$ is self-adjoint and an extension of Q, so it coincides with the closure Q^c of

Q. It follows that σ is the spectrum of Q and hence of A. Then the statement of the proposition follows from (a) and the fact that Q^c is already given in its spectral form.

Remarks: (i) Note that condition (d) does not necessarily imply condition (a). (ii) In specific instances, it might be more convenient to verify condition (b) with a nonfinite σ-finite regular Borel measure. However, this is equivalent to some finite regular Borel measure ν so that, performing the substituitions $\mu \to \nu$ and $G(\lambda) \mapsto (\sqrt{d\mu/d\nu})G(\lambda)$, we recover (b) as stated in the proposition. (iii) For simplicity, we have confined ourselves to the case of simple spectrum, the extension to the general case being straightforward. (iv) Having assumed the existence of a conjugation $h \to \bar{h}$ on $\mathscr{H}$ mapping Φ onto itself continuously and introducing the right eigenvectors $F(\lambda) = \bar{G}(\lambda)$, we can replace (b) with

(b′) $\langle\varphi|\psi\rangle = \int_\sigma \langle\varphi|F(\lambda)\rangle\langle G(\lambda)|\psi\rangle\, d\mu(\lambda) \quad \forall\varphi,\ \psi\in\Phi.$

(v) The functional calculus holds, so that if $f \in L^\infty(\sigma,\mu)$ and $\varphi\in\Phi$ we have

$$[\mathscr{F}(f(A)\varphi)](\lambda) = f(\lambda)\langle G(\lambda)|\varphi\rangle \quad \mu\text{-a.e.} \tag{2.18}$$

or, formally, $\langle G(\lambda)|f(A)\varphi\rangle = f(\lambda)\langle G(\lambda)|\varphi\rangle$ even though $f(A)\varphi$ may not belong to Φ.

Now we turn to the expansions in terms of generalized eigenvectors corresponding to *complex* eigenvalues. In this and the next section conditions allowing such expansions will be assumed. Sections 4 and 5 deal with explicit examples where they are satisfied. Let $A, \Phi, \mu, \{F_j, F(\lambda)\}, \{G_j, G(\lambda)\}$ be a self-adjoint operator, a nuclear space, a regular Borel measure and right and left eigenvectors, respectively, such as to implement Theorem 2.1: for the sake of simplicity, we are assuming that the spectrum of A is simple and we drop the subscript "ac" from the families of eigenvectors pertaining to the absolutely continuous spectrum $\sigma_{ac}(A) \equiv \Lambda$. We assume that Λ is not empty and denote by P the orthogonal projection of $\mathscr{H}$ onto $\mathscr{H}_{ac}$, the subspace of absolute continuity of A. We have $d\mu_{ac}(\lambda) = h(\lambda)\, d\lambda$, where $d\lambda$ denotes the Lebesgue measure. We assume that the function $h(\lambda)$ is the restriction to Λ of a function $h(z)$ analytic in a region Ω_1 with $\Omega_1 \cap \Lambda$ a subset of $\mathbb{R}$ with nonempty interior and that the family $G(\lambda)$ is the restriction to Λ of a family $G(z)$, analytic in a region Ω_2, with $\Omega_2 \cap \Lambda$ an infinite set with a proper accumulation point in Ω_2. Then, $F(\lambda)$ is the restriction to Λ of the family $F(z)$, analytic in $\overline{\Omega}_2$. Let Λ_0 be a subset of Λ and Λ_1 a curve such that $\Lambda_0 \cup \Lambda_1$ is a scroc and $\Lambda_0 \cup \Lambda_1 \subset (\Omega_1 \cap \Omega_2 \cap \overline{\Omega}_2)^c$ (the superscript c denoting closure). Moreover, assume that $G(z)$ is continuous on $\Lambda_0 \cup \Lambda_1$. Set $\Gamma = (\Lambda \setminus \Lambda_0) \cup \Lambda_1$ (see Fig. 1). Then, if $\varphi, \psi \in \Phi$, from

$$\langle\varphi|\psi\rangle = \sum_j \langle\varphi|F_j\rangle\langle G_j|\psi\rangle + \int_\Lambda \langle\varphi|F(\lambda)\rangle\langle G(\lambda)|\psi\rangle h(\lambda)\, d\lambda,$$

we obtain

$$\langle\varphi|\psi\rangle = \sum_j \langle\varphi|F_j\rangle\langle G_j|\psi\rangle + \int_\Gamma \langle\varphi|F(z)\rangle\langle G(z)|\psi\rangle h(z)\, dz + R(A;\varphi,\psi;h), \tag{2.19}$$

where $R(A;\varphi,\psi;h)$ is the contribution from the possible singularities $S_1, S_2, \cdots$ of the integrand in $\mathrm{Int}(\Lambda_0 \cup \Lambda_1)$, the interior of the region bounded by $\Lambda_0 \cup \Lambda_1$. Since Λ contains at least an accumulation point, it follows from the corollary to Proposition 2.1 and from Propositions 2.2 and 2.3 that the $F(z)$ and $G(z)$ appearing in the integrand at the right-hand side (rhs) of (2.19) are respectively right and left eigenvectors of A with complex eigenvalue z. In particular, if the only singularities of $F(z)$, $G(z)$ and $h(z)$ in $\mathrm{Int}(\Lambda_0 \cup \Lambda_1)$ are isolated, say $z_l = 0, 1, 2, \cdots$, we have $R(A;\varphi,\psi;h) = -2\pi i \Sigma_l \mathrm{Res}(z_l)$, where $\mathrm{Res}(z_l)$ is the residue of $\langle\varphi|F(z)\rangle\langle G(z)|\psi\rangle h(z)$ at $z = z_l$. This is evaluated in terms of the coefficients of the Laurent expansions of $\langle\varphi|F(z)\rangle$ and $\langle G(z)|\psi\rangle$, that is to say of $F(z)$ and $G(z)$, and of $h(z)$ about z_l. Therefore, it follows from Proposition 2.4 that

$$R(A;\varphi,\psi;h) = -2\pi i \sum_l \sum_j \langle\varphi|M_j(z_l)\rangle\langle L_j(z_l)|\psi\rangle a_j(z_l), \tag{2.20}$$

where $M_j(z_l)$ and $L_j(z_l)$ are right and left eigenvectors and associated vectors of A with corresponding eigenvalues z_l, and $a_j(z_l)$ are complex coefficients. If f is a function analytic in $\Omega_3 \supset \Lambda$ and bounded on the spectrum, in analogy to (2.19) one has

$$\langle\varphi|f(A)\psi\rangle = \sum_i f(\lambda_i)\langle\varphi|F_i\rangle\langle G_i|\psi\rangle + \int_\Gamma f(z)\langle\varphi|F(z)\rangle\langle G(z)|\psi\rangle h(z)\, dz + R(A;\varphi,\psi;f;h) \quad \forall\varphi,\ \psi\in\Phi. \tag{2.21}$$

Here the third term at the rhs contains also the contribution of the possible singularities of f. The extension to the case

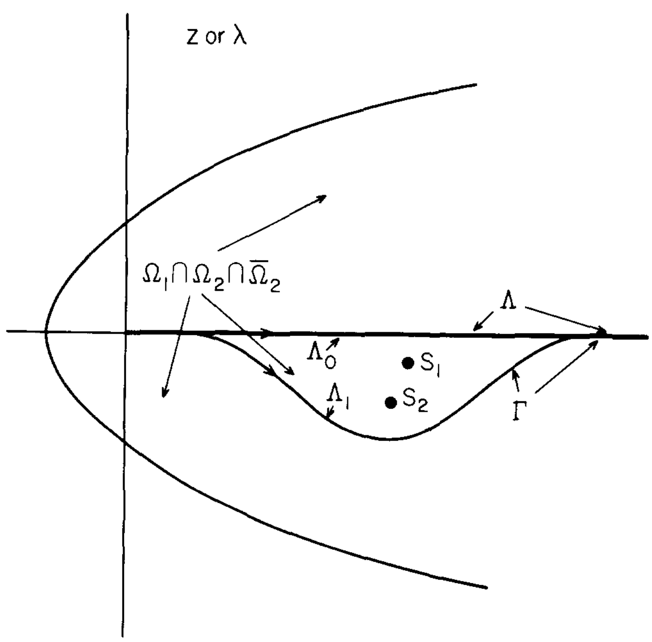


FIG. 1. The region $\Omega_1 \cap \Omega_2 \cap \overline{\Omega}_2$, the deformation Γ of the integration path Λ for the completeness of the eigenvectors, and the isolated singularities S_1 and S_2 of the integrand in the completeness relation for the eigenvectors.

when the path Γ is unbounded is trivial, provided there are no contributions to the integrals at infinity.

Formulas (2.19)–(2.21) provide, for *self-adjoint* operators, expansions in terms of "complete sets" of generalized eigenvectors and associated vectors corresponding to complex eigenvalues. We see that besides the proper eigenvectors, there are in general additional "discrete" eigenvectors and possibly associated vectors [see (2.20)] which contribute to the "complex completeness", which arise from the contributions of the isolated singularities in $\mathrm{Int}(\Lambda_0 \cup \Lambda_1)$. We have here, for a self-adjoint operator, a sort of infinite dimensional analogue of the Jordan canonical form for nonself-adjoint matrices: whereas the canonical form of a self-adjoint matrix is diagonal, the extension of a self-adjoint operator to a rigged Hilbert space may allow for the appearance of "complete sets" formed by eigenvectors and associated vectors, so that, in addition to the diagonal canonical form, there also "canonical" triangular forms.

As an example of application of formula (2.21), which is relevant to the quantum decay problem, let $f(A) = \exp(-iAt)$ and let $F(z)$ and $h(z)$ be holomorphic in $\mathrm{Int}(\Lambda_0 \cup \Lambda_1)$ and $G(z)$ have a first order pole at a point $z = z_0 \in \mathrm{Int}(\Lambda_0 \cup \Lambda_1)$. Then

$$\langle P\varphi | \exp(-iAt) P\psi\rangle = \int_\Gamma \exp(-izt) \langle \varphi | F(z)\rangle \langle G(z)|\psi\rangle h(z)\, dz - 2\pi i \exp(-iz_0 t) \langle \varphi | F(z_0)\rangle \langle C_{-1}(z_0)|\psi\rangle h(z_0) \quad \forall \varphi,\ \psi \in \Phi, \tag{2.22}$$

where [compare (2.7) and Proposition 2.4]

$$C_{-1}(z_0) = \frac{1}{2\pi i} \int_C G(z)\, dz, \tag{2.23}$$

is a left eigenvector of A with corresponding eigenvalue z_0.

Another important instance of application of (2.21) is provided by the choice of $f(A) = (A - \xi)^{-1} \equiv R(\xi;A)$. In this case, if ξ does not belong to the spectrum nor to the closure of $\mathrm{Int}(\Lambda_0 \cup \Lambda_1)$ we have

$$\langle \varphi | R(\xi;A)\psi\rangle = \sum_j \frac{1}{\lambda_j - \xi} \langle \varphi | F_j\rangle \langle G_j|\psi\rangle + \int_\Gamma \frac{1}{z-\xi} \langle \varphi | F(z)\rangle \langle G(z)|\psi\rangle h(z)\, dz + R\left(A; \varphi, \psi; \frac{1}{z-\xi}; h\right) \quad \forall \varphi,\ \psi \in \Phi. \tag{2.24}$$

The rhs of (2.24) is well defined also when $\xi \in (\Lambda_0 \setminus \sigma_s(A)) \cup \mathrm{Int}(\Lambda_0 \cup \Lambda_1)$ [$\sigma_s(A)$ denoting the singular spectrum of A] and defines a function in the plane cut along Γ, analytic everywhere except at the singularities of $R(A; \varphi, \psi; 1/(z-\xi); h)$. This shows that the (bilinear) matrix elements of the resolvent between vectors $\varphi, \psi \in \Phi$ can be continued analytically as functions of ξ, across $(\Lambda_0 \setminus \sigma_s(A))$ into $\mathrm{Int}(\Lambda_0 \cup \Lambda_1)$ and exhibit therein the same singularities that $F(z)$, $G(z)$ and $h(z)$ do. When ξ is in the region $(\Lambda_0 \setminus \sigma_s(A)) \cup \mathrm{Int}(\Lambda_0 \cup \Lambda_1)$, the expression (2.24) can be put in a more transparent form, which shows explicitly that the resolvent itself can be continued analytically as an operator in $\mathscr{L}(\Phi, \Phi')$, the space of continuous linear operators from Φ into Φ'. Indeed, let φ, $\psi \in \Phi$ and consider the matrix element

$$\langle \varphi | R(\xi;A)\psi\rangle = \sum_j \frac{1}{\lambda_j - \xi} \langle \varphi | F_j\rangle \langle G_j|\psi\rangle + \int_\Lambda \frac{1}{\lambda - \xi} \langle \varphi | F(\lambda)\rangle \langle G(\lambda)|\psi\rangle h(z)\, d\lambda. \tag{2.25}$$

Due to our assumptions on the analyticity properties of the integrand, the limits $\lim_{\eta \downarrow 0} \langle \varphi | R(x \pm i\eta;A)\psi\rangle$ exist for all φ, $\psi \in \Phi$ and for all x in the interior of $\Lambda_0 \setminus \sigma_s(A)$. Then, since Φ is barreled, the principle of uniform boundedness[30] ensures the existence of the limits $R(x \pm i0;A)$ as operators in $\mathscr{L}(\Phi,\Phi')$ and the application of Plemelj formulas gives

$$R(x+i0) = R(x-i0) + 2\pi i |F(x)\rangle \langle G(x)| h(x). \tag{2.26}$$

Therefore as an operator in $\mathscr{L}(\Phi, \Phi')$, the resolvent can be continued across $\Lambda_0 \setminus \sigma_s(A)$ into its second Riemann sheet as

$$R_{II}(\xi;A) = R(\xi;A) + 2\pi i |F(\xi)\rangle \langle G(\xi)| h(\xi). \tag{2.27}$$

Furthermore, the singularities of the continued resolvent are those of $F(\xi)$, $G(\xi)$ and $h(\xi)$.

It goes without saying that the continuation (2.27) is possible provided that the interior of $\Lambda_0 \setminus \sigma_s(A)$ is not empty.

In conclusion, in the framework above one is able to associate to each other: (1) singularities of analytic families of eigenvectors of a self-adjoint operator; (2) eigenvectors and associated vectors with corresponding eigenvalues at the locations of the singularities; (3) discrete contributions to expansions in terms of "complete sets" of eigenvectors corresponding to complex eigenvalues; and (4) singularities of the continued resolvent.

3. ANALYTICITY OF THE *S* MATRIX AND OF EIGENVECTOR FAMILIES OF THE HAMILTONIAN

The analytic structure of the eigenvector families of a self-adjoint operator depends on the choice of the nuclear space Φ (for more comments on this topic, see Sec. 6). This implies that, if one wants to associate the resonances of a given physical problem with the poles of the resolvent and of the eigenvector families of the Hamiltonian, and with discrete exponential contributions to survival amplitudes, one must choose Φ in a proper way. An inappropriate choice of Φ may indeed cancel the physical resonances and/or introduce spurious ones. In this section, to overcome this difficulty, we tie our formalism to the S matrix. We assume that resonances are associated to second sheet poles of the S matrix, in the lower half-plane. Poles of the eigenvector families and of the resolvent of the Hamiltonian, and discrete exponential contributions will be located exactly at such points.

Suppose that we are given two self-adjoint operators H^0 and H on a Hilbert space $\mathscr{H}$, which we may think of as the free and interacting Hamiltonians of a quantum system, say a free particle and the same particle interacting with a potential. We denote by P^0 and P the orthogonal projections onto the subspaces $\mathscr{H}^0_{ac}$ and $\mathscr{H}_{ac}$ of absolute continuity of H^0 and H, respectively. We assume the fulfillment of conditions[17] which ensure the existence and completeness of the Møller wave operators, i.e., the strong operator limits

$$W_{\pm} = \text{s-lim}_{t\to\mp\infty} e^{-itH} e^{-itH^0} P^0, \tag{3.1}$$

as partial isometries with initial space $\mathscr{H}^0_{ac}$ and final space $\mathscr{H}_{ac}$. Then the following relations hold

$$\text{(i)} \ W^*_{\pm} W_{\pm} = P^0, \quad W_{\pm} W^*_{\pm} = P,$$
$$\text{(ii)} \ PE(\lambda)W_{\pm} = E(\lambda)W_{\pm} = W_{\pm}E^0(\lambda) = W_{\pm}E^0(\lambda)P^0, \tag{3.2}$$
$$\text{(iii)} \ e^{-iHt}W_{\pm} = W_{\pm}e^{-iH^0t},$$

where $E^0(\lambda)$ and $E(\lambda)$ are the spectral families associated to H^0 and H, respectively. The scattering operator S is defined as

$$S = W^*_{-} W_{+}. \tag{3.3}$$

We assume in the following that the singular continuous spectra of H^0 and H are empty. Let $\mathscr{F}^0$ be the generalized Fourier transform associated to the direct intergral decomposition of $\mathscr{H}$ wrt H^0:

$$\mathscr{F}^0: \mathscr{H} \to \mathscr{H}^0 = \mathscr{H}^0_s \oplus \int_{\oplus\Lambda} \mathscr{H}(\lambda)\, d\mu(\lambda), \tag{3.4}$$

where $\Lambda \equiv \sigma_{ac}(H^0)$ and where $\mathscr{H}^0_s$ is the transform of the singular subspace of H^0, which will be of no interset to us in the following. There are two choices for the generalized Fourier transform associated to H, which are relevant to scattering theory and which we denote by $\mathscr{F}^{(+)}$ and $\mathscr{F}^{(-)}$, respectively. They are chosen in some definitive way and equal to each other on the singular subspace $\mathscr{H}_s$ of H, whereas on $\mathscr{H}_{ac}$ they satisfy

$$\mathscr{F}^{(\pm)}P = \mathscr{F}^0 P^0 W^*_{\pm}, \tag{3.5}$$

so that

$$\mathscr{F}^{(\pm)}: \mathscr{H} \to \mathscr{H}^{\pm} = \tilde{\mathscr{H}} \oplus \int_{\oplus\Lambda} \mathscr{H}(\lambda)\, d\mu(\lambda). \tag{3.6}$$

The physical meaning of $\mathscr{F}^{(+)}$ and $\mathscr{F}^{(-)}$ is well known[33]: it follows from (3.5) and (3.2) (iii) that the restrictions of $\mathscr{F}^{(+)}$ and $\mathscr{F}^{(-)}$ to $\mathscr{H}_{ac}$ correspond respectively to the incoming and outgoing state expansions of the formal theory of scattering. In particular, there is a function $\lambda \to S(\lambda) = \{S^{ij}(\lambda)\}$ defined μ-a.e. on $\Lambda \equiv \sigma_{ac}(H)$ and with $S(\lambda)$ a unitary operator in $\mathscr{H}(\lambda)$, such that

$$[\mathscr{F}^{(-)}(Ph)]^i(\lambda) = \sum_{j=1}^{N(\lambda)} S^{ij}(\lambda)[\mathscr{F}^{(+)}(Ph)]^j(\lambda) \quad \forall h\in\mathscr{H}, \tag{3.7}$$

where $N(\lambda) = \dim\mathscr{H}(\lambda)$. $S^{ij}(\lambda)$ is the usual expression of the S matrix in the H^0 representation. Indeed, let $g, h \in \mathscr{H}_{ac}$ and set $g_{(f)} = W^*_{-}g, h_{(i)} = W^*_{+}h$ [(i) and (f) stand respectively for initial and final]. Then from (3.3), (3.7), and (3.5) one has

$$\begin{aligned}(g_{(f)}, Sh_{(i)}) &= (g,h) \\ &= \int_\Lambda d\mu(\lambda) \sum_i \overline{[\mathscr{F}^{(-)}(g)]^i(\lambda)}\,[\mathscr{F}^{(-)}(h)]^i(\lambda) \\ &= \int_\Lambda d\mu(\lambda) \sum_{i,j} \overline{[\mathscr{F}^{(-)}(g)]^i(\lambda)}\, S^{ij}(\lambda)[\mathscr{F}^{(+)}(h)]^j(\lambda) \\ &= \int_\Lambda d\mu(\lambda) \sum_{i,j} \overline{[\mathscr{F}^0(g_{(f)})]^i(\lambda)}\, S^{ij}(\lambda)[\mathscr{F}^0(h_{(i)})]^j(\lambda).\end{aligned} \tag{3.8}$$

For our purposes we must now assume that a single nuclear space Φ exists such that $\mathscr{F}^0$, $\mathscr{F}^{(+)}$, and $\mathscr{F}^{(-)}$ are implemented via Theorem 2.1. We denote by G^0_n, F^0_n, G_n, F_n (omitting for simplicity the eigenvalue and degeneracy lables) the corresponding eigenvectors which pertain to the singular spectra of H^0 and H and by $G^{0i}(\lambda)$, $F^{0i}(\lambda)$, $G^{(\pm)i}(\lambda)$, $F^{(\pm)i}(\lambda)$ the eigenvectors relative to the absolutely continuous spectrum. According to the commonly employed terminology in scattering theory, we refer to the $G^{(\pm)i}(\lambda)$, $F^{(\pm)i}(\lambda)$ as "incoming" and "outgoing" eigenvectors (stationary scattering states), respectively. In practice, Φ will be constructed in general as a space of wavefunctions in the H^0-representation which have suitable "smoothness" properties, in such a way that the implementation of $\mathscr{F}^0$ is immediate. Then one is confronted with the problem of finding "in" and "out" eigenfunctions of H in Φ' which implement $\mathscr{F}^{(+)}$ and $\mathscr{F}^{(-)}$. To this purpose, one could for instance verify that such eigenfunctions meet the conditions of Proposition 2.5. The following proposition shows that checking these conditions for either a set of "in" or of "out" eigenfunctions is sufficient (obviously, whenever it is appropriate, the assertions in the statement and proof of the proposition must be understood as valid μ-a.e.).

Proposition 3.1: Suppose that a nuclear space Φ exists such that there exists in Φ' a set of left eigenvectors of H, $G^{(+)j}_n$ and $G^{(+)j}(\lambda)$ which implements $\mathscr{F}^{(+)}$ via Proposition 2.5. Then there exists in Φ' another family of left eigenvectors of H, $G^{(-)j}_n$ and $G^{(-)i}(\lambda)$, defined on $\sigma(H)$, which implements $\mathscr{F}^{(-)}$. One has

$$\text{(i)} \ G^{(-)i}(\lambda) = \sum_{j=1}^{N(\lambda)} S^{ij}(\lambda) G^{(+)j}(\lambda), \quad \text{(ii)} \ G^{(-)l}_n = G^{(+)l}_n, \tag{3.9}$$

where $\{S^{ij}(\lambda)\}$ is the unitary matrix appearing in (3.7).

Proof: That we must set $G^{(-)l}_n = G^{(+)l}_n$ is a consequence of the definition of $\mathscr{F}^{(-)}$ and $\mathscr{F}^{(+)}$. If $\varphi\in\Phi$ we have from (3.7) and (2.17)

$$[\mathscr{F}^{(-)}(P\varphi)]^i(\lambda) = \sum_{j=1}^{N(\lambda)} S^{ij}(\lambda)\langle G^{(+)j}(\lambda)|\varphi\rangle. \tag{3.10}$$

By construction

$$\left[\sum_{j=1}^{N(\lambda)} \langle G^{(+)j}(\lambda)|\varphi\rangle|^2\right]^{1/2} = \|[\mathscr{F}^{(+)}(P\varphi)](\lambda)\|_\lambda,$$

where $\|\cdot\|_\lambda$ denotes the norm in $\mathscr{H}(\lambda)$. Therefore, by Schwarz inequality,

$$\begin{aligned}&\sum_{j=1}^{N(\lambda)} |S^{ij}(\lambda)\langle G^{(+)j}(\lambda)|\varphi\rangle| \\ &\quad \leqslant \left[\sum_{j=1}^{N(\lambda)} |S^{ij}(\lambda)|^2\right]^{1/2} \left[\sum_{j=1}^{N(\lambda)} |\langle G^{(+)j}(\lambda)|\varphi\rangle|^2\right]^{1/2} \\ &\quad = \|[\mathscr{F}^{(+)}(P\varphi)](\lambda)\|_\lambda,\end{aligned}$$

so that the series $\Sigma_{j=1}^{N(\lambda)} S^{ij}(\lambda)\langle G^{(+)j}(\lambda)|\varphi\rangle$ is absolutely convergent $\forall\varphi\in\Phi$. Then, weak completeness of Φ' ensures that $\Sigma_{j=1}^{N(\lambda)} S^{ij}(\lambda) G^{(+)j}(\lambda)$ converges to an element $G^{(-)i}(\lambda)$ of Φ'. We have

$$\langle G^{(-)i}(\lambda)|H\varphi\rangle$$

$$= \sum_{j=1}^{N(\lambda)} S^{ij}(\lambda)\langle G^{(+)j}(\lambda)|H\varphi\rangle$$
$$= \lambda \langle G^{(-)i}(\lambda)|\varphi\rangle, \quad \forall\varphi\in\Phi,$$

so that $G^{(-)i}(\lambda)$ is a family of left eigenvectors of H. Also, $\forall\varphi, \psi\in\Phi$,

$$(\varphi,\psi) = \sum_{n,l} \overline{\langle G_n^{(+)l}|\varphi\rangle} \langle G_n^{(+)l}|\psi\rangle + \int_\Lambda \sum_i \overline{\langle G^{(+)i}(\lambda)|\varphi\rangle} \langle G^{(+)i}(\lambda)|\psi\rangle \, d\mu_{ac}(\lambda)$$
$$= \sum_{n,l} \overline{\langle G_n^{(-)l}|\varphi\rangle} \langle G_n^{(-)l}|\psi\rangle + \int_\Lambda \sum_{i,j,k} \overline{(S^{-1}(\lambda))^{ij} \langle G^{(-)j}(\lambda)|\varphi\rangle} \times (S^{-1}(\lambda))^{ik} \langle G^{(-)k}(\lambda)|\psi\rangle \, d\mu_{ac}(\lambda)$$
$$= \sum_{n,l} \langle G_n^{(-)l}(\lambda)|\varphi\rangle \langle G_n^{(-)l}(\lambda)|\psi\rangle + \int_\Lambda \sum_j \overline{\langle G^{(-)j}(\lambda)|\varphi\rangle} \langle G^{(-)j}(\lambda)|\psi\rangle \, d\mu_{ac}(\lambda),$$

so that the family of left eigenvectors (3.9) satisfies (b) of Proposition (2.5). Moreover, let $h(\lambda)\in\Sigma_{i\oplus} L^2(\Lambda,\mu_{ac})$ be such that $\int_\Lambda \Sigma_i h^i(\lambda)\langle G^{(-)i}(\lambda)|\varphi\rangle \, d\mu_{ac}(\lambda) = 0$, $\forall\varphi\in\Phi$. Then, $\int_\Lambda \Sigma_i S^{ij}(\lambda) h^i(\lambda)\langle \overline{G^{(+)j}(\lambda)|\varphi}\rangle \, d\mu_{ac}(\lambda) = 0$. By hypothesis, this implies $\Sigma_j S^{ij}(\lambda) h^i(\lambda) = 0$ or $h(\lambda) = 0$. Hence (3.9) also satisfied condition (c) of Proposition (2.5) where, by (3.10), $\mathscr{F} = \mathscr{F}^{(-)}$ (since $\mathscr{F}^{(-)}$ is a generalized Fourier transform, condition (d) of the proposition obviously holds).

Remark: (i) It is not assumed and it is neither true in general that $P^0\Phi\subset\Phi$ and/or that $P\Phi\subset\Phi$ nor that $W_\pm P^0\Phi\subset P\Phi$ and/or that $SP^0\Phi\subset P^0\Phi$. However, on the basis of formulas (2.4), (2.5), and (3.5) we can still define the transpose $W'_\pm$ and the extension $\widehat{W}_\pm$ of $W_\pm$ restricted respectively to the linear span of the $G^{(\pm)i}(\lambda)$ and of the $F^{0j}(\lambda)$ according to

$$W'_\pm G^{(\pm)i}(\lambda) = G^{0i}(\lambda), \tag{3.11}$$

and

$$\widehat{W}_\pm F^{0j}(\lambda) = F^{(\pm)j}(\lambda). \tag{3.12}$$

(ii) it is perhaps worthwhile to recall that in actual problems one works with some explicit representation of the space and operators treated here. For example, in the commonly used $\mathscr{F}^0$ representation, in which the noninteracting Hamiltonian is diagonal, vectors in $\mathscr{H}$ (respectively, in Φ) are represented by functions $[\mathscr{F}^0(h)]^i(\lambda)$ in $\mathring{\mathscr{H}} = \mathscr{F}^0(\mathscr{H})$ [respectively, $\langle G_n^{0j}|\varphi\rangle$, $\langle G^{0i}(\lambda)|\varphi\rangle$ in $\mathring{\Phi} = \mathscr{F}^0(\Phi)$], defined on $\sigma(H^0)$. Then, elements B in Φ' are "concretely expressed" by elements $\mathring{B}$ in $\mathring{\Phi}'$. For instance, suppose the spectrum of H^0 is simple and absolutely continuous. Then $\mathring{G}^0(\lambda) = \delta_\lambda$, the ordinary Dirac delta function.

We are now in position to outline the application of the methods of analytic continuation of eigenvectors developed in Sec. 2 to the characterization of resonant states. Depending on whether one is interested in the transition amplitude between two fully interacting states (as for example in the description of the decay of an unstable system) or in a scattering experiment, the two relevant formulas are obtained from (3.8) by choosing respectively $g,h\in P\Phi$ or $g_{(f)}, h_{(i)}\in P^0\Phi$. It is convenient to reexpress the integrals in terms of the Lebesgue measure by changing the normalization of the eigenvectors upon multiplying them by the square root of the Radon–Nikodym derivative $\sqrt{d\mu_{ac}(\lambda)/d\lambda}$. Thus, assuming the existence of a conjugation as in Sec. 2 and using (2.17), we have:

$$(P\varphi, P\psi) = \int_\Lambda d\lambda \sum_i \langle\overline{\varphi}|F^{(-)i}(\lambda)\rangle\langle G^{(-)i}(\lambda)|\psi\rangle = \int_\Lambda d\lambda \sum_{i,j} \langle\overline{\varphi}|F^{(-)i}(\lambda)\rangle S^{ij}(\lambda)\langle G^{(+)j}(\lambda)|\psi\rangle \quad \forall\varphi, \psi\in\Phi, \tag{3.13}$$

and

$$(P^0\varphi_{(f)}, SP^0\psi_{(i)}) = \int_\Lambda d\lambda \sum_{ij} \langle \overline{\varphi_{(f)}}|F^{0i}(\lambda)\rangle S^{ij}(\lambda)\langle G^{0j}(\lambda)|\psi_{(i)}\rangle \quad \forall\varphi_{(f)}, \psi_{(i)}\in\Phi. \tag{3.14}$$

These two formulas, equivalent in Hilbert space notation, are quite different in the nuclear space approach, in that *either* the fully interacting states at finite time are required to be in $P\Phi$ *or* the asymptotic free states at $t = \pm\infty$ are required to be in $P^0\Phi$. In particular, when Φ is such as to allow for analyticity properties of the eigenvectors of H^0 and H, it is in general impossible to meet these two requirements simultaneously [compare remark (i) above and the examples of Secs. 4 and 5].

We now assume that: (i) the free eigenvetors $G^{0i}(\lambda)$ and $F^{0i}(\lambda)$ are the restriction to Λ of families $G^{0i}(z)$ and $F^{0i}(z)$ analtic in a region Ω_0, with $\Omega_0\cap\Lambda$ a subset of $\mathbb{R}$ with nonempty interior $(\Omega_0\cap\Lambda)^0$; (ii) the left incoming eigenvectors $G^{(+)i}(\lambda)$ and the right outgoing eigenvectors $F^{(-)i}(\lambda)$ are boundary values respectively of families $G^{(+)i}(z)$ and $F^{(-)i}(z)$ analytic in a domain Ω_1 in the lower half-plane, with $\Omega_1^c\cap\Lambda$ a subset of $\mathbb{R}$ with nonempty interior $(\Omega_1^c\cap\Lambda)^0$. Then the left outgoing eigenvectors $G^{(-)i}(\lambda)$ are boundary values of the family $G^{(-)i}(z) = \overline{F}^{(-)i}(\overline{z})$, analytic in $\overline{\Omega}_1$, and they are related to the $G^{(+)i}(\lambda)$ by Eq. (3.9)-(i). Now, in many cases the S matrix in the H^0 representation, $S^{ij}(\lambda)$, is the boundary value from the upper half-plane of an analytic function which can be continued across $\Omega_1^c\cap\Lambda$ into its second Riemann sheet in the lower half-plane to an analytic function $S_{\text{II}}^{ij}(z)$. Then, provided the convergence of the series (3.9) (i) preserves analyticity, the family $G^{(-)i}(z)$ can be continued analytically across $(\Omega_1^c\cap\Lambda)^0$ to some domain Ω_2 included in Ω_1 to a family

$$G_{\text{II}}^{(-)i}(z) = \sum_j S_{\text{II}}^{ij}(z) G^{(+)j}(z). \tag{3.15}$$

It analyticity properties in this domain are controlled by those of $S_{\text{II}}^{ij}(z)$. Therefore, the singularities of $S_{\text{II}}^{ij}(z)$ in Ω_1 in the second Riemann sheet appear as corresponding singularities of the continuation of the left outgoing eigenvectors of H. In particular, a resonance associated to a second sheet simple pole $z_0 = E_0 - i(\Gamma_0/2)$ of the S matrix close to the real axis in the lower half-plane, appears as a corresponding pole of the family $G_{\text{II}}^{(-)i}(z)$, for some value of the index i. Now, it follows from Proposition 2.4 that the first coefficient $D^{(-)i}(z_0)$ of the Laurent expansion of $G_{\text{II}}^{(-)i}(z)$ about z_0 is a left eigenvector of the interacting Hamiltonian, with z_0 as eigenvalue. Therefore, provided there is a nuclear space Φ which meets the required conditions for H and H^0, *to a resonance at $z_0 = E_0 - i(\Gamma_0/2)$ is associated a set of left "outgoing" eigenvectors $D^{(-)i}(z_0)$ [and the corresponding right "outgoing" eigenvectors $F^{(-)i}(z_0)$] of the interacting Hamiltonian, with the complex eigenvalue $z_0 = E - i(\Gamma_0/2)$*. Such objects have the same mathematical meaning of plane waves and of the stationary scattering states, namely, that of distributions on a suitable dense subspace of the Hilbert space. From a physical point of view, one requires that such eigenvectors be approximable in some sense by means of vectors in $\mathscr{H}$. And this is always true, because Φ and $\mathscr{H}$ are dense in Φ' endowed with the weak topology (see also Sec. 6).

Alternatively, starting from the relation

$$F^{(+)j}(\lambda) = \sum_j F^{(-)i}(\lambda) S^{ij}(\lambda), \qquad (3.16)$$

we could continue the family $F^{(+)j}(z) = \overline{G}^{(+)i}(\bar{z})$, analytic in $\bar{\Omega}_1$, across $\Omega_1^c \cap \Lambda$ into the lower half-plane to give a second sheet continuation $F_{\rm II}^{(+)j}(z) = \Sigma_i F^{(-)i}(z) S_{\rm II}^{ij}(z)$ of the right incoming eigenvectors, whose analyticity properties are again determined by those of $S_{\rm II}^{ij}(z)$. In this way, we could equally well associate to the resonances at z_0 a set of right "incoming" eigenvectors $C^{(+)j}(z_0)$ for some j [with corresponding left " incoming" eigenvectors $G^{(+)j}(z_0)$], where $C^{(+)j}(z_0)$ is here the first coefficient of the Laurent expansion about z_0 of $F_{\rm II}^{(+)j}(z)$.

Because of (3.9)-(i) and (3.16), the relations among the "in" and "out" stationary states associated to a resonance are

$$D^{(-)i}(z_0) = \sum_j \operatorname{Res} S_{\rm II}^{ij}(z) \Big|_{z=z_0} G^{(+)j}(z_0), \qquad (3.17)$$

and

$$C^{(+)j}(z_0) = \sum_i F^{(-)i}(z_0) \operatorname{Res} S_{\rm II}^{ij}(z) \big|_{z=z_0}. \qquad (3.18)$$

The association of a resonance at z_0 with a set of eigenvectors of H with z_0 as eigenvalue is made especially clear when one considers survival amplitudes and scattering amplitudes. Recalling (2.17) and (2.18), the survival amplitude of a state $h = P\varphi$, where $\varphi \in \Phi$, is given by

$$\begin{aligned}(h,&\exp(-iHt)h)\\ &= (P\varphi, \exp(-iHt)P\varphi)\\ &= \int_\Lambda d\lambda \exp(-i\lambda t) \sum_i \langle \bar{\varphi} | F^{(-)i}(\lambda)\rangle \langle G^{(-)i}(\lambda)|\varphi\rangle.\end{aligned} \qquad (3.19)$$

With the notations of Sec. 2, let $\Lambda_0 \subset \Omega_1^c \cap \Lambda$ and let Λ_1 be a curve in the lower half-plane such that $\Lambda_0 \cup \Lambda_1$ is a scroc and $\Lambda_0 \cup \Lambda_1 \subset \Omega_2^c$, with $G_{\rm II}^{(-)i}(z)$ assumed to be continuous on $\Lambda_0 \cup \Lambda_1$. Set $\Gamma = (\Lambda \setminus \Lambda_0) \cup \Lambda_1$ and assume that $G_{\rm II}^{(-)i}(z)$ is analytic in $\operatorname{Int}(\Lambda_0 \cup \Lambda_1)$ except for a simple pole at $z_0 = E_0 - i(\Gamma_0/2)$, close to the real axis. For simplicity, we exclude any other kind of singularity and, in particular, second sheet cuts (complex thresholds), which could be easily included in

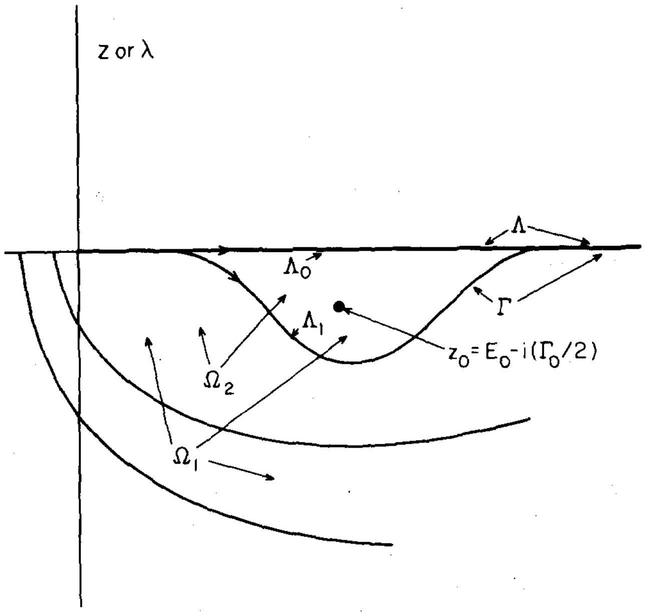


FIG. 2. The domains Ω_1 and Ω_2, and the deformation of the integration path from the spectrum Λ to the path Γ in the lower half plane. The singularity of the eigenvector family $G_{\rm II}^{(-)i}(z)$ at z_0 gives rise to a pure exponential contribution in (3.20).

the background. We can deform the integration contour in the lower half-plane and obtain[21] (see Fig. 2)

$$\begin{aligned}(h,&\exp(-iHt)h)\\ &= (P\varphi, \exp(-iHt)P\varphi)\\ &= \int_\Gamma dz \exp(-izt) \sum_i \langle \bar{\varphi} | F^{(-)i}(z)\rangle \langle G_{\rm II}^{(-)i}(z)|\varphi\rangle\\ &\quad - 2\pi i \exp(-iE_0 t)\\ &\quad \times \exp[-(\Gamma_0/2)t] \sum_i \langle \bar{\varphi} | F^{(-)i}(z_0)\rangle \langle D^{(-)i}(z_0)|\varphi\rangle.\end{aligned} \qquad (3.20)$$

If h is a wave packet which is strongly peaked at the resonant energy E_0 for those values of the index i for which $D^{(-)i}(z_0) \neq 0$, the amplitude is dominated over a large range of time values by the exponential term at the rhs of (3.20), the background integral along Γ being mainly responsible for the deviations from the exponential decay law at short and long times compared to the mean life $1/\Gamma_0$.[1,21] And we see that the purely exponential contribution to the survival amplitude in the "completeness expansion" (3.20) is precisely given by the "discrete" set of complex energy eigenvectors $\{D^{(-)i}(z_0)\}$ whereas the "continuum" $\{G^{(-)i}(z)\}_{z\in\Gamma}$ is responsible for the background corrections.[34] Similarly, provided the convergence of the series $\Sigma_j S^{ij}(\lambda) G^{0j}(\lambda)$ preserves analyticity, the series itself can be analytically continued in the second Riemann sheet in some domain Ω_3 in the lower half-plane, included in Ω_0, and it provides therein a family $G_{\rm II}^{0i}(z) = \Sigma_j S_{\rm II}^{ij}(z) G^{0j}(z)$ of left eigenvectors of H^0. Then we obtain for the scattering amplitude a formula analogous to (3.20) by deformation in the lower half-plane of the integration contour in (3.14) (of course, here the path Γ is not necessarily the same as before):

$$\begin{aligned}(P^0&\varphi_{(f)}, SP^0\psi_{(i)})\\ &= \int_\Gamma dz \sum_{i,j} \langle \bar{\varphi}_{(f)} | F^{0i}(z)\rangle S_{\rm II}^{ij}(z) \langle G^{0j}(z)|\psi_{(i)}\rangle\end{aligned}$$

$$-2\pi i\sum_{i,j}\langle\bar{\varphi}_{(f)}|F^{0i}(z_0)\rangle \mathrm{Res} S^{ij}_{\mathrm{II}}(z)\big|_{z=z_0}\langle G^{0j}(z_0)|\psi_{(i)}\rangle, \tag{3.21}$$

where we have confined ourselves once more to the case when $S^{ij}_{\mathrm{II}}(z)$ is analytic in the interior of $\Gamma\cup\Lambda$ except for a simple pole at $z=z_0$. We remark that relations (3.11) and (3.12) could be generalized in an obvious way by extending the definition of $W'_\pm$ and of $\widehat{W}_\pm$ to complex values of λ in the domain of analyticity of the eigenvectors, as well as to the resonance points.

The situation is slightly more involved in the case of a multiple resonance, when some of the coefficients $S^{ij}_{\mathrm{II}}(z)$ have at $z_0 = E_0 - i(\Gamma_0/2)$ a pole of order $N>1$. In this case, it is well known and a simple exercise to check that the dominant (pole) part of the decay amplitude of the resonant state is of the form $\exp(-iE_0t)\exp[-(\Gamma_0/2)t]P_{N-1}(t)$, where P_{N-1} is an $(N-1)$th order polynomial. This can again be interpreted in our formalism as a "discrete" contribution in an expansion in terms of a "complex completeness" of the Hamiltonian, though in this case associated vectors play a role too [compare the discussion following formulas (2.19)–(2.21)]. Explicitly, expand about z_0 the right and left analytically continued eigenvectors:

$$F^{(-)i}(z) = \sum_{n=0}^{\infty} C^{(-)i}_n(z_0)(z-z_0)^n, \tag{3.22}$$

$$G^{(-)i}_{\mathrm{II}}(z) = \sum_{n=-N}^{\infty} D^{(-)i}_n(z_0)(z-z_0)^n. \tag{3.23}$$

Then, the formula for the survival amplitude which generalizes (3.20) is the following:

$$\begin{aligned}(P\varphi,&\exp(-iHt)P\varphi)\\ &=\int_\Lambda dz\,\exp(-izt)\sum_i\langle\bar{\varphi}|F^{(-)i}(z)\rangle\langle G^{(-)i}_{\mathrm{II}}(z)|\varphi\rangle\\ &\quad-2\pi i\exp(-iE_0t)\exp[-(\Gamma_0/2)t]\sum_{r=0}^{N-1}\frac{(-i)^rt^r}{r!}\\ &\quad\times\sum_i\sum_{l=0}^{N-1-r}\langle\bar{\varphi}|C^{(-)i}_{N-1-l-r}(z_0)\rangle\langle D^{(-)i}_{-N+l}(z_0)|\varphi\rangle,\end{aligned} \tag{3.24}$$

and we may describe the resonance by the two sets $\{D^{(-)i}_\alpha(z_0)\}_{\alpha=-N,\dots,-1}$ and $\{C^{(-)i}_\beta(z_0)\}_{\beta=0,\dots,N-1}$ of left and right eigenvectors and associated vectors of H.

Next we consider the resolvents of H^0 and H. With the respective choices of the families $F^{0i}(z)$, $G^{0i}(z)$ and $F^{(-)j}(z)$, $G^{(-)j}(z)$, we see that formula (2.27) applies. This allows us to continue analytically $R^0(z)$ and $R(z)$, as operators in $\mathscr{L}(\Phi,\Phi')$ from the upper half-plane respectively across $(\Omega_0\cap\Omega)^0\setminus\sigma_s(H^0)$ and $(\Omega^c_1\cap\Lambda)^0\setminus\sigma_s(H)$ into the lower half-plane, seen as part of the second Riemann sheet, to the respective regions of analyticity of the corresponding eigenvectors. The respective continuations are

$$R^0_{\mathrm{II}}(z) = R^0(z) + 2\pi i\sum_i|F^{0i}(z)\rangle\langle G^{0i}(z)|, \tag{3.25}$$

and

$$\begin{aligned}R_{\mathrm{II}}(z) &= R(z) + 2\pi i\sum_i|F^{(-)i}(z)\rangle\langle G^{(-)i}_{\mathrm{II}}(z)|\\ &= R(z) + 2\pi i\sum_{i,j}|F^{(-)i}(z)\rangle S^{ij}(z)\langle G^{(+)j}(z)|.\end{aligned} \tag{3.26}$$

Therefore, the singularities of $R_{\mathrm{II}}(z)$ are exactly the same as those of the continued S matrix (apart from the pathological situations in which poles of $S^{ij}(z)$ might be cancelled by zeros of the eigenvectors), whereas $R^0_{\mathrm{II}}(z)$ is regular at these singularities.[35]

We can sum up the above by saying that the eigenvalues and eigenvectors associated with the poles of the continuation into the second Riemann sheet of the analytic family of left *outgoing* eigenvectors can be interpreted as describing the location of resonances and the actual "stationary" resonant states.

4. THE FRIEDRICHS MODEL

In this and next section we show that the theory set down in Sec. 2 and 3 is not empty by applying it to the example of the Friedrichs model[6] (present section) and to the scattering of a spinless particle by a local central potential (Sec. 5), with suitable assumptions on the corresponding potentials. For pedagogical reasons, and due to its solvability, we shall treat the Friedrichs model in greater detail, whereas in the example of the central potential we shall only state the results, leaving the proofs to the reader. In this model, the Hilbert space $\mathscr{H}$ is taken to be $\mathbb{C}\oplus L^2(0,\infty)$, so that a vector in $\mathscr{H}$ is a pair $\tilde{h}=(h^0,h(\lambda))$, where $h^0\in\mathbb{C}$ and $h(\lambda)\in L^2(0,\infty)\equiv L^2_+$. The free and interacting Hamiltonians are respectively

$$H^0=\begin{pmatrix}M & 0\\ 0 & x\cdot\end{pmatrix}\quad\text{and}\quad H=\begin{pmatrix}M & \epsilon\bar{f}\\ \epsilon f & x\cdot\end{pmatrix},$$

with $M>0$, ϵ a real number and f some function so that

$$H^0\tilde{h}=:\begin{pmatrix}M & 0\\ 0 & x\cdot\end{pmatrix}\begin{pmatrix}h^0\\ h(x)\end{pmatrix}=\begin{pmatrix}Mh^0\\ xh(x)\end{pmatrix}, \tag{4.1a}$$

and

$$h\tilde{h}=:\begin{pmatrix}M & \epsilon\bar{f}\\ \epsilon f & x\cdot\end{pmatrix}\begin{pmatrix}h^0\\ h(x)\end{pmatrix}=\begin{pmatrix}Mh^0+\epsilon\int_0^\infty\bar{f}(x)h(x)\,dx\\ \epsilon h^0f(x)+xh(x)\end{pmatrix}. \tag{4.1b}$$

Assumptions have to be made on the domains and on f in order to make H^0 and H properly defined self-adjoint operators and we will presently make those which are most suitable for our purposes.

We consider $L^2(0,\infty)$ as embedded into $L^2(\mathbb{R})\equiv L^2$ by means of the projection P_+ according to $(P_+h)(x)=\chi_{[0,\infty)}(x)h(x)$, where χ_Δ is the characteristic function of the Borel set Δ of the real axis. We consider the dense subspace $Z\subset L^2$ of functions which are the restriction to $\mathbb{R}$ of the entire functions $\varphi(z)$ with the property that for each φ there is some positive number b and some sequence $\{B_n\}$ of positive numbers such that $(1+|z|)^n|\varphi(z)|\leqslant B_n\exp(b\,|\mathrm{Im} z|),\forall n$. This space can be endowed with a nuclear topology as follows (Ref. 22, Vols. I, II). First, let $Z(b)$ be the submanifold of the entire functions for which the previous inequality holds with a fixed b. Then, define on $Z(b)$ the countable sequence of norms

$$\|\varphi\|_k = \sup_{z\in\mathbb{C}} (1+|z|)^k |\varphi(z)| \exp(-b|\mathrm{Im} z|). \qquad (4.2)$$

In this way, $Z(b)$ becomes a complete countably normed nuclear space. If $a>b$, $Z(a)\supset Z(b)$, and the topology induced by $Z(a)$ on $Z(b)$ coincides with that of $Z(b)$. We have $Z = \bigcup_{a>0} Z(a)$ and we endow Z with the strict inductive limit topology of the spaces $Z(a)$. Then Z is a nuclear space. The canonical embedding $\mathscr{J}$ of Z into L^2, which associates each φ in Z to the L^2 class of functions which are a.e. equal to φ on $\mathbb{R}$ is continuous. The operator $\mathscr{J}_+ = P_+\mathscr{J}$ is a one-to-one continuous linear map of Z ono a dense submanifold of L^2_+. Therefore, by the embedding $\mathscr{J}_+$, a rigged Hilbert space is implemented: $Z\subset L^2_+ \subset Z'$. We choose as conjugation on L^2_+ the ordinary complex conjugation of functions. It acts on Z according to $\varphi(z)\to\tilde\varphi(z)\equiv\overline{\varphi(\bar z)}$ and is continuous wrt the topology of Z. Then L^2_+ is continuously, densely and linearly embedded into Z' by the map $\mathscr{J}'_+$ is defined by $\langle \mathscr{J}'_+ h|\varphi\rangle = (\bar h, \mathscr{J}_+\varphi) = \int_0^\infty h(x)\varphi(x)\,dx$, $h\in L^2_+$, $\varphi\in Z$. Denote by $\tilde L^2_+$ and $\tilde Z$ respectively the Hilbert direct sum $\mathbb{C}\oplus L^2_+$ and the topological direct sum $\mathbb{C}\oplus Z$. $\tilde Z$ is the nuclear space which is the strict inductive limit of the complete countably normed nuclear spaces $\tilde Z(a) = \mathbb{C}\oplus Z(a)$, where the norms on $\tilde Z(a)$ can be chosen for example as $\|(\varphi^0,\varphi)\|_k = (|\varphi^0|^2 + \|\varphi\|_k^2)^{1/2}$, with $\|\varphi\|_k$ given by (4.2). If $\tilde P_+$ is the projection on $\mathbb{C}\oplus L^2$ with range $\tilde L^2_+$ and if $\tilde{\mathscr{J}}$ defined by $\tilde{\mathscr{J}}(\varphi^0,\varphi) = (\varphi^0,\mathscr{J}\varphi)$ is the canonical embedding of $\tilde Z$ into $\mathbb{C}\oplus L^2$, then $\tilde P_+\tilde{\mathscr{J}} = \tilde{\mathscr{J}}_+$ is a one-to-one linear continuous map of $\tilde Z$ onto a dense submanifold of $\tilde L^2_+$, thereby implementing a rigged Hilbert space $\tilde Z\subset\tilde L^2_+\subset\tilde Z'$. The elements of $\tilde Z'$ are pairs $G=(g_0,g)$, where $g_0\in\mathbb{C}$ and $g\in Z'$. The conjugation on $\tilde L^2_+$ is naturally $(h^0,h)\to(\bar h^0,\bar h)$. It leaves $\tilde Z$ invariant and is continuous wrt the topology of $\tilde Z$. The corresponding linear embedding $\tilde{\mathscr{J}}'_+$ of $\tilde L^2_+$ into $\tilde Z'$ is given by $\langle \tilde{\mathscr{J}}'_+(h^0,h)|(\varphi^0,\varphi)\rangle = h^0\varphi^0 + \int_0^\infty h(\lambda)\varphi(\lambda)\,d\lambda$.

After these preliminaries, we require H^0 and H to have both domain $\tilde Z$ and the potential function f in (4.1) to belong to Z. Then, both H^0 and H leave $\tilde Z$ invariant and act with continuity upon $\tilde Z$ (Appendix A). The action of H' and of $\hat H$ on an element $D=(d_0,d)\in\tilde Z'$ is respectively given by

$$H'(d_0,d) = (Md_0 + \epsilon\langle d|f\rangle, \epsilon d_0\bar f + xd), \qquad (4.3)$$

and by

$$\hat H(d_0,d) = (Md_0 + \epsilon\langle d|\bar f\rangle, \epsilon d_0 f + xd), \qquad (4.4)$$

We have that H^0 and H, as operators in $\tilde L^2_+$ with domain $\tilde Z$, are essentially self-adjoint (Appendix C). An immediate remark is that $\tilde L^2_+$ is a direct integral decomposition associated to H^0, [6] that is to say $\mathscr{F}^0$ is the identity operator. Furthermore, $\tilde Z$ is a possible choice for the nuclear space that implements it via Theorem 2.1. Indeed, the left eigenvectors of H^0 are $G^0=(1,0)$ and $G^0(\lambda)=(0,\delta_\lambda)$, namely, $\langle G^0|\tilde\varphi\rangle = \varphi^0$ and $\langle G^0(\lambda)|\tilde\varphi\rangle = \varphi(\lambda)\,\forall\tilde\varphi\in\tilde Z$ so that Proposition (2.5) trivially applies wrt Lebesgue measure on $[0,\infty)$. Since the δ function is analytic on Z, the eigenvectors $G^0(\lambda)$ relative to the a.c. spectrum are continuous boundary values on $[0,\infty)$ of a family of eigenvectors $G^0(z)=(0,\delta_Z)$ which is analytic everywhere.

In order to study the eigenvectors of H, we introduce the inverse of the partial resolvent[6]

$$(G^0, R(z)G^0)^{-1} = \alpha(z) = z - M - \epsilon^2\int_0^\infty \frac{|f(x)|^2}{z-x}\,dx. \qquad (4.5)$$

This function has a cut on $[0,\infty)$ and is otherwise regular, it has zeros at most on the real axis, and under the conditions that $f(0)=0$, $f(M)\neq 0$ and that the coupling constant ϵ is small enough, neither $\alpha(\lambda)$ for $\lambda\leqslant 0$ nor $\alpha(\lambda\pm i0)$ for $\lambda\geqslant 0$ vanish.[36] We assume these conditions to be satisfied. Then we have the following.

Proposition 4.1: With the above assumptions on f, ϵ, every complex number z is a left eigenvalue of H with associated left eigenvector in $\tilde Z'$ which for $Z\notin[0,\infty)$ is given by:

$$G(z) = \left(\epsilon\frac{f(z)}{\alpha(z)}, \epsilon^2\frac{f(z)}{\alpha(z)}\cdot\frac{\bar f(x)}{z-x} + \delta_z\right), \qquad (4.6)$$

and acts on an element $\tilde\varphi\in\tilde Z$ according to

$$\langle G(z)|\tilde\varphi\rangle = \epsilon\frac{f(z)}{\alpha(z)}\varphi^0 + \epsilon^2\frac{f(z)}{\alpha(z)}\int_0^\infty\frac{\bar f(x)\varphi(x)}{z-x}\,dx + \varphi(z). \qquad (4.7)$$

For $z\equiv\lambda\in[0,\infty)$, the boundary values $G^{(\pm)}(\lambda)$ acting according to

$$\langle G^{(\pm)}(\lambda)|\tilde\varphi\rangle = \epsilon\frac{f(\lambda)}{\alpha(\lambda\mp i0)}\varphi^0 + \epsilon^2\frac{f(\lambda)}{\alpha(\lambda\mp i0)} \times\int_0^\infty\frac{\bar f(x)\varphi(x)}{\lambda\mp i0-x}\,dx + \varphi(\lambda), \qquad (4.8)$$

yield the corresponding left eigenvectors. There are no left eigenvectors in $\tilde L^2_+$. *The above eigenvectors are unique* up to a complex factor possibly depending on z; in particular, the following relation holds:

$$G^{(-)}(\lambda) = \frac{\alpha(\lambda-i0)}{\alpha(\lambda+i0)}G^{(+)}(\lambda). \qquad (4.9)$$

In order to prove that (4.7) and (4.8) yield left eigenvectors of H one might proceed as follows. It is easy to check that Friedrichs formulas[6] for the spectral decompositions $\mathscr{F}^{(\pm)}$ wrt H are expressed by operations which, for fixed $\lambda\geqslant 0$, are continuous linear functionals upon $\tilde Z$ which agree with (4.8):

$$[\mathscr{F}^{(\pm)}(\tilde\varphi)](\lambda) = \langle G^{(\pm)}(\lambda)|\tilde\varphi\rangle, \quad \tilde\varphi\in\tilde Z. \qquad (4.10)$$

Next, one checks that such functionals are indeed analytically continuable to the whole complex plane, so that Proposition 2.2 can be applied. However, in order to prove uniqueness, we need to solve the eigenvalue equation explicitly, and we do so in Appendix B.

Corollary: When the above conditions on f and ϵ are satisfied, to every complex number z there is associated a right eigenvector which, for $z\notin[0,\infty)$, acts on the elements of $\tilde Z$ according to

$$\langle\tilde\varphi|F(z)\rangle = \epsilon\frac{\bar f(z)}{\alpha(z)}\varphi^0 + \epsilon^2\frac{\bar f(z)}{\alpha(z)}\int_0^\infty\frac{f(x)\varphi(x)}{z-x}\,dx + \varphi(z), \qquad (4.11)$$

for $z\equiv\lambda\in[0,\infty)$, the boundary values $F^{(\pm)}(\lambda)$ acting according to

$$\langle\tilde\varphi|F^{(\pm)}(\lambda)\rangle = \epsilon\frac{\bar f(\lambda)}{\alpha(\lambda\pm i0)}\varphi^0 + \epsilon^2\frac{\bar f(\lambda)}{\alpha(\lambda\pm i0)} \times\int_0^\infty\frac{f(x)\varphi(x)}{\lambda\pm i0-x}\,dx + \varphi(\lambda), \qquad (4.12)$$

are the corresponding eigenvectors. There are no right eigenvectors in $\tilde{L}^2_+$. The above eigenvectors are unique up to a complex factor possibly depending on z; in particular,

$$F^{(-)}(\lambda) = \frac{\alpha(\lambda + i0)}{\alpha(\lambda - i0)} F^{(+)}(\lambda), \tag{4.13}$$

holds. To prove the corollary, recall that $F(z) = \bar{G}(\bar{z})$.

Remarks: (i) The families of eigenvectors $G(z)$ and $F(z)$ are analytic in the plane with a cut on $[0,\infty)$. Their boundary values on the cut, $G^{(\pm)}(\lambda)$ and $F^{(\pm)}(\lambda)$, $\lambda \geqslant 0$, are continuous. (ii) In particular, $G^{(+)}(\lambda)$ and $F^{(-)}(\lambda)$, $\lambda \geqslant 0$, are boundary values *from below* respectively of the families $G(z)$ and $F(z)$, analytic in the lower half-plane. On the other hand, in this model, $G^{(-)}(\lambda)$ and $F^{(+)}(\lambda)$ are boundary values *from above* respectively of the *same* families $G(z)$ and $F(z)$, so that we can drop the superscripts $(\pm)$ whenever we are not dealing with the boundary values. (iii) The families of eigenvectors $F^{(\pm)}(\lambda)$ and $G^{(\pm)}(\lambda)$ do satisfy (a), (b), and (c) of Proposition 2.5. This need not be shown directly: it is enough to recall (4.10), which trivially implies that $\tilde{Z}$, $G^{(\pm)}(\lambda)$ and $F^{(\pm)}(\lambda)$ implement Theorem 2.1 wrt H and $\mathscr{F}^{(\pm)}$. Regarding condition (d) of the proposition, it follows trivially from the fact that $\tilde{Z}$ is already a domain of essential self-adjointness of H. In particular, the spectrum of H is $[0,\infty)$ simple and the spectral measure is equivalent to the restriction of the Lebesgue measure to $[0,\infty)$. As a matter of fact, the eigenvectors (4.8) and (4.12) are so normalized that the associated measure in (2.16) is Lebesgue. (iv) The superscript $(\pm)$ in (4.8) and (4.12) does not agree with the limiting procedures. Instead, it is apparent from (4.10) and Ref. 6 that the $G^{(+)}(\lambda)$, $F^{(+)}(\lambda)$ and the $G^{(-)}(\lambda)$, $F^{(-)}(\lambda)$ are respectively the "in" and "out" states of the formal theory of scattering. (v) The projections P^0 and P on the a.c. subspaces of H^0 and H respectively are trivially given by $P^0\tilde{h} = (0,h)$ and $P = 1$. It follows that $\tilde{Z}$ is invariant wrt both projections and their action is continuous wrt the nuclear topology of $\tilde{Z}$. This particular feature simplifies formulas (3.13), (3.14), (3.20), (3.21), and (3.24) as we can drop P and take $\varphi_{(f)}, \psi_{(i)} \in P^0\Phi$.

The families $G(z)$ and $F(z)$ can be analytically continued across the cut $[0,\infty)$ and by Proposition 2.2 the continuations preserve their character of being eigenvectors of H. We shall be interested in the continuation of the family $G(z)$ from the first quadrant across the cut into the fourth quadrant of the complex plane, seen as part of the second Riemannian sheet. The continuation of $\alpha(z)$ is

$$\alpha_{\rm II}(z) = z - M - \epsilon^2 \int_0^\infty \frac{|f(x)|^2}{z-x}\,dx + 2\pi i\epsilon^2 \bar{f}(z) f(z)$$
$$= \alpha(z) + 2\pi i \epsilon^2 \bar{f}(z) f(z), \tag{4.14}$$

and for all $\tilde{\varphi}$ in $\tilde{Z}$, the continuation of $\int_0^\infty [\bar{f}(x)\varphi(x)/(z-x)]\,dx$ is $\int_0^\infty [\bar{f}(x)\varphi(x)/(z-x)]\,dx - 2\pi i \bar{f}(z)\varphi(z)$. Therefore, the continuation $G_{\rm II}(z)$ of $G(z)$ is

$$\langle G_{\rm II}(z)|\tilde{\varphi}\rangle$$
$$= \epsilon \frac{f(z)}{\alpha_{\rm II}(z)}\varphi^0 + \epsilon^2 \frac{f(z)}{\alpha_{\rm II}(z)} \int_0^\infty dx \frac{\bar{f}(x)\varphi(x)}{z-x}$$
$$+ \frac{\alpha(z)}{\alpha_{\rm II}(z)}\varphi(z) = \frac{\alpha(z)}{\alpha_{\rm II}(z)} \langle G(z)|\tilde{\varphi}\rangle. \tag{4.15}$$

The last equality, which could have been inferred from (4.9), is not surprising, because of the uniqueness of the left eigenvectors up to a z-dependent factor.

Remarks: (i) $\alpha(\lambda - i0)/\alpha(\lambda + i0) = S(\lambda)$ is the S matrix in the H^0 representation.[6] With the assumptions made for f and ϵ, it can be continued across the cut to the function $S_{\rm II}(z) = \alpha(z)/\alpha_{\rm II}(z)$, which is meromorphic in the cut plane seen as the second Riemannian sheet (see below). (ii) It is not a surprising fact that in $\tilde{Z}'$, together with "in" eigenvectors $G^{(+)}(\lambda)$, $F^{(+)}(\lambda)$, there must appear the "out" ones $G^{(-)}(\lambda)$, because of the preceding remark and of Proposition 3.1. (iii) In this example, it is easily seen that it is not true that the nuclear space, $\tilde{Z}$, is invariant wrt the actions of W_+, W_-, and S.

Whereas $G(z)$ is regular everywhere in the cut plane, $G_{\rm II}(z)$ is not, in general. Indeed, $\alpha(z)$ has no zeros whereas $\alpha_{\rm II}(z)$, while being regular in the cut plane, may vanish. Therefore, it follows from (4.15) that the only possible singularities of the family $G_{\rm II}(z)$ are poles which are in one-to-one correspondence with the zeros of the $\alpha_{\rm II}(z)$. With the assumptions made for f and ϵ, in any strip of the second Riemann sheet with $|{\rm Im}\, z| \leqslant c$, c fixed but otherwise arbitrary, there is at most a finite number of zeros of $\alpha_{\rm II}(z)$. This follows easily from the continuity of $\alpha(z)$ and $\alpha_{\rm II}(z)$ at $z = 0$, and the fact that $\alpha_{\rm II}(0) = \alpha(0)$. Since f is an entire function, the existence of a unique zero of $\alpha_{\rm II}(z)$ in any given half-disk, $\{z \,|\, {\rm Im}\, z < 0, |z - M| < \delta\}$, provided ϵ^2 is small enough, is ensured by Theorem 2.1 of Ref. 12b, and the position of the zero is an analytic function of ϵ^2.

We can now conclude that with the assumption made on f and ϵ, the rigged Hilbert space $\tilde{Z} \subset \tilde{L}^2_+ \subset \tilde{Z}'$ satisfies the assumptions made in Sec. 3 wrt to H^0 and H, so that the conclusions drawn there hold true in this model. In particular, a resonance at $z_0 = E_0 - i(\Gamma_0/2)$, $\Gamma_0 > 0$, i.e., a pole of $S_{\rm II}(z)$, is associated to left and right eigenvectors, and possibly to associated vectors, of H, with z_0 as eigenvalue. These vectors are given by the Laurent expansion of $G_{\rm II}(z)$ about z_0. As stated above, if ϵ is small enough, there is a simple pole close to M, and there appear no associated vectors. The left resonant eigenvector is

$$D(z_0) = \frac{\alpha(z_0)}{d\alpha_{\rm II}(z)/dz\,|_{z=z_0}} G(z_0), \tag{4.16}$$

whereas the corresponding right eigenvector is simply $F(z_0)$. In particular, the survival amplitude under the perturbed evolution law of the state $G^0 = (1,0)$, stable under the unperturbed evolution, can be written as follows [compare (3.20)]:

$$\langle G^0 | e^{-iHt} G^0\rangle$$
$$= \int_\Gamma dz \exp(-izt) \langle G^0|F(z)\rangle \langle G_{\rm II}(z)|G^0\rangle$$
$$-2\pi i \exp(-iE_0 t) \exp[-(\Gamma_0/2)t]$$
$$\times \langle G^0|F(z_0)\rangle \langle D(z_0)|G^0\rangle$$
$$= \epsilon^2 \int_\Gamma dz \exp(-izt) \frac{\bar{f}(z) f(z)}{\alpha(z)\alpha_{\rm II}(z)}$$
$$-2\pi i \exp(-iE_0 t) \exp[-(\Gamma_0/2)t]\,\epsilon^2$$

$$\times \frac{\bar{f}(z_0) f(z_0)}{\alpha(z_0)\, d\alpha_{\rm II}(z)/dz|_{z=z_0}}$$
$$= \epsilon^2 \int_\Gamma dz \exp(-izt) \frac{\bar{f}(z) f(z)}{\alpha(z)\alpha_{\rm II}(z)}$$
$$+ \exp(-iE_0 t)\exp[-(\Gamma_0/2)t] \frac{1}{d\alpha_{\rm II}(z)/dz|_{z=z_0}}, \tag{4.17}$$

where Γ is path in lower half-plane running below the location of the pole, as in (3.20) (see Fig. 2). As is well known, the exponential term at the rhs of (4.17) dominates the amplitude except at short and long times compared to the mean life $1/\Gamma_0$, at which times the contribution of the background integral along Γ becomes important and is responsible for the deviations from the exponential decay law. Formula (4.17) was derived in Ref. 21, but the interpretation there in terms of eigenvector expansion was formal.

The resolvents of H and H^0, as operators in $\mathscr{L}(\tilde{Z},\tilde{Z}')$ can be analytically continued across the cut to yield meromorphic functions in the second Riemann sheet (which here too is the cut plane), as indicated in Secs. 2 and 3. The continuation of $R^0(z)$, $R^0_{\rm II}(z) = R^0(z) + 2\pi i|F^0(z)\rangle\langle G^0(z)|$, has a pole at $z = M$ only, because $F^0(z)$ and $G^0(z)$ are regular everywhere in the present model [of course, $R^0_{\rm II}(z)P^0$ has no poles]. On the other hand, the continuation of $R(z)$, $R_{\rm II}(z) = R(z) + 2\pi i|F(z)\rangle\langle G_{\rm II}(z)|$ has poles precisely where $G_{\rm II}(z)$, i.e., where $S_{\rm II}(z)$, has.

Therefore, in the Friedrichs model we have been able to associate to each other: (1) poles of the analytic continuation of the S matrix; (2) poles of an analytic family of eigenvectors of the total Hamiltonian H; (3) eigenvectors of H with complex eigenvalues equal to the locations of the poles; (4) poles of a suitable continuation of the resolvent of H, at points where the analogously continued resolvent of the free Hamiltonian is regular. What is more important is the fact that expansions in terms of eigenvectors with complex eigenvalues are valid, so that such poles and corresponding eigenvectors in the fourth quadrant are associated in an intuitive way to purely exponentially decaying contributions to the survival amplitudes of unstable states. In this way, the formal treatment given in Ref. 21 is now completely justified in terms of rigged Hilbert spaces.

5. SCATTERING OF A SPINLESS PARTICLE BY A LOCAL CENTRAL POTENTIAL

We consider the nonrelativistic Schrödinger Hamiltonian $H = H^0 + V(r) = -(1/2\mu)\Delta + V(r)$, $r = |\mathbf{x}|$, for a spinless particle interacting with a local central potential. We assume the potential to b C^∞ but possibly at the origin.[37] We require that at the origin $V(r) \sim O(r^{-3/2+\epsilon})$, $\epsilon > 0$, and that at infinity $V(r) \sim O(r^{-3-\delta})$, $\delta > 0$. Scattering theory for such potentials is well known[33]; here we show how it can be recast into the rigged Hilbert space framework of Secs. 2 and 3.

As usual, we write any function in $L^2(\mathbb{R})$ as

$$\psi(\mathbf{x}) = \frac{1}{r}\sum_{l,m} \psi_{lm}(r) Y_{lm}(\hat{\mathbf{x}}), \tag{5.1}$$

where $\hat{\mathbf{x}}$ denotes the angular polar coördinates of $\mathbf{x}$ and $Y_{lm}(\hat{\mathbf{x}})$ are the spherical harmonics. The converse relation is

$$\psi_{lm}(r) = r \int d\hat{\mathbf{x}}\, \overline{Y_{lm}(\hat{\mathbf{x}})}\psi(\mathbf{x}). \tag{5.2}$$

With these conventions, one has

$$(\psi,\varphi) = \sum_{l,m} \int_0^\infty dr\, \overline{\psi_{lm}(r)}\, \varphi_{lm}(r) \equiv \sum_{l,m} (\psi_{lm},\varphi_{lm})_{lm},$$

for all $\psi,\varphi \in L^2(\mathbb{R})$. The function $\psi_{lm}(r)$ is in L^2_+. However, we shall denote L^2_+ by $\mathscr{H}_{lm}$ whenever it is envisaged as the space to which the l-wave component ψ_{lm} belongs. Accordingly, the inner product in L^2_+, $(h,k) = \int_0^\infty dr\, \overline{h(r)} k(r)$, is denoted by $(\cdot,\cdot)_{lm}$ whenever it is intended between the components ψ_{lm}, φ_{lm}. Similar conventions will be made later on. The action of H on the l-wave component $\psi_{lm}(r)$ is formally given by

$$H_l \psi_{lm}(r) \equiv (H\psi)_{lm}(r)$$
$$= -\frac{1}{2\mu}\frac{d^2}{dr^2}\psi_{lm}(r) + \frac{l(l+1)}{2\mu r^2}\psi_{lm}(r) + V(r)\psi_{lm}(r). \tag{5.3}$$

Remark: If H is defined on a core $D(H)$ in such a way that (5.1) decomposes H, the formal expression (5.3) yields an essentially self-adjoint operator H_l, in $\mathscr{H}_{lm}$ on the domain $D(H_l)$ naturally obtained from $D(H)$ by means of (5.2). Conversely, let the formal expressions (5.3) represent essentially self-adjoint operators in $\mathscr{H}_{lm}$ with domains $D(H_l)$ for any l,m. Then the operator H, when defined upon the set $\{\psi(\mathbf{x})|\psi_{lm}(r) \in D(H_l)$, $\psi_{lm}(r)$ vanishing for all but a finite number of $\mathbf{l},\mathbf{m}\}$, is essentially self-adjoint.

Now we determine a nuclear space that will be shown to implement the assumptions made in Sec. 3 wrt H^0 and H, because of the known properties of the Møller operators for the class of potentials considered. Let $C_0^\infty(0,\infty)$ be the space of infinitely differentiable functions with compact support in the open set $(0,\infty)$. Envisage this space as the strict inductive limit of the countably many complete countably normed nuclear spaces $C_0^\infty([1/n,n])$, $n = 1,2,\cdots$, each of these with the topology of Ref. 22, Vol. II. As such, $C_0^\infty(0,\infty)$ is a nuclear space denoted $\mathscr{D}_+$. Here again we shall denote $\mathscr{D}_+$ by $\mathscr{D}_{lm}$ whenever ψ_{lm} is intended to belong to it. Consider the locally convex direct sum $\mathscr{D}_L \equiv \Sigma_{l=0}^L \Sigma_{m=-l}^l \oplus \mathscr{D}_{lm}$ and the strict inductive limit

$$\Phi = \lim_{\rightarrow} \mathscr{D}_L. \tag{5.4}$$

Then, Φ is also a nuclear space. The conjugation $\varphi \rightarrow \bar{\varphi}$ given by the ordinary complex conjugation $\varphi(r) \rightarrow \overline{\varphi(r)}$ is continuous from Φ onto Φ. Then, formula (5.1) gives a natural one to one correspondence between Φ and a dense manifold in $L^2(\mathbb{R}^3)$ which, endowed with the transport topology, we again denote by Φ. Φ is a actually a closed supspace of the space $\mathscr{D}(\mathbb{R}^3)$ of Schwarz test functions, and its topology coincides with that inherited from this space. It is therefore embedded continuously into $L^2(\mathbb{R}^3)$. For any l,m the operator H_l maps $\mathscr{D}_{lm}$ into itself continuously, so that H is a continuous map of Φ into itself. Remark that in Φ there are only vectors with a finite number of non vanishing components $\varphi_{lm}\in\mathscr{D}_{lm}$. The restriction of H_l to $\mathscr{D}_{lm}$ is essentially self-adjoint iff $l\neq 0$,[38] so that H is not essentially self-adjoint upon Φ. The same statements hold true for H^0.

We show that the space Φ meets the requirements of Sec. 3 wrt H^0 and H. We choose for H and H^0 the following common domain of essential self-adjointness:

$$D=\{\psi(\mathbf{x})|\psi_{lm}(r)\in\mathscr{D}_{lm} \text{ for } l\neq 0;\psi_{lm}(r) \text{ vanishes for all but a finite number of } l,m;\ \psi_{00}(r) \text{ is } C^\infty,\ \psi_{00}(0)=0 \text{ and } \exists R>0 \text{ such that } \psi_{00}(r)=0 \text{ for } r>R\}.$$

The singular continuous spectra are empty.

We consider first the free Hamiltonian H^0. The normalized Riccati–Bessel functions

$$\hat{j}_l(x)=x^{1/2}J_{l+1/2}(x),$$

where $J_\lambda(x)$ are the ordinary Bessel functions, satisfy the free radial Schrödinger equation

$$-\frac{1}{2\mu}\frac{d^2}{dr^2}\hat{j}_l(kr)+\frac{l(l+1)}{2\mu r^2}\hat{j}_l(kr)=\frac{k^2}{2\mu}\hat{j}_l(kr), \quad (5.5)$$

and they define continuous functionals upon $\mathscr{D}_{lm}$, as locally summable functions. Then, the functionals $G^0_{lm}(k)\equiv\langle k,l,m|$ defined upon Φ as

$$\langle k,l,m|\chi\rangle=\int_0^\infty dr\,\hat{j}_l(kr)\chi_{lm}(r)\quad\forall\chi\in\Phi, \quad (5.6)$$

are continuous on Φ and left eigenvectors of H^0 with eigenvalue $k^2/2\mu$. Furthermore, they are analytic in the parameter k in the whole complex plane, so that they yield a family of eigenvectors of H^0 which is entire analytic. The completeness relation holds

$$(\chi,\omega)=\int_0^\infty dk\sum_{lm}\overline{\langle k,l,m|\chi\rangle}\langle klm|\omega\rangle\quad\forall\chi,\quad\omega\in\Phi.$$

The map

$$\mathscr{D}_{lm}\ni\chi_{lm}(r)\rightarrow\int_0^\infty dr\,\hat{j}_l(kr)\chi_{lm}(r)\in L^2_+(dk),$$

has a dense image in $L^2_+(dk)$ and hence the map

$$\mathscr{F}^0:\Phi\ni\chi\rightarrow[\mathscr{F}^0(\chi)]_{lm}(k)\equiv\langle k,l,m|\chi\rangle, \quad (5.7)$$

has a dense image in $\Sigma_{l,m\oplus}L^2_+(dk)$. Therefore, conditions (a), (b), and (c) of Proposition 2.5 are valid for the family $\langle k,l,m|$, with $\Lambda=(0,\infty)$, $d\mu(k)=dk$ and $\mathscr{F}=\mathscr{F}^0$. Condition (d) is expressed by the relation

$$[\bar{\mathscr{F}}^0(H^0\psi)]_{lm}(k)=\frac{k^2}{2\mu}[\bar{\mathscr{F}}^0(\psi)]_{lm}(k)\quad\forall\psi\in D\ \text{ and a.e.}$$

which is easily seen to hold by partial integration, noting that for all $\psi\in D$ one has

$$[\bar{\mathscr{F}}^0(\psi)]_{00}(k)=\textstyle\int_0^\infty dr\,\hat{j}_l(kr)\psi_{00}(r)\quad\text{a.e.}$$

Concluding, the family $\langle k,l,m|$ is an entire analytic family of left eigenvectors of H^0 in Φ' that implements Theorem 2.1.

Remark: Note that, because the functions in Φ have support away from the origin, the Riccati–Neumann functions $\hat{n}_l(kr)=(kr)^{1/2}N_{l+1/2}(kr)$ are also eigenvectors of H^0 in Φ'. They provide an example of eigenvectors of a self-adjoint operator with corresponding eigenvalues in the spectrum and yet playing no role in Theorem 2.1.

Next, we consider the total Hamiltonian H. It is well known[39] that the radial Schrödinger equation $H_l\varphi_{l,k}(r)=(k^2/2\mu)\varphi_{l,k}(r)$ has a strong solution that behaves like $(kr)^{l+1}/(2l+1)!!$ as $r\rightarrow 0$. This solution, called the regular solution,[33] is entire analytic in the parameter k for any r under the assumptions made on $V(r)$. In addition, it is not difficult to check by the usual expansions given for it[39] that it defines a left eigenvector of H_l in $\mathscr{D}'_{lm}$ and that the expression $\int_0^\infty dr\,\varphi_{l,k}(r)\chi_{lm}(r)$ is entire analytic in k whenever $\chi_{lm}\in\mathscr{D}_{lm}$. Thus we have an entire analytic family of left eigenvectors of H_l in $\mathscr{D}'_{lm}$.

The relation between the regular solution $\varphi_{l,k}(r)$ and the normalized incoming solution $\psi_{l,k}(r)$ is given by $\varphi_{l,k}(r)=(\pi/2)^{1/2}f_l(k)\psi_{l,k}(r)$. The Jost function[33] $f_l(k)=\bar{f}_l(-\bar{k})$ is an analytic function of k in the half plane $\mathrm{Im}k>0$ and it is continuous for real k except possibly at $k=0$. Therefore, the functionals $G^{(+)}_{lm}(k)$ defined upon Φ as

$$\langle G^{(+)}_{lm}(k)|\chi\rangle=\int_0^\infty dr\bar{\psi}_{l,k}(r)\chi_{lm}(r)=(2/\pi)^{1/2}(1/f_l(-k))\times\int_0^\infty dr\,\varphi_{l,k}(r)\chi_{lm}(r)\quad\forall\chi\in\Phi, \quad (5.8)$$

are left eigenvectors of H, boundary value for $k>0$ of a family analytic in the lower complex k-plane. Denote by G_{nlm} the (left) bound states and by P the orthogonal projection onto the subspace $\mathscr{H}_{ac}$ of absolute continuity of H. Then the family $\{G_{nlm},G^{(+)}_{lm}(k)\}$ satisfies all conditions of Proposition (2.5) and provides for the generalized Fourier transform $\mathscr{F}^+$ of Sec. 3. In particular, the $G^{(+)}_{lm}(k)$ are the l-wave left incoming eigenvectors. We do not need to give a direct proof of these statements, as they follow from the existence and completeness of the Møller wave operators for the class of potentials under consideration.[33] It suffices to note that condition (a) of the proposition holds as a consequence of the definition and that conditions (b) and (c) follow from the relation

$$\int_0^\infty dr\,\bar{\psi}_{l,k}(r)\chi_{lm}(r)=\lim_{R\rightarrow\infty}\int_0^R dr\,\hat{j}_l(kr)(W^*_+P\chi)_{lm}(r)=[\bar{\mathscr{F}}^0(W^*_+P\chi]_{lm}(k)\quad\text{a.e. and }\forall\chi\in\Phi,$$

and from completeness and unitarity of the Møller operators between $L^2(\mathbb{R}^3)$ and $\mathscr{H}_{ac}$. As to condition (d), the proof of its validity runs exactly as it does for H^0. In conclusion, the left incoming eigenvectors (5.8) are boundary values on $[0,\infty)$ of a family of eigenvectors of H analytic in the half plane $\mathrm{Im}\,k<0$. We are in the condition of applying Proposition 3.1 so that the left outgoing eigenvectors are given by

$$G^{(-)}_{lm}(k)=s_l(k)G^{(+)}_{lm}(k),$$

where $s_l(k)=\bar{f}_l(k)/f_l(k)$ is the usual l-wave S matrix. The corresponding right incoming and outgoing eigenvectors are given for $k>0$ respectively by

$$\langle\chi|F^{(+)}_{lm}(k)\rangle = \int_0^\infty \psi_{l,k}(r)\chi_{lm}(r)\,dr = \langle G^{(-)}_{lm}(k)|\chi\rangle \quad \forall\chi\in\Phi,$$

and by

$$\langle\chi|F^{(-)}_{lm}(k)\rangle = \int_0^\infty dr\,\bar\psi_{l,k}(r)\chi_{lm}(r)\,dr = \langle G^{(+)}_{lm}|\chi\rangle \quad \forall\chi\in\Phi.$$

The eigenvectors $G^{(-)}_{lm}(k) = F^{(+)}_{lm}(k)$ are boundary values on $[0,\infty)$ of a family of eigenvectors of H analytic in the half-plane Im $k>0$. All conditions of Sec. 3 are met and the family $G^{(-)}_{lm}(k)$ can be continued across the real axis according to formula (3.15) in the same region as $f_l(k)$ can, whenever the potential is such as to allow for the continuation of the Jost functions. The singularities of the continued family $G^{(-)\mathrm{II}}_{lm}(k)$ in the lower half-plane are poles located at the zeros of the Jost functions. They give rise in the manner described in Sec. 2 to eigenvectors (and possibly to associated vectors) of the Hamiltonian, with complex eigenvalues, and appear as isolated contributions in (3.20). We remark in this context that with our choice (5.4) for the nuclear space Φ we have here that $P\Phi\not\subset\Phi$ contrary to the situation encountered in the Friedrichs model. Therefore, formulas (3.20) and (3.21) allow for the possibilty that $h\neq\varphi$, etc.

The resolvents can also be continued and formulas (3.25) and (3.26) are valid [notice that

$$\Sigma_{l,m}\langle\chi|F^{(-)}_{lm}(k)\rangle\langle G^{(-)\mathrm{II}}_{lm}(k)|\omega\rangle$$

is analytic because the series contains a finite number of terms]. Here too the poles of $R_{\mathrm{II}}(k)$ are the zeros of the continued Jost functions.

6. CONCLUDING REMARKS

We summarize here what has been done in the previous sections and point out the relevant features of our treatment.

It is well know from the work of Maurin, Gel'fand, and collaborators that given a self-adjoint operator A in a Hilbert space $\mathscr{H}$, there exists a (nonunique) rigged Hilbert space $\Phi\subset\mathscr{H}\subset\Phi'$ such that the dual A' of A in Φ' has a complete set of eigenvectors with at least the points of the spectrum of A as eigenvalues. In general the eigenvalue equation for A' in Φ' has also solutions outside the spectrum. This feature has often been considered to be a nuisance (from a mathematical point of view) and conditions have been given in order that the eigenvalue equation in Φ' has solutions corresponding to the points in the spectrum only.[40] However, there are cases when it is desirable to have eigenvalues outside the spectrum. A first example of this fact was met by physicists in connection with the reduction of self-adjoint representations of a noncompact Lie algebra with respect to a noncompact Abelian subalgebra. As shown for example by Mukunda and collaborators,[41] this problem can be solved strictly within the framework of Hilbert space; essentially, the solution amounts to the fact that the direct integral decomposition of the carrier space of the representation with respect to the noncompact generators is such that vectors belonging to the commom domain of all the relevant operators are represented by functions which are analytically continuable. Of course, in a Hilbert space setting eigenvectors are envisaged in a formal sense only. In order to solve the problem, while at the same time giving a rigorous meaning to eigenvectors, one has to use rigged Hilbert space techniques. This way first done explicitly by Iverson[42] and later independently by Lindblad and Nagel.[43]

Also, for the description of resonances it is desiderable to have solutions of the eigenvalue equation outside the spectrum.[9,13] In our approach, which stems from earlier work of some of us about analytic continuation in the Lee–Friedrichs model,[21] we consider rigged Hilbert spaces $\Phi\subset\mathscr{H}\subset\Phi'$ such that both the unperturbed Hamiltonian H^0 and the perturbed one H have eigenvalues in regions of the complex plane. Here we wish to point out that the eigenvectors with eigenvalues outside the spectrum of the Hamiltonian *have the same physical concreteness as stationary scattering states or plane waves*. Indeed, if $G(z_0)$ is such an eigenvector in Φ' with z_0 as eigenvalue, it can be approximated, in a weak sense, by vectors in Φ, i.e., physical wave packets, and the eigenvalue equation can be approximated analogously. To wit, $\forall\epsilon>0$ and $\forall\varphi\in\Phi$, $\exists\psi\in\Phi$ such that $|\langle G(z_0)|\varphi\rangle - \langle\psi|\varphi\rangle| < \epsilon$ and $|z_0\langle\psi|\varphi\rangle - \langle\psi|A\varphi\rangle| < \epsilon$. The families of eigenvectors that we deal with are also required to be analytic with respect to the eigenvalue, considered as a parameter, in some regions of the complex plane. Once we assume that resonances are given by poles of the S matrix continued into the unphysical sheet, if we are able to connect the analytic structure of the eigenvector family of the perturbed Hamiltonian to that of the S matrix, resonances appear as poles of the eigenvector family. The interesting fact is that *the coefficient of the most singular term in the Laurent expansion of the eigenvector family about the pole z_0 is also an eigenvector with z_0 as eigenvalue.* In this way, resonances appear as eigenvectors of H with complex eigenvalues. The family of eigenvectors of H whose analytic structure is linked to that of the S matrix turns out to be the one associated to the "out state" representation of the Hilbert space. In this framework, a rigorous ground is given to the interpretation of the analytic continuation techniques of Ref. 21 in terms of expansions over eigenvectors with complex eigenvalues of the Hamiltonian. As we have eigenvector families which are analytic in some region, we can apply Cauchy techniques to scalar products between elements in Φ. Such scalar products can be expanded in terms of complete systems of eigenvectors with complex eigenvalues of either H^0 or H. There appears an integral along a contour in the complex plane, plus discrete contributions due to the poles of the eigenvector family. The expansion of the survival amplitude of a state in Φ in terms of eigenvectors with complex eigenvalues breaks down into a sum of terms with exponential (or polynomial times exponential) behavior in time in correspondence to the singularities of the eigenvector family, plus a background integral, a fact with stresses again the connection between unstable states and resonances. We have seen also how the resolvents of H^0 and H, reduced to Φ, can be continued across the absolutely continuous spectrum (provided the singular spectrum has good behavior) into the unphysical sheet. The continuation of the resolvent of H has the same singularities as the continuation of the S matrix; in contrast, the continuation of the resolvent of H^0 is regular at these points.

There is a criticism that could be made about the fact that we use a perturbed and an unperturbed Hamiltonian: namely, that since nature ignores the unperturbed Hamiltonian, resonances should be entirely characterized in terms of the total Hamiltonian alone. However, we maintain that this cannot be the case: *In addition to the total Hamiltonian, one needs some further input from physics*. This reflects itself upon the mentioned arbitrariness in the choice of Φ and on the fact that the analytic structure of the eigenvector family depends critically upon such choice. This feature can be easily shown again in the example of the Friedrichs model. Let us consider a Hamiltonian H whose absolutely continuous spectrum is $[0,\infty)$ and Lebesgue; it can be assumed to be simple. Then, the restriction of H to the subspace of absolute continuity is given by the operator of multiplication by x in $L^2(0,\infty)$. Let us consider the operator H^0 given by:

$$\begin{aligned}(H^0h)(\lambda) = {} & M\epsilon^2 \frac{f(\lambda)}{\alpha(\lambda+i0)}\int_0^\infty \frac{\overline{f(x)h(x)}}{\alpha(x-i0)}dx \\ & + \lambda h(\lambda) - \epsilon^2\lambda f(\lambda) \\ & \times \int_0^\infty \frac{\overline{f(x)}h(x)}{\alpha(x-i0)(\lambda+i0-x)} \\ & + \epsilon^2 \frac{f(\lambda)}{\alpha(\lambda-i0)}\int_0^\infty \frac{\overline{f(x)}}{\lambda+i0-x} \\ & \times\bigg(xh(x) - \epsilon^2 x f(x) \\ & \times \int_0^\infty \frac{f(x')h(x')}{\alpha(x'-i0)(x+i0-x')}dx'\bigg)dx \quad (6.1)\end{aligned}$$

where ϵ, M, f are arbitrary provided they satisfy the conditions of Sec. 4. H^0 and H are then immediately recognized to be the Friedrichs unperturbed and perturbd Hamiltonians, respectively, in the "out state" representation. By means of a unitary transformation, they can be given the form (4.1a) and (4.1b), respectively. In this representation, one can repeat all the steps of Sec. 4 to obtain poles of the continued resolvent of H, exponential time behavior, etc., at the zeros of the function $\alpha_{\rm II}(z)$. However, this function depends on the parameters ϵ, M and on the function f, which can be arranged in such a way as to have a zero of $\alpha_{\rm II}(z)$ at any arbitrarily given point of the fourth quadrant.[44] Therefore, given H, one can choose Φ which will depend on the parameters ϵ, M and on the function f, in such a way that the resolvent, as an operator in $\mathscr{L}(\Phi,\Phi')$, can be continued across the spectrum into the unphysical sheet with a pole at any given point of the fourth quadrant; to such a pole are associated all the mathematical features of an unstable state or a resonance and, of course, if the imaginary part of the pole is not too large, the physical features too. However, resonances are *not* seen with any energy and lifetime. This suggests that, for a given Hamiltonian system, not all the states in the corresponding Hilbert space are realizable *in principle*, but that there is a mechanism that selects the physical states; *and this mechanism does not depend on the Hamiltonian alone*. Such a mechanism could possible be the following: The "natural" states for the system are those in the domain of the Hamiltonian; on the other hand, the observer has in the same Hilbert space a "natural" set of his own, depending on the structure of the preparing and measuring devices, and at least in many instances, these can probably be described as states in the domain of a fictitious, unperturbed, Hamiltonian H^0. The observable states should then belong at most to the intersection of such two sets, the system's natural set and the observer's natural set of states, which perhaps could explain why resonances are not found everywhere in the complex energy plane. These considerations reflect themselves upon the arbitrariness of the choice of Φ in the rigged Hilbert space, and on the dependence of the analytic structure of the eigenvector families upon such choice.[45] In this paper, to overcome such arbitrariness, we have considered the fictitious Hamiltonian *as given*, and adhered to the point of view that resonances are associated with poles of the S matrix, being well aware that this is not always the case.[5]

ACKNOWLEDGMENTS

This paper is dedicated to Günther Ludwig on the occasion of his 60th birthday. One of the authors (GP) gratefully acknowledges a Fulbright-Hays Act grant.

APPENDIX A

We show that: (a) $\tilde{Z}$ is invariant wrt the actions of H^0 and H; (b) both H^0 and H act upon $\tilde{Z}$ with continuity.

First of all remark that, as $f\in Z$, there is some positive number a and a sequence $\{C_k\}$, $C_k>0$, such that $(1+|z|)^k|f(z)|\leqslant C_k\exp(a|\mathrm{Im}\,z|)$. (a) Given any vector $\tilde{\varphi}=(\varphi^0,\varphi)$ in $\tilde{Z}$, there is a positive b such that $\varphi\in Z(b)$. Then, there is a sequence $\{B_k\}$, $B_k>0$, such that $(1+|z|)^k|\varphi(z)|\leqslant B_k\exp(b|\mathrm{Im}\,z|)$. Therefore, $(1+|z|)^k|\epsilon\varphi^0 f(z)+z\varphi(z)|\leqslant(1+|z|)^k|\epsilon\varphi^0||f(z)|+(1+|z|)^{k+1}|\varphi(z)|\leqslant|\epsilon\varphi^0|C_k\exp(a|\mathrm{Im}\,z|)+B_{k+1}\exp(b|\mathrm{Im}\,z|)$, so that $H\tilde{\varphi}\in\tilde{Z}(\max[a,b])$. Setting $\epsilon=0$ gives $H^0\tilde{\varphi}\in\tilde{Z}(b)$. (b) In order that linear operator $A:\tilde{Z}\to\tilde{Z}$ be continuous it is necessary and sufficient[30] that (i) A transforms any $\tilde{Z}(b)$ into some $\tilde{Z}(c)$ and (ii) that the action $A\restriction\tilde{Z}(b):\tilde{Z}(b)\to\tilde{Z}(c)$ be continuous. In (a) it is shown that H transforms any $\tilde{Z}(b)$ into $\tilde{Z}(\max[a,b])$, so that it satisfies property (i). If $a>b$, we have $\tilde{Z}(a)\supset\tilde{Z}(b)$ and the topology induced by $\tilde{Z}(a)$ on $\tilde{Z}(b)$ coincides with the topology of $\tilde{Z}(b)$. Therefore, it is enough to prove that H satisfies property (ii) for $b\geqslant a$. In this case, $\tilde{Z}(b)$ is invariant under H and if $\tilde{\varphi}\in\tilde{Z}(b)$, we have

$$\begin{aligned}\|H\tilde{\varphi}\|_k &= \left|M\varphi^0+\epsilon\int_0^\infty \bar{f}(x)\varphi(x)\,dx\right|^2 + \sup_{z\in\mathbb{C}}(1+|z|)^{2k}\exp(-2b\,|\mathrm{Im}\,z|)|\epsilon\varphi^0 f(z)+z\varphi(z)|^2 \\ &\leqslant\left(|M\varphi^0|+\epsilon\int_0^\infty\frac{|f(x)|}{(1+|x|)^{k+1}}|\varphi(x)|(1+|x|)^{k+1}dx\right)^2 + \sup_{z\in\mathbb{C}}(1+|z|)^{2k}\exp(-2b\,|\mathrm{Im}\,z|)\left|\frac{\varphi^0}{(1+|z|)^k}\right.\end{aligned}$$

$$\times \exp(b\,|\mathrm{Im}\,z|)\epsilon f(z)(1+|z|)^k \exp(-b\,|\mathrm{Im}\,z|) + \frac{z}{1+|z|}\varphi(z)(1+|z|)\Big|^2$$

$$\leqslant \Big(|M\varphi^0| + \Big[\sup_{z\in C} |\varphi(z)|(1+|z|)^{k+1}\exp(-b\,|\mathrm{Im}\,z|)\Big]\epsilon\int_0^\infty \frac{|f(x)|}{(1+|x|)^{k+1}}dx\Big)^2$$

$$+ \sup_{z\in C}\Big[(1+|z|)^{2k}\exp(-2b\,|\mathrm{Im}\,z|)(\epsilon^2|f(z)|^2(1+|z|)^{2k}\exp(-2b\,|\mathrm{Im}\,z|) + \frac{|z|^2}{1+|z|^2}\Big]$$

$$\times\Big[\frac{|\varphi^0|^2}{(1+|z|)^{2k}}\exp(2b\,|\mathrm{Im}\,z|) + (1+|z|)^2|\varphi(z)|^2\Big]$$

$$\leqslant \Big[M^2 + \Big(\epsilon\int_0^\infty \frac{|f(x)|}{(1+|x|)^{k+1}}dx\Big)^2\Big]\Big[|\varphi^0|^2 + \sup_{z\in C}(1+|z|)^{2(k+1)}\exp(-2b\,|\mathrm{Im}\,z|)|\varphi(z)|^2\Big]$$

$$+ \sup_{z\in C}\Big(\epsilon^2|f(z)|^2(1+|z|)^{2k}\exp(-2b\,|\mathrm{Im}\,z|) + \frac{|z|^2}{1+|z|^2}\Big)$$

$$\times\Big[|\varphi^0|^2 + \sup_{z\in C}(1+|z|)^{2(k+1)}\exp(-2b\,|\mathrm{Im}\,z|)|\varphi(z)|^2\Big]$$

$$= A_{k+1}\,\|\tilde\varphi\|^2_{k+1},$$

where

$$A_k = M^2 + \epsilon\int_0^\infty \frac{f(x)}{(1+|x|)^k}dx + \sup_{z\in C}\Big(\epsilon^2|f(z)|^2(1+|z|)^{2(k-1)}$$

$$\times\exp(-2b\,|\mathrm{Im}\,z|) + \frac{|z|^2}{(1+|z|)^2}\Big).$$

Setting $\epsilon = 0$ gives that H^0 also satisfies property (ii).

APPENDIX B

By (4.3), the left eigenvalue equation for H, $H'G(\xi) = \xi G(\xi)$, $G(\xi) = (g_{0\xi}, g_\xi)$, writes explicitly

$$Mg_{0\xi} + \epsilon\langle g_\xi | f\rangle = \xi g_{0\xi}, \tag{B1a}$$

$$\epsilon g_{0\xi}\bar f + x\,g_\xi = \xi g_\xi. \tag{B1b}$$

A particular solution of (B1b) is

$$g^p_\xi = \epsilon g_{0\xi}\frac{\bar f(x)}{\xi - x}, \tag{B2}$$

where, if $\xi \equiv \lambda\in[0,\infty)$, we must interpret (B2) as signifying one or the other of the two boundary values $\lim_{\eta\downarrow 0} g^p_{\lambda\pm i\eta}$. The associated homogeneous equation

$$(\xi - x)g^h_\xi = 0, \tag{B3}$$

is the Fourier transform of the equation $(\xi - i\,d/dx)\hat g_\xi = 0$, where $\hat g_\xi$ is in $\mathscr{D}'$, the dual of the space $\mathscr{D}$ of Schwarz test functions. Since for every $\xi\in C$ the latter equation has as only solution $\exp(i\xi x)$ up to an arbitrary multiplicative factor (Ref. 22, Vol. I), the general solution of (B3) is of the form

$$g^h_\xi = a(\xi)\delta_\xi, \quad \forall\xi\in C,$$

where $a(\xi)$ is an arbitrary ξ-dependent factor. Therefore, the general solution of (B1b) is

$$g_\xi = a(\xi)\delta_\xi + g_{0\xi}\frac{\epsilon\bar f(x)}{\xi - x} \tag{B4}$$

Substituting (B4) into (B1a) gives

$$\alpha(\xi)g_{0\xi} = a(\xi)\epsilon f(\xi). \tag{B5}$$

Since $\alpha(\xi)$ does not vanish under our assumptions on f and ϵ we can solve (B5) for $g_{0\xi}$ and substituting into (B4) gives

$$G(\xi) = a(\xi)\Big(\epsilon\frac{f(\xi)}{\alpha(\xi)}, \epsilon^2\frac{f(\xi)}{\alpha(\xi)}\cdot\frac{\bar f(x)}{\xi - x} + \delta_\xi\Big), \quad \forall\xi\in C, \tag{B6}$$

thus proving Proposition 4.1 [in (4.6) we have set $a(\xi) = 1$]. Alternatively, using (4.4), one finds that the solution to the right eigenvalue equation for H, $\hat H F(\xi) = \xi F(\xi)$, is given by

$$F(\xi) = b(\xi)\Big(\epsilon\frac{\bar f(\xi)}{\alpha(\xi)}, \epsilon^2\frac{\bar f(\xi)}{\alpha(\xi)}\cdot\frac{f(\xi)}{\xi - x} + \delta_\xi\Big), \quad \forall\xi\in C, \tag{B7}$$

where $b(\xi)$ is an arbitrary ξ-dependent factor.

APPENDIX C

In order to prove that $\tilde Z$ is a domain of essential self-adjointness for H, it suffices to show that the equations $(H^* \pm i)h = 0$ have no solutions besides $h = 0$ in $D(H^*)$. But this is an immediate consequence of the fact that $\hat H\tilde{\mathscr{F}}'_+ h = \tilde{\mathscr{F}}'_+ H^*h$ for all $h\in D(H^*)$ (see Sec. 2) and that none of the solutions (B7) of the eigenvalue equation $\hat HF(\xi) = \xi F(\xi)$ is in $\tilde L^2_+$. Indeed, the delta function δ_ξ is not in $\tilde L^2_+$ for any $\xi\in C$, whereas the other term in (B7) is in $\tilde L^2_+$. Similarly, one finds that $\tilde Z$ is also a domain of essential self-adjointness for H^0.

[1] L. Fonda, G. C. Ghirardi, and A. Rimini, Rep. Progr. Phys. **41**, 587 (1978).

[2] A. P. Grecos and I. Prigogine, "Irreversible Processes in Quantum Theory," in *Group Theoretical Methods in Physics*, edited by W. Beiglböck, A. Böhm, and E. Takasugi, *Lecture Notes in Physics* (Springer, New York, 1979), Vol. 94, p. 229.

[3] L. A. Khalfin; (a) Zh. Eksp. Teor. Fiz. **33**, 1371 (1957) [Sov. Phys. JETP **6**, 1053 (1958)]; (b) Zh. Eksp. Teor. Fiz. Pis'ma Red. **8**, 106 (1968) [JETP Lett. **8**, 65 (1968)].

[4] F. Haake, in *Springer Tracts in Modern Physics* (Springer, New York, 1973), Vol. 66, p. 98.

[5] G. Calucci, L. Fonda, and G. C. Ghirardi, Phys. Rev. **166**, 1719 (1968); G. Calucci and G. C. Ghirardi, Phys. Rev. **169**, 1339 (1968).

[6] K. O. Friedrichs, Commun. Pure Appl. Math. **1**, 361 (1948).

[7] E. B. Davies, Lett. Math. Phys. **1**, 31 (1975).

[8]J. Schwinger, Ann. Phys. (N. Y.) **9**, 169 (1960).
[9]A. Grossmann, J. Math. Phys. **5**, 1025 (1964).
[10]J. Aguilar and J. M. Combes, Commun. Math. Phys. **22**, 269 (1971); E. Balslev and J. M. Combes, Commun. Math. Phys. **22**, 280 (1971).
[11]M. Reed and B. Simon, *Methods of Modern Mathematical Physics Vol. IV* (Academic, New York 1978), Secs. XII.6, XIII.10, and references quoted therein.
[12]J. S. Howland; (a) Am. J. Math. **91**, 1106 (1969); (b) Pac. J. Math. **29**, 565 (1969); (c) Trans. Am. Math. Soc. **162**, 141 (1971); (d) Bull. Am. Math. Soc. **78**, 280 (1972); (e) J. Math. Anal. Appl. **50**, 415 (1975).
[13]H. Baumgärtel, Math. Nachr. **75**, 133 (1976); "Mathematical Remarks to Resonances and their Eigenfunctionals in Decay-Scattering Systems, demonstrated by means of Friedrichs' Model," in *Group Theoretical Methods in Physics*, edited by W. Beiglböck, A. Böhm, and E. Takasugi, *Lecture Notes in Physics* (Springer, New York, 1979), Vol. 94, p. 244.
[14]L. P. Horwitz and J. P. Marchand, Rocky Moun. J. Math. **1**, 225 (1971).
[15]L. P. Horwitz and J. M. Sigal, Helv. Phys. Acta **51**, 685 (1978).
[16]V. Weisskopf and E. P. Wigner, Z. Phys. **63**, 54 (1930).
[17]T. Kato, *Pertubation Theory for Linear Operators* (Springer, New York, 1966).
[18]T. D. Lee, Phys. Rev. **95**, 1329 (1954).
[19]The Lee model has been studied with dilation analytic potential techniques in: R. A. Weder, J. Math. Phys. **15**, 20 (1974).
[20]C. B. Chiu, E. C. G. Sudarshan, and B. Misra, Phys. Rev. D **16**, 520 (1977).
[21]E. C. G. Sudarshan, C. B. Chiu, and V. Gorini, Phys. Rev. D **18**, 2914 (1978).
[22]I. M. Gel'fand, M. I. Graev, G. E. Shilov, N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1964).
[23]K. Maurin, *General Eigenfunction Expansions and Unitary Representations of Topological Groups* (PWN, Warsaw, 1968).
[24]J. E. Roberts, Commun. Math. Phys. **3**, 98 (1966); J. P. Antoine, J. Math. Phys. **10**, 53 (1969); O. Melsheimer, J. Math. Phys. **15**, 902 (1974) and **15**, 917 (1974).
[25]For a profitable introduction to the use of rigged Hilbert space techniques in quantum mechanics, see A. Böhm, Boulder Lec. Theor. Phys. **9A**, 255 (1966) and *The Rigged Hilbert Space and Quantum Mechanics, Lecture Notes in Physics* (Springer, New York, 1978), Vol. 78.
[26]G. Parravicini, Resonances as complex eigenvectors for degenerate perturbations (in preparation).
[27]T. K. Bailey and W. C. Schieve, Nuovo Cimento A **47**, 231 (1978).
[28]V. Gorini and G. Parravicini, *Unstable Quantum States and Rigged Hilbert Spaces*, in *Group Theoretical Methods in Physics*, edited by W. Beiglböck, A. Böhm, and E. Takasugi, *Lecture Notes in Physics* (Springer, New York, 1979), Vol. 94, p. 219.
[29]Remark that the definition of rigged Hilbert space used in this paper is slightly more general than that originally introduced in Ref. 22.
[30]H. H. Schaefer, *Topological Vector Spaces* (Macmillan, New York, 1966).
[31]See, e.g., Ref. 22, Vol. I, in connection with homogeneous and associated functions.
[32]It should be remarked that, though some of the features of our treatment are better highlighted by introducing a conjugation, one could avoid this by considering the space $\Phi^\times$ of antilinear continuous functionals on Φ, in addition to Φ'. The use of $\Phi^\times$ in our treatment is sketched in Ref. 28. By introducing $\Phi^\times$, the hypothesis of the existence of a conjugation is dropped.
[33]J. R. Taylor, *Scattering Theory* (Wiley, New York, 1972).
[34]In this connection, we note that another point of view is advocated by Prigogine and collaborators (see, e.g., Ref. 2).
[35]We recall that in scattering problems it is often convenient to parametrize the continuum eigenvectors of H^0 and H in terms of some suitable function $p(\lambda)$ of the energy eigenvalue, instead of the eigenvalue itself. Since $p(\lambda)$ is in general an analytic function of λ, the analyticity properties of the eigenvectors, of the S matrix and of the continued resolvents are not destroyed, but get modified accordingly (compare Sec. 5).
[36]E. B. Davies, J. Math. Phys. **15**, 2036 (1974).
[37]The treatment we give could easily be extended to the case when $V(r)$ is C^∞ except possibly at a finite number of finite discontinuities.
[38]W. O. Amrein, J. M. Jauch, and K. B. Sinha, *Scattering Theory in Quantum Mechanics* (Benjamin, Reading, Mass., 1977).
[39]R. G. Newton, J. Math. Phys. **1**, 319 (1960).
[40]W. G. Faris, Helv. Phys. Acta **44**, 930 (1971); D. Babbitt, Rep. Math. Phys. **3**, 37 (1972); K. Napiorkowski, Bull. Acad. Polon. Sci. Sér. Sci. Math. Phys. Astron. Phys. **22**, 1215 (1974) and **23**, 251 (1975); D. Friedricks, Rep. Math. Phys. **8**, 277 (1975).
[41]J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, Commun. Math. Phys. **8**, 204 (1968); J. Math. Phys. **9**, 2100 (1968).
[42]G. J. Iverson, *Unitary Adjoint Representations of the Lorentz Group I, II*, University of Adelaide (1967); Phys. Lett. B **26**, 229 (1968).
[43]G. Lindblad and B. Nagel: Ann. Inst. Henri Poincaré, A **13**, 27 (1970).
[44]Remark that if z_0 is a zero of $\alpha_{11}(z)$ for some ϵ, M, f, then given any positive number b, bz_0 is a zero of $\alpha_{11}(z)$ for $\epsilon' = b^{1/2}\epsilon$, $M' = bM$, $f'(z) = f(z/b)$. Two of us (GP, VG) are indebted to G. C. Ghirardi for pointing out this fact to them.
[45]These considerations will be developed in a paper with G. C. Ghirardi.

A variational problem for the forward or backward scattering amplitude with three constraints

I. A. Sakmar [a),b)]

Université de Nice Laboratoire de Physique Théorique, [c)] *Parc Valrose 06034 Nice Cedex, France*

(Received 29 January 1980; accepted for publication 20 March 1980)

The problem of extremizing the slope of the imaginary part of either the forward or the backward scattering amplitude is treated as a variational problem with three constraints, one of which is the other slope and the other two σ^T and $\sigma_{\rm el,im}$.

I. INTRODUCTION

In the variational problem of extremizing the slope of the forward scattering amplitude of spinless particles the two constraints which have been used are the total cross section σ^T and $\sigma_{\rm el,im}$. This last quantity is the contribution of the imaginary parts of the partial waves a_l to the elastic cross section. Both of these constraints have the form of single series of a_l. Moreover, they are positive. There are not many quantities which can be expressed as a single series of a_l only, which can be used as constraints of the type σ^T and $\sigma_{\rm el,im}$. One such quantity, however, is the slope of the backward scattering. In this paper we have taken this as a third constraint and applied the Lagrange multipliers method to extremize the slope of the forward scattering amplitude. Alternatively, if the forward slope is known the problem can be viewed as a bound for the slope of the backward amplitude. Because the new constraint has the form of an alternating series we have to introduce two types of a_l's, even ones and the odd ones. The unitarity is imposed in the form of positiviness and boundedness on both a_l's.

In Sec. II we define our constraints and with the help of the expressions found for a_l's by the application of the Lagrange multipliers method and the unitarity, express them as functions of the parameters. Elimination of the Lagrange parameters makes it possible to express one of the four quantities as a function of the remaining three. We consider in this section the case $\alpha > 1$, one of the possible values of a Lagrange parameter ($\alpha > 1$ or $\alpha < 1$) which determines the behavior of the imaginary parts of the partial waves. Calculations are done both by summing the series and by integrals. The results are made consistent among themselves by considering the inaccuracies introduced by the transition regions between the different summation domains. All formulas thus reduce to the results of the two-constraint case.

In Sec. III we consider the case $\alpha < 1$ and derive the corresponding equations. Finally we discuss our results.

II. BOUND WHEN $\alpha > 1$

We give the expressions for the following four quantities in terms of the imaginary parts of the partial waves and define them as v, g, h and u.

$$\left.\frac{d\,\mathrm{Im} f}{dt}\right|_{z=+1} = \frac{1}{2k^2}\frac{\sqrt{s}}{k}\sum \tfrac{1}{2} l(l+1)(2l+1)a_l \equiv u, \quad (1)$$

$$\sigma^T = \frac{4\pi}{k^2}\sum (2l+1)a_l \equiv g, \quad (2)$$

$$\sigma_{\rm el,im} = \frac{4\pi}{k^2}\sum (2l+1)a_l^2 \equiv h, \quad (3)$$

$$\left.\frac{d\,\mathrm{Im} f}{dt}\right|_{z=-1} = \frac{1}{2k^2}\frac{\sqrt{s}}{k}\sum \tfrac{1}{2} l(l+1)(2l+1) \times(-1)^{l+1}a_l \equiv v. \quad (4)$$

Three of these can be considered constraints to find a bound on the fourth one. With Lagrange multipliers α, β, and γ we form

$$u - \alpha g - \beta h - \gamma v$$

and differentiate with respect to a_l. After redefining the parameters, one finds

$$a_l = \alpha + \gamma l(l+1)(-1)^{l+1} - \beta l(l+1). \quad (5)$$

We introduce even and odd amplitudes

$$a_{l+} = \alpha - (\beta+\gamma) l(l+1), \quad (6)$$

$$a_{l-} = \alpha - (\beta-\gamma) l(l+1). \quad (7)$$

The unitarity is imposed in the form

$$0 \leqslant a_{l\pm} \leqslant 1;$$

for $\alpha > 1$ we have

$$a_l = 1 \text{ for } l < L_0,$$

$$a_l = 0 \text{ for } l > L_1.$$

In between, a_l is given by Eqs. (6) and (7). L_0 and L_1 will have different values for $l =$ even and $l =$ odd. We show them with L_0^+, L_1^+ for l even, and L_0^-, L_1^- for l odd. They are given approximately by

$$L_0^+(L_0^+ + 1) = \frac{\alpha-1}{\beta+\gamma}, \quad L_1^+(L_1^+ + 1) = \frac{\alpha}{\beta+\gamma} \quad (8)$$

and

$$L_0^-(L_0^- + 1) = \frac{\alpha-1}{\beta-\gamma}, \quad L_1^-(L_1^- + 1) = \frac{\alpha}{\beta-\gamma}. \quad (9)$$

In Ref. 1, integrals over l were used instead of series, because of their simplicity and because the result does not differ much from the one obtained by series. But since the integral does not distinguish between even and odd l's we have to use

[a)] On leave from the Applied Mathematics Department, University of Western Ontario, London, Ontario, Canada.

[b)] Research supported in part by the National Research Council of Canada.

[c)] Equipe de Recherche Associée au C.N.R.S.

series here, at least until the even and odd parts are separated.

Aside from being longer, the series have their own problems. Since in general the L_0 and L_1's found from Eqs. (8) and (9) will not be integers, there are imprecisions introduced at the transition regions. Also, if one starts the summation of the series in the second region with the integer which follows the last index in the previous region, instead of starting with the same index as one does with the integrals, the results become very complicated because the natural combinations of L's as given by Eqs. (8) and (9) are not obtained and their replacement with the parameters becomes difficult. We therefore have started the series with the same integer as the index of the last term of the previous series. The inaccuracies so introduced can easily be traced by comparing the end results with the two-constraint case and requiring that for $\gamma = 0$ they reduce to the former ones. We have also repeated calculations with integrals after the even and odd parts of the series are separated. The calculations are terribly long, especially for the series. We therefore give only the results. We first define

$$U = \frac{\sqrt{s}}{4k^3} u = \sum_{l=\text{even}} l(l+1)(2l+1)a_l + \sum_{l=\text{odd}} l(l+1)(2l+1)a_l, \tag{10}$$

$$\begin{aligned} U = {} & \sum_{\substack{l=0 \\ \text{even}}}^{L_0'} l(l+1)(2l+1) + \sum_{\substack{l=1 \\ \text{odd}}}^{L_0} l(l+1)(2l+1) \\ & + \sum_{L_0'}^{L_1'} l(l+1)(2l+1)[\alpha - (\beta+\gamma)l(l+1)] \\ & + \sum_{L_0}^{L_1} l(l+1)(2l+1)[\alpha - (\beta-\gamma)l(l+1)]. \end{aligned} \tag{11}$$

After long calculations we find

$$U = \frac{1}{6}\frac{\beta^2+\gamma^2}{(\beta^2-\gamma^2)^2}(3\alpha^2 - 3\alpha + 1) - \frac{4}{3}\frac{\beta}{(\beta^2-\gamma^2)}(2\alpha - 1) + 2. \tag{12}$$

Next we find $g = \sigma^T$ in terms of Lagrange parameters. We define

$$G = \frac{k^2}{4\pi} g = \sum_{\text{even}} (2l+1)a_l + \sum_{\text{odd}} (2l+1)a_l, \tag{13}$$

$$\begin{aligned} G = {} & \sum_{\substack{0 \\ \text{even}}}^{L_0'} (2l+1) + \sum_{\substack{1 \\ \text{odd}}}^{L_0} (2l+1) \\ & + \sum_{\substack{L_0' \\ \text{even}}}^{L_1'} (2l+1)[\alpha - (\beta-\gamma)l(l+1)] \\ & + \sum_{\substack{L_0 \\ \text{odd}}}^{L_1} (2l+1)[\alpha - (\beta-\gamma)l(l+1)]. \end{aligned} \tag{14}$$

Here the result is much simpler

$$G = \frac{\beta}{2(\beta^2-\gamma^2)}(2\alpha - 1) - 1. \tag{15}$$

Next we find $h = \sigma_{\text{el,im}}$. We define

$$H = \frac{k^2}{4\pi} h = \sum_{\text{even}} (2l+1)a_l^2 + \sum_{\text{odd}} (2l+1)a_l^2, \tag{16}$$

$$\begin{aligned} H = {} & \sum_{\substack{0 \\ \text{even}}}^{L_0'} (2l+1) + \sum_{\substack{1 \\ \text{odd}}}^{L_0} (2l+1) \\ & + \sum_{\substack{L_0' \\ \text{even}}}^{L_1'} (2l+1)[\alpha - (\beta+\gamma)l(l+1)]^2 \\ & + \sum_{\substack{L_0 \\ \text{odd}}}^{L_1} (2l+1)[\alpha - (\beta-\gamma)l(l+1)]^2. \end{aligned} \tag{17}$$

The result is

$$H = \frac{3\alpha - 2}{3}\frac{\beta}{\beta^2-\gamma^2} + \frac{8}{3}\alpha - 2\beta - \frac{7}{3}. \tag{18}$$

Finally we find v. We define

$$V = \frac{4k^3}{\sqrt{s}} v = \sum l(l+1)(2l+1)(-1)^{l+1}a_l, \tag{19}$$

$$\begin{aligned} V = {} & -\sum_{\substack{0 \\ \text{even}}}^{L_0'} l(l+1)(2l+1) + \sum_{\substack{1 \\ \text{odd}}}^{L_0} l(l+1)(2l+1) \\ & - \sum_{L_0}^{L_1} l(l+1)(2l+1)[\alpha - (\beta+\gamma)l(l+1)] \\ & + \sum_{\substack{L_0 \\ \text{odd}}}^{L_1} l(l+1)(2l+1)[\alpha - (\beta-\gamma)l(l+1)]. \end{aligned} \tag{20}$$

The result is

$$V = \frac{1}{3}\frac{\beta\gamma}{(\beta^2-\gamma^2)^2}(3\alpha^2 - 3\alpha + 1) - \frac{4}{3}\frac{\beta}{(\beta^2-\gamma^2)}(2\alpha - 1) - 2. \tag{21}$$

Equations (12), (15), (18), and (21) are the results obtained with the series for the case $\alpha > 1$. After separating the series into the even and odd parts we also calculate the same quantities by integrating over l and find

$$U = \frac{\beta^2+\gamma^2}{6(\beta^2-\gamma^2)^2}(3\alpha^2 - 3\alpha + 1) - 1, \tag{22}$$

$$G = \frac{\beta}{2(\beta^2-\gamma^2)}(2\alpha - 1) - 1, \tag{23}$$

$$H = \frac{1}{3}\frac{\beta}{(\beta^2-\gamma^2)}(3\alpha - 2) - 1, \tag{24}$$

$$V = \frac{1}{6}\frac{\beta\gamma}{(\beta^2-\gamma^2)^2}(3\alpha^2 - 3\alpha + 1) - 1. \tag{25}$$

Equations (12), (15), (18) and (21) and (22), (23), (24), and (25) should be compared with the formulas for the two-constraint case

$$U = \frac{3\alpha^2 - 3\alpha + 1}{6\beta^2} - \frac{2\alpha - 1}{3\beta}, \qquad U = \frac{3\alpha^2 - 3\alpha + 1}{6\beta^2}, \tag{26}$$

$$G = \frac{2\alpha - 1}{2\beta}, \qquad G = \frac{2\alpha - 1}{2\beta}, \tag{27}$$

$$H = \frac{3\alpha - 2}{3\beta} + \frac{2\alpha - 1}{3}, \qquad H = \frac{3\alpha - 2}{3\beta},$$
$$V = 0, \qquad V = 0. \tag{28}$$

Here the first set is obtained with series and the second set with integrals. Even though these equations also are not exact, one can see the additional inaccuracies introduced (-1 terms) by the separation of the even and odd amplitudes. These we eliminate by requiring that for $\gamma = 0$ the three-constraint equations reduce to the two-constraint equations. The equations obtained from series are too complicated to eliminate the Lagrange parameters. We therefore use the results obtained with integrals. This set is

$$U = \frac{1}{6}\frac{\beta^2 + \gamma^2}{(\beta^2 - \gamma^2)^2}(3\alpha^2 - 3\alpha + 1), \tag{29}$$

$$V = \frac{1}{6}\frac{2\beta\gamma}{(\beta^2 - \gamma^2)^2}(3\alpha^2 - 3\alpha + 1), \tag{30}$$

$$G = \frac{\beta}{2(\beta^2 - \gamma^2)}(2\alpha - 1), \tag{31}$$

$$H = \frac{\beta}{3(\beta^2 - \gamma^2)}(3\alpha - 2). \tag{32}$$

Elimination of the Lagrange parameters leads to the relation

$$U = \frac{1}{2}(4G^2 - 6GH + 3H^2) + \frac{1}{2}\,\frac{1}{4G^2 - 6GH + 3H^2}V^2. \tag{33}$$

For the case $V = 0$ we have only the first term, which is exactly Eq. (12) of Ref. 1 when we replace U, G, and H by their definitions (1), (2), and (3), together with (10), (13), and (16). In terms of physical quantities, Eq. (33) is equivalent to

$$\left.\frac{d\ln A}{dt}\right|_{z=+1} = \frac{1}{32\pi\sigma^T}\left[4(\sigma^T)^2 - 6\sigma^T\sigma_{\rm el,im} + 3\sigma^2_{\rm el,im}\right] + \frac{128\pi^3}{k^2 s\sigma^T}\left[\frac{1}{4(\sigma^T)^2 - 6\sigma^T\sigma_{\rm el,im} + 3\sigma^2_{\rm el,im}}\right]\left(\left.\frac{dA}{dt}\right|_{z=-1}\right)^2, \tag{34}$$

where A is the imaginary part of the amplitude.

III. BOUND WHEN $\alpha < 1$

We now consider the case $\alpha < 1$. In this case

$$a_{l+} = \alpha - (\beta + \gamma)l(l+1) \quad \text{for } l < L_1^+, \tag{35}$$

$$a_{l-} = \alpha - (\beta - \gamma)l(l+1) \quad \text{for } l < L_1^-. \tag{36}$$

L_1^+ and L_1^- are determined by setting both relations equal to 0 with the same considerations we discussed in Sec. II. We form

$$U = \sum_{\substack{l=0\\ \text{even}}}^{L_1^+} l(l+1)(2l+1)\,[\alpha - (\beta + \gamma)l(l+1)] + \sum_{\substack{l=1\\ \text{odd}}}^{L_1^-} l(l+1)(2l+1)\,[\alpha - (\beta - \gamma)l(l+1)]. \tag{37}$$

The result obtained by summing the series is

$$U = \frac{1}{6}\frac{\beta^2 + \gamma^2}{(\beta^2 - \gamma^2)^2}\alpha^3 - \frac{4}{3}\frac{\beta}{\beta^2 - \gamma^2}\alpha^2 + 2\alpha. \tag{38a}$$

Next we find G:

$$G = \sum_{\substack{l=0\\ \text{even}}}^{L_1^+} (2l+1)[\alpha - (\beta + \gamma)l(l+1)] + \sum_{\substack{l=1\\ \text{odd}}}^{L_1^-} (2l+1)[\alpha - (\beta - \gamma)l(l+1)]. \tag{38b}$$

Summation of the series gives

$$G = \frac{1}{2}\frac{\beta}{\beta^2 - \gamma^2}\alpha^2 - \alpha. \tag{39}$$

Next we calculate H:

$$H = \sum_{\substack{l=0\\ \text{even}}}^{L_1^+} (2l+1)[\alpha - (\beta + \gamma)l(l+1)]^2 + \sum_{\substack{l=1\\ \text{odd}}}^{L_1^-} (2l+1)[\alpha - (\beta - \gamma)l(l+1)]^2. \tag{40}$$

Again the series gives

$$H = \frac{1}{3}\frac{\beta}{\beta^2 - \gamma^2}\alpha^3 + \tfrac{1}{3}\alpha^2 - 2\alpha\beta. \tag{41}$$

Finally V is given by

$$V = -\sum_{\substack{l=0\\ \text{even}}}^{L_1^+} l(l+1)(2l+1)\,[\alpha - (\beta + \gamma)l(l+1)] + \sum_{\substack{l=1\\ \text{odd}}}^{L_1^-} l(l+1)(2l+1)\,[\alpha - (\beta - \gamma)l(l+1)]. \tag{42}$$

Summing the series we find

$$V = \frac{1}{6}\frac{2\beta\gamma}{(\beta^2 - \gamma^2)^2}\alpha^3 - \frac{4}{3}\frac{\gamma}{\beta^2 - \gamma^2}\alpha^2. \tag{43}$$

Next we calculate U, G, H, and V with integrals, after we separate the even and odd partial waves. The results are

$$U = \frac{1}{6}\frac{\beta^2 + \gamma^2}{(\beta^2 - \gamma^2)^2}\alpha^3 - \alpha - \tfrac{4}{3}(\beta - \gamma), \tag{44}$$

$$G = \frac{1}{2}\frac{\beta}{\beta^2 - \gamma^2}\alpha^2 - \alpha + (\beta - \gamma), \tag{45}$$

$$H = \frac{1}{3}\frac{\beta}{\beta^2 - \gamma^2}\alpha^3 + 2(\beta - \gamma)[\alpha - \tfrac{2}{3}(\beta - \gamma)], \tag{46}$$

$$V = \frac{1}{6}\frac{2\beta\gamma}{(\beta^2 - \gamma^2)^2}\alpha^3 - \alpha + \tfrac{4}{3}(\beta - \gamma). \tag{47}$$

These equations are still preliminary and contain the inaccuracies discussed before. They should be compared with the two-constraint case results obtained both from the series and the integrals given below:

$$U = \frac{1}{6}\frac{\alpha^3}{\beta^2} - \frac{1}{3}\frac{\alpha^2}{\beta}, \qquad U = \frac{1}{6}\frac{\alpha^3}{\beta^2}, \tag{48}$$

$$G = \frac{1}{2}\frac{\alpha^2}{\beta}, \qquad G = \frac{1}{2}\frac{\alpha^2}{\beta}, \tag{49}$$

$$H = \frac{1}{3}\frac{\alpha^3}{\beta} + \frac{1}{3}\alpha^2, \qquad H = \frac{1}{3}\frac{\alpha^3}{\beta}. \tag{50}$$

The requirement that the three-constraint equations should reduce to the two-constraint equations for $\gamma = 0$ gives the following sets:

$$U = \frac{1}{6}\frac{\beta^2+\gamma^2}{(\beta^2-\gamma^2)^2}\alpha^3 - \frac{1}{3}\frac{\beta}{\beta^2-\gamma^2}\alpha^2, \qquad U = \frac{1}{6}\frac{\beta^2+\gamma^2}{(\beta^2-\gamma^2)^2}\alpha^3, \tag{51}$$

$$G = \frac{1}{2}\frac{\beta}{(\beta^2-\gamma^2)}\alpha^2, \qquad G = \frac{1}{2}\frac{\beta}{\beta^2-\gamma^2}\alpha^2, \tag{52}$$

$$H = \frac{1}{3}\frac{\beta}{\beta^2-\gamma^2}\alpha^3 + \frac{1}{3}\alpha^2, \quad H = \frac{1}{3}\frac{\beta}{\beta^2-\gamma^2}\alpha^3, \tag{53}$$

$$V = \frac{1}{6}\frac{2\beta\gamma}{(\beta^2-\gamma^2)^2}\alpha^3 - \frac{1}{3}\frac{\gamma}{(\beta^2-\gamma^2)^2}, \qquad V = \frac{1}{6}\frac{2\beta\gamma}{(\beta^2-\gamma^2)^2}\alpha^3. \tag{54}$$

The set on the left is too complicated to eliminate the Lagrange multipliers explicitly. But they give, in principle, a relation between U, G, H, and V to determine a lower bound. The second set can be solved explicitly and we find

$$U = \frac{4}{9}\frac{G^3}{H} + \frac{9}{16}\frac{HV^2}{G^3}. \tag{55}$$

When we replace the definitions of U, G, H, and V in this equation it is equivalent to

$$\left.\frac{d\ln A}{dt}\right|_{z=+1} = \frac{1}{36\pi}\frac{(\sigma^T)^2}{\sigma_{\text{el,im}}} + \frac{9(4\pi)^3}{4k^2s}\frac{\sigma_{\text{el,im}}}{(\sigma^T)^4}\left(\left.\frac{dA}{dt}\right|_{z=-1}\right)^2. \tag{56}$$

When the third constraint does not exist, this equation reduces to the first term of Eq. (15) of Ref. 1, because (56) was derived from the simple second set (51), (52), (53), and (54), whereas Eq. (15) of Ref. 1 is derived from series.

Our main results are Eqs. (34) and (56). Even though we have more precise results, the elimination of the Lagrange parameters is complicated for them. The two equations above can be used as bounds either for the forward or the backward slopes provided the other one is known. One possibility is to combine the global quantities like σ^T with the values of some constraint calculated from phase shifts and treat it as a parameter to check the consistency. As a bound for $d\ln A/dt\,|_{z=+1}$ (the way it is written, $\sigma_{\text{el,im}}$ cannot be replaced in Eq. (56) by the experimentally known quantity σ_{el} by using $\sigma_{\text{el}} > \sigma_{\text{el,im}}$ as was possible with two constraints only, because one of the $\sigma_{\text{el,im}}$'s is in the denominator, whereas the other is in the numerator. But if we consider Eq. (56) a bound for $((dA/dt)|_{z=-1})^2$, when the equation is solved for this, $\sigma_{\text{el,im}}$ can be replaced by σ_{el}, because it appears only in the denominators. Now the value of $((dA/dt)|_{z=-1})^2$ as evaluated from phase shifts can be compared with the bound containing σ^T, σ_{el} and $(d\ln A/dt)|_{z=+1}$.

ACKNOWLEDGMENTS

I would like to gratefully acknowledge the hospitality and the support extended to me by the Université de Nice, the Laboratoire de Physique Théorique, and the director of the laboratoire, Professor François Rocca. In particular, I would like to thank Professor M. Le Bellac for making it possible for me to spend my sabbatical leave at his institute.

[1]S. W. MacDowell and A. Martin Phys. Rev. B 135, 960 (1964).

Symmetries of the stationary Einstein–Maxwell field equations. VII. Charging transformations[a)]

William Kinnersley
Department of Physics, Montana State University, Bozeman, Montana 59717

(Received 30 May 1979; accepted for publication 20 September 1979)

The group **B** of transformations which preserve asymptotic flatness is enlarged to include "charging", and other electromagnetic transformations. The new group **B**′ can be applied to the Tomimatsu–Sato solution to produce an Einstein–Maxwell solution with nine free parameters.

1. INTRODUCTION

In paper II of this series,[1] we found the complete symmetry group **K**′ of the stationary axially-symmetric Einstein–Maxwell equations. We showed that **K**′ has an infinite number of generators $\gamma_{AB}^{(k)}, c_A^{(k)}, \sigma^{(k)}, \tau^{(k)}$, each corresponding to a nonlinear transformation on an infinite set of potentials $K^{(m,n)}, L_B^{(m,n)}, M_A^{(m,n)}, N_{AB}^{(m,n)}$. Since that time we have devoted our attention mainly to the subgroup $\mathbf{K} = \{\gamma_{AB}^{(k)}\}$ which preserves vacuum.

In III, we showed that some of the generators of **K** could be "exponentiated", producing certain finite transformations of the group. These transformations were applied to flat space to generate a series of new vacuum solutions, but unfortunately none of them were asymptotically flat.

In IV, we found an infinite subgroup $\mathbf{B}\subset\mathbf{K}$ of transformations which preserve asymptotic flatness. Although the exponentiation of the **B** transformations could not be done in general, we were able to do it in several particular cases, by making a simple choice for the initial metric. Thus, we used **B** to generate Kerr from Schwarzschild, and we also produced a new five-parameter generalization of the $\delta = 2$ Tomimatsu–Sato solution.

In the present paper, we will return to consider the other transformations of **K**′ which do not preserve vacuum. These may be used to generate various electrified solutions. In particular, we will discuss a subgroup $\mathbf{B}'\subset\mathbf{K}'$ which preserves asymptotic flatness, and show how it may be used to generate an electrovac generalization of Tomimatsu–Sato with *nine arbitrary parameters*.

2. CHARGING TRANSFORMATIONS

The action of the infinitesimal generators of **K**′ was given originally in Eqs. (II.3.1)–(II.3.3) and Eq. (III.3.2). For example we had

$$\gamma_{AB}^{(k)} : K^{(m,n)} \to K^{(m,n)} + \gamma^{(k)XY} \sum_s L_X^{(m,n)} M_Y^{(k-s,n)}, \tag{2.1}$$

$$c_A^{(k)} : K^{(m,n)} \to K^{(m,n)} + c^{*(k)X}\Big[M_X^{(m+k-1,n)} + 2i \sum_s K^{(m,s)} M_X^{(k-s,n)}\Big] + c^{(k)X}\Big[L_X^{(m,n+k-1)} - 2i \sum_s L_X^{(m,s)} K^{(k-s,n)}\Big], \tag{2.2}$$

$$\sigma^{(k)} : K^{(m,n)} \to K^{(m,n)} + i\sigma^{(k)} K^{(m+k,n)} - i\sigma^{(k)} K^{(m,n+k)} - 2\sigma^{(k)} \sum_s K^{(m,s)} K^{(k+s-1,n)}, \tag{2.3}$$

with similar equations for the transformation of the other potentials. It is important to note that the ranges of summation are $s = 1,...,k-1$ in the second sum of Eq. (2.2), and $s = 1,...,k$ in all the others. This means that $K^{(0,n)}$ and $L_B^{(0,n)}$ do not ever occur in the transformations, and hence their values may be disregarded.

However it is sometimes convenient to formally include them. For this purpose we choose to set all $K^{(0,n)} = L_B^{(0,n)} = 0$ except for $K^{(0,1)} = \frac{1}{2}i$. With this convention, the same transformation equations may be extended to hold for $m = 0$ as well. The $m = 0$ equations then vanish identically, and their only function is to guarantee that the assigned values of $K^{(0,n)}, L_B^{(0,n)}$ will remain unchanged.

Exponentiation of $c_2^{(k)}$

Although the action of each generator encompases the entire infinite set of potentials, these exists in certain cases an "invariant subspace". That is, there will be a finite subset of potentials on which the action is closed. It is then possible for us to calculate a general expression for the exponentiation. This happens for $\gamma_{22}^{(k)}$ and $c_2^{(k)}$, and having done the first one in III, we now discuss the second. The transformation may be conveniently written in matrix form. Define the $(3k-1)\times(3k-1)$ matrix

$$P = \begin{pmatrix} N_{11}^{(0,1)} & \cdots & N_{11}^{(0,2k-1)} & M_1^{(0,1)} & \cdots & M_1^{(0,k)} \\ \vdots & & \vdots & \vdots & & \vdots \\ N_{11}^{(2k-2,1)} & \cdots & N_{11}^{(2k-2,2k-1)} & M_1^{(2k-2,1)} & \cdots & M_1^{(2k-2,k)} \\ L_1^{(0,1)} & \cdots & L_1^{(0,2k-1)} & K^{(0,1)} & \cdots & K^{(0,k)} \\ \vdots & & \vdots & \vdots & & \vdots \\ L_1^{(k-1,1)} & \cdots & L_1^{(k-1,2k-1)} & K^{(k-1,1)} & \cdots & K^{(k-1,k)} \end{pmatrix}. \tag{2.4}$$

[a)]Supported by National Science Foundation Grants PHY76-12246 and PHY78-12294.

(We have included $K^{(0,k)}, L_1^{(0,k)}$, to make the matrix square.) The infinitesimal transformation may be written

$$\frac{dP}{d\lambda} = PAP + BP + PC, \tag{2.5}$$

where A, B, C are constant matrices. As we found in IV, the corresponding finite transformation is

$$P \rightarrow e^B P [I - DP]^{-1} e^C, \tag{2.6}$$

where

$$D = \int_0^1 e^{\lambda C} A e^{\lambda B}\, d\lambda. \tag{2.7}$$

For $k = 1$, the matrices are

$$P = \begin{pmatrix} N_{11}^{(0,1)} & M_1^{(0,1)} \\ L_1^{(0,1)} & K^{(0,1)} \end{pmatrix},$$

$$A = \begin{pmatrix} 0 & 0 \\ 2ic^* & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ c^* & 0 \end{pmatrix}, \tag{2.8}$$

$$C = \begin{pmatrix} 0 & c \\ 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 \\ 2ic^* & icc^* \end{pmatrix},$$

and we find

$$\begin{aligned} M_1^{(0,1)} &\rightarrow \Delta^{-1}[M_1^{(0,1)} + cN_{11}^{(0,1)}], \\ N_{11}^{(0,1)} &\rightarrow \Delta^{-1}[N_{11}^{(0,1)}], \end{aligned} \tag{2.9}$$

where

$$\Delta = \det[I - DP] = 1 - 2ic^* M_1^{(0,1)} - icc^* N_{11}^{(0,1)}. \tag{2.10}$$

This is the Harrison transformation, already discussed several times elsewhere.[2] It is a charging transformation that may be applied to any stationary solution. For example, it turns Schwarzschild into Reissner–Nordstrom.

For $k = 2$,

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & -2ic^* \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2ic^* & 0 & 0 & 0 \\ 2ic^* & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & c^* & 0 & 0 & 0 \\ 0 & 0 & c^* & 0 & 0 \end{pmatrix}, \tag{2.11}$$

$$C = B^+,$$

$$D = \begin{pmatrix} 0 & 0 & -icc^* & 0 & -2ic^* \\ 0 & icc^* & 0 & 0 & 0 \\ icc^* & 0 & 0 & 0 & 0 \\ 0 & 2ic^* & 0 & 0 & 0 \\ 2ic^* & 0 & 0 & 0 & 0 \end{pmatrix}.$$

To simplify the result, we now assume that the initial metric is vacuum, with $K^{(m,n)} = L_1^{(m,n)} = M_1^{(m,n)} = 0$. We find

$$M_1^{(0,1)} \rightarrow \Delta^{-1}[cN_{11}^{(0,2)} - ic^2c^*(N_{11}^{(0,1)}N_{11}^{(2,2)} - N_{11}^{(0,2)}N_{11}^{(2,1)})], \tag{2.12}$$

$$N_{11}^{(0,1)} \rightarrow \Delta^{-1}[N_{11}^{(0,1)} - icc^*(N_{11}^{(0,1)}N_{11}^{(1,2)} - N_{11}^{(0,2)}N_{11}^{(1,1)})],$$

where

$$\begin{aligned} \Delta = 1 &- icc^*(N_{11}^{(1,2)} - N_{11}^{(2,1)} + N_{11}^{(0,3)}) - c^2c^{*2}(N_{11}^{(0,1)}N_{11}^{(2,3)} \\ &- N_{11}^{(2,1)}N_{11}^{(0,3)} + N_{11}^{(1,2)}N_{11}^{(0,3)} - N_{11}^{(0,2)}N_{11}^{(1,3)} + N_{11}^{(1,1)}N_{11}^{(2,2)} \\ &- N_{11}^{(2,1)}N_{11}^{(1,2)}) + ic^3c^{*3}(N_{11}^{(0,1)}N_{11}^{(1,2)}N_{11}^{(2,3)} + \cdots). \end{aligned} \tag{2.13}$$

The last parenthesis, not written out, is $\det(N_{11}^{(m,n)})$.

It is especially interesting to see what happens when the initial metric is flat space. Using the values of the potentials $N_{11}^{(m,n)}$ for flat space derived in previous papers, we find that $c_2^{(1)}$ reduces to an electromagnetic gauge transformation

$$M_1^{(0,1)} \rightarrow \text{const}, \tag{2.14}$$

while $c_2^{(2)}$ generates the static electrovac solution

$$\begin{aligned} M_1^{(0,1)} &\rightarrow -2icz\Delta^{-1}, \\ N_{11}^{(0,1)} &\rightarrow -i\Delta^{-1}, \\ \Delta &= 1 + 2cc^*(\rho^2 - 2z^2) + c^2c^{*2}\rho^4. \end{aligned} \tag{2.15}$$

This solution is closely related to a twisting vacuum solution found earlier. It is the Bonnor transform[3] of Eq. (III.6.3).

Higher $c_2^{(k)}$'s may also be exponentiated in this way, but the complexity rapidly increases with k, and the transformations all lead to solutions which are not asymptotically flat.

3. THE SUBGROUP B′

We found in IV that the transformations

$$\beta^{(k)} = \gamma_{22}^{(k+2)} + \gamma_{11}^{(k)} \tag{3.1}$$

preserve asymptotic flatness in the vacuum case. Furthermore, the description of their action does not require the full set of potentials, but takes place instead on a smaller set which are certain linear combinations of the $N_{AB}^{(m,n)}$. We would now like to extend these ideas to include electromagnetism. When applied to flat space, the infinitesimal transformations $c_A^{(k)}$ produce weak electromagnetic fields. These are found to be

$$\begin{aligned} c_1^{(k)} &: \varphi_1 \rightarrow -ic(2r)^k P_k(\cos\theta), \\ c_2^{(k)} &: \varphi_1 \rightarrow c(2r)^{k-1} P_{k-1}(\cos\theta). \end{aligned} \tag{3.2}$$

Therefore, the linear combinations

$$b^{(k)} = c_2^{(k+1)} - ic_1^{(k)}, \tag{3.3}$$

produce *no* electromagnetic field. They leave flat space invariant; hence they will be transformations which preserve asymptotic flatness.

To clarify the meaning of this assertion, we consider the case $k = 0$.

$$b^{(0)} = c_2^{(1)} - ic_1^{(0)}. \tag{3.4}$$

As we have seen, the first term $c_2^{(1)}$ is the Harrison transformation. But $c_2^{(1)}$ by itself does *not* preserve asymptotic flatness in the present sense. In the flat regions it produces a gauge transformation on the potentials [Eq. (2.14)]. To maintain asymptotic flatness, with all the potentials at their flat space values, we must undo the gauge transformation.

This is accomplished by the second term $-ic_1^{(0)}$, which insures that $M_1^{(0,1)} \to 0$ at spatial infinity.

To form a closed subgroup, we need to include also the $\sigma^{(k)}$. These are some sort of generalized duality rotations, and also have no effect on flat space. From Eq. (II.3.4), the commutation relations of the subgroup $\mathbf{B}' = \{\beta^{(k)}, b^{(k)}, \sigma^{(k)}\}$, are

$$\begin{aligned}
&[\beta^{(k)}, b^{(l)}] = ib^{(k+l+1)},\\
&[\sigma^{(k)}, b^{(l)}] = -ib^{(k+l)},\\
&[b^{*(k)}, b^{(l)}] = 2i\beta^{(k+l-1)} - 6i\sigma^{(k+l)},\\
&[\beta^{(k)}, \beta^{(l)}] = [\beta^{(k)}, \sigma^{(l)}] = [\sigma^{(k)}, \sigma^{(l)}] = 0.
\end{aligned} \tag{3.5}$$

To describe their action we define a "reduced hierarchy" of potentials

$$\begin{aligned}
N_{mn} &= N_{11}^{(m,n)} - iN_{21}^{(m-1,n)} + iN_{12}^{(m,n-1)} + N_{22}^{(m-1,n-1)},\\
M_{mn} &= M_1^{(m,n)} - iM_2^{(m-1,n)},\\
L_{mn} &= L_1^{(m,n)} + iL_2^{(m,n-1)},\\
K_{mn} &= K^{(m,n)},
\end{aligned} \tag{3.6}$$

and for $m = 0$,

$$\begin{aligned}
N_{0n} &= N_{11}^{(0,n)} + iN_{12}^{(0,n-1)},\\
M_{0n} &= M_1^{(0,n)},\\
L_{0n} &= K_{0n} = 0.
\end{aligned} \tag{3.7}$$

(N_{mn} is what we previously called P_{mn} in paper IV).

The infinitesimal transformations are

$$\begin{aligned}
\beta^{(k)}: K_{mn} &\to K_{mn} + (i\beta K_{m+k+1,n}) - i\beta K_{m,n+k+1} + \beta \sum L_{ms} M_{k+2-s,n} - 2\beta \sum K_{ms} K_{k+2-s,n},\\
L_{mn} &\to L_{mn} + (i\beta L_{m+k+1,n}) - 2i\beta L_{m,n+k+1} + \beta \sum L_{ms} N_{k+2-s,n} - 2\beta \sum K_{ms} L_{k+2-s,n},\\
M_{mn} &\to M_{mn} + (2i\beta M_{m+k+1,n}) - i\beta M_{m,n+k+1} + \beta \sum N_{ms} M_{k+2-s,n} - 2\beta \sum M_{ms} K_{k+2-s,n},\\
N_{mn} &\to N_{mn} + (2i\beta N_{m+k+1,n}) - 2i\beta N_{m,n+k+1} + \beta \sum N_{ms} N_{k+2-s,n} - 2\beta \sum M_{ms} L_{k+2-s,n},
\end{aligned} \tag{3.8}$$

$$\begin{aligned}
b^{(k)}: K_{mn} &\to K_{mn} + (b^* M_{m+k,n}) + bL_{m,n+k} + 2ib^* \sum K_{ms} M_{k+1-s,n} - 2ib \sum L_{ms} K_{k+1-s,n},\\
L_{mn} &\to L_{mn} + (b^* N_{m+k,n}) + 4b^* K_{m,n+k} + 2ib^* \sum K_{ms} N_{k+1-s,n} - 2ib \sum L_{ms} L_{k+1-s,n},\\
M_{mn} &\to M_{mn} + (4bK_{m+k,n}) + bN_{m,n+k} + 2ib^* \sum M_{ms} M_{k+1-s,n} - 2ib \sum N_{ms} K_{k+1-s,n},\\
N_{mn} &\to N_{mn} + (4bL_{m+k,n}) + 4b^* M_{m,n+k} + 2ib^* \sum M_{ms} N_{k+1-s,n} - 2ib \sum N_{ms} L_{k+1-s,n},
\end{aligned} \tag{3.9}$$

$$\begin{aligned}
\sigma^{(k)}: K_{mn} &\to K_{mn} + (i\sigma K_{m+k,n}) - i\sigma K_{m,n+k} - 2\sigma \sum K_{ms} K_{k-s+1,n},\\
L_{mn} &\to L_{mn} + (i\sigma L_{m+k,n}) - 2\sigma \sum K_{ms} L_{k-s+1,n},\\
M_{mn} &\to M_{mn} - i\sigma M_{m,n+k} - 2\sigma \sum M_{ms} K_{k-s+1,n},\\
N_{mn} &\to N_{mn} - 2\sigma \sum M_{ms} L_{k-s+1,n}.
\end{aligned} \tag{3.10}$$

Here the various sums run from $s = 1$ up to $s = k + 2$ and $s = k + 1$ respectively in Eq. (3.8), up to $s = k + 1$ and $s = k$ in Eq. (3.9), and up to $s = k$ in Eq. (3.10). The terms in parentheses are absent when $m = 0$.

All of these transformations can be written in the same form as Eq. (2.5), where P is now the infinite matrix

$$P = \begin{pmatrix}
N_{01} & N_{02} & \cdots & M_{01} & M_{02} & \cdots\\
N_{11} & N_{12} & \cdots & M_{11} & M_{12} & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
L_{01} & L_{02} & \cdots & K_{01} & K_{02} & \cdots\\
L_{11} & L_{12} & \cdots & K_{11} & K_{12} & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{pmatrix}. \tag{3.11}$$

However the transformations making up $\mathbf{B}'$ do not have any invariant subspaces. We have not found a general way to exponentiate them, and we must therefore turn again to a discussion of special cases.

4. THE T–S FAMILY

As we already pointed out in IV, the potentials for the $\delta = 2$ Tomimatsu–Sato (T–S) solution recur beyond $m,n = 2$. This feature permits us to replace P by a finite matrix and thus perform the necessary integration. Application of $\mathbf{B}'$ to T–S then leads to a multiparameter family of related metrics.

For nonrotating T–S, we had the potentials

$N_{01} = 4ix(x+1)^{-2}, \quad N_{02} = -8iy(x+1)^{-2},$

$$N_{11} = 16ixy(x+1)^{-2}, \quad N_{21} = -16i(x^2+y^2)(x+1)^{-2}, \tag{4.1}$$
$$N_{12} = 16i(x+1)^{-1}(x^2-1)^{-1}(x^3+x^2+y^2-3xy^2),$$
$$N_{22} = -64iy(x+1)^{-1}(x^2-1)^{-1}(x-y^2),$$

and the recursion relations

$$N_{0,2l+n} = 4^l N_{0n}, \quad n=1,2, \qquad N_{2k+m,2l+n} = 4^{k+l}N_{mn}, \quad m,n=1,2. \tag{4.2}$$

The relations of Eq. (4.2) were shown in IV to be preserved under $\beta^{(k)}$ transformations. Postulating similar relations for K_{mn}, L_{mn}, and M_{mn} one can show that they are preserved under the entire group $\mathbf{B}'$. Thus we are left to deal this time with a 5×5 matrix:

$$P = \begin{pmatrix} 0 & N_{01} & N_{02} & M_{01} & M_{02} \\ 0 & N_{11} & N_{12} & M_{11} & M_{12} \\ 0 & N_{21} & N_{22} & M_{21} & M_{22} \\ 0 & L_{11} & L_{12} & K_{11} & K_{12} \\ 0 & L_{21} & L_{22} & K_{21} & K_{22} \end{pmatrix}. \tag{4.3}$$

(The extra column of zeros has been added for convenience, to once more make the matrix square).

As in the vacuum case, Eqs. (4.2) induce a recurrence among the transformations themselves. We find

$$\beta^{(2p)} = 4^p(1-p)\beta^{(0)} + 4^{p-1}(p)\beta^{(2)}, \qquad \beta^{(2p+1)} = 4^p(1-p)\beta^{(1)} + 4^{p-1}(p)\beta^{(3)}, \tag{4.4}$$

and similarly for $b^{(k)}$, $\sigma^{(k)}$. All we are left with is $\beta^{(k)}$, $b^{(k)}$, $\sigma^{(k)}$, $k=0,\ldots,3$. These generate a 16-parameter Lie group. (Recall b is complex.) However, even this reduced group is not completely effective in generating new solutions, because its action is not simply transitive on the space in question (the parameter-space of the solutions). Each solution in the family is invariant under a "little group", which turns out to have eight parameters. (For T–S itself, the little group consists of $\sigma^{(0)},\ldots,\sigma^{(3)}$ and the combinations $b^{(0)} = -4b^{(2)}$ and $b^{(1)} = -4b^{(3)}$.) Thus there remain eight effective degrees of freedom. When the mass is counted as another parameter, and the unphysical NUT parameter discarded, we expect to be able to generate an eight-parameter asymptotically flat Einstein–Maxwell solution.

For purposes of actual calculation, the simplest transformations to use are the combinations

$$\alpha_0 = \beta^{(0)} + 4\beta^{(2)}, \quad \alpha_1 = \beta^{(1)} + 8\beta^{(3)}, \qquad \alpha_2 = \beta^{(0)} + 8\beta^{(2)}, \quad \alpha_3 = \beta^{(1)} + 4\beta^{(3)}, \tag{4.5}$$

with similar replacements $b^{(n)} \to a_n$ and $\sigma^{(n)} \to s_n$. The action is

$$\frac{dP}{d\lambda} = PAP + BP + PC, \tag{4.6}$$

where

$$A = \begin{pmatrix} 0 & 4\alpha_3 & \alpha_0 & -2i\alpha_0 & -2i\alpha_3 \\ 4\alpha_3 & \alpha_2 & \alpha_1 & -2i\alpha_1 & \frac{1}{2}i(\alpha_0-\alpha_2) \\ \alpha_0 & \alpha_1 & \frac{1}{4}(\alpha_2-\alpha_0) & \frac{1}{2}i(\alpha_0-\alpha_2) & \frac{1}{2}i(\alpha_3-\alpha_1) \\ 2i\alpha_0{}^* & 2i\alpha_1{}^* & \frac{1}{2}i(\alpha_2{}^*-\alpha_0{}^*) & -2(\alpha_2+s_3) & -\frac{1}{2}(4\alpha_1+s_2-s_0) \\ 2i\alpha_3{}^* & \frac{1}{2}i(\alpha_2{}^*-\alpha_0{}^*) & \frac{1}{2}i(\alpha_1{}^*-\alpha_3{}^*) & -\frac{1}{2}(4\alpha_1+s_2-s_0) & \frac{1}{2}(\alpha_0-\alpha_2+s_3-s_1 \end{pmatrix}, \tag{4.7}$$

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 8i\alpha_3 & 2i\alpha_0 & 4a_0 & 4a_3 \\ 0 & 8i\alpha_0 & 8i\alpha_3 & 16a_3 & 4a_0 \\ 0 & a_0{}^* & a_3{}^* & i(4\alpha_3+s_0) & i(\alpha_0+s_3) \\ 0 & 4a_3{}^* & a_0{}^* & i(4\alpha_0+s_3) & i(4\alpha_3+s_0) \end{pmatrix}, \tag{4.8}$$

and $C = B^+$.

The finite transformation is given by Eqs. (2.6) and (2.7). Even though the various transformations do not commute, we may ignore this fact in the integration. We may integrate using an arbitrary constant linear combination of the eight effective transformations.[3] However, for the sake of simplicity, we will not consider the general case. The vacuum subgroup has been studied in IV. The a_0 is the Harrison transformation, and a_1, a_2 belong to the little group, so we consider only the remaining parameter a_3.

Assuming that the initial solution is vacuum (but not necessarily nonrotating) we calculate

$$N_{01} \to \Delta^{-1}\cosh a[N_{01} - \tfrac{1}{16}i(N_{01}N_{12} - N_{02}N_{11})], \tag{4.9}$$
$$M_{01} \to \tfrac{1}{4}\Delta^{-1}\sinh a[N_{02} - \tfrac{1}{16}i(N_{01}N_{22} - N_{02}N_{21})],$$

where

$$\Delta = 1 - \tfrac{1}{2}i(\cosh a - 1)N_{01} + \tfrac{1}{16}i\sinh^2 a(N_{21} - N_{12}), \\ -\tfrac{1}{32}\sinh^2 a(\cosh a - 1)(N_{01}N_{12} - N_{02}N_{11}), \\ +\tfrac{1}{256}\sinh^4 a(N_{21}N_{12} - N_{11}N_{22}), \tag{4.10}$$
$$a = 4a_3.$$

Using the T–S potentials listed in Eq. (4.1), we find

$$N_{01} \to \frac{-2iy\sinh a}{(x\cosh a+1)^2 - y^2\sinh^2 a}, \\ M_{01} \to \frac{4ix\cosh a}{(x\cosh a+1)^2 - y^2\sinh^2 a}, \\ \mathscr{E} = iN_{01} + 1 = \frac{(x\cosh a - 1)^2 - y^2\sinh^2 a}{(x\cosh a+1)^2 - y^2\sinh^2 a}, \tag{4.11}$$

$$f = \mathscr{E} + \phi\phi^*\left(\frac{x^2\cosh^2 a - y^2\sinh^2 a - 1}{(x\cosh a + 1)^2 - y^2\sinh^2 a}\right)^2.$$

This is the Bonnor solution[4] for a massive nonrotating magnetic dipole. Using the potentials for the rotating T–S metric in Eqs. (4.9) and (4.10) will produce a new rotating generalization of this solution.

As shown in V, the Bonnor transformation is a discrete *automorphism* of our continuous group **K**. Its action on the group is closely mirrored in its action on the families of solutions we can produce. Thus, the Bonnor transform of Eq. (4.11) is the Kerr metric. In other words, a_3 applied to the $\delta = 2$ T–S solution is the Bonnor transform of $\beta^{(0)}$ applied to the $\delta = 1$ (Schwarzschild) solution. This type of relation appears to hold in general between T–S solutions for any values of δ and 2δ.

The other metric we would like to use as an example for application of $\mathbf{B}'$ is the Schwarzschild metric. Unfortunately, so far we have not been able to find any simple set of recursion relations for the Schwarzschild potentials which would be preserved by all of $\mathbf{B}'$. Without this step accomplished, neither the group action nor the matrix P can be made finite-dimensional. The remarks in the preceding paragraph tend to explain our difficulty. At least some of the charged solutions which $\mathbf{B}'$ would generate from Schwarzschild are not expected to be simple. They would be Bonnor transforms of those rotating vacuum solutions which **B** would generate for $\delta = \frac{1}{2}$. The T–S solutions for noninteger δ are known to involve unfamiliar transcendental functions rather than polynomials.[5] We expect that the full group $\mathbf{B}'$ can be used without such difficulty whenever δ is even.

[1] W. Kinnersley, J. Math. Phys. **18**, 1529 (1977) (I); W. Kinnersley and D.M. Chitre, J. Math. Phys. **18**, 1538 (1977) (II); **19**, 1926 (1978) (III); **19**, 2037 (1978) (IV); C. Hoenselaers, J. Math. Phys. **20**, 2529 (1979) (V); C. Hoenselaers, W. Kinnersley, and B.C. Xanthopoulos, J. Math. Phys. **20**, 2530 (1979) (VI); W. Kinnersley and D.M. Chitre, Phys. Rev. Lett. **40**, 1608 (1978); C. Hoenselaers, W. Kinnersley, and B.C. Xanthopoulos, Phys. Rev. Lett. **42**, 481 (1979).

[2] B.K. Harrison, J. Math. Phys. **9**, 1744 (1968); W. Kinnersley, J. Math. Phys. **14**, 651 (1973).

[3] An analogous situation is the three-dimensional rotation group acting on the unit sphere. The sphere can be generated from a single point (the North Pole) by exponentiating fixed linear combinations of J_x and J_y, thus sweeping down along each meridian.

[4] W. Bonnor, Z. Phys. **190**, 444 (1966); Z. Perjes, Nuovo Cimento **55B**, 600 (1968); R. Misra *et al.*, Phys. Rev. D 7, 1587 (1973); W. Kinnersley, in *General Relativity and Gravitation*, edited by G. Shaviv (Wiley, New York, 1975).

[5] C. Cosgrove, J. Phys. A **10**, 1481 (1977).

Canonical connections on Riemannian symmetric spaces and solutions to the Einstein–Yang–Mills equations [a]

J. Harnad and J. Tafel [b]
Centre de Recherches de Mathématiques Appliqués, Université de Montréal, Case Postale 6128, Succ. "A", Montréal H3C 3J7, Canada
S. Shnider
Department of Mathematics, McGill University, 805 Sherbrooke West Montreal H3A2K6 Canada

(Received 21 February 1980; accepted for publication 4 April 1980)

It is shown that for any principal bundle over a Riemannian symmetric space G/G_0 which admits G as automorphism group, the canonical G-invariant connection satisfies the source free gauge field equations. Extending this to product manifolds $V\times G/G_0$ and assuming the metric and gauge fields decompose in a natural way, this result is still valid and the Einstein equations with gauge fields as source may also be satisfied. For G/G_0, this is so automatically, but with a cosmological term present. For $\dim V = 1$ or 2, solutions are found, yielding metrics of the Robertson–Walker and Reissner–Nordstrom type.

1. INTRODUCTION

Nowakowski and Trautman[1,2] have shown that certain natural geometric structures give rise automatically to solutions of the source free gauge field equations. Specifically, if Ω is the curvature of the canonical connections on a Stiefel bundle $V_{n,q}(F)\rightarrow G_{n,q}(F)$ then

$$D\,{}^*\Omega = 0, \qquad (1)$$

where the Hodge operation is taken with respect to the canonical metric. Here F denotes the real numbers R, the complex numbers C, or the quaternions H,

$$V_{n,q}(F) = U_n(F)/U_{n-q}(F)$$

and

$$G_{n,q}(F) = U_n(F)/U_{n-q}(F)\times U_q(F)$$

are the Steifel and Grassmann manifolds, with $U_n(F) = \mathrm{SO}(n)$, $\mathrm{U}(n)$, or $\mathrm{Sp}(n)$ corresponding to $F = R$, C, or H, respectively.

Since the $G_{p,q}(F)$ are all particular types of Riemannian symmetric spaces, it is natural to inquire whether this result may be generalized. In Sec. II it will be shown that this is indeed the case.

If G/G_0 is a Riemannian symmetric space, with G the full group of isometries, then the curvature of the canonical G-invariant connection on the G_0-bundle $G\rightarrow G/G_0$ always satisfies (1). It follows that if G_0 decomposes as $G_0 = G_1\times G_2$, the induced canonical connection on the bundle $G/G_1\rightarrow G/G_1\times G_2$ obtained by projection also satisfies (1). In fact, on any G-homogeneous principal bundle $E\rightarrow G/G_0$ there is a canonical G-invariant connection which we show satisfies (1).

More generally, we may consider a manifold which decomposes as a product $V\times G/G_0$ with G acting only the second factor. Assume the metric has the form

$$g = g_V \oplus f^2 g_M, \qquad (2)$$

where g_V is a metric on V, g_M a G-invariant metric on G/G_0, and f a scalar function on V. Such a decomposition of the metric will always hold if g is G-invariant and G/G_0 an irreducible Riemannian symmetric space, as will be shown in Sec. II. The connection on the bundle $V\times E\rightarrow V\times G/G_0$ obtained by pulling back the canonical connection form on E still satisfies (1) for such a metric, the dependence on directions along V arising only in the metric and not in the gauge fields.

One may next inquire whether g_V and f in (2) may be determined as solutions to the Einstein equations

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = T_{\mu\nu} + \lambda g_{\mu\nu}\,, \qquad (3)$$

with the energy–momentum tensor $T_{\mu\nu}$ of the canonical gauge field as source.

For a metric of the form (2), the Ricci tensor $R_{\mu\nu}$ and the energy–momentum tensor $T_{\mu\nu}$ of the canonical gauge field also split into block diagonal form. If G is semisimple and g_M is (up to a constant) the metric on G/G_0 corresponding to the Killing form on G, then the G/G_0 block for both $R_{\mu\nu}$ and $T_{\mu\nu}$ is a proportional to g_M. Therefore, if V reduces to a point, Eqs. (3) are automatically satisfied for suitable choice of the cosmological constant λ. For a product $V\times G/G_0$, Eq. (3) with $\lambda = 0$ may be solved to determine the metric.

This is done in Sec. 3 for the case $\dim V = 1$ or 2, the resulting metrics being of the Robertson–Walker or Reissner–Nordstrom type, respectively. In low dimensions, the number of inequivalent Riemannian symmetric spaces is very small because of the degeneracies between the different types of low dimensional Lie algebras. The cases with $\dim V + \dim G/G_0 = 4$ are given in Sec. 4, as well as some dicussion of the applicability of the results for higher dimensional spaces to space–time.

[a] Permanent address: Institute of Theoretical Phyics, Warsaw University, Hoza 69, Warsaw, Poland.
[b] Research supported in part by the National Sciences and Engineering Council of Canada and le Ministère de l'Education du Gouvernement du Québec.

2. CANONICAL CONNECTIONS AND GAUGE FIELD EQUATIONS

For our purposes, a Riemannian symmetric space (RSS) will be taken to mean a Riemannian manifold M with a transitive group of isometries G such that, identifying M with G/G_0, the Lie algebra $\mathfrak{g}$ of G admits a symmetric decomposition

$$\mathfrak{g} = \mathfrak{g}_0 + \mathcal{M}, \quad (4)$$

$$[\mathfrak{g}_0,\mathfrak{g}_0] \subset \mathfrak{g}_0, \quad \mathrm{Ad}G_0\mathcal{M} \subset \mathcal{M}, \quad [\mathcal{M},\mathcal{M}] \subset \mathfrak{g}_0, \quad (5)$$

where $\mathfrak{g}_0$ is the Lie algebra of the isotropy subgroup G_0 at an arbitrarily chosen origin 0 and $\mathcal{M}$ is an $\mathrm{Ad}G_0$ invariant subspace complimentary to $\mathfrak{g}_0$ in $\mathfrak{g}$. The space $\mathcal{M}$ may be identified in a standard way[3] with the tangent space TM_0. Through this identification, the metric on M determines an $\mathrm{Ad}G_0$ invariant inner product on $\mathcal{M}$ and conversely, any such inner product determines a G-invariant metric. All the results of the present section are equally valid for indefinite metrics, but we shall mainly be interested in the positive definite case.

Regarding the projection $G \to G/G_0$ as defining a principal G_0-bundle over M, the canonical connection form ω_0 is defined[3] as the $\mathfrak{g}_0$ part in the decomposition of the left-invariant Maurer–Cartan form on G.

$$\omega_{MC} = g^{-1}dg = \omega_0 + \omega_1, \quad (6)$$

where the complement ω_1 has values in $\mathcal{M}$. Given any $\mathrm{Ad}G_0$ invariant inner product K on $\mathcal{M}$, the corresponding G-invariant metric g_M on M may be defined by

$$g_M(X,Y) = K(\theta(X),\theta(Y)), \quad x,y \in TM_p, \quad (7)$$

where $\theta = \sigma^*\omega_1$ is the pull-back of ω_1, under any local section $\sigma: N_p \to G$ defined in a neighborhood N_p of $p \in G/G_0$. The expression is independent of the choice of σ because K is $\mathrm{Ad}G_0$ invariant. Expressing θ relative to a basis $\{T_i\}$ in $\mathcal{M}$ as

$$\theta = \sum_{i=1}^{n} \theta^i T_i, \quad (8)$$

defines a left-invariant co-frame $\{\theta^i\}$. Provided $\mathrm{Ad}G_0$ restricted to $\mathcal{M}$ has positive determinant, this determines a σ-independent orientation on G/G_0 and hence with the metric, a globally defined left-invariant volume form Λ. More generally, if G/G_0 is orientable, the sections σ may always be chosen so that for the transition functions this determinant is positive. If M is not orientable, it is always possible to replace it with an orientable G-homogenous covering space by replacing G_0 by its connected component. Alternatively, Λ can be regarded as defined only up to a sign. The metric and volume define, in standard way, the Hodge star * dual. If M is nonorientable this must also be regarded as defined only up to a sign.

Denoting by $\omega^\sigma = \sigma^*\omega_0$ the pull-back of the canonical connection and

$$\Omega^\sigma = d\omega^\sigma + \tfrac{1}{2}[\omega^\sigma,\omega^\sigma], \quad (9)$$

the pull-back of its curvature, we have the following theorem.

Theorem 1:

$$D{*\Omega^\sigma} = d{*\Omega^\sigma} + [\omega,{*\Omega^\sigma}] = 0, \quad (10)$$

where $*\Omega^\sigma$ denotes the dual under any G-invariant metric g_M on G/G_0.

Proof: Because of the Maurer–Cartan structure equations and Eq. (5), we have

$$\Omega^\sigma = -\tfrac{1}{2}[\theta,\theta], \quad (11)$$

and

$$D\theta = d\theta + [\omega^\sigma,\theta] = 0. \quad (12)$$

From (11), we have,

$$*\Omega^\sigma = \mathcal{P}(\theta,\ldots,\theta) \quad n-2 \text{ terms}, \quad (13)$$

where $\mathcal{P}$ is an alternating multilinear map

$$\mathcal{P}: \Lambda^{m-2}\mathcal{M} \to \mathfrak{g}_0,$$

which is $\mathrm{ad}\mathfrak{g}_0$ invariant because of the $\mathrm{Ad}G_0$ invariance of K:

$$[A,\mathcal{P}(A_1,\ldots A_{n-2})] = \sum_{j=1}^{m-2} \mathcal{P}(A_1,\ldots,[A,A_j],\ldots A_{n-2}), \quad \forall A \in \mathfrak{g}_0, \quad A_1,\ldots A_{n-2} \in \mathcal{M}. \quad (14)$$

This map may be expressed explicitly in tensor index notation as

$$\mathcal{P}(T_{i_1}\cdots T_{i_{n-2}}) = -\frac{1}{4(m-2)!}\eta^{ij}_{i_1\cdots i_{m-2}} f^k_{ij} T_k, \quad (15)$$

where η denotes the Levi-Civita tensor on $\mathcal{M}$ with respect to the metric K, with raised indices obtained by contraction with $K^{ij} = (K^{-1})_{ij}$ and f^k_{ij} the structure constants in

$$[T_i,T_j] = f^k_{ij}T_k. \quad (16)$$

The $\mathrm{Ad}G_0$ invariance (14) follows from the fact that $\mathrm{Ad}G_0$ acts on $\mathcal{M}$ by orthogonal transformations preserving the metric K. Using the fact that d is a derivation and (14), we have

$$D{*\Omega^\sigma} = d{*\Omega^\sigma} + [\omega^\sigma,{*\Omega^\sigma}] = \sum_{j=1}^{n-2}(-1)^{j+1}\mathcal{P}(\theta,\ldots \underset{j\text{th term}}{D\theta}\ldots\theta) = 0 \quad \text{Q.E.D.}$$

Now consider the more general case of any principal H-bundle $\pi: E \to G/G_0$ admitting a left G-action which commutes with the right H-action and projects to the standard left action on the base. Such bundles are completely characterized[4,5] by the homomorphism $\varphi: G_0 \to H$ defined by

$$g \cdot q = q\varphi(g_0), \quad g_0 \in G_0, \quad (17)$$

where q is an arbitrarily chosen point $q \in \pi^{-1}(0)$ and may be identified up to an isomorphism, with the bundle E_φ defined by factoring the trivial H bundle $G \times H \to G$ by the equivalence relation

$$(g,h) \sim (gg_0,\varphi(g_0^{-1})h), \quad g \in G, \quad h \in H, \quad g_0 \in G_0. \quad (18)$$

For each such bundle there is a canonical connection defined by the form

$$\omega_\varphi|_{[(g,h)]} = \mathrm{Ad}\,h^{-1}\varphi_* \circ \omega_0|_g + h^{-1}dh, \quad (19)$$

where $[(g,h)]$ is the point in E_φ determined by the equivalence class of (g,h), $\varphi_*: \mathfrak{g}_0 \to \mathfrak{h}$ is the differential φ at (TG_0), which defines a homomorphism to $\mathfrak{h}$, the Lie algebra of H. Although ω_0 is a form on G and $h^{-1}dh$ the Maurer–Cartan form on H, the left and right translation properties of ω_0 imply the definition (19) of ω_0 passes to the quotient of $G\times H$ by (18) giving a well-defined G-invariant connection form on E_φ (see Ref. 5). Any local section $\sigma:N_p\rightarrow G$ corresponds uniquely to a section $\tilde\sigma:N_p\rightarrow E_\varphi$ defined by $\tilde\sigma(p)=[(\sigma(p),e)]$, under which the pull-back of ω_φ is

$$\tilde\sigma^*\omega_\varphi=\varphi_*\circ\omega^\sigma. \tag{20}$$

Since $\varphi_*:\mathfrak{g}_0\rightarrow\mathfrak{h}$ is a homomorphism, Theorem 1 implies the following.

Corollary 1:

$$D^*\Omega_\varphi^\sigma=0, \tag{21}$$

where $\Omega_\varphi^\sigma=\tilde\sigma^*\Omega_\varphi=\varphi_*\circ\Omega^\sigma$ is the pull-back of the curvature

$$\Omega_\varphi=d\omega_\varphi+\tfrac{1}{2}[\omega_\varphi,\omega_\varphi]. \tag{22}$$

As discussed in the introduction by identifying G, G_0, and H as $U_{n,q}(F)$, $U_{n-q}(F)\times U_q(F)$, and $U_q(F)$, respectively, with the obvious homomorphism

$$\varphi:U_{n-q}(F)\times U_q(F)\rightarrow U_q(F),$$

we recover the canonical connection on the Stiefel bundle and Corollary 1 becomes the Nowakowski–Trautman result.[2]

Now consider a product manifold $V\times G/G_0$ with metric g of the form (2). A sufficient condition implying such a decomposition of the metric is that g be G-invariant and G/G_0 an irreducible RSS. To see this, note that the general form of a G-invariant metric is obtained by solving the linear algebraic condition implied by invariance under the isotropy group G_0 along the cross section $\Sigma(v)=(v,eG_0)$, $v\in V$ and then determining g at an arbitrary point (v,bG_0), $b\in G$ by the invariance relation

$$g(b\ \Sigma(v))=(b^{-1})^*g(\Sigma(v)). \tag{23}$$

Expressing the tensor components of g in a co-frame of the type $\{\phi^a,\theta^i\}$, where $\{\phi^a\}_{a=1,\ldots\dim V}$ is a co-frame for V, as

$$g=(\phi^a,\theta^i)\begin{pmatrix}A & B\\ B^t & D\end{pmatrix}\begin{pmatrix}\phi^a\\ \theta^i\end{pmatrix}, \tag{24}$$

the linear isotropy conditions read

$$RB=B,\quad RDR^T=0, \tag{25}$$

where R is the $m\times m$ Jacobian matrix for $g_0\in G_0$ at $(TM)_0$ defined by

$$\mathrm{Ad}g_0(T_i)=\sum_{j=1}^{m}R_{ji}T_j. \tag{26}$$

Because of the $\mathrm{Ad}G_0$ invariance of K, R is (pseudo)-orthogonal

$$R\widetilde{K}R^T=\widetilde{K}, \tag{27}$$

where $\widetilde K$ is the matrix representing K in the $\{T_i\}$ basis. Since (25) must hold for all $g_0\in G_0$, if $\mathrm{Ad}G_0$ restricted to $\mathscr{M}$ is irreducible, we must have $B=0$, $D\propto\widetilde K$ leading to the expression (2) for the metric. This argument is actually valid for any reductive homogeneous space G/G_0 with a G-invariant metric (in fact, a G-invariant metric can *only* exist if G/G_0 is reductive). But if G/G_0 is an irreducible RSS then $\mathrm{Ad}G_0$ restricted to $\mathscr{M}$ is necessarily irreducible[3] and (2) is the most general form for a G-invariant metric on $V\times G/G_0$.

Now denote by the symbols $\widetilde\omega_\varphi,\widetilde\Omega_\varphi$ the canonical connection and curvature forms on $V\times E_\varphi\rightarrow V\times G/G_0$ defined by pulling back ω_φ and Ω_φ from E_φ under the projection $V\times E_\varphi\rightarrow E_\varphi$. If $\bar\sigma$: $V\times N_p\rightarrow V\times E_\varphi$ is the section defined by

$$\bar\sigma(v,p)=(v,\tilde\sigma(p)), \tag{28}$$

and $\widetilde\Omega_\varphi^\sigma\equiv\bar\sigma^*\widetilde\Omega_\varphi$, we have

$$^*\Omega_\varphi^\sigma=f^m\Lambda_v\wedge\varphi_*\circ\mathscr{P}(\theta,\ldots\theta), \tag{29}$$

where Λ_v is the volume form on V corresponding to the metric g_v. Since f is a function on V only, it follows that:

Corollary 2: Equation (21) is still valid if $\widetilde\Omega_\varphi^\sigma$ is substituted for Ω_φ^σ and the dual* is taken with respect to the metric (2).

3. EINSTEIN EQUATIONS

Denoting by $F_{\mu\nu}$ the tensor components of the gauge field, referred to a local frame, the energy–momentum tensor is of the form

$$T_{\mu\nu}=k(F_{\mu\sigma},F_{\nu\tau})g^{\sigma\tau}-\tfrac{1}{4}g_{\mu\nu}h(F_{\kappa\sigma},F_{\lambda\tau})g^{\kappa\lambda}g^{\sigma\tau}, \tag{30}$$

where k is an $\mathrm{Ad}H$ invariant form on $\mathfrak{h}$. We shall henceforth specialize to the case where the manifold $V\times G/G_0$ with metric (2), the bundle is $V\times E_\varphi\rightarrow V\times G/G_0$, G is semisimple, g_M is, up to a sign, the G-invariant metric corresponding to the restriction of the Killing form on G to $\mathscr{M}$ and the pull-back φ^*k to $\mathfrak{g}_0$ is proportional to the restriction of the Killing form to $\mathfrak{g}_0$. Choosing a co-frame $\{\phi^a,\theta^i\}$ where $\{\phi^i\}_{a=1,\ldots\dim V}$ is any co-frame on V, the metric (2) takes the form

$$g=h_{ab}\phi^a\phi^b+\epsilon f^2k_{ij}\theta^i\theta^j, \tag{31}$$

where k_{ij} is the restriction of the Killing form to $\mathscr{M}$ expressed relative to the basis $\{T_i\}$ and $\epsilon=\pm1$. Taking the gauge field to be Ω_φ^σ, the energy–momentum tensor T is of the form

$$T=\frac{\kappa m}{8f^4}h_{ab}\phi^a\phi^b+\kappa\frac{m-4}{8f^2}k_{ij}\theta^i\theta^j, \tag{32}$$

where κ is a constant, while the Ricci tensor $\mathscr{R}$ also has the block-diagonal form

$$\mathscr{R}=R_{ab}\theta^a\theta^b+\rho k_{ij}\theta^i\theta^j, \tag{33}$$

where

$$R_{ab}=r_{ab}-m\frac{\nabla_a\nabla_b f}{f}, \tag{34}$$

and

$$\rho=-\tfrac{1}{2}-\epsilon(m-1)\nabla_a f\nabla^a f-\epsilon f\nabla_a\nabla^a f, \tag{35}$$

∇_a being covariant differentiaion on V with respect to the Levi–Civita connection for h_{ab} and r_{ab} the corresponding Ricci tensor. It follows that if $\dim V=0$, T and $\mathscr{R}$ are proportional to g and therefore Eq. (3) is satisfied with

$$f = 1, \quad \lambda = \frac{m}{4}\left[1 - \frac{\kappa}{2}\right]. \tag{36}$$

In general, setting $\lambda = 0$, the Einstein equations become

$$r_{ab} - \tfrac{1}{2} r h_{ab} - \frac{m}{f} \nabla_a \nabla_b f + h_{ab} + \left[\frac{m}{f} \nabla_c \nabla^c f \frac{m(m-1)}{2f^2} \nabla_c f \nabla^c f + \frac{m\epsilon}{4f^2}\right] = \frac{\kappa m}{8f^4} h_{ab}, \tag{37}$$

$$-\tfrac{1}{2} r + \frac{m-1}{f} \nabla_c \nabla^c f + \frac{(m-1)(m-2)}{2f^2} \nabla_c f \nabla^c f + \frac{\epsilon(m-2)}{4f^2} = \frac{\kappa(m-4)}{8f^4}, \tag{38}$$

where $r = r_{ab} h^{ab}$.

Provided $\nabla_a f \neq 0$, the second of these is implied by the first in view of the contracted Bianchi identities and covariant conservation of energy–momentum.

If $\dim V = 1$, we may use Gaussian coordinates, in which the metric is

$$g = dt^2 + \epsilon f^2(t) k_{ij} \theta^i \theta^j. \tag{39}$$

Solving Eqs. (37) and (38) gives

$$\epsilon f^2(t) = \frac{\kappa}{2} - \frac{t^2}{2(m-1)}, \tag{40}$$

where the integration constant is absorbed into the definition of t. Note that $m = 1$ is excluded by the assumption that G is semisimple.

If $\dim V = 2$ and $\nabla_a f \neq 0$, we may choose coordinates in which the metric is of the form

$$g = F\,dt^2 - G\,dr^2 + \frac{Gr^2}{2(m-1)} k_{ij} \theta^i \theta^j. \tag{41}$$

Solving Eq. (37), we get

$$F = \frac{1}{G} = 1 + \frac{A}{r^2} - \frac{2M}{r^{m-1}}, \tag{42}$$

where

$$A = \begin{cases} \dfrac{(m-1)^2}{3-m} & \text{for } m \neq 3 \\ -4\kappa \ln r & \text{for } m = 3 \end{cases},$$

and m is an integration constant. If $\nabla_a f = 0$, Eq. (37) implies

$$\epsilon f^2 = \kappa/2, \tag{43}$$

and the solution to Eq. (38), which now defines the scalar curvature on V as

$$r = 2/\kappa, \tag{44}$$

is

$$F = G = \kappa/r^2, \tag{45}$$

or

$$F = 1/G = r^2/\kappa,$$

unique up to a change of coordinates.

4. EXAMPLES AND DISCUSSION

Up to factorization by discrete groups, the compact, irreducible RSS' are[6]: the Grassmann manifolds (over R, C, and H); the compact simple Lie groups, the quotient spaces $SU(2n)/Sp(n)$, $SU(n)/SO(n)$, $Sp(n)/U(n)$ (for $n \geqslant 3$), and $SO(2n)/U(n)$ (for $n \geqslant 4$); and the quotients of the compact expectional groups by their maximal subgroups. To each of these, there corresponds a noncompact analogue obtained by analytic continuation. An arbitrary RSS may be decomposed canonically into products involving the above types together with R^n. However, the irreducible RSS of $\dim \leqslant 4$ are, up to discrete identifications, isomorphic to the Grassmann manifolds $SU(2)/U(1)$, $SO(4)/SO(3)$, $SU(3)/S(U_1 \times U_2)$, $SO(5)/SO(4)$, or their noncompact analogue $SU(1,1)/U(1)$, $SO_0(3,1)/SO(3)$, $SU(2,1)/S(U_1 \times U_2)$, and $SO_0(4,1)/SO(4)$. For these, the canonical connection and metric have been discussed in Ref. 4, therefore we shall only discuss the interpretation in the present context.

For $\dim V = 0$, $m = 4$, the spaces $SO(5)/SO(4)$ and $SO_0(4,1)/SO(4)$ are respectively the sphere and pseudosphere with canonical $SO(5)$ and $SO(4,1)$ invariant metric. The bundle may be identified with that of orthonormal frames and the gauge field with the Riemannian curvature for the Levi–Civita connection. In terms of gauge fields, with the identification $SO(4) \sim SU(2) \times SU(2)$, this is the (unique) $SO(5)$ invariant instanton plus anti-instanton configuration[7] or its analytic continuation. The space $SU(3)/S(U_1 \times U_2)$ is CP^2 with Fubini–Study metric. This and the gauge field for the canonical connection is given explicitly in Ref. 2. The corresponding quantities on the noncompact analogue $SU(2,1)/S(U_2 \times U_1)$ may be obtained by analytic continuation.

For $\dim V = 1$, $m = 3$, the orbits are again either the sphere $SO(4)/SO(3)$ or pseudosphere $SO(3,1)/SO(3)$. The metric (39) becomes precisely the Robertson–Walker metrics for the closed ($\epsilon = \pm 1$) or open ($\epsilon = -1$) universe for the compact and noncompact cases, respectively. For a positive energy density, we must have $\kappa > 0$. The expression (40) is the same as for a universe filled with electromagnetic radiation, but the interpretation is different, since the source is a non-Abelian gauge field. If, for the compact case the bundle is identified instead as

$$R \times SU(2) \times SU(2) \rightarrow R \times SU(2) \times SU(2)/SU(2)_\Delta,$$

where $SU(2)_\Delta$ is the diagonal subgroup, and the base is also identified as $R \times SU(2)$, then under the obvious choice of section $\sigma: g \rightarrow (g,e)$ we obtain

where ω_{MC} is the Maurer–Cartan form on SU(2). This is just the "meron" solution[8] expressed in a curved space which is, however, conformal to Minkowski space. The noncompact analog is $R \times Sl(2,C) \rightarrow R \times Sl(2,C)/SU(2)$ for which the gauge field is obtained by analytic continuation of (46).

A higher dimensional generalization of this example is obtained by replacing SU(2) by any compact, semisimple Lie group. The resulting gauge field, still given by (46) satisfies the source free field equations with respect to the Killing metric and may be interpreted as a higher dimensional analogue of the meron.

For dim $V = 2$, $m = 2$, the manifold is $V \times SU(2)/U(1)$ or $V \times SU(1,1)/U(1)$. For the first, with $\epsilon = +1$, Eqs. (31) and (32) define the Reissner–Nordstrom metric, where again $\kappa > 0$ for positive energy. However, the exterior field must be interpreted as that of a magnetic monopole, symmetric about $r = 0$. For the noncompact case, we have $\epsilon = -1$, $\kappa < 0$ and we must reinterpret the r as a timelike coordinate (for r sufficiently large) and t a spacelike one. The gauge field and metric g_M is obtained again by analytic continuation. The case with constant f, given by Eqs. (43)–(45) gives the Bertotti–Robinson[9] metric as its analytic continuation.

Finally, a word about the relevance of the higher dimensional cases to solutions of field equations in space–time. It has been shown in Refs. 1 and 2 that a class of immersions $CP^1 \rightarrow CP^m$ pull back the canonical connection on the Hopf bundle over CP^n to yield magnetic monopole solutions of strength n. The Einstein equations on $V \times CP^1$ (dim $V = 2$) will also be satisfied provided the contsant κ defining the energy–momentum tensor is suitably redefined. These immersions have a natural group theoretic interpretation in terms of irreducible representations $SU(2) \rightarrow SU(n)$ suggesting that other immersions exist, determined by group homomorphisms, which also pull back the canonical connections on higher dimensional spaces to yield new solutions of the field equations on space–time. An alternative approach involves dimensional reduction[5,10] whereby the submersion of a higher dimensional manifold defined through a group action preserving the connection and metric projects solutions of the source free gauge field equations to solutions of the coupled gauge and Higgs scalar fields equations on the manifold of group orbits. These approaches will be further developed elsewhere.

ACKNOWLEDGMENT

The authors would like to thank Prof. A. Trautman for helpful discussions relating to this work.

[1] A. Trautman, Int. J. Theor. Phys. **16**, 561 (1977).
[2] J. Nowakowski and A. Trautman, J. Math. Phys. **19**, 1100 (1978).
[3] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963), Vol. I.
[4] G. Bredon, *Introduction to Compact Transformation Groups* (Academic, New York, 1972), Chap. 2.
[5] J. Harnad, S. Shnider, and L. Vinet, "Group Actions on Principal Bundles and Invariance Conditions for Gauge Fields," preprint CRMA-899 (1979).
[6] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[7] R. Jackiw and C. Rebbi, Phys. Rev. D. **14**, 517 (1976).
[8] V. de Alfaro, S. Fubini, and G. Furlan, Phys. Lett. B **65**, 163 (1976); Phys. Lett. B **72**, 203 (1977).
[9] C. Misner, K. Thorne, and J. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), p. 845.
[10] J. Harnad, S. Shnider, and J. Tafel, "Group Actions on Principal Bundles and Dimensional Reduction," preprint CRMA-918 (1979).

On a new solution of Einstein's equations[a)]

C. Hoenselaers[b)]

Department of Physics, Montana State University, Bozeman, Montana 59717

(Received 27 June 1979; accepted for publication 28 September 1979)

An analysis of a stationary axisymmetric solution of Einstein's equations recently derived by Hoenselaers, Kinnersley, and Xanthopoulos is given. We derive the metric of the full space–time, the invariants of the Weyl tensor, the first few multipole moments, and discuss the stationary limit surfaces and the axis.

1. INTRODUCTION

The method for dealing with the symmetries of the stationary axisymmetric Einstein–Maxwell equations originally devised by Kinnersley and Chitre has proven very powerful in the otherwise rather tedious search for new solutions. Recently[1,2] we were able to publish several new stationary axisymmetric solutions of the vacuum Einstein equations. In this paper we shall analyze the new solution of Ref. 1. For details of its derivation the reader is referred to Ref. 2 and the papers cited therein.

2. THE SOLUTION

The solution in which we shall be interested is given in polar coordinates $r = (\rho^2 + z^2)^{1/2}$, $\tan\vartheta = z/\rho$, where ρ and z are Weyl canonical coordinates by the Ernst potential

$$\xi = \beta/\alpha, \tag{2.1a}$$

$$\alpha = r^4 - b^2(1-\cos^4\vartheta) + i[2br^2(\cos^2\vartheta - \sin^2\vartheta) - ar^3\cos\vartheta],$$

$$\beta = ar^3 - 2br^2\cos\vartheta - 2ib^2\sin^2\vartheta\cos\vartheta. \tag{2.1b}$$

The full four-dimensional metric is

$$ds^2 = \frac{1}{f}[e^{2\gamma}(dr^2 + r^2 d\vartheta^2) + r^2\sin^2\vartheta d\varphi^2] - f(dt - \omega\, d\varphi)^2, \tag{2.2a}$$

and the various expressions are:

$$f = A/B, \quad e^{2\gamma} = A/r^8, \quad \omega = -2\sin^2\vartheta\,(C/A),$$

$$A = \alpha\alpha^* - \beta\beta^* = r^8 - a^2r^6\sin^2\vartheta + 8abr^5\sin^2\vartheta\cos\vartheta + 2b^2r^4\sin^2\vartheta(9\sin^2\vartheta - 8) + b^4\sin^8\vartheta, \tag{2.2b}$$

$$B = |\alpha+\beta|^2 = [r^4 + ar^3 - 2br^2\cos\vartheta - b^2(1-\cos^4\vartheta)]^2 + [ar^3\cos\vartheta - 2br^2(\cos^2\vartheta - \sin^2\vartheta) + 2b^2\sin^2\vartheta\cos\vartheta]^2, \tag{2.2c}$$

$$C = r^7a^2 - r^6(6ab\cos\vartheta - a^3) - r^5(2b^2(1-5\cos^2\vartheta) + 8a^2b\cos\vartheta) - r^4ab^2(1-21\cos^2\vartheta) - r^3 16b^3\cos^3\vartheta - r2b^4\sin^6\vartheta - ab^4\sin^6\vartheta. \tag{2.2d}$$

[a)]Partially supported by National Science Foundation Grant PHY 78-12294

[b)]Present Address: Max-Planck-Institut für Astrophysik, 8046 Garching, FRG

We have written a and b instead of α_0 and α_1 as in Ref. 1.

For $b = 0$, ξ reduces to the expression for the extreme Kerr solution whereas the limit $b \to \infty$ gives

$$\xi = i\frac{(1+\cos\vartheta)^2 - (1-\cos\vartheta)^2}{(1+\cos\vartheta)^2 + (1-\cos\vartheta)^2}, \tag{2.3}$$

another well known solution which can be obtained from the Voorhees metric[3] by the interchange $x \leftrightarrow y$.

3. WEYL INVARIANTS

For any vacuum metric (2.2a) the tetrad components of the Weyl tensor[4]

$$C_2 = -\frac{1}{8}\left(2\partial\partial\mathscr{E} + \frac{1}{f}(\partial\mathscr{E})^2\right),$$

$$C_0 = \frac{1}{8}\left(\frac{1}{f}\partial\mathscr{E}\partial^*\mathscr{E} - \frac{1}{\rho}(\partial\rho\partial^*\mathscr{E} + \partial^*\rho\partial\mathscr{E})\right),$$

$$C_{-2} = -\frac{1}{8}\left(2\partial^*\partial^*\mathscr{E} + \frac{1}{f}(\partial^*\mathscr{E})^2\right),$$

$$\mathscr{E} = (\alpha-\beta)/(\alpha+\beta), \tag{3.1}$$

give rise to two spin invariant quantities which are for our case

$$I_1 = C_0 = \frac{-r^6}{2(\alpha+\beta)^3}\{r^3a - 6br^2\cos\vartheta - 2b^2 \times [2 + i\cos\vartheta(\cos^2\vartheta - 3)]\}, \tag{3.2a}$$

$$I_2 = C_2C_{-2} - 9C_0^2 = 36[r^{12}/(\alpha+\beta)^5]\, b^2\sin^2\vartheta. \tag{3.2b}$$

They become infinite only if $\alpha + \beta = 0$ (see below). In the limit $r \to 0$ their value depends on the limit of $x = r^{-1}\sin\vartheta$, but there is no way to choose x such that $I_{1,2}$ become infinite. This behavior is very similar to the one of the Tomimatsu–Sato $\delta = 2$ solution[5] and suggests that, as it is the case there, $r = 0$ as approached from above or below are in fact surfaces (see Sec. 7).

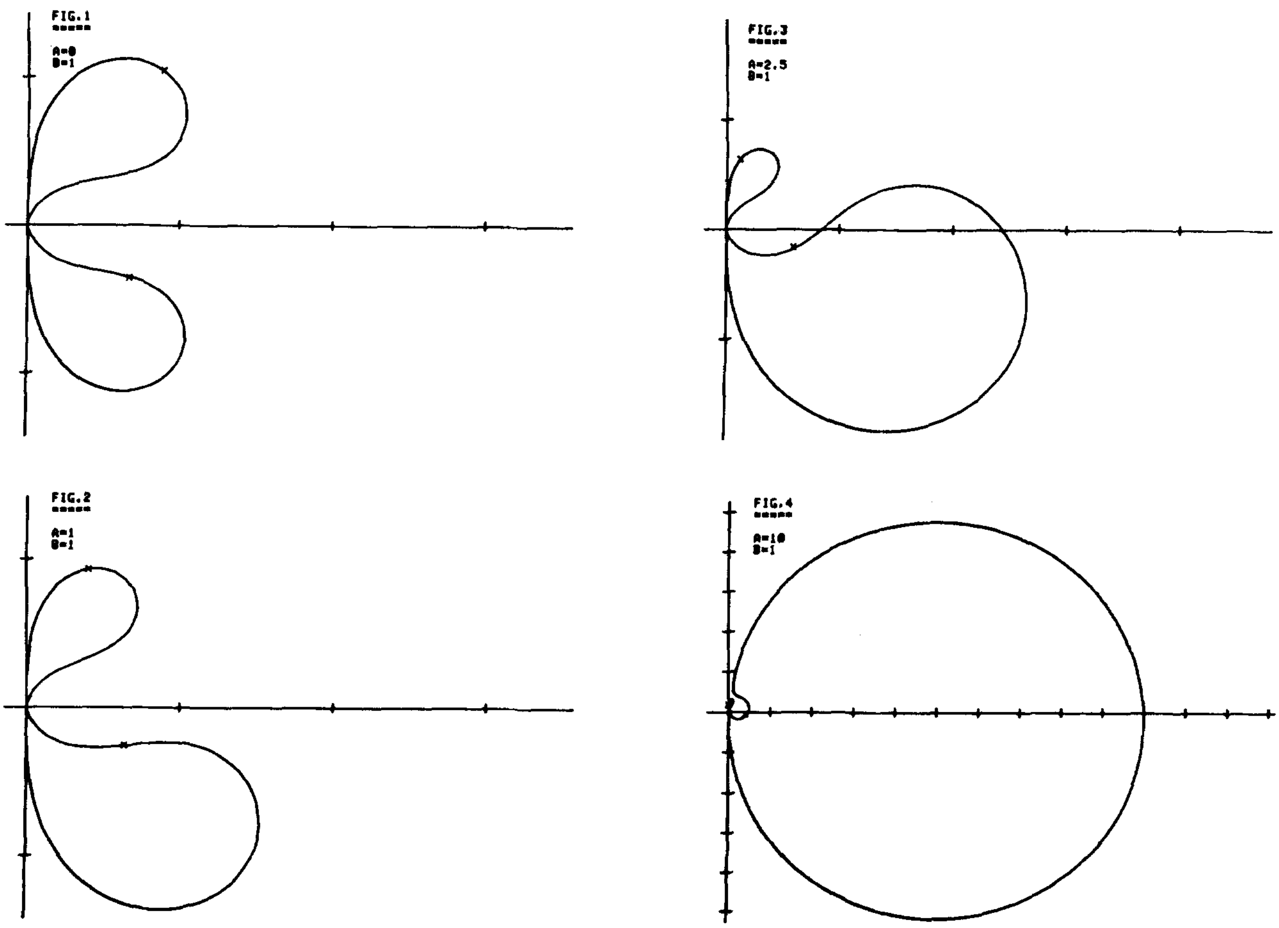
4. STATIONARY LIMIT SURFACES AND SINGULARITIES

The ∂_t-Killing vector becomes a null vector on the surface

$$A = 0. \tag{4.1}$$

At first glance this appears to be an 8th order polynominal

 0022-2488/80/082241-04$1.00



FIGS. 1–4. These show a cross section through the infinite redshift surface and the location of the singularities (indicated by the "x") in Weyl coordinates ρ (horizontal axis) and z (vertical axis).

which cannot be solved analytically. However, upon introducing the variables $x = r^{-1} \sin\vartheta$, $y = r^{-1} \cos\vartheta$ (4.1) becomes for $r \neq 0$

$$(1 + b^2 x^4)^2 - x^2(4by - a)^2 = 0,$$

from which one finds for $b \neq 0$ the parametric form of the stationary limit surface

$$\rho = D^{-1} \cdot 16 b^2 |x^3|,$$
$$z = D^{-1} \cdot 4bx(ax - 1 - b^2 x^4), \quad (4.2)$$
$$D = 16 b^2 x^4 + (ax - 1 - b^2 x^4)^2.$$

For $b = 0$, the solution of (4.1) is, of course, the circle $r = a \sin\vartheta$.

As it has been mentioned above, the singularities are given by

$$\alpha + \beta = 0, \quad (4.3)$$

and hence lie, cf. the definition (2.2b) of A, on the stationary limit surface. The real and imaginary part of this equation can be combined to yield a biquadratic equation whose solution is

$$r^2 = b \tan\vartheta \, [\sin\vartheta \pm (\cos^2\vartheta + 1)], \quad (4.4)$$

with the $\pm$ sign for $\vartheta \lessgtr \pi/2$. Upon substituting this expression into the imaginary part of (4.3) we get after some algebra a 6th order equation for $\sin\vartheta$ which cannot be solved. We thus had to resort to numerical methods.

In Figs. 1–4 we have plotted the shape of the stationary limit surface and the location of the singularities for various values of the parameters.

5. MULTIPOLE MOMENTS

Following the definition of multipole moments of stationary spacetimes given by Hansen[6] we write

$$\phi = \phi_M + i\phi_J = -\alpha^* \beta A^{-1} \quad (5.1)$$

for the combined mass and angular momentum potential. As an appropriate coordinate near "infinity", the point Λ in the conformally completed three manifold, we take $R = r^{-1}$ and choose $\Omega = R^2 e^{-\gamma}$ as conformal factor. This gives

$$d\tilde{\sigma}^2 = dR^2 + R^2 d\vartheta^2 + R^2 \sin^2\vartheta e^{-2\gamma} d\varphi^2,$$
$$\tilde{\phi} = -\alpha^* \beta R^7 (AR^8)^{-3/4} \quad (5.2)$$

for the conformal metric and the transformed potential (5.1). On account of the rotational symmetry of our solution the multipole moments have to be multiples of the symmetric trace-free outer product of the axis vector with itself. The first few moments are

$$M_0 = -a,$$
$$M_1 = -ia^2 + 2b,$$
$$M_2 = a^3 + 4iab,$$
$$M_3 = ia^4 - \frac{40}{3} ib^2 - 10a^2 b. \quad (5.3)$$

For $a \neq 0$ one can shift the origin of the coordinate system by $z \to z + 2b/a$ and find the multipole moments for the center of mass system. They are

$$M'_0 = -a,$$
$$M'_1 = -ia^2,$$
$$M'_2 = a^3 + 8b^2/a,$$
$$M'_3 = ia^4 - \frac{20}{3}(4ib^2 + 8b^3/a^2 + 3a^2 b). \quad (5.4)$$

It is remarkable that in the center of mass system not only the mass dipole moment, but also the angular momentum quadrupole moment vanishes.

6. THE AXIS

We shall now concentrate on the axis, i.e., the two-dimensional geodesic submanifold given by $\vartheta = 0, \pi$. As our solution is not symmetric under $\vartheta \to \pi - \vartheta$, we have to distinguish between the upper ($\vartheta = 0$) and the lower ($\vartheta = \pi$) part of the axis. For both the metric is

$$d\sigma^2 = (B_0/r^4)dr^2 - (r^4/B_0)dt^2,$$
$$B_0 = r^4 + 2ar^3 + 2(a^2 - 2b)r^2 - 8abr + 8b^2, \quad (6.1)$$

and the upper and lower part are connected by $b \to -b$. It can easily be shown that B_0 vanishes nowhere.

Solving the geodesic equation yields

$$\dot{t} = eB_0 r^{-4},$$
$$\dot{r}^2 = e^2 + (r^4/B_0)\epsilon,$$
$$(\epsilon = \pm 1, 0, \quad e = \text{const}). \quad (6.2)$$

Concentrating on timelike geodesics ($\epsilon = -1$) we focus attention on the solutions of

$$e^2 B_0 - r^4 = 0, \quad (6.3)$$

the points at which a particle has to reverse its motion or come to a halt. These r-values are given by

$$r_0 = k^{-1}(-a \pm \sqrt{a^2 + 4kb}),$$
$$k = 1 \pm \sqrt{2e^{-2} - 1}, \quad (6.4)$$

which indicates that a particle encounters a potential barrier of height 2, as there are no real solutions for r if $e^2 > 2$. The allowed and forbidden regions of motion may easily be inferred from (6.4). We just mention that the lowest possible energy for particles on the outside of the potential barrier is

$$e^2_{\min} = \min\left(\frac{2}{(a^2/4b)^2 + 1}, 1\right).$$

Comparing our results with the results Carter[7] obtained for the Kerr metric, we first note that the height of the potential barrier is the same in both cases (n.b., $m = a$, as our solution reduces to the extreme Kerr solution). It should moreover be mentioned that for $b = 0$ the two solutions $r = 0$ of (6.4) are to be omitted, as B_0 contains a factor r^2 which drops out of the discussion and (6.3) reduces to a second order equation.

From (6.2) it can be seen that an extension is necessary. This can, however, be effected in analogy to the case of the extreme Kerr metric, and we shall just sketch the procedure. Introducing double null coordinates u and v by

$$\tfrac{1}{2}(u + v) = F(r) := \int (B_0/r^4)\, dr,$$
$$\tfrac{1}{2}(u - v) = t, \quad (6.5)$$

we find for (6.1)

$$d\sigma^2 = (r^4/B_0)du dv.$$

As $F(r)$ is monotonous for $r > 0$ and $r < 0$, r is given uniquely in terms of u and v, provided one specifies the sign of r. Using coordinates ξ, η, defined by

$$u = \cos\xi/\sin^3\xi, \quad v = \cos\eta/\sin^3\eta,$$

we let α_n, β_n denote the lines

$$\alpha_n : \xi = n\pi, \quad \beta_n : \eta = n\pi \quad (n = 0, \pm 1, \cdots),$$

and Q_{mn} the intersection of the strips bounded by α_n, α_{n+1} and β_n, β_{n+1} respectively. $Q_{n,n}$ is the image of region I ($r > 0$), while $Q_{n+1,n}$ is the one of region II ($r < 0$). We finally find for (6.1)

$$d\sigma^2 = \mathcal{G}(\xi, \eta) d\xi d\eta,$$
$$\mathcal{G} = \frac{r^4}{B_0}(\sin\xi \sin\eta)^{-3}(\sin^2\xi + 3\cos^2\xi) \times (\sin^2\eta + 3\cos^2\eta).$$

It can be shown that $\mathcal{G}$ is continuous and positive definite throughout M^*, consisting of the union of the images of I and II.

7. THE METRIC NEAR $r = 0$

If one wants to consider the metric (2.2a) for small r, i.e., for instance replace $r \to \lambda r$ and take the limit $\lambda \to 0$, one has to rescale the other coordinates as well and perform an appropriate conformal transformation to obtain a nontrivial result. We thus scale the coordinates as

$$r \to \lambda r, \quad \varphi \to \lambda^{-4}\varphi, \quad t \to \lambda^{-3} t + 2a\lambda^{-4}\varphi,$$

perform a conformal transformation with

$$\Omega = \lambda^3,$$

and let $\lambda \to 0$. The resulting metric is

$$ds^2 = B[\sin^4\vartheta r^{-8}(dr^2 + r^3 d\vartheta^2) + r^2 \sin^2\vartheta d\varphi^2] - \sin^4\vartheta B^{-1}(dt - 4r d\varphi)^2,$$
$$B = (1 + \cos^2\vartheta)^2 + 4\cos^2\vartheta. \quad (7.1)$$

A look at the ξ-potential (2.1a) shows that the terms in α and β containing b^2 are just the ones without r. (7.1) thus turns out to be the metric derived from (2.3). We have set the only remaining constant, b, to unity.

It has been pointed out[8] that (7.1) describes the region of the TS 2 metric near $x = 1$ and also arises from the undistinguished limit of that solution.

We shall briefly discuss a few features of geodesics in this solution. The Hamilton–Jacobi equation becomes

$$\epsilon = \frac{r^8}{B\sin^4\vartheta}(s_{,r}^2 + r^{-2}s_{,\vartheta}^2) - \frac{1}{r^2\sin^2\vartheta B} \times\left[-\sin^4\vartheta l^2 + 8\sin^4\vartheta rel + \frac{r^2e^2}{\sin^2\vartheta}(B - 16\sin^6\vartheta)\right], \tag{7.2}$$

where we have already made use of the obvious separation

$$S = \epsilon\tau - et + l\varphi + s(r,\vartheta).$$

As the metric (7.1) is symmetric under $\vartheta \to \pi - \vartheta$ there are geodesics confined to $\vartheta = \pi/2$. Hence setting $s_{,\vartheta} = 0$, $\vartheta = \pi/2$ one finds from (7.2)

$$s_{,r} = r^{-5}\sqrt{-(15e^2 - \epsilon)r^2 + 8elr - l^2}.$$

The motion of a particle is thus confined to $4e - (e^2 + \epsilon)^{1/2} \leqslant (15e^2 - \epsilon)l^{-1}r \leqslant 4e + (e^2 + \epsilon)^{1/2}$. Interestingly there is a circular timelike orbit for particles with unit energy at $r_c = l/4$. Furthermore one can show that null geodesics with $l = 0$ which reach $r = 0$ obey $r \sim \sin\vartheta$.

Returning now to the original solution (2.2a), one can show by a rather lengthy but straightforward calculation that again

$$\sin\vartheta = \alpha r \quad (\alpha = \text{const}), \tag{7.3}$$

is in the approximation of small r a solution of the geodesic equation for null geodesics with vanishing angular momentum. Those geodesics thus approach the axis as they come close to $r = 0$.

On the other hand, the metric (2.2a) restricted to the surfaces $t,\alpha = \text{const}$ becomes in the limit of small r, respectively large $R = r^{-1}$,

$$d\sigma^2 = 8b^2(1 + b^2\alpha^4)[dR^2 + (1 + b^2\alpha^4)^{-2}d\varphi^2], \tag{7.4}$$

i.e., the metric on the surface of a cylinder. Taking now α as a variable, i.e., different values of α label different geodesics, one finds for (2.2a) restricted to $r = \text{const}$ in the limit $r \to 0$

$$d\sigma^2 = 8b^2\{(1 + b^2\alpha^4)d\alpha^2 + [\alpha^2/(1 + b^2\alpha^4)]d\varphi^2\}. \tag{7.5}$$

It should be noted that both (7.4) and (7.5) are independent of whether $\cos\vartheta$ goes to $+1$ or -1 in the limit.

(7.5) shows that $r = 0$ is in fact a null hypersurface. One can furthermore show that it is not isometric to the pole of the TS 2 solution as given by Ernst.[9]

8. CONCLUDING REMARKS

After the analysis of the foregoing paragraphs we may envisage the solution (2.2) as being generated by two rotating objects (maybe rings) with different masses and radii. Coalescing those two objects leads to the extreme Kerr solution. An interesting feature of our solution is the potential barrier which is encountered by a particle on the axis. For small b the barrier becomes narrower, not smaller. Even if one regards (2.2) for infinitesimal b as a perturbation of the extreme Kerr metric, this barrier remains.

ACKNOWLEDGMENTS

I should like to thank Professor W. Kinnersley for helpful discussions and the referee for pointing out a mistake.

[1] C. Hoenselaers, W. Kinnersley, and B. Xanthopoulos, Phys. Rev. Lett. **42**, 481 (1979).
[2] C. Hoenselaers, W. Kinnersley, and B. Xanthopoulos, J. Math. Phys. **20**, 2530, (1979).
[3] B. Voorhees, Phys. Rev. D **2**, 2119 (1970).
[4] F. Ernst, J. Math. Phys. **15**, 1409 (1974).
[5] J. Economou and F. Ernst, J. Math. Phys. **17**, 52 (1976).
[6] R.O. Hansen, J. Math. Phys. **15**, 46 (1974).
[7] B. Carter, Phys. Rev. **141**, 1242 (1966).
[8] W. Kinnersley and E. Kelley, J. Math. Phys. **15**, 2121 (1974).
[9] F. Ernst, J. Math. Phys. **17**, 1091 (1976).

A note on the good cut equation

Robert W. Lind
Department of Physics, Unviersity of Wisconsin-Platteville, Plattevile, Wisconsin 53818

(Received 4 September 1979; accepted for publication 25 January 1980)

The good cut equation ($\eth^2 Z = \sigma$) is the link between a physical asymptotically flat space–time (characterized by σ) and its associated H-space (characterized by Z and interpreted by Penrose as a nonlinear graviton). In this work a class of regular, nontrivial ($\sigma \neq 0$) solutions to the good cut equation is obtained by writing σ in terms of a potential function ϕ.

I. INTRODUCTION

It is not possible, in general, to find asymptotically shear-free null surfaces in real asymptotically flat space–times. If the space–time is analytically extended into the complex in the neighborhood of future null infinity, then *complex* asymptotically shear-free null surfaces can be found. This is how it is done: Use Bondi-type coordinates $(u,\zeta,\tilde{\zeta})$ on complex null infinity with u being complex Bondi time, and ζ and $\tilde{\zeta}$ being complex angular coordinates. (In the real space $\tilde{\zeta} = \bar{\zeta}$, the complex conjugate of ζ.) Choose $u = z(\zeta,\tilde{\zeta})$ such that

$$\eth^2 Z(\zeta,\tilde{\zeta}) = \sigma(Z,\zeta,\tilde{\zeta}). \tag{1.1}$$

The function $\sigma(Z,\zeta,\tilde{\zeta})$ is the analytic extension of the asymptotic Bondi shear $\sigma(u,\zeta,\tilde{\zeta})$ of the Bondi null surfaces for which u = constant. Equation (1.1) is known as the good cut equation and its regular solutions are known as good cut functions. It describes the intersection of complex null infinity with the asymptotically shear-free null surfaces.[1,2,3]

The operator $\eth$ is defined in terms of its action on spin-weighted spherical harmonics,[4]

$$\eth f_s = (1 + \zeta\tilde{\zeta})^{1-s} \frac{\partial[(1+\zeta\tilde{\zeta})^s f_s]}{\partial\zeta}. \tag{1.2}$$

Regular solutions of the good cut equations are those for which $Z(\zeta,\tilde{\zeta})$ is expandable in spherical harmonics. These regular solutions form a four-complex parameter set. These four parameters can be considered as local coordinates in four-complex-dimensional manifolds called H spaces.[1,2,3] Newman views these H spaces as containing information on the related asymptotically flat real space–time and has suggested that they may be used to define the in and out states in an S-matrix formulation of quantum gravity. Penrose has shown that these same complex spaces can be obtained from the structure of a holomorphically deformed twistor space[5,6] and interprets them as nonlinear gravitons, i.e., one-particle helicity eigenstates of a quantum theory of gravity.

In the next section we show how the good cut equation can be written in terms of a potential function for the asymptotic Bondi shear. In Sec. III we find a class of nontrivial solutions to this equation.

II. GOOD CUT EQUATION IN TERMS OF A POTENTIAL FUNCTION

A potential function $\phi(Z,\zeta,\tilde{\zeta})$ can be obtained for $\sigma(Z,\zeta,\tilde{\zeta})$ in two steps.[7] Let $\eth'$ be an operator that acts on spin-weighted spherical harmonics according to Eq. (1.2), but with Z held fixed. Thus

$$\eth' Z = 0. \tag{2.1}$$

First introduce the function $L(Z,\zeta,\tilde{\zeta})$ defined by[7]

$$\sigma \equiv \eth' L + L L_{,Z}, \tag{2.2}$$

where

$$L_{,Z} \equiv \frac{\partial L}{\partial Z},$$

and then the function $\phi(Z,\zeta,\tilde{\zeta})$ defined by[7]

$$L \equiv -\eth'\phi/\phi_{,Z}. \tag{2.3}$$

Now let $\eth$ be an operator acting on spin-weighted spherical harmonics, but with Z not held fixed. The actions of $\eth'$ and $\eth$ on any spin-weighted function $f(Z,\zeta,\tilde{\zeta})$ are then related by

$$\eth' f = \eth f - f_{,Z}\,\eth Z. \tag{2.4}$$

Thus L can be written in the form

$$L = \eth Z - \eth\phi/\phi_{,Z}, \tag{2.5}$$

and σ as

$$\sigma = \eth L - L_{,Z}(L - \eth Z). \tag{2.6}$$

Similarly,

$$L_{,Z} = -\eth^2\phi_{,Z}/\phi_{,Z} + \phi_{,ZZ}\,\eth\phi/(\phi_{,Z})^2, \tag{2.7}$$

and

$$\eth L = \eth^2 Z - \eth^2\phi/\phi_{,Z} + \eth\phi\,\eth\phi_{,Z}/(\phi_{,Z})^2. \tag{2.8}$$

After a little algebra, substitution of Eqs. (2.5), (2.6), (2.7), and (2.8) into Eq. (1.1) yields the following result:

$$\eth[(\phi_{,Z})^2/\eth\phi] = \phi_{,ZZ}. \tag{2.9}$$

This is the good cut equation written in terms of the potential ϕ for σ and its first and second Z derivatives. Specifying the function $\phi(Z,\zeta,\tilde{\zeta})$ leads via Eqs. (2.2) and (2.3) to a particular σ, and to an equation for the good cut function $Z(\zeta,\tilde{\zeta})$ via Eq. (2.9). A simple but nonetheless nontrivial ($\sigma \neq 0$) class of solutions can be found using this approach, as is demonstrated in the next section.

III. A CLASS OF SOLUTIONS TO THE GOOD CUT EQUATION

If the potential has the form

$$\phi = \tfrac{1}{2}Z^2 + \alpha(\zeta,\tilde{\zeta}), \tag{3.1}$$

$$\alpha = \sum_{l,m} \alpha_{lm} Y_{lm}(\zeta,\tilde{\zeta}), \tag{3.2}$$

 0022-2488/80/082245-03$1.00

the good cut equation (2.9) becomes

$$ð[Z^2/(Z ðZ + ðα)] = 1 . \tag{3.3}$$

The regular first integral of this equation is

$$Z^2/(Z ðZ + ðα) = z/ðz ,$$

which can be written in the form

$$z^2ð(Z^2/z^2) = -2ðα , \tag{3.4}$$

where, using the notation of Ref. 3,

$$z = \frac{u + Xζ + Y\bar{ζ} + vζ\bar{ζ}}{1 + ζ\bar{ζ}}, \tag{3.5}$$

with

$$uv - XY \neq 0 .$$

The four arbitrary complex parameters u, v, X, Y can be considered as coordinates in H space.

Equation (3.4) has the general regular solution

$$\frac{Z^2}{z^2} = 2 \sum_{l=1}^{\infty} \left[(-1)^{l-1} \frac{ð^l α}{ðz} ð^{-l+1} \frac{1}{z}\right] + β(\bar{ζ}), \tag{3.6}$$

where the sum is the particular solution to Eq. (3.4) written in terms of $ð^{-1}$, the inverse operator to ð. For example, the first few terms of the sum are

$$\sum_{l=1}^{\infty} (-1)^{l-1} \frac{ð^l α}{ðz} ð^{-l+1} \frac{1}{z} = \frac{ðα}{zðz} - \frac{\ln z ð^2α}{(ðz)^2} + \frac{(z \ln z - z)ð^3α}{(ðz)^3} - \frac{[(1/2)z^2 \ln z - (3/4)z^2]ð^4α}{(ðz)^4} + \cdots . \tag{3.7}$$

Applying ð to z as given by Eq. (3.5) yields

$$ðz = \frac{X - (u - v)\bar{ζ} - Y\bar{ζ}^2}{1 + ζ\bar{ζ}}, \tag{3.8}$$

which has zeros at

$$\bar{ζ} = \frac{-(u - v) \pm [(u - v)^2 + 4XY]^{1/2}}{2Y}. \tag{3.9}$$

Thus the l th term in the expansion (3.7) has an l th-order pole at each of the two values of $\bar{ζ}$ given by Eq. (3.9). These must be eliminated if the solution for Z is to be regular.

$β = β(\bar{ζ})$ is the singular solution of the homogeneous equation, chosen to make the entire solution regular, and must have the form

$$β = \sum_{l=1}^{\infty} β_l , \tag{3.10}$$

where[9]

$$β_l = (β^{a \cdots c} ðl_a \cdots ðl_c)/(ðz)^l , \tag{3.11}$$

with

$$ðl_a = (1/1 + ζ\bar{ζ})[(1 + ζ\bar{ζ}), -(ζ + \bar{ζ}), i(ζ - \bar{ζ}), (1 - ζ\bar{ζ})] . \tag{3.12}$$

$β^{a \cdots c}$ are l th degree three-dimensional trace-free symmetric tensors. Each has $2l + 1$ independent components that can be uniquely determined. The demand that Z be regular yields $2l$ conditions. The final component can be determined by requiring that the $l = 0$ part of the solution appear only in z.

As an example of an explicit solution, consider the special case

$$ð^2α = 0 , \tag{3.13}$$

where $α$ is expandable in terms of $l = 0$ and $l = 1$ spherical harmonics only,

$$α = \sum_{l=0}^{1} \sum_{m=-l}^{l} α_{lm} Y_{lm}(ζ,\bar{ζ}) = \frac{a + bζ + c\bar{ζ} + dζ\bar{ζ}}{1 + ζ\bar{ζ}}. \tag{3.14}$$

The corresponding solution to Eq. (3.4) is

$$\frac{Z^2}{z^2} = \frac{2ðα}{zðz} + β(\bar{ζ}). \tag{3.15}$$

The fact that $β = β(\bar{ζ})$ only can be used to find the particular form of $β$ that will eliminate the poles in the first term of the solution at $\bar{ζ}$ given by Eq. (3.9). Equation (3.15) can be written as

$$Z^2 = zw , \tag{3.16}$$

where

$$w = \frac{2ðα + βzðz}{ðz} \tag{3.17}$$

and

$$ð^2 w = 0 . \tag{3.18}$$

[This can be readily seen by substituting Z^2 given by Eq. (3.16) into Eq. (3.4) and using the fact that $ð^2α = 0$.] Now $β$ must be chosen in such a way that w (hence Z) is regular, i.e., such that w has the form

$$w = \sum_{l=0}^{1} \sum_{m=-l}^{l} w_{lm} Y_{lm}(ζ,\bar{ζ}) = \frac{A + Bζ + C\bar{ζ} + Dζ\bar{ζ}}{1 + ζ\bar{ζ}}. \tag{3.19}$$

This can be accomplished by substituting Eqs. (3.5), (3.14), and (3.19) into Eq. (3.17) and solving for $β$. Doing this we find that

$$β = (A + Bζ + C\bar{ζ} + Dζ\bar{ζ})[X - (u - v)\bar{ζ} - Y\bar{ζ}^2] - 2[b + (d - a)\bar{ζ} - c\bar{ζ}^2](1 + ζ\bar{ζ}) \times \{(u + Xζ + Y\bar{ζ} + vζ\bar{ζ})[X - (u - v)\bar{ζ} - Y\bar{ζ}^2]\}^{-1} . \tag{3.20}$$

The condition that $β = β(\bar{ζ})$ only, i.e.,

$$\frac{\partial β}{\partial ζ} = 0 , \tag{3.21}$$

and the condition that when $α = 0$, then $σ = 0$, and the solution to the good cut equation must reduce to $Z = z$ yield the final results

$$A = u + \frac{2[uv(d - a) + ucX - bvY]}{v(uv - XY)}, \tag{3.22a}$$

$$B = X + \frac{2[vX(d - a) - bv^2 + cX^2]}{v(uv - XY)}, \tag{3.22b}$$

$$C = Y - \frac{2c}{v}, \tag{3.22c}$$

$$D = v. \tag{3.22d}$$

Thus we have proved the following theorem. The class of regular solutions to the good cut equation characterized by

$$\phi = \tfrac{1}{2}Z^2 + \frac{a + b\zeta + c\bar{\zeta} + d\zeta\bar{\zeta}}{1 + \zeta\bar{\zeta}} \tag{3.23}$$

is given by

$$Z^2 = \frac{(A + B\zeta + C\bar{\zeta} + D\zeta\bar{\zeta})\, z}{1 + \zeta\bar{\zeta}}, \tag{3.24}$$

with z given by Eq. (3.5), A, B, C, and D given by Eqs. (3.22).

As an example, the Sparling solution quoted in Ref. 8 has

$$\phi = \tfrac{1}{2}Z^2 + i\lambda^{1/2}\zeta/(1 + \zeta\bar{\zeta}), \tag{3.25}$$

$$\sigma = \lambda/(1 + \zeta\bar{\zeta})^2 Z^3. \tag{3.26}$$

Thus $a = c = d = 0$ and $b = i\lambda^{1/2}$. For this case

$$A = u - \frac{2i\lambda^{1/2} Y}{uv - XY}, \tag{3.27a}$$

$$B = X - \frac{2i\lambda^{1/2} v}{uv - XY}, \tag{3.27b}$$

$$C = Y, \tag{3.27c}$$

$$D = v, \tag{3.27d}$$

and

$$Z^2 = z^2 - \frac{2i\lambda^{1/2}(Y + v\zeta)\, z}{(uv - XY)(1 + \zeta\bar{\zeta})}. \tag{3.28}$$

IV. DISCUSSION

The good cut equation is the direct link between a physical asymptotically flat space (characterized by σ) and its associated H space (characterized by Z). It is therefore important to find regular nontrivial ($\sigma \neq 0$) solutions to this equation. We have shown that the good cut equation can be cast into a different form by simply writing σ in terms of the potential function ϕ. We have also demonstrated that a class of regular nontrivial solutions can be readily obtained in this way. This work is being pursued further and it is hoped that this approach may lead to even broader classes of regular solutions or, at the very least, to a more general first integral of the good cut equation.

ACKNOWLEDGMENT

The author wishes to acknowledge several helpful discussions with Professor E.T. Newman.

[1]E.T. Newman, *General Relativity and Gravitation: Proceedings of the Seventh International Conference* (GR7) Tel Aviv University, June 23–28, 1974 (Wiley, New York, 1975).
[2]E.T. Newman, Gen. Rel. Grav. **7**, 107 (1976).
[3]M. Ko, E.T. Newman, and P. Tod, in *Asymptotic Structure of Space–Time* edited by Paul Esposito and Louis Witten (Plenum, New York, 1977).
[4]J.N. Goldberg *et al.*, J. Math. Phys. **8**, 2155 (1967).
[5]R. Penrose, Gen. Rel. Grav. **7**, 31 (1976).
[6]R. Penrose, *Quantum Gravity*, edited by C.J. Isham, R. Penrose, and D.W. Sciama (Oxford University, New York, 1975).
[7]B. Aronson, R.W. Lind, E.T. Newman, and J. Messmer, J. Math. Phys. **12**, 2462 (1971).
[8]M. Ko, M. Ludvigsen, E.T. Newman, and P.K. Tod, "The Theory of H Space" (unpublished).
[9]E.T. Newman and K.P. Tod, Gen. Rel. Grav. **8**, 679 (1977).

On the stationary Einstein–Maxwell field equations

A. Das and S. Kloster[a)]
Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada

(Received 24 January 1980; accepted for publication 4 April 1980)

The stationary gravitational equations in the presence of the electromagnetic fields, outside charged gravitating sources, are investigated. (i) The action integral of Kramer–Neugebauer–Stephani (K.N.S.) is derived from the Hilbert action integral by using new variational techniques. (ii) It is shown that the classification scheme for the system of partial differential equations of general relativity *depends* on the coordinate system used. In particular, if orthogonal coordinates are chosen for the associated space then the system of Einstein–Maxwell equations is a *hyperbolic* one. (iii) The eigenvalues of the Ricci tensor of associated space are expressed in terms of the invariants of stationary electro-gravitational fields. It is proved that if these eigenvalues are equal then the fields must belong to the class of Peres–Israel–Wilson (PIW) solutions. (iv) The global integrability of some of the stationary Einstein–Maxwell equations and the consequent equilibrium conditions of the "bodies" are investigated. (v) Boundary value problems for some of the field equations are pursued. It is proved that $\omega \equiv \ln|g_{44}|$ is neither subharmonic nor superharmonic and the boundary value problem for this function does *not* yield a unique solution in general. A nontrivial solution of the stationary equations with $\omega \equiv 0$ is given. A special boundary value problem is explicitly solved. (vi) The PIW solutions are generated from the charged Kerr–Tomimatsu–Sato–Yamazaki (KTSY) solutions. The complex axially symmetric harmonic functions of these PIW solutions can be obtained from the real axially symmetric harmonic functions of the static Weyl class of electrovac solutions by a complex scale transformation of the coordinates.

1. INTRODUCTION

In the last decade many papers have dealt with the topic of the stationary Einstein–Maxwell equations. In an important paper[1] KNS expressed these equations using two complex potentials. The KNS action integral has not yet been obtained from the usual Hilbert action integral in general relativity. In a simpler setting of axial symmetry and pure gravity, the problem was solved by Hoenselaers[2] employing Routh's procedure from classical mechanics. In the third section of this paper, using Routh's procedure and a further generalization, the KNS action integral has been derived from the Hilbert action integral.

In the classification scheme[3] of a system of semilinear partial differential equations (p.d.e.s), the unknown functions are treated as scalars under a coordinate transformation. But in general relativity the unknown functions are metric tensor components which undergo the transformation of a second order tensor. That is why the classification scheme for the field equations *does depend* on the chosen coordinate system. For example, it is usually accepted that the static or stationary field equations belong to an elliptic system. However, it was noted by Das[4] that a subset of static vacuum equations in orthogonal coordinates forms a *hyperbolic* system. In the fourth section of this paper, adopting orthogonal coordinates, it is proved that the system of stationary Einstein–Maxwell equations is mixed or *ultrahyperbolic*.

One way to classify the stationary metric algebraically is to investigate the eigenvalues of the Ricci tensor pertaining to the associated space. This approach was initiated[5] by Kloster, Som, and Das and also by Hoenselaers for the case of pure gravity. In the fifth section the eigenvalues of the Ricci tensor are obtained in terms of the invariants of stationary electrogravitational fields. It is proved that the eigenvalues cannot have the same sign. This result is used to show that in case all three eigenvalues coincide, the stationary Einstein–Maxwell fields must belong to the PIW class. The classes $\mathscr{C}_2$ and $\mathscr{C}_3$ for which two eigenvalues are equal and three eigenvalues are distinct respectively remain open for investigation.

Many authors[6] have investigated the global integrability of static and stationary field equations and the consequent equilibrium of the "bodies." In the sixth section several aspects of the known results are generalized. A finite body is defined as a region where at least one of the stationary Einstein–Maxwell equations is violated. Various properties of a "body," like mass density, twist mass density, stress density, etc., are described geometrically. Jump conditions on the bounding surface are provided. The most general integral condition of "equilibrium" of a body is written. A rigid displacement λ^{α} is defined by the condition that the total stress-energy is zero. Examples of such displacements are provided by the solutions of the p.d.e. $\sigma^{\alpha\beta}\lambda_{(\alpha|\beta)} = 0$. A solution of this p.d.e. is furnished in terms of an *arbitrary* C^3-differentiable vector field. In the axially symmetric case a special choice of this vector field (which is not C^3!) yields the known equilibrium condition that "the component of the total force along the axis of symmetry on each body is zero."

[a)]Both authors also belong to the Theoretical Science Inst., Simon Fraser Unviersity.

 0022-2488/80/082248-09$1.00

But in the general case if the vector field is chosen to have an arbitrary power series expansion, then it is not known how many independent equilibrium conditions emerge.

In the next section the boundary value problems for some of the elliptic field equations are looked into. The criterion for the variationally natural[7] boundary conditions is derived. It is known[8] that the function $\omega \equiv \ln|g_{44}|$ is superharmonic for the case of the stationary vacuum and subharmonic for the static Einstein–Maxwell equations. For both of these cases the space–time is flat in case ω is regular everywhere. However, for the stationary Einstein–Maxwell equations ω is neither superharmonic nor subharmonic, in general. This fact has prompted the present authors to find a class of nonflat solutions such that $\omega \equiv 0$. The boundary value problem for the semilinear elliptic equation for ω (in case ω is prescribed on the boundary) does *not* allow a unique solution,[3] in general. However, in special cases, for example the PIW class,[9] the field equations boil down to the ordinary Laplace's equation and a boundary value problem of the first kind allows a unique solution. A particular problem, where the boundary surface is spherical in the base space E_3 of the coordinates, is explicitly solved. *After* the solution has been obtained, it is found that in the physical Riemannian space the corresponding surface is *not* of constant curvature. In general, the boundary of a boundary value problem is *not known* in the physical Riemannian space *before* the solution of the problem. Investigations in this section show certain uncertainties exist even in the classical (unquantized) Einstein–Maxwell field theory.

In the last section the axially symmetric PIW class of solutions is dealt with. The charged versions of KTSV solutions are known[10] already. Each of these solutions gives rise to a PIW solution. This has been explicitly done here. The process is mathematically tricky (some singular transformations are involved) and the uniqueness of the PIW solution thus obtained cannot be claimed. Although Yamazaki's solutions are not completely verified, nevertheless the corresponding PIW solutions as obtained here *certainly* solve the stationary Einstein–Maxwell equations. The complex harmonic potentials of the PIW solutions can be generated from the real harmonic potentials of the static electrovac solutions. The transformation involved is a complex scale transformation of the coordinates. This may be a rationale for the procedure of Newman *et al.*[11]

2. NOTATIONS AND FIELD EQUATIONS

The space–time M_4 is assumed to be a connected semiRiemannian C^3_p-differentiable manifold of signature -2 which is endowed with a paracompact topology. It is also assumed that M_4 admits a timelike Killing motion so that the metric form can be transformed into

$$^{(4)}\Phi \equiv {}^{(4)}g_{ab}dx^a \otimes dx^b = -e^{-\omega(x)}\Phi + e^{\omega(x)}[\theta + dx^4]^2,$$
$$\Phi \equiv g_{\alpha\beta}(x)dx^\alpha \otimes dx^\beta, \quad \theta \equiv a_\alpha(x)dx^\alpha. \tag{2.1}$$

Here the Roman indices take the values 1, 2, 3, 4, the Greek indices take the values 1, 2, 3, and the summation convention is adopted. The associated space M_3 is Riemannian and it has the metric form Φ. Relative to a local chart, a point in M_3 can be mapped in a 1 : 1 fashion to the triple $(x) \equiv (x^1, x^2, x^3)$. The physical units are so chosen that $c = G = (8\pi)^{-1}\kappa = 1$.

The electromagnetic field tensor F_{ab} in M_4 is adapted to the Killing motion ($F_{ab,4} = 0$). Thus one can conclude[1] that there exist potential functions A, B such that

$$\tfrac{1}{2}F_{ab}dx^a \wedge dx^b = A_{,\nu}dx^4 \wedge dx^\nu + [a_{[\mu}A_{,\nu]} - \tfrac{1}{2}e^{-\omega}\eta_{\mu\nu\alpha}B^{|\alpha}]dx^\mu \wedge dx^\nu,$$

where $\wedge$ denotes the wedge product and the comma denotes the partial derivative. A stroke denotes the covariant derivative in M_3. The complex electromagnetic potential can be introduced by $\phi \equiv -A + iB$.

One can define two more vector fields:

$$\tau^\alpha \equiv \tfrac{1}{2}e^{2\omega}\eta^{\alpha\beta\gamma}f_{\beta\gamma},$$
$$\chi_\alpha \equiv \tau_\alpha + (i\kappa/2)(\phi^*\phi_{,\alpha} - \phi\phi^*_{,\alpha}), \tag{2.2a}$$

where $\eta_{\alpha\beta\gamma}$ is the Levi–Civita tensor and

$$\tfrac{1}{2}f_{\mu\nu}dx^\mu \wedge dx^\nu \equiv d\theta = a_{[\mu|\nu]}dx^\mu \wedge dx^\nu.$$

The gravitational equations ${}^{(4)}G_4{}^\alpha = 0$ ensure the existence of a twist potential such that $\chi_\alpha = \chi_{,\alpha}$. One can define an electrogravitational complex potential Γ by

$$\Gamma \equiv e^\omega - (\kappa/2)|\phi|^2 + i\chi. \tag{2.2b}$$

The remaining stationary Einstein–Maxwell equations can be reduced to the following system:

$$\sigma_{\alpha\beta} \equiv G_{\alpha\beta} + \tfrac{1}{2}e^{-2\omega}\mathrm{Re}[(\Gamma^*_{,\alpha} + \kappa\phi\phi^*_{,\alpha})(\Gamma_{,\beta} + \kappa\phi^*\phi_{,\beta}) - \tfrac{1}{2}g_{\alpha\beta}(\Gamma^*_{,\mu} + \kappa\phi\phi^*_{,\mu})(\Gamma^{|\mu} + \kappa\phi^*\phi^{|\mu})] - \kappa e^{-\omega}\mathrm{Re}[\phi^*_{,\alpha}\phi_{,\beta} - (\tfrac{1}{2})g_{\alpha\beta}\phi^*_{,\mu}\phi^{|\mu}] = 0, \tag{2.3}$$

$$\mu \equiv \Delta_2\Gamma - e^{-\omega}\Gamma^{|\alpha}(\Gamma_{,\alpha} + \kappa\phi^*\phi_{,\alpha}) = 0, \tag{2.4}$$

$$\nu \equiv \Delta_2\phi - e^{-\omega}\phi^{|\alpha}(\Gamma_{,\alpha} + \kappa\phi^*\phi_{,\alpha}) = 0, \tag{2.5}$$

where $\Delta_2 V \equiv g^{\alpha\beta}V_{|\alpha\beta}$.

The system of second-order quasilinear partial differential equations has for the number of unkown, real functions $10 = 6(g_{\alpha\beta}) + 2(\Gamma) + 2(\phi)$. The number of equations is $13 = 6(\sigma_{\alpha\beta}) + 2(\mu) + 2(\nu) + 3$(coordinate conditions). But the equations are related by three differential identities:

$$\sigma^{\alpha\beta}{}_{|\beta} \equiv (\kappa/2)f^\alpha,$$
$$f^\alpha \equiv \mathrm{Re}[(\mu\kappa^{-1} + \nu\phi^*)e^{-2\omega}(\Gamma^*_{,\alpha} + \kappa\phi\phi^*_{,\alpha}) - 2\nu e^{-\omega}\phi^*_{,\alpha}]. \tag{2.6}$$

In case $\sigma^{\alpha\beta} = 0$ the above identities yield

$$[\mathrm{Re}(\mu\kappa^{-1} + \nu\phi^*)](e^\omega)_{,\alpha} + [\mathrm{Im}(\mu\kappa^{-1} + \nu\phi^*)]\tau_\alpha \equiv 2e^\omega\mathrm{Re}(\nu\phi^*_{,\alpha}). \tag{2.6'}$$

This may be compared to the equation of motion for a dually charged dust in isometric motion, which can be written as follows:[12]

$$\rho(e^\omega)_{,\alpha} = 2e^\omega\mathrm{Re}(\sqrt{4\pi}\, e^{-\omega/2}\sigma^*\phi^*_{,\alpha}). \tag{2.7}$$

One can conclude that in a certain sense $\rho \equiv \mathrm{Re}(\mu\kappa^{-1} + \nu\phi^*)$ and $\sigma \equiv (4\pi)^{-1/2}e^{\omega/2}\nu^*$ play the roles of the mass density and the complex charge density, respectively.

3. THE VARIATIONAL DERIVATION OF FIELD EQUATIONS

The stationary field equations (2.3)–(2.5) are derivable from the action integral[1]

$$\mathscr{A} = \int_D [R + \tfrac{1}{2}(\mathrm{Re}\Gamma + 4\pi|\phi|^2)^{-2}(\Gamma^*_{,\alpha} + \kappa\phi\phi^*_{,\alpha}) \times(\Gamma^{|\alpha} + \kappa\phi^*\phi^{|\alpha}) - \kappa(\mathrm{Re}\Gamma + 4\pi|\phi|^2)^{-1}\phi^*_{,\alpha}\phi^{|\alpha}]d_3v. \tag{3.1}$$

However, the general Einstein–Maxwell equations are derivable from the Hilbert integral

$${}^{(4)}\mathscr{A} = \int_{(4)_D} [{}^{(4)}R - 4\pi F_{ab}{}^{(4)}F^{ab}]d_4v, \tag{3.2}$$

where $F_{ab} \equiv 2A_{[a,b]}$ must be used.

The action integral (3.1) will be derived from (3.2) in the case of the stationary metric form (2.1), using certain variational techniques.

In analytical mechanics a dynamical variable which is ignorable can be eliminated from the Lagrangian by Routh's procedure. One can go a step further to eliminate a variable for which the equation of motion has been integrated and obtain a modified Lagrangian. This modification can be generalized to any field-theoretic Lagrangian.

Suppose one has a conservative system of N degrees of freedom. The canonical Lagrangian which yields Hamilton's equations is the following[13]:

$$L(q^1,...,q^N,p_1,...,p_N,\dot{q}^1,...\dot{q}^N) = p_A\dot{q}^A + p_N\dot{q}^N - H(q^1,...,q^N,p_1,...,p_N), \tag{3.3}$$

where the capital Roman indices (except N) are summed from 1 to $N-1$. H is assumed to be a quadratic function of $p_A,...,p_N$ such that $\det(\partial^2H/\partial p_A\partial p_B)\neq 0$. Suppose the Nth pair of equations of motion,

$$\left(\frac{\partial L}{\partial \dot{p}_N}\right)^\bullet - \frac{\partial L}{\partial p_N} = -\dot{q}^N + \frac{\partial H}{\partial p_N} = 0,$$

$$\left(\frac{\partial L}{\partial \dot{q}^N}\right)^\bullet - \frac{\partial L}{\partial q^N} = \dot{p}_N + \frac{\partial H}{\partial q_N} = 0,$$

are integrated to obtain the function

$$p_N = \pi_N(t,q^A,p_A). \tag{3.4}$$

One can modify the Lagrangian to

$$\bar{L} \equiv L - p_N\dot{q}^N - \dot{p}_N q^N. \tag{3.5}$$

Here and subsequently the equation (3.4) *must* be used to eliminate p_N. It can be verified using the old equations that

$$\left(\frac{\partial \bar{L}}{\partial \dot{q}^A}\right)^\bullet - \frac{\partial \bar{L}}{\partial q^A} = \left(p_A - \frac{\partial \pi_N}{\partial q^A}q^N\right)^\bullet + \left[\left(\frac{\partial^2\pi_N}{\partial q^A\partial t} + \frac{\partial^2\pi_N}{\partial q^A\partial q^B}\dot{q}^B + \frac{\partial^2\pi_N}{\partial q^A\partial p_B}\dot{p}_B\right)q^N + \frac{\partial H}{\partial q^A} + \frac{\partial H}{\partial p_N}\frac{\partial \pi_N}{\partial q^A}\right] = 0.$$

Similarly, the other equations are also equivalent.

Now consider N twice-differentiable tensor fields $\phi_{(1)}$, ...,$\phi_{(N)}$ of arbitrary ranks in M_3 which satisfy field equations derivable from the canonical Lagrangian

$$\mathscr{L}[\phi_{(1)},...,\phi_{(N)},\pi^{(1)\alpha},...,\pi^{(N)\alpha},\phi_{(1)|\alpha},...,\phi_{(N)|\alpha}] = \pi^{(A)\alpha}\phi_{(A)|\alpha} + \pi^{(N)\alpha}\phi_{(N)|\alpha} - \mathscr{H}(\phi_{(1)},...,\phi_{(N)},\pi^{(1)\alpha},...,\pi^{(N)\alpha}), \tag{3.6}$$

where the index A is summed from 1 to $N-1$.

Suppose that the Nth field equations

$$\left(\frac{\partial \mathscr{L}}{\partial \pi_{(N)\alpha|\beta}}\right)_{|\beta} - \frac{\partial \mathscr{L}}{\partial \pi^{(N)\alpha}} = -\phi_{(N)|\alpha} + \frac{\partial \mathscr{H}}{\partial \pi^{(N)\alpha}} = 0,$$

$$\left(\frac{\partial \mathscr{L}}{\partial \phi_{(N)|\alpha}}\right)_{|\alpha} - \frac{\partial \mathscr{L}}{\partial \phi_{(N)}} = \pi^{(N)\alpha}{}_{|\alpha} + \frac{\partial \mathscr{H}}{\partial \phi_{(N)}} = 0,$$

can be integrated to obtain a known function

$$\pi^{(N)\alpha} = p^{(N)\alpha}(x,\phi_{(A)},\pi^{(A)\beta}). \tag{3.7}$$

The modified Lagrangian in which $\pi^{(N)\alpha}$ is eliminated by substituting (3.7), can be written as

$$\bar{\mathscr{L}} = \mathscr{L} - \pi^{(N)\alpha}\phi_{(N)|\alpha} - \pi^{(N)\alpha}{}_{|\alpha}\phi_{(N)}. \tag{3.8}$$

The new field equations

$$\left(\frac{\partial \bar{\mathscr{L}}}{\partial \phi_{(A)|\alpha}}\right)_{|\alpha} - \frac{\partial \bar{\mathscr{L}}}{\partial \phi_{(A)}} = \left(\pi^{(A)\alpha} - \frac{\partial p^{(N)\alpha}}{\partial \phi_{(A)}}\phi_{(N)}\right)_{|\alpha} + \left[\left(\frac{\partial p^{(N)\alpha}{}_{,\alpha}}{\partial \phi_{(A)}} + \left\{{}^{\alpha}_{\beta\alpha}\right\}\frac{\partial p^{(N)\beta}}{\partial \phi_{(A)}} + \frac{\partial^2 p^{(N)\alpha}}{\partial \phi_{(A)}\partial \phi_{(B)}}\phi_{(B)|\alpha} + \frac{\partial^2 p^{(N)\alpha}}{\partial \phi_{(A)}\partial \pi^{(A)\beta}}\pi^{(A)\beta}_{,\alpha}\right)\phi_{(N)} + \frac{\partial \mathscr{H}}{\partial \phi_{(A)}} + \frac{\partial \mathscr{H}}{\partial \pi^{(N)\beta}}\frac{\partial p^{(N)\beta}}{\partial \phi_{(A)}}\right] = 0$$

are true by virtue of the old equations. Similarly, the other equations are also implied by the old ones.

Now the Hilbert action integral (3.2) will be considered choosing the stationary metric form (2.1). The equations $F_{ab,4} = 0$ are assumed to hold. A domain inside a coordinate neighborhood of M_4 is chosen such that it can be mapped into ${}^{(4)}D = D \times \{x^4|t_1 < x^4 < t_2\} \subset R^4, D \subset R^3$. Then the Hilbert action integral reduces[5] to

$${}^{(4)}\mathscr{A} = -(t_2 - t_1)\int_D [R + \tfrac{1}{2}\omega_{,\alpha}\omega^{|\alpha} - \Delta_2\omega - (\tfrac{1}{4})e^{2\omega}f_{\alpha\beta}f^{\alpha\beta} + 4\pi e^{\omega}\{F_{\alpha\beta}F^{\alpha\beta} + 4a_\alpha F^{\alpha\beta}A_{4,\beta} + 2(g^{\alpha\beta}a_\lambda a^\lambda - a^\alpha a^\beta - e^{-2\omega}g^{\alpha\beta})A_{4,\alpha}A_{4,\beta}\}]d_3v.$$

The term $\int_D \Delta_2\omega d_3v$ can be dropped since it can be converted into a surface integral. The resulting Lagrangian does not contain the potential A_α explicitly. Therefore, $\pi^{\beta\alpha}{}_{|\alpha} = \partial\mathscr{L}/\partial A_\beta = 0$. The modified Lagrangian according to (3.8) is

$$\bar{\mathscr{L}} = \mathscr{L} - \pi^{\alpha\beta}A_{\alpha|\beta} = \mathscr{L} - \frac{\partial \mathscr{L}}{\partial F_{\alpha\beta}}F_{\alpha\beta}. \tag{3.9}$$

Now the field equations for A_α yield the first integral,

$$F_{\mu\nu} = a_\mu A_{4,\nu} - a_\nu A_{4,\mu} - e^{-\omega}\eta_{\mu\nu\alpha}B^{|\alpha},$$

where B is an arbitrary C^2-function of integration.

This expression must be substituted in (3.9) to eliminate $F_{\alpha\beta}$. After a long computation one obtains the modified Lagrangian as

$$\bar{\mathscr{L}} = R + \tfrac{1}{2}\omega_{,\alpha}\omega^{|\alpha} - \tfrac{1}{4}e^{2\omega}f_{\alpha\beta}f^{\alpha\beta} - \kappa[\eta_{\mu\nu\beta}(a^\nu g^{\mu\alpha} - a^\mu g^{\nu\alpha})A_{,\alpha}B^{|\beta} + e^{-\omega}(A_{,\alpha}A^{|\alpha} + B_{,\alpha}B^{|\alpha})], \tag{3.10}$$

where $A \equiv A_4$.

Now the field equations for a_α can be integrated to obtain

$$f^{\alpha\beta} \equiv e^{-2\omega}\eta^{\alpha\beta\gamma}\tau_\gamma = e^{-2\omega}\eta^{\alpha\beta\gamma}[\chi_{,\gamma} + \kappa(BA_{,\gamma} - AB_{,\gamma})], \tag{3.11}$$

where χ is an arbitrary C^2-function.

Modifying the Lagrangian for the second time according to (3.8) one obtains

$$\bar{\bar{\mathscr{L}}} = \bar{\mathscr{L}} - \frac{\partial \bar{\mathscr{L}}}{\partial f^{\alpha\beta}} f^{\alpha\beta} - \frac{\partial \bar{\mathscr{L}}}{\partial a^\alpha} a^\alpha,$$

where (3.11) must be substituted to eliminate $f^{\alpha\beta}$, a^α. After another long computation, using the definition $\phi = -A + iB$ and (2.2b) for Γ, one can finally reduce $\bar{\bar{\mathscr{L}}}$ exactly to the integrand of the action integral (3.1).

4. THE CLASSIFICATION OF THE STATIONARY FIELD EQUATIONS

It would be appropriate to recall some basic definitions for a system of p.d.e.'s. A system of second-order, quasilinear equations[3] can be written as

$$L^M(u^B) + d^M \equiv A^{\alpha\beta M}{}_B u^B_{,\alpha\beta} + d^M = 0, \tag{4.1}$$

where the $k \times k$ matrices $A^{\alpha\beta} = A^{\beta\alpha}$ and the k-vector d^M are given functions of x^α, u^B, $u^B_{,\alpha}$. The characteristic determinant is defined as

$$Q(S) \equiv \det[A^{\alpha\beta M}{}_B S_{,\alpha} S_{,\beta}], \tag{4.2}$$

where $S(x) = 0$ is assumed to be a differentiable surface. In case $Q(S) = 0 \Rightarrow S_{,1} = S_{,2} = S_{,3} = 0$, the system is elliptic. In case $Q(S) = 0$ implies that there exist locally $2k$ distinct surfaces $S(x) = 0$ (characteristic surfaces), the system is totally hyperbolic. In case $Q(S) = 0$ locally allows the existence of l surfaces $S = 0$ such that $0 < l < 2k$, the system is either ultrahyperbolic or parabolic.

The usual transformations under a change of coordinate system are

$$\hat{x}^\alpha = \hat{x}^\alpha(x), \quad \hat{S}(\hat{x}) = S(x), \quad \hat{u}^B(\hat{x}) = u^B(x),$$
$$\hat{A}^{\mu\nu M}{}_B = \frac{\partial \hat{x}^\mu}{\partial x^\alpha}\frac{\partial \hat{x}^\nu}{\partial x^\beta} A^{\alpha\beta M}{}_B. \tag{4.3}$$

Therefore the characteristic determinant transforms as

$$\hat{Q}[\hat{S}(\hat{x})] = Q[S(x)], \tag{4.4}$$

and thus the classification scheme remains unaffected by a coordinate transformation. But in general relativity the p.d.e.'s involve as unknown functions u^B the metric tensor components, and the coefficient functions $A^{\alpha\beta M}{}_B$ involve the metric tensor and its first partial derivatives. These functions undergo transformations which are different than (4.3). That is why the classification scheme in general relativity is coordinate-dependent.

For example, if one uses the harmonic[14] coordinate condition $(g^{1/2}g^{\alpha\beta})_{,\beta} = 0$ for the associated space M_3, the field equations (2.3)–(2.5) go over to

$$\begin{aligned} \sigma^{\mu\nu} &\equiv g^{\alpha\beta}(g^{\mu\nu})_{,\alpha\beta} + \cdots = 0, \\ \mu &\equiv g^{\alpha\beta}\Gamma_{,\alpha\beta} + \cdots = 0, \\ \nu &\equiv g^{\alpha\beta}\phi_{,\alpha\beta} + \cdots = 0, \end{aligned} \tag{4.5}$$

where the dots represent lower-order derivatives. Defining $(u^1, u^2, \ldots, u^{10}) \equiv (g^{11}, g^{12}, \ldots, \mathrm{Im}\phi)$, the characteristic determinant for the above system is $Q(S) = (g^{\alpha\beta}S_{,\alpha}S_{,\beta})^{10}$ and the system is elliptic since $g^{\alpha\beta}$ is positive-definite.

On the other hand, if one chooses orthogonal coordinates for M_3, i.e.,

$$[g_{\alpha\beta}] = \begin{bmatrix} e^{2\alpha} & 0 & 0 \\ 0 & e^{2\beta} & 0 \\ 0 & 0 & e^{2\gamma} \end{bmatrix}. \tag{4.6}$$

the field equations (2.3)–(2.5) yield

$$\begin{aligned} \sigma_{23} &\equiv \alpha_{,23} + \cdots = 0, \\ \sigma_{31} &\equiv \beta_{,31} + \cdots = 0, \\ \sigma_{12} &\equiv \gamma_{,12} + \cdots = 0, \\ \mu &\equiv e^{-2\alpha}\Gamma_{,11} + e^{-2\beta}\Gamma_{,22} + e^{-2\gamma}\Gamma_{,33} = 0, \\ \nu &\equiv e^{-2\alpha}\phi_{,11} + e^{-2\beta}\phi_{,22} + e^{-2\gamma}\phi_{,33} = 0. \end{aligned} \tag{4.7}$$

Defining $(u^1, u^2, \ldots, u^7) \equiv (\alpha, \beta, \ldots, \mathrm{Im}\phi)$, the characteristic determinant of the system is

$$Q(S) = (S_{,1}S_{,2}S_{,3})^2[e^{-2\alpha}(S_{,1})^2 + e^{-2\beta}(S_{,2})^2 + e^{-2\gamma}(S_{,3})^2]^4. \tag{4.8}$$

The system is therefore a mixed or ultrahyperbolic[3] type with three characteristic surfaces locally given by $S_{,1} = 0$, $S_{,2} = 0$, or $S_{,3} = 0$. The remaining equations from (2.3) are

$$\begin{aligned} \sigma_{11} &= e^{2\alpha}(e^{-2\beta}\gamma_{,22} + e^{-2\gamma}\beta_{,33}) + \cdots = 0, \\ \sigma_{22} &= e^{2\beta}(e^{-2\alpha}\gamma_{,11} + e^{-2\gamma}\alpha_{,11}) + \cdots = 0, \\ \sigma_{33} &= e^{2\gamma}(e^{-2\alpha}\beta_{,11} + e^{-2\beta}\alpha_{,22}) + \cdots = 0. \end{aligned} \tag{4.9}$$

In this coordinate system the diagonal equations $\sigma_{11} = \sigma_{22} = \sigma_{33} = \mu = \nu = 0$ give a characteristic determinant which is just twice the expression (4.8) and thus yields the same classification. However, the choice of the system $\sigma_{23} = \sigma_{31} = \sigma_{33} = \mu = \nu = 0$ yields $Q(S) \equiv 0$ and that peculiarity can be traced back to the choice of an overdetermined system so far as the principal parts are concerned.

There exist no algebraic identities for the 10 equations (2.3)–(2.5), although there are three differential identities (2.6). Therefore one has to check the satisfaction of *all* the equations for an exact solution.

The Cauchy problem for the system (4.7) has no solution if the data is prescribed on a characteristic surface ($S_{,1} = 0$, $S_{,2} = 0$, or $S_{,3} = 0$), unless the data satisfy some consistency conditions. In case these conditions are fulfilled, multiple solutions of the problem will exist. Even though the system (4.5) is elliptic, the initial data for it must also satisfy some consistency conditions. These conditions do not come directly from the system itself, but arise from the harmonic coordinate conditions, which must be satisfied by the initial data.

5. EIGENVALUES OF THE RICCI TENSOR

The associated space M_3 can be classified algebraically according to the eigenvalues of the Ricci tensor $R_{\alpha\beta}$. In case all three real eigenvalues of $R_{\alpha\beta}$ coincide, the corresponding class of M_4 will be called stationary-$\mathscr{C}_1$. For the stationary-$\mathscr{C}_2$ class, two eigenvalues coincide and the other one is distinct. The three eigenvalues are all distinct for the class of stationary-$\mathscr{C}_3$ solutions.

For the purpose of determining the eigenvalues, six of the field equations are written as

$$\tilde{\sigma}_{\mu\nu} \equiv R_{\mu\nu} + 2\,\mathrm{Re}(\beta_\mu \beta^*{}_\nu - \alpha_\mu \alpha^*{}_\nu) = 0,$$
$$\beta_\mu \equiv \tfrac{1}{2} e^{-\omega}(\Gamma_{,\mu} + \kappa\phi^*\phi_{,\mu}), \qquad (5.1)$$
$$\alpha_\mu \equiv \sqrt{4\pi}\, e^{-(\omega/2)}\phi_{,\mu}.$$

The eigenvalue equation is

$$-\det(R^\mu{}_\nu - \lambda\delta^\mu{}_\nu) = \lambda^3 - I_1\lambda^2 + I_2\lambda - I_3 = 0,$$
$$I_1 \equiv 2(\alpha^*{}_\mu\alpha^\mu - \beta^*{}_\mu\beta^\mu),$$
$$I_2 \equiv (\alpha^*{}_\mu\alpha^\mu)^2 - |\alpha_\mu\alpha^\mu|^2 + (\beta^*{}_\mu\beta^\mu)^2 - |\beta_\mu\beta^\mu|^2 + 2|\alpha^\mu\beta_\mu|^2 + 2|\alpha^*{}_\mu\beta^\mu|^2 - 4(\alpha^*{}_\mu\alpha^\mu)(\beta^*{}_\nu\beta^\nu), \qquad (5.2)$$
$$I_3 \equiv 2|\eta^{\mu\nu\lambda}\alpha_\mu\beta_\nu\beta^*{}_\lambda|^2 - 2|\eta^{\mu\nu\lambda}\beta_\mu\alpha_\nu\alpha^*{}_\lambda|^2.$$

The above cubic equation can be solved for the real, invariant eigenvalues:

$$\lambda_1 = (I_1/3) + 2\,[(I_1/3)^2 - (I_2/3)]^{1/2}\cos J,$$
$$\lambda_2 = (I_1/3) - [(I_1/3)^2 - (I_2/3)]^{1/2}(\cos J + \sqrt{3}\sin J), \qquad (5.3)$$
$$\lambda_3 = (I_1/3) - [(I_1/3)^2 - (I_2/3)]^{1/2}(\cos J - \sqrt{3}\sin J),$$
$$\cot(3J) \equiv [I_3 - (I_1 I_2/3) + 2(I_1/3)^3]\{4[(I_1/3)^2 - (I_2/3)]^3 - [I_3 - (I_1 I_2/3) + 2(I_1/3)^3]^2\}^{1/2}.$$

In some sense the principal stresses of the stationary Einstein–Maxwell fields are given by $\frac{1}{2}(\lambda_1 - \lambda_2 - \lambda_3)$, $\frac{1}{2}(\lambda_2 - \lambda_3 - \lambda_1)$, $\frac{1}{2}(\lambda_3 - \lambda_1 - \lambda_2)$.

For the stationary-$\mathscr{C}_2$ class, Eqs. (5.3) reduce to

$$\lambda_1 = (I_1/3) + 2[(I_1/3)^2 - (I_2/3)]^{1/2},$$
$$\lambda_2 = \lambda_3 = (I_1/3) - [(I_1/3)^2 - (I_2/3)]^{1/2}. \qquad (5.4)$$

For the stationary-$\mathscr{C}_1$ class the eigenvalues simplify to

$$\lambda_1 = \lambda_2 = \lambda_3 = I_1/3. \qquad (5.5)$$

All the known solutions of the stationary Einstein–Maxwell equations happen to satisfy $I_3 = \det[R^\mu{}_\nu] = 0$. In this case the eigenvalue $\lambda_1 = 0$.

The eigenvalues in (5.3) reduce to those given for the stationary vacuum case[5] by putting $\alpha_\mu \equiv 0$. In that case all the eigenvalues were nonpositive. However, in the electromagnetic generalization, the eigenvalues (5.3) are *not* in general nonpositive. A theorem will be proved to that effect.

Theorem 5.1: The eigenvalues of the Ricci tensor $R^\mu{}_\nu$ as given in (5.3) are real and invariant with respect to coordinate transformations in M_3. Furthermore, these eigenvalues cannot all have the same sign.

Proof: The first part of the theorem is well known. For the second part let $e_E{}^\mu$, $(E = 1,2,3)$ be a set of orthonormal eigenvectors with eigenvalues λ_E. The invariant components are denoted by capital Roman indices for which summation is suspended. Thus $R_{\mu\nu}e_F{}^\nu = \lambda_F e_{F\mu}$. Writing $a_E + ib_E \equiv (\surd 2)\alpha_\mu e_E{}^\mu$, $c_E + id_E \equiv (\surd 2)\beta_\mu e_E{}^\mu$, the invariant form of (5.1) is

$$a_E a_F + b_E b_F - c_E c_F - d_E d_F = \lambda_E \delta_{EF}. \qquad (5.6)$$

Let $u_{(E)} \equiv (a_E, b_E, c_E, d_E)$ and $v_{(E)} \equiv (a_E, b_E, -c_E, -d_E)$. Then there exists a nonzero vector $u_{(4)} = (a_4, b_4, c_4, d_4)$ in R^4 which is orthogonal to the $v_{(E)}$'s for $E = 1,2,3$. If $[A]$ is a 4×4 matrix which has $u_{(E)}$ as its Eth row, one has the following:

$$[A]\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{bmatrix}[A]^T = \begin{bmatrix}\lambda_1&0&0&0\\0&\lambda_2&0&0\\0&0&\lambda_3&0\\0&0&0&l\end{bmatrix}, \qquad (5.7)$$

where $l \equiv a_4^2 + b_4^2 - c_4^2 - d_4^2$ and (5.6) has been used. By way of contradiction suppose all the λ_E's are positive. Then one can show that $[A]$ is nonsingular. For, if $0 = \Sigma_{i=1}^4 m_i u_{(i)}$, multiplying by $v_{(E)}{}^T$ from the right gives $0 = m_E\lambda_E$. Thus $m_1 = m_2 = m_3 = 0$, and since $u_{(4)}$ is nonzero, $m_4 = 0$. Now equation (5.7) shows that our supposition contradicts Sylvester's inertial theorem.[15]

The above theorem is used to obtain the whole class of stationary-$\mathscr{C}_1$ solutions in the next theorem.

Theorem 5.2: The class of stationary-$\mathscr{C}_1$ solutions of the Einstein–Maxwell equation is exactly the class of PIW solutions.[9]

Proof: For the stationary-$\mathscr{C}_1$ class, $\lambda_1 = \lambda_2 = \lambda_3$. By the preceding Theorem 5.1 the eigenvalues must all reduce to zero. Therefore $R_{\mu\nu} = 0$ and M_3 is flat. This is precisely the criterion for the PIW solutions (also see Sec. 8).

A geometrical consequences of this theorem is that M_3 cannot be a space of (nonzero) constant curvature[16] (or an Einstein space, or a projectively flat space).

A physical consequence is that stationary electrogravitational fields never generate (nontrivial, isotropic) pressure.

6. GLOBAL INTEGRABILITY OF THE FIELD EQUATIONS AND THE EQUILIBRIUM CONDITIONS

The integrability of the static and stationary field equations and the implications for the equilibrium of sources have been discussed in many papers.[6] In this section a more general treatment of this topic will be given.

A massive, charged world-tube in a stationary space–time can be sliced by a time-constant hyperplane. This spatial slice can be mapped into a region $\bar{\beta} = \beta \cup \partial\beta \subset R^3$. The region $\bar{\beta}$ will correspond to a regular, isometric body provided the following assumptions hold.

(i) $\bar{\beta}$ is a bounded, connected subset of R^3. The boundary $\partial\beta$ is assumed to be a continuous, piecewise-differentiable, orientable, closed surface (regular). A surrounding domain D-$\bar{\beta}$ is *assumed* to be simply-connected.

(ii) In the corresponding world-tube the isometric conditions ${}^{(4)}F^{ab}{}_{|b} = {}^{(4)}\tilde{F}^{ab}{}_{|b} = {}^{(4)}G^a{}_4 - \kappa\,{}^{(4)}F^{ab}F_{4b} = 0$ hold. These equations, with appropriate conditions of differentiability, guarantee the existence of potentials B, χ inside the body $\bar{\beta}$.

(iii) Inside $\bar{\beta}$ one must have the strict inequality $\sigma_{\alpha\beta}\sigma^{\alpha\beta}$

$+ |\mu|^2 + |\nu|^2 > 0$, or, at least one of the vacuum equations is violated.

(iv) Inside and outside the body $\bar{\beta}$ the differentiability classes of $g_{\alpha\beta}, \Gamma, \phi$ are respectively C^3, C^2, C^2. However, on $\partial\beta$ the functions $g_{\alpha\beta,\mu\nu}, \Gamma_{,\alpha\beta}, \phi_{,\alpha\beta}$ are allowed to have jump discontinuities. But the jump conditions[17]

$$\sigma_{\alpha\beta}n^\beta = [C], \quad \Gamma_{,\alpha}n^\alpha = [C], \quad \phi_{,\alpha}n^\alpha = [C],$$

are assumed to hold across the surface of discontinuity $\partial\beta$.

For the extreme cases where $\bar{\beta}$ degenerates into a surface, a curve, or a point, most of the above assumptions become meaningless. Nevertheless, describing $\sigma_{\alpha\beta}, \mu, \nu$ by Dirac delta functions, the subsequent global integrability conditions can be made mathematically rigorous and physically meaningful.

Now the integrability conditions will be studied. Some of the stationary Einstein–Maxwell equations can be rewritten as

$$2A_{[\alpha,\beta]} = F_{\alpha\beta} \equiv -\eta_{\alpha\beta\gamma}H^\gamma, \tag{6.1a}$$

$$2a_{[\alpha,\beta]} \equiv f_{\alpha\beta} = e^{-2\omega}\eta_{\alpha\beta\gamma}\tau^\gamma, \tag{6.1b}$$

$$\sigma_{\alpha\beta} \equiv G_{\alpha\beta} - S_{\alpha\beta} = 0, \tag{6.1c}$$

$$\begin{aligned} S_{\alpha\beta} \equiv {}& \kappa e^{-\omega}\mathrm{Re}[\phi^*_{,\alpha}\phi_{,\beta} - \tfrac{1}{2}g_{\alpha\beta}\phi^*_{,\mu}\phi^{|\mu}] \\ & - \tfrac{1}{2}e^{-2\omega}\mathrm{Re}[(\Gamma^*_{,\alpha} + \kappa\phi\phi^*_{,\alpha})(\Gamma_{,\beta} + \kappa\phi^*\phi_{,\beta}) \\ & - \tfrac{1}{2}g_{\alpha\beta}(\Gamma^*_{,\mu} + \kappa\phi\phi^*_{,\mu})(\Gamma^{|\mu} + \kappa\phi^*\phi^{|\mu})] = 0. \end{aligned}$$

The *local* integrability conditions for these equations are the following:

$$H^\gamma{}_{|\gamma} \equiv e^{-\omega}\mathrm{Im}(\nu) = 0, \tag{6.2a}$$

$$(e^{-2\omega}\tau^\gamma)_{|\gamma} \equiv e^{-2\omega}\mathrm{Im}(\mu + \kappa\phi^*\nu) = 0, \tag{6.2b}$$

$$2\sigma^{\alpha\gamma}{}_{|\gamma} \equiv \kappa f^\alpha = 0, \tag{6.2c}$$

where the electrogravitational force f^α is defined in (2.6). Under the differentiability assumptions for the functions involved, these local integrability conditions are *identically* satisfied outside bodies. These conditions can be generalized for the neighborhood of a body. Consider a regular, exterior surface Σ that encloses a body $\bar{\beta}$. The relatively global integrability conditions of (6.1a)–(6.1c) are respectively

$$\oint_\Sigma H^\gamma n_\gamma d_2s = \oint_\Sigma [e^{-\omega}B_{,\gamma} - \eta_{\alpha\beta\gamma}a^\alpha A^{|\beta}]n^\gamma d_2s = 0, \tag{6.2d}$$

$$\oint_\Sigma e^{-2\omega}\tau^\gamma n_\gamma d_2s = 0, \tag{6.2e}$$

$$\oint_\Sigma \sigma^{\alpha\gamma}\lambda_\alpha n_\gamma d_2s = 0, \tag{6.2f}$$

where λ_α is an arbitrary vector field. These are *weaker* than (6.2a)–(6.2c). For global integrability, the equations (6.2d)–(6.2f) hold for *every* closed, exterior surface Σ.

In case there are jump discontinuities $[H^\gamma n_\gamma]$, $[e^{-2\omega}\tau^\gamma n_\gamma]$ across the boundary $\partial\beta$ of a body, the equations (6.2d), (6.2e), by Gauss's theorem, imply that

$$\int_\beta e^{-\omega}\mathrm{Im}(\nu)d_3v + \oint_{\partial\beta}[H^\gamma n_\gamma]d_2s = 0, \tag{6.3a}$$

$$\int_\beta e^{-2\omega}\mathrm{Im}(\mu + \kappa\nu\phi^*)d_3v + \oint_{\partial\beta}[e^{-2\omega}\tau^\gamma n_\gamma]d_2s = 0. \tag{6.3b}$$

The equation (6.3a) physically means that the sum of the total volume magnetic charge and the total boundary layer magnetic charge of a body $\bar{\beta}$ is zero. Similarly (6.3b) means that the total volume twist-mass and the total boundary layer twist-mass add up to zero. These two equations need not hold in case the body has a wire (or string) singularity. In that case every enclosing surface would be perforated and Gauss's theorem need not apply.

The integral condition (6.2f) can be expressed using the jump condition $[\sigma^{\alpha\gamma}n_\gamma]_{\partial\beta} = [C]$, as follows.

$$\begin{aligned} 0 &= \oint_\Sigma \sigma^{\alpha\gamma}\lambda_\alpha n_\gamma d_2s = \oint_{\partial\beta}\sigma^{\alpha\gamma}\lambda_\alpha n_\gamma d_2s = \int_\beta(\sigma^{\alpha\gamma}\lambda_\alpha)_{|\gamma}d_3v \\ &= \int_\beta[-S^{\alpha\gamma}{}_{|\gamma}\lambda_\alpha + \sigma^{\alpha\gamma}\lambda_{(\alpha|\gamma)}]d_3v \\ &= \tfrac{1}{2}\int_\beta(\kappa f_\alpha\lambda^\alpha + 2\sigma^{\alpha\gamma}\lambda_{(\alpha|\lambda)})d_3v. \end{aligned} \tag{6.4}$$

The above integral condition implies the equilibrium of a body in the most general language. However, for special choices of λ^α, the above condition can be simplified. Suppose λ^α is a displacement vector for a deformable body $\bar{\beta}$. One possible characterization of a "rigid" displacement is the requirement the total "strain-energy" vanish,

$$\int_\beta \sigma^{\alpha\gamma}\lambda_{(\alpha|\gamma)}d_3v = 0.$$

This condition is implied by the partial differential equation

$$\sigma^{\alpha\gamma}\lambda_{(\alpha|\gamma)} = \eta^{\gamma\alpha\beta}\mu_{[\alpha|\beta]|\gamma}, \tag{6.5}$$

where $\lambda_\alpha \in C^1(\beta)$, $\mu_\alpha \in C^1(\partial\beta) \cap C^2_p(\beta)$. The equation (6.5) is undetermined since it involves six functions and thus infinitely many solutions are expected to exist. For every "rigid" displacement λ^α, Eq. (6.4) reduces to

$$\int_\beta f_\alpha\lambda^\alpha d_3v = 0. \tag{6.6}$$

Corresponding to each λ^α one can define an "infinitesimal rigid displacement" $\delta X^\alpha \equiv \epsilon\lambda^\alpha$, where ϵ is a small positive constant. The last integral goes over to

$$\int_\beta f_\alpha \delta X^\alpha d_3v = 0. \tag{6.7}$$

This is the statement of the principle of virtual work in general relativity for a charged, gravitating, "rigid" body in isometric motion.

Now some special solutions of the p.d.e. (6.5) will be given.

(i) In case[18] there exists a Killing vector field ξ_α in β one can choose the solution $\lambda_\alpha = \xi_\alpha$, $\mu_\alpha \in C^2(\beta)$. In this case the integrand in (6.6) vanishes identically *assuming* $\mathscr{L}_\xi\phi = \mathscr{L}_\xi\Gamma = 0$.

(ii) If λ^α solves[19] $G^{\alpha\beta}\lambda_{(\alpha|\beta)} = 0 = S^{\alpha\beta}\lambda_{(\alpha|\beta)}$ then it solves (6.5) with $\mu_\alpha = 0$. The general solution of $G^{\alpha\beta}\lambda_{(\alpha|\beta)} = 0$ is $G^{\alpha\beta}\lambda_\beta = \eta^{\alpha\beta\gamma}v_{[\beta|\gamma]}$, where $v \in C^2(\beta)$ and otherwise arbitrary. If, furthermore, $\det[G_{\mu\nu}] \neq 0$, $G^-_{\mu\nu} \equiv [G^{\mu\nu}]^{-1}$, then the equation $0 = S^{\mu\nu}\lambda_{(\mu|\nu)} = S^{\mu\nu}\eta^{\alpha\beta\gamma}v_{[\beta|\gamma]|(\nu}G^-{}_{\mu)\alpha}$ should have infinitely many solutions for v_β. To obtain nontrivial equilibrium conditions, one should choose solutions v_α such that $\eta^{\alpha\beta\gamma}v_{[\beta|\gamma]}$ has at least one singularity on $\partial\beta$.

(iii) Suppose one has $\sigma_{\alpha\beta}\equiv 0$, $\mu\not\equiv 0$, $\nu\not\equiv 0$ in β. Furthermore, let a linear holomorphic relationship $\Gamma=\gamma\phi+\delta$ (where γ,δ are complex constants) exist inside and outside $\bar{\beta}$. One can choose as solutions of (6.5) arbitrary $\lambda^\alpha \in C^1(\beta)$, $\mu_\alpha \in C^2(\beta)$. Then the equilibrium condition (6.6) boils down to

$$(|\gamma|^2-2\kappa\mathrm{Re}\delta)\int_\beta \mathrm{Re}[\mu\Gamma^*_{,\alpha}]e^{-2\omega}\lambda^\alpha d_3v=0. \qquad (6.8)$$

The above condition can be trivially satisfied by choosing $|\gamma|^2=2\kappa\mathrm{Re}\delta$ and outside β this class of solutions corresponds precisely to the PIW class (Sec. 8). The other integrability conditions (6.3a), (6.3b), by using the definitions after (2.7), reduce to

$$\sqrt{4\pi}\int_\beta(\mathrm{Im}\sigma^*)e^{-3\omega/2}d_3v+\oint_{\partial\beta}[H^\gamma n_\gamma]d_2s=0,$$

$$\kappa\int_\beta(\sigma^*\sigma-\rho^2)^{1/2}e^{-2\omega}d_3v+\oint_{\partial\beta}[e^{-2\omega}\tau^\gamma n_\gamma]d_2s=0.$$

In the dually charged dust model[12] the above conditions were satisfied by putting $[H^\gamma n_\gamma]=[e^{-2\omega}\tau^\gamma n_\gamma]=0$, $\rho^2=\sigma^*\sigma$.

(iv) In the axially symmetric case, the WLP coordinates are *assumed* to exist in β. One can choose a vector field $v_\alpha=-(\frac{1}{2})e^{-2u}\delta_{\alpha 3}$ which is not well defined on the x^2-axis. One can construct $\lambda^\alpha=G^{-\alpha}{}_\mu\eta^{\mu\beta\gamma}v_{(\beta|\gamma)}$, which happens to have a removable type of singularity on the x^2 axis. If that behavior is removed, then $\lambda_\alpha=\delta_{\alpha 2}$ and satisfies[6] $G^{\alpha\beta}\lambda_{(\alpha|\beta)}=S^{\alpha\beta}\lambda_{(\alpha|\beta)}=0$. Putting this λ_α into (6.6), one obtains

$$\int_\beta \mathrm{Re}[(\mu\kappa^{-1}+\nu\phi^*)(\Gamma^*_{,2}+\kappa\phi\phi^*_{,2})-2\nu e^\omega\phi^*_{,2}]$$
$$\times e^{-2(\omega+u)}x^1dx^1\wedge dx^2\wedge dx^3=0, \qquad (6.9)$$

on in other words "the x^2th component of the total force on the body must vanish." In this symmetry there eixsts the Killing vector $\xi^\alpha=\delta^\alpha{}_3=\lambda^\alpha$. Substituting this vector into (6.6) the integral becomes the x^2th component of total moment and that vanishes identically.

In view of the arbitrary functions v_α which can generate λ^α, (either in the axially symmetric or in the general case) there may be several *independent* relatively global integrability conditions of the field equations. The investigations of these conditions remain open.

7. POTENTIAL THEORY AND BOUNDARY VALUE PROBLEMS

The action integral for the stationary Einstein–Maxwell equations (2.3)–(2.5) is given by (3.1). If one allows boundary variations for that action integral, then the following equation, besides the field equations, are also obtained:

$$[g^{\alpha\beta}\delta\{{}^{\gamma}_{\alpha\gamma}\}-g^{\alpha\gamma}\delta\{{}^{\beta}_{\alpha\gamma}\}+(\mathrm{Re}\Gamma+4\pi|\phi|^2)^{-2}$$
$$\times\mathrm{Re}\{(\Gamma^{|\beta}+\kappa\phi^*\phi^{|\beta})\delta\Gamma^*\}$$
$$-2\kappa(\mathrm{Re}\Gamma+4\pi|\phi|^2)^{-1}\mathrm{Re}(\phi^{|\beta}\delta\phi^*)]n_{\beta|\partial D}$$
$$=0. \qquad (7.1)$$

The boundary conditions which satisfy (7.1) are called variationally natural boundary conditions. One example is

$$\{{}^{\alpha}_{\beta\gamma}\}_{|\partial D}=\Gamma^{|\beta}n_{\beta|\partial D}=\phi^{|\beta}n_{\beta|\partial D}=0.$$

The field equations, together with a set of variationally natural boundary conditions, constitute a self-adjoint problem for the semilinear system.

The real part of the complex potential equation (2.4) is

$$M\equiv\Delta_2\omega+e^{-2\omega}\tau_\alpha\tau^\alpha-\kappa e^{-\omega}\phi^*_{,\alpha}\phi^{|\alpha}=0. \qquad (7.2)$$

In case of stationary vacuum $\phi_{,\alpha}\equiv 0$ and $\Delta_2\omega\leqslant 0$, and thus ω is superharmonic. On the other hand, if M_4 is static then $\tau_\alpha\equiv 0$ and $\Delta_2\omega\geqslant 0$ and thus ω is subharmonic. In both these special cases if ω is regular everywhere and attains a constant value κ at infinity, then M_4 is flat.[8] In the general case, however, ω is neither superharmonic nor subharmonic and its regularity everywhere does not imply flatness of M_4. In fact, there exist nontrivial solutions of the stationary Einstein–Maxwell equations with $\omega\equiv\ln|g_{44}|=0$ *everywhere*. A class of these solutions is given by[20]

$$^{(4)}g_{ab}dx^a\otimes dx^b=-\Phi+(\theta+dx^4)\otimes(\theta+dx^4),$$
$$\Phi=e^{2u(x^1,x^2)}(dx^1\otimes dx^1+dx^2\otimes dx^2)+(x^1)^2dx^3\otimes dx^3,$$
$$\theta=a(x^1,x^2)dx^3,$$
$$\nabla^2H\equiv H_{,11}+(x^1)^{-1}H_{,1}+H_{,22}=0,$$
$$\Gamma=(\tfrac{1}{2}),\quad \phi=(\kappa)^{-1/2}e^{iH}, \qquad (7.3)$$
$$u=(\tfrac{1}{4})\int x^1[\{(H_{,2})^2-(H_{,1})^2\}dx^1-2H_{,1}H_{,2}dx^2],$$
$$a=\int x^1[-H_{,2}dx^1+H_{,1}dx^2].$$

One can pose a boundary value problem for the quasilinear elliptic equation (7.2). The uniqueness of the solution is guaranteed by the following criteria[3]: (i) $[\partial M/\partial\omega_{,\alpha\beta}]=[g^{\alpha\beta}]$ is positive definite and (ii) $\partial M/\partial\omega=\kappa e^{-\omega}\phi^*_{,\alpha}\phi^{|\alpha}-2e^{-2\omega}\tau_\alpha\tau^\alpha\leqslant 0$. The second criterion is not necessarily true and the uniqueness of a solution (7.2) *cannot* be expected in general. However, for the PIW class of solutions (Sec. 8), the four quasilinear elliptic equations reduce to two ordinary potential equations written in the complex language as

$$\nabla^2U=0. \qquad (7.4)$$

Here $U\equiv(2\kappa)^{1/2}(\gamma^*+\kappa\phi)^{-1}$, and ∇^2 is the Laplacian in E_3. The first boundary value problem (Dirichlet) will certainly have a unique solution. This solution can be used to construct the unique semi-Riemannian manifold M_4. However, there is a peculiarity of the boundary value problems in general relativity. The continuous boundary values and the boundary itself are prescribed in the coordinate space E_3. Before actually solving the problem, the corresponding boundary in the (physical) Riemannian space is *completely unknown*. An example might clarify this matter. Let the boundary be the surface of the sphere

$$r=a/2>0,$$

in the coordinate space E_3 using spherical polar coordinates. At this point the corresponding surface in the (physical) Riemannian space is *unknown*. The Dirichlet problem for the complex potential U attaining the continuous boundary value

$$u(\theta,\phi)=1-2ima^{-1}(3+4i\cos\theta)^{-1/2},$$

can be written as the Fourier–Legendre series[3]

$$U(r,\theta,\phi) = 1 - ima^{-1}\left[\sum_{n=0}^{\infty}(2n+1)(2ra^{-1})^n\alpha_n P_n(\cos\theta)\right],$$

$$\alpha_n \equiv \int_0^{\pi}(3+4i\cos\theta')^{-1/2}P_n(\cos\theta')\sin\theta' d\theta', \\ = (2n+1)^{-1}(-i/2)^n. \tag{7.5}$$

The above series is absolutely convergent in the interior of the sphere in E_3 and uniformly convergent in any closed subset of the interior. Using the generating function $(1+t^2-2tx)^{-1/2} = \Sigma_{n=0}^{\infty} t^n P_n(x)$, the series converges to the complex potential

$$U(r,\theta,\phi) = 1 + m(r^2 - a^2 - 2iar\cos\theta)^{-1/2}. \tag{7.6}$$

The above generates the PIW subset of Kerr–Newman solutions. The boundary $r = a/2$ still remains a sphere in the associated space M_3, which is flat in this case. However, in the (physical) Riemannian space, the boundary surface is *known now* and given by the complicated line element

$$dl^2 = a^2\{(B+\tfrac{5}{4}-2s^2)(\tfrac{5}{4}-2s^2)/(1-4s^2)(\tfrac{5}{4}-s^2)\} \\ \times(d\theta\otimes d\theta) + a^2s^2[\{B^2+2B(\tfrac{5}{4}-s^2)+(\tfrac{5}{4}-2s^2) \\ \times(\tfrac{5}{4}-s^2)\}/\{B+\tfrac{5}{4}-2s^2\}](d\phi\otimes d\phi),$$

$$B \equiv (m/a)^2 + 2(m/a)\{\tfrac{1}{4}-s^2\}^{1/2}, \tag{7.7}$$

$$s \equiv \sin\theta.$$

It has been determined that this is *not* a sphere by the computer computations of $\hat{R}_{1212}/(\hat{g}_{11}\hat{g}_{12} - \hat{g}_{12}{}^2)$. Not only is there no uniqueness in the solution of boundary value problems in general, there is no way of selecting *a priori* a geometrically prescribed boundary (like a surface of constant curvature) in the physical space.

8. THE PIW CLASS OF SOLUTIONS

The stationary Einstein–Maxwell equations (2.3)–(2.5) admit a flat M_3 iff

$$\Gamma = \gamma\phi + (2\kappa)^{-1}|\gamma|^2 + id, \tag{8.1}$$

where γ, d are arbitrary complex and real constants, respectively. In this case field equations boil down to[9]

$$\nabla^2 U = 0, \tag{8.2a}$$

$$\mathrm{curl}\vec{a} = i|U|^2\mathrm{grad}[\ln(U^*/U)], \tag{8.2b}$$

where $U \equiv (2\kappa)^{1/2}(\gamma^* + \kappa\phi)^{-1}$ and the above equations are in E_3. These equations characterize the PIW class. An alternate criterion for this class would be to require that the eigenvalues of $R_{\alpha\beta}$ are all equal.

In the axially symmetric case one can use an oblate spheroidal coordinate system in E_3 as specified by

$$x^1 = \sqrt{(x'^2+a^2)(1-y^2)}, \quad x^2 = x'y, \quad x^3 = \phi, \tag{8.3}$$

where x^1, x^2, x^3 are cylindrical coordinates and $a^2 > 0$ is a real parameter. In these coordinates the most general solutions of (8.2a), (8.2b) (except due to an infinite rod) are[21]

$$U(x',y,a) \\ = \pi^{-1}\int_0^{\pi} f\left[x'y + i\sqrt{(x'^2+a^2)(1-y^2)}\cos\phi\right]d\phi,$$

$$a_\phi = 2\mathrm{Im}\left[\int[(1-y^2)U^*U_{,y}dx' \\ + (x'^2+a^2)UU^*_{,x'}dy] + \lambda_{,\phi}\right], \tag{8.4}$$

$$a_{x'} = \lambda_{,x'}, \quad a_{y'} = \lambda_{,y'}.$$

In the above, f is an arbitrary holomorphic function of its complex argument and λ is an arbitrary C^2-function of integration. The function λ can be absorbed in the metric.

Some solutions of the stationary Einstein–Maxwell equations can be generated[1] from the stationary vacuum metrics by the action of the eight-parameter group SU(2,1). Some PIW solutions can be obtained from a vacuum metric by a singular case of these transformations. Although a PIW solutions is thus obtained (this is a tricky process already) it cannot be claimed to be the unique PIW generalization of the vacuum metric. The known asymptotically flat stationary solutions[10] are of the class Kerr–Tomimatsu–Sato–Yamazaki (in short, KTSY). The charged version has also been constructed.[10] In the following, a PIW subset of the charged KTSY solutions will be generated. The KTSY class is given by the Ernst potential $\xi(x,y;p,q,\delta) = N/D$,

$$N \equiv \sum_{r=1}^{\delta}\left\{d(r)[px(x^2-1)^{r-1} - iqy(1-y^2)^{r-1}] \\ \times\left[\sum_{r'=r}^{\delta}C(\delta,r')F(\delta^2-r')\right]\right\},$$

$$D \equiv \sum_{r=1}^{\delta}C(\delta,r)F(\delta^2-r),$$

$$d(r) \equiv \{[(-1)^{r-1}(2r-2)!]/[2^{r-1}(r-1)!]^2\}, \tag{8.5}$$

$$C(\delta,r) \equiv \{[2^{2r-1}\delta(\delta+r-1)!]/[(\delta-r)!(2r)!]\},$$

$$F(\delta^2,r) \equiv \left\{\frac{[(-1)^r\delta!(\delta+1)!\cdots(2\delta-1)!]}{[2^r d(r+1)C(\delta,r)(2!3!\cdots(\delta-1)!)^3]}\right\}\det[M_{st}],$$

$$M_{st} \equiv [p^2(x^2-1)^{s+t-1} + q^2(1-y^2)^{s+t-1}]/(s+t-1),$$

where x, y represent prolate spheroidal coordinates, δ is a positive integer, and p, q are parameters.

A corresponding PIW solution can be constructed by choosing

$$U(x',y;a,\delta) = (ia\pi)^{-1}\int_0^{\pi}\xi\{(ia)^{-1} \\ \times\left[x'y + i\sqrt{(x'^2+a^2)(1-y^2)}\cos\phi\right],1;1,i,\delta\}^{-1}d\phi, \tag{8.6}$$

and obtaining a_ϕ from (8.4b).

It will be instructive to work out the integral (8.6) for the cases $\delta = 1,2,3$. The simplified versions of the integrands are the following:

$$[\xi(x,y;1,i,1)]^{-1} = (x+y)^{-1}, \tag{8.7a}$$

$$[\xi(x,y;1,i,2)]^{-1} \\ = [2x(x^2-1) + 2y(1-y^2)]/[(x^4-1) - (y^4-1) \\ + 2xy(x^2-y^2)] \\ = 2(x+y)^{-1} - 2(x+y)^{-3}(xy+1), \tag{8.7b}$$

$$[\xi(x,y;1,i,3)]^{-1} \\ = 3(x+y)^{-1} - 6(x+y)^{-3}(xy+1) \\ + 2(x+y)^{-5}(-x^2+4xy-y^2+3x^2y^2+3), \tag{8.7c}$$

All of the above functions are *harmonic* in prolate sphroidal coordinates. Moreover, the equation (8.7a) is the potential due to monopole, the equation (8.7b) is the potential due to a monopole plus a dipole, and so on. The corresponding complex potentials are furnished by the following:

$$U(x',y;a,1) = (x' + iay)^{-1},$$

$$U(x',y;a,2) = [2x'(x'^2 + a^2) - 2ia^3y(1 - y^2)]/[x'^4 - a^4y^4 + 2iax'y(x'^2 + a^2y^2)], \tag{8.8}$$

$$U(x',y;a,3) = [3(x' + iay)^4 - 6iax'y(x' + iay)^2 + 8a^2(x' + iay)^2 - 6a^2x'^2y^2 - 12ia^3x'y + 6a^4]/(x' + iay)^5$$

The function $U(x',y;a,1)$ generates a PIW subset of Kerr–Newman solutions. In general, the complex harmonic function $U(x',y,a,\delta)$ which gives rise to a stationary P.I.W. solution can be generated from the *real* harmonic function $[\xi(x,y;1,i,\delta)]^{-1}$ which is associated with a static Weyl class electrovac solution. The transformation involved is a complex one,[11] namely,

$$U(x',y;a,\delta) = [ia\xi(-ia^{-1}x',y;1,i,\delta)]^{-1}. \tag{8.9}$$

In the limiting case $\lim_{a\to 0} U(x',y;a,\delta) = \delta/x'$, which yields a spherically symmetric static electrovac solution. (The statements of the last three sentences have been explicitly verified for $\delta = 1,2,3$.)

It should be mentioned that the solutions $\xi(x,y;p,q,\delta)$ for the Ernst's equation are not yet fully checked (except for $\delta = 1,2,3,4$). Nevertheless, the complex harmonic functions $U(x',y;a,\delta)$ in (8.6) will certainly generate P.I.W. solutions even if for some δ the function $\xi(x,y;p,q,\delta)$ is *not* an Ernst's potential. The reason is that equation (8.6) is a special case of equation (8.4).

ACKNOWLEDGMENT

The authors are indebted to the N.R.C. Grant No. 3993.

[1]D. Kramer, G. Neugebauer, and H. Stephani, Fortschr. Phys. **20**, Heft 1, 1 (1972).

[2]C. Hoenselaers, J. Math. Phys. **17**, 1264 (1976).

[3]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1966), Vol. II, pp. 172, 580; 322.

[4]A. Das, J. Math. Phys. **14**, 1099 (1973).

[5]S. Kloster, M. M. Som, and A. Das, J. Math. Phys. **15**, 1096 (1974); C. Hoenselaers, Prog. Theor. Phys. **57**, 1223 (1977).

[6]S. D. Majumdar, Phys. Rev. **72**, 390 (1947); H. Bondi and T. Morgan, Proc. Roy. Soc. Lond. A **320**, 277 (1970); H. Levy, Nuovo Cimento, LV1 B 253 (1968); H. Muller Zum Hagen, Proc. Cambridge Philos. Soc. **75**, 249 (1974); J. P. Ward, Int. J. Theor. Phys. **15**, 293 (1976).

[7]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1966), Vol. I, p. 208.

[8]A. Lichnerowicz, *Théories relativistes de la gravitation et de l'électromagnétisme* (Masson, Paris, 1955).

[9]Z. Perjes, Phys. Rev. Lett. **27**, 1668 (1971); W. Israel and G. A. Wilson, J. Math. Phys. **13**, 865 (1972).

[10]R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963); A. Tomimatsu and H. Sato, Phys. Rev. Lett. **29**, 1344 (1972) and Prog. Theor. Phys. **50**, 95 (1973); F. J. Ernst, Phys. Rev. **168**, 1415 (1968) and Phys. Rev. D, **7**, 2520 (1973); M. Yamazaki, Prog. Theor. Phys. **57**, 1951 (1977) and J. Math. Phys. **19**, 1377 (1978).

[11]E. T. Newman and A. Janis, J. Math. Phys. **6**, 915 (1965); E. T. Newman *et al.*, J. Math. Phys. **6**, 918 (1965).

[12]A. Das and S. Kloster, J. Math. Phys. **18**, 586 (1977).

[13]C. Lanczos, *The Variational Principles of Mechanics* (University of Toronto Press, Toronto, 1977), p. 169.

[14]S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time* (Cambridge U. P., Cambridge, 1973), p. 230.

[15]F. Hohn, *Elementary Matrix Algebra* (Macmillan, New York, 1958), p. 250.

[16]A. Das, J. Math. Phys. **12**, 1136 (1971). (Compare Theorem 5 of this reference.)

[17]This is one possible set of jump conditions. For a single-layer source distribution these conditions would not hold. For a double-layer source distribution discontinuities of the field variables have to be allowed.

[18]A. Das, P. S. Florides, and J. L. Synge, Proc. Roy. Soc. A **263**, 451, (1961). Also see Refs. 16 and 5.

[19]For the four-dimensional version of this equation see J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1964), p. 233.

[20]These solutions are obtained by putting parameters $a = b = 0$ in the solutions given by W. B. Bonnor, Commun. Math. Phys. **34**, 77 (1973).

[21]H. Bateman, *Partial Differential Equations of Mathematical Physics* (Dover, New York, 1944).

Kerr-like neutrino-gravitational solutions

Anthony W. -C. Lun[a)]
Department of Mathematical Studies, Hong Kong Polytechnic, Hung Hom, Kowloon, Hong Kong

(Received 27 Novermber 1979; accepted for publication 28 March 1980)

Two distinct stationary axisymmetric Kerr-like neutrino-gravitational space–times are presented. They reduce to the Kerr solution when the neutrino field vanishes. In addition, the first solution possesses physical line singularities along the axis of symmetry; whereas the second solution is asymptotically flat and its locally bounded Riemann tensor is discontinuous at the equatorial plane. Both metrics are of Petrov type II and their neutrino fields belong to type $[2N\text{-}2S]_{(2-1)}$ in the Plebański classification.

1. INTRODUCTION

Recently a lot of interest has been generated in connection with investigations into the possibilities of gravitational effects on the structure of elementary particles by solving the Einstein–Dirac equations.[1-3] Neutrino-gravitation interactions are governed by the zero rest-mass Einstein–Dirac equations[4] which can be expressed in the two-component spinor formalism as follows:

$$R_{\mu\nu} = -i[\sigma_{\mu AB'}(\Phi^A\Phi^{B'}{}_{;\nu} - \Phi^A{}_{;\nu}\Phi^{B'}) + \sigma_{\nu AB'}(\Phi^A\Phi^{B'}{}_{;\mu} - \Phi^A{}_{;\mu}\Phi^{B'})],$$
$$\sigma^{\mu}{}_{AB'}\Phi^A{}_{;\mu} = 0. \tag{1}$$

Here $\{\sigma_{\mu AB'}\}, \mu = 1,...,4$, are the generalized Pauli spin matrices and Φ^A, $A = 1,2$, are components of a one-spinor which describes the neutrino field. Exact solutions of equations (1) are given by a number of authors (see Kuchowicz's review[5] and references cited therein). These known solutions imply that spherically symmetric space–times are not compatible with the presence of neutrino-gravitational fields which possess a shearfree geodetic null congruence.[6-8] Furthermore, Madore[9] proved that there exists no neutrino-gravitational field which is static and axisymmetric. In a recent paper, Herrera and Jiménez[2] using asymptotic series expansions showed that the assumptions of axisymmetry and asymptotic flatness would ultimately lead to physically singular solutions of the Einstein–Dirac neutrino field equations. One should note that all the results mentioned above on neutrino-gravitational space–times rest implicitly on the neutrino field being locally smooth[10] in an appropriate coordinate neighborhood where the corresponding metric is defined.

In this paper we present two distinct twisting neutrino-gravitational solutions both of which are reducible to the Kerr metric when the corresponding neutrino field vanishes. In terms of the Kerr coordinates (u,r,θ,ϕ), the first metric (G 1) is stationary axisymmetric, locally smooth, and possesses physical singularities along the axis of symmetry. These physical singularities are induced by the presence of the neutrino field. Thus the G 1 metric does not contradict the existing known theorems on neutrino-gravitational solutions. In contrast with G 1, the second metric (G 2) represents a stationary, axisymmetric asymptotically flat neutrino-gravitational field and it is of class C^{2-}.[11,12] In this case the Riemann tensor remains locally bounded but it is discontinuous at the equatorial plane. This is due directly to the existence of a continuous locally bounded current 4-vector[13] of the neutrino field which has a discontinuous θ-coordinate derivative. The apparent paradox between the G 2 solution and known results can be explained by the order of differentiability of the metric.

2. REDUCED NEUTRINO-GRAVITATIONAL FIELD EQUATIONS

The solutions in here are obtained by assuming that (i) the principal null congruence of the Weyl tensor coincides with the principal null congruence of the neutrino field[13] (we denote such a principal null vector by **L**); (ii) **L** is geodesic and shearfree with nonvanishing twist; (iii) the neutrino field Φ^A is time independent; and (iv) these solutions reduce to the Kerr solution when the neutrino field Φ^A vanishes. Assumptions (i) and (ii) imply that the Weyl tensor is algebraically special. In addition, if a space–time admits a Killing vector of the type $\mathbf{K} = \partial_{u'}$, then, using a theorem by Kerr and Debney,[14] there exists local coordinates $(u',r,\zeta,\bar{\zeta})$ such that the metric is given by

$$ds^2 = 2\Theta^{(1}\Theta^{2)} - 2\Theta^{(3}\Theta^{4)}. \tag{2}$$

Here the basis 1-forms are

$$\Theta^1 = dr + 2\,\mathrm{Im}(\rho^0{}_{\zeta}d\zeta) - U\Theta^2,$$
$$\Theta^2 = du' + bd\zeta + \bar{b}d\bar{\zeta},$$
$$\Theta^3 = -e^p(r - i\rho^0)d\bar{\zeta}, \quad \Theta^4 = \bar{\Theta}^3, \tag{3}$$

where

$$U = R^{(2)} + \frac{mr + \rho^0(M - \Phi^0\bar{\Phi}^0)}{r^2 + \rho^{0^2}}. \tag{4}$$

The tetrad variables ρ^0, p, m, and M are real-valued functions of ζ and $\bar{\zeta}$. The complex function $b(\zeta,\bar{\zeta})$ is defined by

$$b_{\bar{\zeta}} - \bar{b}_{\zeta} = -2ie^{2p}\rho^0. \tag{5}$$

$R^{(2)}$ is the Gaussian curvature of the 2-surface with metric $dl^2 = e^{2p}\,d\zeta\,d\bar{\zeta}$, i.e.,

$$R^{(2)} = e^{-2p}p_{\zeta\bar{\zeta}}\,. \tag{6}$$

[a)]Part of the research was carried out in the Department of Mathematics, Monash University, Clayton, Victoria 3168, Australia.

 0022-2488/80/082257-04$1.00

TABLE I. Functions Φ^0, $\Phi^0\overline{\Phi^0}$, N_1, N_2 and N_3 in $(u', r, \zeta, \bar{\zeta})$ coordinates.

	G1 Solution	G2 solution[a]
Φ^0	$\eta e^{in}[\zeta^{-1}(1+\frac{1}{4}\zeta\bar{\zeta})]^{1/2}$	$\eta e^{in}[\frac{1}{2}(1+\frac{1}{4}\zeta\bar{\zeta})]^{1/2}, \quad 0\leqslant \lvert\zeta\rvert\leqslant 2$ $\eta e^{in}\zeta^{-1}[2(1+\frac{1}{4}\zeta\bar{\zeta})]^{1/2}, \quad 2<\lvert\zeta\rvert<\infty$
$\Phi^0\overline{\Phi^0}$ [b]	$\dfrac{\eta^2(1+\frac{1}{4}\zeta\bar{\zeta})}{(\zeta\bar{\zeta})^{1/2}}$	$2^{-1}\eta^2(1+\frac{1}{4}\zeta\bar{\zeta}), \quad 0\leqslant\lvert\zeta\rvert\leqslant 2$ $2^{-1}\eta^2\left(1+4\dfrac{1}{\zeta\bar{\zeta}}\right), \quad 2<\lvert\zeta\rvert<\infty$
N_1	$\dfrac{2\eta^2(\zeta\bar{\zeta})^{1/2}}{(1+\frac{1}{4}\zeta\bar{\zeta})}$	$\eta^2(1+\frac{1}{4}\zeta\bar{\zeta})^{-1}\left\{\zeta\bar{\zeta}-2(1+\frac{1}{4}\zeta\bar{\zeta})\ln(1+\frac{1}{4}\zeta\bar{\zeta})+(1-\frac{1}{4}\zeta\bar{\zeta})\left[\operatorname{diln}(\frac{1}{4}\zeta\bar{\zeta})+\dfrac{\pi^2}{12}\right]\right\}, \quad 0\leqslant\lvert\zeta\rvert\leqslant 2$ $\eta^2(1+\frac{1}{4}\zeta\bar{\zeta})^{-1}\left\{4-2(1+\frac{1}{4}\zeta\bar{\zeta})\ln\left(1+4\dfrac{1}{\zeta\bar{\zeta}}\right)-(1-\frac{1}{4}\zeta\bar{\zeta})\left[\operatorname{diln}\left(4\dfrac{1}{\zeta\bar{\zeta}}\right)+\dfrac{\pi^2}{12}\right]\right\}, \quad 2<\lvert\zeta\rvert<\infty$
N_2	$\dfrac{\eta^2(\bar{\zeta})^{1/2}(1-\frac{1}{4}\zeta\bar{\zeta})}{(\zeta)^{1/2}(1+\frac{1}{4}\zeta\zeta)^2}$	$\eta^2\bar{\zeta}[2(1+\frac{1}{4}\zeta\bar{\zeta})^2]^{-1}\left\{(1-\frac{1}{4}\zeta\bar{\zeta})-\dfrac{2}{\zeta\bar{\zeta}}\left(1-\dfrac{\zeta^2\bar{\zeta}^2}{16}\right)\ln(1+\frac{1}{4}\zeta\bar{\zeta})-\left[\operatorname{diln}(\frac{1}{4}\zeta\bar{\zeta})+\dfrac{\pi^2}{12}\right]\right\}, \quad 0\leqslant\lvert\zeta\rvert\leqslant 2$ $\eta^2\bar{\zeta}[2(1+\frac{1}{4}\zeta\bar{\zeta})^2]^{-1}\left\{-\left(1-\dfrac{4}{\zeta\bar{\zeta}}\right)-\dfrac{2}{\zeta\bar{\zeta}}\left(1-\dfrac{\zeta^2\bar{\zeta}^2}{16}\right)\ln\left(1+4\dfrac{1}{\zeta\bar{\zeta}}\right)+\left[\operatorname{diln}\left(4\dfrac{1}{\zeta\bar{\zeta}}\right)+\dfrac{\pi^4}{12}\right]\right\}, \quad 2<\lvert\zeta\rvert<\infty$
N_3	$-2i\eta^2\zeta^{-1}\left[\tan^{-1}\dfrac{(\zeta\bar{\zeta})^{1/2}}{2}-\dfrac{(\zeta\bar{\zeta})^{1/2}(1-\frac{1}{4}\zeta\bar{\zeta})}{2(1+\frac{1}{4}\zeta\bar{\zeta})^2}\right]$	$i\eta^2\bar{\zeta}[2(1+\frac{1}{4}\zeta\bar{\zeta})^2]^{-1}\left\{(1+\frac{1}{4}\zeta\bar{\zeta})-\dfrac{2(1+\frac{1}{4}\zeta\bar{\zeta})^2}{\zeta\bar{\zeta}}\left[\dfrac{2\ln(1+\frac{1}{4}\zeta\bar{\zeta})}{(1+\frac{1}{4}\zeta\bar{\zeta})}-\ln 2\right]-\left[\operatorname{diln}(\frac{1}{4}\zeta\bar{\zeta})+\dfrac{\pi^2}{12}\right]\right\}, \quad 0\leqslant\lvert\zeta\rvert\leqslant 2$ $i\eta^2\bar{\zeta}[2(1+\frac{1}{4}\zeta\bar{\zeta})^2]^{-1}\left\{-\left(1-4\dfrac{1}{\zeta\bar{\zeta}}\right)+\dfrac{2(1+\frac{1}{4}\zeta\bar{\zeta})^2}{\zeta\bar{\zeta}}\left[\dfrac{2\ln[1+(4/\zeta\bar{\zeta})]}{[1+(4/\zeta\bar{\zeta})]}-\ln 2\right]+\left[\operatorname{diln}\left(4\dfrac{1}{\zeta\bar{\zeta}}\right)+\dfrac{\pi^2}{12}\right]\right\}, \quad 2<\lvert\zeta\rvert<\infty$

[a]The function diln(z) in the G2 solution is defined by $\operatorname{diln}(z) = -\int_0^z[\ln(1+t)/t]\,dt = \Sigma_{n=1}^{\infty}[(-1)^n/n^2]z^n$, $\lvert z\rvert\leqslant 1$. It is known as the Dilogarithm function (see Ref. 19).

[b]The neutrino flux $j^\mu = \rho\bar{\rho}\Phi^0\overline{\Phi^0}L^\mu$.

TABLE II. Functions Φ^0, $\Phi^0\overline{\Phi^0}$, N'_1, N'_2 and N'_3 in the Kerr coordinates.

	G1 Solution	G2 Solution[a]
Φ^0	$\dfrac{\eta e^{\frac{1}{2}i(\phi-2n)}}{(\sin\theta)^{1/2}}$	$\dfrac{\eta e^{in}}{(1+\cos\theta)^{1/2}}, \quad 0\leqslant\theta\leqslant\dfrac{\pi}{2}$ $\dfrac{\eta e^{-i(\phi-n)}}{(1-\cos\theta)^{1/2}}, \quad \dfrac{\pi}{2}<\theta\leqslant\pi$
$\Phi^0\overline{\Phi^0}$	$\dfrac{\eta^2}{\sin\theta}$	$\dfrac{\eta^2}{1+\lvert\cos\theta\rvert}$
N'_1	$2\eta^2\sin\theta$	$2\eta^2\left\{(1-\lvert\cos\theta\rvert)+[\ln(1+\lvert\cos\theta\rvert)-\ln 2]+\dfrac{\lvert\cos\theta\rvert}{2}\left[\operatorname{diln}\left(\dfrac{1-\lvert\cos\theta\rvert}{1+\lvert\cos\theta\rvert}\right)+\dfrac{\pi^2}{12}\right]\right\}$
N'_2	$2\eta^2\sin\theta\cos\theta$	$2\eta^2\cos\theta\left\{(1-\lvert\cos\theta\rvert)+[\ln(1+\lvert\cos\theta\rvert)-\ln 2]-\dfrac{\sin^2\theta}{2\lvert\cos\theta\rvert}\left[\operatorname{diln}\left(\dfrac{1-\lvert\cos\theta\rvert}{1+\lvert\cos\rvert}\right)+\dfrac{\pi^2}{12}\right]\right\}$
N'_3	$2\eta^2(\sin\theta\cos\theta-\theta)$	$2\eta^2\cos\theta\left\{(1-\lvert\cos\theta\rvert)+\dfrac{(1+\lvert\cos\theta\rvert)}{\lvert\cos\theta\rvert}\left[\ln(1+\lvert\cos\theta\rvert)-\dfrac{(\ln 2)\lvert\cos\theta\rvert}{1+\lvert\cos\theta\rvert}\right]-\dfrac{\sin^2\theta}{2\lvert\cos\theta\rvert}\left[\operatorname{diln}\left(\dfrac{1-\lvert\cos\theta\rvert}{1+\lvert\cos\theta\rvert}\right)+\dfrac{\pi^2}{12}\right]\right\}$

[a]The function diln(z) is the Dilogarithm function defined in footnote a in Table I (see Ref. 19).

The complex function $\Phi^0(\zeta,\bar\zeta)$ gives the neutrino field via the equations

$$\Phi = -\Phi^0 \rho, \quad \Phi^A = \Phi \xi^A, \quad L^\mu = \sigma^\mu{}_{AB'} \xi^A \xi^{B'}. \quad (7)$$

Here (ξ^A, η^A) denote the basis of two-component spinors.[15] The reduced neutrino-gravitational field equations corresponding to assumption (iii) and Eqs. (2) and (3) are

$$(m + iM)_{\bar\zeta} = i e^{p/2} \Phi^0 (e^{-p/2} \bar\Phi^0)_{\bar\zeta}, \quad (8)$$

$$(e^{-2p} p_{\zeta\bar\zeta})_{\zeta\bar\zeta} = 0, \quad (9)$$

$$e^{-2p} \rho^0{}_{\zeta\bar\zeta} - 2R^{(2)} \rho^0 = M, \quad (10)$$

$$(e^{p/2} \Phi^0)_{\bar\zeta} = 0. \quad (11)$$

These equations can also be derived from the paper by Trim and Wainwright[6,16] by making appropriate assumptions. General solutions of the reduced field equations (8)–(11) are not available. However, assumption (iv) implies that the 2-surfaces $dl^2 = 2e^{2p}\, d\zeta d\bar\zeta$ has constant negative Gaussian curvature. In general one can then choose $R^{(2)} = -\frac{1}{2}$.

3. KERR-LIKE NEUTRINO-GRAVITATIONAL SOLUTIONS

Integration of the above field equations with $R^{(2)} = -\frac{1}{2}$ leads to Kerr-NUT[17,18] type neutrino-gravitational space–times characterized by four essential parameters: m^0, M^0, a and η. Here m^0 corresponds to mass, M^0 is the NUT parameter, a is the Kerr parameter while η determines the neutrino field up to a constant phase change $\Phi^A \to e^{in}\Phi^A$, n = constant. Both the $G1$ and $G2$ metrics are given via Eqs. (2) and (3), where the unknown functions $m, M, p, \rho^0, \rho^0{}_\zeta, b$, and U are written as follows:

$$\begin{aligned} m &= m^0, \quad M = M^0 + \Phi^0\bar\Phi^0, \quad e^{-2p} = 2(1 + \tfrac{1}{4}\zeta\bar\zeta)^2, \\ \rho^0 &= -a\left(\frac{1 - \frac{1}{4}\zeta\bar\zeta}{1 + \frac{1}{4}\zeta\bar\zeta}\right) + M^0 + N_1, \\ \rho^0{}_\zeta &= \frac{a\bar\zeta}{2(1 + \frac{1}{4}\zeta\bar\zeta)^2} + N_2, \\ b &= \frac{ia\bar\zeta}{2(1 + \frac{1}{4}\zeta\bar\zeta)^2} + \frac{aiM^0}{\zeta(1 + \frac{1}{4}\zeta\bar\zeta)} + N_3, \\ U &= -\frac{1}{2} + \frac{m^0 r + M^0\rho^0}{r^2 + \rho^{0^2}}. \end{aligned} \quad (12)$$

The real-valued function N_1 and the complex functions Φ^0, N_2, N_3 are directly attributable to the presence of neutrino fields. Together with $\Phi^0\bar\Phi^0$, these functions are listed for the $G1$ and $G2$ metrics in Table I.

Introducing angular coordinates θ, ϕ by $\zeta = 2e^{i\phi} \times \tan(\theta/2)$ and a new u coordinate by $u = u' - 2M^0\phi$, one transforms $(u', r, \zeta, \bar\zeta)$ to the Kerr coordinates (u, r, θ, ϕ). The line element defined by Eqs. (2) and (3) then assumes the form

$$\begin{aligned} ds^2 &= (r^2 + \rho^{0^2})^{-1}\{(\Delta - a^2\sin^2\theta)du^2 + 2[aR\sin^2\theta - H\Delta \\ &\quad + N_2'(r^2 + \rho^{0^2})]du d\phi - (R^2\sin^2\theta + 2HN_2' \\ &\quad - H^2\Delta)d\phi^2\} + 2du dr - 2H dr d\phi - (r^2 + \rho^{0^2})d\theta^2, \end{aligned} \quad (13)$$

with

$$\begin{aligned} \rho^0 &= -a\cos\theta + M^0 + N_1', \\ H &= a\sin^2\theta + 2M^0\cos\theta + N_3', \\ \Delta &= r^2 - 2m^0 r + a^2 - M^{0^2} + N_1'(N_1' - 2a\cos\theta), \\ R &= r^2 + a^2 + M^{0^2} + N_1'(N_1' - 2a\cos\theta + 2M^0) + aN_3'. \end{aligned} \quad (14)$$

In Table II, the functions Φ^0, $\Phi^0\bar\Phi^0$, N_1', N_2', and N_3' for the $G1$ and $G2$ solutions are listed. Note that N_1', N_2' and N_3' vanish when the neutrino field is set equal to zero. The resulting vacuum Kerr solution is in the form given by Kinnersley.[18]

4. CONCLUDING REMARKS

From Table II it is obvious that both $G1$ and $G2$ admit a pair of commuting Killing vectors, $\mathbf{K}_1 = \partial_u$ and $\mathbf{K}_2 = \partial_\phi$. Using the Newman–Penrose formalism[15] one can show that these solutions possess the properties mentioned in the opening paragraphs (see Appendix). Their Weyl tensors are of Petrov type II and their respective neutrino fields belong to type $[2N\text{-}2S]_{(2\text{-}1)}$ in the Plebański classification.[20,21] These solutions are stationary axisymmetric.

From Eqs. (A3), the asymptotic behavior[22] of the nonzero tetrad components of the Ricci tensor is

$$\begin{aligned} \Phi_{11} &= \Phi^0_{11} r^{-4}, \\ \Phi_{21} &= \Phi^0_{21} r^{-3} + O(r^{-4}) = \bar\Phi_{12}, \\ \Phi_{22} &= -e^{-2p}[(e^p\Phi^0_{12})_\zeta + (e^p\Phi^0_{21})_{\bar\zeta}]r^{-3} + O(r^{-4}), \end{aligned}$$

where Φ^0_{11} and Φ^0_{21} are given by Eqs. (A2). Consequently, at large r

$$\Phi_{11}\Phi_{22} - \Phi_{12}\Phi_{21} < 0. \quad (15)$$

For a neutrino field subject to assumptions (i) and (ii), the energy momentum tensor assumes the form

$$\begin{aligned} T_{\mu\nu} &= 2\Phi_{22}L_\mu L_\nu + 2\Phi_{11}[4L_{(\mu}n_{\nu)} - g_{\mu\nu}] \\ &\quad + 4\Phi_{21}L_{(\mu}m_{\nu)} + 4\Phi_{12}L_{(\mu}\bar m_{\nu)}. \end{aligned}$$

Here the vector fields L^μ, n^μ, m^μ and $\bar m^\mu$ are pseudo-orthonormal vectors dual to the basis 1-form (3). Using a theorem by Wainwright,[23] Eq. (15) implies that both $G1$ and $G2$ violate the energy conditions E_1, E_2 and E_3 of Wainwright.[5,23] Moreover, one can show that the condition

$$Q_\mu(u)Q^\mu(u) \geqslant 0 \quad (16)$$

may not be valid with reference to these solutions, where

$$Q_\mu(u) = T_{\mu\nu}u^\nu$$

is the energy flow vector of the neutrino field with respect to an observer moving along a future-pointing unit velocity u^μ.[24]

The author would like to thank the referee's suggestions and for pointing out the work of Wainwright[24] on the nonexistence of physical solutions to the Einstein–Weyl equations in static space–times.

APPENDIX

In the Newman–Penrose formalism,[15] the nonvanishing spin coefficients with respect to the basis 1-form (3) and $R^{(2)}$ = constant are

$$\rho = -(r+i\rho^0)^{-1}, \quad \beta = \tfrac{1}{2} e^{-P} p_{\zeta}\bar{\rho},$$
$$\bar{\alpha} + \beta = 0, \quad \gamma = -\tfrac{1}{2}\Psi^0_2 \rho^2,$$
$$\mu = -R^{(2)}\bar{\rho} - \tfrac{1}{2}\Psi^0_2(\rho^2 + \rho\bar{\rho}) + \Phi^0_{11}\rho^2\bar{\rho},$$
$$\nu = -\tfrac{1}{2}e^{-P}(\Psi^0_2)_{\zeta}\rho^2 - ie^{-P}\rho^0{}_{\zeta}\Psi^0_2\rho^3 - \Phi^0_{21}\rho\bar{\rho} + e^{-P}(\Phi^0_{11})_{\zeta}\rho^2\bar{\rho} + 2ie^{-P}\rho^0{}_{\zeta}\Phi^0_{11}\rho^3\bar{\rho}. \quad \text{(A1)}$$

Here

$$\Psi^0_2 = -(m+iM), \quad \Phi^0_{11} = \rho^0\Phi^0\bar{\Phi}^0,$$
$$\Phi^0_{21} = \frac{i}{2}e^{-P}(\Phi^0\bar{\Phi}^0)_{\zeta}. \quad \text{(A2)}$$

The nonvanishing tetrad components of the Riemann tensor are

$$\Psi_2 = -\Psi^0_2\rho^3 + 2\Phi^0_{11}\rho^3\bar{\rho},$$
$$\Psi_3 = -e^{-P}(\Psi^0_2)_{\zeta}\rho^3 - 3ie^{-P}\rho^0{}_{\zeta}\Psi^0_2\rho^4 - \Phi^0_{21}\rho^2\bar{\rho} + 2e^{-P}(\Phi^0_{11})_{\zeta}\rho^3\bar{\rho} + 6ie^{-P}\rho^0{}_{\zeta}\Phi^0_{11}\rho^4\bar{\rho},$$
$$\begin{aligned}\Psi_4 = &-\tfrac{1}{2}[e^{-2P}(\Psi^0_2)_{\zeta}]_{\zeta}\rho^3 - i[\Psi^0_2(e^{-2P}\rho^0{}_{\zeta})_{\zeta} + 3e^{-2P}\rho^0{}_{\zeta}(\Psi^0_2)_{\zeta}]\rho^4 + 6e^{-2P}(\rho^0{}_{\zeta})^2\Psi^0_2\rho^5 \\ &- (e^{-P}\Phi^0_{21})_{\zeta}\rho^2\bar{\rho} + [(e^{-2P}(\Phi^0_{11})_{\zeta})_{\zeta} + 2ie^{-P}\rho^0{}_{\zeta}\Phi^0_{21}]\rho^3\bar{\rho} + 2i[\Phi^0_{11}(e^{-2P}\rho^0{}_{\zeta})_{\zeta} \\ &+ 3e^{-2P}\rho^0{}_{\zeta}(\Phi^0_{11})_{\zeta}]\rho^4\bar{\rho} - 12e^{-2P}(\rho^0{}_{\zeta})^2\Phi^0_{11}\rho^5\bar{\rho},\end{aligned}$$
$$\Phi_{11} = \Phi^0_{11}\rho^2\bar{\rho}^2,$$
$$\Phi_{21} = -\Phi^0_{21}\rho\bar{\rho}^2 + e^{-P}(\Phi^0_{11})_{\zeta}\rho^2\bar{\rho}^2 + 2ie^{-P}\rho^0{}_{\zeta}\Phi^0_{11}\rho^3\bar{\rho}^2,$$
$$\begin{aligned}\Phi_{22} = &[e^{-2P}(e^P\Phi^0_{12})_{\zeta}\rho^2\bar{\rho} - 2ie^{-P}\rho^0{}_{\zeta}\Phi^0_{12}\rho^3\bar{\rho} + \tfrac{1}{2}e^{-2P}(\Phi^0_{11})_{\zeta\bar{\zeta}}\rho^2\bar{\rho}^2 + 2ie^{-2P}\rho^0{}_{\zeta}(\Phi^0_{11})_{\bar{\zeta}}\rho^3\bar{\rho}^2 \\ &+ 2e^{-2P}\rho^0{}_{\zeta}\rho^0{}_{\bar{\zeta}}\Phi^0_{11}\rho^3\bar{\rho}^3] + \text{complex conjugate}. \quad \text{(A3)}\end{aligned}$$

[1]M. J. Hamilton and A. Das, J. Math. Phys. **18**, 2026 (1977).
[2]L. Herrera and J. Jiménez, J. Math. Phys. **20**, 195 (1979).
[3]K. R. Pechenick and J. M. Cohen, Phys. Rev. D **19**, 1635 (1979).
[4]D. R. Brill and J. A. Wheeler, Rev. Mod. Phys. **29**, 465 (1957).
[5]B. Kuchowicz, Gen. Rel. Grav. **5**, 201 (1974).
[6]D. W. Trim and J. Wainwright, J. Math. Phys. **12**, 2494 (1971).
[7]J. Audretsch, Lett. Nuovo Cimento **4**, 339 (1972).
[8]J. B. Griffiths, Gen. Rel. Grav. **5**, 361 (1973).
[9]J. Madore, Lett. Nuovo Cimento **5**, 48 (1972).
[10]A smooth (C^∞) function has continuous partial derivatives of any order.
[11]A C^n metric is one whose metric coordinate components have continuous $n-1$ coordinate derivatives which satisfy a local Lipschitz condition (see Ref. 12).
[12]S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space–Time* (Cambridge U. P., London, 1973), Secs. 3.1 and 8.4.
[13]The current 4-vector of the neutrino field $j^\mu = \sigma^\mu{}_{AB'}\Phi^A\bar{\Phi}^{B'}$ whose covariant divergence vanishes by virtue of the Dirac equation represents the neutrino flux (see Refs. 3 and 4). It is also known as the principal null vector of the neutrino field (see Ref. 6).
[14]R. P. Kerr and G. C. Debney, Jr., J. Math. Phys. **11**, 2807 (1970).
[15]E. T. Newman and R. Penrose, J. Math. Phys. **11**, 565 (1962).
[16]D. W. Trim and J. Wainwright, J. Math. Phys. **15**, 535 (1974).
[17]M. Demianski and E. T. Newman, Bull. Acad. Pol. Sci. Ser. Math. Phys. **14**, 653 (1966).
[18]W. Kinnersley, J. Math. Phys. **10**, 1195 (1969).
[19]M. Abramowitz and I. A. Segun, *Handbook of Mathematical Functions* (Dover, New York, 1968), p. 1004.
[20]J. Plebański, Acta Phys. Pol. **26**, 963 (1964).
[21]J. B. Griffiths and R. A. Newing, J. Phys. A **4**, 208 (1971).
[22]In Ref. 16, it is proved that there are no "Kerr-like" neutrino-gravitational solutions if the space–time is algebraically special and the Ricci tensor has asymptotic behaviour $R_{\mu\nu} = R_{\mu\nu}r^{-4} + O(r^{-5})$. On the other hand, the Ricci tensor of the solutions here has a leading nonzero term in the inverse third power of r.
[23]J. Wainwright, J. Math. Phys. **12**, 828 (1971).
[24]J. Wainwright, Nuovo Cimento B **22**, 131 (1974).

Some electrovac models of homogeneous gravitational force fields in general relativity

K. N. Srinivasa Rao and A. V. Gopala Rao
Department of Physics, University of Mysore, Manasagangotri, Mysore 570 006, India

(Received 6 June 1979; accepted for publication 19 September 1979)

In this paper we consider the problem of defining and constructing the general relativistic analog of the Newtonian homogeneous gravitational force field. With this purpose, the Newtonian definition of homogeneity is first reformulated in a manner suitable for nonflat space sections and is then shown to lead to a covariant characterization of static space–times admitting uniform gravitational force fields. This covariant characterization is both necessary and sufficient and is consistent with the necessary conditions of homogeneity introduced by Goodinson. Further, a whole class of static solutions of the Einstein–Maxwell equations, all admitting homogeneous gravitational force fields, is obtained and some of their properties are discussed. In particular, this class of electrovac solutions appears to exhibit the remarkable property that the metrics and the source electromagnetic fields are free of singularities.

1. INTRODUCTION

This paper is devoted to an analysis of the concept of the homogeneous gravitational force field in the general theory of relativity (GTR). The homogeneous gravitational force field is perhaps the simplest of all Newtonian gravitational fields and is invariably studied in all elementary introductions to the subject. Even though the concept of a homogeneous gravitational force field is very simple in the Newtonian theory, one cannot say the same thing in GTR. The problem of carrying over the concept of the homogeneous gravitational force field into GTR is not trivial and it would be interesting to study how far GTR can accommodate this Newtonian concept.

With this aim, we introduce here a certain definition of a homogeneous gravitational force field in a static space–time, leaning heavily on the Newtonian definition. Confining our attention only to static space–times, we then show that this definition of homogeneity can be cast into a covariant form yielding thus a covariant criterion which is both necessary and sufficient for the existence of a homogeneous gravitational force field in static space–times. In Sec. 3, using this criterion of homogeneity, we give a simple prescription to construct exact static solutions of the Einstein equations admitting uniform gravitational force fields. In the rest of this paper, we construct and analyze such a class of electrovac solutions of the Einstein equations.

2. THE HOMOGENEITY CRITERION

In the Newtonian theory, a uniform gravitational force field is defined as a vector field which is homogeneous both in space and time. Since the gravitational force is regarded as an absolute "physical" force in Newtonian theory and is distinguished from the forces of inertia which appear in noninertial reference frames, this definition of uniformity or homogeneity retains its meaning in all reference frames. However, in GTR, it is impossible to give an observer-independent meaning to the homogeneity of a gravitational force field because, here, the "gravitational forces" are essentially the accelerations of the observer himself. Thus, in GTR, the concept of homogeneity can be defined only with respect to an observer and this problem has been discussed recently by Goodinson.[1] In the same paper,[1] Goodinson has also given a set of necessary local conditions in order that the gravitational 3 force measured by an observer may appear homogeneous to him.

Here, we introduce a definition of homogeneity which is greatly akin to the Newtonian definition and also satisfies all the three necessary conditions of Goodinson.[1] However, we confine our attention only to static space–times as the concept of a homogeneous gravitational force field is essentially Newtonian and static Einstein fields bear the closest relation to Newtonian gravitational fields.

In GTR, a test particle, which is otherwise free, moves along a geodesic. Thus, according to an observer who is also moving geodesically, the gravitational force on the test particle vanishes and this is commonly referred to as the elimination of gravitational force in GTR. However, it is evident that the gravitational force on test particles is not zero according to an observer who is not following a geodesic; in fact, an observer moving with a 4 velocity $U^i(U_i U^i = -1)$ ascribes the gravitational force $a_i \equiv \dot{U}_i$ to a test particle of unit mass.[1-3] Therefore, a 4 velocity field (corresponding to a set of observers) serves to define a gravitational force field throughout space–time. In particular, the gravitational force measured by observers at rest relative to the frame of reference[4] of a space–time coordinate system x^i is given by $a_i \equiv U_{i;k} U^k$, where U_i is the field of tangents to the time lines ($x^\alpha = \text{const.}$) of the coordinate system. We shall refer to this force field as the gravitational force field existing in the reference frame provided by the coordinate system x^i.

Let us now examine gravitational force fields in static reference frames (of static space–times). By definition, a frame of reference is said to be static if it is attached to a coordinate system x^i in which the line element assumes the form

$$ds^2 \equiv -d\tau^2 = g_{\alpha\beta}dx^\alpha dx^\beta - e^{2\psi}dt^2, \qquad (2.1)$$

where $g_{\alpha\beta}$ and ψ are independent of $x^0 = t$.

In such a coordinate system, a direct calculation yields the following quantities [see also Synge[5] (pp. 388 to 340)]: (i) the tangent vector field of the time lines $x^\alpha = \text{const.}$:

$$U^a \equiv \frac{dx^a}{d\tau} = (e^{-\psi},0,0,0),$$
$$U_a = (-e^{\psi},0,0,0), \qquad U^a U_a = -1, \tag{2.2}$$

where we have chosen τ such that t is an increasing function of τ along the time lines $x^\alpha = \text{const}$; (ii) the nonvanishing Christoffel symbols

$$\Gamma^{\alpha}_{\mu\nu} = \gamma^{\alpha}_{\mu\nu}, \quad \Gamma^{0}_{0\alpha} = \psi_{,\alpha}, \quad \Gamma^{\alpha}_{00} = g^{\alpha\beta} e^{2\psi}\psi_{,\beta}, \tag{2.3}$$

where $\gamma^{\alpha}_{\mu\nu}$ denote the Christoffel symbols constructed from the spatial metric $g_{\mu\nu}$, and the comma denotes ordinary partial differentiation; (iii) the tensor $U_{a;b}$: the only nonzero component of the tensor $U_{a;b}$ obtained by covariant differentiation of U_a is given by

$$U_{\alpha;0} = -\Gamma^{0}_{\alpha 0} U_0 = e^{\psi}\psi_{,\alpha}; \tag{2.4}$$

(iv) the acceleration field a^i:

$$a_i \equiv \dot{U}_i = U_{i;k} U^k = (0,\psi_{,1},\psi_{,2},\psi_{,3}); \; a_i U^i = 0, \tag{2.5}$$

$$a^i = (0, g^{1\alpha}\psi_{,\alpha}, g^{2\alpha}\psi_{,\alpha}, g^{3\alpha}\psi_{,\alpha}), \tag{2.6}$$

$$a^2 \equiv a_k a^k = g^{\alpha\beta}\psi_{,\alpha}\psi_{,\beta}; \tag{2.7}$$

(v) the tensor $a_{i;k}$:

$$a_{0;0} = -\Gamma^{\alpha}_{00} a_\alpha = -e^{2\psi} g^{\alpha\beta}\psi_{,\alpha}\psi_{,\beta} = -a^2 e^{2\psi}, \tag{2.8}$$

$$a_{0;\alpha} = 0, \quad a_{\alpha;0} = 0, \tag{2.9}$$

$$a_{\alpha;\beta} = \psi_{,\alpha\beta} - \Gamma^{\sigma}_{\alpha\beta}\psi_{,\sigma}; \tag{2.10}$$

(vi) the vector $\dot{a}_i \equiv a_{i;j} U^j$:

$$\dot{a}_i = (a_{0;0} U^0,0,0,0) = (-a^2 e^{\psi},0,0,0) = -a^2 U_i. \tag{2.11}$$

From Eqs. (2.2), (2.5), and (2.11) we obtain the relations

$$U_{i;j} = -a_i U_j, \quad \ddot{U}_{[i} U_{j]} \equiv \dot{a}_{[i} U_{j]} = 0, \tag{2.12}$$

showing that U_a is a static congruence.[6] These equations (2.12) in turn imply

$$\theta \equiv g^{ij} U_{i;j} = -g^{ij} a_i U_j = 0, \tag{2.13}$$

$$\sigma_{ij} \equiv U_{(i;j)} + a_{(i} U_{j)} + \tfrac{1}{3}\theta (g_{ij} + U_i U_j) = 0, \tag{2.14}$$

$$\omega_{ij} \equiv -U_{[i;j]} - a_{[i} U_{j]} = 0, \tag{2.15}$$

which show that a static congruence is necessarily rigid and normal.

Coming back to the acceleration vector a_i, we see that (as $a_i U^i = 0$) it lies entirely in the infinitesimal 3-rest space of an observer whose 4 velocity is U_i. Moreover, in the case of a static reference frame, like the one we are considering, the infinitesimal 3-rest spaces of observers traveling along the time lines mesh into a finite spacelike hypersurface called a space section.[6] Thus, at any moment of time $t = t_0$ in a static reference frame, a_i lies entirely in the space section $t = t_0$ which has the metric $g_{\alpha\beta}$ and the spatial components a_α of a_i form the covariant components of a 3 vector **a** defined throughout this space section. The contravariant components of **a** are then given $a^\alpha = g^{\alpha\beta} a_\beta$ and **a** gives the intensity of the gravitational 3 force existing in the static reference frame. Here we must mention another interesting derivation of **a** given in Landau and Lifshitz.[7]

Observe that **a** is already independent of time by virtue of the static nature of the reference frame. Following the Newtonian definition of homogeneity, we now define **a** to be spatially homogeneous if it is constant throughout a space section. Since the space sections are in general non-Euclidean, the constancy of **a** with respect to to the space metric $g_{\alpha\beta}$ has to be defined only as[8,9]

$$a_{\alpha:\beta} = a_{\alpha,\beta} - \gamma^{\sigma}_{\alpha\beta} a_\sigma = 0, \tag{2.16}$$

where $\gamma^{\sigma}_{\alpha\beta}$ are the Christoffel symbols of the space metric $g_{\alpha\beta}$. An immediate consequence of Eq. (2.16) is that

$$(a^\alpha a_\alpha)_{,\beta} = 0, \tag{2.17}$$

which shows that the magnitude of **a** is a constant throughout the space section. This lends more support to Eq. (2.16) as a definition of homogeneity of **a**.

Note that for the metric (2.1) $\gamma^{\sigma}_{\alpha\beta} \equiv \Gamma^{\sigma}_{\alpha\beta}$ and hence Eq. (2.16) requires only the spatial components of the covariant derivative $a_{i;j}$ to be zero. Thus, Eq. (2.16) is not a covariant condition. However, it can be cast in a covariant form easily. To see this consider the projection tensor

$$h_{ab} = g_{ab} + U_a U_b \tag{2.18}$$

associated with U_a. This tensor "projects" into the 3-rest space of an observer with 4 velocity U_a. In particular, for the metric (2.1) and the vector field (2.2) we have

$$h_{00} = h_{0\alpha} = 0, \quad h_{\alpha\beta} = g_{\alpha\beta},$$
$$h^{0}_{\;0} = h^{0}_{\;\alpha} = h^{\alpha}_{\;0} = 0, \quad h^{\alpha}_{\;\beta} = \delta^{\alpha}_{\beta}. \tag{2.19}$$

Because of this, the vanishing of the spatial components of $a_{i;j}$ makes the tensor $h^{i}_{\;a} h^{j}_{\;b} a_{i;j}$ vanish completely, i.e., *for a static congruence* U_a, the noncovariant condition (2.16) is equivalent to the covariant requirement

$$h^{i}_{\;a} h^{j}_{\;b} a_{i;j} = 0. \tag{2.20}$$

Note that Eq. (2.20), when true of a static congruence U_a, is also sufficient to guarantee Eq. (2.16), for, in a special coordinate system with the congruence U_i as the time lines, Eqs. (2.2) and (2.19) are true and hence Eq. (2.20) reduces to

$$h^{\alpha}_{\;\mu} h^{\beta}_{\;\nu} a_{\alpha;\beta} = \delta^{\alpha}_{\mu}\delta^{\beta}_{\nu} a_{\alpha;\beta} = a_{\mu;\nu} = 0,$$

which is Eq. (2.16). *Thus, Eq.* (2.20) *is the necessary and sufficient, covariant, condition for the homogeneity of the gravitational* 3-*force field existing in a static reference frame.* This leads to the following theorem: "A space–time admits a homogeneous gravitational force field if it possesses a static congruence U_a satisfying Eq. (2.20)."

We will now show that Eq. (2.20) is consistent with the necessary covariant conditions of homogeneity introduced by Goodinson.[1] First, we observe that the equation (2.20), or its equivalent equation (2.16), implies the constancy of the scalar $a \equiv (a_k a^k)^{1/2}$, i.e.,

$$a_{,k} = 0. \tag{2.21}$$

We only have to differentiate a^2 to see this; we then get

$$a^2_{,k} = (a_i a^i)_{,k} = 2 a_{i;k} a^k,$$

which vanishes by virtue of Eqs. (2.6), (2.9), and (2.16). The condition (2.21) is one of the Goodinson conditions of homogeneity. To get the others, we have to consider the spacelike congruence defined by the unit spacelike vector field

$$K_i = \frac{1}{a} a_i, \quad K_i K^i = +1 \tag{2.22}$$

and its kinematical quantities[1]

expansion scalar $\theta^* \equiv P^{ij} K_{i;j}$, (2.23)

shear tensor $\sigma^*_{ij} \equiv P^a_i P^b_j K_{(a;b)} - \frac{1}{2}\theta^* P_{ij}$, (2.24)

rotation tensor $\omega^*_{ij} \equiv P^a_i P^b_j K_{[a;b]}$, (2.25)

where

$$P_{ij} \equiv g_{ij} + U_i U_j - K_i K_j = h_{ij} - K_i K_j = P_{ji}. \tag{2.26}$$

This projection tensor P_{ij} associated with K_i has the following properties:

$$P^a_a = 2, \quad P^{ab} K_b = 0,$$
$$P^{ab} U_b = 0, \quad P_{ac} P^{bc} = P^b_a. \tag{2.27}$$

Particularly for the K_i congruence of the static congruence (2.2), the components of P^i_j are given by [see Eqs. (2.5), (2.6), (2.19), (2.21), and (2.22)]

$$P^0_0 = 0, \quad P^0_\alpha = 0, \quad P^\alpha_0 = 0,$$
$$P^\alpha_\beta = h^\alpha_\beta - K^\alpha K_\beta = \delta^\alpha_\beta - \frac{1}{a^2} a^\alpha a_\beta. \tag{2.28}$$

Now, using Eq. (2.28) and observing that $K_{0;0} = (1/a) a_{0;0}$ is the only nonzero component of $K_{i;j}$ for a homogeneous gravitational force field [this follows from Eqs. (2.9), (2.16), (2.21), and (2.22)], it is easy to see that all the three kinematical quantites defined in Eqs. (2.23)–(2.25) vanish for the K_i congruence of a homogeneous gravitational force field. We thus obtain the necessary conditions

$$\theta^* = 0, \tag{2.29}$$

$$\sigma^*_{ij} = 0, \tag{2.30}$$

$$\omega^*_{ij} = 0, \tag{2.31}$$

which, in addition to Eq. (2.21), characterize a homogeneous gravitational force field. To these conditions we may add three more conditions, namely Eq. (2.13)–(2.15), which describe the static nature of the U_i congruence. Thus, we have, in all, seven necessary conditions characterizing a homogeneous gravitational force field. Out of these, the four conditions (2.13), (2.14), (2.21), and (2.29) are the Goodinson necessary conditions of homogeneity.

This discussion shows that the condition (2.20) is a stronger, but consistent, condition of homogeneity for static gravitational 3-force fields. Moreover, it has the advantage of being a simple covariant condition which is both necessary and sufficient.

3. PRESCRIPTION FOR CONSTRUCTING HOMOGENEOUS GRAVITATIONAL FORCE FIELDS

It is now obvious that the problem of construction of a model of a homogeneous gravitational force field is simply the problem of construction of a suitable static congruence in space–time. However, we must note that it is not sufficient to have a mere mathematical solution, namely, a real static coordinate system x^i whose time line congruence satisfies Eq. (2.20); we must also see that the corresponding space–time geometry satisfies the Einstein field equations for some physically reasonable situation. Therefore, one way to find models of homogeneous gravitational force fields is to check whether the known static Einstein fields admit static congruences satisfying Eq. (2.20), but there is little chance of finding such solutions. The other way is to construct anew solutions admitting homogeneous gravitational force fields. In doing this, as we do not know what sort of matter–energy distributions lead to homogeneous gravitational force fields, we have to adopt what Synge [Ref. 5 (p. 189)] calls the *g method* in which a set of ten sufficiently smooth functions $g_{ij}(x)$, having everywhere the signature $(- + + +)$, is chosen and the energy tensor is *calculated* from the Einstein equations. The energy tensor so obtained is then examined for physical acceptability, by considering the algebraic sign of its energy density (which must be positive) and the other eigenvalues. This *g* method, in conjuction with the condition (2.16), leads to a simple prescription for the construction of exact static solutions which admit homogeneous gravitional force fields. It is as follows: "With a suitable static spatial 3 metric $g_{\alpha\beta}$, find a solution ψ, of the differential equation

$$\psi_{,\alpha\beta} - \gamma^\sigma_{\alpha\beta} \psi_{,\sigma} = 0, \tag{3.1}$$

where $\gamma^\sigma_{\alpha\beta}$ are the Christoffel symbols of $g_{\alpha\beta}$. Then construct the energy tensor

$$T_{ij} = -(8\pi)^{-1} \{R_{ij} - \tfrac{1}{2} R g_{ij} - \Lambda g_{ij}\} \tag{3.2}$$

for the static line element

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta - e^{2\psi} dt^2. \tag{3.3}$$

If this energy tensor T_{ij} represents a "reasonable" physical situation, then Eq. (3.3) is an example of a space–time admitting a homogeneous gravitational force field."

4. ELECTROVAC MODELS OF HOMOGENEOUS GRAVITATIONAL FORCE FIELDS

We now proceed to construct a class of static electrovac solutions which admit homogeneous gravitational force fields. The key point is to apply the prescription of the previous section to a spatial line element of the form

$$d\sigma^2 = e^{2L} dx^2 + e^{2M} dy^2 + e^{2N} dz^2, \tag{4.1}$$

where $L = L(x)$, $M = M(y,z)$, and $N = N(y,z)$. As L is a function of x alone, it is possible to absorb e^{2L} into dx^2 by a simple scale change in x, but we shall retain this term for the time being as it is a simple matter to set $L = 0$ whenever we want it. The nonvanishing components of the Christoffel symbols for the metric (4.1) are

$$\gamma^1_{11} = L_1, \quad \gamma^2_{22} = M_2, \quad \gamma^2_{23} = M_3, \quad \gamma^2_{33} = -N_2 e^{2(N-M)},$$
$$\gamma^3_{22} = -M_3 e^{2(M-N)}, \quad \gamma^3_{23} = N_2, \quad \gamma^3_{33} = N_3, \tag{4.2}$$

where we have used the notation $f_{,\alpha} \equiv f_\alpha$ so that $L_1 \equiv L_{,1}$, $M_3 \equiv M_{,3}$, etc. With these Christoffel symbols, we now set out to solve the partial differential equation (3.1) in the unknown ψ. The problem is pretty hard if ψ is dependent on all the three coordinates (x,y,z) and so we look for special solutions. Fortunately, a solution exists when ψ is a function of x alone. Then all the derivatives of ψ, except $\psi_{,1} \equiv \psi_1$ and $\psi_{,11} \equiv \psi_{11}$, vanish and Eq. (3.1) becomes

$$\psi_{11} - L_1\psi_1 = 0.$$

This is an ordinary differential equation in ψ and has the general solution

$$\psi(x) = m + k\int e^{L(x)}dx, \tag{4.3}$$

where m and k are two (real) constants of integration. With this ψ, we now form the static space–time metric

$$ds^2 = e^{2L}dx^2 + e^{2M}dy^2 + e^{2N}dz^2 - e^{2\psi}dt^2, \tag{4.4}$$

which admits a homogeneous gravitational force field of intensity [see Eq. (2.5)]

$$a_\alpha = \psi_{,\alpha} = (ke^L,0,0). \tag{4.5}$$

It only remains to see what sort of physical situation is represented by Eq. (4.4). For this we have to construct the energy tensor of Eq. (4.4). First, regarding the Christoffel symbols of Eq. (4.4), we observe that the spatial components of Γ^i_{jk} are already given in Eq. (4.2) and of the other components, only

$$\Gamma^0_{10} = \psi_1, \quad \Gamma^1_{00} = \psi_1 e^{2(\psi - L)} \tag{4.6}$$

are nonzero. From this, we obtain the nonvanishing components of the Riemann tensor

$$R_{abcd} = \tfrac{1}{2}(g_{ad,bc} + g_{bc,ad} - g_{ac,bd} - g_{bd,ac}) + g_{mn}(\Gamma^m_{ad}\Gamma^n_{bc} - \Gamma^m_{ac}\Gamma^n_{bd}) \tag{4.7}$$

to be

$$R_{0101} = k^2\exp\left(2L + 2k\int e^{L(x)}dx\right) = k^2 e^{2(L+\psi)}, \tag{4.8}$$

$$R_{2323} = -e^{2M}(M_{33} + M_3^2 - M_3N_3) - -e^{2N}(N_{22} + N_2^2 - M_2N_2), \tag{4.9}$$

where we have set the constant m appearing in Eq. (4.3) equal to zero, as this can always be absorbed into t. Contracting R_{abcd}, we find that the Ricci tensor

$$R_{ab} = g^{mn}R_{mabn} \tag{4.10}$$

has only the following nonzero components:

$$R_{00} = -k^2e^{2\psi}, \quad R_{11} = k^2e^{2L},$$
$$R_{22} = e^{2M-2N}(M_{33} + M_3^2 - M_3N_3) + N_{22} + N_2^2 - M_2N_2, \tag{4.11}$$
$$R_{33} = e^{2N-2M}(N_{22} + N_2^2 - M_2N_2) + M_{33} + M_3^2 - M_3N_3.$$

Using R_{ab} and g_{ab}, we find the components of the tensor $M^a_b = R^a_b - \frac{1}{2}R\delta^a_b - \Lambda\delta^a_b$ to be

$$M^0_0 = M^1_1 = e^{-2M-2N}R_{2323} - \Lambda,$$
$$M^2_2 = M^3_3 = -k^2 - \Lambda, \tag{4.12}$$

where we have shown only the nonvanishing components. This tensor is important because it is related to the energy tensor T^a_b by

$$T^a_b = -(8\pi)^{-1}M^a_b \tag{4.13}$$

and we shall now examine whether it is possible to choose the arbitrary real functions L,M and N (ψ is already determined, as a function of L, by the homogeneity criterion) in the metric (4.4) such that the energy tensor (4.13) has a reasonable physical model. First of all, we shall check whether a mechanical fluid model exists for Eq. (4.4). In a general energy tensor, to get pressure rather than tensions and an energy density which is positive {this is even more important; see Synge [Ref. 5 (p. 316)]}, T^a_b should have

$$T^0_0 < 0, \quad T^1_1 > 0, \quad T^2_2 > 0, \quad T^3_3 > 0,$$

which follows from the fact that T^a_b is diagonal in our coordinate system with eignevalues appearing along the diagonal [see Eqs. (4.12) and (4.13)]. Equivalently, M^a_b should have

$$M^0_0 > 0, \quad M^1_1 < 0, \quad M^2_2 < 0, \quad M^3_3 < 0.$$

However, this is impossible in view of Eq. (4.12) and hence *tensions are also present* in the energy tensor we are considering. Thus, we have to rule out fluids as possible models for Eq. (4.4). Moreover, the presence of tensions and the pairwise equality of the eigenvalues of T^i_j actually suggests that we look for a free electromagnetic field as a model of Eq. (4.4), as the electromagnetic field energy tensor also has the same properties. In that case, the M^i_j of Eq. (4.12) must satisfy the Rainich–Misner–Wheeler (RMW) equations[10–12]

$$M^a_a = 0, \tag{4.14}$$

$$M^a_m M^m_b = \tfrac{1}{4}(M_{mn}M^{mn})\delta^a_b, \tag{4.15}$$

$$\alpha_{a,b} - \alpha_{b,a} = 0, \tag{4.16}$$

$$\alpha_a \equiv \frac{\sqrt{-g}\,\epsilon_{abcd}M^{bm;c}M^d_m}{M^2}, \tag{4.17}$$

$$M^2 \equiv M_{mn}M^{mn} \neq 0,$$

$$M_{ab}V^aV^b > 0, \tag{4.18}$$

where the last equation must be true for all timelike vectors V^a. We shall now check whether these equations are satisfied by Eq. (4.12).

Of the RMW equations, the differential equation (4.16) is trivially satisfied in this case as α_i vanishes identically by virtue of the diagonality of g_{ij} and M_{ij}. The equation (4.18) requires

$$M^0_0 > 0 \tag{4.19}$$

and the other two equations (4.14) and (4.15) are satisfied if

$$M^0_0 + M^2_2 = 0, \tag{4.20}$$

which, when written completely, reads

$$2\Lambda + k^2 + e^{-2N}(M_{22} + M_3^2 - M_3N_3) + e^{-2M}(N_{22} + N_2^2 - M_2N_2) = 0. \tag{4.21}$$

Thus, in Eqs. (4.19) and (4.21), we have two conditions to determine the three unknowns L, M, and N, indicating that solutions of the RMW equations do exist. Of these two equations, Eq. (4.19), combined with Eq. (4.20), conditions only the two constants k^2 and Λ such that

$$\Lambda + k^2 > 0. \tag{4.22}$$

The other equation (4.21) determines only M and N and hence L is left undetermined. Thus, we can choose any convenient functional form for $L(x)$. Regarding M and N, as there is only one Eq. (4.21) to determine them, we choose any one of them arbitrarily and determine the other. With each such solution (M,N) and a convenient choice of L (which only amounts to a choice of a suitable scale for the x coordinate) we obtain an electrovac solution admitting a homogeneous gravitational force field of intensity **a** given in Eq. (4.5). Thus, solutions of Eqs. (4.21) and (4.22) form a whole class of static electrovac solutions admitting homogeneous gravitational force fields.

5. SOLUTIONS OF THE DIFFERENTIAL EQUATION (4.21)

Depending on the constant parameter

$$p^2 \equiv k^2 + 2\Lambda \qquad (5.1)$$

which satisfies, by virtue of Eq. (4.22),

$$p^2 > \Lambda, \qquad (5.2)$$

a number of cases arise with the differential equation (4.21). They can be collected under three groups corresponding to $\Lambda \gtreqless 0$ as follows:

case I: $\Lambda > 0, \quad p^2 > \Lambda > 0, \quad k^2 \geqslant 0;$

case II: $\Lambda = 0, \quad p^2 = k^2 > 0;$

case III: $\Lambda < 0, \quad p^2 > \Lambda, \quad p^2 \gtreqless 0.$

Further, there are two subcases in I corresponding to $k^2 \geqslant 0$ and there are three subcases in III corresponding to $p^2 \gtreqless 0$. We shall designate them as follows:

case Ia: $\Lambda > 0, \quad p^2 > \Lambda > 0, \quad k^2 > 0;$

case Ib: $\Lambda > 0, \quad p^2 > \Lambda > 0, \quad k^2 = 0;$

case IIIa: $\Lambda < 0, \quad p^2 > 0 > \Lambda, \quad k^2 > -2\Lambda;$

case IIIb: $\Lambda < 0, \quad p^2 = 0 > \Lambda, \quad k^2 = -2\Lambda;$

case IIIc: $\Lambda < 0, \quad 0 > p^2 > \Lambda, \quad k^2 < -2\Lambda.$

The parameter p is real in all the cases with the exception of the case IIIc, in which it is purely imaginary. We now give some solutions of Eq. (4.21). As already remarked, Eq. (4.21) alone does not determine both the functions M and N and we have to supplement it with an extra condition which is at our choice. Thus we consider the following special solutions.

A. Solutions with $M = N$

In this case Eq. (4.21) reduces to

$$M_{22} + M_{33} = -p^2 e^{2M}. \qquad (5.3)$$

Even this is not a simple equation and we have not found any solutions for $p \neq 0$. However, when $p = 0$, which corresponds to case IIIb, we obtain the familiar two-dimensional Laplacian equation

$$M_{22} + M_{33} = 0, \qquad (5.4)$$

which has a whole class of solutions. For example, we may choose $M = M(y,z)$ to be the real or imaginary part of any complex function which is analytic in the complex variable $y + iz$.

B. Solutions with $N = 0$

In this case Eq. (4.21) becomes

$$M_{33} + M_3^2 + p^2 = 0. \qquad (5.5)$$

Since this does not detemine the y dependence of M, we shall obtain some solutions by spcifically assuming that M is independent of y. Then, writing

$$e^M = f(z), \qquad (5.6)$$

we obtain from Eq. (5.5)

$$\frac{d^2 f}{dz^2} + p^2 f = 0. \qquad (5.7)$$

Thus, the y-independent solutions of Eq. (5.5) is

$$e^M = A\cos(pz) + B\sin(pz), \qquad (5.8)$$

where A and B are two constants of integration. Solutions of this form exist for all the types except IIIb and IIIc. We must choose the solution of Eq. (5.7) as

$$e^M = A\cosh(p'z) + B\sinh(p'z), \quad p \equiv ip', \qquad (5.9)$$

in the Case IIIc, where $p^2 < 0$, and as

$$e^M = A + Bz \qquad (5.10)$$

in the Case IIIb, where $p = 0$.

Apart from these solutions, we can, in principle, obtain other solutions by determining M for a given functional form of N, as in that case Eq. (4.21) becomes a partial differential equation in M only. We shall however be satisfied with the special solutions that we have obtained above as they contain solutions of all types Ia to IIIc.

6. PROPERTIES OF THE SOLUTIONS

Here we shall study some properties of the electrovac solutions that we obtained in the previous section. The properties studied include the electromagnetic fields and their sources, the Petrov types, and the singularities of the metrics.

A. An orthonormal tetrad

For our future calculations we need an orthonormal tetrad (OT) of vectors. For Eq. (4.4), the unit tangent vectors of the coordinate curves of the coordinate system (t,x,y,z) form a convenient OT. This OT is given by

$$e^i_{(0)} = (e^{-\psi},0,0,0), \quad e^i_{(1)} = (0,e^{-L},0,0),$$
$$e^i_{(2)} = (0,0,e^{-M},0), \quad e^i_{(3)} = (0,0,0,e^{-N}), \qquad (6.1)$$

where the indices in the brackets label the vectors.

B. The electromagnetic field

Since the complexion scalar α is a constant throughout the space–time represented by Eq. (4.4) (this follows from the vanishing of $\alpha_i \equiv \alpha_{,i}$), the electromagnetic field F_{ab} of Eq. (4.4) is related to its Maxwell root f_{ab} [10–12] by a constant duality rotation, i.e.,

$$F_{ab} = \cos\alpha f_{ab} - \sin\alpha\, {}^*f_{ab}, \quad \alpha = \text{const}, \qquad (6.2)$$

where ${}^*f_{ab} = \frac{1}{2}(-g)^{1/2}\epsilon_{abcd}f^{cd}$ is the dual of f_{ab}. The arbitrariness of the constant α can be used to give any desired complexion to the field F_{ab}. The choices $\alpha = 0$ and $\pi/2$ correspond, respectively, to purely electric and magnetic complexions whereas any other choice of α in $(0,\pi/2)$ gives F_{ab} a mixed complexion. To find the Maxwell root of Eq. (4.4) we use the prescription of Misner and Wheeler[11] which leads to

$$(f_{01})^2 = q^2 e^{2L+2\psi}, \quad q^2 \equiv \Lambda + k^2 > 0. \tag{6.3}$$

Thus,

$$f_{01} = q\, e^{L+\psi}, \tag{6.4}$$

where we have chosen the positive root. The negative root, if desired, can be obtained from Eq. (6.4) by a duality rotation through $\alpha = \pi$. It is easy to check that all other components of f_{ab} are zero and thus Eq. (6.4) gives the complete Maxwell root. Using Eqs. (6.4) and (6.2), we get the electromagnetic fields of Eq. (4.4). The tetrad components $f_{(ab)}$ of f_{ab} in the OT (6.1) are given by

$$f_{(01)} = q. \tag{6.5}$$

All other components vanish.

C. Field singularities and sources

It is believed that field singularities are the seat of the sources of the electromagnetic field. These singularities may, in general, appear as singular points, lines, or surfaces and to locate them we need the invariants of the field. The two invariants of the electomagnetic field (6.2) are given by

$$F_1 = \tfrac{1}{2}F_{ab}F^{ab} = -q\cos 2\theta,$$
$$F_2 = \tfrac{1}{2}F^{*}_{ab}F^{ab} = -q\sin 2\theta. \tag{6.6}$$

As both these invariants are constants, it appears that there are no intrinsic (i.e., coordinate-independent) singularities for the field (6.2).

D. Petrov classification

The Petrov classification is based on the invariants of the Weyl conformal tensor C_{abcd}. For electrovac space–time in the presence of the cosmological constant Λ, C_{abcd} is defined as[13]

$$C_{abcd} = R_{abcd} + M_{abcd} - \frac{2\Lambda}{3} g_{abcd}, \tag{6.7}$$

where

$$M_{abcd} = \tfrac{1}{2}(g_{ac}M_{bd} + g_{bd}M_{ac} - g_{ad}M_{bc} - g_{bc}M_{ad}) \tag{6.8}$$

and

$$g_{abcd} = \tfrac{1}{2}(g_{ac}g_{bd} - g_{ad}g_{bc}). \tag{6.9}$$

Referred to the OT (6.1), C_{abcd} has the nonvanishing components

$$-C_{(0101)} = C_{(2323)} = 2C_{(0202)} = 2C_{(0303)} = -2C_{(1212)}$$
$$= -2C_{(1313)} = 2\Lambda/3. \tag{6.10}$$

From this it follows that (see, for example, Anderson[14]) C_{abcd} is of Petrov type D. For solutions with $\Lambda = 0$, C_{abcd} is conformally flat.

E. Complete system of second order differential invariants and singularities of the metric

We now determine the basis of the complete system of second order invariant functions of the metric tensor g_{ab} of Eq. (4.4). We do this in an attempt to locate intrinsic singularities of the metric.

The basis for the absolute scalar invariants of order 2 consists of 14 independent invariants [Ref. 15,16, and 13 (pp. 126 to 131)]. Here we follow the scheme of enumeration given by Narlikar and Karmarkar.[15] According to this scheme, the 14 invariants are

$$I_1 = R^a_{\ a}, \quad I_2 = R^b_{\ a}R^a_{\ b}, \quad I_3 = R^a_{\ b}R^b_{\ c}R^c_{\ a},$$
$$I_4 = R^a_{\ b}R^b_{\ c}R^c_{\ d}R^d_{\ a}, \tag{6.11}$$

$$J_1 = g^{hj}g^{ik}A_{hijk}, \quad J_2 = g^{hj}g^{ik}B_{hijk},$$
$$J_3 = g^{hj}g^{ik}E_{hijk}, \quad J_4 = g^{hj}g^{ik}F_{hijk}, \tag{6.12}$$

$$K_1 = R^{hj}R^{ik}C_{hijk}, \quad K_2 = R^{hj}R^{ik}A_{hijk},$$
$$K_3 = R^{hj}R^{ik}\widehat{D}_{hijk},$$
$$K_4 = Q^{hj}Q^{ik}C_{hijk}, \quad K_5 = Q^{hj}Q^{ik}A_{hijk},$$
$$K_6 = Q^{hj}Q^{ik}D_{hijk}, \tag{6.13}$$

where C_{hijk} is the Weyl tensor, and

$$A_{hijk} = C_{hipq}C_{jkrs}g^{pr}g^{qs}, \tag{6.14}$$
$$B_{hijk} = C_{hipq}A_{jkrs}g^{pr}g^{qs}, \tag{6.15}$$
$$D_{hijk} = B_{hijk} - \frac{1}{12}J_2(g_{hj}g_{ik} - g_{hk}g_{ij}) - \tfrac{1}{4}J_1 C_{hijk}, \tag{6.16}$$
$$\widehat{D}_{hijk} = (|J_3|)^{-1/2}D_{hijk}, \tag{6.17}$$
$$E_{hijk} = C_{hipq}D_{rsjk}g^{pr}g^{qs}, \tag{6.18}$$
$$F_{hijk} = C_{hipq}E_{rsjk}\, g^{pr}g^{qs}, \tag{6.19}$$
$$Q^i_{\ j} = R^i_{\ k}R^k_{\ j}. \tag{6.20}$$

Since we are evaluating invariants, we can use any coordinate system and we shall use a locally Galilean coordinate system in which $g_{ij} = g^{ij} = \mathrm{diag}(-1,1,1,1)$. The OT (6.1) provides such a coordinate system and the components of any tensor in this coordiante system are simply the tetrad components with respect to Eq. (6.1). In view of the special diagonal structure of the metric tensor g^{ij}, the formulas (6.12)–(6.20) simplify considerably and we get after a straightforward computation

$$I_1 = -4\Lambda, \ I_2 = 2k^4 + 2(2\Lambda + k^2)^2,$$
$$I_3 = 2k^6 - 2(2\Lambda + k^2)^3,$$
$$I_4 = 2k^8 + 2(2\Lambda + k^2)^4, \ J_1 = 8\Lambda^2/9,$$
$$J_2 = 8\Lambda^3/9, \ J_3 = -8\Lambda^4/9,$$
$$J_4 = -8\Lambda^5/27, \ K_1 = (16\Lambda/3)(\Lambda + k^2)^2,$$
$$K_2 = (8\Lambda^2/9)\,[(2\Lambda + k^2)^2 - 2\Lambda k^2],$$
$$K_3 = \frac{-2\sqrt{8}}{3}\Lambda(\Lambda + k^2)^2,$$
$$K_4 = (64/3)\Lambda^3(\Lambda + k^2)^2,$$
$$K_5 = (8\Lambda^2/9)\,[k^8 + k^4(k^2 + 2\Lambda)^2 + (2\Lambda + k^2)^4],$$
$$K_6 = -(64/9)\Lambda^5(\Lambda + k^2)^2.$$

Thus, we see that all the 14 basis, absolute, invariants of order 2 associated with the metric (4.4) are constants throughout space–time and hence are completely free of singularities. This tempts us to conclude that the space–time domain in which Eq. (4.4) is valid is itself free of singularities, but we cannot be sure of this as these invariants give only a partial information on the nature of the space–time geometry.

Summing up, we see that the electrovac solutions associated with Eq. (4.4) appear to exhibit a remarkable property in that all the second order differential invariants and the electromagnetic-field invariants are completely free of singularities.

F. Discussion of some special solutions

To conclude, we now display some typical special solutions and discuss a few of their properties. The solutions with the corresponding choice of the functions and the constant parameters Λ and k are collected below. All these solutions admit a homogeneous gravitational 3-force field of intensity **a** given by

$$a_\alpha = (k,0,0). \tag{6.21}$$

Solution I is $\Lambda > 0$, $k^2 > 0$, $p = (2\Lambda + k^2)^{1/2}$, $L = 0$, $\psi = kx$, $N = 0$, $M = A\cos(pz) + B\sin(pz)$, $ds^2 = dx^2 + dy^2\exp[2A\cos(pz) + 2B\sin(pz)] + dz^2 - e^{2kx}dt^2$, Petrov type D. Solution II is $\Lambda > 0$, $k = 0$, $p = (2\Lambda)^{1/2}$, $L = 0$, $\psi = 0$, $N = 0$, $M = A\cos(pz) + B\sin(pz)$, $ds^2 = dx^2 + dy^2 \times\exp[2A\cos(pz) + 2B\sin(pz)] + dz^2 - dt^2$, Petrov type D; the coordinate system is static and synchronous. Solution III is $\Lambda = 0$, $k^2 > 0$, $L = 0$, $\psi = kx$, $N = 0$, $M = A\cos(kz) + B\sin(kz)$, $ds^2 = dx^2 + dy^2\exp[2A\cos(kz) + 2B\sin(kz)] + dz^2 - e^{2kx}dt^2$, conformally flat solution without the cosmological term Λ. Solution IV is $\Lambda < 0$, $k^2 > -2\Lambda > 0$, $p = (2\Lambda + k^2)^{1/2}$, $L = 0$, $\psi = kx$, $N = 0$, $M = A\cos(pz) + B\sin(pz)$, $ds^2 = dx^2 + dy^2\exp[2A\cos(pz) + 2B\sin(pz)] + dz^2 - e^{2kx}dt^2$, Petrov type D. Solution V is $\Lambda < 0$, $k^2 = -2\Lambda > 0$, $L = 0$, $\psi = kx$, $M = N = 0$, $ds^2 = dx^2 + dy^2 + dz^2 - e^{2kx}dt^2$, Petrov type D; space sections $t =$ constant are all Euclidean. Solution VI is $\Lambda < 0$, $-2\Lambda > k^2 > 0$, $p = (-2\Lambda - k^2)^{1/2}$, $L = 0$, $\psi = kx$, $N = 0$, $M = A\cosh(pz) + B\sinh(pz)$, $ds^2 = dx^2 + dy^2\exp[2A\cosh(pz) + 2B\sinh(pz)] + dz^2 - e^{2kx} \times dt^2$, Petrov type D.

In all these solutions A and B denote two arbitrary constants of integration.

Lastly, we wish to make a few remarks on the solutions II, III, and V given above. The solution II corresponds to a coordinate system in which the gravitational 3 force is zero. Moreover, it provides an example of a nonflat metric which admits a *static synchronous* reference system. (A coordinate system in which $g_{00} = -1$ and $g_{0\alpha} = 0$ is said to be a synchronous coordinate system.) This appears to contradict, at first sight, the theorem quoted in Landau and Liftshitz [Ref. 7 (p. 292)] that "a synchronous reference system cannot be stationary." This simply means that there cannot be a nonflat metric in which $g_{00} = -1$, $g_{0\alpha} = 0$, and the other metric tensor components are independent of the time coordinate t. However, there is no contradiction as the Landau–Liftshitz theorem holds only when $\Lambda = 0$ and in solution II, $\Lambda \neq 0$. Also, this metric admits an *inertial* reference frame in the sense of Audretsch.[17] According to Audretsch, a reference frame which is both static and synchronous is the analogue of the inertial frame in GTR and such a reference frame is provided by a timelike congruence U_a which is covariantly constant, i.e.,

$$U_{a;b} = 0. \tag{6.22}$$

In view of $\psi = 0$, the time line congruence of this solution (II) clearly satisfies Eq. (6.22) [see Eq. (2.4)]. Audretsch has also shown that if an electrovac solution admitting an inertial reference frame exists, then it must be a Petrov D-type solution corresponding to a nonzero negative value of Λ. This is clearly consistent with the solution II. (Note that Audretsch works with a metric of signature -2 and his condition $\Lambda < 0$ when translated into signature $+2$ with which we are working, becomes $\Lambda > 0$.)

Solution III is worth noting as it is an example of a space–time admitting a homogeneous gravitational force field without the cosmological constant Λ. Note that all the other solutions we have found need a nonzero Λ and it appears as if a homogeneous gravitational force field can be found only when $\Lambda \neq 0$. Thus, it is satisfying that we have at least one example of a homogeneous gravitational force field with $\Lambda = 0$.

The solution V is essentially the solution

$$ds^2 = dx^2 + dy^2 + dz^2 - e^{\mathbf{k}\cdot\mathbf{r}}dt^2, \tag{6.23}$$

which we obtained some time ago.[18] Here, **k** and **r** denote respectively the 3 vectors (k_1, k_2, k_3) and (x,y,z). On choosing the x axis along the direction of **k**, Eq. (6.23) reduces to the solution V. The metric V also bears an interesting relationship with the "Godel-type" metric[19]

$$ds^2 = dx^2 + \tfrac{1}{2}e^{2kt}dy^2 + dz^2 - (dt + e^{kt}dy)^2, \tag{6.24}$$

where $k = (-2\Lambda)^{1/2}$, $\Lambda < 0$. This metric can be interpreted as the geometry of an electromagnetic field with a removable gravitational field in the background. If we replace e^{kt} by e^{kx} in Eq. (6.24), we get the well known Godel solution [Ref. 8 (p. 438)] and that is why we have called Eq. (6.24) a Godel-type solution. However, the appearance of Eq. (6.24) as a stationary solution is only superficial and in Ref. 19 this metric was *erroneously* called a nonstatic metric. In fact, Eq. (6.24) is precisely the Solution V written in a new coordinate system. The coordinate transformation

$$t = \frac{x'}{\sqrt{2}} + \frac{\sqrt{2}}{k}e^{-kt'}, \quad x = t', \quad y = y', \quad z = z'$$

sends the solution V into

$$ds^2 = \tfrac{1}{2}e^{2kt'}dx'^2 + dy'^2 + dz'^2 - (dt' + e^{kt'}dx')^2,$$

which is Eq. (6.24) except for a trivial renaming of y' as x' and *vice versa*.

[1]P.A. Goodinson, Int. J. Theor. Phys. **8**, 179 (1973).
[2]A. Trautman, in *Lectures on General Relativity*, (Prentice Hall, Englewood Cliffs, 1964), Sec. 6.6.
[3]H. Dehnen, Int. J. Theor. Phys. **3**, 509 (1970).
[4]C. Møller, *The Theory of Relativity* (Oxford U.P., London, 1972), 2nd ed.
[5]J.L. Synge, *Relativity, The General Theory* (North-Holland, Amsterdam, 1960).

[6]J. Ehlers and W. Kundt, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962), Chap. 2.
[7]L.D. Landau and E.M. Liftshitz, *The Classical Theory of Fields* (Pergamon, Oxford, 1971), 3rd edition.
[8]R. Adler, M. Bazin and M. Schiffer, *Introduction to General Relativity* (McGraw-Hill, New York, 1975), 2nd edition.
[9]D. Lovelock and H. Rund, *Tensors, Differential Forms, and Variational Principles*, Secs. 2.5 and 3.7 (Wiley, New York, 1975).
[10]G.Y. Rainich, Trans. Am. Math. Soc. **27**, 106 (1925).
[11]C.W. Misner and J.A. Wheeler, Ann. Phys. (N.Y.) **2**, 525 (1957).
[12]L. Witten, in *Gravitation, An Introduction to Current Research*, edited by L. Witten: (Wiley, New York 1962), Chap. 9.
[13]A.Z. Petrov, *Einstein Spaces* (Pergamon, Oxford, 1969).
[14]J.L. Anderson, *Principles of Relativity Physics* (Academic, New York, 1967), pp. 65–7.
[15]V.V. Narlikar and K.R. Karmarkar, Proc. Indian Acad. Sci. A **29**, 91 (1949).
[16]J. Geheniau and R. Debever, Bull. Cla. Sci. Acad. R. Belg. **42**, 114 (1956).
[17]J. Audresch, Int. J. Theor. Phys. **4**, 1 (1971).
[18]K.N. Srinivasa Rao and A.V. Gopala Rao, Phys. Lett. A **39**, 291 (1972).
[19]A.V. Gopala Rao, Ind. J. Pure Appl. Phys. **11**, 351 (1973).

Exact solutions of Brans–Dicke theory for irrotational barotropic models with stiff matter

V. B. Johri and G. K. Goswami
Department of Mathematics, Univeristy of Gorakhpur, Gorakhpur 273001 India

(Received 1 May 1979; accepted for publication 3 October 1979)

In the previous work, we have obtained classes of exact solutions of Brans–Dicke (BD) theory when only a scalar field is present and also in the presence of an electromagnetic field. In the present paper, we have generated a new class of solutions of BD theory corresponding to irrotational barotropic perfect fluid with pressure equal to energy density. The main result of the paper may be stated as follows: "Corresponding to any diagonalizable solution of Einstein vacuum field equations, in which fields and metric tensors are functions of not more than three variables, we can generate a solution of the Brans–Dicke theory for the irrotational barotropic perfect fluid with pressure equal to energy density."

1. INTRODUCTION

In a recent paper,[1] it was found that the Einstein equations for a self-gravitating fluid, with pressure equal to energy density and four-velocity u_a, are equivalent to the field equations

$$R_{ab} = -2\theta_a\theta_b, \quad (1.1)$$

$$\Box\theta = \theta^k{}_{;k} = 0. \quad (1.2)$$

In case of irrotational fluids, we have

$$u_a = \theta_a / \sqrt{\theta_c\theta^c}. \quad (1.3)$$

The pressure and density of the universe are given by

$$p = \rho = \theta_c\theta^c, \quad (1.4)$$

and the energy–momentum tensor is

$$T_{ab} = 2\theta_a\theta_b - g_{ab}\theta_c\theta^c. \quad (1.5)$$

The general solutions of the field equations (1.1) and (1.2) for plane symmetric and cylindrically symmetric metrics have been obtained by the authors given in Refs. 1 and 2. It was also pointed out that the field equations are related to BD-field equations in vacuum. The latter equations are reducible to the form

$$R_{ab} = \theta_a\theta_b,$$

by defining

$$\exp[\theta/(w+3/2)^{1/2}] = \phi_{BD},$$

$$g_{ab} = \exp[\theta/(w+3/2)^{1/2}] g_{ab(BD)}.$$

The present paper is a continuation of our previous work[3,4] in which we have obtained classes of exact solutions of BD-theory when only the scalar field is present and also in the presence of an electromagnetic field. The solutions were obtained by transforming BD-field equations into Einstein vacuum field equations assuming the functional relationship amongst scalar fields, one component of the metric tensor and field potential. Here we have considered the energy–momentum tensor in the form of a perfect fluid given by (1.4) and (1.5). In this way we have obtained one more class of solutions of BD-field equations by transforming them into Einstein field equations.

In Sec 2, we have set up BD-field equations in a suitable form by assuming a functional relationship amongst the scalar field ϕ, one component of metric tensor, i.e., g_{kk}, and field potential θ, then we have established the main result of the paper. In Sec. 3, BD-solutions have been obtained corresponding to the Kasner solution and a solution obtained by Mishra and Radhakrishna[5] of the Einstein theory. Sec 4 contains some concluding remarks.

2. DERIVATION OF BD-FIELD EQUATIONS FOR PERFECT FLUIDS FROM EINSTEIN VACUUM FIELD EQUATIONS

The BD-field equations are

$$R_{ij} = -1/\phi\,[T_{ij} - (w+1)/(2w+3)Tg_{ij}] - w/\phi^2\phi_i\phi_j - \phi_{;ij}/\phi, \quad (2.1)$$

$$\phi^k{}_{;k} = T/(2w+3). \quad (2.2)$$

We consider the energy–momentum tensor as that of a perfect fluid given by (1.4) and (1.5). We attempt to solve the field equations (2.1) and (2.2) for the universe defined by

$$ds^2 = \epsilon\exp(2u)[(dx^k)^2 + \Sigma_{\alpha\beta}dx^\alpha dx^\beta], \quad (2.3)$$

where $\epsilon = \mathrm{sgn}(g_{kk})$, and k is either 0,1,2, or 3. Greek letters take the values 0,1,2, and 3 except k. BD-field equations (2.1), (2.2), and (1.2) in terms of metric (2.3) may be written as

$$P_{\alpha\beta} + \Sigma_{\alpha\beta}\Delta_2 u - 2u_\alpha u_\beta + 2u_{\beta;\alpha} + 2\Sigma_{\alpha\beta}\Delta_1(u) = -\exp(-\delta)[2\theta_\alpha\theta_\beta - 1/(2w+3)\Sigma_{\alpha\beta}\Delta_1(\theta)] - (w+1)\delta_\alpha\delta_\beta - \delta_{\alpha;\beta} + (\delta_\alpha u_\beta + \delta_\beta u_\alpha) - \Sigma_{\alpha\beta}\Delta_1(\delta,u), \quad (2.4)$$

$$\Delta_2(u) + 2\Delta_1(u) = \exp(-\delta)\Delta_1(\theta)/(2w+3) - \Delta_1(u,\delta), \quad (2.5)$$

$$\Delta_2(\delta) + 2\Delta_1(u,\delta) = -2\exp(-\delta)\Delta_1(\theta)/(2w+3) - \Delta_1(\delta), \quad (2.6)$$

and

$$\Delta_2(\theta) = -2\Delta_1(u,\theta), \quad (2.7)$$

 0022-2488/80/082269-03$1.00

where

$\Delta_1(u,v) = \Sigma_{\alpha\beta} u^\alpha v^\beta$,

$\Delta_2(u) = u^\alpha_{;\alpha}$,

$\phi = \exp(\delta)$,

and u, $\Sigma_{\alpha\beta}$ being functions of x^α. $P_{\alpha\beta}$ is the Ricci tensor formed w.r.t. $\Sigma_{\alpha\beta}$ and covariant derivatives are also taken w.r.t. $\Sigma_{\alpha\beta}$. From (2.5) and (2.6), we get the following linear relationship between u and δ:

$$\delta = -2u. \tag{2.8}$$

If we further assume the functional relationship between u and θ, i.e, if we assume

$$v = \exp(-u), \quad v = v(X), \quad \theta = \theta(X), \text{ and } X(x^\alpha), \tag{2.9}$$

then with the help of Eqs. (2.5), (2.6), and (2.8), it can be proved that

$$v = f\,\text{sech}(X), \qquad \theta = c\tanh(X), \tag{2.10}$$

f and c are constants satisfying

$$c^2 = (2w+3)f^2 \tag{2.11}$$

therefore field equations (2.4)–(2.6) convert to

$$P_{\alpha\beta} + 2(2w+3)X_\alpha X_\beta = 0, \qquad \Delta_2(X) = 0, \tag{2.12}$$

which are the Einstein field equations $R_{ij} = 0$, for the metric

$$ds^2 = \epsilon[\exp(2K)(dx^k)^2 + \exp(-2K)\Sigma_{\alpha\beta}dx^\alpha dx^\beta], \tag{2.13}$$

where we have assumed

$$K = \sqrt{(2w+3)}\,X. \tag{2.14}$$

Thus we have established the following result: "Corresponding to any diagonalizable solution of Einstein vacuum field equations in which fields and metric coefficients are functions of not more than three variables we can generate a solution of BD-theory for irrotational barotropic perfect fluid with pressure equal to energy density."

Mathematically, suppose the metric (2.13) satisfies Einstein vacuum field equations, then the metric (2.3) will satisfy BD-field equations for barotropic perfect fluid. The scalar field is given by

$$\phi = \exp(-2u) = f^2\text{sech}^2[K/\sqrt{(2w+3)}], \tag{2.15}$$

field potential by

$$\theta = \sqrt{(2w+3)}f\tanh[K/\sqrt{(2w+3)}],$$

the density and pressure of the universe by

$$p = \rho = f^4\text{sech}^6[K/\sqrt{(2w+3)}]K_\alpha K^\alpha, \tag{2.16}$$

and the four velocity by

$$u_a = \theta_a/\sqrt{(\theta_c\theta^c)} = \cosh|K/\sqrt{(2w+3)}|K_a/f\sqrt{(K_\alpha K^\alpha)}. \tag{2.17}$$

3. CYLINDRICALLY SYMMETRIC EXPANDING BD-UNIVERSE

We apply the result obtained in the previous section to the Kasner vacuum solution of Einstein theory.

The homogeneous and anisotropic Kasner model is given by

$$ds^2 = dt^2 - t^{2p_1}dx^2 - t^{2p_2}dy^2 - t^{2p_3}dz^2, \tag{3.1}$$

where the constants satisfy

$$\Sigma p_i = 1 \text{ and } \Sigma p_i^2 = 1. \tag{3.2}$$

The corresponding BD-solution will be given by

$$ds^2 = V^2[t^{2p_1}dt^2 - dx^2 - t^{2(p_1+p_2)}dy^2 - t^{2(p_1+p_3)}dz^2],$$

with the scalar field

$$\phi = V^{-2} = \frac{4f^2 t^{2p_1/\sqrt{(2w+3)}}}{(1+t^{2p_1/\sqrt{2w+3}})^2},$$

the pressure and density

$$p = \rho = \frac{\rho_0 t^{(6p_1/\sqrt{(2w+3)} - 2(p_1+1))}}{[(1+t^{2p_1/\sqrt{(2w+3)}}]^6},$$

ρ_0 being constant and four-velocity are $(0,0,0,\sqrt{g_{44}})$. Using the coordinate transformation

$$(p_1+1)t^{p_1}dt \to dT,$$
$$(p_1+1)dx \to dX,$$

the above solution is transformed to

$$ds^2 = V^2[L^2(dT^2 - dX^2) - T^{1+a}dy^2 - T^{1-a}dz^2] \tag{3.3}$$

with

$$\phi = 4f^2T^{2p}/(1+T^{2p})^2, \tag{3.4}$$

$$p = \rho = \rho_0\phi^3/T^2, \tag{3.5}$$

$$L = 1/(1+p_1),$$

$$a = (p_3 - p_2)/(1+p_1),$$

$$p = p_1/(1+p_1)\sqrt{(2w+3)}.$$

The solution (3.3) represents cylindrically symmetric expanding BD-universe which corresponds to the solution obtained by Roy and Singh[6] in relativistic theory ($\phi =$ const and $w \to \infty$).

A more general solution: A more general cylindrically symmetric solution can be generated from the Einstein vacuum solution of Mishra and Radhakrishna[5]

$$ds^2 = \exp(m^2r^2/4 + mt)(dt^2 - dr^2) - \exp(mt)r^2d\theta^2 - \exp(-mt)dz^2. \tag{3.6}$$

The corresponding BD-solution will be given by

$$ds^2 = V^2[\exp(m^2r^2/4)(dt^2 - dr^2) - (dz^2 + r^2d\theta^2)], \tag{3.7}$$

with scalar field

$$\phi = \frac{4f^2\exp(-mt/\sqrt{(2w+3)})}{[1+\exp(-mt/\sqrt{(2w+3)})]^2},$$

pressure and density

$$p = \rho$$

$$= \rho_0 \exp\left[-mt/\sqrt{(2w+3)} + m^2r^2/4 \right] \times \{ [\exp(-mt/\sqrt{(2w+3)}) + 1]^6 \times [\exp(mt/\sqrt{(2w+3)}) + 1]^2 \}^{-1},$$

and the four-velocity $u_\alpha = (0,0,0,\sqrt{g_{44}})$.

This solution represents a cylindrically symmetric and inhomogeneous expanding BD-universe.

4. CONCLUSION

The immediate use of the result derived in this paper is in obtaining exact solutions of Brans–Dicke theory for the Zeldovich universe from the known vacuum relativistic solutions.

In conclusion, we at least hope that some physical insight can be gained from these solutions in determining the implications for the primordial state of the universe.

[1]R. Tabensky and A.H. Taub, Comm. Math. Phys. **29**, 61 (1973).
[2]P.S. Letelier and R. Tabensky, Nuovo Cimento, B **28**, 407 (1975).
[3]G.K. Goswami, J. Math. Phys. **19**, 442 (1978).
[4]V.B. Johri and G.K. Goswami, J. Math. Phys. **19**, 987 (1978).
[5]M. Mishra and L. Radhakrishna, Proc. Nat. Inst. Sci. India, A **28**, 632 (1962).
[6]S.R. Roy and P.N. Singh, J. Phys. A: Math. Nucl. Gen. **9**, 255 (1976).

An alternative solution of the Bogoliubov's equation for the two-particle distribution function in the kinetic theory of gases

M. Chen
Department of Mathematics, Vanier College, Montreal, Quebec, Canada H4L 3X9

(Received 21 November 1978; accepted for publication 27 June 1979)

In this paper we consider an alternative solution of the Bogoliubov's equation for the two-particle distribution functions F_2 under the assumption that it is a function of the one-particle distribution function F_1. In terms of the density expansion $F_2 = \Sigma_{i=0}^{\infty} \rho^i F_2^{(i)}$, we show that $F_2^{(n)}$ satisfies a first order linear partial differential equation. The existence of a unique solution for $F_2^{(n)}$ can then be proved under some general conditions. Moreover, $F_2^{(n)}$ reduces to the result in the classical virial expansions at equilibrium. For the discussion of the kinetic equation, a special solution for $F_2^{(0)}$ can be obtained by the method of separation of variables. Following Grad's method, a kinetic equation can then be derived, which reduces to the ordinary Boltzmann equation for the hard sphere potential. It can be shown that the equilibrium solution of the kinetic equation is just the Maxwell–Boltzmann distribution. Under some restricted conditions the kinetic equation can be proved to satisfy the H theorem.

I. INTRODUCTION

In 1946 Bogoliubov[1] presented his well known dynamical theory of gases. Since then much work has been developed in its various applications; however, a rigorous study on the existence and uniqueness solution of the two-particle distribution function F_2 still seems lacking in the literature. In this paper we consider an alternative solution for F_2 under the assumption that F_2 is a function of F_1 (rather than a functional of F_1 as in Bogoliubov's method). By expanding F_2 in powers of the density ρ, $F_2 = F_2^{(0)} + \rho F_2^{(1)} + \rho^2 F_2^{(2)} + \cdots$, $F_2^{(n)}$ can be shown to satisfy a first order linear partial differential equation. Under some general conditions we can prove the existence of a unique solution for $F_2^{(n)}$. Furthermore, $F_2^{(n)}$ reduces to the classical virial expansion at equilibrium.

For the discussion of kinetic equation, we propose a special solution of $F_2^{(0)}$ of the form $F_2^{(0)} = \xi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{p}_1,\mathbf{p}_2) \times F_1(\mathbf{r}_1,\mathbf{p}_1,t) F_1(\mathbf{r}_2,\mathbf{p}_2,t)$ by the method of separation of variables in solving the partial differential equation for $F_2^{(0)}$. It is interesting to note that ξ represents the correlation of the two particles through their dynamical motions. Following Grad's method,[2] a general kinetic equation can then be obtained, which reduces to the ordinary Boltzmann equation for the hard sphere potential. Also, the equilibrium solution of the kinetic equation is just the Maxwell–Boltzmann distribution. When ξ is independent of $\mathbf{p}_1$ and $\mathbf{p}_2$, the kinetic equation then satisfies the H theorem.

II. REVIEW OF THE BOGOLIUBOV'S EQUATION FOR F_n

Consider a classical system of N particles with pairwise intermolecular potential v in a volume V. The state of the system can be described by a normalized probability distribution function f_N in Γ space. Let $x_i = (\mathbf{r}_i, \mathbf{p}_i) \in R^6$ be the position and momenta of the ith particle. The temporal development of f_N is governed by the Liouville equation

$$\frac{\partial f_N}{\partial t} = \{H_N, f_N\}$$

$$= \sum_{i=1}^{N} \left(\frac{\partial H_N}{\partial \mathbf{r}_i}\cdot\frac{\partial f_N}{\partial \mathbf{p}_i} - \frac{\partial H_N}{\partial \mathbf{p}_i}\cdot\frac{\partial f_N}{\partial \mathbf{r}_i}\right), \tag{1}$$

with Hamiltonian

$$H_N = \sum_{i=1}^{N} \tfrac{1}{2}\mathbf{p}_i\cdot\mathbf{p}_i + \sum_{i<j} v(|\mathbf{r}_i - \mathbf{r}_j|),$$

where for simplicity the particle mass m has been assumed to be 1.

Suppose the intermolecular potential v consists of a hard core of diameter σ and an attractive potential ϕ of finite range R, where $\phi = \phi(|\mathbf{r}_1 - \mathbf{r}_2|)$ is a function of class C^1 defined on the domain $D_\phi = \{r \mid r = |\mathbf{r}_1 - \mathbf{r}_2| \geqslant \sigma, (\mathbf{r}_1,\mathbf{r}_2) \in R^3 \times R^3\}$. Assume that $\rho = \lim_{\substack{N\to\infty\\V\to\infty}} (N/V) < \infty$, and denote $\mathbf{G}_{ij} = -\partial\phi_{ij}/\partial\mathbf{r}_i$, $\phi_{ij} = \phi(|\mathbf{r}_i - \mathbf{r}_j|)$. Define Grad's truncated functions $F_1, F_2, \ldots, F_k, \cdots$ by

$$F_1(x_1,t) = V \int_{D_2\times D_3\times\cdots\times D_N} f_N(x_1,x_2,\ldots,x_N,t)\, dx_2\cdots dx_N,$$

$$F_k(x_1,x_2,\ldots,x_k,t) = V^k \int_{D_{k+1}\times\cdots\times D_N} f_N(x_1,x_2,\ldots,x_N,t)\, dx_{k+1}\cdots dx_N, \quad k\geqslant 2,$$

where D_k is the subdomain of the γ_k space defined by $|\mathbf{r}_k - \mathbf{r}_1| \geqslant \sigma$ for all $\mathbf{P}_k$. By Grad's method we can obtain from Eq. (1) the following Grad's hierachy:

$$\frac{\partial F_1}{\partial t} = -\hat{L}_1 F_1 + \rho \widehat{M}_1 F_2,$$

$$\frac{\partial F_k}{\partial t} = -\hat{L}_k F_k + \rho \widehat{M}_k F_{k+1}, \quad k\geqslant 2, \tag{2}$$

where

$$\hat{L}_k = \sum_{i=1}^{k} \mathbf{p}_i\cdot\frac{\partial}{\partial \mathbf{r}_i} - \sum_{i<j}^{k}\left(\frac{\partial\phi_{ij}}{\partial\mathbf{r}_i}\cdot\frac{\partial}{\partial\mathbf{p}_i} + \frac{\partial\phi_{ij}}{\partial\mathbf{r}_j}\cdot\frac{\partial}{\partial\mathbf{p}_j}\right),$$

 0022-2488/80/082272-06$1.00

$$\widehat{M}_k F_{k+1} = \sum_{i=1}^{k} \left(\int d\mathbf{p}_{k+1} \oint_{|\mathbf{r}_i - \mathbf{r}_{k+1}| = \sigma} d\mathbf{S}\cdot(\mathbf{p}_{k+1} - \mathbf{p}_i) F_{k+1} - \int_E dx_{k+1}\, \mathbf{G}_{i,k+1}\cdot \frac{\partial}{\partial \mathbf{p}_i} F_{k+1} \right), \quad k \geqslant 1,$$

and E is the subdomain defined by $|\mathbf{r}_i - \mathbf{r}_{k+1}| \geqslant \sigma$ for $i = 1,2,...,k$.

In the following discussions we shall apply Bogoliubov's dynamical theory to Grad's hierarchy. Let τ be the time of a collision, t_0 the time between two collisions, and T the macroscopic relaxation time. For slowly varying macroscopic phenomena and not too dense system it is known that $\tau \ll t_0 \ll T$. From Eq. (2) it can be seen that F_k (for $k \geqslant 2$) changes rapidly in a time of order τ on account of intermolecular interaction [assuming $\phi(r) \to 0$ rapidly as $r \to \infty$], while F_1 changes slowly during this period of time since ϕ does not affect F_1 directly. The time scale for F_1 is t_0. According to Bogoliubov, after an initial chaotization time of order τ a kinetic stage is reached, in which all of the F_k for $k \geqslant 2$ depend on time t only through F_1. Let F_k be a function of $x_1, x_2, ..., x_k, F_1(x_1,t), F_1(x_2,t), ..., F_1(x_k,t)$:

$$F_k(x_1, x_2, ..., x_k, t) = F_k(x_1, x_2, ..., x_k, F_1(x_1,t), F_1(x_2,t), ..., F_1(x_k,t)), \quad (3)$$

where $F_1(x_i,t)$ is a function of class C^1 defined on $R^6 \times R^1$, whose temporal development is in turn governed by the following equation

$$\frac{\partial F_1}{\partial t} = -\hat{L}_1 F_1 + \rho \widehat{M}_1 F_2 = A(x_i, F_1(x_i,t)). \quad (4)$$

In order to obtain a general equation for F_k we consider the following series expansion of F_k and A in powers of the density ρ:

$$A = A^{(0)} + \rho A^{(1)} + \rho^2 A^{(2)} + \cdots,$$
$$F_k = F_k^{(0)} + \rho F_k^{(1)} + \rho^2 F_k^{(2)} + \cdots. \quad (5)$$

Substituting Eqs. (3), (4), and (5) into Eq. (2) yields the following results:

$$A^{(0)} = -\hat{L}_1 F_1,$$
$$D^{(0)} F_k^{(0)} = -\hat{L}_k F_k^{(0)}, \quad (6)$$

....

$$A^{(n)} = \int d\mathbf{p}_2 \oint_{|\mathbf{r}_1 - \mathbf{r}_2| = \sigma} d\mathbf{S}\cdot(\mathbf{p}_2 - \mathbf{p}_1) F_2^{(n-1)} - \int_{D_2} dx_2\, \mathbf{G}_{1,2}\cdot \frac{\partial}{\partial \mathbf{p}_1} F_2^{(n-1)}, \quad (7)$$

$$D^{(0)} F_k^{(n)} + \hat{L}_k F_k^{(n)} = -\sum_{i=1}^{n} D^{(i)} F_k^{(n-i)} + \sum_{i=1}^{k} \int d\mathbf{p}_{k+1} \times \oint_{|\mathbf{r}_i - \mathbf{r}_{k+1}| = \sigma} d\mathbf{S}\cdot(\mathbf{p}_{k+1} - \mathbf{p}_i) F_{k+1}^{(n-1)} - \sum_{i=1}^{k} \int_E dx_{k+1}\, \mathbf{G}_{i,k+1}\cdot \frac{\partial}{\partial \mathbf{p}_i} F_{k+1}^{(n-1)}, \quad n \geqslant 2, \quad (8)$$

where

$$D^{(0)} = \sum_i A^{(0)}(x_i, F_1(x_i,t)) \frac{\partial}{\partial F_1(x_i,t)},$$
$$D^{(n)} = \sum_i A^{(n)}(x_i, F_1(x_i,t)) \frac{\partial}{\partial F_1(x_i,t)}.$$

By considering F_1 as a known function (but completely arbitrary for the moment, which will be determined later by the kinetic equation), it is evident from Eqs. (6)–(8) that we can solve Eq. (6) for $F_k^{(0)}$, especially $F_2^{(0)}$, which then according to Eq. (7) enables us to obtain $A^{(1)}$. This in turn allows us to find $D^{(1)} F_k^{(0)}$ and then by Eq. (8) $F_k^{(1)}$ can be obtained, particularly $F_2^{(1)}$, and so on. In Sec. III we consider the solutions of $F_2^{(0)}$ and $F_2^{(n)}$.

III. SOLUTIONS OF $F_2^{(0)}$ AND $F_2^{(n)}$.

In order to simplify the notation we denote $u_1 = F_1(x_i,t)$, $\boldsymbol{\phi}_i = -\partial\phi/\partial\mathbf{r}_i$, $\mathbf{b}(\mathbf{r}_i, \mathbf{p}_i, u_i) = \partial u_i/\partial \mathbf{p}_i$, $g_i = \boldsymbol{\phi}_i \cdot \mathbf{b}(\mathbf{r}_i, \mathbf{p}_i, u_i)$, $i = 1,2$, $F_2^{(0)} = G(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1\ \mathbf{p}_2, u_1, u_2)$, and $F_2^{(n)} = H(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2, u_1, u_2)$, for $n \geqslant 1$. Equations (6) and (8) then become

$$\mathbf{p}_1 \cdot \frac{\partial G}{\partial \mathbf{r}_1} + \mathbf{p}_2 \cdot \frac{\partial G}{\partial \mathbf{r}_2} + \boldsymbol{\phi}_1 \cdot \frac{\partial G}{\partial \mathbf{p}_1} + \boldsymbol{\phi}_2 \cdot \frac{\partial G}{\partial \mathbf{p}_2} + g_1 \frac{\partial G}{\partial u_1} + g_2 \frac{\partial G}{\partial u_2} = 0 \quad (9)$$

and

$$\mathbf{p}_1 \cdot \frac{\partial H}{\partial \mathbf{r}_1} + \mathbf{p}_2 \cdot \frac{\partial H}{\partial \mathbf{r}_2} + \boldsymbol{\phi}_1 \cdot \frac{\partial H}{\partial \mathbf{p}_1} + \boldsymbol{\phi}_2 \cdot \frac{\partial H}{\partial \mathbf{p}_2} + g_1 \frac{\partial H}{\partial u_1} + g_2 \frac{\partial H}{\partial u_2} = \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2, u_1, u_2), \quad (10)$$

where ψ represents the right-hand side of Eq. (8).

If G and H are considered as functions of the independent variables $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{p}_1$, $\mathbf{p}_2$, u_1, and u_2, defined on the domain $D = R^{12} \times C^1(R^7) \times C^1(R^7)$ with t as a parameter, Eqs. (9) and (10) then become first order linear partial differential equations. Their solutions can be constructed from the corresponding characteristic equations

$$\frac{dr_{1i}}{p_{1i}} = \frac{dr_{2j}}{p_{2j}} = \frac{dp_{1i}}{\phi_{1i}} = \frac{dp_{2j}}{\phi_{2j}} = \frac{du_1}{g_1} = \frac{du_2}{g_2} = d\lambda, \quad i,j = 1,2,3 \quad (11)$$

and

$$\frac{dr_{1i}}{p_{1i}} = \frac{dr_{2j}}{p_{2j}} = \frac{dp_{1i}}{\phi_{1i}} = \frac{dp_{2j}}{\phi_{2j}} = \frac{du_1}{g_1} = \frac{du_2}{g_2} = \frac{dH}{\psi} = d\lambda, \quad i,j = 1,2,3. \quad (12)$$

From Eqs. (11) and (12) we note that the following systems of equations

$$\frac{dr_{1i}}{d\lambda} = p_{1i}, \quad \frac{dr_{2i}}{d\lambda} = p_{2i}, \quad \frac{dp_{1i}}{d\lambda} = \phi_{1i}, \quad \frac{dp_{2i}}{d\lambda} = \phi_{2i}, \quad i = 1,2,3, \quad (13)$$

can be solved first, and then the solutions substituted into the equations

$$\frac{du_1}{d\lambda} = g_1, \frac{du_2}{d\lambda} = g_2, \frac{dH}{d\lambda} = \psi, \tag{14}$$

to solve for u_1, u_2, and H.

For convenience we let $y = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2, u_1, u_2)$, $A = (\mathbf{p}_1, \mathbf{p}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, g_1, g_2)$, $z = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2, u_1, u_2, H)$, and $B = (\mathbf{p}_1, \mathbf{p}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, g_1, g_2, \Psi)$. The characteristic equations (13) and (14) respectively can be written as

$$\frac{dy}{d\lambda} = A \tag{15}$$

and

$$\frac{dz}{d\lambda} = B. \tag{16}$$

The characteristics of Eq. (15) form a 13-parameter family of curves in a 14-dimensional space, whereas the characteristics of Eq. (16) form a 14-parameter family of curves. Since an integral surface of Eq. (10) is a 14-dimensional surface lying in a 15-dimensional space, an integral surface of Eq. (10) must be generated by a 13-parameter family of characteristic curves, i.e., solutions of Eq. (16). Let the parameters of Eqs. (15) and (16) be $\tau_1, \tau_2, \cdots, \tau_{13}$. In principle, the solutions of Eqs. (15) and (16) can be written as

$$y_i = z_i = f_i(\lambda, \tau_1, \cdots, \tau_{11}), \quad i = 1,2,\cdots,12,$$
$$y_i = z_i = f_i(\lambda, \tau_1, \cdots, \tau_{13}), \quad i = 13,14,$$
$$H = z_{15} = f_{15}(\lambda, \tau_1, \cdots, \tau_{13}).$$

Let $y(0)$ be given at $\lambda = 0$ and assume that $|\mathbf{r}_1(0) - \mathbf{r}_2(0)| \gg R$, where $\phi(R) = 0$. Since the correlation between two particles can be neglected when their separation is much greater than R, thus $F_2(0) = u_1(0)\, u_2(0)$. By Eq. (5) it then follows that

$$G(0) = G(\mathbf{r}_1(0), \mathbf{r}_2(0), \mathbf{p}_1(0), \mathbf{p}_2(0), u_1(0), u_2(0)) = u_1(0)u_2(0) \tag{17}$$

and

$$H(0) = H(\mathbf{r}_1(0), \mathbf{r}_2(0), \mathbf{p}_1(0), \mathbf{p}_2(0), u_1(0), u_2(0)) = 0. \tag{18}$$

By Eq. (9) we have $dG/d\lambda = 0$ on the characteristics. The initial condition of Eq. (19) can be set up as

$$\Gamma_1 : G(\lambda, \tau_1, \cdots, \tau_{13}) = G(0) = u_1(0)u_2(0) = F(\tau_1, \tau_2, \cdots, \tau_{13}). \tag{19}$$

Similarly the initial condition of Eq. (10) can be set up as

$$\Gamma_2 : \begin{cases} z_i(0) = y_i(0) = \Phi_i(\tau_1, \tau_2, \cdots, \tau_{13}), & i = 1,2,\cdots,14, \\ z_{15}(0) = H(0) = 0. \end{cases} \tag{20}$$

We are now ready to consider the solution of Eq. (9) with initial condition (19) and the solution of Eq. (10) with initial condition (20).

A. Solution of $F_2^{(0)}$

Let D_1 and D_2 be two compact subdomains of R^{12} and $C^1(R^7) \times C^1(R^7)$ respectively. Denote $\alpha = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2)$, $\alpha_0 = (\mathbf{r}_1(0), \mathbf{r}_2(0), \mathbf{p}_1(0), \mathbf{p}_2(0))$, $u = (u_1, u_2)$, $u_0 = (u_1(0), u_2(0))$, and $A_1 = (\mathbf{p}_1, \mathbf{p}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2)$. Define the norms $|\alpha| = \Sigma_{i=1}^2 (|\mathbf{r}_i| + |\mathbf{p}_i|)$, $|u| = |u_1| + |u_2|$, and $|A_1| = \Sigma_{i=1}^2 [|\mathbf{p}_i| + |\boldsymbol{\phi}_i|]$. Let $\alpha, \alpha', \alpha_0$ be in D_1 and u, u', u_0 be in D_2. The following lemma is easy to prove.

Lemma 1: Suppose $\boldsymbol{\phi}_i$ satisfies the following condition

$$|\boldsymbol{\phi}_i(\mathbf{r}_1, \mathbf{r}_2) - \boldsymbol{\phi}_1(\mathbf{r}_1', \mathbf{r}_2')| \leqslant m_1 |\mathbf{r}_1 - \mathbf{r}_1'| + m_2 |\mathbf{r}_2 - \mathbf{r}_2'|, \quad i = 1,2, \tag{21}$$

where m_1 and m_2 are positive constants. Then A_1 satisfies a Lipschitz condition in D_1. Consequently, there exists a unique solution in D_1 for the vector differential Eq. (13), $d\alpha/d\lambda = A_1$ for $\lambda \geqslant 0$, which satisfies the initial condition α_0.

The solutions of Eq. (13) can be written as

$$\alpha_i = y_i = f_i(\lambda, \tau_1, \tau_2, \cdots, \tau_{11}), \quad i = 1,2,\cdots,12. \tag{22}$$

Substituting the results of Eq. (22) into Eq. (14) yields

$$\frac{du_i}{d\lambda} = g_i(\lambda, \tau_1, \cdots, \tau_{11}, u_i), \quad i = 1,2. \tag{23}$$

Similar to Lemma 1, we now have

Lemma 2: Suppose g_i satisfies the following Lipschitz condition in D_2,

$$|g_i(\lambda, u_i) - g_i(\lambda, u_i')| \leqslant l_i |u_i - u_i'|, \quad i = 1,2, \tag{24}$$

where l_1 and l_2 are positive constants. Then there exists in D_2 a unique solution for the differential Eq. (23) for $\lambda \geqslant 0$, which satisfies the initial condition $u_i(0)$.

The solution of Eq. (14) can be written as

$$u_1 = f_{13}(\lambda, \tau_1, \tau_2, \cdots, \tau_{11}, \tau_{12}),$$
$$u_2 = f_{14}(\lambda, \tau_1, \tau_2, \cdots, \tau_{11}, \tau_{13}).$$

If the Jacobian

$$J = \frac{\partial(f_1, f_2, \cdots f_{14})}{\partial(\lambda, \tau_1, \tau_2, \cdots \tau_{13})} \neq 0$$

on Γ_1, then $\lambda, \tau_1, \tau_2, \cdots, \tau_{13}$ can be inverted in the neighborhood of Γ_1 as functions of $\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2, u_1$ and u_2,

$$\lambda = \tilde{f}_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2, u_1, u_2),$$
$$\tau_i = \tilde{f}_{i+1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2, u_1, u_2), \quad i = 1,2,\cdots,13.$$

Since

$$\begin{aligned} d\tau_i = 0 &= \frac{\partial \tilde{f}_{i+1}}{\partial \mathbf{r}_1} dr_1 + \frac{\partial \tilde{f}_{i+1}}{\partial \mathbf{r}_2} \cdot dr_2 + \frac{\partial \tilde{f}_{i+1}}{\partial \mathbf{p}_1} \cdot d\mathbf{p}_1 \\ &\quad + \frac{\partial \tilde{f}_{i+1}}{\partial \mathbf{p}_2} \cdot d\mathbf{p}_2 + \frac{\partial \tilde{f}_{i+1}}{\partial u_1} du_1 + \frac{\partial \tilde{f}_{i+1}}{\partial u_2} du_2 \\ &= \Big(\mathbf{p}_1 \cdot \frac{\partial}{\partial \mathbf{r}_1} + \mathbf{p}_2 \cdot \frac{\partial}{\partial \mathbf{r}_2} + \boldsymbol{\phi}_1 \cdot \frac{\partial}{\partial \mathbf{p}_1} + \boldsymbol{\phi}_2 \cdot \frac{\partial}{\partial \mathbf{p}_2} \\ &\quad + g_1 \frac{\partial}{\partial u_1} + g_2 \frac{\partial}{\partial u_2} \Big) \tilde{f}_{i+1} \, d\lambda, \end{aligned}$$

thus $\tilde{f}_{i+1}$ is a solution of Eq. (9), and

$F(\tau_1, \tau_2, \cdots, \tau_{13}) = u_1(0)\; u_2(0) = F(\tilde{f}_2, \tilde{f}_3, \cdots \tilde{f}_{14})$ is a local solution of Eq. (9) with initial condition (19).

Theorem 1: Assume that conditions (21) and (24) are satisfied, and $J \neq 0$ on Γ_1. Then in the compact subdomain $\bar{D} = D_1 \times D_2$ of D, there exists a unique solution for Eq. (9)

in the neighborhood of Γ_1, which satisfies the initial condition (19).

B. Solution of $F_2^{(n)}$

By the solutions of α_i and u_i, ψ then becomes a function of λ and the parameters $\tau_1, \tau_2, \cdots, \tau_{13}$. The differential equation $dH/d\lambda = \psi$ with the initial condition $H(0) = 0$ amounts to a simple integration. Hence, $H = f_{15}(\lambda, \tau_1, \tau_2, \cdots, \tau_{13})$ can be uniquely determined. Consequently, by Lemmas 1 and 2 there exists a unique solution for Eq. (16) in the compact subdomain $\overline{D}_1$ which satisfies the initial condition (20). Again, if the Jacobian $J \neq 0$ on Γ_2, then $\lambda, \tau_1, \tau_2, \cdots, \tau_{13}$ can be inverted in the neighborhood of Γ_2, which can be substituted into the solution $H = f_{15}(\lambda, \tau_1, \cdots, \tau_{13})$ to yield a unique solution for Eq. (10).

Corollary: Assume that conditions (21) and (24) are satisfied, and $J \neq 0$ on Γ_2. Then in the compact subdomain $\overline{D}$ of D, there exists a unique solution for Eq. (10) in the neighborhood of Γ_2, which satisfies the initial condition (20).

We next consider two special solutions of $F_2^{(0)}$.

(I) Hard sphere potential

For the hard sphere potential defined by

$$v(r) = \begin{cases} \infty, & \text{if } r < \sigma, \\ 0, & \text{if } r \geqslant \sigma, \end{cases}$$

we can easily obtain $F_2^{(0)} = G = u_1 u_2$ from Eq. (9).

(II) Solution of Eq. (9) by the method of separation of variables.

By virtue of Eqs. (13) and (14), we try a solution of $F_2^{(0)}$ of the following form

$$G = \xi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2)\, \gamma(u_1, u_2), \tag{25}$$

with initial conditions $\xi(0) = 1$ and $\gamma(0) = u_1(0) u_2(0)$. Equation (9) then becomes

$$\frac{1}{\xi}\left(\mathbf{p}_1 \cdot \frac{\partial \xi}{\partial \mathbf{r}_1} + \mathbf{p}_2 \cdot \frac{\partial \xi}{\partial \mathbf{r}_2} + \boldsymbol{\phi}_1 \cdot \frac{\partial \xi}{\partial \mathbf{p}_1} + \boldsymbol{\phi}_2 \cdot \frac{\partial \xi}{\partial \mathbf{p}_2}\right)$$
$$= -\frac{1}{\gamma}\left(\boldsymbol{\phi}_1 \cdot \mathbf{b}(\mathbf{r}_1, \mathbf{p}_1, u_1) \frac{\partial \gamma}{\partial u_1} + \boldsymbol{\phi}_2 \cdot \mathbf{b}(\mathbf{r}_2, \mathbf{p}_2, u_2) \frac{\partial \gamma}{\partial u_2}\right), \tag{26}$$

which can be solved by the method of separation of variables if and only if

$$\mathbf{b}(\mathbf{r}_i, \mathbf{p}_i, u_i) = \mathbf{q}(\mathbf{r}_i, \mathbf{p}_i), h(u_i)$$

and

$$h(u_1) \frac{\partial \gamma}{\partial u_1} = h(u_2) \frac{\partial \gamma}{\partial u_2} = c\,\gamma, \tag{27}$$

where c is a constant. It then follows from Eq. (27) that $\gamma(u_1, u_2) = u_1 u_2$ is the only solution which satisfies the initial condition $\gamma(0) = u_1(0) u_2(0)$. The solution of G can therefore be written as

$$G = \xi u_1 u_2, \tag{28}$$

where ξ satisfies the following equation

$$\left(\mathbf{p}_1 \cdot \frac{\partial}{\partial \mathbf{r}_1} + \mathbf{p}_2 \cdot \frac{\partial}{\partial \mathbf{r}_2} + \boldsymbol{\phi}_1 \cdot \frac{\partial}{\partial \mathbf{p}_1} + \boldsymbol{\phi}_2 \cdot \frac{\partial}{\partial \mathbf{p}_2}\right) \ln(\xi)$$
$$= -\left[\boldsymbol{\phi}_1 \cdot \mathbf{q}(\mathbf{r}_1, \mathbf{p}_1) + \boldsymbol{\phi}_2 \cdot \mathbf{q}(\mathbf{r}_2, \mathbf{p}_2)\right], \tag{29}$$

with initial condition $\xi(0) = 1$.

Although $\mathbf{q}$ is still unknown (which will be determined by the kinetic equation); however, ξ at least in principle can be solved in terms of $\mathbf{q}$. Since

$$\frac{d}{d\lambda} \ln(\xi) = -\left[\boldsymbol{\phi}_1 \cdot \mathbf{q}(\mathbf{r}_1, \mathbf{p}_1) + \boldsymbol{\phi}_2 \cdot \mathbf{q}(\mathbf{r}_2, \mathbf{p}_2)\right],$$

by Theorem 1 there exists a unique solution for ξ, which can be written as

$$\xi = \exp\left(-\sum_{i=1}^{2} \int_0^{\lambda} \boldsymbol{\phi}_i(s, \tau_1, \cdots, \tau_{11}) \cdot \mathbf{q}(s, \tau_1, \cdots \tau_{11})\, ds\right).$$

The special solution of $G = F_2^{(0)} = \xi u_1 u_2$ will be useful in the discussion of kinetic equations.

In the next section we consider the solutions of $F_2^{(0)}$ and $F_2^{(n)}$ at equilibrium which can be shown to be identical to the classical results.

IV. EQUILIBRIUM SOLUTION

Let $u_1 = F_1(\mathbf{r}_1, \mathbf{p}_1, t)$ be in an equilibrium state, i.e., $u_1 = n(\mathbf{r}_1)(\beta/2\pi)^{3/2} \exp(-\beta p_1^2/2)$, where $\beta = 1/KT$, K is the Boltzmann constant, T is the temperature, and $n(\mathbf{r}_1)$ satisfies an equation to be specified by the kinetic equation. Then $\partial u_1/\partial \mathbf{p}_1 = -\beta \mathbf{p}_1 u_1$, and the characteristic equations (11) becomes

$$\frac{dr_{1i}}{p_{1i}} = \frac{dr_{2j}}{p_{2j}} = dp_{1i} \Big/ -\frac{\partial \phi}{\partial r_{1i}} = dp_{2j} \Big/ \frac{-\partial \phi}{\partial r_{2j}}$$
$$= d\ln(u_1) \Big/ -\beta \sum_{i=1}^{3} \frac{\partial \phi}{\partial r_{1i}} p_{1i}$$
$$= d\ln(u_2) \Big/ -\beta \sum_{i=1}^{3} \frac{\partial \phi}{\partial r_{2i}} p_{2i}. \tag{30}$$

From Eq. (30) we can construct a unique solution satisfying the initial condition (19) as

$$F_{2\,\mathrm{eq}}^{(0)} = e^{-\beta\phi} (\beta/2\pi)^{3/2} n(\mathbf{r}_1) n(\mathbf{r}_2) e^{-\beta(p_1^2 + p_2^2)/2}. \tag{31}$$

The results we have just obtained can easily be generalized to $F_{k\,\mathrm{eq}}^{(0)}$ for $k \geqslant 2$. By Eq. (8), ψ finally reduces to

$$\psi = F_{2\,\mathrm{eq}}^{(0)} \left(\mathbf{p}_1 \cdot \frac{\partial}{\partial \mathbf{r}_1} + \mathbf{p}_2 \cdot \frac{\partial}{\partial \mathbf{r}_2}\right) \int_E d\mathbf{r}_3\, n(\mathbf{r}_3) f_{13} f_{23}, \tag{32}$$

where E denotes $|\mathbf{r}_1 - \mathbf{r}_3| \geqslant \sigma$, $|\mathbf{r}_2 - \mathbf{r}_3| \geqslant \sigma$, and f_{13}, and f_{23} are the Mayer functions. Equation (10) then becomes

$$\mathscr{L} F_{2\,\mathrm{eq}}^{(1)} = \sum_{i=1}^{2} \left(\mathbf{p}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} + \boldsymbol{\phi}_i \frac{\partial}{\partial \mathbf{p}_i} + \beta\, \mathbf{p}_i \cdot \boldsymbol{\phi}_i u_i \frac{\partial}{\partial u_i}\right)$$
$$\times F_{2\,\mathrm{eq}}^{(1)} = \psi, \tag{33}$$

with initial condition (20).

Let $F_{2\,\mathrm{eq}}^{(1)} = F_{2\,\mathrm{eq}}^{(0)}\, \eta$. Then

$$\mathscr{L}\, \eta = \left(\mathbf{p}_1 \cdot \frac{\partial}{\partial \mathbf{r}_1} + \mathbf{p}_2 \cdot \frac{\partial}{\partial \mathbf{r}_2}\right) \int_E d\mathbf{r}_3\, n(\mathbf{r}_3) f_{13} f_{23}$$

and $\eta(0) = 0$.

Consequently,

$$\eta = \int_E d\mathbf{r}_3\, n(\mathbf{r}_3) f_{13} f_{23}.$$

When u_1 is a thermodynamically stable equilibrium state, i.e., $u_1 = (\beta/2\pi)^{3/2} e^{-(\beta p_1^2/2)}$, then $F^{(0)}_{2\,\mathrm{eq}} = e^{-\beta\phi} u_1 u_2$ and $F^{(1)}_{2\,\mathrm{eq}} = F^{(0)}_{2\,\mathrm{eq}} \int_E d\mathbf{r}_3 f_{13} f_{23}$. Let $f_0 = e^{-\beta\phi}$ and $f_1 = e^{-\beta\phi} \int_E d\mathbf{r}_3 f_{13} f_{23}$. Then f_0 and f_1 are the well known classical results in the virial expansion of the radial distribution function. With the solutions of $F^{(0)}_2$ and $F^{(1)}_2$ we can successively solve $F^{(n)}_2$ from Eqs. (6)–(8), and show that $F^{(n)}_{2\,\mathrm{eq}}$ is identical to the result in the classical virial expansion. Since a detailed proof is very complicated and lengthy, we shall just summarize the result in the following.

Theorem 2: If u_1 is in a thermodynamically stable equilibrium state, then $F^{(n)}_{2\,\mathrm{eq}}$ becomes the nth term in the virial expansion of the equilibrium two-particle distribution function.

V. DISCUSSION OF THE KINETIC EQUATION

To the order of ρ Eq. (14) reduces to

$$\frac{\partial F_1}{\partial t} + \mathbf{p}_1\cdot\frac{\partial F_1}{\partial \mathbf{r}_1} = \rho\int d\mathbf{p}_2 \oint_{|r_1 - r_2| = \sigma} d\,\mathbf{S}\cdot(\mathbf{p}_2 - \mathbf{p}_1) F^{(0)}_2 + \rho\int_{D_2} dx_2 \frac{\partial\phi}{\partial \mathbf{r}_1}\cdot\frac{\partial F^{(0)}_2}{\partial \mathbf{p}_1}. \tag{34}$$

Following Grad, we firstly impose the assumption of the binary collision. Secondly we assume that F_1 is a spatially slowly varying function over an interval of length σ. Equation (34) then becomes

$$\begin{aligned}\frac{\partial F_1}{\partial t} + \mathbf{p}_1\cdot\frac{\partial F_1}{\partial \mathbf{r}_1} &= \rho\sigma^2 \int d\mathbf{p}_2 \int_{\hat{\Omega}\cdot\mathbf{p}>0} d\hat{\Omega}\,\hat{\Omega}\cdot\mathbf{p} \\ &\times [F^{(0)}_2(\mathbf{r}_1,\mathbf{r}_1 + \sigma\hat{\Omega},\mathbf{p}_1,\mathbf{p}_2,F_1(1'),F_1(2')) \\ &- F^{(0)}_2(\mathbf{r}_1,\mathbf{r}_1 - \sigma\hat{\Omega},\mathbf{p}_1,\mathbf{p}_2,F_1(1),F_1(2)] \\ &+ \rho\int_{D_2} dx_2 \frac{\partial\phi}{\partial \mathbf{r}_1}\cdot\frac{F^{(0)}_2}{\partial \mathbf{p}_1},\end{aligned} \tag{35}$$

where $F_1(i') = F_1(\mathbf{r}_1,\mathbf{p}'_i,t)$, $F_1(i) = F_1(\mathbf{r}_1,\mathbf{p}_i,t)$, $i = 1,2$, $\mathbf{p} = \mathbf{p}_2 - \mathbf{p}_1$, $\hat{\Omega}$ is a unit radial vector, and the binary collision is denoted by $\mathbf{p}_1 + \mathbf{p}_1 \to \mathbf{p}'_1 + \mathbf{p}'_2$ with $\mathbf{p}'_i = \mathbf{p}_i + \hat{\Omega}(\hat{\Omega}\cdot\mathbf{p})$.

In case of the hard sphere potential, Eq. (35) reduces to the ordinary Boltzmann equation

$$\begin{aligned}\frac{\partial F_1}{\partial t} + \mathbf{p}_1\cdot\frac{\partial F_1}{\partial \mathbf{r}_1} &= \rho\sigma^2 \int d\mathbf{p}_2 \int_{\hat{\Omega}\cdot\mathbf{p}>0} d\hat{\Omega}\hat{\Omega}\cdot\mathbf{p}\,[F_1(1')F_1(2') - F_1(1)F_1(2)] \\ &= J(f,f).\end{aligned}$$

On the other hand, in terms of the special solution $F^{(0)}_2 = \xi u_1 u_2$, Eq. (35) becomes

$$\begin{aligned}\frac{\partial F_1}{\partial t} + \mathbf{p}_1\cdot\frac{\partial F_1}{\partial \mathbf{r}_1} &= \rho\sigma^2 \int d\mathbf{p}_2 \int_{\hat{\Omega}\cdot\mathbf{p}>0} d\hat{\Omega}\hat{\Omega}\cdot\mathbf{p}\,[\,\xi(\mathbf{r}_1,\mathbf{r}_1 + \sigma\hat{\Omega},\mathbf{p}'_1,\mathbf{p}'_2) \\ &\times F_1(1')F_1(2') - \xi(\mathbf{r}_1,\mathbf{r}_1 - \sigma\hat{\Omega},\mathbf{p}_1,\mathbf{p}_2) F_1(1)F_1(2)] \\ &+ \rho\int_{D_2} dx_2 \frac{\partial\phi}{\partial \mathbf{r}_1}\cdot\frac{\partial}{\partial \mathbf{p}_1}[\xi F_1(1)F_1(2)].\end{aligned} \tag{36}$$

It can easily be shown that the Maxwell–Boltzmann distribution

$$F_1(1) = n(\mathbf{r}_1)(\beta/2\pi)^{3/2} e^{-\beta p_1^2/2}$$

is a solution of Eq. (36) provided $n(r_1)$ satisfies the following equation

$$\ln[n(\mathbf{r}_1)] - \rho\int_{D_2} d\mathbf{r}_2\, e^{-\beta\phi} n(\mathbf{r}_2) = \text{const}. \tag{37}$$

Conversely, by setting

$$\begin{aligned}\frac{\partial F_1}{\partial t} = 0,\ &\xi(\mathbf{r}_1,\mathbf{r}_1 + \sigma\hat{\Omega},\mathbf{p}'_1,\mathbf{p}'_2) F_1(1') F_1(2') \\ &= \xi(\mathbf{r}_1,\mathbf{r}_1 - \sigma\hat{\Omega},\mathbf{p}_1,\mathbf{p}_2) F_1(1) F_1(2),\end{aligned} \tag{38}$$

we can recover the Maxwell-Boltzmann distribution.[3]

An equilibrium state can therefore be defined as the solution of Eqs. (37) and (38). Particularly, the thermodynamically stable equilibrium state satisfies Eqs. (36), (37), and (38). In this sense, Eq. (36) can be considered as a generalized kinetic equation. Unfortunately, so far we have not been able to prove the H theorem from Eq. (36).

Recently Braun and Flores[4] had proposed a convergent kinetic theory; however, their method is equivalent to approximating ξ by $1 - g_2(\mathbf{r}_1,\mathbf{r}_2)$ where g_2 is the local equilibrium two-particle correlation function. At any rate, suppose ξ is only a function of $\mathbf{r}_1$ and $\mathbf{r}_2$. Equation (36) then becomes

$$\frac{\partial F_1}{\partial t} + \mathbf{p}_1\cdot\frac{\partial F_1}{\partial \mathbf{r}_1} = \rho\sigma^2\xi(\mathbf{r}_1,\sigma) J(f,f) + \rho\frac{\partial F_1(1)}{\partial \mathbf{p}_1}\cdot\int_{D_2} dx_2 \frac{\partial\phi}{\partial \mathbf{r}_1} F_1(2)\xi. \tag{39}$$

Let $\bar{R}$ be a region such that $F_1(1) \to 0$ as $|\mathbf{p}_1| \to \infty$ on the boundary Σ of $\bar{R}$ and $\oint_\Sigma \mathscr{H}\cdot d\mathbf{s} = 0$, where $\mathscr{H} = \int \mathbf{p}_1 F_1(1) \times \ln[F_1(1)]\, d\mathbf{p}_1$. Multiplying through Eq. (39) by $1 + \ln[F_1(1)]$ and then integrating with respect to $\mathbf{p}_1$, yields the following result

$$\frac{\partial \mathscr{H}}{\partial t} + \nabla\cdot\mathscr{H} = -\xi(r_1,\sigma)\bar{\alpha}, \tag{40}$$

where

$$\mathscr{H} = \int F_1(1)\ln[F_1(1)]\,d\mathbf{p}_1,$$

$$\begin{aligned}\bar{\alpha} = \rho\sigma^2 \int d\mathbf{p}_1 \int d\mathbf{p}_2 \int_{\hat{\Omega}\cdot\mathbf{p}>0} d\hat{\Omega}\,\hat{\Omega}\cdot\mathbf{p}\,&[F_1(1')F_1(2') \\ &- F_1(1)F_1(2)]\ln\left(\frac{F_1(1')F_1(2')}{F_1(1)F_1(2)}\right).\end{aligned}$$

Since $\bar{\alpha} \geqslant 0$, and by Eq. (29), ξ is always a positive function, we can integrate Eq. (40) over the region $\bar{R}$ and prove the H theorem.

[1]N.N. Bogoliubov, "Problems In Dynamical Theory in Statistical Mechanics," translated by E.K. Gora in *Studies In Statistical Mechanics*, edited by J. de Boer and G.E. Uhlenbeck (North Holland, Amsterdam, 1962), Vol. 1.

[2]H. Grad, "Principles Of The Kinetic Theory of Gases," *Handbuch der Physik*, edited by S. Flugge (Springer, New York, 1958), pp. 205–94, Bd. 12.

[3]Since Eq. (38) is true for all $\mathbf{r}_1$, we can set $\mathbf{r}_1 = \mathbf{0}$. Thus

$$\xi(\sigma\widehat{\Omega},\mathbf{p}_1',\mathbf{p}_2')\,F_1(\mathbf{p}_1')\,F_1(\mathbf{p}_2') = \xi(-\sigma\widehat{\Omega},\mathbf{p}_1,\mathbf{p}_2)\,F_1(\mathbf{p}_1)\,F_1(\mathbf{p}_2).$$

Consequently ξ must be an even function of σ, which implies that

$$\xi(\sigma\mathbf{\Omega},\mathbf{p}_1',\mathbf{p}_2') = \xi(\sigma\mathbf{\Omega},\mathbf{p}_1,\mathbf{p}_2)$$

and

$$F_1(\mathbf{p}_1',t)\,F_1(\mathbf{p}_2',t) = F_1(\mathbf{p}_1,t)\,F_1(\mathbf{p}_2,t).$$

It then follows that $F_1(i) = cn(\mathbf{r}_1)\,e^{-\beta(\mathbf{r}_1)p_i^2/2}$, where c is a constant. By the characteristic equation (30) we can obtain

$$d\ln(u_1u_2) = -\sum_{i=1}^{3}\left(\frac{\partial\phi}{\partial\mathbf{r}_{1i}}\beta(\mathbf{r}_1)dr_{1i} + \frac{\partial\phi}{\partial\mathbf{r}_{2i}}\beta(\mathbf{r}_2)dr_{2i}\right)$$
$$= \beta(\mathbf{r}_1)\sum_{i=1}^{3}p_{1i}dp_{1i} + \beta(\mathbf{r}_2)\sum_{i=1}^{3}p_{2i}dp_{2i}.$$

But this is a total differential if and only if $\beta(\mathbf{r}_1) = \beta(\mathbf{r}_2) = \text{const}$. Let $\beta(\mathbf{r}_i) = -\beta$. Then $F_1(i) = n(\mathbf{r}_i)(\beta/2\pi)^{3/2}\,e^{-\beta p_i^2/2}$ and $\xi = e^{-\beta\phi}$.

[4]E. Braun and A. Flores, J. Stat. Phys. **8**, 155 (1973).

Analyticity in the coupling constant of the $\lambda\ P(\phi)$ lattice theory [a)]

Florin Constantinescu

Institut für angewandte Mathematik der Johann Wolfgang Goethe-Universität Frankfurt am Main, Robert-Mayer Strasse 10, West Germany

(Received 28 February 1979; accepted for publication 17 July 1979)

Consider the lattice $\lambda\ P(\phi)$ theory in an arbitrary number of dimensions. The single spin distribution is given by $\lambda\ P(\phi) + (m^2/2)\phi^2$. We show that if $\theta < n + 1/2\pi$, then there is a constant $\epsilon > 0$ such that the correlations are analytic in $|\lambda| < \epsilon, |\arg\lambda| < \theta$ provided m is sufficiently large. This result implies Borel summability of the lattice (Euclidean) $\lambda\ P(\phi)$ quantum field theory for small coupling constants.

1. INTRODUCTION

We study the lattice theory with single spin distributions given by $\lambda P(\phi) + (m^2/2)\phi^2$ in arbitrary number of dimensions. We prove two results. The first one refers to the particular case $P(\phi) = \phi^{2n}, n \geqslant 2$. The correlation functions in this case are shown to be analytic in the entire λ-plane provided m is sufficiently large. Our second result refers to the general case in which $P(\phi)$ is only a semibounded polynomial $P(\phi) = \Sigma_{l=2}^{n} a_l \phi^{2l}$. For $\theta < [(n+1)/2]\pi$ we prove that there is a constant ϵ (possible θ-dependent) such that the correlations are analytic in $|\lambda| < \epsilon, |\arg| < \theta$ provided m is sufficiently large. This analytic structure is similar to the analytic structure of the eigenvalues for the x^{2n} quantum mechanical anharmonic oscillator.[1] This result implies Borel summability of the correlation functions of the $P(\phi)$ lattice model. The factorial bound needed in the Borel–Leroy transform follows by results in Ref. 2 or more directly in Ref. 3. For proving this result we use the Mayer (cluster) expansion and a scaling argument in ϕ. We recall that in the continuum case analyticity of correlations in the restricted domain $|\lambda| < \epsilon, |\arg\lambda| < \pi/2 + \eta$ with ϵ and $\eta > 0$ small and m large is known.[2] Recently[4] detailed information on the partition function of the $P(\phi) = \phi^4$ lattice model as function of the coupling constant λ was obtained.

2. FORMULATION OF THE PROBLEM

Suppose $A(\phi) = \phi_{i_1}^{k_1} \ldots \phi_{i_r}^{k_r}$ is a monomial of total degree $k_1 + \ldots + k_r = k$ localized in the cube $\Lambda \subset \mathbb{R}^\nu$ where $\partial\Lambda$ lies on the dual lattice $(\mathbb{Z}^\nu)^*$. The finite volume expectations (correlations) of the model are defined by

$$\langle A \rangle_\Lambda = \frac{1}{Z(\Lambda)} \int A(\phi) \exp\Big(\sum_{\overline{ij} \in \Lambda} \phi_i \phi_j\Big) \prod_{i \in \Lambda} d\mu_i^0, \quad (1)$$

where $\overline{ij}$ is a bound on $\mathbb{Z}^\nu$ and

$$d\mu_i^0 = \frac{\exp[-\lambda P(\phi_i) - (m^2/2)\phi_i^2]\, d\phi_i}{\int \exp[-\lambda P(\phi_i) - (m_2/2)\phi_i^2]\, d\phi_i}, \quad (2)$$

$$Z(\Lambda) = \int \exp\Big(\sum_{\overline{ij} \in \Lambda} \phi_i \phi_j\Big) \prod_{i \in \Lambda} d\mu_i^0. \quad (3)$$

In (1) we substitute $\phi_i = \alpha^{1/2}\varphi_i$ with $\alpha > 0$. We get

$$\langle A \rangle_\Lambda = \alpha^{k/2} \frac{\int A(\varphi) \exp(\alpha \Sigma_{\overline{ij} \in \Lambda} \varphi_i \varphi_j) \Pi_{i \in \Lambda} \exp[-\lambda \alpha^n \widetilde{P}(\varphi_i, \alpha) - (\alpha m^2/2)\varphi_i^2]\, d\varphi_i}{\int \exp(\alpha \Sigma_{\overline{ij} \in \Lambda} \varphi_i \varphi_j) \Pi_{i \in \Lambda} \exp[-\lambda \alpha^n \widetilde{P}(\varphi_i, \alpha) - (\alpha m^2/2)\varphi_i^2]\, d\varphi_i}, \quad (4)$$

where

$$\widetilde{P}(\varphi, \alpha) = \sum_{l=2}^{n} \frac{a_l}{\alpha^{n-l}} \varphi^{2l}.$$

Let us denote[2]

$$f_\Lambda(\lambda) = \langle A \rangle_\Lambda, \quad g_\Lambda(\lambda, \alpha) = \frac{\int A(\varphi) \exp(\alpha \Sigma_{\overline{ij} \in \Lambda} \varphi_i \varphi_j) \Pi_{i \in \Lambda} \exp[-\lambda \alpha^n \widetilde{P}(\varphi_i, \alpha) - (\alpha m^2/2)\varphi_i^2]\, d\varphi_i}{\int \exp(\alpha \Sigma_{\overline{ij} \in \Lambda} \varphi_i \varphi_j) \Pi_{i \in \Lambda} \exp[-\lambda \alpha^n \widetilde{P}(\varphi_i, \alpha) - (\alpha m^2/2)\varphi_i^2]\, d\varphi_i}.$$

For $\lambda > 0, \alpha > 0$ we have

$$f_\Lambda(\lambda) = \alpha^{k/2} g_\Lambda(\lambda, \alpha). \quad (5)$$

The function $g_\Lambda(\lambda, \alpha) = \alpha^{-k/2} f_\Lambda(\lambda)$ can be analytically extended to $D = \{(\lambda, \alpha); \mathrm{Re}\,\alpha > \delta,\ 1 - \delta < |\alpha| < 1 + \delta,\ \mathrm{Re}\,\lambda\alpha^n > 0, |\lambda| < \epsilon\}$ for ϵ, δ sufficiently small. Indeed

$$\int \exp\Big(\alpha \sum_{\overline{ij} \in \Lambda} \varphi_i \varphi_j\Big) \prod_{i \in \Lambda} \exp[-(\alpha m^2/2)\varphi_i^2]\, d\varphi_i = \alpha^{-|\Lambda|/2} \int \exp\Big(\sum_{\overline{ij}} \phi_i \phi_j\Big) \prod_{i \in \Lambda} \exp[-(m^2/2)\phi_i^2]\, d\phi_i \neq 0.$$

By continuity arguments

[a)]Supported in part by the deutsche Forschungsgemeinschaft.

$$\int \exp\Big(\alpha \sum_{\overline{ij}\in\Lambda} \varphi_i\varphi_j\Big)\prod_{i\in\Lambda} \exp[-\lambda\alpha^n\widetilde{P}(\varphi_i,\alpha)-(\alpha m^2/2)\varphi_i^2]\,d\varphi_i \neq 0 \tag{6}$$

because $|\lambda\alpha^n| < \epsilon(1+\delta)^n$ is small.

Suppose we are able to prove by means of the Mayer (cluster) expansion that for m large the thermodynamic limit $g(\lambda,\alpha)$ of $g_\Lambda(\lambda,\alpha)$ exists for complex values of λ and α in D. In particular the thermodynamic limit $f(\lambda)$ of $f_\Lambda(\lambda)$ exists and is analytic in $\mathrm{Re}\lambda > 0$, $|\lambda| < \epsilon$. For real α and λ in D we have

$$f(\lambda) = \alpha^{k/2} g(\lambda,\alpha).$$

With α real we can analytically continue (6) in λ to $\mathrm{Re}\lambda > 0$. Now fix λ with $\mathrm{Re}\lambda > 0$ and analytically continue the right-hand side of (6) in α to $\mathrm{Re}\alpha > \delta$, $1-\delta < |\alpha| < 1+\delta$ such that $\mathrm{Re}\lambda\alpha^n > 0$. A last analytic continuation of $g(\lambda,\alpha)$ in λ to $\mathrm{Re}\lambda\alpha^n > 0$ brings us to the optimal analyticity domain of $f(\lambda)$: $|\arg\lambda| < \theta < (n+1)\pi/2, |\lambda| < \epsilon$ where $\theta = (n+1)\pi/2 - n\delta$. In fact we will prove in Sec. (6) that in the particular case $P(\phi) = \phi^{2n}, n \geqslant 2, f(\lambda)$ is analytic in the entire plane $|\arg\lambda| < \theta < (n+1)\pi/2$ (for m sufficiently large).

3. THE MAYER (CLUSTER) EXPANSION

We study the thermodynamic limit of

$$h_\Lambda(\mu,\alpha) \equiv g_\Lambda(\lambda\alpha^n,\alpha) = \frac{\int A(\varphi)\exp(-\alpha\Sigma_{\overline{ij}\in\Lambda}\varphi_i\varphi_j)\Pi_{i\in\Lambda}\exp[-\mu\widetilde{P}(\varphi_i,\alpha)-(\alpha m^2/2)\varphi_i^2]\,d\varphi_i}{\int \exp(-\alpha\Sigma_{\overline{ij}\in\Lambda}\varphi_i\varphi_j)\Pi_{i\in\Lambda}\exp[-\mu\widetilde{P}(\varphi_i,\alpha)-(\alpha m^2/2)\varphi_i^2]\,d\varphi_i}, \tag{7}$$

where $\mu = \lambda\alpha^n$. We use the Mayer expansion in the explicit series form given in Ref. 5 (see also Refs. 6, 7, and 8). In (7) we take

$$(\mu,\alpha) \in D_1 = \{(\mu,\alpha): \mathrm{Re}\alpha > \delta, 1-\delta < |\alpha| < 1+\delta, \mathrm{Re}\mu > 0, |\mu| < \epsilon_1\},$$

where ϵ_1 is small. We also suppose that m is sufficiently large. Let us denote $m_1^2 = m^2 - 2\nu > 0$ where ν is the number of dimensions and take $m_1\varphi = q$. Then

$$h_\Lambda(\mu,\alpha) = m_1^{-k}\frac{\int A(q)\exp[(\alpha/2m_1^2)\Sigma_{\overline{ij}\in\Lambda}(q_i-q_j)^2]\Pi_{i\in\Lambda}\exp[-\mu Q(q_i,\alpha,m_1)-(\alpha/2)q_i^2]\,dq_i}{\int \exp[(\alpha/2m_1^2)\Sigma_{\overline{ij}\in\Lambda}(q_i-q_2)^2]\Pi_{i\in\Lambda}\exp[-\mu Q(q_i,\alpha,m_1)-(\alpha/2)q_i^2]\,dq_i},$$

where

$$Q(q,\alpha,m_1) = \sum_{l=2}^{n}\frac{a_l}{m_1^{2l}\alpha^{n-l}}q^{2l}.$$

Let us denote $\eta = \alpha/2m_1^2$. We have $|\eta| < (1+\delta)/2m_1^2$ such that $|\eta|$ is small if m^2, i.e., $m_1^2 = m^2 - 2\nu$ is large. In this section we denote

$$\langle A\rangle_\Lambda = \frac{\int A(q)\exp[\eta\Sigma_{\overline{ij}\in\Lambda}(q_i-q_2)^2]\Pi_{i\in\Lambda}\exp[-\mu Q(q_i,\alpha,m)-(\alpha/2)q_i^2]\,dq_i}{\int \exp[\eta\Sigma_{ij\in\Lambda}(q_i-q_j)^2]\Pi_{i\in\Lambda}\exp[-\mu Q(q_i,\alpha,m)-(\alpha/2)q_i^2]\,dq_i} \tag{8}$$

and by $\langle A\rangle$ the thermodynamic limit of $\langle A\rangle_\Lambda$ (if it exists!). Here $A(q) = q_{i_1}^{k_1}\dots q_{i_r}^{k_r}$, $k = k_1 + \dots + k_r$.
Let[5]

$$k_b \equiv k_{\overline{ij}} = \exp[\eta(q_i-q_j)^2]-1, \quad k_\Gamma = \prod_{b\in\Gamma}k_b,$$

where $b \equiv \overline{ij}$ is a bound on $\mathbb{Z}^\nu$ and $\Gamma = (b_1,\dots,b_k)$ is a collection of bonds b_i, $i = 1,\dots,l$ which are mutually different. Let us denote $\Omega = \times_{i\in\mathbb{Z}^\nu}\mathbb{R}_i$, $\Sigma = \times_{i\in\mathbb{Z}^\nu}\Sigma_i$ and $d\mu^0 = \times_{i\in\mathbb{Z}^\nu}d\mu_i^0$ where in this section

$$d\mu_i^0 = \frac{\exp[-\mu Q(q_i,\alpha,m_1)-(\alpha/2)q_i^2]\,dq_i}{\int\exp[-\mu Q(q_i,\alpha,m_1)-(\alpha/2)q_i^2]\,dq_i}$$

and Σ_i is the Borel τ-algebra of $\mathbb{R}_i$. Expectations with respect to $d\mu^0$ will be denoted by $\langle\cdot\rangle_0$. We use here the Mayer expansion in its explicit form[5]:

$$\langle A\rangle = \sum_\gamma a_\gamma, \tag{9}$$

$$a_\gamma = \langle Ak_{\Gamma_1(\gamma)}\rangle_0\langle -k_{\Gamma_2(\gamma)}\rangle_0\cdots\langle -k_{\Gamma_q(\gamma)}\rangle_0, \tag{10}$$

where we have to specify the summation variable γ and the sets $\Gamma_i \equiv \Gamma_i(\gamma)$. For this we introduce some graph theory notations. Let $\mathscr{V}$ be the collection of all finite nonempty sets in $\mathbb{Z}$. To each $V\in\mathscr{V}$ we associate a point $t_V\in V$ and organize $\{\phi\}\cup\mathscr{V}$ as a graph such that there is a line between V_1 and $V_2\in\mathscr{V}$ iff $V_1 = V_2 - t_{V_2}$ or vice versa. Then $\{\phi\}\cup\mathscr{V}$ is a connected tree with ϕ as lowest vertex.

We assume that the mapping $V \to t_V$ has the following properties:

(i) t_V lies on the boundary of V;

(ii) if $V = V_1 \cup ... \cup V_n$ is a partition of $V \in \mathscr{V}$ into its connected components, then $t_V = t_{V_i}$ for some i. We now specify γ and $\Gamma_i(\gamma)$ in (9) and (10). The summation in (9) goes over sets

$$\gamma = \{(V_1,\Gamma_1),...,(V_q,\Gamma_q)\}, \quad q \geqslant 1,$$

where $V_i \equiv V_i(\gamma) \in \mathscr{V}$, $\Gamma_i \equiv \Gamma_i(\gamma)$ are collections of bonds such that the following conditions are satisfied:

(a) For $q \geqslant p > 1$, V_p lies (in the graph sense) below $V_{p-1} \times \cup \widetilde{\Gamma}_{p-1}$.

(b) For $q \geqslant p > 1$, Γ_p is nonempty and connected and

$$t_V \in \widetilde{\Gamma}_p \subset \mathbb{Z}^\nu - (V_p - t_{V_p}).$$

(c) $V_1(\gamma) = V$ and $\Gamma_1(\gamma)$ may be empty or such that (V_1, Γ_1) is connected.

In (a)–(c) $\widetilde{\Gamma}$ means the set of lattice points contained in Γ. We state the result of this section as a theorem.

Theorem 1: The Mayer series (9) converges absolutely and uniformly in $(\mu,\alpha,\eta) \in D_2$ where $D_2 = (\mu,\alpha,\eta)$:Re$\mu > 0$, $\Lambda - \delta < |\alpha| < 1 + \delta$, $|\eta| < \delta_1 \}$ with $\epsilon_1, \delta, \delta_1 > 0$ sufficiently small.

Proof: Suppose we can prove that there is a constant $C > 0$ such that for all $\Gamma = \{b_1,...,b_k\}$ we have

$$|\langle k_\Gamma \rangle_0| \leqslant (C_1\delta_1)^{|\Gamma|} \tag{11}$$

uniformly in $(\mu,\alpha,\eta) \in D_2$ and similarly

$$|\langle A k_\Gamma \rangle_0| \leqslant (C_2\delta_1)^{|\Gamma|}, \tag{12}$$

where C_2 can depend on A but not on the parameters in D_2. From (10), (11), and (12) we get with a constant $C > 0$

$$|a_\gamma| \leqslant (C\delta_1)^{|\Gamma_1| + \cdots + |\Gamma_q|} \tag{13}$$

uniformly with respect to the parameters in D_2. In order to complete the proof observe that the combinatoric factor in (9) by holding $|\Gamma_1| + \cdots + |\Gamma_q|$ fixed does not exceed $4^{|\Lambda|} \times K^{|\Gamma_1| + \cdots + |\Gamma_q|}$ (see Refs. 5 and 6) where K is a constant.

In the next section we will prove the central estimates (11) and (12).

4. ESTIMATES

We prove (11) by adapting to the complex case a simple argument in Ref. 5. We have by integration along the segment $[0,\eta]$ in the complex η variable

$$|\langle k_\Gamma \rangle_0| = \left| \left\langle \prod_{\overline{ij} \in \Gamma} \int_0^\eta \exp[-\eta'(q_i - q_j)^2][-(q_i - q_j)^2]\, d\eta' \right\rangle_0 \right|$$
$$\leqslant |\delta_1|^{|\Gamma|} \frac{\int \Pi_{\overline{ij}\in\Gamma}(q_i - q_j)^2|\exp[-\mu Q(q_i,\alpha,m_1) - (\alpha/2)q_i^2]|\, dq_i}{|\int \Pi_{\overline{ij}\in\Gamma} \exp[-\mu Q(q_i,\alpha,m_1) - (\alpha/2)q_i^2]\, dq_i|}. \tag{14}$$

Let us study the integral quotient on the right-hand side of (14). It reduces to a sum of terms

$$\frac{\int q^m|\exp[-\mu Q(q,\alpha,m_1) - (\alpha/2)q^2]|\, dq}{|\int \exp[-\mu Q(q,\alpha,m_1) - (\alpha/2)q^2]\, dq|}. \tag{15}$$

For $\mu = 0$, (15) equals

$$\frac{\int q^m \exp[-(\mathrm{Re}\alpha/2)q^2]\, dq}{|\int \exp[-(\alpha/2)q_2]\, dq|} \tag{16}$$

which is uniformly bounded in α for Re$\alpha > \delta$, $1 - \delta < |\alpha| < 1 + \delta$. Taking into account that $|\mu|$ is small, $|\alpha| > 1 - \delta > 0$, and m_1 large it follows that (14) is also uniformly bounded in D_2. This together with (13) proves the estimate (11).

In order to prove (12) we write

$$|\langle Ak_\Gamma \rangle_0| = \left| \left\langle \prod_{\overline{ij}\in\Gamma} \int_0^\eta \exp\left[-\eta' \sum_{ij\in\Gamma} (q_i - q_j)^2\right] \right.\right.$$
$$\times A(q)[-(q_i - q_j)^2]\, d\eta' \rangle_0 |$$

and proceed as above. We use the fact that $A(q)$ is a monomial in q and get the estimate (12).

5. ANALYTICITY IN THE COUPLING CONSTANT

In this section we state our first result.

Theorem 2: Let $P(\phi)$ be a semibounded polynomial of degree $2n$ and consider the weakly coupled lattice $\lambda P(\phi) + (m^2/2)\phi^2$ (Euclidean) quantum field theory. If $\theta < [(n+1)/2]\pi$ then there is a constant $\epsilon > 0$ such that the (generalized) Schwinger functions are analytic in the coupling constant λ in the domain $|\lambda| < \epsilon$, $|\arg\lambda| < \theta$ if the mass m is sufficiently large.

Proof: The proof follows from the results stated in Secs. 2–4.

We remark that this result can be extended for the model with statistical sum

$$Z(\Lambda) = \int \exp\Big(\beta \sum_{\overline{ij}\in\Lambda} \phi_i\phi_j\Big) \prod_{i\in\Lambda} d\mu_i^0 \tag{17}$$

instead of (3) where $\beta = 1/kT$ is the inverse temperature. In this case we have to require the β/m^2 is sufficiently small. The analyticity domain obtained in Theorem 2 is similar to that obtained in the case of the one-dimensional quantum-mechanically anharmonic oscillator with potential x^{2n}. In the particular case $n = 2$ detailed information on analytic properties of the eigenvalues is available. There is a three-sheeted structure around $\lambda = 0$ and $\lambda = 0$ is a limit of square root points with asymptotic phase $\pm 3\pi/2$. On the first sheet there is no singularity. We do not have such detailed information on the analytic properties of our model but in the particular case $P(\phi) = \phi^{2n}$, $n \geqslant 2$ we have some results. This case is discussed in the next section.

6. THE CASE $P(\phi) = \phi^{2n}$, $n \geqslant 2$

Our second result is the following:

Theorem 3: Consider the lattice model $\lambda\phi^{2n} + (m^2/2)\phi^2$ where $n \geqslant 2$ and λ is small. If $\theta < [(n+1)/2]\pi$

and m is sufficiently large then the Schwinger functions of this model are analytic in the coupling constant λ in the domain $|\arg\lambda| < \theta$.

Proof: As in the case of Theorem 2, the proof of this theorem depends on the estimates (11) and (12) which in turn reduces to a study of the one-dimensional integrals in (14). We have to bound the quotient in (15) for all μ with $\mathrm{Re}\mu > 0$ and α such that $\mathrm{Re}\alpha > \delta$, $1 - \delta < |\alpha| < 1 + \delta$. We remark that the numerator in (15) is bounded for $\mathrm{Re}\mu \geqslant 0$, $|\mu| \leqslant R$ where R is a positive constant arbitrarily large. If we can prove that the denominator is strictly positive for $\mathrm{Re}\mu \geqslant 0$, $|\mu| \leqslant R$, $\mathrm{Re}\alpha \geqslant \delta$, $1 - \delta \leqslant |\alpha| \leqslant 1 + \delta$ it will follow that (15) is bounded for $\mathrm{Re}\mu \geqslant 0$ and $|\mu| \leqslant R$. For this we prove the following elementary result.

Lemma 4: Consider the integral

$$I(\lambda,\mu) = \int_0^\infty \exp(-\lambda t^{2n} - \mu t^2)\, dt,$$

where $\mathrm{Re}\lambda \geqslant 0$, $|\lambda| \leqslant R$, $\mathrm{Re}\mu \geqslant \delta$, $1 - \delta \leqslant |\mu| \leqslant 1 + \delta$ where $R > 0$ and $0 < \delta < 1$. Then for all λ, μ with these properties

$$|I(\lambda,\mu)| > C,$$

where C is a positive constant.

Proof: We write

$$\lambda = \lambda_1 + i\lambda_2, \quad \mu = \mu_1 + i\mu_2,$$

and

$$\mathrm{Re}I(\lambda,\mu) = \int_0^\infty \exp(-\lambda_1 t^{2n} - \mu_1 t^2) \times \cos(\lambda_2 t^4 + \mu_2 t^2)\, dt$$

and substitute $f(t) = \lambda_1 t^{2n} + \lambda_2 t^2 = u$. The function $f'(t) = 2n\lambda_1 t^{2n-1} + 2\lambda_2 t$ is increasing for $t \in (0, \infty)$. Let $t = \varphi(u)$ be the inverse function. Then $\varphi'_u = 1/f'_t$ is decreasing on $(0, \infty)$. We have

$$\mathrm{Re}I(\lambda,\mu) = \int_0^\infty \exp[-\lambda_1 \varphi^{2n}(u) - \mu_1 \varphi^2(u)] \times \varphi'_u \cos u\, du.$$

The function $\lambda_1 \varphi^{2n}(u) + \mu_1 \varphi^2(u)$ is strictly increasing on $(0, \infty)$ because $\varphi(u)$ is so. The product $\exp[-\lambda_1 \varphi^{2n}(u) - \mu_1 \varphi^2(u)]\varphi'_u$ is strictly decreasing in $u \in (0, \infty)$. It is clear that there is a positive constant C such that

$$\mathrm{Re}[I(\lambda,\mu)] > C.$$

Then $|I(\lambda,\mu)| \geqslant \mathrm{Re}[I(\lambda,\mu)] > C$ and the lemma is proved.

We have now to study the behavior of (15) for $|\mu|$ large, $\mathrm{Re}\mu > 0$. The generality of our result regarding the analytic continuation in the coupling constant λ in Sec. 2 is not influenced if we restrict to the case $|\arg\mu| \leqslant \pi/2 - \omega$ where ω is arbitrary small. For $|\arg\mu| \leqslant \pi/2 - \omega$, $|\mu| \to \infty$ we can determine the asymptotic behavior of (15) by applying the complex Laplace–Watson[9] method (a particular case of the steepest descent method. By direct application of the Laplace formula[9] we find that for $m > 0$ and $|\arg\mu| \leqslant \pi/2 - \omega$, $|\mu| \to \infty$ the quotient (15) goes uniformly to zero. This shows that (15) is uniformly bounded in the region $|\arg\mu| \leqslant \pi/2 - \omega$, $\mathrm{Re}\alpha \geqslant \delta$, $1 - \delta \leqslant |\alpha| \leqslant 1 + \delta$. The proof of the convergence of the Mayer expansion follows now as in Sec. 4. This completes the proof of Theorem 3.

7. CONCLUSIONS

We have studied some analytic properties in λ of the lattice $\lambda P(\phi) + \frac{1}{2}m^2\phi^2$ models in arbitrary number of dimensions for large m. From the results of this paper it follows that the lattice models have Schwinger functions which are Borel summable in the classical sense. The required relaxed factorial bound can be proved as in Ref. 2 or more directly as in Ref. 3.

ACKNOWLEDGMENTS

I would like to thank Professor J. Fröhlich (I.H.E.S.), B. Ströter (Frankfurt) for discussions, and Professor G.V. Efimov (Dubna) for sending me a copy of his work prior to publication.

[1]M. Reed and B. Simon, *Methods of Modern Mathematical Physics IV: Analysis of Operators*, (Academic, New York, 1978), and references given there; in particular B. Simon, Ann. Phys. **58**, 76 (1970).

[2]J.-P. Eckmann, J. Magnen, and R. Sénéor, Commun. Math. Phys. **39**, 251 (1975).

[3]F. Constantinescu and B. Ströter (in preparation).

[4]G.V. Efimov, Dubna Preprint E2-11462 (1978) (unpublished).

[5]V.A. Malyshev, Commun. Math. Phys. **64**, 131 (1979).

[6]J. Glimm, A. Jaffe, and T. Spencer, in *Lecture Notes, Physics* (Springer, Berlin, 1973), Vol. 25.

[7]G.S. Sylvester, "Weakly coupled Gibbs measure," Rockefeller University preprint (1977) (unpublished).

[8]C. Gruber and H. Kunz, Commun. Math. Phys. **22**, 133 (1971); and H. Kunz, Commun. Math. Phys. **59**, 53 (1978).

[9]F. W. J. Olver, *Asymptotics and Special Functions* (Academic, New York, 1974).

Theory of involutional transformations applied to the Dirac theory of the electron I. Remarks on the Dirac plane waves

Shoon K. Kim
Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122

(Received 4 February 1980; accepted for publication 4 April 1980)

The involutional transformations are shown to play a fundamental role in the Dirac relativistic theory of the electron. The symmetric plane waves in the momentum representation, which are the simultaneous eigenfunctions of the energy and helicity operator, are given by the columns of a unitary matrix defined by $U = Y_E Y_s$, where Y_E and Y_s are IUH (involutional, unitary, and Hermitian) matrices which diagonalize the Dirac Hamiltonian and the helicity of a free particle respectively. It is shown that all the known properties of the Dirac plane waves stem from the properties of the involutional transformations Y_E and Y_s.

1. INTRODUCTION

Recently, the author has developed a general theory of matrix transformation[1,2] which gives an explicit solution for the matrix equation

$$AX = XB, \tag{1.1}$$

for a given pair of square matrices A and B of any given order satisfying the same polynomial equation. In a special case when A and B are involutional, i.e., $A^2 = B^2 = I$, where I is the unit matrix, the theory takes a particularly simple form. The simplest special case of the theorem (Theorem 2 of Ref. 1) has been introduced as the basic lemma for the recent work on the general theory of the spinors.[3] We shall, however, restate the lemma in a slightly modified form suitable to the present work and introduce a corollary to it.

Lemma: Let A and B be involutional matrices of a given order satisfying $A^2 = B^2 = I$. If their anticommutator is a c-number ($\neq -2$),

$$AB + BA = 2c\, I, \quad c \neq -1, \tag{1.2}$$

then there exists an involutional transformation which interchanges A and B via

$$A = YBY, \quad Y^2 = I, \tag{1.3}$$

where

$$Y = (A + B)/(2 + 2c)^{1/2}.$$

The direct proof of this lemma is also very simple. A simple extension of this lemma leads to the following corollary.

Corollary: The most general transformation V, which connects the involutional matrices A and B of the lemma, via a similarity transformation

$$A = VBV^{-1},$$

is given by

$$V = F_A Y = Y F_B, \tag{1.4}$$

where F_A (F_B) is a nonsingular matrix which commutes with A (B).

The fundamental nature of this lemma on the Dirac theory of the electron can be seen from that any linear combinations of the Dirac γ-matrices are involutional and their anticommutators are c-numbers. It has been shown in the previous work[3] that this lemma leads to the general theory of the spinor representations of the group of orthogonal transformations O (d, $\mathbb{C}$) in a d-dimensional Euclidean space $V^{(d)}$ over the complex field. In a special case of the Lorentz group, the lemma gives the complete parametrization of the basic spinor representations of order 2×2. It should be mentioned[3] that the geometric interpretations of the lemma and the corollary with the simplest choice of F_A ($= A$) lead to the general expressions of the axial involution (twofold rotation) and plane rotation in $V^{(d)}$ respectively.

In the present work we shall use the lemma to explicitly construct the Dirac plane waves and analyze their properties. Since the Dirac Hamiltonian for a free particle is involutional, one can immediately write down the involutional transformation Y_E which diagonalizes the Hamiltonian. Moreover, the Hamiltonian is Hermitian so that Y_E is an IUH (involutional, unitary, and Hermitian) matrix.[1] Thus, the four columns of Y_E provide the complete set of orthonormalized energy eigenspinors in the momentum representation. One of the satisfying features of the solution is that it can easily be shown to be a Lorentz transform of the spinor at the rest frame (Sec. 2). We shall introduce the mean operator of an observable A by the involutional transform $A_E = Y_E A Y_E$ and study its properties. This concept has first been introduced by Foldy and Wouthuysen[4] based on a unitary transformation which is not involutional. We shall find that the above definition is more satisfactory owing to the characteristic property of an involutional transformation that the inverse transformation is the same as the original one (Sec. 3). Using the corollary of the lemma we shall next construct the symmetric Dirac plane waves which are the simultaneous eigenfunctions of the energy and helicity. The plane waves are described by the columns of a unitary matrix defined by $U = Y_E Y_s$ where Y_s is another IUH matrix which diagonalizes the helicity operator. It will be shown that all the symmetry properties and orthogonality relations of the plane waves follow directly from those of the IUH matrices Y_E and Y_s (Secs. 4 and 5).

It is emphasized that the purpose of the present communication is to show the fundamental roles played by the involutional transformations in the Dirac relativistic theory of the electron. In the forthcomming communication we

 0022-2488/80/082282-04$1.00

shall show that the Dirac–Coulomb problem can be reduced to the nonrelativistic Coulomb problem by a single involutional transformation. It seems that the simple lemma introduced in the beginning of this introduction provides the most important mathematical tool thus far missing in handling the Dirac relativistic theory of the electron.

2. THE INVOLUTIONAL TRANSFORMATION OF THE DIRAC HAMILTONIAN

The Dirac Hamiltonian for a free particle is written as

$$H = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta m, \quad c = \hbar = 1. \tag{2.1}$$

It is involutional,

$$H^2 = E^2, \quad E = \pm (m^2 + p^2)^{1/2}, \tag{2.2}$$

where $p = |\mathbf{p}|$ and the sign of E is left arbitrary to ensure the greater symmetry of the Dirac plane wave solutions. From the lemma, we can diagonalize the Hamiltonian via an involutional transformation,

$$Y_E H Y_E = \beta E, \quad Y_E{}^2 = I, \tag{2.3}$$

where

$$Y_E = Y_E(\mathbf{p}) = N_E(H + \beta E),$$

$$N_E = \operatorname{sgn}(E)(2E^2 + 2Em)^{-\frac{1}{2}}.$$

We may rewrite Y_E in a form similar to that of the Hamiltonian itself,

$$Y_E = Y_E(\mathbf{p}) = p'\alpha_p + m'\beta, \quad \alpha_p = (\boldsymbol{\alpha}\cdot\hat{\mathbf{p}}), \tag{2.4}$$

where $\hat{\mathbf{p}} = \mathbf{p}/p$ and

$$p' = \operatorname{sgn}(E)[\tfrac{1}{2}(1 - (m/E))]^{1/2}, \quad m' = [\tfrac{1}{2}(1 + (m/E))]^{1/2}.$$

Since the Hamiltonian is Hermitian, Y_E is an IUH (involutiona, unitary, and Hermitian) matrix.[1] Thus, the four columns of Y_E give a complete set of orthonormalized eigenvectors of H. Let the set of column vectors be $u_1 \sim u_4$, then we can write

$$Hu_\nu = \epsilon_\nu E u_\nu, \quad \nu = 1, 2, 3, 4, \tag{2.5}$$

where $u_\nu = u_\nu(E,\mathbf{p}) = Y_E \chi_\nu$ with

$$\chi_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \quad \chi_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \quad \chi_3 = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \quad \chi_4 = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}, \tag{2.6}$$

$$\epsilon_\nu = \begin{cases} 1 & \text{for } \nu = 1, 2, \\ -1 & \text{for } \nu = 3, 4. \end{cases}$$

Now, one has to show that the Dirac spinor $u_\nu\,(E,\mathbf{p})$ is a Lorentz transform of χ_ν, which is the spinor belonging to the Hamiltonian $m\beta$ at the rest frame. We can easily achieve this by rewriting $u_\nu\,(E,\mathbf{p})$ in the form

$$u_\nu(E,\mathbf{p}) = (1 - v_0^2)^{1/4}\epsilon_\nu \exp[\tfrac{1}{2}\epsilon_\nu \chi(\boldsymbol{\alpha}\cdot\hat{\mathbf{p}})]\chi_\nu, \quad E > 0, \tag{2.7}$$

and noting that the spinor representation of the pure Lorentz transformation S_{Lor} corresponding to a uniform motion of the coordinate system $(\mathbf{r}, it)$ with the velocity $\mathbf{v}_0 = -\mathbf{p}/E$ is given by[3]

$$S_{\mathrm{Lor}} = \pm \exp[\tfrac{1}{2}\chi(\boldsymbol{\alpha}\cdot\hat{\mathbf{p}})], \quad \tanh\chi = |\mathbf{v}_0|. \tag{2.8}$$

The factor $(1 - v_0^2)^{1/4}$ in front of (2.7) accounts for the Lorentz contraction of the spacial volume element. To understand the involutory nature of Y_E, it would be even better to compare (2.8) with the spinor representation of an improper Lorentz transformation called the axial involution γ_h of χ_ν about a vector h ($= i\sinh(\chi/2)\hat{\mathbf{p}}, \cosh(\chi/2)$) in the Lorentz frame, for γ_h is given by[3]

$$\gamma_h = \Sigma\gamma_\nu h_\nu = S_{\mathrm{Lor}}\beta. \tag{2.9}$$

So far, we have considered the involutional transformation which diagonalizes the Hamiltonian H. According to the lemma, we can transform H into various other forms. An interesting special case is to transform H into $E\alpha_p$ via

$$Y_\alpha H Y_\alpha = E\alpha_p, \quad Y_\alpha^2 = I, \tag{2.10}$$

where

$$Y_\alpha = [H + E\alpha_p]/[2E(E+p)]^{1/2},$$

which is again an IUH matrix. The above transformation is interesting since it brings the Dirac Hamiltonian for a particle with mass m into that of a massless particle. The above form of the Hamiltonian has been considered by Cini and Touschek[5] and by Bose *et al.*[5] Their transformation matrix appears somewhat involved being a modification of the Foldy–Wouthuysen transformation[4] (see 3.7). It can, however, be shown that their transformation is simply equal to $\alpha_p Y_E$.

One may transform H further into the form of the Weyl Hamiltonian of a massless particle,

$$H_w = E\beta\Sigma_p, \quad \Sigma_p = (\boldsymbol{\Sigma}\cdot\hat{\mathbf{p}}), \tag{2.11}$$

through

$$T_w^\dagger H T_w = H_w, \quad T_w = Y_\alpha(\rho_1 + \rho_3)/\sqrt{2}, \tag{2.12}$$

where T_w is still a unitary matrix. If one uses Theorem 2 of Ref. 1, one can transform H into H_w also by

$$S^{-1} H S = E\beta\Sigma_p, \tag{2.13}$$

where

$$S = H + E\beta\Sigma_p,$$

$$S^{-1} = i\rho_2 (H - E\beta\Sigma_p)/(2pE).$$

Here S is not unitary but Hermitian. It is noted that the determinant of S is given by

$$\det S = (2pE)^2. \tag{2.14}$$

3. MEAN OPERATORS

Let us define the mean operator A_E of an observable A by the involutional transform of A with Y_E which diagonalizes H,

$$A_E = Y_E A Y_E. \tag{3.1}$$

Then A_E gives the matrix representation of A in the representation which diagonalizes H. A few important examples are

$$H_E = \beta E, \quad \beta_E = H/E,$$

$$\boldsymbol{\alpha}_E = (\mathbf{p}/E)\beta - [1 - \mathbf{pp}(E^2 + Em)^{-1}]\cdot\boldsymbol{\alpha}. \tag{3.2}$$

In the Dirac standard representations, α is an odd operator while β is an even operator. Thus, within the manifold of the positive or negative energy states the mean operators H_E and $\boldsymbol{\alpha}_E$ behave like classical observables, being constants equal to the respective classical expectation values. In the case of the velocity operator, $\boldsymbol{\alpha} = \dot{\mathbf{r}}$, the mean operator $\boldsymbol{\alpha}_E$ equals $\epsilon_\nu \mathbf{p}/E$ in each energy manifold. The odd operator part of $\boldsymbol{\alpha}_E$ describes the so-called Zitterbewegung which, however, can be observed only in the nonstationary states.

On account of the involutional nature of Y_E, if A_E is the mean operator of A, then A is also the mean operator of A_E,

$$A = Y_E A_E Y_E. \tag{3.3}$$

An example of such a pair is H and βE as given by (3.2). It is of interest to see how the mean operator changes in time. The equation of motion of A_E is given by

$$\dot{A}_E = i[H,A_E] = iY_E[\beta E,A]\,Y_E. \tag{3.4}$$

Here, it should be kept in mind that E is a function of the momentum operator $\mathbf{p}$, as given by (2.2). In a special case when A is the position operator, we have

$$\dot{\mathbf{r}}_E = (\mathbf{p}/E)(H/E), \tag{3.5}$$

which yields, in the energy–momentum representation,

$$(\dot{\mathbf{r}}_E)_E = \beta\mathbf{p}/E. \tag{3.6}$$

The rhs is completely diagonal and hence free of Zitterbewegung in contrast to $\boldsymbol{\alpha}_E = (\dot{\mathbf{r}})_E$.

The equation of motion (3.4) can also be used to find constants of motion. Suppose that A commutes with βE, then A_E is a constant of motion. Examples of such A are the orbital angular momentum $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ and the spin angular momentum $\frac{1}{2}\boldsymbol{\Sigma}$. Accordingly, $\mathbf{L}_E = Y_E\mathbf{L}Y_E$ and $\frac{1}{2}\boldsymbol{\Sigma}_E = \frac{1}{2}Y_E\boldsymbol{\Sigma}Y_E$ are constants of motion. For the case of the Dirac operator, $K = \beta(\boldsymbol{\Sigma}\cdot\mathbf{L} + 1)$, which is one of the well-known constants of motion, we have $K = K_E$.

The concept of the mean operators was first introduced by Foldy and Wouthuysen by means of a unitary matrix.

$$T_{\rm FW} = \exp[\tfrac{1}{2}\beta(\boldsymbol{\alpha}\cdot\hat{\mathbf{p}})\tan^{-1}(p/m)], \tag{3.7}$$

which brings the Dirac Hamilitonian into a form which is free of odd operators, i.e., $T_{\rm FW}HT_{\rm FW}^{\dagger} = \beta E$ which is diagonal in this case. It can easily be shown that the matrix $T_{\rm FW}$ is simply equal to βY_E. Since $T_{\rm FW}$ is not involutional, their arguments on the mean operators are somewhat involved.

4. THE SYMMETRIC DIRAC PLANE WAVES

The Dirac plane waves described by $u_\nu = Y_E\chi_\nu$ are degenerate. One may easily remove this degeneracy by introducing the simultaneous eigenspinors of the energy and helicity operator $\Sigma_p = \boldsymbol{\Sigma}\cdot\hat{\mathbf{p}}$. For this purpose, we shall first introduce one more involutional transformation which diagonalizes the helicity. From the lemma, we have

$$Y_s\Sigma_p Y_s = s\Sigma_3, \quad \Sigma_s^2 = I, \tag{4.1}$$

where $s = \pm 1$ is inserted for convenience, and

$$Y_s = Y_s(\hat{\mathbf{p}}) = N_s(\Sigma_p + s\Sigma_3), \quad N_s = (2 + 2s\hat{\mathbf{p}}_3)^{-1/2}.$$

By definition $Y_s(\hat{\mathbf{p}})$ satisfies the symmetry property

$$Y_s(\hat{\mathbf{p}}) = -Y_{-s}(-\hat{\mathbf{p}}). \tag{4.2}$$

Now, from the corollary of the lemma in Sec. 1, one can see that a unitary matrix which diagonalizes H and Σ_p via

$$U^\dagger HU = BE, \quad U^\dagger\Sigma_p U = s\Sigma_3, \tag{4.3}$$

is given by

$$U = U(E,\mathbf{p},s) = Y_E(\mathbf{p})Y_s(\hat{\mathbf{p}}). \tag{4.4}$$

Thus, the simultaneous eigenspinor belonging to the energy $\epsilon_\nu E$, the momentum $\mathbf{p}$ and the helicity $(-1)^{\nu+1}s$ is given by the column vector of U, denoted by U_ν,

$$U_\nu(E,\mathbf{p},s) = U(E,\mathbf{p},s)\chi_\nu, \tag{4.5a}$$

in the momentum representation. For later use we shall write the eigenspinor in the coordinate representation as follows:

$$\psi_\nu(E,\mathbf{p},s;\mathbf{r},t) = U\chi_\nu\exp[i\mathbf{p}\cdot\mathbf{r} - i\epsilon_\nu Et]. \tag{4.5b}$$

The unitary matrix U defined by (4.4) is highly symmetric. To see this, we first rewrite (4.4) in the form

$$U = Y_s(s\alpha_3 p' + \beta m'), \tag{4.6a}$$

from which it follows that

$$U = N_s\begin{bmatrix} m'\begin{bmatrix}\hat{p}_3+s & \hat{p}_- \\ \hat{p}_+ & -(\hat{p}_3+s)\end{bmatrix} & sp'\begin{bmatrix}\hat{p}_3+s & -\hat{p}_- \\ \hat{p}_+ & \hat{p}_3+s\end{bmatrix} \\ sp'\begin{bmatrix}\hat{p}_3+s & \hat{p}_- \\ \hat{p}_+ & \hat{p}_3+s\end{bmatrix} & -m'\begin{bmatrix}\hat{p}_3+s & \hat{p}_- \\ \hat{p}_+ & -(\hat{p}_3+s)\end{bmatrix}\end{bmatrix}, \tag{4.6b}$$

where N_s is the normalization constant for Y_s of (4.1). We see immediately that the positive and negative energy states are related by

$$U_3(E,\mathbf{p},s) = \operatorname{sgn}(sE)U_1(-E,\mathbf{p},s),$$
$$U_4(E,\mathbf{p},s) = -\operatorname{sgn}(sE)U_2(-E,\mathbf{p},s), \tag{4.7}$$

and the large and small components of each U_ν denoted by $U_\nu^>$ and $U_\nu^<$ respectively satisfy

$$U_\nu^<(E,\mathbf{p},s) = (-1)^{\nu+1}\epsilon_\nu sU_\nu^>(-E,\mathbf{p},s), \tag{4.8}$$

which holds for positive or negative E. It can be shown that a symmetry relation analogous to (4.8) exists also for the Dirac–Coulomb waves.[3] On account of these symmetries one may call ψ_ν of (4.5) the symmetric Dirac plane waves.

We shall next discuss the transformation properties of the ψ_ν under the various symmetry operations. Let $C = \gamma_2 K$, $T = -i\Sigma_2 K$, and $P = -i\gamma_4$ be the charge conjugation, time reversal and parity operation respectively, where the Dirac operators are in the standard representation and K is the complex conjugation. Then, from the symmetry properties of the involutional matrices of Y_E and Y_s under these symmetry operations, one can easily obtain the following result,

$$C\psi_\nu(E,\mathbf{p},s;\mathbf{r},t) = (-1)^\nu\epsilon_\nu\psi_{5-\nu}(E,-\mathbf{p},-s;\mathbf{r},t),$$
$$T\psi_\nu(E,\mathbf{p},s;\mathbf{r},-t) = (-1)^\nu\psi_{\tilde{\nu}}(E,-\mathbf{p},-s;\mathbf{r},t), \tag{4.9}$$
$$P\psi_\nu(E,\mathbf{p},s;-\mathbf{r},t) = i\epsilon_\nu\psi_\nu(E,-\mathbf{p},-s;\mathbf{r},t),$$

where $\tilde{\nu} = \nu - (-1)^\nu$, i.e., $\tilde{1} = 2, \tilde{2} = 1$, etc. Using (4.7), we may rewrite the first equation of (4.9) as follows:

$$C\psi_\nu(E,\mathbf{p},s;\mathbf{r},t) = \operatorname{sgn}(sE)\psi_{\tilde{\nu}}(-E,-\mathbf{p},-s;\mathbf{r},t). \tag{4.10}$$

Thus, under these symmetry operations, C, T, and P, the wavefunctions ψ_ν transform within the subset ψ_1, ψ_2 or ψ_3, ψ_4 with modification of the parameters.

5. THE ORTHOGONALITY AND PROJECTIVE PROPERTIES OF THE DIRAC PLANE WAVES

The orthogonality of the set of the column vector U_ν is a direct consequence of the unitary nature of the transformation matrix U,

$$U_\nu^\dagger U_\mu = \chi_\nu^\dagger U^\dagger U \chi_\mu = \delta_{\nu\mu}, \quad \nu,\mu = 1,2,3,4, \tag{5.1}$$

$$\sum_{\nu=1}^{4} U_\nu U_\nu^\dagger = \sum_\nu \chi_\nu U U^\dagger \chi_\nu^\dagger = I.$$

For the Dirac adjoint spinors $\bar{U}_\nu = U_\nu^\dagger \beta$ we have

$$\bar{U}_\nu U_\mu = \chi_\nu^\dagger Y_s H Y_s \chi_\mu / E = \chi_\nu^\dagger (m\beta + sp\alpha_3)\chi_\mu / E, \tag{5.2}$$

$$\sum_\nu U_\nu \bar{U}_\nu = \sum_\nu U \chi_\nu \chi_\nu^\dagger U^\dagger \beta = \beta.$$

Since α_3 is an odd operator, we have the following orthogonality relations in the positive or negative energy manifolds:

$$\bar{U}_\nu U_{\nu'} = \epsilon_\nu (m/E)\delta_{\nu\nu'}, \quad \nu,\nu' = 1,2 \text{ or } 3,4.$$

These relations (5.1) and (5.2) hold for the set of spinors U_ν with fixed parameters. If one uses the relation

$$\beta Y_E(\mathbf{p})\beta = Y_E(-\mathbf{p}),$$

one obtains an alternate set of orthogonality relations and its closure,

$$\bar{U}_\nu(E, -\mathbf{p}, -s) U_\mu(E,\mathbf{p},s) = \chi_\nu^\dagger Y_{-s}(-\hat{\mathbf{p}})\beta Y_s(\hat{\mathbf{p}})\chi_\mu = -\epsilon_\nu \delta_{\nu\mu}, \tag{5.3}$$

$$\sum_{\nu=1}^{4} -\epsilon_\nu U_\nu(E,\mathbf{p},s)\bar{U}_\nu(E, -\mathbf{p}, -s) = I.$$

Next, we shall construct the projection operators from the orthonormalized complete set $\{U_\nu\}$. The projection operator P_z onto a subset z of the set of the sates $\{U_\nu\}$ is defined by the partial sum,

$$P_z = \sum_{\nu\epsilon z} U_\nu U_\nu^\dagger, \quad P_z^2 = P_z. \tag{5.4}$$

In the calculation of P_z the following identity plays the essential role

$$\chi_\nu \chi_\nu^\dagger = \tfrac{1}{4}(1 + \epsilon_\nu \beta)[1 - (-1)^\nu \Sigma_3], \tag{5.5}$$

from which it follows that

$$\sum_{\nu=1,2} \chi_\nu \chi_\nu^\dagger = \tfrac{1}{2}(1+\beta), \quad \sum_{\nu=3,4} \chi_\nu \chi_\nu^\dagger = \tfrac{1}{2}(1-\beta)$$

$$\sum_{\nu=1,3} \chi_\nu \chi_\nu^\dagger = \tfrac{1}{2}(1+\Sigma_3), \quad \sum_{\nu=2,4} \chi_\nu \chi_\nu^\dagger = \tfrac{1}{2}(1-\Sigma_3) \tag{5.6}$$

It is an immediate consequence of these identities that the projection operator onto the states of a given energy E is given by

$$P_E(\mathbf{p}) = \sum_{\nu=1,2} U_\nu(E,\mathbf{p},s)U_\nu^\dagger(E,\mathbf{p},s) = Y_E Y_s \tfrac{1}{2}(1+\beta)Y_s Y_E = \tfrac{1}{2}(1 + (H/E)), \tag{5.7}$$

where E can be positive or negative. Likewise, the projection operator onto the states with the helicity s is given by

$$P_s(\hat{\mathbf{p}}) = \sum_{\nu=1,3} U_\nu(E,\mathbf{p},s)U_\nu^\dagger(E,\mathbf{p},s) = Y_E Y_s \tfrac{1}{2}(1+\Sigma_3)Y_s Y_E = \tfrac{1}{2}[1 + s(\mathbf{\Sigma}\cdot\hat{\mathbf{p}})]. \tag{5.8}$$

It is noted here that $P_E(\mathbf{p})$ and $P_s(\mathbf{p})$ commute with each other since the helicity is a constant of motion. Finally, the projection operator onto the state with the energy E and the helicity s is given by

$$P_{E,s}(\mathbf{p}) = U_1(E,\mathbf{p},s)U_1^\dagger(E,\mathbf{p},s) = \tfrac{1}{4}Y_E Y_s(1+\beta)(1+\Sigma_3)Y_s Y_E = P_E(\mathbf{p})P_s(\hat{\mathbf{p}}). \tag{5.9}$$

It is well known[6] that the proofs of these relations (5.2), (5.3), and (5.7)~(5.9) for the Dirac plane waves are very messy, without use of the factorized form of the unitary matrix, $U = Y_E Y_s$.

It is worthwhile to note that the difference between the actions of the unitary operator $U^\dagger$ and the projection operator $P_{E,s}(\mathbf{p})$ on an arbitrary spinor state $U_a = \Sigma_\nu a_\nu U_\nu$ may be characterized by

$$U^\dagger U_a = \sum_{\nu=1}^{4} a_\nu \chi_\nu, \quad P_{E,s} U_a = a_1 U_1, \tag{5.10}$$

that is, the column elements of the former describe the probability amplitudes of all spinor states while the latter projects out the state with the energy E and the helicity s.

Under the charge conjugation C, time reversal T, and parity operation P, the projection operator $P_{E,s}$ transforms as follows:

$$CP_{E,s}(\mathbf{p})C^{-1} = P_{-E,s}(-\mathbf{p}),$$
$$TP_{E,s}(\mathbf{p})T^{-1} = P_{E,s}(-\mathbf{p}), \tag{5.11}$$
$$PP_{E,s}(\mathbf{p})P^{-1} = P_{E,-s}(-\mathbf{p}),$$

where $\mathbf{p}$ is the momentum vector which is real.

Through these proofs of the well-known properties of the Dirac plane waves one may be convinced with the effectiveness of the factorized form of the unitary matrix $U = Y_E Y_s$ in describing the Dirac plane waves.

ACKNOWLEDGMENTS

The author wishes to thank Professor J. Franklin at Temple University and Professor S. S. Yun at Ohio State University, Athens for helpful suggestions.

[1]Shoon K. Kim, J. Math. Phys. **20**, 2153 (1979).
[2]Shoon K. Kim, J. Math. Phys. **20**, 2159 (1979).
[3]Shoon K. Kim, J. Math. Phys. **21**, 1299 (1980).
[4]L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).
[5]M. Cini and B. Touschek, Nuovo Cimento **7**, 422 (1958); see also S. K. Bose, A. Gamba, and E. C. G. Sudarshan, Phys. Rev. **113**, 1661 (1959).
[6]See for example, M. E. Rose, *Relativistic Electron Theory* (John Wiley & Sons, Inc., New York, 1961), p. 92. See also S. S. Schweber, *Relativistic Quantum Field Theory*, (Harper and Row, New York, 1962), p. 85.

Theory of involutional transformations applied to the Dirac theory of the electron. II. Remarks on the Dirac–Coulomb waves

Shoon K. Kim
Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122

(Received 4 February 1980; accepted for publication 2 May 1980)

By means of a single involutional transformation, the Dirac equation for a charged particle in a central Coulomb field is reduced to a second-order differential equation for the radial part which has the same form as the radial equation for the nonrelativistic Coulomb problem. The general Dirac–Coulomb waves thus obtained are highly symmetric with respect to signs of the energies and other constants of motion. It gives the precise phase shift for the continuum spectrum without any additive multiple of π in contrast to the existing solutions. The symmetry properties of the Dirac–Coulomb waves are discussed.

1. INTRODUCTION

In the previous work[1] (referred as I), the theory of involutional transformation has been applied to construct the Dirac plane wave solutions. It is based on a simple lemma which explicitly gives an involutional transformation which connects two involutional matrices. This simple lemma has previously been applied to construct the general theory of the spinor representation of the group of the orthogonal transformations in arbitrary dimensions,[2] a special case of which is the Lorentz group.

In the present work we shall apply the involutional transformation to solve the Dirac–Coulomb problem. The method is so very effective that the exact solution follows almost immediately from knowledge of the nonrelativistic Coulomb problem. Since the Hamiltonian is not involutional in the present problem, it is not possible to write down the spinor solution as in the case of the free Dirac particle. We can, however, introduce an involutional transformation Y_λ which diagonalizes an involutional operator Λ contained in the Hamiltonian. This operator was first introduced by Temple,[3] and then later by Martin and Glauber,[4] and by others.[5,6] Its eigenvalues λ play a role similar to that of the orbital angular momentum quantum numbers in the nonrelativistic problem. It will be shown that the involutional transformation Y_λ immediately reduces the Dirac equation to a pair of first-order differential equations for the radial parts known as the Infeld–Hull[7] factorized form of a second-order differential equation. The latter has the same form as the radial equation for the nonrelativistic Coulomb problem. Since the solutions of this equation are completely known, the rest is a simple matter of transforming the solutions back into the original representation by the involutional transformation Y_λ. Despite a great deal of work on this famous old problem (especially by Darwin[8] and by others[3–7,9,10]), it still seems to have some unrecognized simple features.

The present method gives the solution which seems thus far the most satisfactory in the sense that it is simple and holds for both positive and negative energies as well as for attractive or repulsive Coulomb field. The so-called large and small components of the spinor interchanges with each other by simple changes of the signs of the energy E, charge z, and κ the eigenvalue of the Dirac operator K. The solution clearly exhibits the proper symmetries with respect to the charge conjugation, time reversal and parity operators. It is smoothly transformed into the force-free case as well as into the nonrelativistic case without any further calculation. There exists little difficulty in calculating the normalization constant or the phase shift of the solution. This is in marked contrast to Darwin's solution, for which it is well known that the normalization requires a formidable task for the discrete case and that the phase shift can be determined up to an additive multiple of π for the continuum case. We shall also discuss the so-called Coulomb helicity operator first introduced by Johnson and Lippman[11] and later by Biedenharn.[5] The operator will be given in its most symmetrized form.

2. THE DIRAC–COULOMB WAVES

The Dirac equation for a charged particle in a central Coulomb field may be written as

$$(H - E)\psi = 0, \tag{2.1}$$

with

$$H = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta m - zr^{-1}, \quad c = \hbar = 1.$$

We shall leave the signs of the energy E and the potential parameter z arbitrary, in order to achieve the greater symmetry of the wavefunction ψ with respect to the signs of these quantities. It is assumed, however, that $m \geqslant 0$ and $|z| < 1$. For the electron in a nuclear field, z equals the atomic number of the nucleus multiplied by the fine structure constant. In the spherical coordinates we have

$$H = \beta m - i\alpha_r\left(\frac{\partial}{\partial r} + \frac{1}{r} + \frac{\Lambda}{r}\right), \tag{2.2}$$

where $\alpha_r = r^{-1}(\boldsymbol{\alpha}\cdot\mathbf{r})$ and Λ is an operator[3–6] defined by

$$\Lambda = \beta K - i\alpha_r z, \quad K = -\beta(\boldsymbol{\Sigma}\cdot\mathbf{L} + 1). \tag{2.3}$$

Here $\mathbf{L}$ is the usual orbital angular momentum, $\boldsymbol{\Sigma}$ is the Dirac spin matrices, and K is the Dirac operator as originally defined except that the sign has been changed for convenience.

As is well known, the Dirac operator K is a constant of motion and commutes with β, α_r, $\mathbf{L}$, and $\mathbf{J}$ (the total angular

 0022-2488/80/082286-05$1.00

momentum). The operator Λ commutes with K and $\mathbf{J}$ but it is not a constant of motion. Both K and Λ are involutional in the subspace where $\mathbf{J}^2 = j(j+1)$, since

$$K^2 = \mathbf{J}^2 + \tfrac{1}{4}, \quad \Lambda^2 = K^2 - z^2. \tag{2.4}$$

Thus, the eigenvalues of K are given by $\kappa = \pm 1, \pm 2, \cdots$ and the eigenvalues of Λ are given by $\pm\lambda$ with

$$\lambda = \kappa(1 - (z/\kappa)^2)^{1/2} \tag{2.5}$$

where we have adopted the positive square root convention and hence sign $(\lambda) =$ sign (κ). In view of the definition of K, one may introduce a two-component spinor χ_κ^μ which satisfies

$$-(\sigma\cdot\mathbf{L} + 1)\,\chi_\kappa^\mu = \kappa\,\chi_\kappa^\mu, \quad J_z\,\chi_\kappa^\mu = \mu\chi_\kappa^\mu. \tag{2.6}$$

Then, χ_κ^μ is also an eigenfunction of J^2 and L^2 belonging to their respective quantum numbers given by

$$j(\kappa) = |\kappa| - \tfrac{1}{2}, \quad l(\kappa) = |\kappa| + \tfrac{1}{2}[\text{sign}(\kappa) - 1]. \tag{2.7}$$

The well-known explicit form of χ_κ^μ and its symmetry properties are given in the Appendix for convenience.

With the preparation given above we shall let $\psi_{E,\kappa}^\mu$ be a simultaneous eigenfunction of H, K, and J_z with the respective eigenvalues E, κ, and μ and set

$$\phi_{E,\kappa}^\mu = Y_\lambda\,\psi_{E,\kappa}^\mu, \tag{2.8}$$

where Y_λ is the involutional transformation which diagonolizes Λ via

$$Y_\lambda\,\Lambda\,Y_\lambda = \beta\lambda, \quad Y_\lambda{}^2 = 1, \tag{2.9.1}$$

where, from the lemma of I,

$$Y_\lambda = \frac{1}{2\lambda}\left(\frac{2\lambda}{k+\lambda}\right)^{1/2}(\Lambda + \beta\lambda). \tag{2.9.2}$$

It should be noted that $\phi_{E,\kappa}^\mu$ is still an eigenfunction of K and J_z but of $H' = Y_\lambda H\,Y_\lambda$. The equation for $\phi_{E,\kappa}^\mu$ may be written in the form

$$Y_\lambda\, i\alpha_r\,(H - E)\,Y_\lambda\,\phi_{E,\kappa}^\mu = 0,$$

or explicitly

$$\left\{\frac{\partial}{\partial r} + \beta\left(\frac{\lambda}{r} - \frac{zE}{\lambda}\right) + i\alpha_r\left(\frac{\kappa E}{\lambda} - \beta m\right)\right\} r\,\phi_{E,\kappa}^\mu = 0. \tag{2.10}$$

If one gives the matrix representation of this equation one can see that it is best to set

$$\phi_{E,\kappa}^\mu = \begin{pmatrix} c_1\,R_{E,\lambda}\,\chi_\kappa^\mu \\ -is_1\,R_{E,-\lambda}\,\chi_{-\kappa}^\mu \end{pmatrix}, \tag{2.11}$$

where $R_{E,\pm\lambda}$ are properly normalized radial functions and c_1, s_1 are the normalization constants satisfying $c_1{}^2 + s_1{}^2 = 1$. Substitution of (2.11) into (2.10) followed by elimination of the angular parts, using $(\sigma\cdot\mathbf{r}/r)\chi_\kappa^\mu = -\chi_{-\kappa}^\mu$, yields

$$\left(\frac{d}{dr} + \frac{\lambda}{r} - \frac{zE}{\lambda}\right) r\,R_{E,\lambda} = N_\lambda\, r\,R_{E,-\lambda},$$

$$\left(-\frac{d}{dr} + \frac{\lambda}{r} - \frac{zE}{\lambda}\right) r\,R_{E,-\lambda} = N_\lambda\, r\,R_{E,\lambda}, \tag{2.12.1}$$

with

$$N_\lambda = \frac{s_1}{c_1}\left(\frac{E\kappa}{\lambda} + m\right) = \frac{c_1}{s_1}\left(\frac{E\kappa}{\lambda} - m\right) = \lambda^{-1}((\kappa E)^2 - (m\lambda)^2)^{1/2};$$

$$(E\kappa/m\lambda)^2 \geqslant 1. \tag{2.12.2}$$

In the Appendix, an alternate expression of N_λ is given by (A5). The inequality given above which limits the allowed energy range, follows from that the two operators on LHS of (2.12.1) are mutually Hermitian conjugate. The self-consistent set of the constants c_1 and s_1 may be written as

$$c_1 = \left[\left(1 + \frac{m\lambda}{E\kappa}\right)\Big/2\right]^{1/2},$$

$$s_1 = \text{sign}(\kappa E)\left[\left(1 - \frac{m\lambda}{E\kappa}\right)\Big/2\right]^{1/2}, \tag{2.12.3}$$

with the convention of the positive square root.

The set of Eq. (2.12.1) is known as the Infeld–Hull factorized form[7] of the differential equation

$$\left\{\frac{d^2}{dr^2} - \frac{\lambda(\lambda+1)}{r^2} - \frac{2zE}{r} - (m^2 - E^2)\right\} r\,R_{E,\lambda} = 0, \tag{2.13}$$

which has the same form as the radial equation for the nonrelativistic Coulomb problem. Since the solutions of this equation are completely known, one may consider that the problem is solved. The rest is a simple matter of writing down the solutions for $\phi_{E,\kappa}^\mu$ and transforming it back to $\psi_{E,\kappa}^\mu$ using Y_λ given by (2.9). For this purpose we shall first introduce a set of parameters,

$$q = (m^2 - E^2)^{1/2}, \quad 0 \leqslant \arg q \leqslant \frac{\pi}{2}, \quad \eta = zE/q, \tag{2.14.1}$$

where q is defined to be real and positive when $|E| < m$ and pure imaginary with the positive imaginary part when $|E| > m$. In any case, E takes the following familiar form

$$E = \pm m(1 + (z/\eta)^2)^{-1/2}, \tag{2.14.2}$$

where the sign has to be consistent with $E = q\eta/z$. From the boundary condition at the origin, $rR_{E,\lambda} \to 0$ as $r \to 0$, the initial exponent of the acceptable solution is given by

$$l(\lambda) = |\lambda| + \tfrac{1}{2}(\text{sign}(\lambda) - 1). \tag{2.15}$$

Thus, it is more convenient to denote the radial solution by $R_{E,l(\lambda)}$ instead of $R_{E,\lambda}$. Then,

$$R_{E,l(\lambda)}(r) = N_{E,l(\lambda)}\,(2|q|r)^{l(\lambda)}\,e^{-qr} \times {}_1F_1\,(l(\lambda) + 1 - \eta,\, 2l(\lambda) + 2,\, 2qr) \tag{2.16.1}$$

where ${}_1F_1$ is the confluent hypergeometric series[12] and $N_{E,l(\lambda)}$ is the normalization constant to be determined. The radial function is real even when q is imaginary, since then it is invariant for $q \to -q$ and $\eta \to -\eta$ (Kummer's transformation[12]). The function $R_{E,l(\lambda)}$ is further characterized for the discrete and continuum spectra separately by the convergence condition at $r \to \infty$:

(A) The discrete spectrum where $m \geqslant |E| \geqslant m(\lambda/\kappa)$. The radial function (2.16.1) is characterized by a set of discrete values of η and the corresponding normalization constants:

$$\eta = l(\lambda) + 1 + n_r \geqslant n_r, \quad n_r = 0, 1, 2, \cdots,$$

$$N_{E,l(\lambda)} = N_{\eta,l(\lambda)} \equiv \left[\frac{4q^3}{\eta}\frac{\Gamma(\eta+l(\lambda)+1)}{\Gamma(\eta-l(\lambda))}\right]^{1/2} \Big/ \Gamma(2l(\lambda)+2). \tag{2.16.2}$$

The discrete solution denoted by $R_{\eta,l(\lambda)}(r)$ can be expressed alternatively in terms of the associate Laguerre polynomial of the degree n_r defined by

$$L_{n_r}{}^{(s)}(2qr) = \binom{n_r+s}{n_r}\, {}_1F_1(-n_r, 1+s, 2qr).$$

It is noted that the sign of E is determined by the sign of z since $zE \geqslant 0$. The minimum value $|E| = m\lambda/\kappa$ for a given value of $|\kappa|$ occurs at $n_r = 0$, $\kappa < 0$ and $\eta = |\kappa|$. The corresponding wavefunctions take particularly simple forms, which will be shown in the next section.

(B) The continuum spectrum where $|E| \geqslant m$. The radial solution (2.16.1) is characterized by imaginary q and η, and the corresponding normalization constant,

$$q = ik,\ k>0,\quad \eta = -i\zeta,\quad \zeta = zE/k,$$

$$N_{E,l(\lambda)} = N_{k,l(\lambda)} \equiv (2\pi)^{1/2} e^{(1/2)\pi\zeta}\, |\Gamma(l(\lambda)+1-i\zeta)| / \Gamma(2l(\lambda)+2) \tag{2.16.3}$$

where the normalization constant is given in the wave number scale. The continuum wave denoted by $R_{k,l(\lambda)}$ has the following asymptote

$$R_{k,l(\lambda)} \sim (2/\pi)^{1/2} r^{-1} \sin(kr - (\pi/2)l(\lambda) + \zeta \ln(2kr) + \delta_{l(\lambda)}),$$
$$\delta_{l(\lambda)} = \arg\Gamma(l(\lambda)+1-i\zeta). \tag{2.16.4}$$

Since there exists no constraint on ζ except that it be real, the energy can be positive or negative for any given value of z.

It should be noted that the radial solutions (2.16.1) characterized by (2.16.2) and (2.16.3) have the same forms as those of the nonrelativistic Coulomb problem with a trivial exception of "the noninteger orbital angular momentum" $l(\lambda)$. It is also noted that the normalization constants are consistent with the constant factor N_λ of (2.12.2) which may be rewritten in the form of (A5) for the comparison.

Now, we shall transform $\phi^\mu_{E,\kappa}$ back to $\psi^\mu_{E,\kappa}$ by the involutional transformation Y_λ. Using (2.9.2) and (2.11) we may write the solution in the form,

$$\psi^\mu_{E,\kappa}(z,\mathbf{r}) = \begin{pmatrix} g_\kappa \chi^\mu_\kappa \\ if_\kappa \chi^\mu_{-\kappa} \end{pmatrix} = N_\kappa (\lambda/\kappa)^{1/2} Y_\lambda \phi^\mu_{E,\kappa}(z,\mathbf{r}), \tag{2.17.1}$$

where an extra normalization factor is introduced, for Y_λ is not unitary. The radial parts are given by

$$g_\kappa \equiv g_\kappa(z,E,\mathbf{r}) = N_\kappa\,[c_1c_2R_{E,l(\lambda)} + s_1s_2R_{E,l(-\lambda)}], \tag{2.17.2}$$

$$f_\kappa \equiv f_\kappa(z,E,\mathbf{r}) = N_\kappa\,[c_1s_2R_{E,l(\lambda)} + s_1c_2R_{E,l(-\lambda)}],$$

with the coefficients c_i and s_i $(i=1,2)$,

$$c_1 = 2^{-1/2}(1+(m\lambda/E\kappa))^{1/2},$$
$$s_1 = 2^{-1/2}\mathrm{sign}(\kappa E)(1-(m\lambda/E\kappa))^{1/2} \tag{2.17.3}$$
$$c_2 = 2^{-1/2}(1+(\lambda/\kappa))^{1/2},\ s_2 = 2^{-1/2}\mathrm{sign}(\kappa z)\,(1-(\lambda/\kappa))^{1/2}.$$

Note that $c_i{}^2 + s_i{}^2 = 1$ and the set (c_1,s_1) is from (2.12.3) and (c_2, s_2) is from Y_λ written in the form

$$Y_\lambda = (\kappa/\lambda)^{1/2}(c_2\beta - is_2\alpha_r). \tag{2.17.4}$$

The normalization factor N_κ is given by

$$N_\kappa = \begin{cases} |E\kappa/m\lambda| & \text{for the discrete spectrum,} \\ |\kappa/\lambda| & \text{for the continuum spectrum,} \end{cases} \tag{2.17.5}$$

where use has been made of the integrals given by (A3) and the identity (A4).

The solution given by (2.17) is completely general; it holds for arbitrary signs of the energy E and charge z. According to (2.17.2), we have the following symmetry properties,

$$f_\kappa(z,E;r) = \mathrm{sign}(\kappa E)\, g_{-\kappa}(-z,-E;r), \tag{2.18.1}$$

noting that $R_{E,l(\lambda)}$ depends on E through zE. Thus, the large component $\psi^\mu_{E,\kappa}(z,r)^>$ and small component $\psi^\mu_{E,\kappa}(z,r)^<$ of (2.17.1) are related by

$$\psi^\mu_{E,\kappa}(z,r)^< = i\,\mathrm{sign}(\kappa)\,\psi^\mu_{-E,\kappa}(-z,r)^>. \tag{2.18.2}$$

A similar symmetry has been obtained for the symmetric Dirac plane waves in I.

We shall next explicitly show the symmetry properties of the spinor solution (2.17) with respect to the charge conjugation $C\,(=\lambda_2K$, K being the complex conjugation) and the time reversal $T\,(=-i\Sigma_2K)$, and parity $P=-i\gamma_4$. Using (2.18.1) and the symmetry properties of χ^μ_κ given by (A2) we have

$$C\,\psi^\mu_{E,\kappa}(z;\mathbf{r}) = i(-1)^{\mu+1/2}\mathrm{sign}(E)\,\psi^{-\mu}_{-E,-\kappa}(-z,\mathbf{r}),$$
$$T\,\psi^\mu_{E\kappa}(z;\mathbf{r}) = (-1)^{\mu+1/2}\mathrm{sign}(E)\,\psi^{-\mu}_{E,\kappa}(z,\mathbf{r}), \tag{2.19}$$
$$P\,\psi^\mu_{E\kappa}(z;-\mathbf{r}) = i(-1)^{l(\kappa)}\,\psi^\mu_{E,\kappa}(z,\mathbf{r}).$$

These symmetry properties (except for the phase factors) are consistent with the transformation properties of the operators H, K, and J under the respective symmetry operations.

We shall now calculate the phase shift for the continuum spectrum. From (2.17.2) and (2.16.4) we obtain for the asymptotes of the radial functions g_κ and f_κ,

$$g_\kappa \sim (2/\pi)^{1/2} c_1(0) r^{-1}\sin(kr + \zeta\ln(2kr) - (\pi/2)l(\kappa) + \delta_\kappa), \tag{2.20.1}$$

$$f_\kappa \sim (2/\pi)^{1/2} s_1(0) r^{-1}\sin(kr + \zeta\ln(2kr) - (\pi/2)l(-\kappa) + \delta_\kappa),$$

where

$$c_1(0) = 2^{-1/2}(1+(m/E))^{1/2},$$
$$s_1(0) = 2^{-1/2}\mathrm{sign}(\kappa E)\,(1-(m/E))^{1/2} \tag{2.20.2}$$

and δ_κ is the phase shift, except for the logarithmic term, as one can see from that $\lambda \to \kappa$ as $z \to 0$. It is given by

$$\delta_\kappa = (\pi/2)|\kappa - \lambda| + \arg\Gamma(l(\lambda)+1-i\zeta) + \Delta_\kappa,$$
$$\Delta_\kappa = \tan^{-1}\left(\frac{1-(m/E)}{\kappa+\lambda}\zeta\right) = \tan^{-1}\left(\frac{\kappa-\lambda}{1+(m/E)}\zeta^{-1}\right),$$
$$-\frac{\pi}{2} \leqslant \Delta_\kappa \leqslant \frac{\pi}{2}, \tag{2.20.3}$$

where the identity (A4) has been used. It can easily be shown that $\delta_\kappa \equiv \delta_\kappa(z,E)$ satisfies the symmetry property,

$$\delta_\kappa(z,E) = \delta_{-\kappa}(-z,-E), \tag{2.20.4}$$

which can be shown by a direct calculation or from the general symmetry property (2.18.1).

In the existing continuum solutions[9,10] of Darwin's type the phase shift δ_κ is given through $\exp(2i\delta_\kappa)$ so that δ_κ is known only to within an additive multiple of π. This gives the sign ambiguity in Darwin's solution. One can transform the present solution to Darwin's solution[9] by using the contiguous relation of the confluent hypergeometric functions. For the comparison one may need the following relation

$$\exp[2i\Delta_\kappa] = [\kappa - i(m\zeta/E)]/(\lambda - i\zeta), \qquad (2.21)$$

which follows from the definition of Δ_κ given in (2.20.3).

Before concluding this section it is worthwhile to mention another trivial symmetry property for the radial functions with respect to the sign of λ. So far, we have used the convention that $\text{sign}(\lambda) = \text{sign}(\kappa)$. We may change this convention and regard their signs independent by redefining g_κ and f_κ as functions of κ and λ as they are written in (2.17.2). Let $g_\kappa = g_\kappa(\lambda,z,E)$, $f_\kappa = f_\kappa(\lambda,z,E)$, then we have

$$\begin{aligned} g_\kappa(-\lambda,z,E) &= \text{sign}(zE)\, g_\kappa(\lambda,z,E), \\ f_\kappa(-\lambda,z,E) &= \text{sign}(zE)\, f_\kappa(\lambda,z,E), \end{aligned} \qquad (2.22)$$

i.e., these are "invariant" for the sign change $\lambda \to -\lambda$ keeping the rest of the parameters unchanged. Thus, one can simply put $\lambda = |\lambda|$ in the radial functions. This property may conveniently be used when one compares the present result with the existing solutions[6,9,10] where $\lambda = |\lambda|$ is used. This convention, however, introduces an inconvenience when we calculate the phase shift δ_κ since $\lambda \to |\kappa|$ instead of $\lambda \to \kappa$ as $z \to 0$ (compare (2.22) and (3.5)). We shall not use this convention in the present work.

Comment on Dirac's Procedure: It is of interest to look into the ingeneous representation which Dirac introduced in his original paper[13] for the Dirac–Coulomb problem. It will help us to understand a certain feature of the Dirac–Coulomb problem. Based on the facts that α_r and β anticommute with one another and commute with the constant of motion K and the rest of the variables occurring in H of (2.2), Dirac assumed the existence of "a canonical transformation" which brings α_r into ρ_2 without changing β and the other variables in H. The canonical transformation automatically maps off the angular dependence of the Dirac equation in a subspace where $K = \kappa$ and hence leads to a set of radial equations, which Dirac has solved. Evidently, this approach does not give the angular part of the wavefunction as it stands unless the canonical transformation is explicitly known. So far, however, this has never been reported to the knowledge of the author. Thanks to the lemma and its corrollary introduced in I, it is now a simple matter to write down such a unitary matrix which brings the desired transformation since α_r and β are involutional. In the following we shall discuss some of the properties of this transformation.

Let us define a unitary matrix U_D by a product of two IUH (involutional, unitary, and Hermitian) matrices as follows

$$U_D = \rho_2(\alpha_r + \rho_2)/\sqrt{2} = (1 - i\beta\Sigma_r)/\sqrt{2}. \qquad (2.23)$$

Then, we have immediately

$$U_D\, \alpha_r\, U_D^{\dagger} = \rho_2, \quad U_D \beta\, U_D^{\dagger} = \beta. \qquad (2.24)$$

By direct calculations, one can show that the transformation reduces the Dirac–Coulomb wave $\psi^\mu_{E,\kappa}$ of (2.17.1) into a direct product of the radial and angular parts, as expected,

$$\psi_D = U_D\, \psi^\mu_{E,\kappa} = \begin{pmatrix} g_\kappa \\ f_\kappa \end{pmatrix} \otimes \chi^\mu_{\kappa,D}, \qquad (2.25)$$

where, $\chi^\mu_{\kappa,D}$ is a two component spinor defined by

$$\chi^\mu_{\kappa,D} = (\chi^\mu_\kappa + i\chi^\mu_{-\kappa})/\sqrt{2}. \qquad (2.26)$$

An analogous expression holds for $\phi_D = U_D\, \phi^\mu_{E,\kappa}$. The Dirac operator K, on which the angular part of $\psi^\mu_{E,\kappa}$ depends, is also reduced into the form

$$K_D = U_D\, K\, U_D^{\dagger} = i\Sigma_r(\mathbf{\Sigma}\cdot\mathbf{L} + 1), \qquad (2.27)$$

which is consistent with the direct product form of ψ_D. The transform of Y_λ takes the form,

$$Y_{\lambda D} = U_D Y_\lambda U_D^{\dagger} = (\kappa/\lambda)^{1/2}(c_2\beta - is_2\rho_2). \qquad (2.28)$$

With use of this expression, we can read off (2.17.2) from (2.17.1). Analogously, we can read off (2.12.1) from (2.10) written in the representation U_D. There is no doubt that Dirac's representation simplifies the actual calculations, if not essential. It can be very useful, particularly since the transformation U_D is now known.

3. SPECIAL CASES

We shall discuss some of the special cases of the general solution given by (2.17):

(i) When $|E| = m\lambda/\kappa$, the minimum value of $|E|$ for a given value of $|\kappa|$, we have

$$\eta = |\lambda|, \quad n_r = 0, \quad \kappa < 0, \quad \lambda < 0,$$

so that the Laguerre polynomial reduces to unity, and $c_1 = 1, s_1 = 0$. Thus,

$$\psi^\mu_{E,\kappa<0} \to \left[\frac{(2q)^3}{\Gamma(2|\lambda|+1)}\right]^{1/2} (2qr)^{|\lambda|-1} e^{-qr} \begin{pmatrix} c_2\chi^\mu_\kappa \\ is_2\chi^\mu_{-\kappa} \end{pmatrix}. \qquad (3.1)$$

For these energy levels, $\kappa > 0$ is forbidden.

(ii) In the nonrelativistic limit, z^2 can be neglected compared to 1 so that

$$\lambda = \kappa, \quad \eta = l(\kappa) + 1 + n_r, \quad c_1 = c_2 = 1, \quad s_1 = s_2 = 0.$$

Thus,

$$\psi^\mu_{E,\kappa} \to R_{E,l(\kappa)} \begin{pmatrix} \chi^\mu_\kappa \\ 0 \end{pmatrix}, \qquad (3.2)$$

which is indeed the well-known classical limit expressed in a compact manner.

(iii) In the case of a free Dirac particle, there remain only the continuum solutions. We let $z = 0$, then $\eta = -i\zeta = 0$, $\lambda = \kappa$, and

$$R_{k,l(\kappa)}(z=0) = (2/\pi)^{1/2} k j_{l(\kappa)}(kr), \qquad (3.3)$$

where $j_{l(\kappa)}(kr)$ is the spherical Bessel function. Thus,

$$\psi^\mu_{E,\kappa}(z=0) = (2/\pi)^{1/2} k \begin{pmatrix} c_1(0) j_{l(\kappa)} \chi^\mu_\kappa \\ is_1(0) j_{l(-\kappa)} \chi^\mu_{-\kappa} \end{pmatrix}, \qquad (3.4)$$

where $c_1(0)$ and $s_1(0)$ have already been defined in (2.20.2). The asymptote of this solution follows from (2.20.1) or directly from the asymptote of the spherical Bessel function,

$$j_{l(\kappa)}(kr) \sim (kr)^{-1} \sin(kr - (\pi/2)l(\kappa)). \qquad (3.5)$$

(iv) The approximation $\lambda \to \kappa$ corresponds in energy wise to the approximation due to Sommerfeld and Maue. The quantum number η takes the nonrelativistic integral values $l(\kappa) + 1 + n_r$, while the energy E is still given by the relativistic expression (2.14.2). The solution (2.17) becomes

$$\psi^{\mu}_{E,\kappa} \to \begin{pmatrix} c_1(0) R_{E,l(\kappa)} \chi^{\mu}_{\kappa} \\ i s_1(0) R_{E,l(-\kappa)} \chi^{\mu}_{-\kappa} \end{pmatrix}, \qquad (3.6)$$

which is very similar to $\phi^{\mu}_{E,\kappa}$ of (2.11). This is due to the fact that $Y_\lambda \to \beta$ in the same approximation. A similar expression has been considered by Biedenharn and Swamy[14] as an eigenfunction of approximate Hamiltonian which they called the symmetrized Hamiltonian. The above approximation indicates that the present solution is also very suitable for approximate calculations as well.

4. THE COULOMB HELICITY OPERATOR

There exists one more constant of motion y which is involutional. It changes the sign of the eigenvalue κ of the Dirac operator K as follows,

$$y \psi^{\mu}_{E,\kappa} = \psi^{\mu}_{E,-\kappa}, \quad y^2 = I. \qquad (4.1)$$

Such an operator was first introduced by Johnson and Lippman[11] and later reconstructed by Biedenharn.[5] Following Biedenharn, we use the Infeld–Hull factorized from (2.12.1) and $\sigma_r \chi^{\mu}_{\kappa} = -\chi^{\mu}_{-\kappa}$ to obtain

$$y = i\rho_1 \left(\frac{1 - (mA/KH)}{1 + (mA/KH)} \right)^{1/2} \operatorname{sign}(KH). \qquad (4.2)$$

One can directly verify that this operator indeed satisfies (4.1) by using the following relations,

$$[K, \rho_1]_+ = 0, \quad [H, \rho_1]_- = 2im\rho_2,$$
$$[H, A]_- = -2\beta(KH - mA), \qquad (4.3)$$

where $[\ldots]_+$ denotes the anticommutator and $[\ldots]_-$ the commutator. The branch point of the operator y occurs when $|E| = m\lambda/\kappa$. This is related to that $\kappa > 0$ is forbidden for the minimum value $|E| = m\lambda/\kappa$. For a free particle $y \to -i(\Sigma \cdot \hat{\mathbf{p}})\operatorname{sign}(K)$, which is essentially the helicity operator. On account of this property, Biedenharn named y the Coulomb helicity operator. The expression of y given by (4.2) is the most symmetric one of the kind.

APPENDIX

For completeness, we may define the angular part of the two-component spinor χ^{μ}_{κ} by (cf. Ref. 10, p. 26)

$$\chi^{\mu}_{\kappa}(\hat{\mathbf{r}}) = -\operatorname{sign}(\kappa)\left(\frac{1}{2} - \frac{\mu}{2\kappa + 1} \right)^{1/2} Y^{\mu - 1/2}_{l(\kappa)}(\hat{\mathbf{r}}) \chi^{1/2} + \left(\frac{1}{2} + \frac{\mu}{2\kappa + 1} \right)^{1/2} Y^{\mu + 1/2}_{l(\kappa)}(\hat{\mathbf{r}}) \chi^{-1/2}, \qquad (A1)$$

where $\hat{r}$ is the unit vector $\mathbf{r}/r$. The symmetry properties of $\chi^{\mu}_{\kappa}(\hat{r})$ are

$$\sigma_r \chi^{\mu}_{\kappa}(\hat{\mathbf{r}}) = -\chi^{\mu}_{-\kappa}(\hat{\mathbf{r}}), \quad \chi^{\mu}_{\kappa}(-\hat{\mathbf{r}}) = (-1)^{l(\kappa)} \chi^{\mu}_{\kappa}(\hat{\mathbf{r}}), \qquad (A2)$$
$$-i\sigma_2 \chi^{\mu}_{\kappa}(\hat{\mathbf{r}})^* = (-1)^{\mu + 1/2} \operatorname{sign}(\kappa) \chi_{\kappa}^{-\mu}(\hat{r}).$$

The normalization of the spinor $\psi^{\mu}_{E\kappa}$ which gives N_κ of (2.17.5) requires the following integrals for the radial functions $R_{\eta, l(\lambda)}$ and $R_{k, l(\lambda)}$ belonging to the discrete and continuum spectra respectively,

$$\int_0^\infty R_{\eta, l(\lambda)} R_{\eta, l(-\lambda)} r^2 \, dr = -(1 - (\lambda/\eta)^2)^{1/2}$$
$$\int_0^\infty R_{k', l(\lambda)} R_{k'', l(-\lambda)} r^2 \, dr = -\operatorname{sign}(zE)(1 + (\lambda/\zeta)^2)^{-1/2} \delta(k' - k'') \qquad (A3)$$

and the identity

$$1 - (m\lambda/E\kappa)^2 = (z/\kappa)^2 (1 - (\lambda/\eta)^2) \geqslant 0. \qquad (A4)$$

In terms of this identity, N_λ of (2.12.1), can be written in the form

$$N_\lambda = \lambda^{-1} |\kappa E| [1 - (\lambda/\eta)^2]^{1/2}. \qquad (A5)$$

ACKNOWLEDGMENTS

The author wishes to thank Professor J. Franklin at Temple University and Professor S. S. Yun at Ohio State University, Athens, for helpful suggestions.

[1] Shoon K. Kim, J. Math. Phys. **21**, 2282 (1980).
[2] Shoon K. Kim, J. Math. Phys. **21**, 1299 (1980).
[3] G. Temple, Proc. R. Soc. London Ser. A **127**, 349 (1930); *The General Principles of Quantum Mechanics* (Methuen, New York, 1948), p. 92.
[4] P. C. Martin and R. J. Glauber, Phys. Rev. **109**, 1307 (1958).
[5] L. C. Biedenharn, Phys. Rev. **126**, 845 (1962).
[6] E. A. Hylleraas, Z. Phys. **140**, 626 (1955); **164**, 493 (1961).
[7] L. Infeld and T. E. Hull, Rev. Mod. Phys. **23**, 21 (1951).
[8] C. G. Darwin, Proc. R. Soc. London, Ser. A **118**, 654 (1928).
[9] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One and Two Electron Atoms*, (Academic, New York, 1957), p. 68, and references therein.
[10] M. E. Rose, *Relativistic Electron Theory* (Wiley, New York, 1961), p. 157.
[11] M. E. Johnson and B. A. Lippman, Phys. Rev. **78**, 329 (1950).
[12] *Higher Transcendental Functions*, Bateman Manuscript Project, (McGraw-Hill, New York, 1953), Vol. 1, Chapter VI.
[13] P. A. M. Dirac, Proc. R. Soc. Ser. A **117**, 610 (1928); *The Principles of Quantum Mechanics*, 3rd ed. (Clarendon, Oxford, England, 1947).
[14] L. C. Biedenharn and N. V. V. J. Swamy, Phys. Rev. B **133**, 1353 (1964).

Global existence of Maxwell–Klein–Gordon fields in (2 +1)-dimensional spacetime [a)]

Vincent Moncrief
Department of Physics, Yale University, 217 Prospect St. New Haven, Connecticut 06520

(Received 25 February 1980; accepted for publication 18 April 1980)

We study the global existence problem for the Maxwell–Klein–Gordon equations in (2 +1)-dimensional, Minkowski spacetime. We first establish local existence, in a suitable Sobolev space, by specializing to the Lorentz gauge and applying standard techniques. We then prove global existence by showing that an appropriate norm of the solutions cannot blow up in a finite time. An essential step in the proof involves showing that a certain second order "energy" does not blow up.

I. INTRODUCTION

In this paper we shall study the global existence problem for the (classical) Maxwell–Klein–Gordon (MKG) equations in (2 +1)-dimensional spacetime. We first establish local existence (in a suitable Sobolev space) by specializing to the Lorentz gauge and applying standard techniques. We then extend this result to a global one by proving that an appropriate norm of the potentials and their velocities does not blow up in a finite time.

Our local existence argument requires that we work in the $H_2 \times H_1$ space of potentials and their time derivatives (here H_s is the Sobolev space of functions in L^2 with derivatives up to order s also in L^2). To extend the local result to a global one requires that we prove that the $H_2 \times H_1$ norm of a solution does not blow up in a finite time. Energy conservation is clearly insufficient for this purpose since, even with suitable gauge conditions, the energy could at most bound the $H_1 \times L^2$ norm of a solution. We shall show, however, that a suitably defined "pseudoenergy", though not strictly conserved, does not blow up in a finite time. This pseudoenergy will provide the needed bound on the higher derivatives in the $H_2 \times H_1$ norm.

Our global argument is somewhat indirect for the following reason. The local argument is facilitated by using the Lorentz gauge but the energy and the pseudoenergy provide bounds on only the "transverse parts" of the fields. This follows from the fact that both the energy and the pseudoenergy are gauge invariant. To sidestep this complication, we shall use the no blowup result for the pseudoenergy to show that the H_1 norms of the electric charge and current densities do not blow up in a finite time. Since these quantities are gauge invariant, the no blowup result holds, in particular, in the Lorentz gauge. Using this no blowup result for the currents, one can then show that the $H_2 \times H_1$ norm of the Lorentz gauge potentials does not itself blow up.

In Sec. II we derive our local existence result for the MKG equations in the Lorentz gauge and define the "Coulomb transform" of any particular Lorentz gauge solution. In Sec. III we define the pseudoenergy and prove that it does not blow up in a finite time. We extract from this result a number of bounds upon the Coulomb transform of any given solution and use these to show that the H_1 norms of the charge and current densities do not blow up in a finite time.

[a)]Research supported in part by NSF grant PHY76-82353.

In Sec. IV we complete the global existence proof by showing that the $H_2 \times H_1$ norm of a (Lorentz gauge) solution cannot blow up in a finite time.

II. LOCAL EXISTENCE

The Maxwell–Klein–Gordon field consists of a vector potential A_μ and a complex scalar field ψ (with complex conjugate ψ^*). To write the MKG equations in first order form, we define the momenta

$$P_\mu = \partial_t A_\mu, \quad \pi^* = \partial_t \psi + ie A_0 \psi, \quad \pi = \partial_t \psi^* - ie A_0 \psi^*, \tag{2.1}$$

where e is the coupling constant (charge) of the scalar field. In terms of these, the electric field **E** is given by

$$E^i = \partial_i A_0 - \partial_t A_i = \partial_i A_0 - P_i \tag{2.2}$$

and, in the Hamiltonian formalism, $(A_i, -E^i)$, (ψ, π), and (ψ^*, π^*) are canonically conjugate pairs.

We introduce the notation

$$\phi = (A_0, P_0, A_1, P_1, A_2, P_2, \psi, \pi^*), \tag{2.3}$$

(regarding ϕ below as a column matrix) and define the linear differential operator $\widetilde{A}$ by

$$\widetilde{A} = i \begin{vmatrix} \Gamma & 0 & 0 & 0 \\ 0 & \Gamma & 0 & 0 \\ 0 & 0 & \Gamma & 0 \\ 0 & 0 & 0 & \Gamma \end{vmatrix}, \quad \Gamma = \begin{pmatrix} 0 & I \\ \Delta - m^2 & 0 \end{pmatrix}, \tag{2.4}$$

where $\Delta - m^2 = \partial_i \partial_i - m^2$ for a positive (mass) constant m and where each zero in $\widetilde{A}$ stands for a 2×2 matrix of zeros. The MKG evolution equations (in Lorentz gauge) may be expressed as

$$\frac{d\phi(t)}{dt} = -i \widetilde{A} \phi(t) + J(\phi(t)) \tag{2.5}$$

where J is the nonlinear map given by

$$J(\phi) = \begin{bmatrix} 0 \\ m^2 A_0 + ie(\psi^* \pi^* - \psi \pi) \\ 0 \\ m^2 A_1 + ie(\psi^* \nabla_1 \psi - \psi \nabla_1 \psi^*) \\ 0 \\ m^2 A_2 + ie(\psi^* \nabla_2 \psi - \psi \nabla_2 \psi^*) \\ -ie A_0 \psi \\ +ie\, \partial_j(A_j \psi) + ie A_j \nabla_j \psi - ie A_0 \pi^* \end{bmatrix}, \tag{2.6}$$

where $\nabla_j \psi$ is the gauge covariant derivative

$$\nabla_j \psi = \partial_j \psi + ie A_j \psi,$$
$$\nabla_j \psi^* = (\nabla_j \psi)^* = \partial_j \psi^* - ieA_j \psi^*. \tag{2.7}$$

Note that we have added and subtracted a mass term $m^2 A_\mu$ to the $\partial_t P_\mu$ components of the Maxwell equations. This procedure regularizes the linear operator $\widetilde{A}$ (by ensuring a bounded inverse) but leaves the full Maxwell field massless as required.

We supplement the evolution equations with two initial value constraints: the Lorentz condition

$$\partial_\mu A^\mu = \partial_i A_i - P_0 = 0 \tag{2.8}$$

and the Gauss equation

$$\Delta A_0 - \partial_i P_i = ie(\psi \pi - \psi^* \pi^*), \tag{2.9}$$

which may be reexpressed as

$$\partial_i E^i = ie(\psi \pi - \psi^* \pi^*). \tag{2.10}$$

We shall show below that these two conditions are preserved by the evolution equations.

As a "rough phase space" for the field ϕ we take the Hilbert space $\mathscr{H}$ of all $A_\mu \in H_2$, $\psi \in H_2$, $P_\mu \in H_1$, and $\pi^* \in H_1$, with the inner product on $\mathscr{H}$ defined by

$$\langle \phi, \phi \rangle_{\mathscr{H}} = ||\phi||^2_{\mathscr{H}} = \sum_\mu (||A_\mu||^2_{H_2} + ||P_\mu||^2_{H_1}) + ||\psi||^2_{H_2} + ||\pi^*||^2_{H_1}, \tag{2.11}$$

where H_s is the Sobolev space of functions f for which

$$\langle f,f \rangle_{H_s} = ||f||^2_{H_s} = \int_{R^3} d^2x \{f^*f + \partial_i f^* \partial_i f + \cdots + (\partial_{i_1} \cdots \partial_{i_s} f^* \partial_{i_1} \cdots \partial_{i_s} f)\} < \infty. \tag{2.12}$$

One can show that $\widetilde{A}$ is a self-adjoint operator on $\mathscr{H}$ with a domain $D(\widetilde{A})$ consisting of all $A_\mu \in H_3$, $\psi \in H_3$, $P_\mu \in H_2$ and $\pi^* \in H_2$. Furthermore, one can show, by means of the Sobolev estimates discussed below, that J maps $D(\widetilde{A})$ to $D(\widetilde{A})$. We may thus apply the general methods of Browder[1] and Segal[2] to investigate the local and global existence of solutions. In the following we shall appeal to the formulation of this general theory given by Reed and Simon.[3]

For local existence and uniqueness of solutions it suffices to verify the following inequalities for all $\phi, \eta \in D(\widetilde{A})$:

$$||J(\phi)|| \leqslant C(||\phi||)||\phi||,$$
$$||\widetilde{A}J(\phi)|| \leqslant C(||\phi||)||\widetilde{A}\phi||,$$
$$||J(\phi) - J(\eta)|| \leqslant C(||\phi||, ||\eta||)||\phi - \eta||, \tag{2.13}$$
$$||\widetilde{A}(J(\phi) - J(\eta))|| \leqslant C(||\phi||, ||\eta||, ||\widetilde{A}\phi||, ||\widetilde{A}\eta||) \times ||\widetilde{A}\phi - \widetilde{A}\eta||,$$

where each $C(\,)$ is a monotone increasing everywhere finite function of all its arguments (where we have written $||\ ||$ for $||\ ||_{\mathscr{H}}$ to simplify notation). Given these estimates, which we shall verify below, the general theory assert that for any $\phi_0 \in D(\widetilde{A})$ there is $T_- < 0$ and a $T_+ > 0$ such that (2.5) has a unique continuously differentiable $D(\widetilde{A})$-valued solution with $\phi(0) = \phi_0$. Furthermore, if on any finite interval of existence the solution $\phi(t)$ has the property that $||\phi(t)||$ is bounded from above, then $\phi(t)$ exists and is strongly differentiable for all $t \in (-\infty, \infty)$. (We have here appealed to the time reversal invariance of the MKG equations to extend the solution backwards in time.)

To verify the above Lipshitz conditions, one needs to use the Nirenberg[4]–Gagliardo[5] inequalities

$$||f||_{L^p} \leqslant K ||D^m f||^a_{L^r} ||f||^{1-a}_{L^q} \tag{2.14}$$

where $1/p = a[(1/r) - (m/n)] + (1-a)(1/q)$, $0 \leqslant a \leqslant 1$, (if $m - (n/r)$ is a non-negative integer, only $a < 1$ is allowed) and $f: R^n \to R^k$. Here K stands for a constant which depends only on the values of n, p, q, k, r, and m and which is independent of the function f. In particular, one needs (for $n = 2$) the estimates

$$||f||_{L^4} \leqslant K(||D^1 f||^{1/2}_{L^2})||f||^{1/2}_{L^2},$$
$$||f||_{L^6} \leqslant K(||D^2 f||^{1/3}_{L^2})||f||^{2/3}_{L^2}, \tag{2.15}$$

and

$$||f||_{L^\infty} = \text{ess sup}|f| \leqslant K||f||_{H_2} \tag{2.16}$$

with these and standard tools like Hölder inequality, one can show that

$$||J(\phi)|| \leqslant (C_1||\phi|| + C_2||\phi||^2)||\phi||,$$
$$||\widetilde{A}J(\phi)|| \leqslant (C_0 + C_1||\phi|| + C_2||\phi||^2)||\widetilde{A}\phi||,$$
$$||J(\phi) - J(\eta)|| \leqslant (C_0 + C_1(||\phi|| + ||\eta||) + C_2(||\phi||^2 + ||\eta||^2))||\phi - \eta||, \tag{2.17}$$
$$||\widetilde{A}J(\phi) - \widetilde{A}J(\eta)|| \leqslant (C_0 + C_1(||\widetilde{A}\phi|| + ||\widetilde{A}\eta||) + C_2(||\phi||^2 + ||\eta||^2))||\widetilde{A}(\phi - \eta)||),$$

for some positive constants C_0, C_1, and C_2 and for all ϕ, $\eta \in D(\widetilde{A})$.

To show that the supplementary initial value equations (2.8) and (2.9) are propagated by the evolution equations, we write

$$U \equiv P_0 - \partial_i A_i,$$
$$V \equiv \partial_i E^i + ie(\psi^* \pi^* - \psi \pi) = \partial_i(\partial_i A_0 - P_i) + ie(\psi^* \pi^* - \psi \pi) \tag{2.18}$$

and compute, using (2.5), that

$$\frac{d}{dt}\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix}. \tag{2.19}$$

This is just the first order form for the massless linear wave equation for a "field" U with velocity $dU/dt = V$. We note from the expressions (2.18) for U and V that $U(t) \in H_2$ and $V(t) \in H_1$ whenever $\phi(t) \in D(\widetilde{A})$. However, since the linear wave equation has unique global solutions on $H_2 \times H_1$, it follows that $U(t)$ and $V(t)$ vanish throughout the interval of existence of $\phi(t)$ provided they vanish at any instant on this interval.

In Sec. III we shall need to estimate various gauge invariant quantities constructed from the canonical variables $(A_i, -E^i, \psi, \pi)$. These estimates are facilitated by first gauge-transforming the fields to the Coulomb-gauge, estimating the transformed expressions in terms of other gauge invariant quantities, and finally reverting to the original (Lorentz) gauge. To show that this procedure is well defined, we conclude this section with a discussion of gauge transformations and define the Coulomb transform of any set of canonical fields.

At any fixed time t let $A_i \in H_2$, $\psi \in H_2$, $E^i \in H_1$, $\pi \in H_1$ and assume that λ is a real valued, locally square integrable function whose gradient $\partial_i \lambda$ is in H_2 (note that λ need not be in L^2). We define the corresponding gauge-transformed fields (indicated by a prime) at time t to be

$$A'_i = A_i + \partial_i \lambda, \quad E^{i\prime} = E^i,$$
$$\psi' = \exp(-ie\lambda)\psi \quad \pi' = \exp(ie\lambda)\pi. \tag{2.20}$$

Clearly $A'_i \in H_2$ and $E^{i\prime} \in H_1$. To show that $\pi' \in H_1$ we expand the norm

$$\|\pi'\|^2_{H_1} = \|\pi\|^2_{H_1} + ie\int_{R^2} \partial_i\lambda\,(\pi\,\partial_i\pi^* - \pi^*\partial_i\pi)$$
$$+ e^2\int_{R^2}(\partial_i\lambda)(\partial_i\lambda)\,\pi^*\pi \tag{2.21}$$

and use the estimate $\|\partial_i\lambda\|_{L^\infty} \leqslant C\|\partial_i\lambda\|_{H_2}$ to show that

$$\|\pi'\|^2_{H_1} \leqslant \|\pi\|^2_{H_1}(1 + |e|C\|\partial_i\lambda\|_{H_2})^2. \tag{2.22}$$

In a similar way one shows that

$$\|\psi'\|^2_{H_2} \leqslant K\|\psi\|^2_{H_2}(1 + e^2C^2\|\partial_i\lambda\|^2_{H_2})^2 \tag{2.23}$$

for a suitable constant K. One can clearly extend this argument to show that gauge transformations preserve $A_i \in H_s$, $\psi \in H_s$ $E^i \in H_{s-1}$, $\pi \in H_{s-1}$ (for $s \geqslant 2$) provided $\partial_i \lambda \in H_s$.

We now wish to show that any (Lorentz gauge) solution may be transformed (at any instant of its interval of existence) to the Coulomb gauge. The solutions discussed above lie in $D(\widetilde{A})$ so they yield canonical data which satisfy $A_i \in H_3$, $E^i \in H_2$, $\psi \in H_3$, $\pi \in H_2$. A "Coulomb transform" of this data is defined to be a gauge transform of these fields which satisfies the Coulomb condition

$$\partial_i A'_i \equiv \partial_i A^c_i = 0. \tag{2.24}$$

Without an extra condition this transform will not be unique, since the function λ which generates it will be determined only up to an additive constant. We may ignore this lack of uniqueness, however, since the expressions we shall need to estimate are all independent of a constant change of phase of ψ and π^*.

To show that a Coulomb transform always exists, we must show that any $A_i \in H_3$ admits a decomposition

$$A_i = A^c_i - \partial_i\lambda, \tag{2.25}$$

in which $\partial_i A^c_i = 0$ and $\partial_i\lambda \in H_3$. That any L^2 vector field (on R^2 or R^3) may be uniquely decomposed into an L^2 divergence-free field and an L^2 gradient (with the summands, in fact, always L^2 orthogonal) was shown by Ladyzhenskaya.[6] As remarked by Cantor,[7] this proof extends to the H_s case. For $s \geqslant 3$ we can give a simpler argument as follows:

Let $A_i \in H_s$ and let $\widehat{A}_i$ designate the Fourier transform of A_i. We can decompose $\widehat{A}_i$ as

$$\widehat{A}_i = \widehat{A}^c_i + \widehat{A}^l_i, \tag{2.26}$$

where

$$\widehat{A}^c_i = \left(\delta_{ij} - \frac{k_i k_j}{k\cdot k}\right)\widehat{A}_j,$$
$$\widehat{A}^l_i = \frac{k_i k_j}{k\cdot k}\widehat{A}_j. \tag{2.27}$$

Since

$$\int_{R^2} d^2k\,(1+k^2)^s\widehat{A}^c_i\widehat{A}^c_i \leqslant \int_{R^2} d^2k\,(1+k^2)^s\widehat{A}_i\widehat{A}_i,$$
$$\int_{R^2} d^2k\,(1+k^2)^s\widehat{A}^l_i\widehat{A}^l_i \leqslant \int_{R^2} d^2k\,(1+k^2)^s\widehat{A}_i\widehat{A}_i, \tag{2.28}$$
$$\int_{R^2} d^2k\,\widehat{A}^c_i\widehat{A}^l_i = 0,$$

it follows that A^c_i and A^l_i (the inverse transforms of $\widehat{A}^c_i$ and $\widehat{A}^l_i$, and respectively) obey

$$A^c_i \in H_s,\ A^l_i \in H_s,$$
$$\int_{R^2} d^2x\,(A^c_i A^l_i) = 0, \tag{2.29}$$

whenever $A_i \in H_s$. Furthermore, the equations

$$k_i\widehat{A}^c_i = 0, \quad k_i\widehat{A}^l_j - k_j\widehat{A}^l_i = 0 \tag{2.30}$$

imply that

$$\partial_i A^c_i = 0, \quad \partial_i A^l_j - \partial_j A^l_i = 0, \tag{2.31}$$

and we recall (from the Sobolev embedding lemma) that A^c_i and A^l_i are both at least C^1 for $s \geqslant 3$. Since A^l_i has vanishing curl (i.e., $\mathbf{A}^l = A^l_i dx^i$ has vanishing exterior derivative) one may construct a C^2 function λ with gradient $\partial_i\lambda = A^l_i$ by means of the argument used in proving the Poincaré lemma. This λ is determined from A_i only up to an arbitrary additive constant, as we have mentioned.

We shall designate the Coulomb transform of (A_i, E^i, ψ, π) by $(A^c_i, E_i, \psi^c, \pi^c)$ or, whenever it's clear from the context which gauge is intended, by (A^c_i, E_i, ψ, π).

III. ENERGY INEQUALITIES

The conserved, gauge invariant energy for the MKG field is

$$E = \int_{R^2} d^2x\{\pi^*\pi + \tfrac{1}{2}E^iE^i + m^2\psi^*\psi$$
$$+ \tfrac{1}{4}(\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i)$$
$$+ (\partial_j\psi^* - ieA_j\psi^*)(\partial_j\psi + ieA_j\psi)\}. \tag{3.1}$$

For any particular solution we have $E = E_0 =$ const and thus get *a priori* bounds on the L^2 norms of ψ, π, E^i, $F_{ij} = \partial_i A_j - \partial_j A_i$, and $\nabla_j\psi = \partial_j\psi + ieA_j\psi$.

Now consider the gauge invariant pseudo-energy $\mathscr{E}_2$ defined by

$$\mathscr{E}_2 \equiv \int_{R^2} d^2x\{(\nabla_j\pi^*)(\nabla_j\pi) + \tfrac{1}{2}(\partial_j E^i)(\partial_j E^i)$$
$$+ m^2(\nabla_j\psi^*)(\nabla_j\psi) + (\nabla_j\nabla_i\psi^*)(\nabla_j\nabla_i\psi)$$
$$+ \tfrac{1}{4}(\partial_j F_{ik})(\partial_j F_{ik})\}, \tag{3.2}$$

where

$$\nabla_i\pi = \partial_i\pi - ieA_i\pi,$$

$$\nabla_j \nabla_i \psi = (\partial_j + ieA_j)[(\partial_i + ieA_i)\psi)], \tag{3.3}$$

$$\nabla_i \pi^* = (\nabla_i \pi)^*, \quad \nabla_j \nabla_i \psi^* = (\nabla_j \nabla_i \psi)^*.$$

Both $\mathscr{C}_1 \equiv E$ and $\mathscr{C}_2$ converge for any $\{(\psi, \mathbf{A}), (\pi, \mathbf{E})\} \in H_2 \times H_1$ and we shall show that $\mathscr{C}_2$ does not blow up in a finite time.

Using the equations of motion, one computes that

$$\begin{aligned}\frac{d\mathscr{C}_2}{dt} = \int_{R^2} d^2x \{ & ie[(\nabla_j \pi^*) E^j \pi - (\nabla_j \pi) E^j \pi^*] \\ & + ie\, \partial_j E^i [(\nabla_j \psi)(\nabla_i \psi^*) - (\nabla_j \psi^*)(\nabla_i \psi)] \\ & + iem^2 [E^j \psi^* (\nabla_j \psi) - E^j \psi (\nabla_j \psi^*)] \\ & + 2ieE^j [(\nabla_i \psi^*)(\nabla_j \nabla_i \psi) - (\nabla_i \psi)(\nabla_j \nabla_i \psi^*)] \\ & + ieF_{jk} [(\nabla_j \pi)(\nabla_k \psi) - (\nabla_j \pi^*)(\nabla_k \psi^*)] \\ & + \tfrac{1}{2} e^2 F_{jk} F_{jk} (\pi \psi + \pi^* \psi^*)\}. \end{aligned} \tag{3.4}$$

Each of the terms of the right-hand side is separately gauge invariant. Therefore, to estimate any of these terms, we may evaluate it in the Coulomb gauge as discussed in Sec. II. If, as we shall see, each term is estimable in terms of gauge invariant quantities, then we may revert to the Lorentz gauge at the end of the estimate. This procedure will give us a differential inequality for the function $\mathscr{C}_2$.

We shall need a bound on the L^2 norm of $\mathbf{A}^c(t)$, the Coulomb transform of $\mathbf{A}(t)$. To derive this, consider the integral

$$||\widetilde{\mathbf{A}}(t)||^2_{L^2} \equiv \int_{R^2} d^2x (A_i(t) A_i(t) + A_0(t) A_0(t)), \tag{3.5}$$

defined for A_μ in the Lorentz gauge. Computing the time derivative of this integral and using the Lorentz gauge condition to reexpress the result, we obtain

$$\begin{aligned}\frac{d}{dt}(||\widetilde{\mathbf{A}}(t)||^2_{L^2}) &= -2 \int_{R^2} d^2x (A_i E^i) \\ &\leqslant 4 ||\mathbf{A}||_{L^2} ||\mathbf{E}||_{L^2} \leqslant 4(2E_0)^{1/2} ||\widetilde{\mathbf{A}}(t)||_{L^2},\end{aligned} \tag{3.6}$$

where E_0 is the energy of the particular solution considered. It follows that, for $t \geqslant 0$,

$$||\widetilde{\mathbf{A}}(t)||_{L^2} \leqslant ||\widetilde{\mathbf{A}}(0)||_{L^2} + 2(2E_0)^{1/2} t. \tag{3.7}$$

However, $\mathbf{A}^c(t)$, the Coulomb transform of $\mathbf{A}(t)$, is simply the transverse part of $\mathbf{A}(t)$ in the sense of the decomposition (2.25) and thus satisfies

$$||\mathbf{A}^c(t)||_{L^2} \leqslant ||\mathbf{A}(t)||_{L^2} \leqslant ||\widetilde{\mathbf{A}}(t)||_{L^2}. \tag{3.8}$$

It follows that

$$||\mathbf{A}^c(t)||_{L^2} \leqslant ||\widetilde{\mathbf{A}}(0)||_{L^2} + 2(2E_0)^{1/2} t. \tag{3.9}$$

In fact, it is always possible (using a gauge transformation which does not disturb the Lorentz condition) to impose the initial condition

$$\mathbf{A}(0) = \mathbf{A}^c(0), \quad A_0(0) = 0, \tag{3.10}$$

So that one may also write

$$||\mathbf{A}(t)||_{L^2} \leqslant ||\mathbf{A}^c(0)||_{L^2} + 2(2E_0)^{1/2} t \tag{3.11}$$

Recall from (2.16) that the L^∞ norm of $\mathbf{A}^c$ obeys

$$\begin{aligned}||\mathbf{A}^c(t)||_{L^\infty} &\leqslant C ||\mathbf{A}^c(t)||_{H_2} \\ &\leqslant C \{ ||\mathbf{A}^c||^2_{L^2} + \int_{R^2} d^2x (\partial_j A^c_i)(\partial_j A^c_i) \\ &\quad + \int_{R^2} d^2x (\partial_j \partial_k A^c_i)(\partial_j \partial_k A^c_i)\}^{1/2}.\end{aligned} \tag{3.12}$$

Using the Coulomb gauge condition $\partial_i A^c_i = 0$ to reexpress the integrals (e.g., $\int_{R^2} d^2x (\partial_j A^c_i \partial_j A^c_i) = \frac{1}{2} \int_{R^2} d^2x (\partial_j A^c_i - \partial_i A^c_j)(\partial_j A^c_i - \partial_i A^c_j)$), one finds that

$$||\mathbf{A}^c(t)||_{L^\infty} \leqslant C \{ ||\mathbf{A}^c(t)||^2_{L^2} + 2E_0 + 2\mathscr{C}_2 \}^{1/2}, \tag{3.13}$$

and thus, recalling (3.11), that

$$\begin{aligned}||\mathbf{A}^c(t)||_{L^\infty} &\leqslant 2^{1/2} C \{ (||\mathbf{A}^c(0)||^2_{L^2} + E_0) + 8E_0 t^2 \\ &\quad + \mathscr{C}_2(t)\}^{1/2} \\ &\leqslant K \{ D + D' t^2 + \mathscr{C}_2(t)\}^{1/2},\end{aligned} \tag{3.14}$$

where D and D' are positive constants depending upon the values of E_0 and $||\mathbf{A}^c(0)||_{L^2}$.

We may now proceed to estimate the terms on the right-hand side of Eq. (3.4). The terms in the first bracket may be estimated as follows:

$$\left| \int_{R^2} d^2x \, [(\nabla_j \pi) E^j \pi^*] \right| \leqslant 2 ||\pi||_{L^4} ||\mathbf{E}||_{L^4} \times \left(\int_{R^2} d^2x (\nabla_j \pi)(\nabla_j \pi^*) \right)^{1/2},$$

$$||\mathbf{E}||_{L^4} \leqslant C \left(\tfrac{1}{2} \int_{R^2} d^2x (\partial_j E^i)(\partial_j E^i) \right)^{1/4} \times \left(\tfrac{1}{2} \int_{R^2} d^2x \, E^k E^k \right)^{1/4} \leqslant C E_0^{1/4} \mathscr{C}_2^{1/4}, \tag{3.15}$$

$$||\pi||_{L^4} \leqslant C E_0^{1/4} \left(\int_{R^2} d^2x (\partial_i \pi)(\partial_i \pi^*) \right)^{1/4},$$

where we have used the Nirenberg–Gagliardo inequalities (2.14) twice. Now, reexpressing $\partial_i \pi$ as $(\nabla_i \pi + ie A^c_i \pi)$, we have

$$\begin{aligned}\int_{R^2} d^2x (\partial_j \pi \partial_i \pi^*) &\leqslant 2 \left\{ \left(\int_{R^2} d^2x (\nabla_j \pi \nabla_j \pi^*) \right)^{1/2} + \left(e^2 \int_{R^2} d^2x \, A^c_i A^c_i \pi^* \pi \right)^{1/2} \right\}^2 \\ &\leqslant 4 \left\{ \left(\int_{R^2} d^2x (\nabla_i \pi^*)(\nabla_i \pi) \right) + e^2 (||\mathbf{A}^c||_{L^\infty})^2 E_0 \right\} \\ &\leqslant 4 \{ \mathscr{C}_2 + e^2 E_0 K^2 (D + D' t^2 + \mathscr{C}_2) \},\end{aligned} \tag{3.16}$$

where we have used Eq. (3.14) to estimate $||\mathbf{A}^c||_{L^\infty}$. Combining Eqs. (3.15) and (3.16) one thus gets

$$\left| \int_{R^2} d^2x \, [(\nabla_j \pi) E^{j\pi^*}] \right| \leqslant C E_0^{1/2} \mathscr{C}_2^{3/4} [\mathscr{C}_2 (1 + e^2 E_0 K^2) + e^2 E_0 K^2 (D + D' t^2)]^{1/4}. \tag{3.17}$$

In a completely analogous way one shows that

$$\begin{aligned}&\left| \int_{R^2} d^2x (\partial_j E^i)(\nabla_j \psi^*)(\nabla_i \psi) \right| \\ &\leqslant C \mathscr{C}_2^{1/2} E_0^{1/2} \left(\int_{R^2} d^2x (\partial_j \nabla_i \psi)(\partial_j \nabla_i \psi^*) \right)^{1/2} \\ &\leqslant C \mathscr{C}_2^{1/2} E_0^{1/2} [\mathscr{C}_2 (1 + e^2 E_0 K^2) + e^2 E_0 K^2 (D + D' t^2)]^{1/2},\end{aligned}$$

$$\left| \int_{R^2} d^2x (E^j (\nabla_j \psi^*) \psi \right|$$

$$\leqslant CE_0^{5/4}\mathscr{E}_2^{1/4}\{1+K'[\mathscr{E}_2+(D+D't^2)]^{1/2}\}^{1/2},$$

$$\left|\int_{R^2} d^2x\,((\nabla_j\nabla_i\psi^*)E^j(\nabla_i\psi))\right|$$
$$\leqslant CE_0^{1/2}\mathscr{E}_2^{3/4}\{\mathscr{E}_2(1+E_0e^2K^2)+E_0K^2e^2(D+D't^2)\}^{1/4},$$

$$\left|\int_{R^2} d^2x\,(\nabla_j\pi)(\nabla_k\psi)F_{jk}\right|$$
$$\leqslant CE_0^{1/2}\mathscr{E}_2^{3/4}\{\mathscr{E}_2(1+E_0e^2K^2)+e^2K^2E_0(D+D't^2)\}^{1/4}, \tag{3.18}$$

and finally that

$$\left|\int_{R^2} d^2x\,(F_{jk}F_{jk}\,\pi\psi)\right|$$
$$\leqslant CE_0^{5/4}\mathscr{E}_2^{1/2}\{\mathscr{E}_2(1+e^2K^2E_0)+e^2K^2E_0(D+D't^2)\}^{1/4}\times\{1+K'[(D+D't^2)+\mathscr{E}_2]^{1/2}\}^{1/2}. \tag{3.19}$$

It follows that

$$\frac{d\mathscr{E}_2(t)}{dt}\leqslant K_0+K_1t^2+K_2\mathscr{E}_2(t) \tag{3.20}$$

where K_0, K_1, and K_2 are positive constants which depend on the fields only through the values of E_0 and $||\mathbf{A}^c(0)||_{L^4}$. Thus, for $t\geqslant 0$

$$\mathscr{E}_2(t)\leqslant\mathscr{E}_2(0)+K_0t+\tfrac{1}{3}K_1t^3+K_2\int_0^t dt'\,\mathscr{E}_2(t'), \tag{3.21}$$

and from Gronwall's inequality it follows that

$$\mathscr{E}_2(t)\leqslant\left[\mathscr{E}_2(0)+K_0t+\frac{K_1}{3}t^3\right]+\int_0^t dt'\,K_2|\mathscr{E}_2(0)+K_0t'+\tfrac{1}{3}K_1t'^3|\times\exp[K_2(t-t')]. \tag{3.22}$$

Thus $\mathscr{E}_2$ does not blow up in a finite time $t\geqslant 0$. (A similar argument holds for $t\leqslant 0$ by virtue of the time reversal invariance of the MKG equations.)

From conservation of energy, the L^2 bound on $\mathbf{A}^c(t)$ given in Eq. (3.9), and the no blow-up result for $\mathscr{E}_2$ we obtain

$$||\mathbf{A}^c(t)||_{H_2}<\infty,\quad ||\mathbf{E}(t)||_{H_1}<\infty,\quad ||\psi(t)||_{L^2}<\infty,$$
$$||\pi(t)||_{L^2}<\infty\ ||\nabla_i\psi(t)||_{L^2}<\infty,$$
$$||\nabla_i\nabla_j\psi(t)||_{L^2}<\infty,\quad ||\nabla_i\pi(t)||_{L^2}<\infty, \tag{3.23}$$

for all finite t.

Recalling Eq. (3.16) we thus get for the Coulomb transformed fields, that

$$||\pi(t)||_{H_1}<\infty,$$

and, by similar argument, that

$$||\psi(t)||_{H_1}<\infty$$

and

$$\int_{R^2} d^2x\,(\partial_j\nabla_i\psi^*)(\partial_j\nabla_i\psi)(t)\leqslant C\,[\mathscr{E}_2^{1/2}(t)+|e|\,E_0^{1/2}||\mathbf{A}^c(t)||_{L^\infty}]^2<\infty. \tag{3.24}$$

A straightforward sequence of estimates gives

$$\int_{R^2} d^2x\,(\partial_j\partial_i\psi^*)(\partial_j\partial_i\psi)$$
$$\leqslant C_1\int_{R^2} d^2x\{(\partial_j\nabla_i\psi^*)(\partial_j\nabla_i\psi)+(\partial_jA_i^c)(\partial_jA_i^c)\psi^*\psi+A_i^cA_i^c(\partial_j\psi^*)(\partial_j\psi)\}$$
$$\leqslant C_2[\mathscr{E}_2^{1/2}+|e|E_0^{1/2}||\mathbf{A}^c(t)||_{L^\infty}]^2+C_3E_0||\mathbf{A}^c||_{L^\infty}^2(1+\left|\frac{e}{m}\right|\,||\mathbf{A}^c(t)||_{L^\infty})^2+C_4(||\partial_jA_i^c||_{L^4})^2(||\psi||_{L^4})^2. \tag{3.25}$$

But, using the Nirenberg–Gagliardo estimates we get

$$||\psi||_{L^4}\leqslant C'E_0^{1/4}||\partial_i\psi||_{L^2}^{1/2}\leqslant C'E_0^{1/2}\times[1+|e/m|\,||\mathbf{A}^c||_{L^\infty}]^{1/2} \tag{3.26}$$

and

$$||\partial_jA_i^c||_{L^4}\leqslant C'||\partial_j\partial_kA_i^c||_{L^2}^{1/2}||\partial_jA_i||_{L^2}^{1/2}\leqslant C'E_0^{1/4}\mathscr{E}_2^{1/2}. \tag{3.27}$$

It follows from (3.23) and (3.25)–(3.27) that

$$||\psi(t)||_{H_2}<\infty$$

for all finite t.

To summarize the results of this section, we have shown that the Coulomb–transformed fields obey

$$||\psi(t)||_{H_2}<\infty,\ ||\pi(t)||_{H_1}<\infty,\ ||\mathbf{A}^c(t)||_{H_2}<\infty,$$
$$||\mathbf{E}(t)||_{H_1}<\infty, \tag{3.28}$$

for all finite t. We also obtained the gauge invariant bounds

$$||\psi(t)||_{L^2}<\infty,\ ||\nabla_i\psi(t)||_{L^2}<\infty,\ ||\nabla_i\nabla_j\psi(t)||_{L^2}<\infty,$$
$$||\pi(t)||_{L^2}<\infty,\ ||\nabla_i\pi(t)||_{L^2}<\infty, \tag{3.29}$$

for all finite t.

To complete the global existence argument, we shall need to show that the H_1 norms of the charge density $ie(\psi\pi-\psi^*\pi^*)$ and current density $ie(\psi\nabla_i\psi^*-\psi^*\nabla_i\psi)$ do not blow up. Since these quantities are gauge invariant, it suffices to estimate their H_1 norms in the Coulomb gauge. We have

$$||(\psi\pi-\psi^*\pi^*)||_{H_1}^2$$
$$\leqslant 4\int_{R^2} d^2x\,|\psi|^2|\pi|^2+8\int_{R^2} d^2x\,[|\partial_i\psi|^2|\pi|^2+|\psi|^2|\partial_i\pi|^2]$$
$$\leqslant 4E_0||\psi||_{L^\infty}^2+8||\psi||_{L^\infty}^2\int_{R^2} d^2x|\partial_i\pi|^2+8||\pi||_{L^4}^2||\partial_i\psi||_{L^4}^2$$
$$\leqslant C||\psi||_{H_2}^2(E_0+2||\partial_i\pi||_{L_2}^2)+C'E_0^{1/2}||\partial_i\pi||_{L^2}\times||\partial_k\partial_i\psi||_{L^2}||\partial_j\psi||_{L^2}. \tag{3.30}$$

Thus $||(\psi\pi-\psi^*\pi^*)(t)||_{H_1}$ is bounded on bounded time intervals by virtue of the forgoing bounds on $||\psi(t)||_{H_2}$ and $||\pi(t)||_{H_1}$.

In a similar way we get

$$||\psi\nabla_i\psi^*-\psi^*\nabla_i\psi||_{H_1}$$
$$\leqslant 4\int_{R^2} d^2x\,(|\psi|^2|\nabla_i\psi|^2)$$

$$+ 8\int_{R^2} d^2x\,(|\partial_j \psi|^2|\nabla_i \psi|^2 + |\psi|^2|\partial_j(\nabla_i \psi)|^2)$$
$$\leqslant 4||\psi||^2_{L^\infty}(||\nabla_i \psi||^2_{L^2} + 2||\partial_j(\nabla_i \psi)||^2_{L^2}) + 8||\partial_j \psi||^2_{L^4}||\nabla_i \psi||^2_{L^4}$$
$$\leqslant C||\psi||^2_{H_2}(E_0 + 2||\partial_j(\nabla_i \psi)||^2_{L^2}) + C'||\partial_j \psi||_{L^2}||\partial_k\partial_l \psi||_{L^2} 2E_0^{1/2}||\partial_m(\nabla_i \psi)||_{L^2}. \quad (3.31)$$

Recalling (3.24)–(3.27), we have that $||\psi(t)||_{H_2}$ and $||\partial_j(\nabla_i \psi(t))||_{L^2}$ are bounded for finite t. It follows that $||(\psi\nabla_i \psi^* - \psi^*\nabla_i \psi)(t)||_{H_1}$ is bounded on bounded time intervals.

IV. COMPLETION OF THE GLOBAL EXISTENCE PROOF

The Maxwell equations in Lorentz gauge are

$$\frac{\partial^2 A_0}{\partial t^2} - \Delta A_0 = ie(\psi^*\pi^* - \psi\pi) \equiv \mathscr{J}_0 \quad (4.1)$$

$$\frac{\partial^2 A_i}{\partial t^2} - \Delta A_i = ie(\psi^*\nabla_i \psi - \psi\nabla_i \psi^*) \equiv \mathscr{J}_i.$$

We already know from (3.7) and its time reversed extension that

$$||\widetilde{\mathbf{A}}(t)||_{L^2} \leqslant ||\widetilde{\mathbf{A}}(0)||_{L^2} + 2(2E_0)^{1/2}|t| \quad (4.2)$$

for all t. To show that the full $H_2\times H_1$ norm of $[\widetilde{\mathbf{A}}, \widetilde{\mathbf{P}}(t)]$ does not blow up, we define

$$\Gamma(t) = \int_{R^2} d^2x\,\{\dot{A}_0^2 + \dot{A}_i^2 + A_{0,i}A_{0,i} + A_{j,i}A_{j,i} + \dot{A}_{0,i}\dot{A}_{0,i} + \dot{A}_{j,i}\dot{A}_{j,i} + A_{0,ij}A_{0,ij} + A_{k,ij}A_{k,ij}\} \quad (4.3)$$

(where $\dot{A}_\mu \equiv P_\mu$) and compute

$$\frac{d\Gamma(t)}{dt} = 2\int_{R^2} d^2x\,\{\dot{A}_0\mathscr{J}_0 + \dot{A}_i\mathscr{J}_i + (\dot{A}_{0,k}(\partial_k \mathscr{J}_0) + (\dot{A}_{j,k})(\partial_k J_j)\}$$
$$\leqslant C\{||\widetilde{\mathbf{P}}(t)||_{L^2}||\widetilde{\mathscr{J}}(t)||_{L^2} + ||\partial_k\widetilde{\mathbf{P}}(t)||_{L^2}||\partial_k \widetilde{\mathscr{J}}(t)||_{L^2}\}$$
$$\leqslant C\{||\widetilde{\mathbf{P}}(t)||^2_{L^2} + ||\partial_k\widetilde{\mathbf{P}}(t)||^2_{L^2} + ||\widetilde{\mathscr{J}}(t)||^2_{L^2}||\partial_k \widetilde{\mathscr{J}}(t)||^2_{L^2}\}.$$
$$\leqslant C\{\Gamma(t) + ||\widetilde{\mathscr{J}}(t)||^2_{H_1}\}. \quad (4.4)$$

From the no blow-up result of Sec. III, we know that $||\widetilde{\mathscr{J}}(t)||_{H_1}$ is bounded on finite time intervals. It follows from Gronwall's inequality that $\Gamma(t)$ cannot blow in a finite time and thus (since we already know that $||\widetilde{\mathbf{A}}(t)||_{L^2} < \infty$) that

$$||\widetilde{\mathbf{A}}(t)||_{H_2} < \infty, \quad ||\widetilde{\mathbf{P}}(t)||_{H_1} < \infty, \quad (4.5)$$

for all finite t. It follows that $||\widetilde{\mathbf{A}}(t)||_{L^\infty} < \infty$ for all finite t since $||\widetilde{\mathbf{A}}(t)||_{L^\infty} < C||\widetilde{\mathbf{A}}(t)||_{H_2}$.

To show that the $H_2\times H_1$ norm of (ψ, π^*) does not blow up (in Lorentz gauge) we define

$$\Sigma(t) = \int_{R^2} d^2x\,\{\pi^*\pi + m^2\psi^*\psi + (\partial_i \pi^*)(\partial_i \pi) + (\partial_i \psi^*)(\partial_i \psi) + (\partial_i\partial_j \psi^*)(\partial_i\partial_j \psi)\}, \quad (4.6)$$

and compute $d\Sigma(t)/dt$ using (2.5). After a lengthy but straightforward sequence of estimates, one shows that

$$\frac{d\Sigma(t)}{dt} \leqslant (C_1 + C_2||\widetilde{\mathbf{A}}(t)||^2_{H_2})\Sigma(t). \quad (4.7)$$

Since $||\widetilde{\mathbf{A}}(t)||_{H_2}$ does not blow up in a finite time, it follows (again using Gronwall's inequality) that $\Sigma(t)$ does not blow up in a finite time. This result completes our global existence proof.

An alternative argument for the last step may be given as follows. We know from Sec. III that the $H_2\times H_1$ norm of the Coulomb-transformed fields (ψ^c, π^c) does not blow up in a finite time. From Eqs. (2.22) and (2.23) we know that the corresponding norm of the Lorentz-transformed fields will not blow up provided $||\partial_i \lambda||_{H_2}$ does not blow up. However, $\partial_i \lambda$ is simply the longitudinal part of the spatial, Lorentz gauge potential A_i. That this does not blow up follows from Eq. (3.9) and the properties of the decomposition discussed in Sec. II.

ACKNOWLEDGMENTS

To a considerable extent this work was carried out in parallel with a joint effort on the Yang–Mills equations with Douglas Eardley. I have benefited greatly from many conversations with Eardley. I am also grateful to Jerrold Marsden for several helpful conversations.

[1]F. Browder, Math. Z. 80, 249 (1962).
[2]I. Segal, Ann. Math. **78**, 339 (1963).
[3]M. Reed and B. Simon, *Method of Modern Mathematical Physics. II. Fourier Analysis and Self-Adjointness* (Academic, New York, 1975).
[4]L. Nirenberg, Ann. Scuola Norm. Sup. Pisa **13**, 115 (1959).
[5]E. Gagliardo, Ric. Mat. **8**, 24 (1959).
[6]O. A. Ladyzhenskaya, *The Mathematical Theory of Viscous Incompressible Flow* (Gordon and Breach, New York, 1963).
[7]M. Cantor, Indiana. Math. J. **24**, 897 (1975).

Existence of Hartree–Fock solutions

G. Rosensteel [a)]
Department of Physics, Tulane University, New Orleans, Louisiana 70118
Edwin Ihrig
Department of Mathematics, Arizona State University, Tempe, Arizona 85281

(Received 22 January 1980; accepted for publication 18 March 1980)

For a finite-dimensional space with only a mild restriction on the Hamiltonian, it is shown that there exist at least as many Hartree–Fock states as the dimension of the many-fermion space. The index of the random phase approximation matrix is determined for these HF states and the relationship between that index and the number of real and complex excitation energies established.

1. INTRODUCTION

The Hartree–Fock (HF) self-consistent field approximation is fundamental to the theory of finite many-fermion systems. Nevertheless, because the HF equations are nonlinear, the question of existence of HF solutions is nontrivial and only recently have some results been obtained. For the simpler Hartree problem, the existence of the ground state solution (minimum minimorum) has been proved using various methods, albeit often restricted to the case of the Helium atom.[1-5] Using a more powerful method, Lieb and Simon[6] proved the existence of the HF minimum for any neutral atom or positive ion. In contrast to the above paper, we take the state space to be finite-dimensional. However, no assumption is made here restricting either the strength of the interaction or its type, i.e., two-body.

It will be shown that there are at least as many HF states as the dimension of the many-fermion state space (Theorem 1). Furthermore, the index of the RPA matrix at these HF states is given. In Sec. 3, the relationship between the index of the RPA matrix and the number of real and imaginary RPA excitation energies is obtained (Theorem 2).

The existence proof is based on Morse theory.[7] This theory places a minimum on the number of critical points of a smooth real-valued function F on a compact manifold M due to the topology of M. Specifically, if $m \in M$ is a critical point of F and $(x^1, x^2, \dots, x^n)$ is a chart about m with $x^i(m) = 0$, then

$$\begin{aligned} F(x^1,x^2,\dots,x^n) &= F(m) + \sum_{i=1}^{n} x^i \frac{\partial F}{\partial x^i}\Big|_{x=0} + \tfrac{1}{2}\sum_{i,j=1}^{n} x^i x^j \frac{\partial^2 F}{\partial x^i \partial x^j}\Big|_{x=0} + \cdots \\ &= F(m) + \tfrac{1}{2}\sum_{i,j=1}^{n} x^i x^j G_{ij} + \cdots, \end{aligned} \tag{1}$$

where the matrix $G_{ij} = (\partial^2 F / \partial x^i \partial x^j)|_{x=0}$ is called the Hessian. The index of the matrix G regarded as a bilinear form on $\mathbb{R}^n$ is the dimension of the largest subspace of $\mathbb{R}^n$ on which G is negative definite; at a critical point, this index is independent of the chart used to compute G. A critical point is nondegenerate if the Hessian is nonsingular. If the index of a nondegenerate critical point is λ, then a chart $(y^1, y^2, \dots, y^n)$ can be found for which

$$\begin{aligned} F(y^1, y^2, \dots, y^n) &= F(m) - (y^1)^2 - (y^2)^2 - \cdots - (y^\lambda)^2 + (y^{\lambda+1})^2 \\ &\quad + \cdots + (y^n)^2 + \cdots . \end{aligned} \tag{2}$$

Let C_λ be the number of critical points of F for which the index of the Hessian is λ. Then, Morse theory claims that if F has no degenerate critical points, then $C_\lambda \geqslant \beta_\lambda$, where β_λ are the Betti numbers of the manifold. Somewhat stronger inequalities can be given, see Ref. 7.

2. EXISTENCE OF HARTREE–FOCK STATES

Suppose $\mathscr{H}$ is an n-dimensional Hilbert space and $\wedge\mathscr{H}$ is the exterior product of A copies of $\mathscr{H}$. Let H be the Hamiltonian, a self-adjoint operator on $\wedge\mathscr{H}$, not necessarily the sum of one-body and two-body operators. By definition, a Hartree–Fock state is a critical point of the energy function[8,9]

$$\mathrm{E}:\mathrm{S}\to\mathbb{R}, \quad \mathrm{E}(\phi) = \langle \phi \mid \mathrm{H}\phi \rangle, \tag{3}$$

where the set of states S is the Slater determinants,

$$\mathrm{S} = \{\phi = \phi_1 \wedge \phi_2 \wedge \cdots \wedge \phi_A \mid \langle \phi_i | \phi_j \rangle = \delta_{ij}\}. \tag{4}$$

Morse theory cannot be applied directly to the energy function since every point is degenerate, $\mathrm{E}(\lambda\phi) = \mathrm{E}(\phi)$ for all $\phi \in \mathrm{S}$ and $\lambda \in \mathbb{C}$, $|\lambda| = 1$. In order to remove this obvious degeneracy, define an equivalence relation on S by $\phi \sim \psi$ for $\phi, \psi \in \mathrm{S}$ iff $\phi = \lambda\psi$ for some $\lambda \in \mathbb{C}$, $|\lambda| = 1$. The set of equivalence classes in S is identified with the complex Grassman variety $\mathbb{C}G(A, n-A)$, the set of A-dimensional hyperplanes in an n-dimensional complex vector space.[10,11] The energy function is well defined on $\mathbb{C}G(A, n-A)$,

$$\mathrm{E}:\mathbb{C}G(A, n-A) \to \mathbb{R}. \quad \mathrm{E}(\Phi) = \langle \phi \mid \mathrm{H}\phi \rangle, \tag{5}$$

where $\Phi = \{\lambda\phi \mid \phi \in \mathrm{S}, \lambda \in \mathbb{C}, |\lambda| = 1\} \in \mathbb{C}G(A, n-A)$.

Before computing the Hessian of E, coordinates must be given for the complex Grassman variety in a neighborhood of a fixed critical point Φ in $\mathbb{C}G(A, n-A)$.[12] First pick a $\phi = \phi_1 \wedge \phi_2 \wedge \cdots \wedge \phi_A \in \Phi$, and then augment the set of A

a) This material is based upon work supported by the National Science Foundation under Grant PHY-7906534.

 0022-2488/80/082297-05$1.00

states $\{\phi_1,\phi_2,\ldots,\phi_A\}$ with $(n-A)$ additional orthonormal states $\{\phi_{A+1},\phi_{A+2},\ldots,\phi_n\}$. Conventionally, the first A vectors are hole states indexed by h, h', while the last $(n-A)$ vectors are particle states indexed by p, p'. Let b_α^+ and b_α denote the fermion creation and annihilation operators for the state ϕ_α, and set

$$iQ_{ph} = (i/\sqrt{2})(b_p^+ b_h + b_h^+ b_p),$$
$$iP_{ph} = (1/\sqrt{2})(b_p^+ b_h - b_h^+ b_p). \quad (6)$$

Then, a chart containing Φ for the $2m \equiv 2A(n-A)$ dimensional space $\mathbb{C}G(A,n-A)$ is given by

$$(q_{ph}, p_{ph}) \to \exp(X)\Phi, \quad (7)$$

where

$$X = i\sum_{ph}(p_{ph}Q_{ph} - q_{ph}P_{ph}).$$

We are now able to calculate the Hessian from

$$\begin{aligned}\mathsf{E}(\exp(-X)\Phi) - \mathsf{E}(\Phi) &= \langle\phi\,|e^X H e^{-X}\phi\rangle - \langle\phi\,|\,H\phi\rangle\\ &= \langle\phi\,|\,[X,H]\phi\rangle + \tfrac{1}{2}\langle\phi\,|\,[X,[X,H]]\phi\rangle + \cdots. \quad (8)\end{aligned}$$

Hence, Φ is a critical point iff $\langle\phi\,|\,[X,H]\phi\rangle = 0$ for all. Thus, if Φ is a Hartree–Fock state,

$$\begin{aligned}\mathsf{E}(\exp(-X)\Phi) - \mathsf{E}(\Phi) &= \tfrac{1}{2}\langle\phi\,|\,[X,[X,H]]\phi\rangle + \cdots\\ &= \tfrac{1}{2}\langle\phi\,|\,[X,H,X]\phi\rangle + \cdots, \quad (9)\end{aligned}$$

where the double commutator $2[A,B,C] = [A,[B,C]] + [[A,B],C]$. Introducing the coordinates (q_{ph}, p_{ph}), the Hessian G given by

$$\mathsf{E}(\exp(-X)\Phi) - \mathsf{E}(\Phi) = \tfrac{1}{2}(q\;p)G\binom{q}{p} + \cdots \quad (10)$$

is unitarily equivalent to the random phase approximation (RPA) matrix,[13]

$$G = C\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}C^+, \quad C = \frac{1}{\sqrt{2}}\begin{pmatrix} I_m & I_m \\ iI_m & -iI_m \end{pmatrix}, \quad (11)$$

where I_m is the $m\times m$ unit matrix. The submatrices A and B of the RPA matrix are given by

$$A_{php'h'} = \langle\phi\,|\,[a_h^+ a_p, H, a_{p'}^+ a_{h'}]\phi\rangle \quad \text{(Hermitian)}$$
$$B_{php'h'} = -\langle\phi\,|\,[a_h^+ a_p, H, a_{h'}^+ a_{p'}]\phi\rangle \quad \text{(Symmetric)}. \quad (12)$$

The condition for the applicability of Morse Theory is that every critical point of the energy function be nondegenerate. However, a Hartree–Fock state is nondegenerate if and only if the determinant of the RPA matrix is nonzero or that there be no nonzero solutions v to $Gv = 0$. Such solutions v are called spurious (see Ref. 14) and are exceptional in a finite-dimensional space. Indeed, the Hamiltonians with a degenerate Hartree–Fock state form a nowhere dense set in the space of all self-adjoint operators. As a practical matter, the only case where spurious solutions arise is when an exact symmetry of the Hamiltonian is violated by the HF state. This problem is fixed, say in the case of rotational and isospin symmetry, by restricting states to fixed m and τ_z.

Theorem 1: Suppose that the Hamiltonian admits no spurious RPA states. Then there exist at least $\binom{n}{A}$ Hartree–Fock states.

Proof: The condition of the theorem is just that which permits Morse theory to be applied to the energy function. Let C_λ be the number of Hartree–Fock states for which the index of the Hessian G equals λ. Then $C_\lambda \geqslant \beta_\lambda$, where β_λ are the Betti numbers of the complex Grassman variety. These Betti numbers are known[10,11]:

$$\beta_{2k+1} = 0,$$
$$\beta_{2k} = o\Big(\Big\{(p_1,p_2,\ldots,p_A)\,|\,p_i\in\mathbb{Z}^{\geqslant 0},\ \sum_i p_i \leqslant n-A,\ \sum_i ip_i = k\Big\}\Big).$$

In order to complete the proof, it must be shown that

$$\sum_\lambda \beta_\lambda = \sum_k \beta_{2k} = \binom{n}{A},$$

or

$$\binom{n}{A} = o\Big(\Big\{(p_1,p_2,\ldots,p_A)\,|\,p_i\in\mathbb{Z}^{\geqslant 0},\ \sum_i p_i \leqslant n-A\Big\}\Big).$$

The order N of this set will be calculated from a generating polynomial. Let

$$P(x) = \sum_{p_i\geqslant 0} x^{p_1+p_2+\cdots+p_A} = (1-x)^{-A}.$$

The coefficient a_r in $P(x) = \Sigma_{r=0}^{\infty} a_r x^r$ is the number of A-tuples $(p_1,p_2,\ldots,p_A)$ with $\Sigma_i p_i = r$; hence, $N = \Sigma_{r=0}^{n-A} a_r$ is the order of the set. But

$$a_r = \frac{1}{r!}\frac{d^r}{dx^r}P(x)\Big|_{x=0} = \binom{A+r-1}{r}.$$

Now $N(s) \equiv \Sigma_{r=0}^{s} a_r$ satisfies

$$N(s+1) = N(s) + \binom{A+s}{s+1} \quad \text{and } N(0) = 1.$$

Since $\binom{s+A}{A}$ satisfies these two conditions, we have

$$N(s) = \binom{s+A}{A}.$$

Therefore, $N = N(n-A) = \binom{n}{A}$. Q.E.D.

Morse theory yields still further information about the number of critical points. One has[7]

$$\sum(-)^\lambda C_\lambda = \sum(-)^\lambda \beta_\lambda = \sum_k \beta_{2k} = \binom{n}{A} \quad (13)$$

and the strong Morse inequalities

$$C_\lambda - C_{\lambda-1} + - \cdots \pm C_0 \geqslant \beta_\lambda - \beta_{\lambda-1} + - \cdots \pm \beta_0. \quad (14)$$

From the equality,

$$\sum C_\lambda = \binom{n}{A} + 2\sum_\lambda \frac{1-(-1)^\lambda}{2}C_\lambda \quad (15)$$

and from the inequality for $\lambda = 1$,

$$C_1 \geqslant C_0 - 1, \quad (16)$$

we have

$$\sum C_\lambda \geqslant \binom{n}{A} + 2C_1 \geqslant \binom{n}{A} + 2(C_0 - 1). \qquad (17)$$

Hence, if $C_0 > 1$, then the number of HF states is greater than the dimension of the many-fermion state space, $\dim \wedge \mathscr{H} = \binom{n}{A}$.

Nuclear HF calculations for even–even nuclei in the region $4 \leqslant A \leqslant 40$ have been made by Bassichis, Kerman, and Svenne[15] with the space $\mathscr{H}$ spanned by the 1s, 1p, 2s − 1d and 2p − 1f oscillator shells. They discovered for the nuclei C^{12}, Si^{28}, and S^{32} two HF states with index zero (prolate and oblate solutions). Hence, the total number of critical points for these nuclei is at least two more than the dimension of $\wedge \mathscr{H}$.

3. RPA SOLUTIONS

The aim of this section is to determine the relationship between the index λ of the Hessian G and the spectrum of the RPA equations

$$Gv_j = -i\omega_j J v_j, \qquad (18)$$

where J is the $2m \times 2m$ symplectic form

$$J = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}. \qquad (19)$$

The solutions v_j define particle-hole excitation operators belonging to the excitation energies ω_j.

The result to be proved here applies to a more general situation than the RPA, e.g., the equations of motion of Rowe.[14] Thus, the theorem will be formulated in a general setting. First, a lemma must be proved.

Lemma: Let V be a finite-dimensional vector space and V^* its dual. Suppose $G: V \to V^*$ is symmetric and nondegenerate, and $T: V \to V$ is antisymmetric in G, i.e., $G(Tv,v') = -G(v,Tv')$, $v, v' \in V$. If W is a subspace of V invariant under T, $T(W) \subset W$, and $G|_W$ is nondegenerate, then

(a) $V = W \oplus W^\perp$;

(b) $T = T|_W \oplus T|_{W^\perp}$;

(c) $G = G|_W \oplus G|_{W^\perp}$.

Proof: (a) $W^\perp = \{ y \in V \mid G(y,w) = 0 \text{ for all } w \in W \}$. If $w \in W \cap W^\perp$, then

$$w \in W^\perp \Rightarrow G(w,w') = 0 \text{ for all } w' \in W.$$

But $w \in W$ and $G|_W$ nondegenerate $\Rightarrow w = 0$. Hence, $W \cap W^\perp = \{0\}$. The linear map

$$\rho: V \to W^*, \quad \rho(v)(w) = G(v,w)$$

is surjective with kernal $W^\perp$. Thus,

$$\dim W = \dim V - \dim W^\perp,$$

so that

$$\dim(W \oplus W^\perp) = \dim W + \dim W^\perp, \text{ and}$$

since $W \cap W^\perp = \{0\}$, it also equals $\dim V$. Therefore, $V = W \oplus W^\perp$.

(b) $T(W^\perp) \subset W^\perp$, since for $y \in W^\perp$

$$G(Ty,w) = -G(y,Tw) = 0 \text{ for all } w \in W.$$

$$\therefore \; T = T|_W \oplus T|_{W^\perp}.$$

(c) By (a), we have for all $v_1, v_2 \in V$,

$$v_1 = w_1 + y_1, \quad v_2 = w_2 + y_2, \quad w_1, w_2 \in W, \quad y_1, y_2 \in W^\perp.$$

Thus, $G(v_1,v_2) = G(w_1,w_2) + G(y_1,y_2)$.

Theorem 2: Let V be a finite-dimensional real vector space. Suppose $G: V \to V^*$ is symmetric and nondegenerate with index λ, and $J: V \to V^*$ is skew-symmetric and nondegenerate. Assume that $T = G^{-1}J: V \to V$ has a complete set of eigenvectors v_j in the complexification $C \otimes V$ belonging to the eigenvalues $\mu_j = (i/\omega_j)$, $Tv_j = \mu_j v_j$, $j = 1,2,\ldots,\dim V$. Let Λ_C denote the number of nonreal, nonimaginary ω_j; Λ_I, the number the pure imaginary ω_j; Λ_R^+, the number of real ω_j with eigenvectors v_j with positive length in G, i.e., $G(v_j,v_j) > 0$; Λ_R^-, the number of real ω_j with negative length in G; Λ_R^0, the number of real ω_j with zero length in G. Then,

(a) $\lambda = \Lambda_R^- + \frac{1}{2}(\Lambda_R^0 + \Lambda_I + \Lambda_C)$,

(b) $2 | \Lambda_R^\pm, 4 | \Lambda_R^0, 2 | \Lambda_I, 4 | \Lambda_C$.

Conversely, given an index λ and an even dimensional vector space V with $\lambda \leqslant \dim V$ together with five numbers $\Lambda_R^\pm$, Λ_R^0, Λ_I, Λ_C satisfying (a), (b), and $\dim V = \Lambda_R^+ + \Lambda_R^- + \Lambda_R^0 + \Lambda_I + \Lambda_C$, then there are maps G and J as above with these five numbers acting as $\Lambda_R^\pm$, Λ_R^0, Λ_I, Λ_C for $G^{-1}J$.

Proof: We begin with three observations. Firstly, since T is antisymmetric in G,

$$G(Tv_j,v_k) + G(v_j,Tv_k) = 0,$$

or $(\mu_j + \bar{\mu}_k)G(v_j,v_k) = 0$ for all eigenvectors v_j, v_k. Secondly, since T is real

$$Tv_j = \mu_j v_j \Rightarrow T\bar{v}_j = \bar{\mu}_j \bar{v}_j.$$

Thus, if $v_j = x_j + iy_j$, $\bar{v}_j = x_j - iy_j$ with x_j and y_j real, then

$$0 = (\mu_j + \bar{\mu}_j)G(v_j,\bar{v}_j), \\ = 2\mu_j G(v_j,\bar{v}_j).$$

But $\mu_j \neq 0$, since T is nondegenerate. Therefore,

$$0 = G(v_j,\bar{v}_j) \\ = G(x_j,x_j) - G(y_j,y_j) + 2iG(x_j,y_j),$$

or $G(x_j,x_j) = G(y_j,y_j)$ and $G(x_j,y_j) = 0$. Hence, $G(v_j,v_j) = 2G(x_j,x_j) = 2G(y_j,y_j)$. Thirdly, recall that J skew-symmetric and nondegenerate implies V is even dimensional.[16] There are two main cases to consider.

I. $G(v_j,v_j) \neq 0$.

Then $\mu_j + \bar{\mu}_j = 0$, or ω_j is real. In this case, $W = \text{span}\{x_j, y_j\}$ is invariant under T and $G|_W$ is nondegenerate. [It is either positive or negative definite depending on the sign of $G(v_j,v_j)$.] Hence, there are two real eigenvalues (ω_j and $-\omega_j$) in W and, from the lemma

$$\lambda|_V = 0 + \lambda|_{W^\perp},$$

with two eigenvalues in Λ_R^+,

$$\lambda|_V = 2 + \lambda|_{W^\perp},$$

with two eigenvalues in Λ_R^-.

If W_I is the span of all x_j, y_j for which $G(v_j,v_j) \neq 0$, then by repeated application of the above result we are able to reduce the proof to the study of $W_I^\perp$ since

$$V = W_I \oplus W_I^\perp, \quad T = T|_{W_I} \oplus T|_{W_I^\perp},$$
$$G = G|_{W_I} \oplus G|_{W_I^\perp}.$$

II. $G(v_j, v_j) = 0$

Since $G|_{W_I^\perp}$ nondegenerate, there is a $v_k \in W_I^\perp$ with $G(v_j, v_k) \neq 0$. But $(\mu_j + \bar{\mu}_k) = 0 \Rightarrow \mu_k = -\bar{\mu}_j$. Moreover, $v_k \in W_I^\perp$ implies $G(v_k, v_k) = 0$. Observe that if μ_j is not real, then $G(v_j, \bar{v}_k) = 0$, since

$$0 = (\mu_j + \mu_k) G(v_j, \bar{v}_k)$$

$$= (\mu_j - \bar{\mu}_j) G(v_j, \bar{v}_k).$$

We have three subcases to consider.

1. *μ_j is neither real nor pure imaginary.*

The subspace $W = \operatorname{span}\{x_j, y_j, x_k, y_k\}$ is invariant under T and $C \otimes W$ contains four eigenvalues in Λ_C: $\mu_j, \bar{\mu}_j, -\bar{\mu}_j, -\mu_j$. By the second observation,

$$0 = G(x_j + iy_j, x_k - iy_k)$$

$$= G(x_j, x_k) - G(y_j, y_k) + iG(y_j, x_k)$$

$$+ iG(x_j, y_k),$$

$$\therefore G(x_j, x_k) = G(y_j, y_k) = a,$$
$$G(x_j, y_k) = -G(y_j, x_k) = b.$$

Hence, the Hessian in W is given in the ordered basis $\{x_j, y_j, x_k, y_k\}$ by

$$G|_W = \begin{pmatrix} 0 & 0 & a & b \\ 0 & 0 & -b & a \\ a & -b & 0 & 0 \\ b & a & 0 & 0 \end{pmatrix}.$$

So $\det G|_W = (a^2 + b^2)^2 = \frac{1}{16} |G(v_j, v_k)|^4 \neq 0$. Thus, $G|_W$ is nondegenerate and the lemma may be applied to W. We must only determine the signature of $G|_W$ to be done. Since W contains a two-dimensional null subspace, $\operatorname{span}\{x_j, y_j\}$, the signature of W cannot be 0, 4, 1, 3, i.e., it cannot be positive or negative definite, or Lorentz. This leaves only signature $= 2$. For this subcase, we have

$$\lambda|_{W_I^\perp} = 2 + \lambda|_{W^\perp}$$

with four eigenvalues in Λ_C.

2. *μ_j is pure imaginary.*

Subcase 1 applies here with one modification: The four eigenvalues are pure imaginary, $\mu_j, \mu_j, -\mu_j, -\mu_j$. Thus,

$$\lambda|_{W_I^\perp} = 2 + \lambda|_{W^\perp}$$

with four eigenvalues in Λ_R^0.

3. *μ_j is real.*

Observe that the eigenvectors v_j belonging to real eigenvalues μ_j may be chosen real, $v_j = x_j$. If v_j were pure imaginary, then iv_j is real. If v_j is neither real nor pure imaginary, then $\frac{1}{2}(v_j + \bar{v}_j)$ and $(1/2i)(v_j - \bar{v}_j)$ are two real eigenvectors in the two-dimensional subspace spanned by v_j and $\bar{v}_j$.

In this subcase, $v_j = x_j$ belongs to the real eigenvalue μ_j and $v_k = x_k$ belongs to the real eigenvalues $-\mu_j$ with $G(x_j, x_k) = a \neq 0$. The subspace $W = \operatorname{span}\{x_j, x_k\}$ is invariant under T and

$$G|_W = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}.$$

Since $\det G|_W = -a^2 \neq 0$, we have $G|_W$ is nondegenerate and the lemma applies to W. Clearly, the signature of $G|_W$ is one. Therefore,

$$\lambda|_{W_I^\perp} = 1 + \lambda|_{W^\perp}$$

with two eigenvalues in Λ_I.

By combining these various cases, we complete the first half of the theorem and show claims (a) and (b).

Conversely, give the various numbers, we want to construct G and J. This can be done by taking the direct sum of examples of I and II in the first half of the theorem:

I.

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mu = \pm i,$$

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mu = \pm i,$$

II.1

$$J = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

$$G = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix},$$

$$T = \frac{1}{2} \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

$\mu = \frac{1}{2}(\pm 1 \pm i)$ (four values);

II.2

$$J=\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

$$G=\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

$$T=\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix},$$

$\mu = \pm i, \pm i$ (four values);

II.3.

$$J=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad G=\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$T=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mu = \pm 1.$$

This completes the proof of the entire theorem.

We would like to point out that the assumption that $G^{-1}J$ is diagonalizable is necessary. For example, if

$$J=\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

$$G=\begin{pmatrix} 0 & 0 & -\lambda & 0 \\ 0 & 0 & -1 & -\lambda \\ -\lambda & -1 & 0 & 0 \\ 0 & -\lambda & 0 & 0 \end{pmatrix},$$

then $G^{-1}J$ is not diagonalizable for any λ.

As useful corollaries, one has the following results:

Corollary 1: A HF state has index zero, i.e., is a minimum, iff all the RPA energies are real and the RPA states have positive length in G. A HF state has index one iff there are exactly two pure imaginary RPA energies, while the remaining energies are real with eigenvectors with positive length in G.

Corollary 2: If there is a HF state for which the number of pure imaginary RPA energies is not divisible by four, then the total number of HF states is at least two more than the dimension of the many fermion state space.

Proof: There are three claims to be verified.

(a) A HF state has index zero, $\lambda = 0$, iff $\Lambda_R^- = \Lambda_R^0 = \Lambda_I = \Lambda_C = 0$ and, thus, $\dim V = \Lambda_R^+$.[17]

(b) For index one, $\lambda = 1$, we must have $\Lambda_R^- = 0$, since $2|\Lambda_R^-, \Lambda_R = \Lambda_C = 0$, since $4|\Lambda_R^0$ and $4|\Lambda_C$, and $\Lambda_I = 2$. Thus, $\Lambda_R^+ = \dim V - 2$. Conversely, $\Lambda_I = 2$, $\Lambda_R^+ = \dim V - 2 \Rightarrow \Lambda_R^- = \Lambda_R^0 = \Lambda_C = 0$ and $\lambda = 1$.

(c) If there is a HF state for which Λ_I is not divisible by four, then the index of that state is odd. Hence, $C_{2k+1} \geqslant 1$ for some k and from the strong Morse inequality, Eq. (15),

$$C = \dim \wedge \mathscr{H} + 2 \sum_{\lambda} \frac{1-(-)^{\lambda}}{2} C_{\lambda}$$
$$\geqslant \dim \wedge \mathscr{H} + 2C_{2k+1} \geqslant \dim \wedge \mathscr{H} + 2.$$

ACKNOWLEDGMENT

We would like to thank D. J. Rowe for helpful remarks.

[1]M. Reeken, J. Math. Phys. **11**, 2505 (1970).
[2]K. Gustafson and D. Sather, Rend. Mat. **4**, 1 (1971); Arch. Ratl. Mech. Anal. **48**, 109 (1972).
[3]J. Wolkowisky, Indiana U. Math. J. **22**, 551 (1972).
[4]C. Stuart, Arch. Ratl. Mech. Anal. **51**, 60 (1973); J. Math. Anal. Appl. **49**, 725 (1975).
[5]C. D. Luning and W. L. Perry. J. Math. Phys. **17**, 1156 (1976).
[6]Elliot H. Lieb and Barry Simon, Commun. Math. Phys. **53**, 185 (1977).
[7]J. Milnor, *Morse Theory* (Princeton U. P., Princeton, New Jersey, 1963).
[8]D. R. Hartree, Proc. Camb. Phil. Soc. **24**, 89 (1927–8).
[9]V. Fock, Z. Phys. **61**, 126 (1930).
[10]A. Lundell and S. Weingram, *The Topology of CW Complexes* (Van Nostrand–Reinholt, Toronto, 1969), p. 161.
[11]W. Greub, S. Halperin, and R. Vanstone, *Connections, Curvature and Cohomology* (Academic, New York, 1973), p. 193.
[12]D. J. Rowe, A. Ryman and G. Rosensteel, "Many-body quantum mechanics as a symplectic dynamical system" (to be published).
[13]D. J. Rowe, *Nuclear Collective Motion* (Methuen, London, 1970), pp. 247, 269.
[14]D. J. Rowe, Rev. Mod. Phys. **40**, 153 (1968).
[15]W. H. Bassichis, A. K. Kerman, and J. P. Svenne, Phys. Rev. **160**, 746 (1967).
[16]S. Lang, *Algebra* (Addison–Wesley, Reading, Massachusetts, 1965), p. 371.
[17]D. J. Thouless, Nucl. Phys. **22**, 78 (1961).

Periodic solutions of the classical polaron and bipolaron systems

F. Peeters [a)]
Department Natuurkunde, Universitaire Instelling Antwerpen (U.I.A.), Universiteitsplein 1, B-2610 Wilrijk, Belgium

J. T. Devreese [b)]
Institute of Applied Mathematics, Rijksuniversitair Centrum Antwerpen, Groenenborgerlaan 171, B-2020 Antwerpen, Belgium

(Received 13 August 1979; accepted for publication 18 April 1980)

It is shown that the exact solutions for the classical polaron model derived by Evrard *et al.* (the electron rotates about a fixed polaron centre with frequency $\Omega \gg \omega$) exist also for the discrete frequencies $\Omega = \omega/(n + \frac{1}{2})$($\omega$ is the longitudinal optical phonon frequency and n is an integer). For an infinite Brillouin zone radius the radius of the electron orbit is finite; this is in contrast with the behavior of the solutions found by Evrard *et al.*, where the radius diverges with the Brillouin zone radius. Ringwood and Devreese have obtained the same set of solutions as proposed here with a different method. The present calculation shows clearly that the discrete frequencies arise as a consequence of the fact that the electron has to move in such a way that the different divergent contributions to the self-interaction cancel each other. For the bipolaron system similar exact solutions are found. It is shown that the orbit frequency of the electrons, in the bipolaron system, can have the values $\Omega = \omega/(2n + 1)$.

1. INTRODUCTION

In recent years there has been increasing interest in solutions of classical field theories. As is well known, some of these solutions possess specific properties which cannot be derived via perturbation techniques. A study of a classical field theory can serve as a basis for a semi-classical approximation to the corresponding nontrivial quantum field theory.[1]

In the present paper a classical mechanical study is made of the interaction of a nonrelativistic particle with a scalar field. The interaction is nonlinear in the particle coordinates but linear in the field coordinates. This model is analogous to a proton interacting with chargeless, spinless mesons. In solid state physics the model corresponds to the polaron system,[2] which describes an electron interacting with the polarization field of an ionic crystal. The case in which two electrons are involved is called the biopolaron system.

Over the years, two fundamentally different approaches have been developed in the study of this field-theoretical problem. The first one relies on an elimination of the electron variable. Indeed, for the quantum-mechanical polaron Lee, Low, and Pines[3] have shown, via a canonical transformation, how the electron coordinate can be eliminated in the Hamiltonian. In the classical mechanical problem such a procedure has been followed by Gross.[4] The other approach uses the property that the interaction is linear in the field coordinates; this means that the field variables can be formally eliminated. In the quantum-mechanical theory this idea has been followed by Feynman,[5] using path-integral techniques, and by Devreese and others[6] using the Heisenberg equations of motion. Evrard *et al.*[7] have performed such a formal elimination of the phonon coordinates in the classical mechanical theory. The latter approach will be followed in the present paper.

The structure of the present paper is as follows. In Sec. 2 the field coordinates, in the electron equations of motion, are formally eliminated. In Ref. 7 the problem was then regularized via the introduction of a Gaussian distribution function $\exp(-k^2/K^2)$ in the $\mathbf{k}$-space. The parameter K corresponds roughly to the radius of the spherical symmetrical Brillouin zone. At the end of the calculations one takes the limit $K \to \infty$. In the present paper another distribution function was introduced in the $\mathbf{k}$-space which allowed us to perform the integrations over this $\mathbf{k}$-space. This distribution function contains a parameter ξ, which corresponds with the inverse of the parameter K of Ref. 7. Finally, the equation of motion for the electron position coordinate becomes an integro-differential equation in which the self-interaction contains only one integral, namely an integral over the time (the corresponding self-interaction in Ref. 7 contains a time integral and an integral over the $\mathbf{k}$-space). At zero temperature this equation is compared with the equation of motion resulting from a minimalization of the action $S[\dot{\mathbf{r}}(t), \mathbf{r}(t)]$ in Feynman's theory. A similar formal elimination, of the phonon coordinates, is carried through in the Hamiltonian.

In Sec. 3 the trial solution of Ref. 7 (an electron rotating about a fixed polaron center) is studied. The solutions of Ref. 7 are rederived. Further, the set of exact solutions [the electron frequency has the discrete value $\Omega = \omega/(n + \frac{1}{2})$, with n an integer] recently obtained by Ringwood and Devreese[10] is rederived by another method. The formulation of the self-interaction, as one time integral (presented in Sec. 2), allowed us to give an intuitive argument for the appearance of this set of solutions. Namely, the discrete frequencies arise as a consequence of the fact that the electron has to move in

[a)]Aspirant of the National Foundation of Scientific Research NFWO.
[b)]Also at UIA.

such a way that the different divergent contributions to the self-interaction cancel each other. This physical explanation is not apparent in Ref. 10.

As a direct generalization of the preceding sections, the bipolaron is studied in Sec. 4. Similar manipulations are performed as in the case of the polaron system. A set of exact solutions is found in which the two electrons rotate about the same center with the same frequency and radius. The allowed frequencies are just half those of the set of exact solutions for the polaron, namely $\Omega = \omega/(2n+1)$. An intuitive explanation, similar to that in Sec. 3, is given for this set of solutions.

2. THE EQUATION OF MOTION FOR THE ELECTRON

The system consisting of a free electron in interaction with the optical modes of a polar crystal is described by the Fröhlich Hamiltonian[2]:

$$H = \mathbf{p}^2/2m + \sum_{\mathbf{k}} \hbar\omega a_{\mathbf{k}}^* a_{\mathbf{k}} + \sum_{\mathbf{k}} (V_k a_{\mathbf{k}} \exp[i\mathbf{k}\cdot\mathbf{r}] + V_k^* a_{\mathbf{k}}^* \exp[-i\mathbf{k}\cdot\mathbf{r}]), \tag{2.1}$$

where $(\mathbf{r},\mathbf{p})$ and $(a_{\mathbf{k}}, a_{\mathbf{k}}^*)$ are the conjugate variables of, respectively, the electron and the phonon subsystems. The coupling strength between the electron and the lattice mode $\mathbf{k}$ is $V_k = -i(\hbar\omega/k)(\hbar/2m\omega)^{1/4}(4\pi\alpha/V)^{1/2}$, with V the volume of the crystal, α the dimensionless coupling constant, ω the frequency of the longitudinal-optical (L.O.) mode, and m the band mass of the electron.

Following Ref. 7, the phonon coordinates in the electron equation of motion can formally be eliminated; this results in the integro-differential equation:

$$m\ddot{\mathbf{r}} = \mathbf{F}_d + \mathbf{F}_s \tag{2.2}$$

with

$$\mathbf{F}_d = -i\sum_{\mathbf{k}} \mathbf{k} V_k a_+(\mathbf{k}) \exp[i(\mathbf{k}\cdot\mathbf{r}(t) - \omega t)] e^{-\epsilon|t|} + \text{c.c.} \tag{2.3}$$

$$\mathbf{F}_s = -\sum_{\mathbf{k}} \mathbf{k} \frac{|V_k|^2}{\hbar} \exp[i(\mathbf{k}\cdot\mathbf{r}(t) - \omega t)] e^{-\epsilon|t|} \times \int_{-\infty}^{t} dt' \exp[-i(\mathbf{k}\cdot\mathbf{r}(t') - \omega t')] e^{-\epsilon|t'|} + \text{c.c.} \tag{2.4}$$

where c.c. denotes the complex conjugate of the foregoing terms. A positive parameter ϵ ($\epsilon \to +0$) is introduced in the theory. In fact, one has replaced V_k by $V_k e^{-\epsilon|t|}$, which means that the electron-phonon coupling is switched on adiabatically. Further, the following notation has been used: $a_+(k) = \lim_{t_0 \to -\infty} a_{\mathbf{k}}(t_0) e^{+i\omega t_0}$.

In the classical theory $\mathbf{r}(t)$ is a function (not an operator) and therefore commutes with $\mathbf{r}(t')$. Thus the summation over the $\mathbf{k}$-vectors, in Eq. (2.4), can be performed. In doing this, special attention has to be paid to the interchange of the sum and the time integral. Indeed, the sum over the $\mathbf{k}$-vectors is not defined for the points $\mathbf{r}(t) = \mathbf{r}(t')$. Therefore a new regularization parameter ξ will be defined by introducing a distribution function $d(\xi k)$ in the $\mathbf{k}$-space. It is convenient to use the following distribution function: $d(x) = 1 - xf(x)$, where $f(x) = \text{Ci}(x)\cdot\sin(x) - \text{si}(x)\cdot\cos(x)$, with $\text{Ci}(x)$ and $\text{Si}(x) = \text{si}(x) + \pi/2$, the cosine and the sine integral, respectively.[8] In the limit $\xi \to 0$ this distribution function tends to 1, while for large k it behaves like $(\xi k)^{-2}$.

Inserting this distribution function into Eq. (2.4) and performing the summation (in the limit of an infinite crystal this sum is replaced by an integral) results in the expression

$$\mathbf{F}_s = \frac{-b^2}{2\pi\hbar} \lim_{\substack{\epsilon\to 0\\ \xi\to 0}} \int_{-\infty}^{t} dt' \, e^{-\epsilon(t-t')} \times \frac{\mathbf{r}(t) - \mathbf{r}(t')}{|\mathbf{r}(t) - \mathbf{r}(t')|} \cdot \frac{\sin\omega(t-t')}{(|\mathbf{r}(t) - \mathbf{r}(t')| + \xi)^2}, \tag{2.5}$$

where the notation $|V_k|^2 = b^2/Vk^2$ has been used. In the limit of zero temperature, this means $a_+(\mathbf{k}) = 0$ and $a_+^*(\mathbf{k}) = 0$ (thus $\mathbf{F}_d = 0$). Equation (2.2) describes an electron, with coordinate $\mathbf{r}(t)$, interacting, via an oscillating Coulomb force, with its positions $\mathbf{r}(t')$ in the past.

This integro-differential equation [Eqs. (2.2) and (2.5) at zero temperature] is slightly different from the one which can be derived from a minimalization of the action (for real times) in Feynman's theory[5,9] [see Eq. (4), Ref. 5]. This apparent discrepancy is a consequence of the different boundary conditions imposed on the electron and phonon coordinates in the theory of Feynman and in Eqs. (2.2) and (2.5). Namely, in Feynman's theory one imposed the condition that the electron and phonon coordinates are in the same phase-space point at the times $t = -\infty$ and 0. After eliminating the phonon variables, via path-integral techniques, this condition induces the self-interaction of the electron with its past and its future, while in deriving Eqs. (2.2) and (2.5) the phonon coordinates are only fixed at $t = -\infty$. Thus an elimination of the phonon coordinates induces a self-interaction of the electron with its past only.

The phonon coordinates can also be eliminated from the Hamiltonian (2.1). The phonon energy, given by the second term in Eq. (2.1), splits up into three terms:

$$H_f = H_{f_1} + H_{f_2} + H_{f_3}, \tag{2.6}$$

where

$$H_{f_1} = \sum_{\mathbf{k}} \hbar\omega a_+^*(\mathbf{k}) a_+(\mathbf{k}), \tag{2.7}$$

$$H_{f_2} = \sum_{\mathbf{k}} \omega |V_k| \lim_{\epsilon\to 0} \int_0^{\infty} d\tau \, e^{-\epsilon\tau} (a_+^*(\mathbf{k}) e^{-i[\mathbf{k}\cdot\mathbf{r}(t-\tau) + \omega\tau]} + a_+(\mathbf{k}) e^{i[\mathbf{k}\cdot\mathbf{r}(t-\tau) + \omega\tau]}), \tag{2.8}$$

$$H_{f_3} = \frac{b^2\omega}{2\pi\hbar} \lim_{\substack{\epsilon\to 0\\ \xi\to 0}} \int_{\infty}^{t} dt' \, e^{2\epsilon t'} \times \int_0^{\infty} d\tau \frac{\cos\omega\tau \cdot e^{-\epsilon\tau}}{|\mathbf{r}(t') - \mathbf{r}(t'-\tau)| + \xi}. \tag{2.9}$$

The first term H_{f_1} corresponds to the phonon energy of the real phonons, H_{f_2} represents the interaction energy of the real phonons with the virtual phonons, and H_{f_3} is the energy of the virtual phonons. The interaction energy, given by the last term in Eq. (1.1), splits up into two terms:

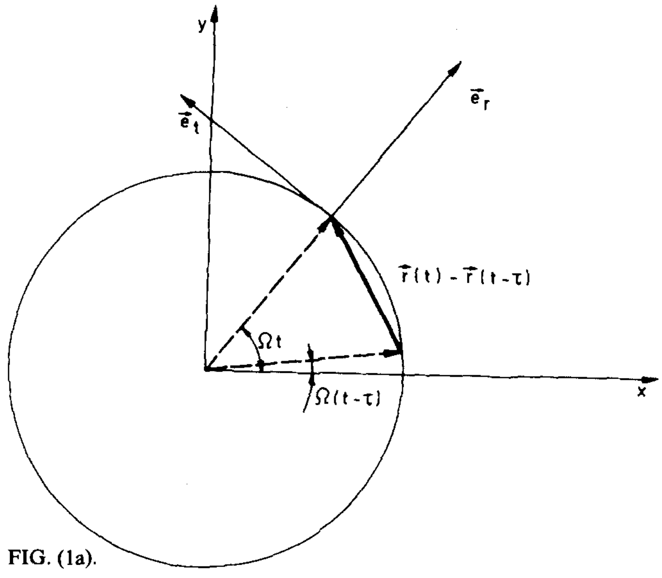


FIG. (1a).

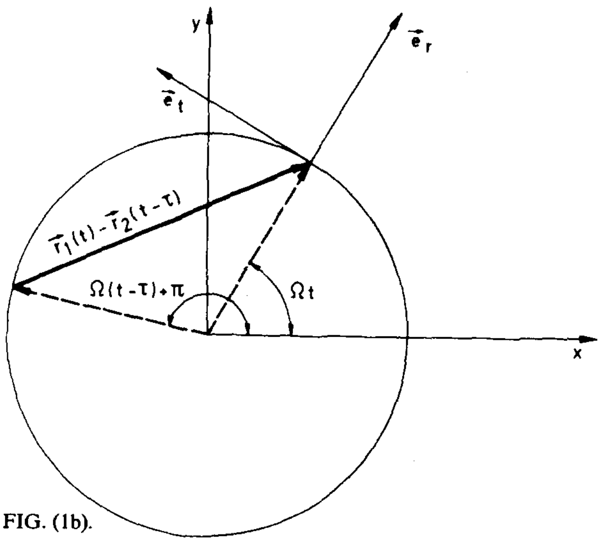


FIG. (1b).

FIG. 1. Coordinate transformation for the solution of the present classical polaron model, (a) [see Eq. (3.1)] and bipolaron model, (b) [see Eqs. (4.5) and (4.6)].

$$H_i = H_{i_1} + H_{i_2}, \tag{2.10}$$

where

$$H_{i_1} = \sum_{\mathbf{k}} (V_k a_+(\mathbf{k}) e^{i[\mathbf{k}\cdot\mathbf{r}(t) - \omega t]} + V_k^* a_+^*(\mathbf{k}) e^{-i[\mathbf{k}\cdot\mathbf{r}(t) - \omega t]}), \tag{2.11}$$

$$H_{i_2} = -\frac{b^2}{2\pi\hbar} \lim_{\substack{\epsilon\to 0\\ \xi\to 0}} \int_0^\infty d\tau \frac{\sin\omega\tau\cdot e^{-\epsilon\tau}}{|\mathbf{r}(t) - \mathbf{r}(t-\tau)| + \xi}. \tag{2.12}$$

The first term H_{i_1} is the interaction energy resulting from the direct interaction of the electron with the real polarization field. The second term H_{i_2} is the self-interaction energy of the electron.

3. STATIONARY SOLUTIONS OF THE CLASSICAL POLARON

In a paper by Evrard *et al.*[7] a trial solution is suggested in which the electron rotates in a circle, about a fixed polaron center, with radius u and frequency Ω:

$$\mathbf{r}(t) = u \cos\Omega t\, \mathbf{e}_x + u \sin\Omega t\, \mathbf{e}_y. \tag{3.1}$$

It is convenient to describe this motion in a uniform rotating coordinate frame fixed on the electron (see Fig. 1a). In this coordinate frame the following expressions are valid:

$$\mathbf{r}(t) - \mathbf{r}(t-\tau) = 2u \sin^2(\Omega\tau/2)\, \mathbf{e}_r + u \sin\Omega\tau\, \mathbf{e}_t, \tag{3.2}$$

$$|\mathbf{r}(t) - \mathbf{r}(t-\tau)| = 2u|\sin(\Omega\tau/2)|, \tag{3.3}$$

$$\ddot{\mathbf{r}}(t) = -u\Omega^2 \mathbf{e}_r. \tag{3.4}$$

Inserting these expressions into Eqs. (2.2) and (2.5), the following two integral equations result for the unknowns, u and Ω:

$$0 = \lim_{\substack{\epsilon\to 0\\ \xi\to 0}} \int_0^\infty d\tau \frac{\sin\Omega\tau}{|\sin(\Omega\tau/2)|} \cdot \frac{\sin\omega\tau\cdot e^{-\epsilon\tau}}{[2u|\sin(\Omega\tau/2)| + \xi]^2}, \tag{3.5}$$

$$mu\Omega^2 = \frac{b^2}{2\pi\hbar} \lim_{\substack{\epsilon\to 0\\ \xi\to 0}} \int_0^\infty d\tau\, |\sin(\Omega\tau/2)| \cdot \frac{\sin\omega\tau\cdot e^{-\epsilon\tau}}{[2u|\sin(\Omega\tau/2)| + \xi]^2}, \tag{3.6}$$

which are, respectively, the force along the tangential and the radial directions. The time integral over the semi-infinite interval can be replaced by a sum of time integrals over a finite time interval. After performing this summation, Eqs. (3.5) and (3.6) reduce to one equation:

$$u^3\Omega^3 = \frac{ib^2}{4\pi\hbar m} \lim_{\substack{\epsilon\to 0\\ \xi\to 0}} \int_0^\pi dt \frac{e^{-it} e^{-(2\epsilon/\Omega)t}}{\{\sin t + \xi/2u\}^2} \times \frac{\sin(2\omega t/\Omega) + \sin\{2\omega(\pi - t)/\Omega\} e^{-(2\pi/\Omega)\epsilon}}{1 - 2\cos(2\pi\omega/\Omega) e^{-(2\pi/\Omega)\epsilon} + e^{-(4\pi/\Omega)\epsilon}}. \tag{3.7}$$

The parameter ϵ, which regularizes the problem at time minus infinity, can now be put equal to zero

$$u^3\Omega^3 = \frac{ib^2}{8\pi\hbar m} \frac{1}{\sin(\pi\omega/\Omega)} \times \lim_{\xi\to 0} \int_0^\pi dt \frac{e^{-it} \cos[2\omega(t - \pi/2)/\Omega]}{(\sin t + \xi/2u)^2}. \tag{3.8}$$

For frequencies $\Omega = \omega/n$, with n an integer, the radius u diverges. The right-hand side (RHS) of Eq. (3.8) is real for $\Omega \neq \omega/n$; this is a consequence of the symmetry of the integrand. The leading term, in an asymptotic expansion about small ξ, of the integral in Eq. (3.8) is

$$u^3\Omega^3 \approx (b^2/4\pi\hbar m)\cot(\pi\omega/\Omega)\ln(u/\xi). \tag{3.9}$$

In Ref. 7 solutions with frequencies $\Omega \gg \omega$ are considered. In this frequency region the relation between the radius u and the frequency Ω is given by

$$u^3\Omega^3 \sim (2\alpha/\pi)(\hbar\omega/2m)^{3/2} \ln(u/\xi). \tag{3.10}$$

When the identification $\xi = 1/K$ is made (K is the radius of the spherical symmetrical Brillouin zone; this parameter is the ultraviolet cutoff parameter in Ref. 7) Eq. (36b) of Ref. 7 is reobtained.

In the limit of $\xi \to 0$ the relation between the radius u and the frequency Ω is logaritmically divergent, except for the frequency $\Omega = \omega/(n+\frac{1}{2})$, with n an integer. The corresponding radius u is given by $u\Omega = v$, with $v^3 = b^3/8\hbar m = \pi\alpha(\hbar/2m\omega)^{3/2}\omega^3$. This is the same set of solutions recently obtained by one of us (JTD) and others.[10]

The energy of this last set of solutions can be calculated via Eqs. (2.6) and (2.10). The energy of the total system can be split into three parts: $E = E_e + E_f + E_i$, with E_e the kinetic energy of the electron, E_f the phonon energy, and E_i the interaction energy. In the limit of small ξ one finds

$$E_e = \tfrac{1}{4}(\pi\alpha)^{2/3}\hbar\omega;\quad E_i = -(\pi\alpha)^{2/3}\hbar\omega,$$
$$E_f \sim (n+\tfrac{1}{2})(\pi\alpha)^{2/3}\hbar\omega\ \ln(u/\xi). \tag{3.11}$$

It is apparent that the kinetic energy E_e and the interaction energy E_i are both finite and do not depend on the electron frequency. Further they satisfy the following relations:

$$E_i = -4E_e;\quad E_i = 2\alpha(d/d\alpha)(E_e + E_i). \tag{3.12}$$

The quantum polaron ground state energy satisfies similar relations,[11] except that in the quantum mechanical case the last relation has to be replaced by $E_i = 2\alpha\, dE/d\alpha$.

Returning to the self-interaction expression (2.5), one can give an intuitive explanation for the appearance of the condition $\Omega = \omega/(n+\frac{1}{2})$ on the orbit frequencies. In the case of a periodic motion, the integral (2.5) contains an infinite contribution for every time $t' = t - mT$; with $T = 2\pi/\Omega$ the period and m an integer. Indeed, for every period in the past the electron position coordinate $\mathbf{r}(t') = \mathbf{r}(t - mT)$ equals $\mathbf{r}(t)$. This leads to a divergent Coulomb interaction: $\sin(m\omega T)/\xi^2$ ($\xi \to 0$). The sign of this divergence is determined by the factor $\sin\omega(t-t') = \sin(m\omega T)$. Consequently, the total self-interaction, which is a sum of such infinite contributions, will be finite when these divergent contributions cancel each other, or, in other words, when these divergent terms have an alternating sign. This means $\omega T = (2n+1)\pi$ or $\Omega = \omega/(n+\frac{1}{2})$. The foregoing explicit evaluation of Eq. (2.5), for the circle orbit motion, shows that this is the only set of frequencies which are allowed in the limit $\xi \to 0$.

4. THE CLASSICAL BIPOLARON

In this section the interaction of two electrons with each other and with the polarization field will be studied.[9] Such a system is described by the following Hamiltonian:

$$H = \sum_{j=1}^{2} \frac{\mathbf{p}_j^2}{2m} + \sum_{\mathbf{k}} \hbar\omega a_{\mathbf{k}}^* a_{\mathbf{k}} + \sum_{j=1}^{2}\sum_{\mathbf{k}} (V_k a_{\mathbf{k}} \exp[i\mathbf{k}\cdot\mathbf{r}_j] + V_k^* a_{\mathbf{k}}^* \exp[-i\mathbf{k}\cdot\mathbf{r}_j]) + e^2/|\mathbf{r}_1 - \mathbf{r}_2|. \tag{4.1}$$

The equation of motion for the electron with position vector $\mathbf{r}_1(t)$ can be obtained in the same way as in Sec. 2 (again the temperature is taken equal to zero):

$$m\ddot{\mathbf{r}}_1 = \mathbf{F}_d + \mathbf{F}_s + \mathbf{F}_{ic} + \mathbf{F}_c, \tag{4.2}$$

where $\mathbf{F}_d$ and $\mathbf{F}_s$ are, respectively, given by Eqs. (2.3) and (2.4). The other electron, with position vector $\mathbf{r}_2(t)$, is responsible for the forces

$$\mathbf{F}_c = \frac{\mathbf{r}_1(t) - \mathbf{r}_2(t)}{|\mathbf{r}_1(t) - \mathbf{r}_2(t)|}\cdot\frac{e^2}{|\mathbf{r}_1(t) - \mathbf{r}_2(t)|^2}, \tag{4.3}$$

$$\mathbf{F}_{ic} = -\frac{b^2}{2\pi\hbar}\lim_{\substack{\epsilon\to 0\\ \xi\to 0}}\int_{-\infty}^{t} dt' \frac{\mathbf{r}_1(t) - \mathbf{r}_2(t')}{|\mathbf{r}_1(t) - \mathbf{r}_2(t')|} \times \frac{\sin\omega(t-t')\cdot e^{-\epsilon(t-t')}}{[\,|\mathbf{r}_1(t) - \mathbf{r}_2(t')| + \xi\,]^2}. \tag{4.4}$$

$\mathbf{F}_c$ is the direct Coulomb force between the two electrons and $\mathbf{F}_{ic}$ is an integral of an oscillating Coulomb interaction (which is mediated by the polarization field) between electron 1 and the past of electron 2.

The set of exact solutions for the one electron system, found in the preceding section, suggests trying a similar solution for the bipolaron system. Therefore, consider the two electrons rotating in a circle about the same fixed point, with frequency Ω and radius u:

$$\mathbf{r}_1(t) = u\cos\Omega t\ \mathbf{e}_x + u\sin\Omega t\ \mathbf{e}_y, \tag{4.5}$$
$$\mathbf{r}_2(t) = u\cos\Omega t\ \mathbf{e}_x - u\sin\Omega t\ \mathbf{e}_y. \tag{4.6}$$

As in Sec. 3, consider a uniform rotating coordinate frame fixed on the electron (see Fig. 1b). For $\mathbf{r}_1(t) - \mathbf{r}_1(t-\tau)$, the same expression as (3.2) can be written down. To calculate (4.4) we need the following expressions:

$$\mathbf{r}_1(t) - \mathbf{r}_2(t-\tau) = 2u\cos^2(\Omega\tau/2)\ \mathbf{e}_r - u\sin\Omega\tau\ \mathbf{e}_t, \tag{4.7}$$

$$|\mathbf{r}_1(t) - \mathbf{r}_2(t-\tau)| = 2u|\cos(\Omega\tau/2)|. \tag{4.8}$$

Equation (4.2) then reduces to the following integral equations:

$$0 = \lim_{\substack{\epsilon\to 0\\ \xi\to 0}}\left\{\int_0^\infty d\tau \frac{\sin\Omega\tau}{|\sin(\Omega\tau/2)|}\cdot\frac{\sin\omega\tau\cdot e^{-\epsilon\tau}}{(2u|\sin(\Omega\tau/2)| + \xi)^2} - \int_0^\infty d\tau \frac{\sin\Omega\tau}{|\cos(\Omega\tau/2)|}\cdot\frac{\sin\omega\tau\cdot e^{-\epsilon\tau}}{(2u|\cos(\Omega\tau/2)| + \xi)^2}\right\}; \tag{4.9}$$

$$mu\Omega^2 = \frac{b^2}{2\pi\hbar}\lim_{\substack{\epsilon\to 0\\ \xi\to 0}}\left\{\int_0^\infty d\tau|\sin(\Omega\tau/2)|\cdot\frac{\sin\omega\tau\cdot e^{-\epsilon\tau}}{(2u|\sin(\Omega\tau/2)| + \xi)^2} + \int_0^\infty d\tau|\cos(\Omega\tau/2)|\cdot\frac{\sin\omega\tau\cdot e^{-\epsilon\tau}}{(2u|\cos(\Omega\tau/2)| + \xi)^2}\right\} - \frac{e^2}{4u^2}, \tag{4.10}$$

which correspond, respectively, to the force along the tangential and the radial direction. These integrals are computed in the same way as in Sec. 3. Introducing the notation $\nu = 2\omega/\Omega$ and $\xi' = \xi/2u$, for $\nu \neq 2n$, Eq. (4.9) reduces to

$$0 = \lim_{\xi\to 0}\int_0^{\pi/2} dx\,\sin(x - \pi/4)/(\sin x + \xi')^2\,\sin x, \tag{4.11}$$

and for $\nu = 2n$

$$0 = \lim_{\xi\to 0}\int_0^{\pi/2} dx \frac{\cos x - \sin 2nx}{(\sin x + \xi')^2}\cdot\frac{1 + (-1)^n}{1 - (-1)^n}. \tag{4.12}$$

It is evident that Eq. (4.11) cannot be satisfied because the RHS is equal to infinity. On the other hand, the RHS of Eq. (4.12) is identically zero for n an odd integer. Thus the electron frequencies are restricted to the set $\Omega = \Omega_n = \omega(2n+1)$, with n an integer which is different from those of Eq. (4.12).

Equation (4.10) gives the relation between the radius u and the frequency Ω. For the foregoing set of frequencies, Eq. (4.10) reduces to

$$u^3\Omega^3 = \frac{b^2}{4\pi\hbar m}\lim_{\xi\to 0}\int_0^{\pi/2} dx\left\{\frac{\sin x}{(\sin x+\xi')^2} + \frac{\cos x}{(\cos x+\xi')^2}\right\}\sin\nu x - e^2\Omega/4m. \quad (4.13)$$

An explicit calculation of the integral results in the expression

$$u_n^3\Omega_n^3 = -\frac{e^2}{4m}\Omega_n + \frac{b^2}{8m\hbar}\cdot A(n), \quad (4.14)$$

where

$$A(n) = \frac{4}{\pi}\cdot\frac{1}{4n+3}F(1,2n+\tfrac{3}{2};2n+\tfrac{5}{2};-1)+1, \quad (4.15)$$

with $F(a,b;c;z)$ the confluent hypergeometric function.[12] This function can be written as a finite sum:

$$A(n) = \frac{4}{\pi}\left[1-2\sum_{m=1}^{n}\frac{1}{(4m-1)(4m+1)}\right]. \quad (4.16)$$

It is interesting to take the limit for large n:

$$\lim_{n\to\infty}(u_n\Omega_n) = v, \quad (4.17)$$

which reduces to the result of the one electron case. It is not hard to understand this result, because in the limit $n\to\infty$ one has $\Omega\to 0$, and thus $u\to\infty$. This means that the electrons are far apart, the Coulomb interaction will be very weak in comparison with the electron-phonon interaction, and thus may be neglected. In this limit the system reduces to two noninteracting polarons; each of these polarons can be described by the Hamiltonian of Sec. 3.

The kinetic energy of the two electrons is $E_e = mu_n^2\Omega_n^2$, with $u_n\Omega_n$ given by Eq. (4.14). Observe that, in contrast to the exact solutions of Sec. 3, the kinetic energy is different for the different states; those states are labeled by the integer n. The interaction energy is equal to $E_i = -4mu_n^2\Omega_n^2$. This interaction energy contains both the interaction energy of the electrons with the polaron field and the interaction energy of the Coulomb interaction between the two electrons. These two energies, E_e and E_i, satisfy he same conditions [Eq. (3.12)] as the energies of the set of exact solutions found in Sec. 3. The phonon energy is again logarithmically divergent.

An intuitive argument similar to that in Sec. 3 can be given for the appearance of the discrete set of frequencies $\Omega_n = \omega/(2n+1)$. For each half-period in the past there is an infinite contribution to the self-interaction. Indeed, the integrand of $\mathbf{F}_s$ [Eq. (2.5)] is divergent for $t' = t - mT$, while the integrand $\mathbf{F}_{ic}$ [Eq. (4.4)] is divergent for $t' = t-(m+\frac{1}{2})T$, with $T = 2\pi/\Omega$ the period and m an integer. These divergences have the following form: $\sin(m\omega T/2)/\xi^2$ $(\xi\to 0)$. The sum of these divergences will be finite when the sign of the terms is alternating. This results in the condition $\omega T/2 = (2n+1)\pi$ or $\Omega = \omega/(2n+1)$, which is the set obtained above.

ACKNOWLEDGMENTS

One of us (F.P.) wishes to thank the I.W.O.N.L. for a scholarship during the academic year 1977–78, and L. Lemmens and F. Brosens for stimulating discussions. This work was performed in the frame work of the joint project E.S.I.S. (Electronic structure in solids) of the University of Antwerp U.I.A., and the University of Liège, and a CDC grant.

[1] R. Jackiw, Rev. Mod. Phys. **49**, 681 (1977).
[2] H. Fröhlich, Adv. Phys. **3**, 325 (1954).
[3] T. D. Lee, F. E. Low, and D., Pines, Phys. Rev. **90**, 297 (1953).
[4] E. P. Gross, Phys. Rev. **100**, 1571 (1955).
[5] R. P. Feynman, Phys. Rev. **97**, 660 (1955).
[6] J. T. Devreese, R. Evrard, and E. Kartheuser, Phys. Rev. B **12**, 3353 (1975).
[7] R. Evrard, J. Devreese, E. Kartheuser, and C.C. Grosjean, Nuovo Cimento B **12**, (1972).
[8] M. Abramowitz and I. A. Stegund, *Handbook of Mathematical Functions* (Dover, New York, 1966), p. 232.
[9] T. D. Schultz, in *Polarons and Excitons*, edited by C. Kuper and G. Whitfield (Oliver and Boyd, London, 1963).
[10] G. A. Ringwood and J. T. Devreese (to be published).
[11] L. F. Lemmens, J. De Sitter, and J. T. Devreese, Phys. Rev. B **8**, 2717 (1973).
[12] I. S. Gradshteyn and I. M. Rynhik, *Table of Integrals, Series and Products* (Academic, New York, 1965), p. 1040.

Thermal excitations in Heisenberg-xy systems

M. W. Puga[a)]

Department of Physics, University of Virginia, Charlottesville, Virginia 22901
and
Institut de Physique, Université de Neuchâtel, Rue A.-L. Breguet 1, CH-2000 Neuchâtel, Switzerland

(Received 17 December 1980; accepted for publication 4 April 1980)

The finite temperature excitations of the anisotropic chain in a magnetic field are calculated in the region $|\Delta| < 1$. Some special solutions are given, which point out several changes with respect to the $T = 0$ case.

I. INTRODUCTION

The electronic and magnetic properties of one-dimensional materials is one of the topics most widely studied at the present time. It is also very well known that a rather wide class of these materials present nearest neighbors spin coupling as the leading mechanism for the interaction between their atoms, and even though the spin can be of any value, at low enough temperatures, it is characteristic of some materials to present a doublet state, such that an effective spin $\frac{1}{2}$ Hamiltonian is sufficient.

Some complications, however, (in general due to crystal field effects)[1] introduce certain anisotropy into the system, leading consequently to the study of the following anisotropic-spin $\frac{1}{2}$-nearest neighbor Hamiltonian

$$\mathscr{H} = \tfrac{1}{2} \sum_{i=1}^{N} [\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta \sigma_i^z \sigma_{i+1}^z] - H \sum_i \sigma_i^z , \tag{1.1}$$

where the σ's are the usual Pauli matrices, Δ is defined as the anisotropy of the system, and the last term is due to the energy of interaction of the magnetic atoms with an external magnetic field along the direction of anisotropy. This Hamiltonian has been widely used to describe several magnetic materials,[2] and it is also useful to explain electronic properties in some one-dimensional organic materials.[3]

The ground state energy for this model has been previously calculated by Yang and Yang,[4] and the present author[5] has calculated, in some previous work, the elementary excitations for any value of anisotropy and magnetization y defined by

$$y = \sum_{i=1}^{N} \sigma_i^z / N.$$

The free energy of Hamiltonian (1.1), in the region $|\Delta| > 1$, has been calculated by Gaudin,[6] as a solution to an infinite set of certain nonlinear integral equations. One year later, Johnson and Mccoy[7] solved this set of equations in several limiting cases. Later on, Johnson[8] made use of Gaudin's equations to calculate the elementary excitations at finite temperature, in the region mentioned above ($|\Delta| > 1$).

The region $|\Delta| < 1$ has been independently studied in a series of papers by Takahashi *et al.*,[9] where they point out some differences between the two regions. In particular they find that for certain values of Δ in the region $|\Delta| < 1$, the set of equations becomes finite and can be solved, at least by numerical methods. However, no effort has been made as finite temperature for this region. In this paper we attempt a solution to this problem in order to extend the excitation spectrum to all values of anisotropy, magnetic field and temperature.

The work is divided into five sections. Section II introduces the notation and reviews the work of Takahashi and Suzuki.[9] In Sec. III we find by quadrature the elementary excitations spectrum at finite temperature. Section IV is devoted to checking the limit of zero temperature, and the method used there is generalized in Sec. V to explicitly solve the equations found in Sec. III for small values of temperature.

II. NOTATIONAL PRELIMINARIES

In this section we reproduce the equations given by Takahashi and Suzuki[9] for the thermodynamics of Hamiltonian (1.1). The symmetry $\mathscr{H}(\Delta) \mapsto -\mathscr{H}(-\Delta)$ allows the constraint $0 \leqslant \Delta < 1$. We first introduce the quantity μ, defined by

$$\mu = \cos^{-1}\Delta, \tag{2.1}$$

and expand this variable in a continued fraction, namely,

$$\frac{\mu}{\pi} = \frac{1}{\nu_1 +} \; \frac{1}{\nu_2 +} \; \frac{1}{\nu_3 + \cdots} . \tag{2.2}$$

These integers ν_i, defined by this expansion, are now used to defined the set of integers m_i and y_i by the following relations:

$$m_0 = 0, \quad m_i = \sum_{k=1}^{i} \nu_k , \tag{2.3}$$

$$y_{-1} = 0, \quad y_0 = 1, \quad y_i = y_{i-2} + \nu_i y_{i-1} . \tag{2.4}$$

Since the third component of the magnetization, y, is a good quantum number, we can study the system with a constant number (M) of spins down. Obviously the relation between y and M is trivially given by

$$y = 1 - 2M/N. \tag{2.5}$$

The energy and momentum of an excited state is given in terms of some quantities p_k by[9]

$$E = 2 \sum_{k=1}^{M} (\cos P_k - \Delta) - (N - 2M)H, \tag{2.6}$$

$$K = \sum_k P_k , \tag{2.7}$$

where the P_k can be found by imposing cyclic boundary conditions on Hamiltonian (1.1).

If we parametrize the P_k by the new quantities x_k by

$$\cot(P_k/2) = \cot(\mu/2)\tanh(\mu x_k/2), \tag{2.8}$$

Eqs. (2.6) and (2.7) become

[a)]Present address: Institut de Physique, Université de Neuchâtel, Rue A.-L. Breguet 1, CH-2000 Neuchâtel, Switzerland

$$E = \sum_{k=1}^{M} (-4\pi\mu^{-1}\sin\mu a_1(x_k) + 2H) - NH, \quad (2.9)$$

$$K = \sum_{k=1}^{M} \frac{1}{i} \ln \frac{\sinh[\frac{1}{2}\mu(x_k + i)]}{\sinh[\frac{1}{2}\mu(x_k - i)]}, \quad (2.10)$$

where the function $a_1(x)$ is defined as follows:

$$a_1(x) = \frac{\mu}{2\pi} \frac{\sin\mu}{\cosh\mu x - \cos\mu}. \quad (2.11)$$

It was argued by Gaudin[6] and Takahashi[9] that as $N \to \infty$ the solution to the x_k are grouped in strings along lines parallel to the imaginary axis. Therefore, we define a complex of order n and real part x as

$$C_n(x) \equiv \{z \in C;\ z = x + i(n + 1 - 2r) \quad r = 1,\dots,n\},$$

and argue that the x_k's will then be given by an infinite set of complexes, M_j of which have order n_j, parity v_j, and real parts x_α^j ($\alpha = 1,\dots,M_j; j = 1,2,\dots$). For a given j, n_j and v_j are determined by the sequences[9]

$$n_j = y_{r_i - 1} + (j - m_{r_i}) y_{r_i} \quad \text{with} \quad m_{r_i} \leqslant j \leqslant m_{r_i+1} \quad \text{for} \quad j = 1,2,\cdots, \quad (2.12)$$

$$v_1 = 1, \quad v_{m_1} = -1, \quad v_j = \exp\left\{\pi i \left[\frac{n_j - 1}{\pi} \right]\right\} \quad \text{for} \quad j \neq 1, m_j, \quad (2.13)$$

where the symbol $[x]$ denotes the maximum integer less than or equal to x. The set $\{x_\alpha\}$ is then replaced by the set $\{x_\alpha^j\}$ of complexes with unknown real parts.

The cyclic boundary conditions then give the following equations for x_α

$$N t_j(x_\alpha^j) = 2\pi I_\alpha^j + \sum_{k=1}^{\infty} \sum_{\beta=1}^{M_k} \theta_{jk}(x_\alpha^j - x_\beta^k) \quad \alpha = 1,2,\dots,M_j, \quad (2.14)$$

where t and θ are defined by

$$t_j(x) = f(x, n_j, v_j), \quad (2.15a)$$

$$\theta_{jk}(x) = f(x, |n_j - n_k|, v_j v_k) + f(x, n_j + n_k, v_j v_k) + 2 \sum_{l=1}^{\min(n_j, n_k) - 1} f(x, |n_j - n_k| + 2l, v_j v_k), \quad (2.15b)$$

with

$$f(x,n,v) = \begin{cases} 0 & \text{for} \quad n\mu/\pi = \text{integer}, \\ 2v\tan^{-1}\{(\cot n\mu/2)^v \tanh x\mu/2\}, \end{cases} \quad (2.16)$$

[notice that in the complex z plane, this function has branch points at the places: $\mathrm{Re}z = 0$, $\mathrm{Im}z = \pm(2/\mu) \times \arctan(\tan n\mu/2)^v$, $(n\mu/2) \neq$ integer. The real axis has no branch points, and one can choose the branch cuts connecting every other couple of branch points along the imaginary axis such that no cut crosses the real axis].

In the limit $N \to \infty$, the thermodynamics of the system is given by the distribution functions $\rho_j(x)$ and $\eta_j(x)$ which obey[9]

$$a_j(x) = (-1)^{r_j}(1 + \eta_j)\rho_j(x) + \sum_{k=1}^{\infty} \int_{-\infty}^{\infty} T_{jk}(x - x')\rho_k(x')\, dx', \quad (2.17)$$

$$\ln\eta_j(x) = \frac{-Aa_j(x) + 2n_j H}{T} + \sum_{k=1}^{\infty} \int_{\infty}^{\infty} dx' (-1)^{r_k} T_{jk}(x - x') \times \ln(1 + \eta_k^{-1}(x')), \quad (2.18)$$

where we have used the notation

$$T_{jk}(x) = \frac{1}{2\pi} \frac{d\theta_{jk}(x)}{dx}, \quad a_j(x) = \frac{1}{2\pi} \frac{dt_j(x)}{dx},$$

and the constant A is given by

$$A = 4\pi\mu^{-1} \sin\mu.$$

The energy, entropy, and other thermodynamic quantities can then be expressed in terms of the distribution functions $\rho_j(x)$ and $\eta_j(x)$. However, here we have given a general introduction to the equations we will use later on to develop the thermal excitation spectrum. The interested reader is referred to the original papers[9] for a complete analysis of the thermodynamic properties.

Equations (2.14), (2.17), and (2.18) are used in Sec. III to study the temperature effects on the energy and momentum of these excitations.

III. EXCITATION SPECTRUM

We create an elementary excitation of the system by simply removing or adding a new x_k to the equilibrium set given in Eq. (2.14). The new set $(x_\alpha'^j)$ will obey a similar equation, and upon subtracting both equations we arrive at the following:

$$N(t_j(x_\alpha'^j) - t_j(x_\alpha^j)) = 2\pi(I_\alpha'^j - I_\alpha^j) \mp \theta_{jk}(x_\alpha^j - x_k) + \sum_i \sum_\beta [\theta_{ji}(x_\alpha'^j - x_\beta'^i) - \theta_{ji}(x_\alpha^j - x_\beta^i)], \quad (3.1)$$

where the upper sign indicates that we have removed x_k from the equilibrium set while the lower is used when we insert it. The term with x_β^i equal to x_k is, of course, not included in the sum.

Since for an elementary excitation $x_\alpha'^j$ is very close to x_α^j at position different from x_k, we expand t_j and θ_{ji} about their equilibrium set x_α^j and obtain

$$a_j(x)\chi_j(x) = l_j + \sum_{i=1}^{\infty} \int_{-\infty}^{\infty} T_{ji}(x - x')[\chi_j(x) - \chi_i(x')]\rho_i(x')\, dx' \mp \frac{1}{2\pi}\theta_{jk}(x - x_k), \quad (3.2)$$

where we have defined

$$l_j = I_j' - I_j, \quad \chi_j(x_\alpha^j) \equiv N(x_\alpha'^j - x_\alpha^j),$$

and we have taken the thermodynamic limit with $\rho_j(x)$ being the density of x_α^j at position x. In this limit, as proved by Takahashi and Suzuki,[9] the ρ_j and η_j are given by the integral equations (2.17) and (2.18).

To express our equations in a compact notation, we define an operator B_{ij} such that

$$B_{ij}\cdot A_j = \sum_{j=1}^{\infty}\int_{-\infty}^{\infty}\Big\{\delta_{ij}\delta(x-x') + \frac{1}{S_i(x)}T_{ij}(x-x')\Big\}A_j(x')\,dx',$$

with

$$S_j(x) = (-1)^{r_j}(1+\eta_j(x)).$$

Notice that the operator D_{ij} defined by

$$D_{ij} = \frac{1}{S_j}B_{ij},$$

is symmetric in the sense that we have the following relation

$$A_i\cdot D_{ij}\cdot Q_j = Q_j\cdot D_{ji}\cdot A_i, \tag{3.3}$$

whatever the vectors Q and A are.

Multiplying Eq. (2.17) by $\chi_j(x)$ and subtracting from Eq. (3.2) gives

$$B_{ji}\cdot Q_i = L_j, \tag{3.4}$$

where the vectors Q and L are defined by

$$Q_j = \rho_j(x)\chi_j(x),\quad L_j = \frac{1}{S_j(x)}\Big(l_j \mp \frac{1}{2\pi}\theta_{jk}(x-x_k)\Big).$$

We can now proceed to the derivation of the energy and the momentum as functionals of Q.

The energy of the equilibrium set x_α^i is given from Eq. (2.9) by[9]

$$E = \sum_{j=1}\sum_{\alpha=1}^{M_j}\{-Aa_j(x_\alpha^j)+2n_jH\} - NH.$$

The difference in energy between this set and the new $x_\alpha'^i$, keeping y and Δ constant, is given by

$$\epsilon = -Aa_i'\cdot I_{ij}\cdot Q_j \mp (-Aa_k(x_k)+2n_kH),$$

with I_{ij} as the unit operator, and $a_i' = \partial a_i(x)/\partial x$.

To evaluate the integral over Q we use Eq. (3.3) and get

$$a_i'\cdot I_{ij}\cdot Q_j = (S_tL_t)\cdot B_{ti}^{-1}\cdot(a_i'/S_i).$$

The right-hand side can be calculated by means of the derivative of Eq. (2.18). A straightforward calculation gives

$$-\frac{A}{T}\frac{a_j'}{S_j} = B_{jt}\cdot\frac{\eta_t'}{\eta_tS_t}.$$

With the help of this identity the energy becomes

$$\epsilon = TL_s\cdot I_{st}\cdot\frac{\eta_t'}{\eta_t} \pm (Aa_k(x_k)-2n_kH).$$

By doing an integration by parts, we switch the derivative to L_s and the remaining integral is easily evaluated by using Eq. (2.18); we obtain

$$\epsilon = \pm T\sum_{j=1}^{\infty}(-1)^{r_j}\Big\{\frac{\theta_{jk}(\infty)}{\pi}\ln(1+\eta_j^{-1}(\infty))\Big\} \mp T\ln\eta_k(x_k). \tag{3.5}$$

To calculate the change in momentum, we start with the initial one that can be calculated from Eq. (2.10) to read[9]

$$K = \sum_{j=1}^{\infty}\sum_{\beta=1}^{M_j}\{t_j(x_\alpha^j)+\pi n_j\}.$$

In a similar way the change of momentum due to the excitation is given by

$$P = 2\pi a_i\cdot I_{ij}\cdot Q_j \mp \pi n_k \mp t_k(x_k),$$

and by using formulas (2.17) and (3.3) we get

$$a_i\cdot I_{ij}\cdot Q_j = (S_pL_p)\cdot I_{pt}\cdot \rho_t.$$

ρ_t can now be calculated by applying the operator

$$H\frac{\partial}{\partial H} + T\frac{\partial}{\partial T},$$

to Eq. (2.18) and, comparing the result with Eq. (2.17), we get

$$\rho_j(x) = -\frac{T}{A}(-1)^{r_j}\Big(H\frac{\partial}{\partial H}+T\frac{\partial}{\partial T}\Big)\ln(1+\eta_j^{-1}(x)).$$

Using this expression, the integral involved in the calculation of P can be evaluated as an integral form of Eq. (2.18). The result is

$$P = P_0 \mp \frac{\mu T}{2\sin\mu}\int_{-\infty}^{x_k}\Big[H\frac{\partial}{\partial H}+T\frac{\partial}{\partial T}\Big]\ln\eta_k(x)\,dx, \tag{3.6}$$

where P_0 does not depend on x_k and is given by

$$P_0 = -\frac{\mu T}{2\sin\mu}\sum_{i=1}^{\infty}(-1)^{r_i}\Big(l_i \mp \frac{\theta_{ik}(\infty)}{2\pi}\Big) \times\int_{-\infty}^{\infty}\Big[H\frac{\partial}{\partial H}+T\frac{\partial}{\partial T}\Big]\ln(1+\eta_i^{-1}(x))\,dx \pm (t_k(\infty)-\pi n_k).$$

Equations (3.5) and (3.6) give the energy–momentum relation for the elementary excitation at finite temperature, in parametric form, and with the η_j given by Eq. (2.18).

A check of the calculations can be achieved by taking the limit $T\to 0$. In Sec. IV we show that, in this limit, we obtain the expected equation.[5] Part of our reason for doing this calculation is to show some explicit results that we need in Sec. V, when we analyze the spectrum found here, for small values of T.

IV. THE ZERO TEMPERATURE LIMIT

In this section we prove explicitly that the elementary excitation spectrum at finite temperature, calculated in Sec. III, goes to the proper limit[5] as $T\to 0$.

Equation (3.5) gives the exictation energy. We can easily calculate the term $\eta_j^{-1}(\infty)$ at any temperature, by taking the limit $x\to\infty$ in Eq. (2.18). Using the fact that

$$a_j(\infty)=0,\quad T_{jk}(\infty)=0,$$

we get

$$\ln\eta_j(\infty) = \frac{2n_jH}{T}.$$

As the temperature approaches zero with H being constant we then see that the quantity $\eta_j^{-1}(\infty)$ vanishes. Notice, however, that the combination $T\ln\eta_j(x)$ remains finite as $T\to 0$. We therefore define

$$\epsilon_j^0(x) = \lim_{T\to 0}T\ln\eta_j(x).$$

Since[9] $\epsilon_j^0>0$ for $j\geqslant 2$, the only relevant integral equation in (2.18) is the one corresponding to $j=1$. ϵ_1^0 is an even function of x positive at ∞ and negative at zero. It has two zeros that we denote by $\pm k$ (we come back to this point later). The integral equation for $\epsilon_1^0(x)$ then becomes

$$\epsilon_1^0(x) = -Aa_1(x) + 2H - \int_{-k}^{k} T_{11}(x-x')\epsilon_1^0(x')\,dx', \quad (4.1)$$

where we have used the fact that $n_1 = 1$ and $\eta_1^{-1}(x')$ vanishes outside the region $[-k,k]$.

We now remove an $x_1 = \hat{x}$ from string 1 and add an $x_1 = k$ in the same string. The energy then becomes

$$\epsilon = \epsilon_1^0(k) - \epsilon_1^0(\hat{x}). \quad (4.2)$$

To calculate the momentum, we define the function $\hat{R}(x)$ by the relation

$$\hat{R}(x) = T\left(H\frac{\partial}{\partial H} + T\frac{\partial}{\partial T}\right)\frac{\epsilon_1^0(x)}{T}, \quad (4.3)$$

so that P can now be expressed as

$$P = -\frac{2\pi}{A}\int_{-k}^{\hat{x}} \hat{R}(x)\,dx + \frac{2\pi}{A}\int_{-k}^{k}\hat{R}(x)\,dx, \quad (4.4)$$

where we have taken the integer l_1 equal to zero to match the case $\Delta = 0$.

We now substitute Eq. (4.3) into Eq. (4.1) and find the following integral equation for $\hat{R}(x)$

$$\hat{R}(x) = Aa_1(x) - \int_{-k}^{k} T_{11}(x-x')\hat{R}(x')\,dx'.$$

With the change of variables $x = \alpha/\mu$, $b = \mu k$ we obtain

$$P = \int_{\mu\hat{x}}^{b} R(\alpha)\,d\alpha, \quad (4.5)$$

where the function

$$R(\alpha) = \frac{\hat{R}(\alpha/\mu)}{2\sin\mu}$$

satisfies the following integral equation

$$R(\alpha) = \frac{\sin\mu}{\cosh\alpha - \cos\mu} - \frac{1}{2\pi} \times \int_{-b}^{b}\frac{\sin 2\mu}{\cosh(\alpha-\beta) - \cos 2\mu}R(\beta)\,d\beta. \quad (4.6)$$

To transform the energy equation we first use the fact that k is a zero of ϵ_1^0, so that Eq. (4.2) simply reads

$$\epsilon = -\epsilon_1^0(\hat{x}).$$

Performing the change of variables mentioned above in Eq. (4.1) we get

$$\epsilon = 2\sin\mu S(\hat{\beta}), \quad (4.7)$$

with $S(\alpha)$ satisfying

$$S(\alpha) = \frac{\sin\mu}{\cosh\alpha - \cos\mu} - \frac{2H}{2\sin\mu} - \frac{1}{2\pi}\int_{-b}^{b}\frac{\sin 2\mu}{\cosh(\alpha-\beta) - \cos 2\mu}S(\beta)\,d\beta.$$

We now introduce the function $Z(\alpha)$ as a solution of the equation

$$Z(\alpha) = 1 - \frac{1}{2\pi}\int_{-b}^{b}\frac{\sin 2\mu}{\cosh(\alpha-\beta) - \cos 2\mu}Z(\beta)\,d\beta; \quad (4.8)$$

and with the use of this equation together with Eq. (4.6) we can write $S(\alpha)$ as

$$S(\alpha) = R(\alpha) - \frac{2H}{2\sin\mu}Z(\alpha)$$

so that Eq. (4.7) can be written as

$$\epsilon = 2\sin\mu\left\{R(\hat{\beta}) - \frac{R(b)}{Z(b)}Z(\hat{\beta})\right\}, \quad (4.9)$$

after using the fact that b is a zero of $S(\alpha)$, as implied by Eq. (4.7).

Equations (4.5) and (4.9) are the solution of the elementary excitation spectrum at zero temperature.[10] In Sec. V we use these techniques to solve the excitation equations for small values of temperature.

V. SPECIAL SOLUTIONS

Equations (3.5) and (3.6) are extremely difficult to solve in closed form, and a general solution looks rather impossible. Some idea of the effects of temperature on these excitations can be achieved, however, if one expands the solution for small values of temperature.

Using the results obtained in Sec. IV, we can create an elementary excitation at finite magnetization by removing x_1 from string 1 and placing a new x_1' in the same string. These changes (performed in the same string), keep constant the magnetization of the system, and we can therefore simplify the equations considerably. In fact, Eq. (3.5) now reads

$$\epsilon = -T\ln\eta_1(x_1) + T\ln\eta_1(x_1'), \quad (5.1)$$

while the momentum becomes

$$P = \frac{\mu T}{2\sin\mu}\int_{x_1}^{x_1'}\left(H\frac{\partial}{\partial H} + T\frac{\partial}{\partial T}\right)\ln\eta_1(x)\,dx, \quad (5.2)$$

after taking all the $l_i = 0$ to match the proper curve in the limit of zero temperature.

Defining the function $\epsilon_1(x)$ as

$$\epsilon_1(x) = T\ln\eta_1(x),$$

we can see from Eq. (2.18) that this function obeys[9]

$$\epsilon_1(x) + Aa_1(x) - 2H = T\int_{-\infty}^{\infty} a_2(x-x')\ln(1 + e^{-\epsilon_1(x')/T})\,dx', \quad (5.3)$$

where

$$a_j(x) = \frac{\mu}{2\pi}\frac{\sin j\mu}{\cosh\mu x - \cos j\mu}, \quad (5.4)$$

and we have assumed $T \ll H$.

If $T = 0$, the right-hand side of Eq. (5.3) vanishes wherever $\epsilon_1(x)$ is positive, so denoting by $\epsilon_1^0(x)$ the solution of (5.3) for zero temperature (see also Sec. IV), we can write

$$\epsilon_1^0(x) = -Aa_1(x) + 2H - \int_{\epsilon_1^0<0} a_2(x-x')\epsilon_1^0(x')\,dx'. \quad (5.5)$$

Since $a_1(x)$ and $a_2(x)$ are even functions of x, so is $\epsilon_1^0(x)$.

Moreover, at $\pm\infty$ we have

$$\epsilon_1^0(\pm\infty) = 2H.$$

Thus in the region of $y \neq 0,1$ $[H < 2(1-\Delta)]$ ϵ_1^0 has necessarily two zeros, denoted by $\pm k$ in Sec. IV. As the temperature increases slightly the zeros of $\epsilon_1(x)$ move to locations that we denote by k', and Takahashi[9] proves that for small temperature (and $\Delta \neq -1$) we have

$$\epsilon_1(x) = \epsilon_1^0(x) + \frac{\pi\mu U(x)}{24\sin\mu W(k)} T^2, \tag{5.6}$$

where U and W satisfy

$$U(x) = a_2(x+k) + a_2(x-k) - \int_{-k}^{k} a_2(x-x')U(x')\,dx', \tag{5.7}$$

$$W(x) = \frac{\mu^2}{2\pi}\frac{\sin\mu\,\sinh\mu x}{(\cosh\mu x - \cos\mu)^2} - \int_{-k}^{k} a_2(x-x')W(x')\,dx'. \tag{5.8}$$

We can now proceed to use Eq. (5.6) to evaluate the exictation energy and momentum. From Eq. (5.3) it becomes obvious that $\epsilon_1(x)$ is an even function of x so that if we make a move such that $x_1' = -x_1$, the energy will be zero and the momentum finite. That allows many zero-energy excitations at finite temperature due to the number of holes created by exciting particles above the "Fermi surface".

Out of these many possibilities we will concentrate on the excitation such that $x_1' = k'$, which is the one we can easily compare to the only one allowed at zero temperature.

Since k' is a zero of $\epsilon_1(x)$, Eq. (5.1) becomes

$$\epsilon = -\epsilon_1^0(x_1) - \frac{\pi\mu}{24\sin\mu}\frac{U(x_1)}{W(k)}T^2, \tag{5.9}$$

where $U(x)$ and $W(x)$ are given by Eq. (5.7) and (5.8).

To calculate the momentum change due to temperature effects we first divide Eq. (5.3) by T and apply the operator

$$T\left(T\frac{\partial}{\partial T} + H\frac{\partial}{\partial H}\right)$$

to both sides. A simple calculation yields

$$L(x) = Aa_1(x) - \int_{-\infty}^{\infty} a_2(x-x')\frac{1}{1+\exp(\epsilon_1(x')/T)} \times L(x')\,dx', \tag{5.10}$$

where $L(x)$ has been defined by

$$L(x) = T\left(T\frac{\partial}{\partial T} + H\frac{\partial}{\partial H}\right)\frac{\epsilon_1(x)}{T}.$$

With the help of Eq. (5.6) we can expand Eq. (5.10) as

$$\left[\frac{L(x)}{2\sin\mu}\right] = \frac{\sin\mu}{\cosh\mu x - \cos\mu} - \frac{1}{2\pi} \times \int_{-k'}^{k'} \frac{\sin 2\mu}{\cosh\mu(x-x') - \cos 2\mu}\left[\frac{L(x')}{2\sin\mu}\right]\mu\,dx'.$$

Changing variables to $\alpha = \mu x$ we can express the momentum as

$$P = \int_{\mu x_1}^{b} [R(\alpha) - K(\alpha)R(b)(b'-b)]\,d\alpha,$$

$$K(\alpha) = \frac{1}{2\pi}\left\{\frac{\sin 2\mu}{\cosh(\alpha+b) - \cos 2\mu} + \frac{\sin 2\mu}{\cosh(\alpha-b) - \cos 2\mu}\right\}, \tag{5.11}$$

where $b' = \mu k'$ and $R(\alpha)$ is defined in Sec. IV.

Expanding (5.11) around μk we have

$$P = P^0 + R(b)\mu(k'-k)\left(1 - \int_{\mu x_1}^{b} K(\alpha)\,d\alpha\right), \tag{5.12}$$

where P^0 is the momentum obtained at zero temperature. $(k'-k)$, being of order T^2, is now calculated by using the fact that k' is a zero $\epsilon_1(x)$. Equation (5.6) for $x = k'$ reads

$$\epsilon_1^0(k') + \frac{\pi\mu U(k')}{24\sin\mu W(k)}T^2 = 0. \tag{5.13}$$

Expanding the function about k we have

$$\epsilon_1^{0\prime}(k)\cdot(k'-k) = -\frac{\pi\mu U(k)}{24\sin\mu W(k)}T^2, \tag{5.14}$$

after using the fact that k is a zero of $\epsilon_1^0(x)$.

Taking the derivative in Eq. (5.5) and integrating by parts we get

$$\epsilon_1^{0\prime}(x) = AW(x),$$

which transforms Eq. (5.14) into

$$k' - k = -\frac{\mu^2 U(k)}{96\sin^2\mu W^2(k)}T^2.$$

Equation (5.12) now gives the final result

$$P = P^0 - \frac{R(b)U(k)\mu^3}{96\sin^2\mu W^2(k)}T^2\left(1 - \int_{\mu x_1}^{b} K(\alpha)\,d\alpha\right). \tag{5.15}$$

Equations (5.9) and (5.15) give us the momentum and energy changes due to a small temperature effect.

We can see by inspection that for $x_1 = k'$ we have $P = 0$ and $\epsilon = 0$, so that the curve starts at the origin. For $x_1 = -k'$ the energy is again zero by the symmetry argument mentioned above, while the momentum is shifted with respect to the one at zero temperature by the amount

$$\xi = P_{\max}(T=T) - P_{\max}(T=0) = -\frac{R(b)\mu^3}{96\sin^2\mu}\frac{U(k)}{W^2(k)}T^2\left(2 - \int_{-b}^{b} K(\alpha)\,d\alpha\right),$$

the factor of two coming from the surface effect at the lower end of the integral.

The eigenvalues of the integral operator a_2 were shown by Yang and Yang[4] to lie in the strip $[-1,1]$. The eigenvalues of the resolvent $1/(1+a_2)$ are then positive, and consequently the functions $U(x)$ and $W(x)$ have the same sign as their inhomogeneous terms [Eq. (5.7) and (5.8)]. $W(x)$ is then always positive and

$$\operatorname{sign} U(x) = \operatorname{sign}(\sin 2\mu) = \operatorname{sign}\Delta.$$

On the other hand, we also have the result[4]

$$R(\alpha) > 0, \quad \forall \Delta \epsilon [-1,1],$$

$$\left|\int_{-b}^{b} K(\alpha)\,d\alpha\right| \leqslant \int_{-\infty}^{\infty} |K(\alpha)|\,d\alpha \leqslant 2.$$

So we can conclude that this shift ξ, and the slight energy difference, are negative for positive Δ and positive for negative Δ. The shift of the momentum at which the energy vanishes is then proportional to T^2 and moves toward the right or the left according to the sign of Δ (negative or positve).[11] A not so surprising result is that for $\Delta = 0$, there is not such a

T^2 correction. This can easily be seen by comparing Eqs. (5.3) and (5.5), that yield the same solution for $a_2 = 0$, which is the case for $\Delta = 0$. This is in good agreement with the intuitive result that we would expect from the known relation of the $\Delta = 0$ case to a free particle gas.

Equation (2.18) can also be used to discuss other order complexes. Keeping in mind that all the ϵ_j for $j \geqslant 2$ are positive in the limit of zero temperature, one can approximate this equation at small T. A straightforward calculation yields:

$$\epsilon_j(x) = -Aa_j(x) + 2n_j H - \int_{-\infty}^{\infty} T_{j1}(x - x') \ln(1 - e^{-\epsilon_1(x')/T})\, dx'.$$

Expanding the integral on the right-hand side for low temperature, and using the fact that k is a zero of $\epsilon_1^0(x)$, we obtain

$$\epsilon_j(x) = \epsilon_j^0(x) - \left[\frac{\pi\mu T^2}{24 \sin\mu\, W(k)}\right] \times \left[\int_{-k}^{k} T_{j1}(x - x') U(x')\, dx' - T_{j1}(x - k) - T_{j1}(x + k)\right]$$

which shows a first order correction of T^2 for all complexes.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor Hans Beck for his critical reading of the work, as well as for his suggestion to increase the range of validity of these results. I also thank Dr. Jill Bonner and Dr. Gerhard Müller for many stimulating discussions. This work was supported by the University of Virginia and the Swiss National Science Foundation.

[1]See, for instance, U. Tellenbach, J. Phys. C **11**, 2287 (1978).
[2]L. J. de Jongh and A. R. Miedena, Adv. Phys. **23**, 1 (1974); Y. Endoh, A. Shirane, R. J. Birgeneau, P. M. Richards, and S. L. Holt, Phys. Rev. Lett. **32**, 170 (1974).
[3]J. Hubbard, Phys. Rev. B **17**, 494 (1978).
[4]C. N. Yang and C. P. Yang, Phys. Rev. **150**, 321, 327 (1966).
[5]M. W. Puga, Phys. Rev. Lett. **42**, 405 (1979).
[6]M. Gaudin, Phys. Rev. Lett. **26**, 1301 (1971).
[7]J. D. Johnson and B. M. Mccoy, Phys. Rev. A **6**, 1613 (1972).
[8]J. D. Johnson, Phys. Rev. A **9**, 1743 (1974).
[9]M. Takahashi and M. Suzuki, Prog. Theor. Phys. **48**, 2187 (1972); M. Takahashi, Prog. Theor. Phys. **50**, 1519 (1973); **51**, 1348 (1974).
[10]The equations given in Ref. 5 can be put in this form by using the function $Z(\alpha)$ defined in Eq. (4.9).
[11]This is in perfect agreement with the recent result that magnetic field and temperature have similar effects on the excitation spectrum of the Heisenberg antiferromagnet ($\Delta = 1$); J. Bonner and G. Müller (private communication).

Erratum: On the sound field due to a moving source in a superfluid [J. Math. Phys. 20, 1409 (1979)]

J. C. Murray
Department of Mathematical Sciences, University of Petroleum and Minerals, Dhahran, Saudi Arabia

(Received 24 April 1980; accepted for publication 2 May 1980)

Equations (3.2), (3.4), (3.28), and (4.12) should be corrected to read

$$\alpha \frac{\partial^2 p^+}{\partial t^2} - \Delta p^+ - \gamma \frac{\partial^2 T^+}{\partial t^2} = \frac{q_0}{4U\pi} \frac{\partial}{\partial t} [\exp(i\omega_0 t)\delta(t - x/U)] \frac{\delta(r)}{r}, \qquad (3.2)$$

$$\alpha \frac{\partial^2 p^-}{\partial t^2} - \Delta p^- - \gamma \frac{\partial^2 T^-}{\partial t^2} = \frac{q_0}{4U\pi} \frac{\partial}{\partial t} [\exp(-i\omega_0 t)\delta(t - x/U)] \frac{\delta(r)}{r}, \qquad (3.4)$$

$$s_j^{\pm} = \frac{-i(1 - M_j^2)^{1/2}}{U} (\omega \pm a_j)^{1/2} (\omega \pm b_j)^{1/2}, \qquad (3.28)$$

$$\Pi = -\frac{q_0^2 \omega_0^2}{8\pi(u_1^2 - u_2^2)^2} \Bigg[\frac{s_e(\rho_{n_e} - \rho_{s_e})}{\rho_{n_e}} \mu u_1^2 u_2^2 \times \left(\frac{u_2^2(\beta u_1^2 - 1)}{u_1(1 - M_1^2)^2} + \frac{u_1^2(\beta u_2^2 - 1)}{u_2(1 - M_2^2)^2} \right) - \frac{2}{\rho_e} \left(\frac{u_2^4(\beta u_1^2 - 1)^2}{u_1(1 - M_1^2)^2} + \frac{u_1^4(\beta u_2^2 - 1)^2}{u_2(1 - M_2^2)^2} \right) \Bigg]. \qquad (4.12)$$

Erratum: The Planck integral cannot be evaluated in terms of a finite series of elementary functions [J. Math. Phys. 21, 14 (1980)]

Richard Pavelle
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02173

(Received 1 May 1980; accepted for publication 2 May 1980)

In Ref. 3, terms of the equation are not clear, especially superscripts and subscripts. The equation should read as follows:

$$(e^x - 1)^{-1} = e^{-x}(1 - e^{-x})^{-1} = e^{-x} \sum_{n=0}^{\infty} e^{-nx} = \sum_{n=1}^{\infty} e^{-nx}.$$

 0022-2488/80/082313-01$0.00

Erratum: Integration of near-resonant systems in slow-fluctuation approximation [J. Math. Phys. 21, 462 (1980)]

M. F. Augusteijn and E. Breitenberger
Department of Physics, Ohio University, Athens, Ohio 45701

(Received 7 May 1980; accepted for publication 16 May 1980)

Several disfigured statements appear in the paper, which should be corrected as follows:

(1) After Eq. (4.3), in the sentence beginning "General conclusions...": instead of "can fortunately be drawn" read "can unfortunately not be drawn".

(2)On p. 467, top left: instead of "with *n* which are not more simple additive" read "with *n* more which are simple additive".

(3) In Ref. 14, in the reference to Korteweg: instead of "5, 10 (1897)" read "5, No. 8 (1879), especially p. 10".

(4) In Ref. 17: instead of "Oxford University, New York" read "Clarendon Press, Oxford".

 0022-2488/80/082314-01$0.00