N-level systems in contact with a singular reservoir. II

Alberto Frigerio*

Istituto di Fisica dell'Università, Milano, Italy

Vittorio Gorini

Istituto di Fisica dell'Università, Milano, Italy
and Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Italy
(Received 9 December 1975)

We study an N-level system coupled linearly to an infinite quasifree Fermi or Bose reservoir in the vacuum state or in a state corresponding to an arbitrary temperature. We show that the singular reservoir limit can be performed in the vacuum state and at infinite temperature, thus leading to a completely positive Markovian reduced time evolution for the system, which, in the infinite temperature case, preserves the central state. On the other hand, no such limit is possible for KMS states (finite temperature) and at zero temperature. Some extension to norm-continuous semigroups of an infinite-dimensional $\mathcal{B}(\mathcal{H})$ is possible.

1. INTRODUCTION

In a previous paper,[1] hereafter referred to as I, one of us and Kossakowski have studied the reduced dynamics of an N-level atom coupled linearly to a set of N^2 quasifree Bose fields in the vacuum state, under the assumptions that the one-particle energy spectrum of the reservoir is unbounded from above and from below, and that the reservoir particles emitted and absorbed by the atom have a Gaussian energy distribution. It was shown in I that in the limit when the inverse width of the Gaussian tends to zero, thus leading to no interference between particles emitted and/or absorbed at different times, the reduced dynamics of the atom becomes a completely positive dynamical semigroup. The most general form of such a semigroup[2,3] can be obtained in this way by a suitable choice of the interaction parameters. An implicit indication that some result of this kind should hold can be found also in Ref. 4, Sec. 4, where an isometric (but in general nonunitary) dilation of a completely positive dynamical semigroup is constructed on the tensor product of the (possibly infinite-dimensional) Hilbert space of the system times a boson or fermion Fock space.

Complete positivity is believed to be a general feature of the dynamics of quantum open systems.[2,3,5,6] As regards the semigroup property, it is often used as an approximation, but cannot be derived exactly, if reasonable reservoirs are considered, unless some limiting procedure is performed. The standard procedures are the weak coupling limit, which in particular allows one to derive dynamical semigroups describing the approach of a system to thermal equilibrium with its surroundings,[7-9] and the limit of singular reservoirs, such as the one performed in I. Singular reservoirs have been considered also in Ref. 10, in the study of laser models. The aim of this note is to investigate some possible generalizations of I, coupling the N-level system to a set of quasifree Fermi or Bose fields in the vacuum state or in a state corresponding to an arbitrary temperature. Some norm-continuous semigroups of $\mathcal{B}(\mathcal{H})$, with an infinite-dimensional $\mathcal{H}$, can be studied in the same way. We show that in the vacuum state a fermion reservoir yields the same results that were found in I with the use of bosons (Sec. 2). For finite-temperature states ($0 \leq T < \infty$) no generalization holds. On the other hand, the singular reservoir limit can again be performed at infinite temperature, leading to a semigroup which leaves the central state invariant. This procedure is straightforward in the fermion case, while in the boson case the limit $T \to \infty$ has to be performed considering an interaction Hamiltonian with an appropriate temperature dependence, in order to compensate for the indefinite increase of the expected number of quanta per unit energy. The case of infinite temperature has the advantage that the one-particle energy spectrum can be chosen to be positive (Sec. 3). We conclude with a few remarks dealing with the case of a Fermi system whose creation and annihilation operators anticommute with the Fermi fields of the reservoir (Sec. 4).

2. GENERAL FEATURES AND VACUUM STATE CASE

In analogy to I, we consider an N-level system S, coupled to N^2 Fermi fields (the reservoir R) by a linear interaction

$$V^\epsilon = \sum_{\alpha=1}^{N^2} F_\alpha \otimes \varphi_\alpha(f^\epsilon), \tag{2.1}$$

where $\{F_\alpha \mid \alpha = 1, \ldots, N^2\}$ is an orthonormal basis of self-adjoint matrices in $M(N)$, $F_{N^2} = \mathbb{1}/\sqrt{N}$,

$$\varphi_\alpha(f^\epsilon) = \sum_{\beta=1}^{N^2} [\bar{\mu}_\beta \bar{\lambda}_\alpha^\beta a_\beta(f^\epsilon) + \mu_\beta \lambda_\alpha^\beta a_\beta(f^\epsilon)^*], \tag{2.2}$$

$$\sum_{\alpha=1}^{N^2} \bar{\lambda}_\alpha^\rho \lambda_\alpha^\sigma = \delta_{\rho\sigma}, \tag{2.3}$$

$$f^\epsilon(\omega) = (2\pi)^{-1/2} \exp[-\epsilon^2\omega^2/8]. \tag{2.4}$$

The free evolution of the fields is given by

$$\alpha_t a_\alpha(f) = a_\alpha(f_{-t}),$$
$$f_t = \exp(-iht)f, \quad [\exp(-iht)f](\omega) = \exp(-i\omega t)f(\omega). \tag{2.5}$$

We choose the state of the reservoir to be a quasifree gauge-invariant state,[11-13] which means that its correlation functions are given by

$$\omega_A(a_{\alpha_n}(f_n)^* \cdots a_{\alpha_1}(f_1)^* a_{\beta_1}(g_1) \cdots a_{\beta_m}(g_m))$$
$$= \delta_{mn} \det\{\omega_A(a_{\alpha_i}(f_i)^* a_{\beta_j}(g_j))\},$$

$$\omega_A(a_\alpha(f)^* a_\beta(g)) = \delta_{\alpha\beta}(g, Af). \tag{2.6}$$

A is an operator on the one-particle space, $0 \leq A \leq 1$. $A = 0$ gives the vacuum state, the (unique) KMS state at inverse temperature β and with chemical potential μ is given by[12]

$$A = (\exp[\beta(h - \mu)] + 1)^{-1}. \tag{2.7}$$

For $\beta \to 0$ (infinite temperature), A converges to $\frac{1}{2}$. For $\beta \to \infty$, A tends to the projection on the functions whose support is contained in $(-\infty, \mu]$ {this would correspond to the vacuum if we had chosen the one-particle energy spectrum to be contained in $[\mu, \infty)$}.

We take as Hilbert space the tensor product between $\mathbb{C}^N$ and the GNS space $\mathcal{H}_{\omega_A}$ of the state ω_A. Working in the interaction picture and letting

$$\begin{aligned} F_\alpha(t) &= \exp(-iH_S t) F_\alpha \exp(iH_S t), \\ \varphi^\epsilon_\alpha(t) &= \alpha_{-t}\varphi_\alpha(f^\epsilon) = \varphi_\alpha(f^\epsilon_t), \end{aligned} \tag{2.8}$$

where H_S is the Hamiltonian of the free N-level system, we are interested in computing the limit as $\epsilon \to 0$ of

$$\left(x_t \otimes \Omega \middle| \left[\sum_{m=0}^{\infty} \int_{0 \leq t_m \leq \cdots \leq t_1 \leq t} dt_1 \cdots dt_m M^\epsilon(t_1) \cdots M^\epsilon(t_m) \times (A \otimes \mathbb{1})\right] y_t \otimes \Omega\right), \tag{2.9}$$

where $x, y \in \mathbb{C}^N$, $x_t = \exp(-iH_S t)x$,

$$M^\epsilon(t) = i\left[\sum_{\alpha=1}^{N^2} F_\alpha(t) \otimes \varphi^\epsilon_\alpha(t), (\cdot)\right] = i[V^\epsilon(t), (\cdot)],$$

where, because of the gauge-invariance of ω_A, only those terms survive for which $m = 2n$. Expression (2.9) depends on the reservoir through its (even-point) correlation functions, which are[13]

$$\begin{aligned} &\omega_A(\varphi^\epsilon_{\alpha_1}(t_1) \cdots \varphi^\epsilon_{\alpha_{2n}}(t_{2n})) \\ &= \sum_{p \in \mathcal{P}_n} \operatorname{sgn} p \prod_{r=1}^{n} \omega_A(\varphi^\epsilon_{\alpha_{p(2r-1)}}(t_{p(2r-1)}) \varphi^\epsilon_{\alpha_{p(2r)}}(t_{p(2r)})), \end{aligned} \tag{2.10}$$

where $\mathcal{P}_n$ is the set of those permutations p of $\{1 \cdots 2n\}$ for which $p(2r-1) < p(2r)$ and $p(2r-1) < p(2r+1)$, and $\operatorname{sgn} p$ is the parity of p with respect to $\{1 \cdots 2n\}$.

The two-point correlation functions are

$$\begin{aligned} &\omega_A(\varphi^\epsilon_\alpha(s)\varphi^\epsilon_\beta(t)) \\ &= \sum_{\gamma=1}^{N^2} |\mu_\gamma|^2 [\overline{\lambda^\gamma_\alpha} \lambda^\gamma_\beta (f^\epsilon_s, (1-A) f^\epsilon_t) + \lambda^\gamma_\alpha \overline{\lambda^\gamma_\beta}(f^\epsilon_t, A f^\epsilon_s)] \\ &= c_{\alpha\beta}(f^\epsilon_t, A f^\epsilon_s) + c_{\beta\alpha}(f^\epsilon_s, (1-A) f^\epsilon_t), \end{aligned} \tag{2.11}$$

where we have set

$$c_{\alpha\beta} = \sum_{\gamma=1}^{N^2} |\mu_\gamma|^2 \lambda^\gamma_\alpha \overline{\lambda^\gamma_\beta}. \tag{2.12}$$

$\{c_{\alpha\beta}\}$ is the general form of a positive matrix in $M(N^2)$.

Using a boson reservoir would lead to an analogous result, with the following modifications (see Refs. 14 and 9):

the determinant in (2.6) replaced by a permanent,

no alternating sign in (2.10),

$(1 - A)$ replaced by $(1 + A)$ in (2.11),

A is a positive, not necessarily bounded, operator,

the gauge-invariant KMS state obtained by setting $A = (\exp[\beta(h - \mu)] - 1)^{-1}$

(there exist also KMS states which are not gauge-invariant and we will not consider them). Note that for A to be positive the energy spectrum has to be contained in $[\mu, +\infty)$.

(2.11) is the generalization of (2.10) of I.

Vacuum state case: The vacuum state is given by $A = 0$ and (2.11) reduces to (2.10) of I. The alternating sign in (2.10) does not affect the validity of Appendices I and II of I (see Appendix), in which it is shown that:

(I) the series (2.9) [(2.20) in I] converges uniformly in $0 \leq \epsilon \leq \epsilon_0 \neq 0$;

(II) as $\epsilon \to 0$, the series (2.9) converges term by term to

$$\left(x_t \middle| \left[\sum_{n=0}^{\infty} \int_{0 \leq t_n \leq \cdots \leq t_1 \leq t} dt_1 \cdots dt_n S(t_1) \cdots S(t_n) A\right] y_t\right) \tag{2.13}$$

[(2.27) of I], where

$$S(t) = \tfrac{1}{2} \sum_{\alpha,\beta=1}^{N^2} c_{\alpha\beta}\{[F_\beta(t), (\cdot)]F_\alpha(t) + F_\beta(t)[(\cdot), F_\alpha(t)]\}.$$

Therefore, as in I, going over to the Heisenberg picture, one obtains a completely positive identity-preserving semigroup, in the limit $\epsilon \to 0$. Its generator is given by

$$\begin{aligned} L^* A &= i[H_S, A] + \tfrac{1}{2} \sum_{\alpha,\beta=1}^{N^2} c_{\alpha\beta}\{[[F_\beta, A]F_\alpha + F_\beta[A, F_\alpha]\} \qquad (2.14) \\ &= i[H_S + H_1, A] + \tfrac{1}{2} \sum_{i,j=1}^{N^2-1} c_{ij}\{[[F_j, A]F_i + F_j[A, F_i]\}, \qquad (2.14') \end{aligned}$$

with

$$H_1 = -N^{-1/2} \sum_{k=1}^{N^2-1} (\operatorname{Im} c_{kN^2}) F_k. \tag{2.15}$$

In the Schrödinger picture,

$$L\rho = -i[H_S + H_1, \rho] + \tfrac{1}{2} \sum_{i,j=1}^{N^2-1} c_{ij}\{[[F_i, \rho F_j] + [F_i \rho, F_j]\}. \tag{2.16}$$

This is the general form of the generator of a completely positive dynamical semigroup of an N-level system.

Some generalization to dynamical semigroups of a system with infinitely many levels is possible. Consider the generator of a norm-continuous completely positive identity-preserving semigroup of $\mathcal{B}(\mathcal{H})$[3]

$$L^* A = i[H, A] - \tfrac{1}{2} \sum_{i \geq 1} \{A_i^* A_i, A\} + \sum_{i \geq 1} A_i^* A A_i, \tag{2.17}$$

where $H, A_i \in \mathcal{B}(\mathcal{H})$ and $\sum_{i \geq 1} A_i^* A_i$ converges ultraweakly.[15] It can be transformed to a form similar to (2.14) by defining:

$$B_{+i}=(1/\sqrt{2})(A_i^*+A_i),\quad B_{-i}=(i/\sqrt{2})(A_i^*-A_i),$$
$$F_\alpha=B_\alpha/\|B_\alpha\|,\quad (\alpha=\pm i,\ i\geq 1), \tag{2.18}$$

in terms of which L^* can be written as

$$L^*A=i[H,A]+\tfrac{1}{2}\sum_{\alpha,\beta}c_{\alpha\beta}\{[F_\beta,A]F_\alpha+F_\beta[A,F_\alpha]\} \tag{2.17'}$$

with

$$c_{\alpha\beta}=\tfrac{1}{2}\delta_{|\alpha|\,|\beta|}i^{(\mathrm{sgn}\beta-\mathrm{sgn}\alpha)/2}\|B_\alpha\|\,\|B_\beta\|. \tag{2.19}$$

In this case, the interaction V^ϵ can be written as

$$V^\epsilon=\sum_{\alpha\neq 0}F_\alpha\otimes\varphi_\alpha(f^\epsilon)=\sum_{i\geq 1}A_i\otimes a_i(f^\epsilon)^*+A_i^*\otimes a_i(f^\epsilon) \tag{2.20}$$

with

$$\varphi_{+i}(f)=(1/\sqrt{2})\|B_{+i}\|\,(a_i(f)^*+a_i(f)),$$
$$\varphi_{-i}(f)=(i/\sqrt{2})\|B_{-i}\|\,(a_i(f)^*-a_i(f)). \tag{2.21}$$

Under the additional condition $\sum_{\alpha\beta}|c_{\alpha\beta}|<\infty$, which implies that the series in (2.17) are norm convergent, the arguments of the Appendix and of Appendices I and II of I extend to this case, thus leading again to a completely positive dynamical semigroup in the limit $\epsilon\to 0$ (note that no "orthonormality" of the F_α's is needed in the course of the proof).

3. KMS STATES AND INFINITE TEMPERATURE LIMIT

In the sequel, for the sake of simplifying the notation, we shall absorbe μ into h.

Substituting (2.7) for A in (2.11) and setting $g^\epsilon(\omega)=|f^\epsilon(\omega)|^2$, we get

$$\omega_A(\varphi_\alpha^\epsilon(s)\varphi_\beta^\epsilon(t))=\int_{-\infty}^{+\infty}d\omega\,\exp[-i\omega(s-t)]\{g^\epsilon(\omega)c_{\alpha\beta}+g^\epsilon(-\omega)c_{\beta\alpha}\}\times[\exp(\beta\omega)+1]^{-1}. \tag{3.1}$$

For $\beta\neq 0$ as well as in the limit $\beta\to\infty$ the expression (3.1) cannot be made to tend to a δ function as $\epsilon\to 0$, since the limit of its Fourier transform is not a constant, no matter how one chooses f^ϵ (see, e.g., Ref. 16). In other words, the correlation functions of the reservoir have a finite decay time even in the limit of "white noise".

In the infinite temperature limit ($\beta\to 0$), A tends to $\frac{1}{2}$, and we are left with the expression

$$\tfrac{1}{2}\{c_{\beta\alpha}(f_s^\epsilon,f_t^\epsilon)+c_{\alpha\beta}(f_t^\epsilon,f_s^\epsilon)\}=\mathrm{Re}c_{\alpha\beta}\,\mathrm{Re}(f_s^\epsilon,f_t^\epsilon)-\mathrm{Im}c_{\alpha\beta}\,\mathrm{Im}(f_s^\epsilon,f_t^\epsilon). \tag{3.2}$$

If f^ϵ is chosen according to (2.4), this is the same result as the one found in the vacuum state case, apart from the replacement of $\{c_{\alpha\beta}\}$ by $\{\mathrm{Re}c_{\alpha\beta}\}$, a real and symmetric matrix. Note that, if we choose from the outset $\{c_{\alpha\beta}\}$ to be real and symmetric, we can take, in place of f^ϵ, a function with a positive energy spectrum, e.g.,

$$f^{\epsilon(+)}(\omega)=\pi^{-1/2}\theta(\omega)\exp[-\epsilon^2\omega^2/8]. \tag{3.3}$$

Indeed, $\mathrm{Re}(f_s^{\epsilon(+)},f_t^{\epsilon(+)})=(f_s^\epsilon,f_t^\epsilon)$, whereas $\mathrm{Im}(f_s^{\epsilon(+)},f_t^{\epsilon(+)})$, which, as $\epsilon\to 0$, behaves like $(s-t)^{-1}$, disappears from (3.2) because of the special choice of $\{c_{\alpha\beta}\}$.

For bosons, the two-point correlation function is given by

$$\int_{-\infty}^{+\infty}d\omega\,\exp[-i\omega(s-t)]\{g^\epsilon(\omega)c_{\alpha\beta}-g^\epsilon(-\omega)c_{\beta\alpha}\}\times(\exp(\beta\omega)-1)^{-1}, \tag{3.4}$$

where $g^\epsilon(\omega)$ has to vanish for $\omega\leq 0$, in order for the operator $A=(\exp(\beta h)-1)^{-1}$ to be positive. An analogous argument as for fermions applies here to exclude that (3.4) tends to a δ function as $\epsilon\to 0$, for $\beta\neq 0$ as well as for $\beta\to\infty$.

The limit of infinite temperature can be performed also in this case, with the one-particle energy spectrum chosen as $[0,+\infty)$. However, one must compensate for the divergence of $A=(\exp(\beta h)-1)^{-1}$ as $\beta\to 0$ by giving the interaction Hamiltonian an appropriate temperature dependence which ensures that the integrand in (3.4) remains finite in the limit. This can be achieved by replacing f^ϵ with the following temperature-dependent function which tends to zero pointwise as $\beta\to 0$:

$$f^{\epsilon,\beta}(\omega)=(2\pi)^{-1/2}\theta(\omega)(\beta\omega)^{1/2}\exp[-\epsilon^2\omega^2/8]. \tag{3.5}$$

Then

$$g^{\epsilon,\beta}(\omega)=(2\pi)^{-1}\theta(\omega)\beta\omega\exp[-\epsilon^2\omega^2/4] \tag{3.5'}$$

and upon insertion of (3.5') into (3.4), supposing $\{c_{\alpha\beta}\}$ real and symmetric, the Fourier transform of the correlation function becomes

$$c_{\alpha\beta}(2\pi)^{-1}\exp[-\epsilon^2\omega^2/4][\beta\omega/(\exp(\beta\omega)-1)]\times(\theta(\omega)+\theta(-\omega)), \tag{3.6}$$

which tends to the Fourier transform of $c_{\alpha\beta}\delta(s-t)$ as $\epsilon\to 0$, $\beta\to 0$. If we choose $\beta=\pi\epsilon^2$, then $\|f^{\epsilon,\beta}\|=1$ for all $\epsilon\neq 0$. The general form of a generator with a real and symmetric matrix $\{c_{ij}\}$ and no shift to the Hamiltonian is found again.

The singular reservoir limit at infinite temperature is less unphysical than the limit in the vacuum state, because the one-particle energy spectrum is the positive half-line. However, while in the vacuum state any completely positive dynamical semigroup can be obtained (for N-level systems), at infinite temperature only those semigroups are found, whose generator can be expressed by means of a real symmetric $\{c_{\alpha\beta}\}$. Such generators satisfy $L\mathbb{1}=0$, which means (when $\mathcal{H}$ is finite-dimensional) that they leave the central state invariant. This is not surprising, as the central state is the one in thermal equilibrium with the reservoir at infinite temperature. However, if $\dim\mathcal{H}>2$, the semigroup obtained in this way is not the most general one, among those leaving the central state invariant.

Comparing this result with Sec. 3 of I, one sees that the coupling of an N-level system to a reservoir at infinite temperature is equivalent to the addition of a stochastic term to the Hamiltonian.

4. A REMARK ON FERMI SYSTEMS

If S represents a system of Fermi particles and R is itself a fermion reservoir, we should require that the creation and annihilation operators $A_i^\#$ of S anticommute with the fields of R.

With the hypothesis that the free evolution of S preserves the total fermion number, this can be realized as follows.

Let $\mathcal{A}^e$ (respectively, $\mathcal{A}^0$) be the vector span of the even (respectively, odd) polynomials in the creation and annihilation operators $A_i^{\#}$ of S. Perform the following identifications:

$$\mathcal{A} \to \mathcal{A} \otimes \mathbb{1}_R; \quad \varphi_i \to (-1)^{N_S} \otimes \varphi_i; \tag{4.1}$$

where N_S is the total fermion number operator of S.

We write the coupling V^ϵ as

$$V^\epsilon = \sum_\alpha F_\alpha^e \hat{\varphi}_\alpha^\epsilon + i \sum_\beta F_\beta^0 \hat{\varphi}_\beta^\epsilon, \tag{4.2}$$

or, equivalently, as

$$V^\epsilon = \sum_\alpha G_\alpha^e \otimes \varphi_\alpha^\epsilon + \sum_\beta G_\beta^0 \otimes \varphi_\beta^\epsilon, \tag{4.3}$$

where

$$\begin{aligned} &F_\alpha^e = F_\alpha^{e*} \in \mathcal{A}^e, \quad F_\beta^0 = F_\beta^{0*} \in \mathcal{A}^0, \\ &\hat{\varphi}_\gamma^\epsilon = (-1)^{N_S} \otimes \varphi_\gamma^\epsilon \quad (\gamma = \alpha, \beta), \\ &G_\alpha^e = G_\alpha^{e*} = F_\alpha^e (-1)^{N_S} \in \mathcal{A}^e, \\ &G_\beta^0 = G_\beta^{0*} = i F_\beta^0 (-1)^{N_S} \in \mathcal{A}^0. \end{aligned} \tag{4.4}$$

Clearly, the singular reservoir limit leads again to a completely positive dynamical semigroup, whose generator is expressed in terms of the orthonormal basis $\{G_\alpha^e\} \cup \{G_\beta^0\}$.

In the special case when the generator splits into an "even" and an "odd" part

$$\begin{aligned} L^* A = {} & i[H_S, A] + \tfrac{1}{2} \sum_{\alpha,\beta} c_{\alpha\beta}^{(e)} \{[G_\beta^e, A] G_\alpha^e + G_\beta^e [A, G_\alpha^e]\} \\ & + \tfrac{1}{2} \sum_{\gamma,\delta} c_{\gamma\delta}^{(0)} \{[G_\delta^0, A] G_\gamma^0 + G_\delta^0 [A, G_\gamma^0]\}, \end{aligned} \tag{4.5}$$

and if one restricts consideration to the time evolution of even observables, then L^* can be written in the usual form (2.14), with the same operators F_α^e and F_β^0 which appear in (4.2). Hence in this case no new feature is introduced by the assumption of anticommutativity between the odd operators of the system and the Fermi fields of the reservoir.

We can also construct a coupled time evolution which preserves the total relative fermion number of the compound system by suitably coupling the even gauge-invariant operators of S to Bose fields φ_B (to be identified with $\mathbb{1}_S \otimes \varphi_B$) and the odd, linear in the $A_i^{\#}$'s, operators of S to Fermi fields φ_F [to be identified with $(-1)^{N_S} \otimes \varphi_F$]. Again, for what concerns the time evolution of the even observables, nothing is changed with respect to the models of the previous sections.

ACKNOWLEDGMENTS

We are indebted to A. Kossakowski for discussions and suggestions.

APPENDIX

We want to show that the replacement of the formula valid for Bose fields

$$\omega(\varphi_{\alpha_1}^\epsilon(t_1) \cdots \varphi_{\alpha_{2n}}^\epsilon(t_{2n})) = \sum_{p \in \rho_n} \prod_{r=1}^n \omega(\varphi_{\alpha_{p(2r-1)}}^\epsilon(t_{p(2r-1)}) \varphi_{\alpha_{p(2r)}}^\epsilon(t_{p(2r)})) \tag{A1}$$

by the formula valid for Fermi fields

$$\omega(\varphi_{\alpha_1}^\epsilon(t_1) \cdots \varphi_{\alpha_{2n}}^\epsilon(t_{2n})) = \sum_{p \in \rho_n} \operatorname{sgn} p \prod_{r=1}^n \omega(\varphi_{\alpha_{p(2r-1)}}^\epsilon(t_{p(2r-1)}) \varphi_{\alpha_{p(2r)}}^\epsilon(t_{p(2r)})) \tag{A2}$$

does not affect the validity of the proofs in Appendices I and II of I.

The uniform convergence of the Dyson series: Consider

$$\begin{aligned} &(\Omega | [M^\epsilon(t_1) \cdots M^\epsilon(t_{2n})(A \otimes \mathbb{1})] \Omega) \\ &\quad = (-1)^n \sum_{[i,k,2n]} (-1)^{2n-k} (\Omega | [V^\epsilon(t_{i_1}) \cdots V^\epsilon(t_{i_k}) \\ &\qquad \times (A \otimes \mathbb{1}) V^\epsilon(t_{i_{k+1}}) \cdots V^\epsilon(t_{i_{2n}})] \Omega), \end{aligned} \tag{A3}$$

where the summation is extended over all 2^{2n} partitions $[i,k,2n] = (i_1 \cdots i_k)(i_{k+1} \cdots i_{2n})$ of $\{1 \cdots 2n\}$, where $i_1 < \cdots < i_k$ and $i_{k+1} > \cdots > i_{2n}$. The norm of a summand of (A3) is majorized by

$$\begin{aligned} &\sum_{\alpha_1 \cdots \alpha_{2n}} \|A\| F^{2n} | \omega(\varphi_{\alpha_1}^\epsilon(t_{i_1}) \cdots \varphi_{\alpha_k}^\epsilon(t_{i_k}) \varphi_{\alpha_{k+1}}^\epsilon(t_{i_{k+1}}) \cdots \\ &\qquad \times \varphi_{\alpha_{2n}}^\epsilon(t_{i_{2n}})) | \leq \sum_{\alpha_1 \cdots \alpha_{2n}} \|A\| F^{2n} \sum_{p \in \rho_n} \prod_{r=1}^n \\ &\qquad \times | \omega(\varphi_{\alpha_{p(2r-1)}}^\epsilon(t_{i_{p(2r-1)}}) \varphi_{\alpha_{p(2r)}}^\epsilon(t_{i_{p(2r)}})) |, \end{aligned}$$

where $F = \sup_\alpha \|F_\alpha\|$, independently of which of the decompositions (A1), (A2) is used. Then the arguments of Appendix I of I apply, provided $\sum_{\alpha\beta} |c_{\alpha\beta}| = C^2 < \infty$, which is trivial for N-level systems, and is to be assumed otherwise.

The reduced dynamics in the limit $\epsilon \to 0$: We have to show that the contributions from permutations with the minus sign in (A1) vanish in the limit $\epsilon \to 0$. Expand

$$\omega(\varphi_{\alpha_1}^\epsilon(t_{i_1}) \cdots \varphi_{\alpha_k}^\epsilon(t_{i_k}) \varphi_{\alpha_{k+1}}^\epsilon(t_{i_{k+1}}) \cdots \varphi_{\alpha_{2n}}^\epsilon(t_{i_{2n}}))$$

as a summation over all possible manners of pairing the field operators. A contribution with the minus sign can only come from a permutation containing "overlapping pairings," of the form

$$\omega(\varphi_{\alpha_m}^\epsilon(t_{i_m}) \varphi_{\alpha_p}^\epsilon(t_{i_p})) \omega(\varphi_{\alpha_n}^\epsilon(t_{i_n}) \varphi_{\alpha_q}^\epsilon(t_{i_q})) \tag{A4}$$

with $m < n < p < q$. According to whether none, one, two, three, or all of the numbers m,n,p,q, are larger than k, one has, respectively, one of the following inequalities:

$$\begin{aligned} &t_{i_m} \geq t_{i_n} \geq t_{i_p} \geq t_{i_q}; \quad t_{i_m} \geq t_{i_n} \geq t_{i_p}; \quad t_{i_m} \geq t_{i_n}, \quad t_{i_p} \leq t_{i_q}; \\ &t_{i_n} \leq t_{i_p} \leq t_{i_q}; \quad t_{i_m} \leq t_{i_n} \leq t_{i_p} \leq t_{i_q}. \end{aligned}$$

In the limit $\epsilon \to 0$, we have a product $\delta(t_{i_m} - t_{i_p}) \delta(t_{i_n} - t_{i_q})$. In any of the listed cases, all the four t's have to be equal. Hence the integral over $dt_1 \cdots dt_{2n}$ vanishes. We are left only with permutations not containing such overlapping pairings, which have a plus sign, like in the boson case, and the proof in the Appendix II of I is still valid.

*A fellowship from the Italian Ministry of Public Education is acknowledged.
[1]V. Gorini and A. Kossakowski, J. Math. Phys. **17**, 1298 (1976).
[2]V. Gorini, A. Kossakowski, and E.C.G. Sudarshan, J. Math. Phys. **17**, 821 (1976).
[3]G. Lindblad, "On the Generators of Quantum Dynamical Semigroups," Preprint TRITA-TFY-75-1 (Royal Institute of Technology, Stockholm, 1975).
[4]E.B. Davies, Z. Wahrscheinlichkeitstheorie 23, 261 (1972).
[5]K. Kruas, Ann. Phys. (N.Y.) **64**, 311 (1971).
[6]L. Accardi, Non-Relativistic Quantum Mechanics as a Noncommutative Markoff Process, preprint Lab. di Cibernetica del C.N.R., Arco Felice, Naples, 1975.
[7]E.B. Davies, Comm. Math. Phys. **33**, 171 (1973).
[8]E.B. Davies, Comm. Math. Phys. **39**, 91 (1974).
[9]J.V. Pulè, Comm. Math. Phys. **38**, 241 (1974).
[10]K. Hepp and E. Lieb, Helv. Phys. Acta **46**, 573 (1973).
[11]R.T. Powers and E. Størmer, Comm. Math. Phys. **16**, 1 (1970).
[12]F. Rocca, M. Sirugue, and D. Testard, Comm. Math. Phys. **13**, 317 (1969).
[13]A. Verbeure and E. Balslev, Comm. Math. Phys. **7**, 55 (1968).
[14]F. Rocca, M. Sirugue, and D. Testard, Comm. Math. Phys. **19**, 119 (1970).
[15]Note that, contrary to the case of an N-level system, no unique separation of (2.17) into a Hamiltonian and a dissipative part exists if H is infinite-dimensional.
[16]K. Yosida, in *Functional Analysis, II Ed.* (Springer-Verlag, Berlin-Heidelberg-New York, 1968), Chap. VI, Cor. 2, p. 152.

Some properties of canonical invariantly relaxing (CIR) systems

Tak-Yuen Chong*

Department of Physics, St. Louis University, St. Louis, Missouri 63103
(Received 20 June 1975; final revised manuscript received 26 February 1976)

In this paper, two general properties of the conserved discrete states CIR systems have been found: (1) the equal spacing of and a general expression for the characteristic roots of their transition-rates-matrices; (2) a general formulation for their conditional probabilities. The discussion is also extended to the disappearing discrete-states CIR systems which include the infinite level harmonic oscillators as a special case.

I. INTRODUCTION

A canonical invariantly relaxing (CIR) system is defined as a system which, when it relaxes from one (thermal) canonical equilibrium to another, will preserve the canonical distribution, i.e., will be describable all the time by a distribution of the same form as that for canonical equilibrium but with a time-dependent temperature. Thus if a $(N+1)$-discrete-states CIR system, i.e., a CIR system with $N+1$ discrete allowable energy states—$0=\epsilon_0<\epsilon_1<\cdots<\epsilon_N$, where N may be a finite or enumerable integer—is allowed to relax from one canonical equilibrium of temperature T_0 to a final one of temperature T_∞, its distribution at any time t can then be expressed as

$$P_{\epsilon_i}(t)=\frac{g_i\exp[-\beta(t)\epsilon_i]}{\sum_{i=0}^{N}g_i\exp[-\beta(t)\epsilon_i]},\quad i=0,1,\ldots,N, \tag{1.1}$$

where g_i is the degeneracy of the ith energy state, $\beta(t)=1/kT(t)$ with k standing for the Boltzmann constant, and $T(t)$ the time-dependent temperature such that $T(0)=T_0$ and $T(\infty)=T_\infty$. The nonequilibrium states of a CIR system, when it relaxes from one canonical equilibrium to another, therefore can be described exactly by time-dependent thermodynamic functions, which are the same as the usual equilibrium thermodynamic functions with only an additional property of being time-dependent.[1]

Examples of CIR systems, in addition to the trivial two-states systems, are: the vibrational relaxation of a set of harmonic oscillators, classical as well as quantal, subjected to Landau–Teller transition probabilities and in contact with a heat bath; the transitional relaxation of a classical hard-sphere Rayleigh gas; the transitional relaxation of a classical hard-sphere Lorentz gas obeying the Maxwellian force law[2]; the vibrational relaxation of two coupled systems of harmonic oscillators[3]; the stochastic model of Hill and Plesner, which can satisfactorily describe the relaxation problems of quantal hard-sphere Rayleigh gas, ideal gas, two-component and three-component lattice gasses.[4] Chandheri and Scheigger have also confirmed the validity of canonical invariance in hydrodynamic phenomena under conditions analogous to those obtained by Shuler and co-workers in Ref. 2.[5]

A long standing problem in the development of exact time-dependent thermodynamics is that there still does not exist a general set of functions which can be used to satisfactorily describe the properties of nonequilibrium physical systems. Thus, further investigation into the properties of CIR systems is of interest in the sense that it may give some new insight into the development of nonequilibrium thermodynamics and the generalization of the time-dependent thermodynamic functions. For example, if one can somehow compare the relaxation phenomena of a CIR system when it relaxes from an unknown initial nonequilibrium state to that when it relaxes from an initial canonical state, both with the same initial average energy and to the same final canonical state, one may possibly obtain some information about that initial nonequilibrium state. Research work has been carried out based on this thinking; particular attention was paid to the discrete-states CIR systems since it is well known that the relaxation behavior of a system in a time process is predominantly determined by the characteristic roots of the transition rates matrix (TRM) formed by its transition rates,[6] and that it has been shown by Shuler and co-workers that the transition rates of the discrete-states CIR systems are of some specific form [Eq. (2.1)].

In this paper, two general properties of the conserved $(N+1)$-discrete-states CIR systems will be presented:

(1) The characteristic roots of their TRM's, as found in Sec. II, are equally spaced, and can be expressed in the general form: $\lambda_i=-A_{10}(1+|a|d)i$, $i=0,1,\ldots,N$ [Eq. (2.23)].

(2) A general expression for the conditional probabilities $P_{ij}(0,t)$, $i,j=0,1,\ldots,N$, is formulated in Sec. III as the coefficients of a generating function [Eq. (3.9)]. A final section (Sec. IV) is devoted to the discussion of "disappearing" discrete-states CIR systems, which includes the infinite-level harmonic oscillators as a special case.

II. THE CHARACTERISTIC ROOTS

As shown by Shuler and co-workers, the transition rates of a $(N+1)$-discrete-states CIR system have in the form[2]

$$A_{ij}=A_{10}[i\Delta(i-1,j)+(ai+b)\,d\Delta(i+1,j)],\quad i,j=0,1,\ldots,N, \tag{2.1}$$

where A_{ij} denotes the transition rate of the system from the ith to the jth state,

$$\Delta(i,j)=\begin{cases}1, & i=j,\\ 0, & \text{otherwise},\end{cases}$$

i.e., a delta function and

$$a=\frac{2g_2}{g_1}-\frac{g_1}{g_0},\quad b=\frac{g_1}{g_0},\quad d=\exp[-\beta(\infty)\epsilon_1]. \tag{2.2}$$

It should be noticed that Eq. (2.1) is a necessary and sufficient condition for a discrete-states system to be CIR, which thus requires its transition rates be of nearest-neighbor type, energy levels be equally spaced, and degeneracy ratio be in the form of $g_{k+1}/g_k = a(k+f)/k+1$ with $f=b/a$.

An $(N+1)$-discrete-states CIR system will be called conserved if it cannot disappear, which requires

$$A_{0,-1}=0 \quad \text{and} \quad A_{N,N+1}=aN+b=0 \tag{2.3}$$

and therefore

$$b/a=-N. \tag{2.4}$$

Equation (2.4) then implies that

$$a<0 \quad \text{since } 0<b. \tag{2.5}$$

For a conserved discrete-states CIR system, the transition rates (2.1) can then be expressed in a tridiagonal matrix $\mathbf{A}^{(N)}$:

$$\begin{pmatrix} -a_0 & a_0 & & & & & \\ b_1 & -(a_1+b_1) & a_1 & & & & \\ & b_2 & -(a_2+b_2) & a_2 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & a_{N-2} & & & \\ & & & b_{N-1} & -(a_{N-1}+b_{N-1}) & a_{N-1} & \\ & & & & b_N & -b_N \end{pmatrix} \tag{2.6}$$

where the representations $a_i=A_{i,i+1}$ and $b_i=A_{i,i-1}$ have been used; it should be noticed that $A_{i,i}=-(a_i+b_i)$.

Finding the characteristic roots of a matrix generally is not an easy task, especially when exact and explicit expressions are sought which at the same time can be generalized to matrices of the same form but of arbitrary orders. In trying to find the characteristic roots of the CIR matrix (2.6) of any order N, the major difficulties with the conventional determinant expansions method obviously lie in obtaining the characteristic polynomial, and most of all, in finding the roots of this polynomial. The alternative mathematical induction approach does not readily give a straight proof of the final result either, although one can easily obtain for the simple cases of $N=1$ and $N=2$ the characteristic roots as $\lambda_i=-A_{10}(1+|a|d)i$ with $i=0,1$ for $N=1$ and $i=0,1,2$ for $N=2$, and anticipate a general solution in the form of $\lambda_i=-A_{10}(1+|a|d)i$ with $i=0,1,\ldots,N$.

A successful method of solution employing finite difference technique was discovered after noticing some properties of the CIR matrix which were observed by Reuter and Ledermann on the analogous birth–death matrix.[7] Defining the nth modified section of $\mathbf{A}^{(N)}$ as the matrix $\mathbf{A}_m^{(n)}$ obtained by deleting the rows and columns of $\mathbf{A}^{(N)}$ larger than the $(n+1)$th, one can then obtain the following properties of the CIR matrices:

(a) Denote the characteristic polynomial of $\mathbf{A}_m^{(n)}$ by $\psi_n(\lambda)$ (notice that $\psi_N(\lambda)$ is the characteristic polynomial of $\mathbf{A}^{(N)}$). Then the $\psi_n(\lambda)$'s satisfy the following recurrence relation [cf. Ref. 7, Eq. (1.55)]:

$$\psi_n(\lambda)-(\lambda+a_{n-1}+b_n)\psi_{n-1}(\lambda)+a_{n-1}b_{n-1}\psi_{n-2}(\lambda)=0 \quad \text{with } \psi_{-1}(\lambda)=0,\ \psi_0(\lambda)=\lambda. \tag{2.7}$$

(b) The characteristic roots of $\psi_N(\lambda)$ are all distinct; one of them is zero and the others are negative (cf. Ref. 7. Theorem 2). Denote them as

$$\lambda_N<\lambda_{N-1}<\cdots<\lambda_1<\lambda_0=0; \tag{2.8}$$

one can then write

$$\psi_N(\lambda)=|\lambda\mathbf{I}-\mathbf{A}^{(N)}|=\prod_{r=0}^{N}(\lambda-\lambda_r). \tag{2.9}$$

Thus the major task in finding the desired characteristic roots is to solve the set of finite difference equations (2.7), and thereupon obtain the roots of $\psi_N(\lambda)$.

The technique of finding characteristic roots of a matrix via solving a set of finite difference equations, though not new, is seldom used. In the following it can be seen that this technique yields results in a rather fancy way, and thus should deserve more attention in the future. Some previous interesting work employing this technique has been done by Elliot, who worked on second-order finite-difference equations with constant coefficients.[8] The set of finite difference equations dealt with here are also of second order but have variable coefficients, which renders them more complicated.

On substituting the transition rates (2.1) into the recurrence relation (2.7) and letting $n\to n+1$, one obtains for a $(N+1)$-discrete-states CIR system the following set of finite difference equations:

$$\psi_{n+2}(\lambda) - [\lambda + A_{10}(1+ad)(n+1) + A_{10}(1+bd)]\psi_{n+1}(\lambda) + (A_{10})^2 d[an + (a+b)](n+1)\psi_n(\lambda) = 0, \quad n = 0, 1, \ldots, N. \tag{2.10}$$

With the transform

$$\psi_n(\lambda) = \Gamma(n+1)\phi_n(\lambda), \tag{2.11}$$

where $\Gamma(n)$ denotes a gamma function, Eq. (2.10) then, after being divided by $(n+1)\Gamma(n+1)$, appears in the form of the Laplace difference equation:

$$(n+2)\phi_{n+2}(\lambda) - [A_{10}(1+ad)(n+1) + A_{10}(1+bd) + \lambda]\phi_{n+1}(\lambda) + (A_{10})^2 d[an + (a+b)]\phi_n(\lambda) = 0, \tag{2.12}$$

for which techniques of a solution have been well established.[9]

Let

$$\phi_n(\lambda) = \int_c z^{n-1} F(z, \lambda)\,dz \tag{2.13}$$

and thus

$$n\phi_n(\lambda) = [z^n F(z, \lambda)]_c - \int_c z^n F'(z, \lambda)\,dz, \tag{2.14}$$

where z is a complex variable, $F'(z, \lambda) = (d/dz)F(z, \lambda)$ and c is an integration path which will be specified in the following.

Equation (2.12) then becomes, on making use of the above transform and rearranging,

$$[(z^2 - A_{10}(1+ad)z + (A_{10})^2 ad)z^n F(z, \lambda)]_c - \int_c \{[z^2 - A_{10}(1+ad)z + (A_{10})^2 ad]zF'(z, \lambda) + ([A_{10}(1+bd) + \lambda]z - (A_{10})^2(a+b)d)F(z, \lambda)\} z^{n-1}\,dz = 0. \tag{2.15}$$

The conditions applied to Eq. (2.15) are:

(a) The function $F(z, \lambda)$ is chosen so that the integrand of its second term becomes zero.

(b) The integration path c is chosen so that its first term becomes zero.

Condition (a) requires that

$$\frac{F'(z, \lambda)}{F(z, \lambda)} = -\frac{(A_{10}(1+bd) + \lambda)z - (A_{10})^2(a+b)d}{z(z^2 - A_{10}(1+ad)z + (A_{10})^2 ad)} = \left(1 + \frac{b}{a}\right)\frac{1}{z} - \left(1 + \frac{\lambda}{A_{10}(1-ad)}\right)\frac{1}{z - A_{10}} - \left(\frac{b}{a} - \frac{\lambda}{A_{10}(1-ad)}\right)\frac{1}{z - A_{10}\,ad}$$

which therefore implies that

$$F(z, \lambda) = z^{-\alpha_0}(z - \mu_1)^{\gamma_1}(z - \mu_2)^{\gamma_2}, \tag{2.16}$$

where

$$-\alpha_0 = 1 + b/a = 1 - N, \tag{2.17}$$

$$\gamma_1 = -1 - \lambda/A_{10}(1 - ad), \tag{2.18}$$

$$\gamma_2 = N + \lambda/A_{10}(1 - ad), \tag{2.19}$$

and

$$\mu_1 = A_{10}, \quad \mu_2 = A_{10}\,ad \tag{2.20}$$

It should be noticed that μ_1 and μ_2 are the roots of the auxiliary equation

$$z^2 - A_{10}(1+ad)z + (A_{10})^2 ad = 0$$

and that $\mu_1 \neq \mu_2$.

On substituting Eq. (2.16) into Eq. (2.13) and assigning two integration paths satisfying condition (b), one can obtain two particular solutions for $\phi_n(\lambda)$; a linear combination of these two solutions with appropriate coefficients will then give a general solution of $\phi_n(\lambda)$. For the case $n = N$ of interest, one has (cf. Ref. 9, Sec. 15.1)

$$\phi_N(\lambda) = \frac{C_{N1}(\lambda)}{2\pi i}\oint_{c_1}(z - \mu_1)^{\gamma_1}(z - \mu_2)^{\gamma_2}\,dz + \frac{C_{N2}(\lambda)}{2\pi i}\oint_{c_2}(z - \mu_1)^{\gamma_1}(z - \mu_2)^{\gamma_2}\,dz \tag{2.21}$$

where $i = \sqrt{-1}$, and the integration path c_1 is a closed loop from the origin and round μ_1 in a positive sense (counterclockwise), with μ_1 as an inner point but not enclosing μ_2; a comparable description holds for c_2.

Without explicitly evaluating the integrations and the coefficients $C_{N1}(\lambda)$ and $C_{N2}(\lambda)$ of Eq. (2.21), one can discuss the roots of $\phi_N(\lambda)$, and thus of $\psi_N(\lambda)$ as follows.

Suppose that both γ_1 and γ_2 are nonintegers. With the transforms $t = \mu_1 z$ for the first integral and $t = \mu_2 z$ for the second one, and on denoting $k_1 = \mu_1/\mu_2$ and $k_2 = \mu_2/\mu_1$, Eq. (2.21) can easily be rewritten as (cf. Ref. 9, Sec. 15.52)

$$\phi_N(\lambda) = \frac{C_{N1}(\lambda)}{2\pi i}\mu_1^{\gamma_1+1}\mu_2^{\gamma_2}(1 - \exp(2\pi\gamma_1 i))\frac{\Gamma(\gamma_1+1)}{\Gamma(\gamma_1+2)} \times H(1, -\gamma_2; \gamma_1+2, k_1) + \frac{C_{N2}(\lambda)}{2\pi i}\mu_2^{\gamma_2+1}\mu_1^{\gamma_1} \times (1 - \exp(2\pi\gamma_2 i))\frac{\Gamma(\gamma_2+1)}{\Gamma(\gamma_2+2)}H(1, -\gamma_1; \gamma_2+2, k_2)$$

where $H(\,)$ denotes a hypergeometric function. (2.22)

Since $\mu_1 \neq \mu_2$, either $|k_1|$ or $|k_2|$ will be greater than 1, and thus either $H(1, -\gamma_2; \gamma_1+2, k_1)$ or $H(1, -\gamma_1; \gamma_2+2, k_2)$ will become a divergent hypergeometric series in λ; the assumption of γ_1 and γ_2 both being nonintegers therefore gives a $\phi_N(\lambda)$ which contradicts the fact that it is a polynomial of degree $N+1$ in λ.

It should be noticed that the validity of Eq. (2.22) requires $0 < R(\gamma_1+1)$ and $0 < R(\gamma_2+1)$; these can easily be seen to be satisfied on noticing that the values of λ important for locating the roots of $\phi_N(\lambda)$ are $-NA_{10}(1-ad) \leq \lambda \leq 0$; the upper bound has been verified as stated in Eq. (2.8), and the lower bound can be justified by the final result. [$R(\,)$ indicates the real part of the quantity enclosed.]

From the above one comes to the conclusion that the characteristic roots of $\psi_N(\lambda)$ can only possibly occur when either γ_1 or γ_2 is an integer, which further implies, as from Eqs. (2.18), (2.19), that only when γ_1, γ_2 and $\lambda/A_{10}(1-ad)$ all are integers. Since by Cauchy's integral theorem,[10] the vanishing of the integrals in

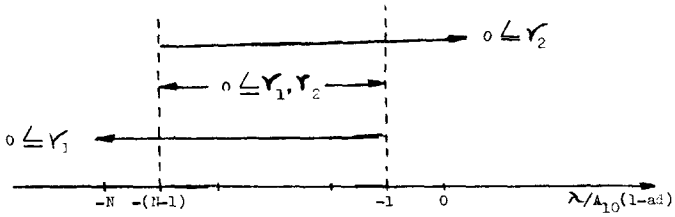

FIG. 1. Integral values of $\lambda/A_{10}(1-ad)$ which make both γ_1 and γ_2 positive.

Eq. (2.21) requires both γ_1 and γ_2 to be positive, one thus finally can locate N roots of $\psi_N(\lambda)$ at integral values of $\lambda/A_{10}(1-ad)$ which makes both γ_1 and γ_2 positive (cf. Fig. 1), i.e., at

$$\lambda_i = -A_{10}(1-ad)i, \quad i=1,2,\ldots,N,$$

which, together with the fact that $\lambda=0$ must be a root [cf. Eq. (2.8)], finally gives a general expression of the characteristic roots of the TRM of a conserved $(N+1)$-discrete-states CIR system:

$$\lambda_i = -A_{10}(1+|a|d)i, \quad i=0,1,\ldots,N. \tag{2.23}$$

III. THE CONDITIONAL PROBABILITIES

It is well known that, under the weak interaction of an external system, the instantaneous probability distribution $[\mathbf{P}(t)]$ (square-bracket notation is used to designate a row vector; unless otherwise stated, $[\mathbf{P}(t)]$ will be of dimension $N+1$) of a conserved $(N+1)$-discrete-states CIR system will be determined by the master equation[6]

$$\frac{d}{dt}[\mathbf{P}(t)] = [\mathbf{P}(t)]\mathbf{A}^{(N)}, \tag{3.1}$$

where $\mathbf{A}^{(N)}$ is given by Eq. (2.6).

Equation (3.1) can be equivalently expressed by the following set of differential-difference equations:

$$\begin{aligned}\frac{d}{dt}P_j(t) &= A_{10}[a(j-1)+b]dP_{j-1}(t) \\ &\quad - A_{10}[j+(aj+b)d]P_j(t) + A_{10}(j+1)P_{j+1}(t), \\ &\qquad\qquad j=0,1,\ldots,N-1 \\ \frac{d}{dt}P_N(t) &= A_{10}[a(N-1)+b]dP_{N-1}(t) - A_{10}NP_N(t),\end{aligned} \tag{3.2}$$

where it is to be noticed that for a conserved system, $P_j(t)=0$ for $j<0$ and $N<j$, and it will be assumed that the process starts at $t=0$.

A solution of the distribution problem can be achieved by using generating function technique as follows.

With the generating function

$$G(z,t) = \sum_{j=0}^{N} P_j(t)z^j \tag{3.3}$$

the above set of difference equations can be combined into a partial differential equation

$$\begin{aligned}&\frac{1}{A_{10}}\frac{\partial}{\partial t}G(z,t) - d(az-d^{-1})(z-1)\frac{\partial}{\partial z}G(z,t) \\ &\quad = bd(z-1)G(z,t) - d(aN+b)(z-1)P_N(t)z^N\end{aligned}$$

which finally gives, on noticing that $aN+b=0$ [Eq. (2.3)],

$$\begin{aligned}&\frac{1}{A_{10}}\frac{\partial}{\partial t}G(z,t) - d(az-d^{-1})(z-1)\frac{\partial}{\partial z}G(z,t) \\ &\quad = bd(z-1)G(z,t),\end{aligned} \tag{3.4}$$

whose solution can be obtained via solving its auxiliary equation[11]:

$$\frac{dt}{1/A_{10}} = \frac{dz}{-d(az-d^{-1})(z-1)} = \frac{dG(z,t)}{bd(z-1)} \times \frac{1}{G(z,t)}. \tag{3.5}$$

The first equation of the simultaneous equation (3.5) gives a solution

$$C_1 = (z-1)(az-d^{-1})^{-1}\exp[-A_{10}(1-ad)t] \tag{3.6a}$$

while the second one gives

$$C_2 = G(z,t)(az-d^{-1})^{b/a}, \tag{3.6b}$$

where C_1 and C_2 are constants. Thus the general solution of Eq. (3.5) can be expressed as

$$C_2 = f(C_1),$$

i.e.,

$$\begin{aligned}G(z,t) &= (az-d^{-1})^{-b/a} \\ &\quad \times f((z-1)(az-d^{-1})^{-1}\exp[-A_{10}(1-ad)t]),\end{aligned} \tag{3.7}$$

where the form of the function $f(\,)$ will be determined by initial conditions.

Following the procedures employed by Montroll and Shuler in treating the special case of infinite level harmonic oscillators with $a=b=1$,[12] one can easily obtain, on noticing $G(1,t)=1$, that under the initial condition

$$P_j(0) = \begin{cases}1, & \text{if } j=i, \\ 0, & \text{otherwise,}\end{cases} \tag{3.8}$$

the solution (3.7) can be explicitly determined as

$$G(z,t)\Big|_{P_i(0)=1} = \left(\frac{a-d^{-1}}{ay-d^{-1}}\right)^{b/a}\left(\frac{v}{ad}\right)^i \frac{(1-wz)^i}{(1-vz)^{i+b/a}}, \tag{3.9}$$

where

$$y = \exp[-A_{10}(1-ad)t],$$

$$v = \frac{a(y-1)}{ay-d^{-1}}, \quad w = \frac{y-ad}{y-1}.$$

The conditional probability $P_{ij}(0,t)$ (the probability of a system at state j at time t under the condition that it is at state i at time $t=0$) can then be identified as the coefficient of z^j of Eq. (3.9) when expanded into a polynomial of z. The probability distribution can finally be obtained via the relation

$$P_j(t) = \sum_{i=0}^{N} P_i(0)P_{ij}(0,t).$$

IV. THE DISAPPEARING CASES

The discussions and results in Secs. II and III hold for conserved discrete-states CIR systems which have

$$A_{N,N+1} = aN+b = 0, \tag{2.3}$$

and thus

$$a<0. \tag{2.5}$$

However, there also exist discrete-states CIR systems which do not obey Eq. (2.3). Examples of these are the quantal harmonic oscillators with infinite number of allowable states and with $a=b=1$,[12] and the decompositions of molecules at their highest allowable states. These systems can well be called "disappearing" or "escaping" systems so as to account for the reasonable explanation that they have a probability of "disappearing" or "escaping" at their highest allowable states; thus the infinite level harmonic oscillators can be said to have a probability of escaping to infinity, while the decomposition of molecules at their highest allowable states can be accounted for as due to a probability of disappearing (the cases of systems with $A_{0,-1}\neq 0$ will not be discussed here). While the "conserved" discrete-states CIR systems have their instantaneous probability distribution determined by Eq. (3.2), the "disappearing" discrete-states CIR systems follow a somewhat different set of differential-difference equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}P_j(t)=A_{10}[a(j-1)+b]dP_{j-1}(t)$$
$$-A_{10}[j+(aj+b)d]P_j(t)+A_{10}(j+1)P_{j+1}(t),$$
$$j=0,1,\ldots,N-1,$$
$$\frac{\mathrm{d}}{\mathrm{d}t}P_N(t)=A_{10}[a(N-1)+b]dP_{N-1}(t)$$
$$-A_{10}[N+(aN+b)d]P_N(t). \qquad (4.1)$$

The applying of the generating function (3.3) will then combine the above set of equations into a partial differential equation which is quite different from Eq. (3.4):

$$\frac{1}{A_{10}}\frac{\partial}{\partial t}G(z,t)-d(az-d^{-1})(z-1)\frac{\partial}{\partial z}G(z,t)$$
$$=bd(z-1)G(z,t)-(aN+b)dz^{N+1}P_N(t). \qquad (4.2)$$

Since the requirement of $|z|<1$ for the validity of the generating function (3.3) implies

$$\lim_{N\to\infty}(aN+b)z^{N+1}dP_N(t)=0 \qquad (4.3)$$

one sees that when N is large enough, Eq. (4.2) becomes practically the same as Eq. (3.4), and thus can conclude that the "disappearing" discrete-states CIR systems behave as if they are conserved; however, a similar conclusion generally cannot be ensured when N is small.

The finite $(N+1)$-discrete-states CIR systems with escaping probability at their highest allowable states and with $a=b=1$ can be called "finite level harmonic oscillators," since when N increases to infinity, their partial differential equation (3.3) will become the same as that for the well-known infinite level harmonic oscillators discussed by Montroll and Shuler.[12] This idea of finite level with escaping probability seems to be of interest in the sense that it may replace the somewhat vague idea of infinite levels, and will be further explored in the future.

ACKNOWLEDGMENTS

The author wishes to thank the referees for their kindly pointing out some confusing places in the original manuscript, and their valuable suggestions in preparing this writing.

*Corresponding address: 9 Bailey Cres., Scarborough, Ontario, Canada M1G 2P4

[1]H. C. Andersen, I. Oppenheim, K. E. Shuler, and G. H. Weiss, J. Chem. Phys. 41, 3012 (1964).
[2]H. C. Andersen, I. Oppenheim, K. E. Shuler, and G. H. Weiss, J. Math. Phys. 5, 522 (1964).
[3]R. Herman and R. J. Rubin, Phys. Fluids 2, 547 (1959). K. E. Shuler and G. H. Weiss, J. Chem. Phys. 45, 1105 (1966).
[4]T. L. Hill and I. W. Plesner, J. Chem. Phys. 43, 267 (1965).
[5]N. M. Chandheri and A. E. Scheigger, Can. J. Phys. 43, 1776 (1968).
[6]I. Oppenheim, K. E. Shuler, and G. H. Weiss, Advan. Mol. Relaxation Processes 1, 13 (1967–68).
[7]W. Ledermann and G. E. H. Reuter, Phil. Trans. Roy. Soc. A 246, 321 (1954).
[8]J. F. Elliot, "The Characteristic Roots of Certain Symmetric Matrices," Master's thesis (University of Tennesse, 1953).
[9]L. M. Milne-Thomson, *The Calculus of Finite Difference* (McMillan, London, 1933).
[10]L. L. Pennisi, *Elements of Complex Variables* (Holt-Rinehart & Winston, New York, 1963).
[11]A. Cohen, *An Elementary Treatise on Differential Equations* (Heath, Boston, 1933).
[12]E. W. Montroll and K. E. Shuler, J. Chem. Phys. 26, 454 (1957).

A note on energy bounds for boson matter*

David Brydges† and Paul Federbush

Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48104
(Received 15 September 1975)

Two new proofs are given of the Dyson and Lenard lower bound for the energy of matter with boson electrons. Another result is a new inequality for the two-point correlation function.

1. INTRODUCTION AND RESULTS

We consider a system of like positively charged particles described by a field ϕ and like negatively charged particles described by a field ψ. The charge is denoted by ϵ, any masses that appear satisfy either $m = \infty$ or $2m = 1$, and the total number of particles is $2N$. $\rho'(x, y)$ is defined by

$$\rho'(x, y) = \langle \bar{\phi}\phi(x)\, \bar{\phi}\phi(y) \rangle \tag{1.1}$$

$$= \langle :\bar{\phi}\phi(x)\, \bar{\phi}\phi(y): \rangle + \delta(x-y)\langle \bar{\phi}(x)\, \phi(x) \rangle \tag{1.2}$$

and thus is simply related to the usual two-point correlation function, the first term in (1.2). We note the following two theorems.

Theorem 1: If both sets of particles are bosons and not both masses are ∞, then there is a constant c such that the ground state energy, E_N, satisfies

$$E_N \geq -cN^{5/3}. \tag{1.3}$$

Theorem 2: There is a constant c such that if $f(r)$ is a right-continuous monotonically decreasing nonnegative function and

$$\int_{|x-y|<\alpha} \rho' = D, \tag{1.4}$$

then

$$\int_{|x-y|\geq\alpha} \rho' f(|x-y|) \leq c(D/\alpha^3)\left[\alpha^3 f(\alpha) + \int_\alpha^\infty f(r)\, r^2\, dr\right]. \tag{1.5}$$

Theorem 1 is a slightly strengthened form of a theorem of Dyson and Lenard.[1] (The original theorem requires both masses to be finite.[2] It is believed that the best exponent is 7/5 rather than 5/3.[6] Curiously both proofs we present of Theorem 1, very different, when pursued, are limited by a configuration of linear size $\sim 1/N^{1/3}$ and average spacing $\sim 1/N^{2/3}$.

Theorem 2 follows from the Packing inequality, Fact 2 of Ref. 4, by an easy argument. It is used in our second proof of Theorem 1. The proof of Theorem 2 is given in Sec. 4.

The body of the paper presents two proofs of Theorem 1 (one proof yields the original theorem) in Sec. 2 and Sec. 3. We feel the techniques of this paper are interesting and aesthetic in their own right—but our motivation is to use these techniques to generalize these theorems to a form where they will be useful in developing cluster expansions. Along this line, work is in progress to extend the results of Refs. 7 and 8.

2. FIRST PROOF OF THEOREM 1

Our first proof requires that both masses be finite:

$$H = -\sum \Delta_i + \tfrac{1}{2}\sum_{i\neq j} (\pm)\epsilon^2/|x_i - x_j|. \tag{2.1}$$

We use the electrostatic inequality, an easy inequality from Ref. 1, to obtain

$$H \geq -\sum \Delta_i - c\sum (1/R_i), \tag{2.2}$$

where R_i is the distance between the ith particle and its nearest neighbor. (For interest we repeat the remark from Ref. 1 that, to improve on the 5/3 power in Theorem 1, one would have to improve on this estimate.)

To prove Theorem 2 from (2.2) we note it is sufficient to show

$$-\Delta - c/R \geq -c_1 N^{2/3}, \tag{2.3}$$

where (2.3) describes the motion of one particle in the "field" of N fixed particles. R is the minimum distance to one of the fixed particles. [For notational reasons the $(2N-1)$ fixed particles each particle in (2.2) sees has been changed to N.] This inequality is implied by

$$\left\| \frac{1}{(-\Delta + c_1 N^{2/3})^{1/2}} \frac{c}{R} \frac{1}{(-\Delta + c_1 N^{2/3})^{1/2}} \right\| \leq 1. \tag{2.4}$$

We note that

$$c/R \leq c_1 N^{2/3}/2 + (c/R)\chi, \tag{2.5}$$

where χ is the characteristic function of the set where

$$c/R \geq \tfrac{1}{2} c_1 N^{2/3}. \tag{2.6}$$

It is enough now to show

$$\left\| \frac{1}{(-\Delta)^{1/2}} \left(\frac{c}{R}\chi\right) \frac{1}{(-\Delta)^{1/2}} \right\|_{\text{H.S.}} \leq \tfrac{1}{2}. \tag{2.7}$$

The subscript indicates the Hilbert–Schmidt norm. We use the Sobolev inequality[9]

$$\left| \int d^3x \int d^3y\, g(x) f(y)/|x-y|^2 \right| \leq c_2 \|f\|_{3/2} \|g\|_{3/2} \tag{2.8}$$

to convert (2.7) to

$$\|(c/R)\chi\|_{3/2} \leq c_3. \tag{2.9}$$

It is easy to see

$$\|(c/R)\chi\|_{3/2} \leq [4\pi N \int_0^{2c/c_1 N^{2/3}} (c/r)^{3/2} r^2\, dr]^{2/3} \tag{2.10}$$

$$\leq [(c^3/c_1^{3/2})c_4]^{2/3}.$$

so if c_1 is large enough, (2.3) holds and the collapse inequality of Ref. 1 has been proven.

3. SECOND PROOF OF THEOREM 1

In this proof, more technical in nature, we lean heavily on the methods of Ref. 4. We make the inessential simplification of considering a system of N positive and N negative particles, the positive particles in fixed classical positions. We also place the system in a unit box (with Neumann or periodic boundary conditions). We need show there is a c with

$$0 \leq cN^{5/3} + H. \tag{3.1}$$

We define

$$A_n = \int_{|x-y|<1/n} \rho' \tag{3.2}$$

and deduce from the Packing inequality[4] that there is a c_2 such that for any $c_1 < 1$ it is possible to find an n satisfying

$$c_1 n^{3/2} \leq A_n \leq c_2 c_1 n^{3/2}. \tag{3.3}$$

This is achieved by starting with $n = N^{2/3}$ and increasing n until (3.3) holds. A and n are now defined as values for which (3.3) holds.

With the notation of Ref. 4 we see

$$H_6 \geq c_3 n(A - N) e^{-1} \tag{3.4}$$

and

$$H_4 \geq -2c_3 n N. \tag{3.5}$$

Considering H_5, it is enough to show, to complete the proof, that

$$(-\Delta - V + c_3 n) \geq 0, \tag{3.6}$$

where

$$V = \epsilon^2 \int [\exp(-nr)/r]\, \overline{\phi}\phi. \tag{3.7}$$

Analagous to (2.4) we find, using an H.S. norm, that (3.6) is implied by

$$(1/n^{1/4}) \| V \|_2 \leq c_4 \tag{3.8}$$

or

$$(1/n^{1/4}) \left[\int \rho'(1/n) \exp(-nr) \right]^{1/2} \leq c_5. \tag{3.9}$$

Using Theorem 2, with $f(r) = \exp(-nr)$, (3.9) is implied by

$$(1/n^{3/4}) A^{1/2} [n^3 \int r^2 \exp(-nr)\, dr]^{1/2} \leq c_6 \tag{3.10}$$

or

$$A/n^{3/2} \leq c_7. \tag{3.11}$$

(3.11) holds if c_1 is small enough. The choices of constants c_i can be made independent of N and the positive particles' configuration.

The devoted reader may observe that if it were desired to prove Theorem 1 for any power larger than 5/3 (instead of exactly 5/3) this second proof could be much simplified, but our intended applications and generalizations require the 5/3 power.

4. PROOF OF THEOREM 2

We start from Fact 2 of Ref. 4 in the form:

Fact 2: There is a constant $c_2 > 0$ such that if $0 \leq R' \leq R$ then

$$\int_{|x-y|<R'} \rho' \geq c_2 (R'/R)^3 \int_{|x-y|<R} \rho'. \tag{4.1}$$

Let $f(r)$ be a right-continuous montonically decreasing nonnegative function. Define $S_\lambda(r)$ by

$$S_\lambda(r) = \begin{cases} 1, & r < \lambda, \\ 0, & r \geq \lambda. \end{cases} \tag{4.2}$$

For $r \geq \alpha$, $f(r)$ may be expressed in the form

$$f(r) = \int_\alpha^\infty d\sigma(\lambda)\, S_\lambda(r) \tag{4.3}$$

with $\sigma(\lambda)$ a positive measure. It is sufficient to prove (1.5) with $f(r) = S_\lambda(r)$, for then this form of (1.5) may be integrated with respect to the measure $\sigma(\lambda)$ to obtain Theorem 2. Substituting, with $\lambda \geq \alpha$, into (1.5), we get

$$\begin{aligned} \int_{\lambda > |x-y| \geq \alpha} \rho' &\leq c(D/\alpha^3)(\alpha^3 + \int_\alpha^\lambda r^2 dr) \\ &\leq cD[(\lambda^3 + 2\alpha^3)/3\alpha^3]. \end{aligned} \tag{4.4}$$

This is implied by

$$\int_{|x-y|<\lambda} \rho' \leq c \cdot \int_{|x-y|<\alpha} \rho' \cdot (\lambda^3/3\alpha^3). \tag{4.5}$$

If $c/3 \geq 1/c_2$, then (4.5) holds by Fact 2, yielding (4.4) and the theorem.

ACKNOWLEDGMENT

We would like to thank Professor B.A. Taylor for the idea of the proof in Sec. 4.

*This work was supported in part by NSF Grant GP-17523.

†Michigan Junior Fellow.

[1] F. J. Dyson and A. Lenard, J. Math. Phys. 8, 423 (1967).
[2] Dyson and Lenard in Ref. 3 obtained a proof of Theorem 1 in conjunction with their proof of stability of matter.
[3] F. J. Dyson and A. Lenard, J. Math. Phys. 9, 698 (1968).
[4] P. Federbush, J. Math. Phys. 16, 347 (1975).
[5] P. Federbush, J. Math. Phys. 16, 706 (1975).
[6] F. J. Dyson, J. Math. Phys. 8, 1538 (1967).
[7] P. Federbush, J. Math. Phys. 17, 200 (1976).
[8] P. Federbush, J. Math. Phys. 17, 204 (1976).
[9] B. Simon, *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms* (Princeton U. P., Princeton, N. J., 1971).

Shock waves in relativistic magnetohydrodynamics under general assumptions

André Lichnerowicz

Collège de France, Paris, France
(Received 15 December 1975)

We present a rigorous theory of magnetohydrodynamical shock waves in the framework of a given curved space–time, under general assumptions corresponding both to a plasma and to a condensed medium. The results can be of use in astrophysics. We prove the timelike character of the wavefronts, the main thermodynamic inequalities, the relative location of the speeds of the shock waves with respect to the magnetosonic and Alfvén speeds, and show some existence and uniqueness theorems. In particular, we show that there can exist initial states giving slow shocks, but no weak shocks.

INTRODUCTION

Hydrodynamic and magnetohydrodynamic shock waves play an important role in different fields of theoretical astrophysics. It is reasonable to study these waves in the relativistic frame given by a curved space–time. There it is necessary for various astronomical applications, and in addition, relativistic magnetohydrodynamics is simpler than the classical theory, from several viewpoints: relativistic dynamics present a natural harmony with Maxwell equations.

Some years ago, I gave[1] a first rigorous study of relativistic magnetohydrodynamic shock waves. The reader can find a more detailed treatment in Ref. 2. This treatment was based on the following assumptions (see notation in Sec. 1):

(1) $\tau'_p < 0$, that is sonic speed $< c$,

(2) $\tau''_{p^2} > 0$, convexity condition connected with the stability of the hydrodynamic shock waves,

(3) $\tau'_S > 0$.

Israël[3] and Lucquiaud[4] have proved these assumptions, by methods of statistical mechanics, for a Boltzmann type of gas. But condensate media appear also in relativistic astrophysics and it is interesting, according to a suggestion of Thorne,[5] to study also the case of media for which τ'_S is less than 0.

The purpose of this paper is to make precise and extend my previous analysis, for relativistic magnetohydrodynamic shock waves, under general assumptions corresponding both to a plasma and to a condensed medium. General questions are the following:

(1) location, with respect to c, of the wavefront speeds;

(2) location of these speeds with respect to the magnetosonic and Alfvên speeds;

(3) thermodynamic inequalities corresponding to a shock;

(4) existence and uniqueness properties for a nontrivial solution of the shock equations.

Usual arguments[6] postulate the connected character of the Hugoniot curve (Taub adiabat[7] for hydrodynamic shock waves, in the terminology of Thorne) and this property is by no means evident *a priori*, in the magnetohydrodynamic case.

Our rigorous study gives complete answers for the first three above questions and partial answers for the fourth. The paper is largely self-contained; calculations which are long and straightforward, are often omitted.

1. THE FLUID

(a) Let $(V_4, \mathbf{g})$ be a given space–time, where $\mathbf{g}$ is a Lorentzian metric of class C^1, with the signature $(+---)$. A perfect fluid is described by an energy tensor,

$$\hat{T}_{\alpha\beta} = (\rho + p)\, u_\alpha u_\beta - p g_{\alpha\beta}, \quad \alpha, \beta, \cdots = 0,1,2,3,$$

where ρ is the proper energy density, p the pressure, and u_α the unit 4-velocity, oriented towards the future. We put

$$\rho = c^2 r(1 + \epsilon/c^2), \quad r > 0,$$

where r is the matter density of the fluid, and ϵ its specific internal energy. We introduce the specific enthalpy

$$i = \epsilon + pV, \quad V = 1/r,$$

and the so-called *index of the fluid*

$$f = 1 + i/c^2, \quad f > 1.$$

The energy tensor can be written

$$\hat{T}_{\alpha\beta} = c^2 r f u_\alpha u_\beta - p g_{\alpha\beta}. \tag{1.1}$$

The proper temperature Θ of the fluid and its specific entropy S satisfy, as in classical hydrodynamics, the differential relation

$$\Theta\, dS = d\epsilon + p\, dV = c^2\, df - V\, dp, \quad \Theta > 0.$$

Therefore

$$c^2\, df = V\, dp + \Theta\, dS. \tag{1.2}$$

In relativity, the thermodynamical variable $\tau = fV$ (dynamical volume) plays an important role and can be substituted for the specific volume V. We consider τ as a given function $\tau = \tau(p, S)$ defining an *equation of state* of the fluid.

(b) Let Σ be a regular hypersurface of V_4, $\varphi = 0$ its

local equation; we set $l_\alpha = \partial_\alpha \varphi$. The speed v^Σ of Σ with respect to the fluid, that is with respect to the time direction $\mathbf{u}$ is given by

$$(v^\Sigma)^2/c^2 = (u^\alpha l_\alpha)^2/\{(u^\alpha l_\alpha)^2 - l^\alpha l_\alpha\}. \tag{1.3}$$

v^Σ is less than c if an only if $l^\alpha l_\alpha < 0$ (Σ is timelike). If v is the sonic speed of the fluid, $\gamma = c^2/v^2$ is given by $c^2\tau'_p = -V^2(\gamma - 1)$; *$v < c$ is equivalent to $\tau'_p < 0$*. We assume in all the following,

Assumption A(1): We have *$\tau'_p < 0$ and $\tau''_{p2} > 0$*.

A(2): *τ'_S is $\neq 0$ and there exists a function $S = S(p, \tau)$ whose unique inverse is $\tau = \tau(p,S)$.*

2. THE MAGNETOHYDRODYNAMICS EQUATIONS

(a) We consider here the case of the *perfect* Magnetohydrodynamics (infinite conductivity and constant magnetic permeability μ). The electromagnetic field is reduced to the magnetic field $\mathbf{h}$ with respect to the fluid, which satisfies

$$u^\rho h_\rho = 0. \tag{2.1}$$

We obtain then for the system (fluid + field) the total energy tensor

$$T_{\alpha\beta} = (c^2 rf + \mu |h|^2) u_\alpha u_\beta - q g_{\alpha\beta} - \mu h_\alpha h_\beta, \tag{2.2}$$

where $|h|^2 = -h^\rho h_\rho$ is strictly positive and where $q = p + \frac{1}{2}\mu |h|^2$.

The differential system of relativistic magnetohydrodynamics (MHD) is given by the following considerations: We suppose, first, that the matter density r (which corresponds to the specific number of particles) is conserved; if ∇ is the operator of covariant differentiation,

$$\nabla_\alpha (r u^\alpha) = 0. \tag{2.3}$$

Maxwell equations give here only

$$\nabla_\alpha (u^\alpha h^\beta - u^\beta h^\alpha) = 0 \tag{2.4}$$

and the equations of relativistic dynamics are

$$\nabla_\alpha T^{\alpha\beta} = 0. \tag{2.5}$$

The system (2.3), (2.4) and (2.5) implies the equation of isentropic flow $u^\alpha \partial_\alpha S = 0$.

(b) The characteristic manifolds of our system are the tangential waves ($u^\alpha l_\alpha = 0$), the magnetosonic waves, solutions of

$$P(l) \equiv c^2 rf(\gamma - 1)(u^\alpha l_\alpha)^4 + (c^2 rf + \mu |h|^2 \gamma) \times (u^\alpha l_\alpha)^2 l^\beta l_\beta - \mu (h^\alpha l_\alpha)^2 l^\beta l_\beta = 0, \tag{2.6}$$

and the Alfvén waves, solutions of

$$D(l) \equiv (c^2 rf + \mu |h|^2)(u^\alpha l_\alpha)^2 - \mu (h^\alpha l_\alpha)^2 = 0. \tag{2.7}$$

Under the assumption $\tau'_p < 0$ (or $\gamma > 1$), (2.6) and (2.7) define three speeds v^{MS}, v^{MF}, and v^A satisfying the inequalities

$$v^{MS} < v^A < v^{MF} < c, \quad v^{MS} < v < v^{MF}$$

(v^{MS} magnetosonic slow speed, v^{MF} magnetosonic fast speed).

3. GENERAL SHOCK EQUATIONS

(a) A *state* Y of the system (fluid + field) at a point x of V_4 is defined by the values of p, S, $\mathbf{u}$, and $\mathbf{h}$ (eight parameters). A magnetohydrodynamic shock wave is a solution of the main system in a weak sense such that there exists a hypersurface Σ (the wavefront) satisfying the following conditions:

(1) On both sides of Σ, the states are continuous functions of x and the main system is satisfied in the usual sense;

(2) The variables defining the states are regularly discontinuous across Σ and, in the neighborhood of Σ the main system is satisfied in the sense of distributions.

I will show that, under the assumption $\tau'_p < 0$, Σ is necessarily timelike. If we decompose $\mathbf{u}$ and $\mathbf{h}$ into a tangential component and a normal component with respect to Σ ($\varphi = 0$; $l = d\varphi$) we obtain

$$u^\beta = v^\beta + (u^\alpha l_\alpha)(l^\alpha l_\alpha)^{-1} l^\beta,$$

$$h^\beta = k^\beta + (h^\alpha l_\alpha)(l^\alpha l_\alpha)^{-1} l^\beta,$$

with $v^\beta l_\beta = k^\beta l_\beta = 0$. We denote as Y_0 the state at $x \in \Sigma$ before the shock, as Y_1 the state after the shock. A bracket corresponds to the discontinuity of a quantity across Σ. From the main system, we obtain by a classical argument, the general shock equations

$$[r u^\alpha] l_\alpha = 0, \quad [h^\alpha u^\beta - u^\alpha h^\beta] l_\alpha = 0, \quad [T^{\alpha\beta}] l_\alpha = 0. \tag{3.1}$$

We add to (3.1) the assumption

$$[S] \geq 0, \tag{3.2}$$

which is the formulation of the so-called Clausius–Duhem inequality. Equation (3.1) expresses the invariance of the scalar

$$a(Y) = r u^\alpha l_\alpha,$$

the invariance of the tangent vector to Σ,

$$V^\beta(Y) = (h^\alpha l_\alpha) u^\beta - (a/r) h^\beta,$$

and the invariance of the vector

$$W^\beta(Y) = \{c^2\tau + \mu |h|^2 r^{-2}\} a r u^\beta - q l^\beta - \mu (h^\alpha l_\alpha) h^\beta.$$

The case $a = 0$ (tangential shock) is trivial and we assume $a \neq 0$ in the following.

(b) If we decompose W^β into a tangential and a normal component, we obtain a scalar and a tangent vector which are invariant under the shock. We deduce, from all these invariants, five scalar invariants of the shock concerning the two thermodynamical variables and the three scalars $|h|^2$, $u^\alpha l_\alpha$, $h^\alpha l_\alpha$ (see the Appendix),

$$r u^\alpha l_\alpha = a, \tag{3.3}$$

$$f h^\alpha l_\alpha = b, \tag{3.4}$$

$$(h^\alpha l_\alpha)^2/a^2 - |h|^2/r^2 = H, \tag{3.5}$$

$$\chi \alpha^2 = L, \tag{3.6}$$

$$\bar{q} - (c^2 a^2 / l^\alpha l_\alpha)\tau = e, \tag{3.7}$$

where

$$\alpha = c^2\tau - \mu H, \quad \chi = |h|^2 + (a^2/l^\alpha l_\alpha)H, \tag{3.8}$$

$$\bar{q} = p + \tfrac{1}{2}\mu\chi.$$

$\alpha = 0$ expresses that Σ is an Alfvên wavefront for the state Y. We have the following lemma (see Ref. 2, p. 151).

Lemma 1: (i) We have $(l^\alpha l_\alpha)\chi \leq 0$;

(ii) If $l^\alpha l_\alpha \neq 0$, we have $\chi = 0$ if and only if l is in the 2-plane $(\mathbf{u}, \mathbf{h})$;

(iii) If $l^\alpha l_\alpha \geq 0$, we have $H \leq 0$ and thus $\alpha > 0$.

(c) Now the tangential components of the velocity and of the magnetic field satisfy

$$(h_1^\alpha l_\alpha)v_1^\beta - (\mu_1^\alpha l_\alpha)k_1^\beta = (h_0^\alpha l_\alpha) - (u_0^\alpha l_\alpha)k_0^\beta, \tag{3.9}$$

$$(c^2\gamma_1 f_1 + \mu|h_1|^2)(u_1^\alpha l_\alpha)v_1^\beta - \mu(h_1^\alpha l_\alpha)k_1^\beta = (c^2\gamma_0 f_0 + \mu|h_0|^2)(u_0^\alpha l_\alpha)v_0^\beta - \mu(h_0^\alpha l_\alpha)k_0^\beta. \tag{3.10}$$

The determinant of the left members of (3.9) and (3.10) with respect to v_1^β, k_1^β is $D_1(l) = a^2\alpha_1$. If $\alpha_1 \neq 0$, (3.9) and (3.10) give v_1^β, k_1^β in terms of the initial state and of quantities which satisfy the five scalar invariance relations (3.3)–(3.7).

Suppose $\alpha_1 = 0$: Σ is then a timelike Alfvên wavefront after the shock. It follows from (3.6) that either $\alpha_0 = 0$ or $\chi_0 = 0$.

(d) An *Alfvên shock* is a shock such that $\alpha_1 = \alpha_0 = 0$. It is easy to prove that, *under the assumption* $\tau_p' < 0$, $\alpha_1 = \alpha_0$ implies that the two thermodynamical variables and the three scalars $|h|^2$, $u^\alpha l_\alpha$, $h^\alpha l_\alpha$ are invariant under the shock. If $\alpha_1 = \alpha_0 \neq 0$, the shock is null. If $\alpha_1 = \alpha_0 = 0$ (Alfvén shock), the direction of the tangential magnetic field after the shock is undetermined, but determines the direction of the tangential velocity.

Moreover, it is possible to prove the following result[2]: A shock wave such that $\alpha_0\alpha_1 = 0$ is compatible with usual Alfvên waves if and only if it is an Alfvên shock ($\alpha_1 = \alpha_0 = 0$). The case $\alpha_1 = 0$, $\chi_0 = 0$, $\alpha_0 \neq 0$ and, symmetrically, $\alpha_0 = 0$, $\chi_1 = 0$, $\alpha_1 \neq 0$ (singular shocks) are *physically forbidden* as unstable.

In the following part, *we consider shocks which are not Alfvên shocks.*

4. HUGONIOT FUNCTION AND MAIN FORMULAS

(a) It is possible to deduce[2] from the five invariants (3.3)–(3.7) the following relation concerning the states Y_0 and Y_1 corresponding to a shock:

$$c^2(f_1^2 - f_0^2) - (\tau_0 + \tau_1)(p_1 - p_0) + (\tau_1 - \tau_0)\cdot\tfrac{1}{2}\mu(\chi_0 + \chi_1 - 2\chi_0\alpha_0/\alpha_1) = 0. \tag{4.1}$$

Equation (4.1) will be called here the *Hugoniot relation* and can be substituted for (3.4), if we assume that $(h_1^\alpha l_\alpha)(h_0^\alpha l_\alpha) \geq 0$.

(b) An initial state Y_0 of the system (fluid + field) being given at $x \in \Sigma$, we consider in the following the set $\mathcal{E}$ of all possible states satisfying the conditions

$$H(Y) = H(Y_0) = H, \quad L(Y) = L(Y_0) = L, \tag{4.2}$$

and

$$(h^\alpha l_\alpha)(h_0^\alpha l_\alpha) \geq 0 \tag{4.3}$$

so that, for $\tau \neq \mu H/c^2$, we have

$$\chi = L/(c^2\tau - \mu H)^2 = \chi_0\alpha_0^2/(c^2\tau - \mu H)^2.$$

We have, between the variables τ, S, and $\bar{q}$ the functional relation

$$\bar{q} = p(\tau, S) + \mu L/2(c^2\tau - \mu H)^2, \tag{4.4}$$

where $p(\tau, S) \geq 0$ is defined by inversion of the equation of state.

We denote as Π the half-plane $(\tau, \bar{q})$ defined by $\tau \geq 0$. Under the conditions (4.2) and (4.3), a thermodynamical state of the fluid, defined for example by (τ, p), determines a point $Z = (\tau, \bar{q})$ of Π such that

$$\bar{q} \geq \mu L/2(c^2\tau - \mu H)^2. \tag{4.5}$$

Equation (4.5) defines in Π one or two connected and *convex* regions $\mathcal{R}$, according to $H \leq 0$ or $H > 0$. Conversely a point $Z = (\tau, \bar{q})$ in $\mathcal{R}$ defines a thermodynamical state of the fluid (that is τ, p, S) and values for α and χ.

We denote as σ the natural map from $\mathcal{E}$ into Π. We are led to introduce the Hugoniot function $\mathcal{H}(Z_0, Z)$ for relativistic MHD, considered as a function of $Z \in \mathcal{R}$, for a given initial point $Z_0 \in \mathcal{R}$,

$$\mathcal{H}(Z_0, Z) = c^2(f^2 - f_0^2) - (\tau + \tau_0)(p - p_0) + (\tau - \tau_0)\cdot\tfrac{1}{2}\mu(\chi + \chi_0 - 2\chi_0\alpha_0/\alpha). \tag{4.6}$$

It is clear that $\mathcal{H}(Z_0, Z_0) = 0$ and that (4.1) can be written $\mathcal{H}(Z_0, Z_1) = 0$ for $\sigma(Y_1) = Z_1$. A detailed study of the behavior of the Hugoniot function will give a great part of the sought after results.

We obtain by differentiation of (4.6), from (4.2) and (1.2),[2]

$$d\mathcal{H} = 2f\Theta\, dS + (\tau - \tau_0)\, d\bar{q} - (\bar{q} - \bar{q}_0)\, d\tau = 2f\Theta\, dS + (\tau - \tau_0)^2\, dm, \tag{4.7}$$

where m is the slope of the straight line (Z_0, Z) of $\mathcal{R}$.

(c) Study the differential of S along the part in $\mathcal{R}$ of a straight line Δ of Π of slope m. We have

$$\tau_S' dS = d\tau - \tau_p' dp$$

and

$$dp + \tfrac{1}{2}\mu\, d\chi = d\bar{q} = m\, d\tau$$

so that

$$\tau_S' dS = (1 - m\tau_p')\, d\tau + \tfrac{1}{2}\mu\tau_p'\, d\chi,$$

but we obtain by differentiation of $\chi\alpha^2 = \text{const}$,

$$\alpha\, d\chi = -2c^2\chi\, d\tau.$$

It follows that

$$\tau_S'\, dS = \tau_p'(\tau_p'^{-1} - c^2\mu\chi\alpha^{-1} - m)\, d\tau. \tag{4.8}$$

Consider, for a point Z of $\mathcal{R}$, the isentropic curve $\mathcal{S}$ of Z corresponding to the value S of the specific entropy

and defined by (4.4). The slope of S at Z is given by

$$\left(\frac{d\bar{q}}{d\tau}\right)_S = \tau_p'^{-1} - c^2\mu\chi\alpha^{-1}$$

We set, for S and for a straightline Δ of slope m issued from Z

$$Q(m) = (d\bar{q}/d\tau)_S - m = \tau_p'^{-1} - c^2\mu\chi\alpha^{-1} - m. \tag{4.9}$$

Formula (4.8) gives

$$\tau_S' dS = \tau_p' Q(m)\, d\tau. \tag{4.10}$$

(d) For a state Y, $P(l)$ can be written

$$P(l) = c^2\gamma f(\gamma-1)(u^\alpha l_\alpha)^4 + \{c^2\gamma f + \mu|h|^2(\gamma-1)\} \times (u^\alpha l_\alpha)^2 l^\beta l_\beta - \mu a^2 H l^\beta l_\beta. \tag{4.11}$$

We obtain, for $l^\alpha l_\alpha \neq 0$

$$P(l) = c^2\gamma f(\gamma-1)(u^\alpha l_\alpha)^4 + \{c^2\gamma f + \mu(\chi - (a^2/l^\alpha l_\alpha)H)(\gamma-1)\} \times (u^\alpha l_\alpha)^2 l^\beta l_\beta - \mu a^2 H l^\beta l_\beta \tag{4.12}$$

and for $l^\alpha l_\alpha = 0$,

$$P(l) = c^2\gamma f(\gamma-1)(u^\alpha l_\alpha)^4. \tag{4.13}$$

We have seen that a point Z of $\mathcal{R}$ defines values for the thermodynamical variables of the fluid and for α and χ. Now, we consider also that Z defines a value of $(u^\alpha l_\alpha)$ by the condition

$$\gamma u^\alpha l_\alpha = \gamma_0 u_0^\alpha l_\alpha = a. \tag{4.14}$$

If such is the case, the right-hand side of (4.12) [respectively (4.13)] defines, for each point $Z \in \mathcal{R}$, a number denoted also as $P(l)(Z)$.

Consider, according to (3.7), the points Z of the straight line Δ_a of slope $m_a = c^2a^2/l^\alpha l_\alpha$ issued from Z_0. It is easy to verify that, for these points and for $l^\alpha l_\alpha \neq 0$, $P(l)$ has for its value

$$P(l) = a^2\tau_p' l^\alpha l_\alpha \circ \alpha Q(m_a). \tag{4.15}$$

It follows from (4.10) that along Δ_a

$$c^2a^4\tau_S'\alpha\, dS = P(l)\, d\bar{q}. \tag{4.16}$$

This formula is also valid for $l^\alpha l_\alpha = 0$.

5. ORIENTATION OF THE SHOCK WAVEFRONTS

(a) Consider at $x \in \Sigma$ a nonvanishing shock $Y_0 \rightarrow Y_1$ which is not an Alfvên shock. We will study, under the assumption $\tau_p' < 0$, the orientation of the wavefront Σ. We set $Z_0 = \sigma(Y_0)$ and $Z_1 = \sigma(Y_1)$.

Consider the points Z of the segment (Z_0, Z_1) of slope m_a and the corresponding values of $u^\alpha l_\alpha$, according to (4.14). *Suppose* $l^\alpha l_\alpha \geq 0$. It follows (see the lemma in Sec. 3) that $H \leq 0$, $\alpha > 0$ and the segment (Z_0, Z_1) belongs to one determined region $\mathcal{R}$. *If* $l^\alpha l_\alpha > 0$, we have $\chi \leq 0$. For the points Z_0 and Z_1, we have

$$\chi_0 - (a^2/l^\alpha l_\alpha)H = |h_0|^2 \geq 0,$$
$$\chi_1 - (a^2/l^\alpha l_\alpha)H = |h_1|^2 \geq 0. \tag{5.1}$$

Along the segment (Z_0, Z_1), we have for example $\tau_0 \leq \tau$ (respectively $\tau_1 \leq \tau$), that is $\alpha_0^2 \leq \alpha^2$ (respectively $\alpha_1^2 \leq \alpha^2$) and $|\chi| \leq |\chi_0|$ (respectively $|\chi| \leq |\chi_1|$). We deduce from (5.1)

$$\chi - (a^2/l^\alpha l_\alpha)H \geq 0,$$

and it follows from (4.12) that $P(l) > 0$. *If* $l^\alpha l_\alpha = 0$, we have also $P(l) > 0$, according to (4.13). We see that, if $l^\alpha l_\alpha \geq 0$, it follows from (4.16) that $dS/d\bar{q} \neq 0$ along (Z_0, Z_1). Along this segment, we have according to (4.7),

$$d\mathcal{H} = 2f\Theta\, dS. \tag{5.2}$$

But $\mathcal{H}(Z_0, Z_0) = \mathcal{H}(Z_0, Z_1) = 0$ and the function $\mathcal{H}(Z_0, Z)$ is stationary at at least one point of the segment (Z_0, Z_1) and the same is true for S according to (5.2). We obtain a contradiction. Therefore $l^\alpha l_\alpha < 0$.

Theorem: Under the assumption $\tau_p' < 0$, each magnetohydrodynamic wavefront Σ is necessarily timelike. If v_0^{E} and v_1^{E} are the speeds of Σ with respect to the fluid before and after the shock, we have $v_0^{\mathrm{E}} < c$ and $v_1^{\mathrm{E}} < c$.

(b) We know that $\chi = \Psi^2$ is positive; (3.6) can be written $\Psi_1\alpha_1 = \Psi_0\alpha_0$. We can substitute for the second condition (4.2),

$$\Psi\alpha = \Psi_0\alpha_0. \tag{5.3}$$

We have then

$$\chi + \chi_0 - 2\chi_0\alpha_0/\alpha = \Psi^2 + \Psi_0^2 - 2\Psi\Psi_0 = (\Psi - \Psi_0)^2.$$

The Hugoniot function can be written

$$\mathcal{H}(Z_0, Z) = c^2(f^2 - f_0^2) - (\tau + \tau_0)(p - p_0) + (\tau - \tau_0) \circ \tfrac{1}{2}\mu(\Psi - \Psi_0)^2. \tag{5.4}$$

6. THERMODYNAMIC INEQUALITIES

According to (1.2), in terms of the variables p and S, we have

$$c^2 f_p' = V > 0, \quad c^2 f_S' = \Theta > 0. \tag{6.1}$$

It follows that

$$c^2(f^2)_p' = 2\tau. \tag{6.2}$$

Consider a shock $Y_0 \rightarrow Y_1$; $Z_0 = \sigma(Y_0)$ and $Z_1 = \sigma(Y_1)$ are connected by the Hugoniot relation which is symmetrical in 0 and 1. *We assume now Assumption A and that* $S_1 \geq S_0$, *according to* (3.2).

(a) *Suppose* $S_1 = S_0$. If $p_1 \neq p_0$, we can suppose for example $p_1 > p_0$. We then have $\tau_1 < \tau_0$, since $\tau_p' < 0$. We deduce from (6.2)

$$c^2\{f^2(p_1, S_0) - f^2(p_0, S_0)\} = 2\int_{p_0}^{p_1} \tau(p, S_0)\, dp.$$

It follows from the convexity condition $\tau_{p^2}'' > 0$ that

$$c^2(f_1^2 - f_0^2) < \{\tau(p_0, S_0) + \tau(p_1, S_0)\}(p_1 - p_0) = (\tau_1 + \tau_0)(p_1 - p_0).$$

We deduce from the Hugoniot relation $\tau_1 > \tau_0$, namely a contradiction. We thus have $p_1 = p_0$ and our shock is null or is an Alfvên shock.

(b) *Consider a shock with* $S_1 > S_0$ *and suppose* $p_1 \geq p_0$. We have

$$c^2\{f^2(p_1, S_1)\} - f^2(p_0, S_1)\} = 2\int_{p_0}^{p_1} \tau(p, S_1)\, dp \geq 2\tau_1(p_1 - p_0)$$

and thus

$$c^2(f_1^2 - f_0^2) - 2\tau_1(p_1 - p_0) > 0.$$

It follows from the Hugoniot relation that

$$(\tau_1-\tau_0)\{(p_1-p_0)+\tfrac{1}{2}\mu(\Psi_1-\Psi_0)^2\}<0.$$

Let $\tau_1<\tau_0$. We see that $p_1\geqslant p_0$ implies $\tau_1<\tau_0$.

(c) *Suppose* $\tau'_S<0$. I shall show that, in this case, $p_1<p_0$ also implies $\tau_1<\tau_0$. In fact

$$c^2\{f^2(p_0,S_1)-f^2(p_1,S_1)\}=2\int_{p_1}^{p_0}\tau(p,S_1)\,dp.$$

It follows from the convexity condition that

$$c^2(f_0^2-f_1^2)<(\tau(p_0,S_1)+\tau(p_1,S_1))(p_0-p_1)<(\tau_0+\tau_1)(p_0-p_1).$$

That is,

$$c^2(f_1^2-f_0^2)-(\tau_0+\tau_1)(p_1-p_0)>0.$$

We deduce from the Hugoniot relation that $\tau_1<\tau_0$.

(d) *Suppose* $\tau'_S>0$. We have in this case $p_1>p_0$. In fact if $p_1\leqslant p_0$, we deduce from

$$c^2\{f^2(p_0,S_0)-f^2(p_1,S_0)\}=2\int_{p_1}^{p_0}\tau(p,S_0)\,dp,$$

by an argument similar to the argument of c, that $\tau_1<\tau_0$. We obtain a contradiction with $\tau'_p<0$, $\tau'_S>0$.

We have proved the following theorem.

Theorem: For a nonvanishing shock which is not an Alfvên shock, we have under Assumption (A)

$$S_1>S_0,\quad \tau_1<\tau_0, \tag{6.3}$$

and the following:

Proposition: If, moreover $\tau'_S>0$, we have

$$p_1>p_0,\quad f_1>f_0,\quad V_1<V_0. \tag{6.4}$$

It follows from the theorem that we have $\alpha_1<\alpha_0$. It is possible to prove[2] that a shock wave, which is not an Alfvên shock wave, is compatible with usual Alfvên waves *if and only if* $\alpha_1\alpha_0>0$. We obtain two types of shocks: the *slow shocks* for which $\alpha_1<\alpha_0<0$ and the *fast shocks* for which $0<\alpha_1<\alpha_0$.

7. HUGONIOT CURVE AND SPEEDS OF THE SHOCK WAVES

(a) Straightforward calculus shows the following results[2]:

(α) We have for each isentropic curve $\mathcal{S}$ of $\mathcal{R}$,

$$(d^2\bar{q}/d\tau^2)_{\mathcal{S}}=-\tau'^{-3}_p M,$$

where

$$M=\tau''_p{}^2-3c^4\mu\psi^2\alpha^{-2}\tau'^3_p>0.$$

Therefore *each isentropic curve is strictly concave.*

(β) Let Δ be a straight line in $\mathcal{R}$. We have the following lemma.

Lemma: Under Assumption (A), we have at each point Z_s of Δ where S is stationary,

$$\{\tau'_S(d^2S/d\tau^2)_\Delta\}(Z_s)=-\{\tau'^{-2}_p M\}(Z_s)<0.$$

(b) An initial state Y_0 being given, we consider, according to Sec. 6, the set $\mathcal{J}$ of the states Y satisfying

$$H(Y)=H(Y_0)=H,\quad \Psi\alpha=\Psi_0\alpha_0,\quad (\alpha\alpha_0>0)$$
$$(h^\alpha l_\alpha)(h_0^\alpha l_\alpha)\geqslant 0.$$

If $H>0$, the straight line $\tau=\mu H/c^2$ is forbidden.

In Π, we consider only the region defined by $\tau\leqslant\tau_0$ in the region $\mathcal{R}$. We call *Hugoniot curve* $\mathcal{H}$ the set of the points Z in this region satisfying $\mathcal{H}(Z_0,Z)=0$.

Let $Z_1\neq Z_0$ be an arbitrary point in $\mathcal{H}$ and Δ the ray (Z_0,Z_1) of slope m; $\mathcal{H}(Z_0,Z)$ is stationary at one point Z_s at least of the segment (Z_0,Z_1); S is stationary at Z_s according to (4.7) and it follows from the lemma of (a) that Z_s is unique and corresponds, according to the sign of τ'_S, to a strict maximum or a strict minimum of S on Δ.

Therefore, we have at Z_1

$$\{\tau'_S(dS/d\tau)_\Delta\}(Z_1)>0 \tag{7.1}$$

and according to (4.10)

$$Q(m,Z_1)<0.$$

In particular $Q(m,Z_1)$ *is not vanishing onto* $\mathcal{H}$. On the other hand

$$\{\tau'_S(dS/d\tau)_\Delta\}(Z_0)<0, \tag{7.2}$$

that is,

$$Q(m,Z_0)>0.$$

If m_0 is the slope at Z_0 of the isentropic curve $\mathcal{S}_0$ of this point, we have

$$m<m_0. \tag{7.3}$$

(c) Suppose that there exists $Z_1\in\mathcal{H}$ such that $(dS/d\tau)_{\mathcal{H}}(Z_1)=0$. The isentropic curve of Z_1 is tangent at Z_1 to $\mathcal{H}$. It follows from (4.7) that the ray $\Delta=(Z_0,Z_1)$ of slope m is also tangent at Z_1 to these curves and $Q(m,Z_1)=0$; but that is impossible. Therefore S *is a strictly monotonic function of* τ *on each connected component of* $\mathcal{H}$.

We have

$$(dS/dp)_{\mathcal{H}}=(dS/d\tau)_{\mathcal{H}}\{\tau'_p+\tau'_S(dS/dp)_{\mathcal{H}}\}$$

and $(dS/dp)_{\mathcal{H}}$ cannot be null; S *is also a strictly monotonic function of* p *on each conneted component of* $\mathcal{H}$.

(d) Let $Y_0\rightharpoonup Y_1$ be a shock, Δ_a the ray (Z_0,Z_1) of slope m_a. We have, according to b, (7.1) and (7.2) and it follows from (4.10) and (4.15) that

$$\alpha_0P(l)_0>0,\quad \alpha_1P(l)_1<0. \tag{7.4}$$

Geometrically, (7.4) expresses that

$$m_0=(d\bar{q}/d\tau)_{\mathcal{S}}\ (Z_0)>m_a,$$
$$m_1=(d\bar{q}/d\tau)_{\mathcal{S}}\ (Z_1)<m_a. \tag{7.5}$$

Equation (7.4) can be interpreted in terms of speeds of the shock wave and of the magnetosonic and Alfvên speeds, before and after the shock. We obtain the following theorem.

Theorem: Under Assumption (A), the speeds v_0^Σ and v_1^Σ of a shock wavefront Σ satisfy the inequalities:

(1) for a fast shock:

$$v_0^{MS}<v_0^A<v_0^{MF}<v_0^\Sigma,\quad v_1^{MS}<v_1^A<v_1^\Sigma<v_1^{MF},$$

(2) for a slow shock:

$$v_0^{MS} < v_0^{\Sigma} < v_0^{A} < v_0^{MF}, \quad v_1^{\Sigma} < v_1^{MS} < v_1^{A} < v_1^{MF}.$$

8. ISENTROPIC CURVES AND WEAK SHOCKS

(a) We consider, in the region $\tau \leq \tau_0$ of $\mathcal{R}$ ($\alpha\alpha_0 > 0$), the isentropic curve $\mathcal{S}_0$ corresponding to $S = S_0$. We have along $\mathcal{S}_0$

$$(d\mathcal{H}/d\tau)_{\mathcal{S}_0} = (\tau - \tau_0)(d\bar{q}/d\tau)_{\mathcal{S}_0} - (\bar{q} - \bar{q}_0)$$

and

$$(d^2\mathcal{H}/d\tau^2)_{\mathcal{S}_0} = (\tau - \tau_0)(d^2\bar{q}/d\tau^2)_{\mathcal{S}_0} = -(\tau - \tau_0)\tau_p'^{-3}M < 0.$$

We see that we have on the considered arc of $\mathcal{S}_0$

$$(d\mathcal{H}/d\tau)_{\mathcal{S}_0} > 0, \quad \mathcal{H}(Z_0, Z) < 0 \text{ for } Z \in \mathcal{S}_0. \tag{8.1}$$

The intersection of $\mathcal{S}_0$ and $\mathcal{H}$ is empty.

(b) We shall prove in the sequel that $\mathcal{H}$ is connected. Study $\mathcal{H}$, for the moment, in the neighborhood of Z_0. A straightforward calculus gives

$$(dS/d\tau)_{\mathcal{H}}(Z_0) = 0, \quad (d^2S/d\tau^2)_{\mathcal{H}}(Z_0) = 0.$$

Thus $\mathcal{H}$ and $\mathcal{S}_0$ have a second order contact at Z_0 and

$$(d^2\bar{q}/d\tau^2)_{\mathcal{H}}(Z_0) = (d^2\bar{q}/d\tau^2)_{\mathcal{S}_0}(Z_0) = -(\tau_p'^{-3}M)(Z_0).$$

Moreover

$$(d^3S/d\tau^3)_{\mathcal{H}}(Z_0) = \{(2f\Theta\tau_p'^3)^{-1}M\}(Z_0).$$

It follows that in the neighborhood of Z_0, we have along $\mathcal{H}$,

$$\begin{aligned} S - S_0 &= \{(12f\Theta\tau_p'^3)^{-1}M\}(Z_0)(\tau - \tau_0)^3 \\ &= \{(12f\Theta)^{-1}M\}(Z_0)(p - p_0)^3. \end{aligned} \tag{8.2}$$

We see that, in a weak shock, the entropy increase $(S_1 - S_0)$ is of third order with respect to the shock strength $(p_1 - p_0)$.

9. THE CASE $\tau_S' > 0$

We have studied this case in Ref. 2 and we recall briefly the methods and results.

(a) We assume that *the equation of state is such that* $p = p(\tau, S_0) \to +\infty$ *when* $\tau \to 0$. We denote as m_0 the slope at Z_0 of $\mathcal{S}_0$ ($\tau \leq \tau_0$). Let Δ be a straight line issued from Z_0 with slope m with

$$m < m_0. \tag{9.1}$$

$\mathcal{S}_0$ being strictly concave, Δ meets $\mathcal{S}_0$ at a unique point $Z_A \neq Z_0$ and, for this point, we have, according to (8.1),

$$\mathcal{H}(Z_0, Z_A) < 0. \tag{9.2}$$

But $S(Z_0) = S(Z_A) = S_0$ and S is necessarily stationary on the segment (Z_0, Z_A) for one and only one point Z_s which is a strict maximum for S along Δ. We have $(dS/d\tau)_{\Delta}(Z_0) < 0$ and so $(d\mathcal{H}/d\tau)_{\Delta}(Z_0) < 0$. When Z goes from Z_0 to Z_A, $\mathcal{H}(Z_0, Z)$ is first *positive* and it follows from (9.2) that there exists, between Z_0 and Z_A, a point Z_1 of Δ, necessarily unique such that $\mathcal{H}(Z_0, Z_1) = 0$.

We see that $\mathcal{H}$ is *connected*. According to (8.2), we have for $Z \in \mathcal{H}$, in the neighborhood of Z_0,

$$(dS/d\tau)_{\mathcal{H}} < 0, \quad (dS/dp)_{\mathcal{H}} > 0.$$

It follows from Sec. 7(c) that these inequalities are true on all the Hugoniot curves.

Proposition: If $\tau_S' > 0$ and under the assumptions on the equation of state, the Hugoniot curve $\mathcal{H}$ is connected. More precisely to $m < m_0$ corresponds a unique point $Z_1 \neq Z_0$ of $\mathcal{H}$ (with $\alpha_1\alpha_0 > 0$), such that $m = (\bar{q}_1 - \bar{q}_0)(\tau_1 - \tau_0)^{-1}$. We have on $\mathcal{H}$ the inequalities

$$(dS/d\tau)_{\mathcal{H}} < 0, \quad (dS/dp)_{\mathcal{H}} > 0.$$

(b) Consider an initial state Y_0 and, for $a = a(Y_0)$, the straight line Δ_a of slope m_a issued from Z_0. We assume

$$m_a < m_0, \tag{9.3}$$

that is, $\alpha_0 P(l)_0 > 0$ [see Sec. 7(c)]. To each shock $Y_0 \to Y_1$ corresponds a point Z_1, the intersection of $\mathcal{H}$ and Δ_a. Conversely, let Z_1 be the point $\neq Z_0$ where Δ_a intersects $\mathcal{H}$; Z_1 being known, $\tau_1 < \tau_0$, $p_1 > p_0$, and Ψ_1 are known; f_1 is given by the Hugoniot relation; $r_1 = \tau_1/f_1$ being known, we deduce from (3.3) the value of $u_1^{\alpha} l_{\alpha}$ and from (3.4) the value of $h_1^{\alpha} l_{\alpha}$; $|h_1|^2$ can be given by

$$|h_1|^2(1 - (u_1^{\alpha} l_{\alpha})^2/l^{\alpha} l_{\alpha}) = \Psi_1^2 - (h_1^{\alpha} l_{\alpha})^2/l^{\alpha} l_{\alpha}. \tag{9.4}$$

We can verify that these quantities satisfy the five invariance relations (3.3)—(3.7). We see that each point Z_1 belonging to $\mathcal{H}$ and Δ_a [with (9.3)] and different from Z_0 defines an unique nontrivial solution of the shock equations such that $\alpha_0\alpha_1 > 0$. We have the following theorem.

Theorem: Suppose $\tau_S' > 0$ and the general assumptions on the equation of state. Let Y_0 be a state satisfying $\alpha_0 P(l)_0 > 0$ (that is $v_0^{MF} < v_0^{\Sigma}$ for a fast shock and $v_0^{MS} < v_0^{\Sigma} < v_0^{A}$ for a slow shock). Then there exists a corresponding unique nontrivial solution of the shock equations such that $v_1^{A} < v_1^{\Sigma}$ for a fast shock and, $v_1^{\Sigma} < v_1^{A}$ for a slow shock. The solution satisfies the thermodynamic inequalities (6.4) and the inequalities of the theorem in Sec. 7.

We note that, for $H > 0$, there exists states Y_0 with $m_0 > 0$ *which can give slow shocks, but not weak shocks.*

10. THE CASE $\tau_S' < 0$

(a) Consider the isentropic curves $\bar{q} = \bar{q}(\tau, S)$ defined by (4.4). For a fixed τ, $\bar{q}$ is decreasing when S is increasing and we have the relative location of the strictly concave isentropic curves. Let $Z_1 \neq Z_0$ be an arbitrary point of $\mathcal{H}$ and Δ the ray (Z_0, Z_1) of slope m. According to $\mathcal{H}(Z_0, Z_0) = \mathcal{H}(Z_0, Z_1) = 0$, $\mathcal{H}$ is stationary at one point Z_s between Z_0 and Z_1. The same is true for S, and Z_1 is a strict minimum for S and $\mathcal{H}(Z_0, Z)$ along Δ.

Suppose Z_1 is in the *upper* region determined by $\mathcal{S}_0$; Δ intersects $\mathcal{S}_0$ at one point Z_A such that Z_1 is between Z_0 and Z_A. When Z goes from Z_0 to Z_A, $\mathcal{H}(Z_0, Z)$ is decreasing from Z_0 to Z_s, has a minimum at Z_s, and is increasing thereafter up to the value $\mathcal{H}(Z_0, Z_A) < 0$. We have then a contradiction. We see that *each point* $Z_1 \neq Z_0$ *of* $\mathcal{H}$ *is in the lower region determined by* $\mathcal{S}_0$. Therefore the isentropic curve $\mathcal{S}_1$ of Z_1 corresponds to an entropy $S_1 > S_0$.

Now, when Z goes from Z_0 to Z_1, S is decreasing from Z_0 to Z_s, has a minimum at Z_1, and is increasing

after this point up to the value $S_1 > S_0$. Thus Δ inserts $\mathcal{S}_0$ at one point Z_A, between Z_0 and Z_1. We know that necessarily

$$m < m_0. \tag{10.1}$$

Each point $\neq Z_0$ of $\mathcal{H}$ belongs to the open region defined by by the isentropic curve $\mathcal{S}_0$ of Z_0 and the tangent at Z_0 to $\mathcal{S}_0$.

If $H(Z_0, Z)$ is expressed in terms of the variable $(\tau, \bar{q})$, we have

$$\partial H/\partial \bar{q} = -2f\Theta \tau'_p \tau'^{-1}_S + (\tau - \tau_0).$$

Therefore, $(\partial H/\partial \bar{q})(Z) < 0$ and $\bar{q}$ is a regular function of τ, without singular points, along a connected component $\mathcal{H}_0$ of $\mathcal{H}$.

(b) We consider only in the sequel *the case* $m_0 \leq 0$ (*existence of weak shocks*). Let $Z_1 \neq Z_0$ be a point of $\mathcal{H}_0$. It follows from Sec. 7(b) that

$$Q(m, Z_1) = (d\bar{q}/d\tau)_{\mathcal{S}}(Z_1) - m < 0.$$

But we have $m < m_0 \leq 0$. Therefore

$$(d\bar{q}/d\tau)_{\mathcal{S}}(Z_1) < 0. \tag{10.2}$$

If $(d\bar{q}/d\tau)_{\mathcal{H}}(Z_1) \geq 0$, it follows from (4.7) that

$$2f\Theta(dS/d\tau)_{\mathcal{H}}(Z_1) = -(\tau_1 - \tau_0)(d\bar{q}/d\tau)_{\mathcal{H}}(Z_1) + (q_1 - q_0) > 0,$$

since $\bar{q}_1 - q_0 = m(\tau_1 - \tau_0) > 0$. But it is clear that $(dS/d\tau)_{\mathcal{H}}(Z_1) > 0$ is contradictory with the location of the isentropic curves. It follows that $(d\bar{q}/d\tau)_{\mathcal{H}_0}$ is always strictly negative and, according to *a*, when τ is increasing, $\mathcal{H}_0$ ends necessarily at Z_0.

Therefore *the Hugoniot curve $\mathcal{H}$ is connected* and the same argument as in Sec. 9 shows that we have on $\mathcal{H}$ the inequalities

$$(dS/d\tau)_{\mathcal{H}} < 0, \quad (dS/dp)_{\mathcal{H}} > 0. \tag{10.3}$$

It follows that, for a shock, *the thermodynamic inequalities* (6.4) *are true again under the assumption* $\tau'_S < 0$.

(c) Compare, under the same assumptions, the values of $(d\bar{q}/d\tau)_{\mathcal{H}}$ and $(d\bar{q}/d\tau)_{\mathcal{S}}$. We deduce from (4.7) that

$$2f\Theta(dS/d\tau)_{\mathcal{H}} + (\tau - \tau_0)(d\bar{q}/d\tau)_{\mathcal{H}} - (\bar{q} - \bar{q}_0) = 0.$$

But it follows from the equation of state and from the definition of $\bar{q}$ that

$$(dS/d\tau)_{\mathcal{H}} = \tau'^{-1}_S - \tau'_p \tau'^{-1}_S \{(d\bar{q}/d\tau)_{\mathcal{H}} + c^2 \mu \Psi^2 \alpha^{-1}\}.$$

We obtain

$$\{(\tau - \tau_0) - 2f\Theta \tau'_p \tau'^{-1}_S\}(d\bar{q}/d\tau)_{\mathcal{H}} + 2f\Theta \tau'_p \tau'^{-1}_S(\tau'^{-1}_p - c^2 \mu \Psi^2 \alpha^{-1}) - (\bar{q} - \bar{q}_0) = 0,$$

where

$$(d\bar{q}/d\tau)_{\mathcal{S}} = \tau'^{-1}_p - c^2 \mu \Psi^2 \alpha^{-1}.$$

We can write

$$\{1 - (\tau - \tau_0)(2f\Theta)^{-1} \tau'^{-1}_p \tau'_S\}(d\bar{q}/d\tau)_{\mathcal{H}} = (d\bar{q}/d\tau)_{\mathcal{S}} - (\bar{q} - \bar{q}_0)(2f\Theta)^{-1} \tau'^{-1}_p \tau'_S.$$

Let

$$(d\bar{q}/d\tau)_{\mathcal{H}} - (d\bar{q}/d\tau)_{\mathcal{S}} = (2f\Theta)^{-1} \tau'^{-1}_p \tau'_S (\tau - \tau_0)\{(d\bar{q}/d\tau)_{\mathcal{H}} - (\bar{q} - \bar{q}_0)(\tau - \tau_0)^{-1}. \tag{10.4}$$

Let $Z \neq Z_0$ be an arbitrary point of $\mathcal{H}$, Δ the straight line (Z_0, Z) of slope $m = (\bar{q} - \bar{q}_0)(\tau - \tau_0)^{-1}$, where $m < m_0 < 0$. Suppose

$$(d\bar{q}/d\tau)_{\mathcal{H}}(Z) \leq (d\bar{q}/d\tau)_{\mathcal{S}}(Z).$$

We have, according to Sec. 7(b),

$$Q(m, Z) = (d\bar{q}/d\tau)_{\mathcal{S}}(Z) - m < 0$$

and so

$$(d\bar{q}/d\tau)_{\mathcal{H}}(Z) < m.$$

It follows from (10.4) that we have $(d\bar{q}/d\tau)_{\mathcal{S}}(Z) < (d\bar{q}/d\tau)_{\mathcal{H}}(Z)$ and we obtain a contradiction. Therefore we have

$$(d\bar{q}/d\tau)_{\mathcal{S}}(Z) < (d\bar{q}/d\tau)_{\mathcal{H}}(Z) \tag{10.5}$$

for each point $Z \neq Z_0$ of $\mathcal{H}$.

(d) Let Z_1 and Z_2 (with $\tau_2 < \tau_1$) be two arbitrary points of $\mathcal{H}$, where Z_2 is in the neighborhood of Z_1; Δ is the chord (Z_1, Z_2) of slope m. We have $H(Z_0, Z_1) = H(Z_0, Z_2) = 0$. Therefore, $H(Z_0, Z)$ is stationary at one point Z_σ of the chord (Z_1, Z_2). It follows from (4.7) that we have along Δ

$$dH = 2f\Theta\, dS - K\, d\tau, \tag{10.6}$$

where K is the constant

$$K = (\bar{q}_1 - \bar{q}_0) - m(\tau_1 - \tau_0). \tag{10.7}$$

We have at the point Z_σ

$$\{2f\Theta(dS/d\tau)_\Delta\}(Z_\sigma) = K. \tag{10.8}$$

Let m_1 be the slope of the ray (Z_0, Z_1). *For* $m > m_1$, we have $K > 0$; if $(dS/d\tau)_\Delta \neq 0$ along the chord (Z_1, Z_2), this derivative is positive according to (10.8) and we have at Z_1

$$\{\tau'_S(dS/d\tau)_\Delta\}(Z_1) < 0. \tag{10.9}$$

For $m < m_1$, we have $K < 0$; if $(dS/d\tau)_\Delta \neq 0$ along the chord (Z_1, Z_2), this derivative is negative and we have $S_2 > S_1$. We obtain a contradiction with the relative location of the isentropic curves. *For* $m = m_1$, S is stationary at Z_σ.

Therefore, we study the case where S is stationary at one point Z_s of the chord (Z_1, Z_2), a point which is a strict minimum, according to the lemma in Sec. 7(a). If such is the case, we have again the inequality (10.9).

In all the cases, we have according to (4.10),

$$Q(m, Z_1) > 0.$$

Let

$$(d\bar{q}/d\tau)_{\mathcal{S}}(Z_1) > m.$$

It follows from (10.5) that

$$(d\bar{q}/d\tau)_{\mathcal{H}}(Z_1) > m = \text{slope of the chord } (Z_1, Z_2). \tag{10.10}$$

We deduce from (10.10) that $\mathcal{H}$ *is concave*.

If Z_2 is an arbitrary point of $\mathcal{H}$ and m_2 the slope of the ray (Z_0, Z_2), each ray Δ issued from Z_0 of slope m

such that

$$m_2 \leqslant m < m_0$$

intersects H in one and only one point Z.

We have proved the following Proposition.

Proposition: If $\tau'_S < 0$, $m_0 \leqslant 0$ and under the general assumptions on the equation of state, the Hugoniot curve is connected and concave. We have on H the inequalities

$$(dS/d\tau)_H < 0, \quad (dS/dp)_H > 0.$$

The same argument as in Sec. 9 gives the following theorem.

Theorem: Suppose $\tau'_S < 0$, $m_0 \leqslant 0$ (existence of weak shocks) for a state Y_0 satisfying $\alpha_0 P(l)_0 > 0$ (that is $v_0^{MF} < v_0^D$ for a fast shock and $v_0^{MS} < v_0^D < v_0^A$ for a slow shock). If Z_2 is a point of H and if $(\bar{q}_2 - \bar{q}_0)(\tau_2 - \tau_0)^{-1} \leqslant c^2 a^2 / l^\alpha l_\alpha$, there exists a corresponding unique non trivial solution of the shock equations such that $v_1^A < v_1^D$ for a fast shock and $v_1^D < v_1^A$ for a slow shock. That solution satisfies the thermodynamic inequalities (6.4) and the inequalities of the theorem of Sec. 7.

The case $m_0 > 0$ remains open.

11. CONCLUSION

Under the general Assumption (A) (Sec. 1), we have obtained the following results: timelike character of the shock wavefronts, thermodynamic inequalities (in particular each shock wave is a compression wave), and location of the speeds of the wavefront with respect to the magnetosonic and Alfvên waves.

We have proved also, by two different methods some existence and uniqueness theorems for nontrivial solutions of the shock equations: (1) if $\tau'_S > 0$, (2) if $\tau'_S < 0$, and if the initial state admits weak shocks.

If $\tau'_S > 0$, there exist initial states giving slow shocks, but not weak shocks. The same result is probably true also for $\tau'_S < 0$.

APPENDIX: CLASSICAL APPROXIMATION OF THE RELATIVISTIC SHOCK EQUATIONS

Study the Newtonian approximation of the main shock equations (3.3)—(3.7). We put

$$u^\alpha l_\alpha = w/c, \quad j = rw, \quad l^\alpha l_\alpha = -1,$$

so that $a = j/c$. We search the main parts of the shock equations, with respect to c^{-2}. We have first, according to (3.3),

$$[j] = [rw] = 0. \tag{A1}$$

The invariance relation (3.4) gives

$$(1 + i_1 c^{-2}) h_1^\alpha l_\alpha = (1 + i_0 c^{-2}) h_0^\alpha l_\alpha,$$

so that

$$h_1^\alpha l_\alpha \approx \{1 - (i_1 - i_0)c^{-2}\} h_0^\alpha l_\alpha$$

up to terms in c^{-4}. We obtain in the Newtonian limit

$$[h^\alpha l_\alpha] = 0. \tag{A2}$$

It follows from (3.7) that

$$[rw^2 + p + \tfrac{1}{2}\mu |h|^2] = 0. \tag{A3}$$

It follows from the expression of $h_l^\alpha l_\alpha$ that the invariance of H gives

$$c^2 \cdot \{1 - 2(i_1 - i_0)c^{-2}\} j^{-2} (h_0^\alpha l_\alpha)^2 - |h_1|^2 r_1^{-2}$$

$$= c^2 j^{-2} (h_0^\alpha l_\alpha)^2 - |h_0|^2 r_0^{-2}.$$

We obtain

$$\left[i + \frac{1}{2} \frac{|h|^2}{(h^\alpha l_\alpha)^2} w^2 \right] = 0. \tag{A4}$$

Let, according to $|h|^2 = |k|^2 + (h^\alpha l_\alpha)^2$,

$$\left[i + \frac{1}{2} w^2 + \frac{1}{2} \frac{|k|^2}{(h^\alpha l_\alpha)^2} w^2 \right] = 0. \tag{A4'}$$

Finally, consider the relation of invariance (3.6). The quantity

$$c^{-2}\alpha = (1 + i/c^2) r^{-1} - \mu \{ (h^\alpha l_\alpha)^2 j^{-2} - |h|^2 c^{-2} r^{-2} \}$$

has for Newtonian limit

$$r^{-1} - \mu (h^\alpha l_\alpha)^2 j^{-2}.$$

On the other hand,

$$\Psi^2 = |k|^2 + j^2 c^{-2} |h|^2 r^{-2} \approx |k|^2.$$

We obtain

$$[|k_\alpha|^2 (r^{-1} - \mu (h^\alpha l_\alpha)^2 j^{-2})] = 0. \tag{A5}$$

The five Eqs. (A1)—(A5) are the shock equations of the classical MHD for a rest frame of the shock.

[1] A. Lichnerowicz, *Relativistic Hydrodynamics and Magnetohydrodynamics* (Benjamin, New York, 1967); *Ondes de choc en magnétohydrodynamique relativiste* (Coll. Int. C.N.R.S. n° 184, Lille, 1969), p. 61—71.

[2] A. Lichnerowicz, *Relativistic fluid dynamics*, Centr. Int. Matem. Est. 1970 (Cremonese, Roma, 1971), p. 144—97.

[3] W. Israël, J. Math. Phys. 4, 1163 (1963).

[4] J.C. Lucquiaud, C.R. Acad. Sci. Paris 270, 85 (1970); 271, 1089 (1970); "Sur l'équation de Boltzmann relativiste et applications," thèse (Dijon, 1975).

[5] K.S. Thorne, *Relativistic shocks: the Taub adiabat*, Preprint series in nuclear, atomic, and relativistic astrophysics Calif. Inst. of Technology (1972).

[6] L.D. Landau and E.M. Lifschitz, *Fluid Mechanics* (Pergamon, London, 1959).

[7] A.H. Taub, Phys. Rev. 74, 328 (1948). It is the first paper on relativistic shock waves.

Differential equations satisfied by Fredholm determinants and application to the inversion formalism for parameter dependent potentials

Henri Cornille

Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, BP n° 2—9110 Gif-sur-Yvette, France
(Received 9 May 1976)

We consider the two first order differential operators $A_x = \mu(x)\partial/\partial x + \lambda(x)$, $B_x = -(\partial/\partial x)\mu(x) + \lambda(x)$, associate two kernels f, g satisfying both well-defined boundary conditions and $A_x f = B_y g$, $A_y f = B_x g$, and construct the Fredholm determinants corresponding to these kernels. From these determinants we can build up solutions of second order differential equations. These solutions have an interpretation in the Schrödinger inversion formalism. For instance, for the inversion at fixed angular momentum l, these solutions for $k = 0$ correspond to the classical Gel'fand–Levitan and Marchenko equations, whereas, for $k \neq 0$, they correspond to k dependent potentials. Similarly, for the inversion at fixed k, these solutions for $l = 0$ correspond to the classical Regge–Newton equations, whereas, for $l \neq 0$, they correspond to l-dependent potentials. More generally we show, in a generalized inversion formalism, how parameter dependent potentials appear very naturally in the theory.

I. INTRODUCTION

In recent years, the same differential equation

$$\left(\frac{\partial^2}{\partial r^2} + 2\frac{\partial^2}{\partial r^2}\log B\right)\left(\frac{A}{B}\right) = 0 \tag{1}$$

for two Fredholm determinants A and B (defined on some interval $[r, a]$), has appeared in different physical contexts Eq. (1) which is still satisfied if $A \longleftrightarrow B$ can be interpreted as a Schrödinger equation for angular moment $l = 0$ and momentum $k = 0$, whereas the form of the potential is suggestive of an inversion formalism.[1] First it was found[2] that two opposite kernels of the S wave Marchenko inversion equations lead to Fredholm determinants (defined on $[r, \infty]$) satisfying Eq. (1). Secondly, in connection with the theory of random matrices, Gaudin[3] found that the Fredholm determinants (defined on $[0, r]$) correcponding to two kernels $\Phi(x-y) \pm \Phi(x+y)$ [$\Phi(l)$ even] also satisfy Eq. (1). Thirdly,[4] in the $l \neq 0$, $k = 0$, Marchenko inversion formalism, Eq. (1) was again found for a pair of functions linked to Fredholm determinants. Finally in the random matrix framework, Dyson in a recent paper[5] takes great advantage of Eq. (1) and notes, for a particular case, that it has an interpretation in the S wave, $k = 0$, Gel'fand–Levitan inversion formalism.

All these facts suggest that Eq. (1) could be deduced from a general theorem which has an interpretation in an inversion formalism at $k = 0$, in such a way that all previous results appear as particular applications. This is the first motivation of the paper. Furthermore, we would like to generalize Eq. (1) in order to include the $k^2 \neq 0$ energy term. Investigating some simple examples it is easy to realize that both A and B must be k dependent, so the generalization of Eq. (1), if it exists, leads certainly to k dependent potentials or to a nonclassical inversion formalism, and this is the second motivation of the paper. We recall that the classical inversion formalism is built up in two successive steps. First, from the data and some comparison potential V_0, we have to construct the scattering data kernel which is the input in the inversion integral equation. Secondly, solving this equation, we obtain a kernel solution giving the possibility of reconstructing both the potential to be added to V^0 and the solution of the whole potential. In this paper we consider mainly the second step which is self contained in itself, because from any scattering data kernel we can construct both the potential and the solution, and so reproduce the data.

In this way we shall see in Sec. II that k-dependent,[6] l-dependent,[6] or more generally parameter-dependent potentials appear very naturally in the inversion framework. Let us consider a second order differential operator $\Delta_0 = \mu^2(r)[\partial^2/\partial r^2 - V_0(r)]$ and build a scattering data kernel with the eigenfunction of Δ_0. The generalized inversion formalism consists of the reconstruction of another second order differential operator $\Delta = \mu^2(\partial^2/\partial r^2 - V)$, where $V - V_0$ depends on a Fredholm determinant associated with the scattering data kernel. Two cases occur depending upon whether or not V_0 contains the potential δ/μ^2 (δ being a constant). If V_0 contains $\delta\mu^{-2}$, then the eigenvalues and the eigenfunctions of Δ_0, V, and Δ are δ dependent, whereas if V_0 is $\delta\mu^{-2}$ independent, then V and Δ are δ independent. For instance, if we take $\mu = 1$ (inversion at fixed l) and $V_0 = l(l+1)r^{-2}$, then V is k independent, whereas if $V_0 = -k^2 + l(l+1)r^{-2}$, then V is k dependent. Similarly, if $\mu = r$ (inversion at fixed k) and $V_0 = k^2$, then V is l independent whereas if $V_0 = k^2 + l(l+1)r^{-2}$, then V is l dependent. Let us consider two comparison potentials V_0 differing only by $\delta\mu^{-2}$. Associated with the case for which V_0 does not contain $\delta\mu^{-2}$ is the solution $\Delta u = \delta u$, whereas with the other V_0 containing $\delta\mu^{-2}$, we associate the solution $\Delta u = 0$. It is clear that the two solutions and the two formalisms coincide for $\delta = 0$. Conversely if we start from the solution corresponding to $\delta = 0$ and try, via the inversion formalism, to extend the solutions for $\delta \neq 0$, we have two possibilities depending upon whether V_0 contains δ or not. In this way, extending to $k \neq 0$ the differential Eq. (1) satisfied by coupled determinants in the inversion at fixed l (or $l \neq 0$ in the fixed k case) shall lead to k dependent (l dependent) potentials. The reason for this is that in the extension of Eq. (1), both determinants remain linked, one of them being necessarily k dependent in order to

describe a k dependent solution; then the other is also k dependent and consequently the potential is k dependent. In Secs. III and IV, independently of the inversion formalism, we establish a general mathematical theorem concerning second order differential equations satisfied by linked Fredholm determinants. Let us consider two kernels $f(x,y)$, $g(x,y)$, their corresponding determinants, $D_f=\det[1+\rho f]_r^a$, $D_g=\det[1+\rho g]_r^a$ (a and ρ being fixed constants), and two arbitrary functions $\lambda(x)$, $\mu(x)$ (f,g,λ,μ can depend on other parameters).

Let us assume:

$$\text{(i) } A_x f=B_y g \text{ and } A_y f=B_x g, \text{ with } A_x=\mu\partial/\partial x+\lambda,\quad B_x=-(\partial/\partial x)\mu(x)+\lambda(x) \tag{2a}$$

which can be written as

$$(\Delta_0^g(x)-\Delta_0^g(y))\bar{g}=0,\quad \bar{g}=\mu(x)\mu(y)g,\quad \Delta_0^g=\mu^2\left(\frac{\partial^2}{\partial x^2}-V_0^g\right),$$

$$(\Delta_0^f(x)-\Delta_0^f(y))\bar{f}=0,\quad \bar{f}=\mu(x)\mu(y)f,\quad \Delta_0^f(x)=\mu^2\left(\frac{\partial^2}{\partial x^2}-V_0^f\right),$$

$$V_0^g=\left(\frac{\lambda}{\mu}\right)^2+\left(\frac{\lambda}{\mu}\right)',\quad V_0^f=\left(\frac{\lambda}{\mu}\right)^2-\frac{(\lambda\mu)'}{\mu^2}+\frac{\mu''}{\mu}; \tag{2b}$$

(ii) either $\bar{f}(x,a)=\bar{f}(a,y)=0$ or $\bar{g}(x,a)=\bar{g}(a,y)=0$, or $\bar{f}(x,a)=\bar{g}(x,a)=0$ or $\bar{f}(a,y)=\bar{g}(a,y)=0$.

If (i) and (ii) hold, then D_g and D_f satisfy

$$\left(\frac{\partial^2}{\partial r^2}-V^g\right)\left(\frac{D_f}{D_g}\exp\int^r\frac{\lambda}{\mu}dx\right)=0,$$

$$V^g=V_0^g-\frac{2}{\mu}\frac{\partial}{\partial r}\left(\frac{\mu}{D_g}\frac{\partial D_g}{\partial r}\right),$$

$$\left(\frac{\partial^2}{\partial r^2}-V^f\right)\left(\mu\frac{D_g}{D_f}\exp\left(-\int^r\frac{\lambda}{\mu}dx\right)\right)=0, \tag{3}$$

$$V^f=V_0^f-\frac{2}{\mu}\frac{\partial}{\partial r}\left(\frac{\mu}{D_f}\frac{\partial D_f}{\partial r}\right).$$

If V_0^g (or V_0^f) contains a term $\delta\mu^{-2}$, then V is δ dependent. This theorem depends on two arbitrary functions among the four functions λ,μ,V_0^f,V_0^g; and if we specify two of them, then the other two can be determined by differentiation or integration. There exists a general method to build kernels (f,g) satisfying (2b), and this is also explained in Sec. IV.

In Sec. V we study the inversion at fixed l corresponding to $\mu=1$. If λ is k independent, then Eqs. (2) and (3) correspond to the classical inversion formalism at fixed l and $k=0$: The Gel'fand–Levitan equation when $a=0$, and the Marchenko equation when $a=\infty$. If $\lambda=0$ we come back to Eq. (1) with with an S wave, whereas if $\lambda=l/x$ we get the $l\neq 0$ case. Furthermore, if λ (or V_0^g or V_0^f) is k dependent, then the reconstructed potentials V are k dependent. Thus the solutions satisfying our general theorem, Eqs. (2) and (3), correspond to the classical inversion equations for $k=0$, and to energy dependent potentials for $k\neq 0$.

In Sec. VI we study the case $\mu=r$ corresponding to the inversion at fixed k. If λ is l independent, then Eqs. (2) and (3) correspond to the solution $l=0$ of the classical inversion formalism.[7] Further, if V_0^g (or V_0^f) including the centrifugal potential is l dependent, then the reconstructed potentials given by Eq. (3) are also l dependent.

By investigating, in the l-dependent case, some simple degenerate scattering data kernels, we show how they are linked to (δ_l), the set of phase shifts. However they are only partly restricted by the set (δ_l) and have less constraints than in the l-independent case. Once the set (δ_l) is given, there is still much arbitrariness in these scattering data kernels. If μ is different from both 1 and r, then Eqs. (2) and (3) correspond to generalized inversion equations. Some simple examples are quoted in Sec. VII.

Our general result, Eqs. (2) and (3), can also be interpreted in another way. If (V_0^f,V_0^g) are given as input (or the second order differential operators $\partial^2/\partial r^2-V_0^g$, $\partial^2/\partial r^2-V_0^f$, one may construct coupled first order linear operators (A_x,B_x) and corresponding kernels (f,g). In this way and due to the fact that (λ,μ) are solutions of differential equations, many different coupled (λ,μ) solutions can be obtained, and this is illustrated in Sec. VII.

II. GENERALIZED INVERSION FORMALISM FOR PARAMETER DEPENDENT POTENTIALS (k DEPENDENT, l DEPENDENT, ...)

We establish some general results (see Appendix A) of a generalized inversion formalism which emphasize the fact that parameter dependent reconstructed potentials appear naturally in the theory.

(i) Define a second order differential operator

$$\Delta_0=\mu^2(r)\left(\frac{\partial^2}{\partial r^2}-V_0\right), \tag{4}$$

where μ and V_0 are arbitrary r functions. Further, V_0 can depend on different parameters $k,l,\cdots$ and V_0, the comparison potential, can be written

$$V_0(r,\delta,\alpha_1,\alpha_2,\ldots,\alpha_i)=\frac{\delta}{\mu^2(r)}+\sum_i V_0^i(r,\alpha_i)$$

with δ and α_i being fixed constants which can vanish.

We associate a kernel $\bar{g}(x,y)$ satisfying both differential relations and boundary conditions at some fixed value a,

$$(\Delta_0(x)-\Delta_0(y))\bar{g}(x,y)=0, \tag{5}$$

$$\lim_{x\to a}\bar{g}(x,y)=\lim_{y\to a}\bar{g}(x,y)=0. \tag{6}$$

There exists a general method of constructing such kernels. Let us consider the eigenfunctions $\Delta_0\Phi_n^0=\delta_n\Phi_n^0$ and associate the kernel

$$\bar{g}(x,y)=\sum_n\Phi_n^0(x)\Phi_n^0(y)c_n,$$

the c_n being constants. If $\lim_{x\to a}\Phi_n^0(x)=0$, then both conditions (5) and (6) are satisfied. (In the classical inversion formalism, $\bar{g}$ is called scattering data kernel because the c_n have to be linked to the observables: phase shifts, bound states, $\cdots$).

Of course δ_n and Φ_n^0 are linked to the potential μ^{-2}. The eigenvalues Φ_n^0 and consequently $\bar{g}$ depend on the parameters α_i and δ. Two different cases occur.

Firstly, δ *is different from zero*, V_0 contains a term $\delta\mu^{-2}$, and the eigenfunctions can be written $(\partial^2/\partial r^2 - \sum_i V_0^i(r, \alpha_i))\Phi_n^0 = (\delta + \delta_n)\mu^{-2}\Phi_n^0$. Then Φ_n and $\bar{g}$ are δ *dependent* and correspond *to fixed* α_i *values*. In this case we shall say that we have a δ-dependent inversion formalism at fixed α_i values. Secondly, if δ *is zero*, V_0 does not contain $\delta\mu^{-2}$, and Φ_n^0 and $\bar{g}$ are δ *independent* but correspond to α_i *fixed values*. We shall say that the inversion is δ independent at fixed α_i values. This second case is the restriction on the limit $\delta \to 0$ of the first case. The justification of this nomenclature will appear later when we shall see that the reconstructed potential is entirely determined by $\bar{g}$ and μ. As an illustration of these different possibilities we consider $V_0 = -k^2 + l(l+1)r^{-2}$ in which μ is equal to one or r.

If $\mu = 1$, which corresponds to inversion at fixed l, $\bar{g}$ can be written

$$\bar{g}_l(k; x, y) = \int \hat{j}_l(\sqrt{\gamma}\, x)\hat{j}_l(\sqrt{\gamma}\, y)C_l(k, k')\,dk',$$

$\gamma = k^2 + k'^2$, $\hat{j}_l(\rho)$ being ρ times the spherical Bessel functions[8] and $C_l(k, k')$ an arbitrary function such that $\lim_{k\to 0} C_l(k, k')$ exists. $\bar{g}_l$ corresponds to a *fixed* l *value* and is k dependent if $k \neq 0$. The case $k = 0$ can be obtained as $\lim_{k\to 0}\bar{g}_l$. If $\mu = r$, which corresponds to inversion at fixed k, $\bar{g}$ can be written

$$\bar{g}_k(l; x, y) = \int \hat{j}_{x(\nu)}(kx)\hat{j}_{x(\nu)}(ky)\tilde{C}_k(l, \nu)\,d\nu,$$

$$x(x+1) = \nu(\nu+1) + l(l+1), \quad x \geq 0.$$

$\tilde{C}_k(l, \nu)$ is also an arbitrary function. $\bar{g}_k$ corresponds to a fixed k value and is in general l dependent, the case $l = 0$ being a particular case.

(ii) We define a kernel $\bar{K}(x, y)$ such that the integral equation

$$\bar{K}(r, \sigma) + \bar{g}(r, \sigma) + \int_r^a \bar{K}(r, t)\bar{g}(t, \sigma)\mu^{-2}(t)\,dt = 0 \tag{7}$$

has one and only one solution. Of course $\bar{K}$ corresponds also to fixed α_i values and is δ dependent if V_0 contains a term $\delta\mu^{-2}$. For instance, for $V_0 = -k^2 + l(l+1)r^{-2}$ and $\mu = 1$, then l is fixed but $\bar{K}$ is k dependent.

(iii) We consider the boundary condition

$$\lim_{x\to a}\left(\bar{K}(r, x)\frac{\partial}{\partial x}\bar{g}(x, y) - \bar{g}(x, y)\frac{\partial}{\partial x}\bar{K}(r, x)\right) = 0. \tag{8}$$

Applying the boundary condition (6) to the integral equation (7), we get $\lim_{x\to a}\bar{K}(r, x) = \lim_{r\to a}\bar{K}(r, x) = 0$. It follows that in general (8) is satisfied. In the following we always assume that this is the case.

(iv) We define a reconstructed second order differential operator

$$\Delta = \Delta_0 + 2\mu\frac{d}{dr}\left(\frac{\bar{K}(r, r)}{\mu(r)}\right) = \mu^2\left(\frac{\partial^2}{\partial r^2} - V\right),$$

$$V = V_0 - \frac{2}{\mu(r)}\frac{d}{dr}\left(\frac{\bar{K}(r, r)}{\mu(r)}\right), \tag{9}$$

where V is the reconstructed potential and V_0 is the comparison potential.

Property I: If (i) Eqs. (4)–(6), (ii) Eq. (7), (iii) Eq. (8), (iv) Eq. (9) are satisfied, then, with some algebra, we get for the kernel $\bar{K}(r, \sigma)$

$$(\Delta(r) - \Delta_0(\sigma))\bar{K}(r, \sigma) = 0. \tag{10}$$

[One key property used is the fact that the homogeneous integral equation corresponding to Eq. (7) has only the trivial, identically zero solution.]

(v) We consider an eigenfunction U_n^0 of Δ_0 such that

$$\Delta_0 U_n^0 = \gamma_n U_n^0. \tag{11}$$

We write for these eigenfunctions (γ_n, U_n^0), a different notation from that of the eigenfunctions (δ_n, Φ_n^0) with which we build our kernel $\bar{g}$, because U_n^0 is not necessarily a function of the set $\{\Phi_n^0\}$ entering into $\bar{g}$.

(vi) We define a function U_n such that

$$U_n(r) = U_n^0(r) + \int_r^a U_n^0(t)\bar{K}(r, t)\mu^{-2}(t)\,dt. \tag{12}$$

(vii) We consider the boundary condition

$$\lim_{x\to a}\left(\bar{K}(r, x)\frac{\partial}{\partial x}U_n^0(x) - U_n^0(x)\frac{\partial}{\partial x}(\bar{K}(r, x))\right) = 0. \tag{13}$$

We recall that $\lim_{x\to a}\bar{K}(r, x) = 0$. If either $\lim_{x\to a}(\partial/\partial x)(\bar{K}(r, x)) = 0$ or $\lim_{x\to a}U_n^0(x) = 0$, then in general (13) is satisfied.

Property II: If the assumptions leading to Property I are satisfied and further, if (v) Eq. (11), (vi) Eq. (12), and (vii) Eq. (13) are verified, then U_n is an eigenvalue of the differential operator Δ,

$$\Delta(r)\,U_n(r) = \gamma_n U_n(r).$$
$$\mu^2\left(\frac{\partial^2}{\partial r^2} - V(r)\right)U_n(r) = \gamma_n U_n(r). \tag{14}$$

As an application, let us compare both cases where V_0 differs by a $\delta\mu^{-2}$ term.

If V_0 contains a $\delta\mu^{-2}$ term, we associate $\Delta_0 U^0 = 0$ and the solution U defined in Eq. (12), where U_n^0 is replaced by U^0. Property II says that $\Delta U = 0$ or

$$\left(\frac{\partial^2}{\partial r^2} - \sum_i V_0^i(r, \alpha_i) - \delta\mu^{-2} + 2\mu^{-1}\frac{d}{dr} \times (\bar{K}_\delta(r, r)\mu^{-1})\right)U(r, \delta) = 0, \tag{14a}$$

where we have written $\bar{K}_\delta$ in order to emphasize that $\bar{K}$ is δ dependent.

If V_0 does not contain the terms $\delta\mu^{-2}$ while the other $V_0^i(r, \alpha_i)$ are the same, let us consider in Eq. (11) the solution $\tilde{\Delta}_0 U^0 = \delta U^0$, and in Eq. (12) the function $\tilde{U}$, where U_n^0 is replaced by U^0. Property II tells us that $\tilde{\Delta}\tilde{U} = \delta\tilde{U}$ or

$$\left(\frac{\partial^2}{\partial r^2} - \sum_i V_0^i(r, \alpha_i) - \delta\mu^{-2} + 2\mu^{-1}\frac{d}{dr} \times (\bar{K}_{\delta=0}(r, r)\mu^{-1})\right)\tilde{U}(r, \delta) = 0, \tag{14b}$$

where now $\bar{K}$ is δ independent. It is clear that both U and $\tilde{U}$ coincide at $\delta = 0$, whereas for $\delta \neq 0$ they correspond to two different extensions of the solutions at $\delta = 0$, one for δ dependent potentials and the other for δ independent potentials.

Property III: If g is the kernel $\mu(x)\mu(y)g(x,y)=\bar{g}(x,y)$, then the solution $K(r,r)$ of Eq. (7) with kernel $\rho\bar{g}$ (ρ being a constant) and $D_g=\det[1+\rho g]_r^a$ are linked in such a way that

$$\bar{K}(r,r)=\mu^2(r)\frac{\partial D_g}{\partial r}\Big/ D_g. \tag{15}$$

For the proof, in Eq. (7) let us define $K(x,y)\mu(x)\mu(y)=\bar{K}(x,y)$, then Eq. (7) with $\bar{g}\to\rho\bar{g}$ becomes an integral equation where only K and g appear in

$$K(r,x)+\rho g(r,x)+\rho\int_r^a K(r,t)\,g(t,x)\,dt=0. \tag{7'}$$

In Ref. 9 it has been shown that such integral equations have the property

$$K(r,r)=\frac{\partial D}{\partial r}\Big/ D, \tag{15'}$$

where D is the Fredholm determinant corresponding to the kernel ρg. The result (15) follows immediately.

Substituting the result (15) in the definition of Δ and V we get

$$\begin{aligned}\Delta&=\Delta_0+2\mu\frac{\partial}{\partial r}\left(\mu\frac{D_g'}{D_g}\right),\\ V&=V_0-\frac{2}{\mu}\frac{\partial}{\partial r}\left(\mu\frac{D_g'}{D_g}\right).\end{aligned} \tag{9'}$$

Property IV: If $\bar{g}=\sum_n\phi_n^0(x)\phi_n^0(y)C_n$, $\Delta_0\phi_n^0=\delta_n\phi_n^0$, and if ϕ_n is the solution U_n defined in Eq. (12), where U_n^0 is replaced by ϕ_n^0, then we have

$$\begin{aligned}&-\bar{K}(x,y)=\sum_n\phi_n(x)\,\phi_n^0(y)\,C_n,\quad \Delta\phi_n=\delta_n\phi_n,\\ &\phi_n(r)=\phi_n^0(r)+\sum_p C_p\,\phi_p(r)\int_r^a\phi_p^0(r)\,\phi_n^0(t)\,\mu^{-2}(t)\,dt.\end{aligned}$$

If Property IV holds, then ϕ_n and $\bar{K}(x,y)$ can be written in terms of $\{C_n\}$ and known functions. Let us substitute this expression of $\bar{K}$ in Eq. (12) and assume that the behavior of $U_n(r)$, for a well-defined r value, is linked to physical quantities. If this happens to be possible, this means that there is a link between the $\{C_n\}$ and the observables. We shall illustrate this point in Sec. VI for $\mu=r$ and a particular class of degenerate $\bar{g}$ kernels.

III. RELATIONS BETWEEN THE FIRST AND THE SECOND DERIVATIVE OF FREDHOLM DETERMINANTS

The results presented in this section are entirely independent of those of the previous one. Let us consider a kernel $f(x,y)$ and its Fredholm determinant on some interval $[r,a]$, $D_f=\mathrm{Det}[1+\rho f]_r^a$, ρ and a being constants. D_f can be written

$$\begin{aligned}&D_f=\exp-\bar{\Delta}_f,\quad \bar{\Delta}_f=\sum_1^\infty\rho^n\frac{A_n^f}{n},\\ &A_n^f=\int_r^a dx_1\cdots\int_r^a dx_n\,f(x_1,x_2)\,f(x_2,x_3)\cdots f(x_{n-1},x_n).\end{aligned} \tag{16}$$

The first and the second derivative with respect to r are easily calculated (we assume, of course, that the Fredholm theory can be applied, for instance that the kernels are square integrable, differentiable, and that all the traces exist):

$$\begin{aligned}\frac{\partial}{\partial r}\bar{\Delta}_f&=-\sum_1^\infty\rho^nB_n^f,\quad B_n^f=\int_r^a dx_1\cdots\int_r^a dx_{n-1}\\ &\quad\times f(r,x_1)f(x_1,x_2)\cdots f(x_{n-1},r),\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial r^2}\bar{\Delta}_f&=\left(\frac{\partial}{\partial r}\bar{\Delta}_f\right)^2-\frac{d}{dr}f(r,r)-\sum_2^\infty\rho^n(\gamma_n^f+\beta_n^f),\\ \gamma_n^f&=\int_r^a dx_1\cdots\int_r^a dx_{n-1}\,f(r,x_1)f(x_1,x_2)\cdots f(x_{n-2},x_{n-1})\\ &\quad\times\frac{\partial}{\partial r}f(x_{n-1},r),\\ \beta_n^f&=\int_r^a dx_1\cdots\int_r^a dx_{n-1}\left(\frac{\partial}{\partial r}f(r,x_1)\right)f(x_1,x_2)\cdots f(x_{n-1},r).\end{aligned} \tag{17}$$

Let us introduce another kernel $g(x,y)$ and $D_g=\mathrm{Det}[1+\rho g]_r^a$. We get from (17)

$$-\frac{\partial^2}{\partial r^2}(\bar{\Delta}_f+\bar{\Delta}_g)+\left(\frac{\partial}{\partial r}(\bar{\Delta}_f-\bar{\Delta}_g)\right)^2\equiv H, \tag{18a}$$

$$\begin{aligned}H\equiv&\frac{d}{dr}(f(r,r)+g(r,r))-2\left(\frac{\partial}{\partial r}\bar{\Delta}_f\right)\left(\frac{\partial}{\partial r}\bar{\Delta}_g\right)\\ &+\sum_2\mu^n(\gamma_n^f+\gamma_n^g+\beta_n^f+\beta_n^g).\end{aligned} \tag{18b}$$

We note that if H is identically zero, then D_g and D_f satisfy Eq. (1). However, we want to generalize Eq. (1) and so H, given by Eq. (18b), must be expressed in terms of (D_f,D_g) or in terms of the derivatives of these determinants. However, $\partial f/\partial r$ and $\partial g/\partial r$ appear in the terms γ_n and β_n. If we want to eliminate these terms, there must exist relations between $\partial f/\partial x_i$ and $\partial g/\partial x_j$ $(i,j=1,2,\ x_1=x,\ x_2=y)$. Such relations are given in the next section.

IV. COUPLED FIRST ORDER DIFFERENTIAL OPERATORS APPLIED TO KERNELS GENERATING FREDHOLM DETERMINANTS, AND SECOND ORDER DIFFERENTIAL EQUATIONS SATISFIED BY THESE FREDHOLM DETERMINANTS

A. General results

Theorem I: If two kernels $f(x,y)$ and $g(x,y)$ satisfy:

(i) $A_xf=B_yg$, $A_yf=B_xg$,

$$A_x=\mu(x)\frac{\partial}{\partial x}+\lambda(x),\quad B_x=\left(-\frac{\partial}{\partial x}\mu(x)\right)+\lambda(x),$$

λ,μ being arbitrary functions, then Eq. (2a) can be written:

$$\begin{aligned}&(\Delta_0^g(x)-\Delta_0^g(y))\bar{g}=0,\quad \bar{g}=\mu(x)\,\mu(y)g,\\ &\Delta_0^g(x)=\mu^2(x)\left\{\frac{\partial^2}{\partial x^2}-V_0^g(x)\right\}=\mu(x)A_xB_x\,\mu^{-1}(x),\\ &(\Delta_0^f(x)-\Delta_0^f(y))\bar{f}=0,\quad \bar{f}=\mu(x)\,\mu(y)f,\\ &\Delta_0^f(x)=\mu^2(x)\left\{\frac{\partial^2}{\partial x^2}-V_0^f(x)\right\}=\mu(x)B_xA_x\,\mu^{-1}(x),\\ &V_0^g=\left(\frac{\lambda}{\mu}\right)^2+\left(\frac{\lambda}{\mu}\right)',\quad V_0^f=\left(\frac{\lambda}{\mu}\right)^2-\frac{(\lambda\mu)'}{\mu^2}+\frac{\mu''}{\mu}.\end{aligned} \tag{2.b}$$

(ii) If either $\bar{f}(x,a)=\bar{f}(a,y)=0$ or $\bar{g}(x,a)=\bar{g}(a,y)=0$, or $\bar{f}(x,a)=\bar{g}(x,a)=0$, or $\bar{f}(a,y)=\bar{g}(a,y)=0$, then $D_f=\det[1+\rho f]_a^r$ and $D_g=[1+\rho g]_a^r$ satisfy:

$$\left(\frac{\partial^2}{\partial r^2}-V^g\right)U_g=0,\quad U_g=U_0^g\frac{D_f}{D_g},\quad V^g=V_0^g-\frac{2}{\mu}\frac{\partial}{\partial r}\left(\mu\frac{\partial D_g}{\partial r}\right),$$

$$\left(\frac{\partial^2}{\partial r^2} - V^f\right)U_f = 0, \quad U_f = U_0^f \frac{D_g}{D_f}, \quad V^f = V_0^f - \frac{2}{\mu}\frac{\partial}{\partial r}\left(\mu \frac{\partial D_f}{\partial r}\right), \tag{3}$$

$$\Delta_0^g U_0^g = \Delta_0^f U_0^f = 0, \quad U_0^g = \exp\left(\int^r \lambda/\mu\, dx\right),$$

$$U_0^f = \mu \exp\left(-\int^r \lambda/\mu\, dx\right).$$

Let us note that if we define the second order differential operators by $\Delta^f = \mu^2(\partial^2/\partial r^2 - V^f)$, $\Delta^g = \mu^2(\partial^2/\partial r^2 - V^g)$, then Eq. (3) could equivalently be written $\Delta^g U_g = 0$, $\Delta^f U_f = 0$. On the other hand, from Eq. (3) we get $\mu = U_0^g U_0^f$, $\lambda = U_0^f(\partial/\partial r\, U_0^g)$. For the proof, we first consider the first relation given by (2a),

$$\mu(x)\frac{\partial f}{\partial x}(x,y) = -\lambda(x) f(x,y) + \lambda(y) g(x,y) - \frac{\partial}{\partial y}(\mu(y) g(x,y)). \tag{2a'}$$

In the expression of H given by Eq. (18b) we substitute for $(\partial f/\partial r)(r, x_1)$ in β_n^f the above relation (2a'). We integrate by part, take into account the boundary condition, (ii), and then $(\partial f/\partial x_1)(x_1, x_2)$ appears:

$$B_n^f = -\frac{\lambda(r)}{\mu(r)} B_n^f + \int_r^a dx \int_r^a \frac{dx_{n-1}}{\mu(r)} g(r, x_1)\, \mu(x_1) \frac{\partial f}{\partial x_1}(x_1, x_2)$$

$$f \cdots f + B_1^g B_{n-1}^f + \int_r^a dx_1 \cdots \int_r^a dx_{n-1} \frac{\lambda(x_1)}{\mu(r)} g(r, x_1) f \cdots f.$$

Applying once more Eq. (2a') and (ii), the last term is cancelled and $(\partial f/\partial x_2)(x_2, x_3)$ appears, and so on $\cdots$. Finally one gets

$$\beta_n^f + \gamma_n^g = -\frac{\lambda(r)}{\mu(r)} B_n^f + B_n^g\left(\frac{\lambda(r)}{\mu(r)} - \frac{\partial}{\partial r}\log\mu\right) + \sum_{m=1}^{m=n-1} B_m^g B_{n-m}^f. \tag{19a}$$

Similarly the second relation (2a) can be written

$$\frac{\partial}{\partial x}(\mu(x) g(x,y)) = \lambda(x) g(x,y) - \mu(y)\frac{\partial f}{\partial y}(x,y) - \lambda(y) f(x,y) \tag{2a''}$$

and applying both (ii) and (2a'') to β_n^g, we get

$$\beta_n^g + \gamma_n^f = -\frac{\lambda(r)}{\mu(r)} B_n^f + B_n^g\left(\frac{\lambda(r)}{\mu(r)} - \frac{\partial}{\partial r}\log\mu\right) + \sum_{m=1}^{m=n-1} B_m^f B_{n-m}^g. \tag{19b}$$

From (2a) we get also

$$\mu(r)\frac{d}{dr}(f(r,r) + g(r,r)) = \frac{2\lambda(r)}{\mu(r)}(B_1^g - B_1^f) - B_1^g \frac{\mu'}{\mu}. \tag{19c}$$

Finally in the rhs of Eq. (18b) if we substitute the results (19) we get

$$H = \frac{2\lambda(r)}{\mu(r)}(\bar{\Delta}_f' - \bar{\Delta}_g') + \frac{2\mu'}{\mu}\bar{\Delta}_g'.$$

It follows that (i) and (ii) imply the differential equation

$$\left(\frac{\partial \bar{\Delta}_g}{\partial r} - \frac{\partial \bar{\Delta}_f}{\partial r}\right)^2 - \frac{\partial^2}{\partial r^2}(\bar{\Delta}_f + \bar{\Delta}_g) - \frac{2\lambda(r)}{\mu(r)}\left(\frac{2\bar{\Delta}_f}{\partial r} - \frac{\partial \bar{\Delta}_g}{\partial r}\right)$$

$$-\frac{2}{\mu}\left(\frac{d\mu}{dr}\frac{\partial \bar{\Delta}_g}{\partial r}\right) = 0, \tag{20}$$

or equivalently

$$\left(\frac{\partial^2 D_f}{\partial r^2}\right) D_g + D_f \frac{\partial^2 D_g}{\partial r^2} - 2\left(\frac{\partial D_g}{\partial r}\right)\left(\frac{\partial D_f}{\partial r}\right) + \frac{2\lambda(r)}{\mu(r)}\left(D_g \frac{\partial D_f}{\partial r} - D_f \frac{\partial D_g}{\partial r}\right)$$

$$+ \frac{2\mu'}{\mu} D_f \frac{\partial D_g}{\partial r} = 0. \tag{20'}$$

With some elementary algebra it can be shown that Eq. (20) is equivalent to Eq. (3). Let us remark that for $\mu = 1$ and $\lambda = 0$, Eq. (3) reduces to Eq. (1), and hence we have obtained a generalization of Eq. (1).

B. The second order differential operators Δ_0^f, Δ^f, Δ_0^g, and Δ^g depend on two arbitrary functions

In other words, among the four functions λ, μ, V_0^g, and V_0^f, if we give two of them, the other two can be obtained by differentiation or by solving integral equations.

Case 1: λ and μ are given, then V_0^g and V_0^f are obtained from (2b). We remark that only one pair (V_0^g, V_0^f) of functions is obtained when we differentiate (λ, μ). On the contrary in all other cases, we always have to integrate some differential equation and there exist different pairs (λ, μ) depending upon our choices of the solutions of these differential equations. We put $\bar{\lambda} = \lambda\mu^{-1} = \psi^g[\psi^g]^{-1} = -\psi^f[\psi^f]^{-1}$. Equation (2b) can be written:

$$V_0^g = \bar{\lambda}^2 + \bar{\lambda}', \tag{20a}$$

$$V_0^f = \bar{\lambda}^2 - \bar{\lambda}' - 2\bar{\lambda}\mu'\mu^{-1} + \mu''\mu^{-1}, \tag{20b}$$

$$\psi^{g''} - V_0^g \psi^g = 0, \tag{20c}$$

$$\psi^{f''} + 2\psi^{f'}\mu'\mu^{-1} + (\mu''\mu^{-1} - V_0^f)\psi^f = 0. \tag{20d}$$

Case 2: V_0^f and V_0^g are given; we get ψ^g and $\bar{\lambda}$ from (20c), μ from (20b), and $\lambda = \bar{\lambda}\mu$. The general solution is given in Sec. VII, Eq. (75).

Case 3: μ and V_0^g are given; we get ψ^g, $\bar{\lambda}$, and $\lambda = \bar{\lambda}\mu$ from (20c), V_0^f from (20b).

Case 4: λ and V_0^g are given; we get ψ^g, $\bar{\lambda}$, and $\mu = \lambda\bar{\lambda}^{-1}$ from (20c), V_0^f from (20b).

Case 5: μ and V_0^f are given; we get ψ^f, $\bar{\lambda}$, and λ from (20d), V_0^g from (20a).

Case 6: where λ and V_0^f are given is the most difficult one because we have first to solve $\mu''\mu - \mu'\lambda - \mu\lambda' - V_0^f\mu^2 + \lambda^2 = 0$. Once μ is obtained, from $\bar{\lambda} = \lambda\mu^{-1}$ and (20a) we get V_0^g. Here also μ is not uniquely determined from its differential equation, hence V_0^g is not uniquely determined, either.

C. General method for the explicit construction of coupled kernels (f, g) satisfying the conditions of Theorem I

We consider the eigenvalues and eigenfunctions of Δ_0^g, Δ_0^f:

$$\Delta_0^g \phi_n^0 = \lambda_n \phi_n^0, \quad \Delta_0^g(x) = \mu A_x B_x \mu^{-1},$$
$$\Delta_0^f \psi_n^0 = \bar{\lambda}_n \psi_n^0, \quad \Delta_0^f(x) = \mu B_x A_x \mu^{-1}. \tag{21}$$

We want to show that for each eigenvalue λ_n of Δ_0^g (or $\bar{\lambda}_n$ of Δ_0^f) we can associate an eigenfunction of Δ_0^f with the same eigenvalue λ_n (or an eigenfunction of Δ_0^g with the same eigenvalue $\bar{\lambda}_n$). In other words, we want to show that λ_n and $\bar{\lambda}_n$ are the same. Let us define

$$\mu B_x \mu^{-1} \phi_n^0 = \tilde{\psi}_n^0, \quad \mu A_x \mu^{-1} \psi_n^0 = \tilde{\phi}_n^0, \tag{22}$$

so that we get

$$\Delta_0^f \tilde{\psi}_n^0 = \mu B_x A_x \mu^{-1} \mu B_x \mu^{-1} \phi_n^0 = \mu B_x \mu^{-1} \Delta_0^g \phi_n^0 = \lambda_n \tilde{\psi}_n^0,$$

$$\Delta_0^g \tilde{\phi}_n^0 = \mu A_x B_x \mu^{-1} \mu A_x \mu^{-1} \psi_n^0 = \mu A_x \mu^{-1} \Delta_0^f \psi_n^0 = \bar{\lambda}_n \tilde{\phi}_n^0.$$

It follows that $\tilde{\psi}_n^0$ (or $\tilde{\phi}_n^0$) is an eigenfunction of Δ_0^f (or Δ_0^g) with eigenvalue λ_n (or $\bar{\lambda}_n$), however we must specify the boundary conditions of ψ_n^0 and ϕ_n^0. If $\lambda_n = \bar{\lambda}_n$ and if, modulo a constant, ψ_n^0 and $\tilde{\psi}_n^0$ correspond to the same boundary conditions (or ϕ_n^0 and $\tilde{\phi}_n^0$), then $\tilde{\psi}_n^0 = \delta_n \psi_n^0$ and $\tilde{\phi}_n^0 = \epsilon_n \phi_n^0$, δ_n and ϵ_n being constants. In order to construct the kernels (f, g), let us define

$$\mu A_x \mu^{-1} \tilde{\psi}_n^0 = \tilde{\phi}_n^0, \quad \mu B_x \mu^{-1} \tilde{\phi}_n^0 = \tilde{\psi}_n^0, \tag{23a}$$

and substituting into Eqs. (21)—(22) we get

$$\tilde{\phi}_n^0 = \Delta_0^g \phi_n^0 = \lambda_n \phi_n^0, \quad \tilde{\psi}_n^0 = \Delta_0^f \psi_n^0 = \bar{\lambda}_n \psi_n^0. \tag{23b}$$

We can build the symmetric kernels (f, g) in two different manners depending upon whether we start with (ϕ_n^0, λ_n) or $(\psi_n^0, \bar{\lambda}_n)$,

$$f = \sum_n \tilde{\psi}_n^0(x) \mu^{-1}(x) C_n \tilde{\psi}_n^0(y) \mu^{-1}(y),$$
$$g = \sum_n \phi_n^0(x) \mu^{-1}(x) D_n \phi_n^0(y) \mu^{-1}(y), \tag{24a}$$

$$f = \sum_n \psi_n^0(x) \mu^{-1}(x) \bar{C}_n \psi_n^0(y) \mu^{-1}(y),$$
$$g = \sum_n \tilde{\phi}_n^0(x) \mu^{-1}(x) \bar{D}_n \tilde{\phi}_n^0(y)\, \mu^{-1}(y). \tag{24b}$$

They obviously satisfy the second order differential conditions written down in Eq. (2b). We would like to link the constants C_n and D_n (or $\bar{C}_n$ and $\bar{D}_n$) in order that they satisfy the first order conditions $A_x f = B_y g$ and $A_y f = B_x g$. With the help of Eqs. (22)—(23), in both cases of Eq. (24) it is easy to verify that Eq. (2a) is satisfied if

$$\lambda_n C_n = D_n \quad \text{or} \quad \bar{C}_n = \bar{\lambda}_n \bar{D}_n. \tag{25}$$

A sufficient condition in order that the boundary condition (ii) of Theorem I be satisfied is either

$$\phi_n^0(x) \underset{x \to a}{\to} 0 \quad \text{or} \quad \psi_n^0(x) \underset{x \to a}{\to} 0.$$

D. Connection with the inversion formalism for parameter dependent potentials studied in Sec. II

Let us assume that V_0^g and Δ_0^g correspond to V_0 and Δ_0 defined in Eq. (4). (We could equivalently choose V_0^f and Δ_0^f). Applying the formalism developed in Sec. II we know that from the kernel g (or $\bar{g}$) and from the integral equation (7) we can obtain the Fredholm type solution of $K(r, t)$ (or $\bar{K}$) whose Fredholm determinant is D_g, and reconstruct the whole potential V^g (which is the same as that given by Theorem I).

Substituting this type of $\bar{K}$ solution in Eq. (12) we see that U can be written as the ratio of two terms such that the denominator is D_g but nothing is known for the numerator. Our general result means that *if there exists coupled kernels* (f, g) satisfying the conditions of Theorem I, then there exists a solution U such that $U(U_0)^{-1}$ can be written as the ratio of two Fredholm determinants.

Now concerning the solutions given by Theorem I in Eq. (3), what are the corresponding solutions when r goes to the limit a? If we recall that when $r \to a$, $D_f \to 1$ and $D_g \to 1$, we see that in Eq. (3), $\lim_{r \to a} U = U_0$ and then the solutions of Theorem I can be written when $r \to a$

$$\lim_{r \to a} \Delta^g \left(U_0^g \frac{D_f}{D_g} \right) = \Delta_0^g U_0^g = 0, \quad U_0^g = \exp\left(\int^r \frac{\lambda}{\mu}\, dx \right),$$
$$\lim_{r \to a} \Delta^f \left(U_0^f \frac{D_g}{D_f} \right) = \Delta_0^f U_0^f = 0, \quad U_0^f = \mu \exp\left(-\int^r \frac{\lambda}{\mu}\, dx \right). \tag{3'}$$

Let us compare the solution of Theorem I written down in Eq. (3') and those corresponding to Property II studied in Sec. II. From Eq. (12), $U = U_0 + \int_r^a U_0 \bar{K} \mu^{-2}\, dl$, We note that here also $\lim_{r \to a} U = U_0$. It follows that the solutions (3) correspond to $\gamma_n = 0$ in Eqs. (11) and (14). The solutions of Theorem I correspond to $\Delta_0 U^0 = 0$ and $\Delta U = 0$, where U could also be obtained from Eq. (12).

Concerning the interpretation of Theorem I, two different cases occur. Firstly, V_0 does not contain a $\delta \mu^{-2}$ term; the eigenfunctions ϕ_n^0 and ψ_n^0 building f and g as well as V and the solution U are δ independent. In this case, Eq. (3) corresponds to the solution $\delta \equiv 0$. Secondly, V_0 contains a $\delta \mu^{-2}$ term. Then ϕ_n^0, ψ_n^0, f, g, $\bar{K}$, V, and U are δ-dependent functions. The solution given by Theorem I is δ dependent but *cannot correspond to the classical inversion* because V is δ dependent.

All this discussion means that not every inversion formalism solution can be written as in Theorem I. For instance, if we consider the fixed l case $\mu = 1$ and the potential V^g for the classical inversion formalism where $(\partial^2/\partial r^2) \log D_g$ is k independent, only for $k = 0$, the solution could be written in the manner of Theorem I. Otherwise, for $k \neq 0$, both f and g kernels must be k dependent in order to satisfy Theorem I. Thus for $k \neq 0$ the solutions of the classical inversion formalism with energy independent potentials do not reduce to the form given by Theorem I. Similarly for the inversion at fixed k, the solutions which reduce to that of Theorem I correspond for $l \neq 0$ to l dependent potentials and for $l = 0$ to the classical inversion formalism.[7]

E. Application to the very simple example corresponding to Eq. (1)

Although this case is a particular application of a more general result which is established in the next section, due to its importance (it includes Gel'fand—Levitan and Marchenko equations for $l = 0$, $k = 0$) and its simplicity, we present here another independent derivation.

In order that the result of Theorem I reduce to Eq. (1), we necessarily have $\lambda = 0$ and $\mu = \text{const}$. We discuss now the conditions of an application of Theorem I.

1. Discussion of condition (i) of Theorem I

Condition (i) of Theorem I gives $\partial f/\partial x = -\partial g/\partial y$ and $\partial f/\partial y = -\partial g/\partial x$ and consequently $\partial^2 f/\partial x^2 + \partial^2 f/\partial y^2 = \partial^2 g/\partial x^2 + \partial^2 g/\partial y^2 = 0$. Putting $x + y = v$, $x - y = u$, we get $(\partial f/\partial u)(\partial f/\partial v) = (\partial g/\partial u)(\partial g/\partial v) = 0$. It follows that the coupled kernels (f, g) are of the type

$$f=\psi(x-y+\text{const})+\phi(x+y+\text{const}),$$
$$g=\psi(x-y+\text{const})-\phi(x+y+\text{const}), \tag{26}$$

where at this stage ϕ and ψ are arbitrary functions (subject of course to conditions warranting the existence of the Fredholm theory).

2. Discussion of condition (ii) of Theorem I, when a is finite

It is easy to verify that among the four boundary conditions, there is only one possibility $f(x,a)=f(a,y)=0$ requiring $\psi(l)=\phi(l)=\phi(-l)$ and so $f=\phi(x-y)-\phi(x+y-2a)$. Thus, if

$$f=\phi(x-y)-\phi(x+y-2a),\quad \phi(l)=\phi(-l),\quad a \text{ finite},$$
$$g=\phi(x-y)+\phi(x+y-2a), \tag{27}$$

then

$$(D_f=\mathrm{Det}[1+\rho f]_r^a,\quad D_g=\mathrm{Det}[1+\rho g]_r^a)$$

satisfy Eq. (1). For $a=0$ we recover Gaudin's Theorem.[3] A representation of kernels satisfying Eq. (27) can be written down. If

$$f(x,y)=\int \sin k(x-a)\sin k(y-a)\,G(k)\,dk,$$
$$g(x,y)=\int \cos k(x-a)\cos k(y-a)\,G(k)\,dk, \tag{28}$$

then the corresponding coupled determinants (D_f,D_g) satisfy Eq. (1).

For $a=0$, f can be identified with the Gel'fand–Levitan scattering data kernel.[1] G is linked as usual to the Jost function. For a particular $G(k)$ function, Eq. (28) has an interpretation in the theory of random matrices.[3-5]

3. Discussion of the boundary condition (III) of Theorem I when a is infinite

In order that the Fredholm theory exists we must put $\psi\equiv 0$ in Eq. (26) and the remaining possibility is $f=\phi$ and $g=-\phi$. The boundary condition (ii) of Theorem I is automatically satisfied because the existence of the Fredholm theory requires at least $\lim_{t\to\infty}\phi(t)=0$. Thus

$$\text{if } \begin{cases} f=\phi(x+y), \\ g=-\phi(x+y), \end{cases} \text{ then } (\mathrm{Det}[1+\rho\phi]_r^\infty,\ \mathrm{Det}[1-\rho\phi]_r^\infty) \text{ satisfies Eq. (1).} \tag{29}$$

The Marchenko inversion formalism for an S wave is an application of this result,

$$\text{if } \begin{cases} f=\int \exp[ik(x+y)]\,M(k)\,dk, \\ g=-f(x+y), \end{cases} \text{ then } (D_f,D_g) \text{ satisfies Eq. (1).} \tag{30}$$

We recall that in Eq. (28) for the $a=0$ case and in Eq. (30) we can have a discrete spectrum (imaginary k values corresponding to the existence of bound states) as well as a continuum spectrum (k real). The connection between $G(k)$, $M(k)$, and the S matrix or the Jost function as well as the conditions on $G(k)$ and $M(k)$ such that the Fredholm theory exists, are given in the literature.[1]

This simple example studied here corresponds to the inversion at fixed $l=0$ and has solutions of the Theorem I type for $k=0$. In the next section, for fixed $l\neq 0$ or $l=0$, we study more generally the solutions of Theorem I corresponding to $k\neq 0$ as well as $k=0$.

V. APPLICATION OF THEOREM I TO THE INVERSION FORMALISM AT l FIXED ($\mu=1$) IN BOTH $k\neq 0$ AND $k=0$ CASES

For $\mu=1$, besides the simple case $\lambda=V_0^g=V_0^f=0$, $k=0$, and $l=0$, which was studied in Sec. IV and which belongs to the classical inversion formalism,[1] there exists the possibility of extending the formalism of Theorem I to the case where V_0 contains a $k\neq 0$ term. We develop a general formalism, explicitly obtaining the coupled kernels (f,g) for $k\neq 0$, $l\neq 0$ and investigate the connection with the classical formalism when $k\to 0$. The extension of the Gel'fand–Levitan formalism (determinants defined on $[0,r]$) is done in detail whereas the extension of the Marchenko case (determinants on $[r,\infty]$) is only briefly sketched.

A. Extension of the Gel'fand-Levitan formalism for $k\neq 0$ leading to k dependent potentials

In the formulation of Theorem I we put $a=0$ *and assume* $\mu=1$ *and* $V_0^f=-k^2+l(l+1)r^{-2}$. As was explained in Sec. IV, Part B, λ can be deduced as a solution of $\lambda^2+\lambda'=V_0^f$. We choose for λ two different solutions

$$\lambda^\pm(l,k,r)=\left(-\frac{\partial}{\partial r}\,\hat h_l^\pm(kr)\right)\,(\hat h_l^\pm(kr))^{-1},$$
$$\lambda^\pm(l=0,k,r)=\mp ik, \tag{31}$$

where $\hat h_l^\pm(\rho)$ are ρ times the lth spherical Hankel functions.[8] For integer l, $\hat h_l^+$ can be written

$$\hat h_l^+(\rho)=\exp(i\rho)\sum_{n=0}^{l}\frac{i^{n-l}}{(2\rho)^n}\,\frac{(l+n)!}{n!\,(l-n)!},\quad \hat h_l^-(\rho)=(\hat h_l^+(\rho))^*$$
$$\text{for real } \rho. \tag{32}$$

Applying the general method developed in Sec. IV C, to build the kernels (f,g), we get for the eigenvalues (ψ_n^0,ϕ_n^0),

$$\psi_{k'}^0=\hat j_l(\sqrt\gamma\, r),\quad \gamma=k^2+k'^2,\quad (\Delta_0^f+k'^2)\,\psi_{k'}^0=0,$$
$$\phi_{k'}^{0\pm}=A_r^\pm\hat j_l(\sqrt\gamma\, r)=\phi_l^\pm.$$

1. Kernels $(f, g_l^\pm)$ satisfying Theorem I, l being an integer ≥ 0

Let us define:

$$f_l(k;x,y)=\int \hat j_l(\sqrt\gamma\, x)\,\hat j_l(\sqrt\gamma\, y)\,C_l(k,k')\,k'^2\,dk'$$

$$g_l^\pm(k;x,y)=\int \phi_l^\pm(k,k',x)\,\phi_l^\pm(k,k',y)\,C_l(k,k')\,dk'$$

$$\phi_l^\pm(k,k',x)=A_x^\pm\,\hat j_l(\sqrt\gamma\, x)=\left(\frac{\partial}{\partial x}+\lambda^\pm\right)\hat j_l(\sqrt\gamma\, x),\quad \gamma=k^2+k'^2 \tag{33}$$

with $\lambda^\pm$ given in Eq. (31). We easily verify that these kernels satisfy condition (i) of Theorem I, taking into account

$$B_y^{\pm}\phi_l = \left(-\frac{\partial}{\partial y} + \lambda^{\pm}\right)\phi_l^{\pm} = k'^2\hat{j}_l(\sqrt{\gamma}\,y) \quad \text{and} \quad A_x^{\pm}\hat{j}_l = \phi_l^{\pm}.$$

On the other hand, $\hat{j}_l(0) = 0$ and the boundary conditions (ii) of Theorem I, are automatically satisfied.

2. Irregular solution when $r \to 0$; $k^l v_l^{\pm}$ $(k, r) \simeq r^{-l}$ $(2l-1)!!$

If we apply the results of Theorem I to the kernels $(f_l, g_l^{\pm})$ defined in Eq. (33), we see that the functions $V_l^{\pm}$,

$$V_l^{\pm} = \hat{h}_l^{\pm}(kr)\frac{\mathrm{Det}\,g_l^{\pm}}{\mathrm{Det}f_l} \underset{r\to 0}{\approx} r^{-l}k^{-l}(2l-1)!!, \tag{34}$$

are two complex solutions irregular when $r \to 0$, of the Schrödinger equation

$$\left(\frac{\partial^2}{\partial r^2} + k^2 - \frac{l(l+1)}{r^2} + \frac{2\partial^2}{\partial r^2}\log \mathrm{Det}f_l\right)(V_l^{\pm}) = 0. \tag{35}$$

If we assume that C_l is real, then V^f is real and $V_l^{\pm}$ are complex conjugate. Since f_l is k dependent, it follows that $\mathrm{Det}f_l$ and V^f are k dependent. From $\mu = 1$ and Eqs. (34) and (35) we see that $U_0^f = \hat{h}_l^{\pm}$ with $\Delta_0^f U_0^f = 0$ and $U_0^g = [\hat{h}_l^{\pm}]^{-1}$ with $\Delta_0^g U_0^g = 0$.

3. Regular solution u_l when $r \to 0$; $u_l \simeq r^{l+1}$ $[(2l+1)!!]^{-1}$

We always consider $k \neq 0$. Let us define

$$u_l = \frac{V_l^+ - V_l^-}{2ik^{l+1}} = \frac{\mathrm{Im}[(\mathrm{Det}\,g_l^{+})\hat{h}_l^{+}(kr)]}{k^{l+1}\mathrm{Det}f_l}. \tag{36}$$

u_l is a real solution (C_l being real) of the Schrödinger equation (35). Taking into account $\mathrm{Det}(g_l^{\pm})_{r\to 0} \to 1$, $\mathrm{Det}(f_l)_{r\to 0} \to 1$, we get

$$u_l \underset{r\to 0}{\approx} \frac{\hat{h}_l^{+} - h_l^{-}}{2ik^{l+1}} \approx \hat{j}_l(kr)\,k^{-l-1} \approx r^{l+1}[(2l+1)!!]^{-1}.$$

B. Classical Gel'fand-Levitan solution for $k = 0$ deduced from the application of Theorem I

In order to test the consistency of the theory we must consider the $k \neq 0$ solutions obtained in Sec V, Part A, take the limit $k \to 0$, and compare them with the direct application of Theorem I to the case $k = 0$. This comparison is very easy for the irregular solutions $V_l^{\pm} k^l$ but more delicate for the regular solution u_l. We assume $\lim_{k\to 0} C_l(k, k') = C_l(0, k')$. We shall find that finally no real difficulty occurs when $k \to 0$, and so the above results presented in Sec. V A, are an extension of those for $k = 0$.

1. Irregular solution when $r \to 0$

(a) We first take the limit $k \to 0$ of different quantities defined above in Sec. V A,

$$\begin{aligned}
&\lim_{k\to 0}\lambda^{\pm}(l, k, r) = l/r,\\
&\lim_{k\to 0}\phi_l^{\pm}(k, k', x) = \left(\frac{\partial}{\partial x} + \frac{l}{x}\right)\hat{j}_l(k'x) = k'\hat{j}_{l-1}(k'x),\\
&\lim_{k\to 0} f_l(k; x, y) = f_l(0; x, y) = \int \hat{j}_l(k'x)\hat{j}_l(k'y)\,k'^2 C_l(0, k')\,dk',\\
&l \neq 0, \quad \lim_{k\to 0} g_l^{\pm}(k; x, y) = g_l(0, x, y)\\
&\quad = \int \hat{j}_{l-1}(k'y)\int \hat{j}_{l-1}(k'x)\hat{j}_{l-1}(k'y)\,k'^2 C_l(0, k')\,dk',\\
&l = 0, \quad g_0(0; x, y) = \int \cos(k'x)\cos(k'y)\,C_0(0, k')\,k'^2\,dk',
\end{aligned} \tag{37, 38}$$

$$\lim_{k\to 0}\frac{k^l V_l^{\pm}(k, r)}{(2l-1)!!} = r^{-l}\frac{\mathrm{Det}(g_l(0; x, y))}{\mathrm{Det}(f_l(0; x, y))}. \tag{39}$$

For $l = 0$ in Eq. (38), both the kernels $(f_0(0; x, y), g_0(0; x, y))$ and the solution D_{g_0}/D_{f_0} are identical with the results established in Sec. IV, Part E, Eq. (28) giving the application of the Theorem I to the Gel'fand–Levitan S wave at $k = 0$. In Eq. (39) we note that $U_0^f = r^{-l}$ and, because $\mu = 1$, $U_0^g = r^l$, leading to $\lambda = U_0^g U_0^f = lr^{-1}$.

(b) *For $l \neq 0$ and $k = 0$* we directly construct the kernels (f_l, g_l) and the solutions given by Theorem I using the general method developed in Sec. IV. We consider $\mu = 1$ and $V_0^f = l(l+1)r^{-2}$. Among the different solutions of the differential equation $\lambda^2 - \lambda' = V^f$ we choose $\lambda = l/r$. It follows that $V_0^g = \lambda^2 + \lambda' = l(l-1)r^{-2}$, $\psi_{k'}^0 = \hat{j}_l(k'x)$, $\phi_{k'}^0 = \hat{j}_{l-1}(k'x)$. Finally the direct construction of the kernels (f_l, g_l) in the $l \neq 0$, $k = 0$ case leads exactly to the same kernels as those obtained from $k \neq 0$ on taking the limit $k \to 0$ as in Eq. (38). Similarly the solution written down in Eq. (39) can be derived directly from Theorem I when $\mu = 1$ and $\lambda = l/r$. Because our formalism, when $k \to 0$, is identical to the classical Gel'fand–Levitan formalism, it follows that $C_l(0, k')$ is directly connected to the Jost function $\mathcal{J}_l(k')$ for energy independent potentials. We recall that the Gel'fand–Levitan scattering data kernel[1] can be written $f_l(0; x, y) = 2\pi^{-1}\int_0^{\infty}\hat{j}_l\hat{j}_l(1 - 1/|\mathcal{J}_l(k')|^2)\,dk'$ and we have only to identify with the expression written down in Eq. (38).

2. Regular solution when $r \to 0$, $u_l r^{-(l+1)} \to \mathrm{const}$

(a) *First, for $k = 0$* we directly apply Theorem I and the method of Sec. IV in order to construct the kernels f_l and g_l (which we call h_l). We again consider $\mu = 1$, $V_0^f = l(l+1)r^{-2}$ but now we choose $\lambda = -(l+1)r^{-1}$ as the solution of the differential equation $V_0^f = \lambda^2 - \lambda'$. We thus apply Theorem I to the pair $(\mu = 1, \lambda = -(l+1)r^{-1})$. We get $V_0^g = (l+1)(l+2)r^{-2}$, $\psi_{k'}^0 = \hat{j}_l(k'r)$, $\phi_{k'}^0 = \hat{j}_{l+1}(k'r)$ and our coupled kernels (f_l, h_l) are

$$\begin{aligned}
f_l(0; x, y) &= \int k'^2\hat{j}_l(k'y)\,\hat{j}_l(k'y)\,C_l(0, k')\,dk',\\
h_l(x, y) &= \int k'^2\hat{j}_{l+1}(k'x)\hat{j}_{l+1}(k'y)\,C_l(0, k')\,dk'.
\end{aligned} \tag{40}$$

We note that the kernel f_l is the same as the one given in Eq. (38). This means that the irregular solution Eq. (39) and the solution obtained by the application of Theorem I to the coupled kernels (f_l, h_l) are solutions of the same Schrödinger equation with $V = V_0 + 2\partial^2/\partial r^2 \times \log \mathrm{Det}(f_l)$. We can verify, using the recurrence relations satisfied by spherical Bessel functions, that $A_x f_l = B_y h_l$, $A_y f_l = B_x h_l$, and the boundary conditions of Theorem I are automatically satisfied. Applying the result of Theorem I, we get

$$\begin{aligned}
&\left(\frac{\partial^2}{\partial r^2} - l(l+1)r^{-2} + 2\frac{\partial^2}{\partial r^2}\log \mathrm{Det}f_l\right)u_l = 0,\\
&u_l = r^{l+1}\left(\frac{\mathrm{Det}h_l}{\mathrm{Det}f_l}\right).
\end{aligned} \tag{41}$$

We note that $U_0^f = r^{l+1}$, $U_0^g = r^{-l-1}$, and $\lambda = U_0^g U_0^f = -(l+1)r^{-1}$.

(b) *Secondly, for $k \neq 0$* we consider the regular solution $u_l(k, r)$ defined in Eq. (36). On the one hand, we

want to show that the $\lim_{k\to 0} u_l(k,r)$ exists, and on the other hand that the limit function behaves as the constant r^{l+1}, when r is small.

Taking into account $\hat{h}_l^+ = \hat{n}_l + i\hat{j}_l$, where $\hat{n}_l$ is ρ times the sperhical Neumann function, and the conditions

$$\frac{\hat{h}_l(\rho)\rho^l}{(2l-1)!!} \underset{\rho\to 0}{\to} 1, \quad \frac{\hat{j}_l(\rho)}{\rho^{l+1}}(2l+1)!! \underset{\rho\to 0}{\to} 1,$$

$$\lim_{k\to 0} u_l(k,r) = [\mathrm{Det}(f_l(k=0))]^{-1} \times \left(\frac{(2l-1)!!}{r^l}\lim_{k\to 0}\frac{\mathrm{Im}(\mathrm{Det}\, g_l^+)}{k^{2l+1}} + \frac{r^{l+1}}{(2l+1)!!} \times \lim_{k\to 0}\mathrm{Re}(\mathrm{Det}\, g_l^+)\right), \tag{42}$$

where $g_l^+ = \int \phi_l^+ \phi_l^+ C_l\, dk'$ is defined in Eq. (33). In Appendix B we show

$$\mathrm{Re}\, g_l^+ \underset{k\to 0}{\approx} \int k_1^2 \hat{j}_{l-1}(k,x)\hat{j}_{l-1}(k,y) C_l(0,k_1)\, dk_1,$$

$$\mathrm{Im}\, g_l^+ \underset{k\to 0}{\approx} -\frac{k^{2l+1}}{[(2l-1)!!]^2}[x^{2l}\int k_1 \hat{j}_l(k_1 x)\hat{j}_{l-1}(k_1 y) \times C_l(0,k_1)\, dk_1 + i\,\mathrm{idem}\, x\to y]. \tag{43}$$

With the results (43) at hand, one can show (with some algebra), from $\mathrm{Det}\, g_l^+$ (which means $\det[1+g_l^+]_0^r$) that both $\lim_{k\to 0} k^{-(2l+1)}\mathrm{Im}(\mathrm{Det}\, g_l^+)$ and $\lim_{k\to 0}\mathrm{Re}(\mathrm{Det}\, g_l^+)$ exist and that the limit in Eq. (42) behaves like r^{l+1} when $r\to 0$. In Sec V C we shall exhibit a very simple example where we shall see that $\lim_{k\to 0} u_l(k,r)$ in Eq. (42) and $u_l(r)$ defined in Eqs. (40) and (41) coincide.

3. Integral relations between the two solutions corresponding to k = 0

The irregular solution Eq. (39) and the regular one Eq. (41) are two different solutions of the same Schrödinger equation. Consequently there must exist relations between them. Let us define in the classical Gel'fand–Levitan formalism the kernels

$$f_{\nu,l}(x,y) = \frac{2}{\pi}\int_0^\infty \hat{j}_\nu(k_1 x)\hat{j}_\nu(k_1 y)\left(1 - \frac{1}{|\mathcal{J}_l(k_1)|^2}\right) dk_1,$$

where $\mathcal{J}_l$ is the lth Jost function. From our above results it follows that

$$y_1 = r^{-l}\frac{\mathrm{Det}f_{l-1,l}}{\mathrm{Det}f_{l,l}} \quad \text{and} \quad y_2 = r^{l+1}\frac{\mathrm{Det}f_{l+1,l}}{\mathrm{Det}f_{l,l}}$$

are solutions of same Schrödinger equation

$$\left(\frac{\partial^2}{\partial r^2} - \frac{l(l+1)}{r^2} + 2\frac{\partial^2}{\partial r^2}\log \mathrm{Det}f_{l,l}\right) y_i = 0, \quad i = 1,2.$$

It follows that we have the following integral relations:

$$y_i = y_j\left(\mathrm{const}_1 + \mathrm{const}_2\int^r \frac{dx}{y_j^2(x)}\right), \quad i,j = 1,2, \quad i\neq j$$

C. A very simple example

We consider a trivial example where the kernels (f_l, g_l) in Eq. (33) and (f_l, h_l) in Eq. (40) are degenerate. Although it corresponds to a pathological potential even when $k\to 0$ where it reduces to a well-known case, it has the advantage of easily checking the theory. One can verify very easily that the solutions satisfy the corresponding Schrödinger equation and one can follow what happens when $k\to 0$.

(a) Let us consider for $k\neq 0$,

$$f_0 = k_1^2 \sin\sqrt{\gamma}\, x\, \sin\sqrt{\gamma}\, y, \quad \gamma = k^2 + k_1^2,$$

$$g_0^\pm = (\sqrt{\gamma}\cos\sqrt{\gamma}\, x \mp ik\sin\sqrt{\gamma}\, x)(\mathrm{idem}\, x\to y).$$

We get

$$\mathrm{Det}f_0 = 1 + k_1^2\left(\frac{\sin 2\sqrt{\gamma}\, r}{4\sqrt{\gamma}} - \frac{r}{2}\right),$$

$$\mathrm{Det}\, g_0^\pm = 1 - \frac{k_1^2 r}{2} - \left(\frac{\gamma + k^2}{4\sqrt{\gamma}}\right)\sin 2\sqrt{\gamma}\, r \mp \frac{ik}{2}(\cos 2\sqrt{\gamma}\, r - 1).$$

The two irregular solutions are $V_0^\pm = \exp(\pm ikr)\,\mathrm{Det}\, g_0^\pm \times (\mathrm{Det}f_0)^{-1}$ and the regular solution is

$$(\mathrm{Det}f_0)\, u_0(k,r) = \frac{\sin kr}{k}\left(1 - \frac{k_1^2 r}{2} - \frac{(\gamma+k^2)}{4\sqrt{\gamma}}\sin 2\sqrt{\gamma}\, r\right) - \tfrac{1}{2}(\cos 2\sqrt{\gamma}\, r - 1)\cos kr. \tag{44}$$

Having the explicit dependence of $\mathrm{Det}f_0$, it is easy to calculate the corresponding dependence of the potential. Taking the limit $k\to 0$ of the solution we get

$$\lim_{k\to 0}[u_0\,\mathrm{Det}f_0] = r\left(1 - \frac{k_1^2 r}{2} - \frac{k_1}{4}\sin 2k_1 r\right) - \tfrac{1}{2}[\cos 2k_1 r - 1].$$

(b) Let us consider for $k=0$ the corresponding kernels of Eq. (40),

$$f_0 = k_1^2 \sin k_1 x \sin k_1 y,$$

$$h_0 = k_1^2\left(\frac{\sin k_1 x}{k_1 x} - \cos k_1 x\right) \quad (\mathrm{idem}\, x\to y).$$

$\mathrm{Det}h_0$ is easily calculated. We find that $u_0(r)\,\mathrm{Det}f_0 = r\,\mathrm{Det}h_0$ is equal to the rhs of Eq. (44) and so is equal to $\lim_{k\to 0}((\mathrm{Det}f_0)u_0(k,r))$, where $\mathrm{Det}f_0$ and $u_0(k,r)$ are the k dependent Fredholm determinant and solution corresponding to $k\neq 0$.

D. Inclusion of a Coulombic potential in the extension of the Gel'fand-Levitan formalism for k dependent potentials

We always consider $a=0$, $\mu=1$, but now $V_0^f = -k^2 + l(l+1)r^{-2} + \eta r^{-1} = \lambda^2 - \lambda'$. From this differential equation we choose the solution

$$\lambda^c = -\left(\frac{\partial}{\partial r}\hat{h}_l^{c,+}(k,r)\right)(\hat{h}_l^{c,+}(k,r))^{-1},$$

where $\hat{h}_l^{c,+}$ is the Coulomb solution such that

$$\hat{h}_l^{c,+} \underset{r\to\infty}{\approx} \exp[i(kr - \eta k^{-1}\log(2kr) - l\pi/2)].$$

In order to construct the kernels (f_l, g_l) we consider (as explained in Sec. IV) the eigenfunctions $\Delta_0^f \psi_n^0 = \lambda_n \psi_n^0$. We choose the regular Coulomb solutions $\psi_{k'}^0 = \hat{j}_l^c(\sqrt{\gamma}\, r)$, $\gamma = k^2 + k'^2$, $(\partial^2/\partial\rho^2 + 1 + \eta^{-1}\rho^{-1} - l(l+1)\rho^{-2})\hat{j}_l^c(\rho) = 0$, $\rho = r\gamma$, $\hat{j}_l^c \underset{r\to 0}{\to} \mathrm{const}\, r^{l+1}$, and for $\phi_{k'}^0$, we apply the operator A_x which is well defined once λ and μ are given, $\phi_{k'}^0 = A_x \hat{j}_l^c$. We get for the coupled kernels:

$$f_l^c = \int \hat{j}_l^c(\sqrt{\gamma}\, x)\hat{j}_l^c(\sqrt{\gamma}\, y)\, C_l^c(k',k)\, k'^2\, dk', \quad \gamma = k^2 + k'^2,$$

$$g_l^c = \int \phi_l^c(k,k',x)\,\phi_l^c(k,k',y)\,C_l^c(k',k)\,dk',$$

$$\phi_l^c(k,k',x) = A_x \hat{f}_l^c = \left(\frac{\partial}{\partial x} + \lambda^c\right) \hat{f}_l^c(\sqrt{\gamma}\,x). \tag{45}$$

We check easily that $A_x f_l^c = B_y g_l^c = \int \phi_l^c \hat{f}_l^c k'^2 C_l^c\,dk'$ and that the boundary condition (ii) of Theorem I is satisfied because $\hat{f}_l^c(0) = 0$.

Finally the solution given by Theorem I is an irregular solution when $r \to 0$,

$$\left(\frac{\partial^2}{\partial r^2} + k^2 - \frac{l(l+1)}{r^2} - \frac{\eta}{r} + 2\frac{\partial^2}{\partial r^2}\log \mathrm{Det} f_l^c\right) \times \left(\hat{h}_l^{c,+} \frac{\mathrm{Det} g_l^c}{\mathrm{Det} f_l^c}\right) = 0 \tag{46}$$

We note that in Eq. (46), $U_0^f = \hat{h}_l^{c,+}$ and consequently $U_0^g = [\hat{h}_l^{c,+}]^{-1}$. We could also obtain the other irregular solution when $r \to 0$ of Eq. (46) by choosing another λ. For this it is sufficient to replace $\hat{h}_l^{c,+}$ by the Coulomb solution $\hat{h}_l^{c,-}$, behaving when $r \to \infty$ as $\exp[-i(kr - \eta k^{-1} \times \log 2kr - l\pi/2)]$. We could also extend to the Coulombic case all that has been done above for the non-Coulombic one.

E. Marchenko type inversion formalism

In this case $\mu = 1$ and $a = \infty$. We do not give a detailed analysis as in the preceding Gel'fand—Levitan case. We only consider some aspects of the theory, the missing points presented in Sec. V. A—D can easily be reconstructed. We always assume that our coupled kernels (f, g) satisfy the conditions of Fredholm theory. It follows that the boundary conditions (ii) of Theorem I are always satisfied,

$$\lim_{y\to\infty} f(x,y) = \lim_{x\to\infty} f(x,y) = \lim_{y\to\infty} g(x,y) = \lim_{x\to\infty} g(x,y) = 0.$$

For the explicit constructions of the coupled kernels (f_l, g_l) we only discuss the coupled differential conditions (i) of Theorem I.

1. S wave: $k \neq 0$ (extension of Marchenko equations) and $k = 0$ (classical Marchenko equations)

We consider $\mu = 1$, $\lambda^\pm = \pm ik$, $V_0^g = V_0^f = -k^2$. In this case the kernels (f_0, g_0) are functions of only one variable $t = x + y$ and we assume (as for the classical Marchenko case[1]) that f and g are absolutely integrable. Let us consider the functions

$$f_0^\pm(k;t) = \int (\sqrt{\gamma} \mp k)\,M(k,k')[\exp(i\sqrt{\gamma}\,t)]\,dk', \quad \gamma = k^2 + k'^2,$$
$$g_0^\pm(k;t) = -\int (\sqrt{\gamma} \pm k)\,M(k,k')[\exp(i\sqrt{\gamma}\,t)]\,dk', \tag{47}$$

and associate the kernels $(f_0^\pm(k;x+y),\ g_0^\pm(k;x+y))$. We easily verify

$$A_x^\pm f_0^\pm = \left(\frac{\partial}{\partial x} \pm ik\right) f_0^\pm = i\int k'^2 M(k,k')\,[\exp(i\sqrt{\gamma}(x+y))]\,dk' = B_y^\pm g_0^\pm = \left(-\frac{\partial}{\partial y} \pm ik\right) g_0^\pm,$$

and so the Fredholm determinants associated with (47) satisfy the result of Theorem I,

$$\left(\frac{\partial^2}{\partial r^2} - k^2 + 2\frac{\partial^2}{\partial r^2}\log \mathrm{Det} f_0^\pm\right)\left(\exp(\mp ikr)\frac{\mathrm{Det} g_0^\pm}{\mathrm{Det} f_0^\pm}\right) = 0. \tag{48}$$

For $k = 0$ we recover the usual Marchenko case for $l = 0$, $k = 0$ discussed previously in Sec. IV E. 3, Eq. (30). In Eq. (48), we check that $U_0^f = \exp(\mp ikr)$, $U_0^g = \exp(\pm ikr)$, and consequently $\lambda = \pm ik$.

However, the $f_0^\pm$ kernels in Eq. (47) are complex, and we want, as in the Gel'fand—Levitan case, to have a real reconstructed potential, so we consider

$$f_0(k;t) = \int_0^\infty k'^2\,dk'[M(k,k')\exp(i\sqrt{\gamma}\,t) + M^*(k,k')\exp(-i\sqrt{\gamma}\,t)],$$
$$g_0(k;t) = -\int_0^\infty dk'[(k+\sqrt{\gamma})^2 M\exp(i\sqrt{\gamma}\,t) + (\sqrt{\gamma}-k)M^* \times \exp(-i\sqrt{\gamma}\,t)], \tag{49}$$

leading to the kernels $(f_0(k;x+y),\ g_0(k;x+y))$. It is easy to verify

$$A_x f_0 = \left(\frac{\partial}{\partial x} + ik\right) f_0 = i\int_0^\infty k'^2\,dk'[M(\sqrt{\gamma}+k) \times \exp(i\sqrt{\gamma}(x+y)) + M^*(-\sqrt{\gamma}+k) \times \exp(-i\sqrt{\gamma}(x+y)] = B_y g_0$$

It follows that $\mathrm{Det} f_0$ and $\mathrm{Det} g_0$ satisfy

$$\left(\frac{\partial^2}{\partial r^2} - k^2 + 2\frac{\partial^2}{\partial r^2}\log \mathrm{Det} f_0\right)(\exp(-ikr)\,\mathrm{Det} g_0(\mathrm{Det} f_0)^{-1}) = 0.$$

When $k \to 0$, f_0 and g_0 reduce to $f_0(t) = -g_0(t)$ and we recover the classical Marchenko case studied in Sec. IV E. 3.

For the extension to the $l \neq 0$, $k \neq 0$ case we must replace $\hat{h}_0$ by $\hat{h}_l$ and construct the corresponding λ, as was done for the extension of the Gel'fand—Levitan case where $\hat{f}_0$ was replaced by $\hat{f}_l$ and the corresponding λ in the $l \neq 0$ case was explicitly constructed.

2. $l \neq 0$ and $k = 0$: classical Marchenko inversion equations for $k = 0$

We consider $\mu = 1$ and $V_0^f = l(l+1)\,r^{-2}$. From the differential equation $V_0^f = \lambda^2 - \lambda' = l(l+1)\,r^{-2}$ we choose the two solutions lr^{-1} and $-(l+1)\,r^{-1}$.

a. $\mu = 1$ and $\lambda = lr^{-1}$: We have the coupled kernels $(f_l, g_l^{(1)})$,

$$f_l(x,y) = \int \hat{h}_l(k'x)\,\hat{h}_l(k'y)\,M_l(k')\,dk',$$
$$g_l^{(1)}(x,y) = \int \hat{h}_{l-1}(k'x)\,\hat{h}_{l-1}(k'y)\,M_l(k')\,dk'. \tag{50}$$

Taking into account the recurrence relations $(\partial/\partial x + lx^{-1})\,\hat{h}_l(kx) = k\hat{h}_{l-1}$ and $(-\partial/\partial x + lx^{-1})\,\hat{h}_{l-1}(kx) = k\hat{h}_l$ we can verify that the relations $A_x f_l = B_y g_l$ are satisfied.

b. $\mu = 1$ and $\lambda = -(l+1)\,r^{-1}$: We get the coupled kernels $(f_l, g_l^{(2)})$,

$$g_l^{(2)}(x,y) = \int \hat{h}_{l+1}(k'x)\,\hat{h}_{l+1}(k'y)\,M_l(k')\,dk', \tag{51}$$

and can also verify that $(f_l, g_l^{(2)})$ satisfy the conditions of Theorem I.

c. a, b, and Theorem I: From *a*, *b*, and the application of Theorem I, it follows that the differential equation

$$\left(\frac{\partial^2}{\partial r^2} - \frac{l(l+1)}{r^2} + 2\frac{\partial^2}{\partial r^2}\log \mathrm{Det} f_l\right) y_l = 0 \tag{52}$$

has two solutions $y_l^{(1)} = r^{-l}\,\mathrm{Det} g_l^{(1)}/\mathrm{Det} f_l$ and $y_l^{(2)} = r^{l+1}\mathrm{Det} g_l^{(2)}/\mathrm{Det} f_l$. Here also we have the integral relations $y_l^{(i)} = y_l^{(j)}[\mathrm{const}_1 + \mathrm{const}_2 \int^r dx/(y_l^j(x))^2]$.

We recall[4] that these relations as well as the significance of these solutions $y_l^{(i)}$ $(i=1,2)$ were previously established in the N/D formalism where the Marchenko inversion formalism was restricted to generalized Yukawa potentials. We recall also that these integral relations were useful in order to understand the physical validity of the inversion (an N/D) formalism: right threshold behavior for the S_l matrix, appearance of ghosts in the theory, ···. Finally we recall that if we identify the f_l kernel in Eq. (50) with the classical Marchenko kernel, then $M_l(k) = \mathrm{const}(S_l(k) - 1)$.

VI. INVERSION FORMALISM AT FIXED k ($\mu = r$) FOR l DEPENDENT POTENTIALS IN BOTH $l \neq 0$ AND $l = 0$ CASES

For the application of Theorem I we do not give a detailed analysis as in the fixed l case. Instead we present only some partial results; the missing ones could be easily reconstructed because the method is the same in both cases. For instance, we consider only the Regge–Newton[7] case $a = 0$ (the equivalent of the Gel'fand–Levitan case for fixed l) where the inversion formalism is defined on $[0, r]$. As in the previous section we encounter an extension of the classical inversion formalism. For $l \neq 0$, the solutions obtained from the application of Theorem I lead to l dependent potentials, whereas for $l = 0$ we mainly recover the Regge–Newton inversion formalism with l independent potentials.

Independently of Theorem I we apply the formalism defined in Sec. II, for parameter dependent potentials, to l dependent potentials. Although in this paper the problem of the construction of the scattering data kernel from the data is not considered, we give, in this fixed k case, some insight in this direction for a particular class of degenerate scattering data kernels.

A. Application of Theorem 1

1. Construction of the kernels (f_k, g_k) for l dependent potentials

We start with $\mu = r$ and $V_0^g = -k^2 + l(l+1)r^{-2}$. From the differential equation $V_0^g = (\lambda/\mu)^2 + (\lambda/\mu)'$ we choose two solutions $\lambda^\pm$,

$$\lambda^\pm = r\left(\frac{\partial}{\partial r}\hat{h}_l^\pm(kr)\right)(\hat{h}_l^\pm(kr))^{-1}, \quad \lambda^\pm(l=0,k,r) = \pm ikr. \tag{53}$$

For each of these choices $\lambda^\pm$ we shall associate the kernels $f_k^\pm$ and obtain the solutions $V_k^\pm$. In order to build up the kernels $(f_k^\pm, g_k)$ from the eigenfunctions ϕ_n^0, ψ_n^0 we apply the general method developed in Sec. III. For ϕ_n^0 we apply $\Delta_0^g\phi_n^0 = \lambda_n\phi_n^0$ given in Eq. (21) and we find

$$r^2\left\{\frac{\partial^2}{\partial r^2} + k^2 - \frac{l(l+1)}{r^2}\right\}\phi^0_{\chi(\nu)}(kr) = \nu(\nu+1)\,\phi^0_{\chi(\nu)}(kr),$$

$$\begin{cases}\chi(\chi+1) = \nu(\nu+1) + l(l+1),\\ \chi \geq l,\end{cases}$$

l integer, $l \geqslant 0$. We choose $\phi_\chi^0 = \hat{j}_\chi$ where $\hat{j}_\chi(\rho) \underset{\rho\to 0}{\approx} \mathrm{const}\rho^{\chi+1}$. For ψ_n^0 we apply the operator $\mu B_x\mu^{-1}\phi_n^0 = \tilde{\psi}_n^0$ given in Eq. (22) and obtain

$$yB_y^\pm(\hat{j}_\chi y^{-1}) = \psi_\chi^\pm,$$

where we write $\psi_\chi^\pm$ as $\psi^\pm(\nu, l, ky)$. Let us consider the coupled kernels $(f_k^\pm, g_k)$:

$$\begin{aligned} f_k^\pm(l;x,y) &= -\int \psi^\pm(\nu,l,kx)\psi^\pm(\nu,l,ky)C(\nu,l)x^{-1}y^{-1}\,d\nu\\ g_k(l;x,y) &= \int \hat{j}_{\chi(\nu)}(kx)\hat{j}_{\chi(\nu)}(ky)\nu(\nu+1)C(\nu,l)x^{-1}y^{-1}\,d\nu,\\ \psi^\pm(\nu,l,kx) &= \left(-x\frac{\partial}{\partial x} + \lambda^\pm\right)\hat{j}_{\chi(\nu)}(kx),\\ \chi(\chi+1) &= \nu(\nu+1) + l(l+1), \quad \chi \geqslant 0, \end{aligned} \tag{54}$$

where $\lambda^\pm$ is given in Eq. (53) and $C(\nu, l)$ is a real function.

The potentials reconstructed from $\mathrm{Det} g_k$ are real. Taking into account $xA_x^\pm(\psi^\pm x^{-1}) = -\nu(\nu+1)\hat{j}_\chi$ and $yB_y^\pm(\hat{j}_\chi y^{-1}) = \psi^\pm$ we easily verify that condition (i) of Theorem I is satisfied. Moreover, $\hat{j}_\chi(0) = 0$ for $\chi > -1$. The boundary condition (ii) of Theorem I is satisfied because $\bar{g}_k(l, 0, y) = \bar{g}_k(l, x, 0) = 0$.

2. Irregular solutions when $r \to 0$, $k^l V_k^\pm \underset{r\to 0}{\simeq} r^{-l}\,(2l-1)!!$

Applying the results of Theorem I to the coupled kernels defined in Eq. (54) we see that the functions

$$V_k^\pm(l,r) = \hat{h}_l^\pm(kr)\left(\frac{\mathrm{Det} f_k^\pm}{\mathrm{Det} g_k}\right) \underset{r\to 0}{\approx} r^{-l}k^{-l}(2l-1)!!$$

are two complex, irregular solutions, when $r \to 0$, of the Schrödinger equation

$$\left(\frac{\partial^2}{\partial r^2} + k^2 - \frac{l(l+1)}{r^2} + \frac{2}{r}\frac{\partial}{\partial r}\left(\frac{r(\partial/\partial r)\mathrm{Det} g_k}{\mathrm{Det} g_k}\right)\right)(V_k^\pm) = 0. \tag{55}$$

Because g_k, defined in Eq. (54), is l dependent, it follows that $\mathrm{Det} g_k$ and V^g are l dependent. Let us not notice that in the representation of the g_k kernel written down in Eq. (54), the term corresponding to $\nu = 0$ is not present; whereas in the general theory presented in Sec. II, paragraph (i), the corresponding term does not vanish. This means that there is a restriction for the kernel g_k (or the potentials) leading to solutions of the type given by Theorem I. For the solution $U_g = V_k^\pm$ we check that $U_0^g = \hat{h}_l^\pm$, $U_0^f = r[\hat{h}_l^\pm]^{-1}$ such that $\lambda^\pm$ is given in Eq. (53).

3. Regular solutions when $r \to 0$, $U_k(l, r) \underset{r\to 0}{\simeq} r^{l+1}[(2l+1)!!]^{-1}$

From the kernels written in Eq. (54) and the corresponding solutions $V_k^\pm(l, r)$ let us define:

$$\begin{aligned} U_k(l,r) &= \frac{k^{-l-1}}{2i}\left(V_k^+(l,r) - V_k^-(l,r)\right)\\ &= k^{-l-1}[\mathrm{Det} g_k]^{-1}\,\mathrm{Im}(\hat{h}_l^+(kr)\mathrm{Det} f_k^+). \end{aligned} \tag{56}$$

Let us recall that in this section k *is fixed* $\neq 0$ and $\lim_{r\to 0}\mathrm{Det} g_k = 1$, $\lim_{r\to 0}\mathrm{Det} f_k^+ = 1$, and $\mathrm{Im}\hat{h}_l^+(kr) = \hat{j}_l(kr)$. It follows from Eq. (56) that $U_k(l, r) \underset{r\to 0}{\approx} r^{l+1}[(2l+1)!!]^{-1}$. A more careful analysis, taking into account $\lim_{r\to 0}(\mathrm{Im}\,\mathrm{Det} f_k^+)\,\mathrm{Re}\hat{h}_l^+$ and $\lim_{r\to 0}(\mathrm{Re}\,\mathrm{Det} f_k^+)\,\mathrm{Im}\hat{h}_l^+$ leads to the same result.

4. Application to $l = 0$ connection with the classical inversion formalism[7]

For $l=0$, we recover the classical inversion formalism at fixed k except that the spectral representation of the kernel g_k has no $\nu=0$ term. Either we put $l=0$ in both kernels of Eq. (54) or we apply directly the method of Theorem I to $\mu=r$, $\lambda^{\pm}=\pm ikr$. In both cases we get the same result,

$$g_k(0;x,y)=\int \hat{j}_\nu(kx)\hat{j}_\nu(ky)\nu(\nu+1)C(\nu,0)(xy)^{-1}\,d\nu,$$
$$f_k^{\pm}(0;x,y)=-\int \psi^{\pm}(\nu,0,kx)\psi^{\pm}(\nu,0,ky)C(\nu,0)(xy)^{-1}\,d\nu, \quad (57)$$
$$\psi^{\pm}(\nu,0,kx)=\left(-x\frac{\partial}{\partial x}\pm ikx\right)\hat{j}_\nu(kx).$$

Starting with the kernels written down in Eq. (57), if we define $V_k^{\pm}(0,r)=\exp(\pm ikr)\,\mathrm{Det}f_k^{\pm}[\mathrm{Det}g_k]^{-1}$ and $U_k(0,r)=(V_k^{+}-V_k^{-})(2ik)^{-1}$, it follows that $V_k^{\pm}$ and U_k are solutions of the differential equation (55) with the centrifugal potential equal to zero. In order to make explicit the connection between our formalism in this case ($l=0$) and Newton's formalism we consider a very simply example in Eq. (57), where the corresponding kernels have only one eigenvalue. Let us put $k=1$ and consider the kernels

$$g=-\hat{j}_\nu(x)\hat{j}_\nu(y)C_\nu(xy)^{-1}, \quad \nu \text{ a positive integer different of zero}, \quad (58)$$
$$f^{\pm}=\frac{C_\nu}{\nu(\nu+1)}\left(-\hat{j}'_\nu(x)\pm i\hat{j}_\nu(x)\right)(\text{idem } x\to y).$$

We get

$$\mathrm{Det}g=1+C_\nu\int_0^r \hat{j}_\nu^2 x^{-2}\,dx,$$
$$\mathrm{Det}f^{\pm}=1-\frac{C_\nu}{\nu(\nu+1)}\int_0^r(\hat{j}_\nu'^2-\hat{j}_\nu^2)\,dx\pm iC_\nu\hat{j}_\nu^2[\nu(\nu+1)]^{-1}. \quad (59)$$

The two irregular solutions are $V^{\pm}=\exp(\pm ikr)\,\mathrm{Det}f^{\pm}\times[\mathrm{Det}g]^{-1}$ and the regular one is

$$U_0\,\mathrm{Det}g=\mathrm{Im}(\exp(ikr)\,\mathrm{Det}f^{+})$$
$$=\frac{C_\nu}{\nu(\nu+1)}\hat{j}_\nu^2\cos r+\sin r\left(1-\frac{C_\nu}{\nu(\nu+1)}\times\int_0^r(\hat{j}_\nu'^2-\hat{j}_\nu^2)\,dx\right)$$

which can be written, with some algebra, as

$$U_0=\sin r-C_\nu\hat{j}_\nu[\mathrm{Det}g]^{-1}\int_0^r \sin x\hat{j}_\nu x^{-2}\,dx, \quad (60)$$

this last expression coinciding exactly with a particular solution written down in Newton's paper.[10]

5. $l=0$, $k=0$ case

We consider $\mu=r$, $\lambda=0$, $V_0^f=V_0^g=0$ and apply the results of Theorem I. Let us consider the coupled kernels

$$f=\int x^\nu y^\nu G(\nu)\,d\nu,$$
$$g=-\int x^\nu y^\nu\left(\frac{\nu}{\nu+1}\right)G(\nu)\,d\nu, \quad (61)$$

and the corresponding Fredholm determinants $D_f=\mathrm{Det}f=\mathrm{Det}[1+\rho f]_0^r$ and $D_g=\mathrm{Det}g=\mathrm{Det}[1+\rho g]_0^r$. For $\nu>-1$, boundary condition (ii) of Theorem I is satisfied, however, in order that the Fredholm determinants exist it is at least necessary that $\nu>-\frac{1}{2}$. We consider $\nu>0$, and the determinants satisfy

$$\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\left(r\frac{D_g'}{D_g}\right)\right)\left(\frac{D_f}{D_g}\right)=0,$$
$$\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\left(r\frac{D_f'}{D_f}\right)\right)\left(r\frac{D_g}{D_f}\right)=0. \quad (62)$$

Similarly the kernels

$$f=\int x^{-(\nu+1)}y^{-(\nu+1)}M(\nu)\,d\nu,$$
$$g=-\int\left(\frac{\nu+1}{\nu}\right)x^{-(\nu+1)}y^{-(\nu+1)}M(\nu)\,d\nu, \quad (63)$$

for $\nu>0$ satisfy the conditions of Theorem I if we take $a=\infty$. Consequently the determinants $D_f=\mathrm{Det}f=\mathrm{Det}[1+\rho f]_r^\infty$, $D_g=\mathrm{Det}g=\mathrm{Det}[1+\rho g]_r^\infty$ also satisfy the differential equation (62). In Eq. (62) we check that $U_0^f=r$, $U_0^g=1$, leadint to $\mu=r$ and $\lambda=0$.

6. Inclusion of the Coulombic potential

We always consider $a=0$, $\mu=r$, but now $V_0^g=-k^2+l(l+1)r^{-2}+\eta r^{-1}$. For the solutions $r^2(\partial^2/\partial r^2-V_0^g)\phi_\chi^c=\nu(\nu+1)\phi_\chi^c$ we choose $\phi_\chi^c=\hat{j}^c_{\chi(\nu)}$, the regular solution which behaves like $\mathrm{const}r^{\chi+1}$ when $r\to 0$. From $V_0^g=(\lambda/\mu)^2+(\lambda/\mu)^1$ we choose $\lambda^{c,\pm}=r((\partial/\partial r)\hat{h}_l^{c,\pm}(k,\eta,r))\times[\hat{h}_l^{c,\pm}(k,\eta,r)]^{-1}$, where $\hat{h}_l^{c,\pm}$ has been defined in the previous section. For the kernel $f_k^{\pm,c}$ we consider the solution $xB_x^{\pm}(\hat{j}_\chi^c x^{-1})=\psi_\chi=\psi^{\pm}(\eta,\nu,l,k,x)$ and finally obtain for the coupled kernels $(f_k^{c,\pm},g_k)$, satisfying Theorem I,

$$f_k^{c,\pm}(l,\eta;x,y)=-\int\psi^{\pm}(\eta,\nu,l,k,x)\psi^{\pm}(\eta,\nu,l,k,y)\times C^c(\nu,l)x^{-1}y^{-1}\,d\nu,\ g_k^c(l,\eta;x,y)$$
$$=\int\hat{j}^c_{\chi(\nu)}(k,\eta,x)\hat{j}^c_{\chi(\nu)}(k,\eta,y)\nu(\nu+1)\times x^{-1}y^{-1}C^c(\nu,l)\,d\nu,\ \psi^{\pm}(\eta,\nu,l,k,x)$$
$$=\left(-x\frac{\partial}{\partial x}+\lambda^{c,\pm}\right)\hat{j}_{\chi(\nu)}(k,\eta,x),$$
$$\chi(\chi+1)=\nu(\nu+1)+l(l+1), \quad \chi>0, \quad \nu>0,$$

$l\geq 0$, l being integer. We can check that $U_0^g=\hat{h}_l^{c,\pm}$ and $U_0^f=r[\hat{h}^{c,\pm}]^{-1}$ in order that $\mu=r$ and $\lambda^{c,\pm}$ is the one given above.

We can apply Theorem I because $xA_x^{\pm}(\psi^{\pm}x^{-1})=-\nu(\nu+1)\hat{j}_\chi^c$, $yB_y^{\pm}(\hat{j}_\chi^c y^{-1})=\psi^{\pm}$, and $\hat{j}^c_{\chi(\nu)}(0)=0$. We get

$$\left(\frac{\partial^2}{\partial r^2}+k^2-\frac{l(l+1)}{r^2}-\frac{\eta}{r}+\frac{2}{r}\frac{\partial}{\partial r}\left(\frac{r(\partial/\partial r)\,\mathrm{Det}g_k^c}{\mathrm{Det}g_k^c}\right)\right)V_k^{c,\pm}=0,$$
$$V_k^{c,\pm}=\hat{h}_l^{c,\pm}(k,\eta,r)\frac{\mathrm{Det}f_k^{c,\pm}}{\mathrm{Det}g_k^c}\underset{r\to 0}{\approx}\mathrm{const}r^{-l}.$$

We could also extend to this case all the results obtained above in the non-Coulombic one.

B. Some results concerning the determination of the scattering data kernel

Forgetting the results of Theorem I we return to the general formalism defined in Sec. II and try to make explicit (always for $\mu=r$) the links between a particular ansatz for the scattering data kernel and the phase

shifts. There exists an arbitrariness in the determination of the spectral function $C_k(\nu,l)$ (k fixed) entering into the scattering data kernel. We have a function of two variables (ν,l) to be linked to a set $\{\delta_l(k)\}$ depending on one variable l. This arbitrariness can be reduced if we require that the same $C(\nu,l)$ occurs for different energies $k_1,k_2,\cdots$.

1. Inversion at one fixed energy k_1

Let us start with $\bar{g}=xyg$ of the type

$$\bar{g}(l;x,y)=-\sum_{\nu}\hat{j}_\chi(x)\hat{j}_\chi(y)C_\chi(\nu,l),\quad \chi(\chi+1)=\nu(\nu+1)+l(l+1), \tag{64}$$

$x\geq\nu$, ν *integer* >0, l integer ≥ 0. From the definition given in Eq. (64) it is easy to verify that $\chi-l$ is never an integer for $l\neq 0$. For $l=0$, $x-l$ is equal to ν, so if we want in the following that $\sin[(\pi/2)(x-l)]\neq 0$ it is sufficient to take odd integer ν values. For simplicity we put $k_1=1$, otherwise we could say (as is usually done) that $k_1x\to x$, $k_1y\to y$. We want to find the link between $C_\chi(\nu,l)$ and the phase shift $\{\delta_l(k_1)\}$.

The solution of the Fredholm linear integral equation (7) [where the kernel $\bar{g}$ is given by Eq. (64)] can be written

$$\bar{K}(r,x)=\sum_{\chi\text{ or }\nu}C_\chi U_\chi(r)\hat{j}_\chi(r), \tag{65}$$

where U_χ is the solution of Eq. (12) when $U_n^0=\hat{j}\chi$,

$$U_\chi(r)=\hat{j}_\chi(r)-\int_0^r\hat{j}_\chi(t)\bar{K}(r,t)t^{-2}\,dt. \tag{66}$$

Thus the knowledge of U_χ determines $\bar{K}$ and conversely. The elimination of $\bar{K}$ between Eqs. (65)–(66) leads to the system

$$U_\chi(r)[1+C_\chi L_{\chi,\chi}(r)]+\sum_{\chi'\neq\chi}U_{\chi'}C_{\chi'}L_{\chi,\chi'}(r)=\hat{j}_\chi(r),$$
$$\chi(\chi+1)=\nu(\nu+1)+l(l+1), \tag{67}$$

where we have put $L_{\chi,\chi'}(r)=\int_0^r\hat{j}_\chi\hat{j}_{\chi'}t^{-2}\,dt$. We have in Eq. (67) a linear system for the determination of the set $\{U_\chi\}$. Note that the determinant of the system [lhs of Eq. (67)] is also the Fredholm determinant of the integral equation (7) giving $\bar{K}$ from $\bar{g}$. This determinant is $D_g(r)=\mathrm{Det}g=\mathrm{Det}[1+g]_0^r=\mathrm{Det}[1+(xy)^{-1}\bar{g}]_0^r$. We recall the following result gotten in Ref. 8: If $\mathrm{Det}[1+\rho g]_0^\infty\neq 0$ for $|\rho|\leq 1$, then $\mathrm{Det}[1+\rho g]_0^r\neq 0$. We assume that we are in this case (otherwise Regge ghosts could occur in the theory) and consequently $D_g(r)\neq 0$ and the set $\{U_\chi(r)\}$ solution of Eq. (67) exists for any r value. Furthermore, V^g has no second order poles for $r\geq 0$. For integer l, the physical solution corresponding to Eq. (12) can be written

$$U_l(r)=\hat{j}_l(r)-\sum_\chi U_\chi(r)C_\chi L_{\chi,l}(r), \tag{68}$$

where we replace $\{U_\chi\}$ by the solution of Eq. (67). It follows that the rhs of Eq. (68) can be written in terms of known functions $L_{\chi,\chi}(r)$, $D_g(r)\hat{j}_l$, $L_{\chi,\chi}(r)$, $L_{\chi,l}(r)$, $\hat{j}_\chi$, and the unknown $C_\chi(\nu,l)$.

If we substitute the asymptotic behavior $U_l(r)\underset{r\to\infty}{\approx}A_l\times\sin(r-l\pi/2+\delta_l)$, $\hat{j}_l\approx\sin(r-l\pi/2)$, $L_{\chi,\chi'}(\infty)$, $D_g(\infty),\cdots$, then from both sides of Eq. (68) we get a set of relations between $\{\delta_l\}$ and $\{C_\chi(\nu,l)\}$.

In order to see more clearly that happens, we consider the case of one and two terms in Eq. (64) or one and two eigenvalues for the kernel g.

Case of one eigenvalue:

$$\bar{g}(l;x,y)=xyg=-\hat{j}_\chi(x)\hat{j}_\chi(y)C_\chi(\nu,l),$$
$$\chi(\chi+1)=\nu(\nu+1)+l(l+1), \tag{69}$$

with $\chi\geq\nu$ and ν an odd integer such that $\sin[(\pi/2)(x-l)]\neq 0$. The solutions for r fixed are

$$U_\chi=\hat{j}_\chi[D_g]^{-1},\quad D_g(r)=1+C_\chi L_{\chi,\chi}(r)$$
$$U_l=\hat{j}_l-C_\chi\hat{j}_\chi L_{l,\chi}[D_g]^{-1}.$$

For $l=0$, we recover for $U_0(r)$ the result given by Eq. (60b). Now we investigate the behavior when $r\to\infty$, taking into account

$$L_{\chi,l}(\infty)=\frac{\sin[(\pi/2)(\chi-l)]}{(\chi-l)(l+\chi+1)},\quad L_{\chi,\chi}(\infty)=\frac{\pi}{2(2\chi+1)}.$$

We get

$$D_g(\infty)A_l\exp(i\delta_l)=D_g(\infty)-C_\chi L_{\chi,l}(\infty)\exp[i\pi(l-\chi)/2].$$

If we define $\tilde{C}_\chi=\tilde{C}_\chi(\nu,l)=C_\chi L_{\chi,l}(\infty)\sin[(\pi/2)(l-\chi)]$ we finally have

$$\tilde{C}_\chi=\tilde{C}(\nu,l)=\frac{D_g(\infty)\tan\delta_l}{-1+\tan\delta_l\cot((\pi/2)(l-\chi))},$$
$$\tan\delta_l=\tilde{C}_\chi[(\cot((\pi/2)(l-\chi)))\tilde{C}_\chi-D_g(\infty)]^{-1}. \tag{70}$$

If we except the exceptional values for which the denominators in Eq. (70) can vanish we see that the set $\{\delta_l\}$ determines the set $\{\tilde{C}(\nu,l)\}$ where ν is fixed and conversely. There remains of course the arbitrariness in the choice of ν. In conclusion for the set $\{\delta_l\}$ and the ansatz kernel Eq. (69) there exists a family (depending of one parameter ν) of potentials leading to the same phase shifts.

Case of two eigenvalues:

$$\bar{g}(l;x,y)=xyg=-\sum_{i=1}^{i=2}C_{\chi_i}\hat{j}_{\chi_i}(x)\hat{j}_{\chi_i}(y),$$
$$\chi_i(\chi_i+1)=\nu_i(\nu_i+1)+l(l+1), \tag{71}$$

ν_i, positive integers but not being necessarily both odd integers. We get for the intermediate solutions $U_{\chi_i}(r)$,

$$U_{\chi_i}D_g=\hat{j}_{\chi_i}+C_{\chi_j}(\hat{j}_{\chi_i}L_{\chi_j,\chi_j}-\hat{j}_{\chi_j}L_{\chi_i,\chi_j}),\quad i\neq j,\ i=1,2,\ j=1,2,$$
$$D_g=[1+C_{\chi_1}L_{\chi_1,\chi_1}][1+C_{\chi_2}L_{\chi_2,\chi_2}]-C_{\chi_1}C_{\chi_2}L^2_{\chi_1,\chi_2}.$$

For the physical solutions $U_l(r)$, we obtain, when $r\to\infty$,

$$D_g(\infty)A_l\sin\delta_l=-\sum_{i=1}^{2}C_{\chi_i}a_i+C_{\chi_1}C_{\chi_2}(a_3+a_4), \tag{72}$$

with $a_i=L_{\chi_i,l}(\infty)\sin[(\pi/2)(l-\chi_i)]$ for $i=1,2$, and $a_3=\sin[(\pi/2)(l-\chi_1)](L_{\chi_1\chi_2}(\infty)L_{\chi_2,l}(\infty)-L_{\chi_2\chi_2}(\infty)L_{\chi_1,l}(\infty)$,

$$a_4=(\sin(\pi/2)(l-\chi_2))(L_{\chi_1,\chi_2}(\infty)L_{\chi_1,l}(\infty)-L_{\chi_1,\chi_1}(\infty)L_{\chi_1,l}(\infty)).$$

Let us define $b_i=a_i\cot[(\pi/2)(l-\chi_1)]$ for $i=1$ and 3, and let $b_i=a_i\cot[(\pi/2)(l-\chi_2)]$ for $i=2$ and 4. We get

$$D_g(\infty)A_l\cos\delta_l=D_g(\infty)-C_{\chi_1}b_1-C_{\chi_2}b_2-C_{\chi_1}C_{\chi_2}(b_3+b_4),$$
$$C_{\chi_2}=\frac{\tan\delta_l D_g(\infty)-C_{\chi_1}(a_1-b_1\tan\delta_l)}{b_2\tan\delta_l-a_2+C_{\chi_1}(a_3+a_4-b_3\tan\delta_l-b_4\tan\delta_l)}. \tag{73}$$

First we consider the problem of transparent potentials or $\{\delta_l\}=\{0\}$ for all physical l values. We can always choose $C(\nu_1,l)$ and $C(\nu_2,l)$ in such a way that for any integer $l\geq 0$, the rhs of Eq. (72) is zero. Secondly we see that if the two sets $\{\delta_l\}$ and $\{C(\nu_1,l)\}$ are given, as well as the ν_2 value fixed, then $C(\nu_2,l)$ is determined by the rhs of Eq. (73). It follows that from any arbitrary $\{C(\nu_1,l)\}$ and arbitrary ν_2, one can determine $C(\nu_2,l)$. If instead of two eigenvalues we take three terms in $\bar{g}$, from $\{\delta_l\}$, $\{C(\nu_1,l)\}$, $\{C(\nu_2,l)\}$, and ν_3 arbitrary, we shall find a relation which determines $C(\nu_3,l)$.

In conclusion, these l-dependent potentials given by an ansatz of the type Eq. (64) are less constrained than the corresponding ones for l-independent potentials (see for comparison Newton's paper).

2. Inversion at two fixed energies k_1 and k_2, at three fixed energies $k_1, k_2, k_3, \ldots$

This arbitrariness, which could appear at first glance as a drawback of the parameter inversion formalism is perhaps an advantage if we try to introduce more observables in the problem.

For instance, let us return to the ansatz of the type written down in Eq. (71) with two terms. We introduce two sets of phase shifts $\{\delta_l(k_j)\}$, $j=1,2$ corresponding to two different energies k_1^2 and k_2^2, and ask if they can lead to the same $C_{\chi_i}(\nu_i,l)$ in both cases. We reintroduce the symbol k_j in the kernels,

$$\bar{g}_{k_j}(l;x,y)=-\sum_{i=1}^{i=2} C_{\chi_i}(\nu_i,l)\hat{j}_{\chi_i}(k_jx)\hat{j}_{\chi_i}(k_jy),\quad j=1,2,\ i=1,2.$$

If we go on the formalism (like in the previous case) we see that Eq. (73) is replaced by two similar equations of the type

$$\begin{aligned} &C_{\chi_1}\alpha_1+C_{\chi_2}\alpha_2+C_{\chi_1}C_{\chi_2}\alpha_3=\alpha_4,\\ &C_{\chi_1}\beta_1+C_{\chi_2}\beta_2+C_{\chi_1}C_{\chi_2}\beta_3=\beta_4, \end{aligned} \tag{73'}$$

where the α_i and β_i are known once $\nu_i, l, k_j, \delta_l(k_j)$ are given, the α_i corresponding to k_1 and the β_i to k_2. Let us assume that ν_1 and ν_2 are fixed (not necessarily both odd integers), then from (73') we could in general determine $C_{\chi_1}(\nu_1,l)$ and $C_{\chi_2}(\nu_2,l)$ for any l value. We have thus reduced the arbitrariness by requiring that the potentials reproduce the phase shifts for all physical l values and two different energies. Of course, from the fact that $C_{\chi_i}(\nu_i,l)$ are the same for k_1 and k_2, does not follow that the potentials are the same because the weight functions, $\hat{j}_{\chi_1}(k_1x)$ and $\hat{j}_{\chi_2}(k_2x)$ in the kernel g_{k_1} and g_{k_2} are different. If in $\bar{g}$ we consider three terms, we could in principle introduce three energies, $\cdots$, and so on.

VII. SOME EXAMPLES OF THE APPLICATION OF THEOREM I TO CASES NOT CORRESPONDING TO $\mu=1$ OR $\mu=r$

Theorem I is not restricted to inversion at fixed l or at fixed k (it has perhaps even a more general meaning than inversion scattering). We give some other examples which must be considered only as mathematical illustrations of the large domain of application of this theorem.

A. Determination of pairs (λ,μ) from pairs (V_0^g, V_0^f)

The link between these two pairs of functions is given in Theorem I and can be written

$$\Delta_0^g U_0^g=0,\quad U_0^g=\exp\int^r \lambda/\mu\,dx, \tag{74a}$$

$$\Delta_0^f U_0^f=0,\quad U_0^f=\mu\exp(-\int^r \lambda/\mu\,dx), \tag{74b}$$

$$U_0^f U_0^g=\mu,\quad U_0^f\frac{\partial}{\partial r}U_0^g=\lambda. \tag{74c}$$

Let us consider $U_{0,1}^g$ and $U_{0,2}^g$, two independent solutions of (74a), and $U_{0,1}^f$ and $U_{0,2}^f$, two independent solutions of (74b). The general pair of solutions (λ,μ) depend on four arbitrary constants $\alpha,\beta,\gamma,\delta$ and can be written from (74c) as

$$\begin{aligned} &\mu(r)=(\alpha U_{0,1}^g(r)+\beta U_{0,2}^g(r))(\gamma U_{0,1}^f(r)+\delta U_{0,2}^f(r)),\\ &\lambda(r)=\left(\alpha\frac{\partial U_{0,1}^g(r)}{\partial r}+\beta\frac{\partial}{\partial r}U_{0,2}^g\right)(\gamma U_{0,1}^f+\delta U_{0,2}^f). \end{aligned} \tag{75}$$

(a) As a first application we consider $V_0^f=V_0^g=0$. From Eq. (75) we get $(\mu=(\alpha r+\beta)(\gamma r+\delta),\ \lambda=\alpha(\gamma r+\delta))$. For instance, as particular examples we get $(\mu=1,\ \lambda=0)$, $(\mu=r,\ \lambda=0)$, $(\mu=r,\ \lambda=0)$, $(\mu=r^2,\lambda=r),\cdots$.

(b) As a second application we consider $V_0^f=2r^{-2}$, $V_0^g=0$. The general solution given by Eq. (75) is $(\mu=(\alpha r+\beta)(\gamma r^2+\delta r^{-1}),\ \lambda=\alpha(\gamma r^2+\delta r^{-1}))$, and as particular examples, we give: $(\mu=1,\lambda=r^{-1})$, $(\mu=r^3,\lambda=r^2)$, $(\mu=r^2,\lambda=0)$, $(\mu=r^{-1},\lambda=0),\cdots$. In all these cases we could construct the corresponding kernels (f,g) and get the solutions given by Theorem I.

B. Examples: explicit kernels (f, g) in two cases where $\mu\neq 1$ and $\mu\neq r$

(a) $(\mu=r^2,\lambda=0)$ which, as we have seen, is a particular solution of $V_0^g=0$, $V_0^f=2r^{-2}$. The coupled kernels

$$f=\int (1+\nu/x)\exp(-\nu/x)(1+\nu/y)\exp(-\nu/y)C(\nu)\,d\nu,\quad \nu>0, \tag{76}$$

$$g=-\int \exp(-\nu[x^{-1}+y^{-1}])(xy)^{-1}\nu^2C(\nu)\,d\nu,$$

satisfy the conditions of Theorem I. It follows that their determinants defined on $[0,r]$ verify:

$$\Delta^g\left(U_0^g\frac{D_f}{D_g}\right)=0\ \text{ or }\ \left(\frac{\partial^2}{\partial r^2}-V^g\right)\left(U_0^g\frac{D_f}{D_g}\right)=0,$$

$$U_0^g=1,\quad V^g=-2r^{-2}\frac{\partial}{\partial r}\left(r^2\frac{D_g'}{D_g}\right),\quad \Delta^f\left(U_0^f\frac{D_g}{D_f}\right)=0,\quad U_0^f=r^2,$$

$$V^f=2r^{-2}-2r^{-2}\frac{\partial}{\partial r}\left(r^2\frac{D_f'}{D_f}\right).$$

(b) $(\mu=r^{n+1},\ \lambda=[(n+1)/2]r^n)$ which is a pair of solutions corresponding to $V_0^f=V_0^g=(n^2-1)r^{-2}$. The coupled kernels

$$f=\int \exp[-\nu(x^{-n}+y^{-n})](xy)^{-(n+1)/2}D(\nu)\,d\nu,\quad \nu>0,\ n>0, \tag{77}$$

$$g=-\int \exp[-\nu(x^{-n}+y^{-n})](xy)^{-(n+1)/2}D(\nu)\,d\nu,$$

satisfy the conditions of Theorem I, consequently their determinants defined on $[0,r]$ verify

$$\Delta^g\left(U_0^g\frac{D_f}{D_g}\right)=0,\quad \Delta^f\left(U_0^f\frac{D_g}{D_f}\right)=0,\quad U_0^g=U_0^f=r^{(n+1)/2},$$

$$V^g=V_0^g-2r^{-n-1}\frac{\partial}{\partial r}\left(r^{n+1}\frac{D_g'}{D_g}\right),\quad V^f=V_0^f-2r^{-n-1}\frac{\partial}{\partial r}\left(r^{n+1}\frac{D_f'}{D_f}\right).$$

VIII. CONCLUSION

Our main result is a mathematical theorem (Theorem I) which connects two arbitrary second order differential operators $(\partial^2/\partial r^2 - V_0^g(r))$ and $(\partial^2/\partial r^2 - V_0^f(r))$ with second order differential equations for two coupled Fredholm determinants in the following way. First we take two eigenfunctions (u_0^g and u_0^f with eigenvalues zero) of these differential operators and with them associate a pair of functions ($\mu = u_0^f u_0^g$, $\lambda = u_0^f(g/\partial r)u_0^g$). Secondly, we construct a pair of symmetric kernels (f, g) satisfying both coupled first order partial differential relations associated with (μ, λ) and well-defined boundary conditions. Thirdly, we associate with these kernels their corresponding Fredholm determinants D_g, D_f. The final result is that $u_0^g D_f D_g^{-1}$ or $u_0^f D_g D_f^{-1}$ are solutions of a Schrödinger-inversionlike equation for the potentials

$$V^g = V_0^g - 2\mu^{-1}\frac{\partial}{\partial r}(\mu D_g^{-1} D_g'), \quad V^f = V_0^f - 2\mu^{-1}\frac{\partial}{\partial r}(\mu D_f^{-1} D_f').$$

We have focused our attention on the interpretation of this theorem in the inversion formalism although it has certainly a larger domain of application. In this way we have seen that the classical inversion formalism at fixed l, where the potential is k independent (or at fixed k with l-independent potentials) can be generalized in order to include k-dependent (or l dependent) potentials.

Let us remark that for an lth partial wave, the problem of the determination of a unique, local, k-independent potential, pioneered twenty-five years ago by Gel'fand—Levitan although clearly a respectable academic mathematical problem (which recently appears as a useful tool in the problem of nonlinear partial differential equations) nevertheless has had very few practicable applications in nuclear physics. However, it is now phenomenologically and theoretically widely accepted that such two-body forces are either nonlocal or k dependent.[11]

Parameter dependent potentials are in general less constrained than the corresponding independent-parameter ones, however, we could perhaps take advantage of this arbitrariness in order to introduce *more observables* into the problem. For instance, the scattering data kernels g_l (or g_k) generating k-dependent (or l-dependent) potentials are of the type

$$g_l = \int \hat{j}_l(\sqrt{k^2 + k'^2}x)\hat{j}_l(\sqrt{k^2 + k'^2}y) C_l(k, k')\, dk'$$

[or $g_k = \int \hat{j}_{\chi(\nu)}(kx)\hat{j}_{\chi(\nu)}(ky) C_k(\nu, l)\, d\nu$]. The arbitrariness comes from the fact that instead of obtaining at fixed l a function of *one variable* $C(0, k')$ [or $C_k(\nu, 0)$ at fixed k] to be linked to a function of one variable $\delta_l(k')$ [or a set $\{\delta_l(k)\}$ at fixed k] as in the conventional inversion formalism in Ref. 1 (or in Ref. 6), we have at our disposal a function of *two variables* $C_l(k, k')$ [or $C_k(\nu, l)$]. A very rough counting argument would suggest that it is perhaps more convenient to introduce as data a function of two variables $\delta(l, k)$ in order to try to determine a function of two variables $C(k, k')$, l independent [or $C(\nu, l)$ k independent]. Practically we could try to see if similar C_l (l independent or weakly l dependent) could reproduce two or more phase shifts $\delta_l(k)$. In the same way, perhaps similar C_k (k independent or weakly k dependent) could reproduce the sets $\{\delta_l\}$ for two or more k values; for instance one small and another large. Let us remark that even the same $C_l(k, k')$, l independent [or the same $C_k(\nu, l)$, k independent] does not lead to the same potential because, in the kernel generating the potential, the weight function $\hat{j}_l((k^2 + k'^2)^{1/2}x)$ [or $\hat{j}_{\chi(\nu)}(kx)$] are different for different l values (or k values).

Let us add two final remarks. From the potential theory point of view one may remark that outside the local potential context, very few results have been established, and thus this work could be the starting point for future theoretical investigations of k-dependent potentials, for instance, the general construction of the scattering data kernel from the physical data. In a recent paper[9] it was emphasized that there exists a great unity in the inversion framework, and that inversion at fixed l or at fixed k are two different investigations of the same theory. The fact that our Theorem I can be applied in both cases is also an illustration of this point of view.

ACKNOWLEDGMENTS

I am very indebted to Professor F. J. Dyson for his question about the significance of Theorem I when the energy is different of zero. I also wish to thank Drs. R. Balian, H. Fried, M. Gaudin, M. L. Mehta, and P. Moussa for discussions.

APPENDIX A

For the reader not familiarized with inversion formalism we give a brief sketch of the proofs of Properties I and II.

Property I: We apply $\Delta_0(r) - \Delta_0(\sigma)$ to both sides of Eq. (7), with Δ_0 given in Eq. (4), and take into account Eq. (5),

$$(\Delta_0(r) - \Delta_0(\sigma))[\bar{K}(r, \sigma) + \int_r^a \mu^{-2}(t)\bar{g}(t, \sigma)\bar{K}(r, t)\, dt] = 0. \tag{A1}$$

It is easy to get by differentiation and integration,

$$\Delta_0(r)\int_r^a \bar{g}\bar{K}\mu^{-2}\, dt = \int_r^a \bar{g}\bar{\mu}^{-2}\Delta_0(r)\bar{K}(r, t)\, dt + A, \tag{A2}$$
$$A = -\mu^2(r)\frac{d}{dr}\left(\frac{\bar{g}(r, \sigma)\bar{K}(r, r)}{\mu^2(r)}\right) - \mu(r)\bar{g}(r, \sigma)\left(\frac{\partial}{\partial r}\frac{\bar{K}(r, t)}{\mu(r)}\right)_{t=r}.$$

From $\int \bar{K}(r, t)\mu^{-2}(t)(\Delta_0(\sigma) - \Delta_0(t))\bar{g}(t, \sigma)\, dt = 0$ and the boundary condition given by Eq. (8), we get

$$\Delta_0(\sigma)\int_r^a \bar{g}\bar{K}\mu^{-2}\, dt = \int_r^a \bar{g}\mu^{-2}\Delta_0(r)\bar{K}(r, t)\, dt + B, \tag{A3}$$
$$B = -\bar{K}(r, r)\left(\frac{\partial}{\partial t}\bar{g}(t, \sigma)\right)_{t=r} + \bar{g}(r, \sigma)\left(\frac{\partial}{\partial t}\bar{K}(r, t)\right)_{t=r}.$$

It follows that (A1) can be written

$$(\Delta_0(r) - \Delta_0(\sigma))(\bar{K}(r, \sigma)) + \int_r^a \bar{g}\mu^{-2}(\Delta_0(r) - \Delta_0(t))\bar{K}(r, t)\, dt = B - A,$$

$$B - A = 2\mu(r)\bar{g}(r, \sigma)\frac{d}{dr}\left(\frac{\bar{K}(r, r)}{\mu(r)}\right),$$

or using the definition of Δ given in Eq. (9),

$$(\Delta(r) - \Delta_0(\sigma))(\bar{K}(r, \sigma)) + \int_r^a \bar{g}(t, \sigma)\mu^{-2}(t)(\Delta(r) - \Delta_0(t)\bar{K}(r, t)\, dt = 0.$$

Since homogeneous integral equation of Eq. (7) has only the trivial solution, we get the result of Property I,

$$(\Delta(r) - \Delta_0(\sigma))\bar{K}(r,\sigma) = 0. \tag{A4}$$

Property II: We apply $\Delta(r)$ to both sides of Eq. (12),

$$\Delta U_n = U_n^0\left[\gamma_n + 2\mu\frac{d}{dr}\left(\frac{\bar{K}(r,r)}{\mu(r)}\right)\right] + \mu^2\left(\frac{\partial^2}{\partial r^2} - V\right) \times \int_r^a U_n^0\mu^{-2}\bar{K}(r,t)\,dt, \tag{A5}$$

and we differentiate the last term of the rhs of (A5) and get

$$\Delta U_n = D + \int_r^a U_n^0(t)\mu^{-2}(t)\Delta(r)\bar{K}(r,t)\,dt,$$
$$D = U_n^0\left\{\gamma_n + \mu\frac{d}{dr}\left(\frac{\bar{K}(r,r)}{\mu(r)}\right) - \mu\left(\frac{\partial}{\partial r}\left(\frac{\bar{K}(r,t)}{\mu(t)}\right)\right)_{t=r} + \bar{K}(r,r)\mu^{-1}\frac{d\mu}{dr}\right\} - \bar{K}(r,r)\frac{dU_n^0}{dr}. \tag{A6}$$

From the identity $0 = \int_r^a \bar{K}(r,t)\mu^{-2}(t)(\Delta_0(t) - \gamma_n)U_n^0(t)\,dt$ and the boundary condition (13) we get

$$0 = \int_r^a U_n^0\mu^{-2}(\Delta_0(t) - \gamma_n)\bar{K}(r,t)\,dt + C,$$
$$C = -\bar{K}(r,r)\frac{dU_n^0}{dr} + U_n^0(r)\left(\frac{\partial\bar{K}}{\partial t}(r,t)\right)_{t=r}. \tag{A7}$$

It follows that (A6) can be written as

$$\Delta U_n = \int_r^a U_n^0\mu^{-2}(\Delta(r) - \Delta_0(t) + \gamma_n)\bar{K}(r,t)\,dt + D - C,$$

with $D - C = \gamma_n U_n^0$. If we apply property I we get

$$\Delta U_n = \gamma_n\{U_n^0 + \int_r^a U_n^0\mu^{-2}\bar{K}\,dt\} = \gamma_n U_n,$$

which is property II.

APPENDIX B

We want to investigate the behavior of the kernel g_l^+ when k is small, so we consider

$$g_l^+(k;x,y) = \int \phi_l^+(k,k_1,x)\phi_l^+(k,k_1,y)C_l(k,k_1)\,dk_1 \tag{B1}$$

with

$$\phi_l^+(k,k_1,x) = \left(\frac{\partial}{\partial x} + \lambda^+(kx)\right)\hat{j}_l(\sqrt{\gamma}x), \quad \gamma = k^2 + k_1^2,$$
$$\lambda^+ = -\left(\frac{\partial}{\partial x}\hat{h}_l^+(kx)\right)(\hat{h}_l^+(kx))^{-1}. \tag{B2}$$

Taking into account the recurrence relation $(\partial/\partial x + l/x) \times \hat{h}_l^+(kx) = k\hat{h}_{l-1}^+(kx)$ we get

$$\lambda^+ = \frac{l}{x} - k\frac{\hat{h}_{l-1}(kx)}{\hat{h}_l(kx)}. \tag{B3}$$

Writing $\hat{h}_l(\rho) = \hat{n}_l(\rho) + i\hat{j}_l(\rho)$, and taking into account

$$\lim_{\rho\to 0}\rho^l\hat{n}_l(\rho)[(2l-1)!!]^{-1} = 1, \quad \lim_{\rho\to 0}(2l+1)!!\,\hat{j}_l(\rho)\rho^{-l-1} = 1, \tag{B4}$$

$$\frac{\hat{h}_{l-1}(kx)}{\hat{h}_l(kx)} \underset{(kx)\to 0}{\approx} kx(2l-1)^{-1} + i(kx)^{2l}[(2l-1)!!]^{-2},$$

and substituting (B3) and (B4) in (B2), we get for small k

$$\mathrm{Re}\phi_l^+(k,k_1,x) \underset{k\to 0}{\approx} k_1\hat{j}_{l-1}(k_1x)$$
$$\mathrm{Im}\phi_l^+(k,k_1,x) \underset{k\to 0}{\approx} -\frac{k^{l+1}}{[(2l-1)!!]^2}x^{2l}\hat{j}_l(k_1x). \tag{B5}$$

Finally, for small k, the kernel g_l^+ has the behavior

$$\mathrm{Re}g_l^+(k;x,y) \underset{k\to 0}{\approx} \int k_1^2\hat{j}_{l-1}(k_1x)\hat{j}_{l-1}(k_1x)C_l(0,k_1)\,dk_1, \tag{B6}$$

$$\mathrm{Im}g_l^+(k;x,y) \underset{k\to 0}{\approx} -\frac{k^{2l+1}}{[(2l-1)!!]^2} \times x^{2l}\int k_1\hat{j}_l(k_1x)\hat{j}_{l-1}(k_1y) + \mathrm{idem}\,x\to y\ .$$

[1] I. M. Gel'fand and B. M. Levitan, Am. Math. Soc. Trans. Ser. 2 **1**, 253 (1955); V. A. Marchenko, Dokl. Akad. Nauk SSSR **72**, 457 (1950). For books and review papers see Z. S. Agranovich and V. A. Marchenko, *The Inverse Problem of Scattering Theory* (Gordon and Breach, New York, 1963); R. G. Newton, J. Math. Phys. **1**, 319 (1960); L. D. Faddeyev, J. Math. Phys. **4**, 72 (1963). V. De Alfaro and T. Regge, "Potential Scattering" (Amsterdam, 1965).

[2] H. Cornille, Preprint Orsay IPNO TH 78 (1966); J. Math. Phys. **11**, 615 (1970).

[3] M. Gaudin (unpublished), the proof is quoted in M. L. Mehta, *Random Matrices and Statistical Theory of Energy Levels* (Academic, New York, 1967), see Appendix A31.

[4] H. Cornille and G. Rubinstein, Nuovo Cimento A **56**, 867 (1968), Sec. 3.1; the pair $(B = r[l(l+1)/2]D_l$ and $A = r[l(l-1)/2]D_{l-1})$ satisfy (Eq. 1) of the present paper as well as the pair

$$B = r\frac{(l+1)(l+2)}{2}D_{l+1} \quad \text{and} \quad A = r\frac{l(l+1)}{2}D_l\ .$$

[5] F. J. Dyson, Commun. Math. Phys. **47**, 171 (1976).

[6] For particular l-dependent potentials see: P. C. Sabatier, J. Math. Phys. **9**, 1241 (1968); M. A. Hooshyar, J. Math. Phys. **12**, 2143 (1971); **13**, 1931 (1972); **16**, 157 (1975), and for particular k-dependent potentials, see H. Cornille, J. Math. Phys. **11**, 79 (1970); M. Jaulent and C. Jean, Commun. Math. Phys. **28**, 172 (1972) and the note added in proof therein, p. 219; M. Jaulent, Ann. Inst. Henri Poincaré **17**, 363 (1972).

[7] T. Regge, Nuovo Cimento **14**, 951 (1959); R. G. Newton, J. Math. Phys. **3**, 75 (1962).

[8] A. Messiah, *Mécanique quantique* (Dunod, Paris, 1959), Appendix B, p. 475—the spherical Bessel functions can be also defined for real noninteger l values.

[9] H. Cornille, Nuovo Cimento A **14**, 141 (1973).

[10] R. G. Newton, Ref. 6, p. 80 (IV examples).

[11] R. Vinh-Mau, "Topics in the Nucleon—Nucleon Interaction," Preprint Orsay IPNO TH 76/05.

Uniqueness of perturbation of a Reissner–Nordström black hole

Chul Hoon Lee and W. D. McGlinn

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 20 May 1976)

Coupled gravitational and electromagnetic perturbations of a Reissner-Nordström black hole are analyzed using the Newman-Penrose formalism. It is shown that $\chi_1^B(\equiv 3\psi_2\phi_0^B - 2\phi_1\psi_1^B)$ or $\chi_{-1}^B(\equiv 3\psi_2\phi_2^B - 2\phi_1\psi_3^B)$ determines the perturbations except for those corresponding to an infinitesimal change in the mass, charge, and angular momentum parameters of the black hole.

I. INTRODUCTION

Some time ago, Wald[1] analyzed gravitational perturbations of a Kerr black hole using the Newman–Penrose (NP) formalism and proved that either of the perturbations ψ_0^B [2] or ψ_4^B alone uniquely specifies the nontrivial part of a gravitational perturbation of a Kerr black hole and that with ψ_0^B $(\psi_4^B) = 0$ the only well-behaved solutions of the perturbation equations are those corresponding to a change in the mass and angular momentum parameters.

Since it appears that a black hole can possess a net charge,[3,4] it is of interest to extend the problem to the charged case. In analyzing perturbations of a charged black hole one cannot treat the gravitational and the electromagnetic perturbations separately because one gives rise to the other through a nonzero first order perturbation of the stress-energy tensor. Hence one has to solve for the tetrad components of the Weyl tensor $(\psi_0, \psi_1, \psi_2, \psi_3, \psi_4)$ and of the electromagnetic field tensor (ϕ_0, ϕ_1, ϕ_2) simultaneously.

In Ref. 5 coupled gravitational and electromagnetic perturbation equations connecting ψ_0^B (ψ_4^B) to χ_1^B (χ_{-1}^B) are derived. The particular combinations $\chi_1^B \equiv 3\psi_2\phi_0^B - 2\phi_1\psi_1^B$ and $\chi_{-1}^B \equiv 3\psi_2\phi_2^B - 2\phi_1\psi_3^B$ occur in the equations as they are invariant under infinitesimal tetrad transformations while ϕ_0^B, ϕ_2^B, ψ_1^B, and ψ_3^B are not (see the Appendix).

The ingoing electromagnetic and gravitational radiations at infinity can be calculated immediately from ψ_0^B and χ_1^B alone and the outgoing radiations from ψ_4^B and χ_{-1}^B alone.

When restricted to the nonrotating case (Reissner–Nordström black hole) the coupled equations separate and the problem reduces to solving ordinary differential equations. In this paper we restrict ourselves to discussing perturbations of a nonrotating black hole. In Sec. II, the coupled equations of Ref. 5 are written in a slightly different form because of a different choice of the coordinate system. We use these equations to show that χ_1^B (χ_{-1}^B) uniquely determines ψ_0^B (ψ_4^B).

In Sec. III, it is shown that $\psi_0^B = \chi_1^B = 0$ implies $\chi_{-1}^B = \psi_4^B = 0$. Finally in Sec. IV, we show that with $\psi_0^B = \chi_1^B = \psi_4^B = \chi_{-1}^B = 0$ the only well-behaved solutions for ψ_2^B and ϕ_1^B are the ones corresponding to an infinitesimal change in the mass and charge parameters and an addition of small angular momentum to the black hole. This result is discussed in Sec. V.

In the null coordinate the Reissner–Nordström metric is

$$ds^2 = \left(1 - \frac{2Mr - Q^2}{r^2}\right)du^2 + 2\,du\,dr - r^2\,d\theta^2 - r^2\sin^2\theta\,d\phi^2. \tag{1.1}$$

We use the tetrad whose components are

$$l^\mu = \delta_2^\mu,$$
$$n^\mu = \delta_1^\mu - \frac{\Delta}{2r^2}\delta_2^\mu, \quad (\mu = 1,2,3,4) \tag{1.2}$$
$$m^\mu = \frac{1}{\sqrt{2}r}\left(\delta_3^\mu + \frac{i}{\sin\theta}\,\delta_4^\mu\right),$$

where $\Delta = r^2 - 2Mr + Q^2$.

The resulting tetrad components of the Weyl tensor are[6]

$$\psi_0 = \psi_1 = \psi_3 = \psi_4 = 0 \quad \psi_2 = \frac{-Mr + Q^2}{r^4} \tag{1.3}$$

and of the electromagnetic field tensor are

$$\phi_0 = \phi_2 = 0, \quad \phi_1 = \frac{Q}{2r^2}. \tag{1.4}$$

The spin coefficients are

$$\kappa = \sigma = \lambda = \nu = \epsilon = \tau = \pi = 0,$$
$$\rho = -\frac{1}{r}, \quad \beta = \frac{\cot\theta}{2\sqrt{2}r}, \quad \alpha = -\frac{\cot\theta}{2\sqrt{2}r}, \tag{1.5}$$
$$\mu = -\frac{\Delta}{2r^3}, \quad \gamma = -\frac{\Delta}{2r^3} + \frac{r - M}{2r^2}.$$

II. COUPLED EQUATIONS FOR ψ_0^B AND χ_1^B (ψ_4^B AND χ_{-1}^B)

Coupled perturbation equations connecting ψ_0^B with χ_1^B and ψ_4^B with χ_{-1}^B are derived in Ref. 5 for the Kerr–Newman background metric. As is shown there, these equations can be separated for the nonrotating case by writing

$$\psi_0^B = \exp(-i\omega u)\exp(im\phi)\,{}_2Y_l^m(\theta)R_l^{(2)}(r),$$
$$\chi_1^B = \exp(-i\omega u)\exp(im\phi)\,{}_1Y_l^m(\theta)R_l^{(1)}(r), \tag{2.1}$$
$$\chi_{-1}^B = \exp(-i\omega u)\exp(im\phi)\,{}_{-1}Y_l^m(\theta)\frac{\Delta}{2r^2}\exp(-2i\omega r')R_l^{(-1)}(r),$$
$$\psi_4^B = \exp(-i\omega u)\exp(im\phi)\,{}_{-2}Y_l^m(\theta)\frac{\Delta^2}{4r^4}\exp(-2i\omega r')R_l^{(-2)}(r),$$

where the angular functions, ${}_sY_l^m(\theta)$, are the spin weighted harmonics and $dr' = (r^2/\Delta)\,dr$. The equations for the radial functions are

$$\left[\Delta\frac{d^2}{dr^2} + \left(\pm 2i\omega r^2 + 6(r-m) - \frac{4Q^2\Delta}{r(3Mr-4Q^2)}\right)\frac{d}{dr} \pm i\omega\left(10r - \frac{8Q^2r}{3Mr-4Q^2}\right) + r + \frac{2Q^2}{r^2} - \frac{4Q^2(r^2+2Mr-3Q^2)}{r^2(3Mr-rQ^2)} - \frac{3Mr-4Q^2}{3Mr-2Q^2}(l-1)(l+2)\right]R_l^{(\pm 2)} = \frac{-2\sqrt{2}Q\sqrt{(l-1)(l+2)}r^3}{3Mr-2Q^2}\left(\frac{d}{dr} + \frac{4}{r} - \frac{4Q^2}{r(3Mr-4Q^2)}\right) \times R_l^{(\pm 1)}, \tag{2.2a}$$

$$\left[\Delta\frac{d^2}{dr^2} + \left(\pm 2i\omega r^2 + r(r-M) + \frac{6\Delta}{r} - \frac{2Q^2\Delta}{r(3Mr-2Q^2)}\right)\frac{d}{dr} \pm 12i\omega r + \frac{18r^2-24Mr+2Q^2}{r^2} - \frac{12Q^2\Delta}{r^2(3Mr-2Q^2)} - \frac{3Mr-2Q^2}{3Mr-4Q^2}(l-1)(l+2)\right]R_l^{(\pm 1)} = \frac{\sqrt{2}Q^3\sqrt{(l-1)(l+2)}\Delta}{r^3(3Mr-4Q^2)}\left(\frac{d}{dr} \pm 2i\omega\frac{r^2}{\Delta} - \frac{2}{r} + \frac{4(r-M)}{\Delta} - \frac{2Q^2}{r(3Mr-2Q^2)}\right)R_l^{(\pm 2)}. \tag{2.2b}$$

These equations are slightly different from the radial equations in Ref. (5) due to different choice of the coordinate system in this paper ($X^1 = u = t - r'$). Notice that $l \geq 2$ for $R_l^{(\pm 2)}$ and $l \geq 1$ for $R_l^{(\pm 1)}$. For $l=1$, $R_l^{(\pm 2)}$ is automatically zero so that Eq. (2.2a) gives $0=0$ and Eq. (2.2b) gives

$$\left[\Delta\frac{d^2}{dr^2} + \left(\pm 2i\omega r^2 + r(r-M) + \frac{6\Delta}{r} - \frac{2Q^2\Delta}{r(3Mr-2Q^2)}\right)\frac{d}{dr} \pm 12i\omega r + \frac{18r^2-24Mr+2Q^2}{r^2} - \frac{12Q^2\Delta}{r^2(3Mr-2Q^2)}\right]R^{(\pm 1)} = 0. \tag{2.3}$$

When $R_l^{(\pm 1)} = 0$, Eq. (2.2a) can be rewritten as

$$\left[\left(\Delta\frac{d}{dr} + \frac{2\Delta}{r} + 2(r-M) + \frac{2Q^2\Delta}{r(3Mr-2Q^2)} - \frac{2Q^2\Delta}{r(3Mr-4Q^2)}\right) \times\left(\frac{d}{dr} \pm 2i\omega\frac{r^2}{\Delta} - \frac{2}{r} + \frac{4(r-M)}{\Delta} - \frac{2Q^2}{r(3Mr-2Q^2)}\right) + A\right] \times R_l^{(\pm 2)} = 0, \tag{2.4}$$

where

$$A = \pm i\omega\left(2r - \frac{4Q^2r}{3Mr-2Q^2}\right) - 2 - \frac{3Mr-4Q^2}{3Mr-2Q^2} \times (l-1)(l+2) + \frac{4Q^2}{3Mr-2Q^2}.$$

Eq. (2.2b) becomes

$$\left(\frac{d}{dr} \pm 2i\omega\frac{r^2}{\Delta} - \frac{2}{r} + \frac{4(r-M)}{\Delta} - \frac{2Q^2}{r(3Mr-2Q^2)}\right)R_l^{(\pm 2)} = 0. \tag{2.5}$$

From Eqs. (2.4) and (2.5)

$$AR_l^{(\pm 2)} = 0 \tag{2.6}$$

which yields $R_l^{(\pm 2)} = 0$ as A is not identically zero. This concludes the proof that χ_1^B (χ_{-1}^B) uniquely determines ψ_0^B (ψ_4^B). The converse of the above statement is not true because with ψ_0^B (ψ_4^B) $= 0$ there still exists a solution for χ_1^B (χ_{-1}^B)

$$\chi_1^B = \exp(-i\omega u)\exp(im\phi)\,{}_1Y_1^m(\theta)R_1^{(1)}(r),$$

$$\chi_{-1}^B = \exp(-i\omega u)\exp(im\phi)\,{}_{-1}Y_1^m(\theta)\frac{\Delta}{2r^2} \times\exp(-2i\omega r')R_1^{(-1)}(r), \tag{2.7}$$

with $R_1^{(\pm 1)}$ satisfying Eq. (2.3).

This solution for $R_1^{(+1)}$ with ω nonzero and real corresponds to ingoing dipole electromagnetic radiation at infinity with no ingoing gravitational radiation. It can be shown from the connection relations of Ref. (5) [Eqs. (3.11) and (3.12)] that the resulting outgoing radiation for this solution is pure electromagnetic.

For $l \geq 2$, one can get a completely decoupled equation for $R_l^{(\pm 1)}$ ($R_l^{(\pm 2)}$) by applying a first order differential operator on Eq. (2.2a) [Eq. (2.2b)] and a second order differential operator on Eq. (2.2b) [Eq. (2.2a)] and thus eliminating $R_l^{(\pm 2)}$ ($R_l^{(\pm 1)}$) by subtraction.[5] The resulting equation is a fourth order differential equation whose solution has four degrees of freedom. The asymptotic solutions at infinity ($r \to \infty$) with $\omega \neq 0$ are

$$R_l^{(\pm 2)} \sim \frac{\exp(\mp 2i\omega r')}{r}, \frac{\exp(\mp 2i\omega r')}{r^3}, \frac{1}{r^5}, \frac{1}{r^6},$$

$$R_l^{(\pm 1)} \sim \frac{\exp(\pm 2i\omega r')}{r^4}, \frac{\exp(\pm 2i\omega r')}{r^5}, \frac{1}{r^6}, \frac{1}{r^8}.$$

III. PERTURBATIONS WITH $\psi_0^B = \chi_1^B = 0$

In this section we prove that $\chi_{-1}^B = \psi_4^B = 0$ if $\psi_0^B = \chi_1^B = 0$. First we start with the case where $\omega \neq 0$. One group of quantities (ψ_0^B, χ_1^B) is related to the other (ψ_4^B, χ_{-1}^B) through Bianchi identities. Since the equations connecting them involve the complex conjugate of some quantities we take the following forms for ψ_0^B, χ_1^B, χ_{-1}^B, and ψ_4^B for a given ω:

$$\psi_0^B = \exp(-i\omega u)\exp(im\phi)\,{}_2Y_l^mR_l^{(2)} - \exp(i\omega^* u)\exp(-im\phi)\,{}_2\bar{Y}_l^mP_l^{(2)},$$

$$\chi_1^B = \exp(-i\omega u)\exp(im\phi)\,{}_1Y_l^mR_l^{(1)} - \exp(i\omega^* u)\exp(-im\phi)\,{}_1\bar{Y}_l^mP_l^{(1)},$$

$$\chi_{-1}^B = \exp(-i\omega u)\exp(im\phi)\,{}_{-1}Y_l^m\frac{\Delta}{2r^2}\exp(-2i\omega r')R_l^{(-1)} \tag{3.1}$$

$$-\exp(i\omega^* u)\exp(-im\phi)\,{}_{-1}Y_l^{\sim m}\frac{\Delta}{2r^2}\exp(2i\omega^* r')P_l^{(-1)},$$

$$\psi_4^B=\exp(-i\omega u)\exp(im\phi)\,{}_{-2}Y_l^m\frac{\Delta^2}{4r^4}\exp(-2i\omega r')R_l^{(-2)}$$

$$-\exp(i\omega^* u)\exp(-im\phi)\,{}_{-2}Y_l^{\sim m}\frac{\Delta^2}{4r^4}\exp(2i\omega^* r')P_l^{(-2)},$$

where ω is any complex constant and $P_l^{(\pm1)*}$ and $P_l^{(\pm2)*}$ satisfy the same equations as $R_l^{(\pm1)}$ and $R_l^{(\pm2)}$ do.

Using the asymptotic solutions Eq. (2.8), one gets the following forms at infinity:

$$\begin{aligned}
\psi_0^B&=\exp(-i\omega u)\exp(im\phi)\,{}_2Y_l^m\left[A^{(2)}\exp(-2i\omega r')\left(\frac{1}{r}+\cdots\right)+B^{(2)}\left(\frac{1}{r^5}+\cdots\right)\right]\\
&\quad-\exp(+i\omega^* u)\exp(-im\phi)\,{}_2Y_l^{\sim m}\left[C^{(2)}\exp(2i\omega^* r')\left(\frac{1}{r}+\cdots\right)+D^{(2)}\left(\frac{1}{r^5}+\cdots\right)\right],\\
\chi_1^B&=\exp(-i\omega u)\exp(im\phi)\,{}_1Y_l^m\left[A_l^{(1)}\exp(-2i\omega r')\left(\frac{1}{r^4}+\cdots\right)+B_l^{(1)}\left(\frac{1}{r^6}+\cdots\right)\right]\\
&\quad-\exp(i\omega^* u)\exp(-im\phi)\,{}_1Y_l^{\sim m}\left[C_l^{(1)}\exp(2i\omega^* r')\left(\frac{1}{r^4}+\cdots\right)+D_l^{(1)}\left(\frac{1}{r^6}+\cdots\right)\right],\\
\chi_{-1}^B&=\exp(-i\omega u)\exp(im\phi)\,{}_{-1}Y_l^m\left[\frac{A_l^{(-1)}}{2}\left(\frac{1}{r^4}+\cdots\right)+\frac{B_l^{(-1)}}{2}\exp(-2i\omega^* r')\left(\frac{1}{r^6}+\cdots\right)\right]\\
&\quad-\exp(i\omega^* u)\exp(-im\phi)\,{}_{-1}Y_l^{\sim m}\left[\frac{C_l^{(-1)}}{2}\left(\frac{1}{r^4}+\cdots\right)+\frac{D_l^{(-1)}}{2}\exp(2i\omega^* r')\left(\frac{1}{r^6}+\cdots\right)\right],\\
\psi_4^B&=\exp(-i\omega u)\exp(im\phi)\,{}_{-2}Y_l^m\left[\frac{A_l^{(-2)}}{4}\left(\frac{1}{r}+\cdots\right)+\frac{B_l^{(-2)}}{4}\exp(-2i\omega r')\left(\frac{1}{r^5}+\cdots\right)\right]\\
&\quad-\exp(i\omega^* u)\exp(-im\phi)\,{}_{-2}Y_l^{\sim m}\left[\frac{C_l^{(-2)}}{4}\left(\frac{1}{r}+\cdots\right)+\frac{D_l^{(-2)}}{4}\exp(2i\omega^* r')\left(\frac{1}{r^5}+\cdots\right)\right].
\end{aligned}\tag{3.2}$$

The connections between the coefficients for a real ω are given in Ref. 5. They can be easily generalized to a complex ω case and the results are

$$\begin{pmatrix}
-\frac{l(l+1)}{4\omega^2} & \frac{Q^2}{3i\omega M} & 0 & \frac{\sqrt{2}Q^3\sqrt{(l-1)(l+2)}}{12\omega^2M}\\
\frac{Q^2}{3i\omega M} & -\frac{l(l+1)}{4\omega^2} & \frac{\sqrt{2}Q^3\sqrt{(l-1)(l+2)}}{12\omega^2M} & 0\\
0 & -\frac{\sqrt{2}Q\sqrt{(l-1)(l+2)}}{6\omega^2M} & \frac{(l-1)l(l+1)(l+2)}{16\omega^4} & \frac{3M}{4i\omega^3}+\frac{Q^2(l+1)(l+2)}{12i\omega^3M}\\
-\frac{\sqrt{2}Q\sqrt{(l-1)(l+2)}}{6\omega^2M} & 0 & \frac{3M}{4i\omega^3}+\frac{Q^2(l-1)(l+2)}{12i\omega^3M} & \frac{(l-1)l(l+1)(l+2)}{16\omega^4}
\end{pmatrix}\begin{pmatrix}A_l^{(1)}\\ C_l^{(1)*}\\ A_l^{(2)}\\ C_l^{(2)*}\end{pmatrix}=\begin{pmatrix}B_l^{(-1)}\\ D_l^{(-1)*}\\ B_l^{(-2)}\\ D_l^{(-2)*}\end{pmatrix},\tag{3.3a}$$

$$\begin{pmatrix}
-\frac{l(l+1)}{4\omega^2} & -\frac{Q^2}{3i\omega M} & 0 & \frac{\sqrt{2}Q^3\sqrt{(l-1)(l+2)}}{12\omega^2M}\\
-\frac{Q^2}{3i\omega M} & -\frac{l(l+1)}{4\omega^2} & \frac{\sqrt{2}Q^3\sqrt{(l-1)(l+2)}}{12\omega^2M} & 0\\
0 & -\frac{\sqrt{2}Q\sqrt{(l-1)(l+2)}}{6\omega^2M} & \frac{(l-1)l(l+1)(l+2)}{16\omega^4} & -\frac{3M}{4i\omega^3}-\frac{Q^2(l-1)(l+2)}{12i\omega^3M}\\
-\frac{\sqrt{2}Q\sqrt{(l-1)(l+2)}}{6\omega^2M} & 0 & -\frac{3M}{4i\omega^3}-\frac{Q^2(l-1)(l+2)}{12i\omega^3M} & \frac{(l-1)(l+1)(l+2)}{16\omega^4}
\end{pmatrix}\begin{pmatrix}A_l^{(-1)}\\ C_l^{(-1)*}\\ A_l^{(-2)}\\ C_l^{(-2)*}\end{pmatrix}=\begin{pmatrix}B_l^{(1)}\\ D_l^{(1)*}\\ B_l^{(2)}\\ D_l^{(2)*}\end{pmatrix}.\tag{3.3b}$$

These connection equations are also good for $l=1$ with $A_1^{(\pm2)}=B_1^{(\pm2)}=C_1^{(\pm2)}=D_1^{(\pm2)}=0$. When $\psi_0^B=\chi_1^B=0$, $A_l^{(1)}=A_l^{(2)}=B_l^{(1)}=B_l^{(2)}=C_l^{(1)}=C_l^{(2)}=D_l^{(1)}=D_l^{(2)}=0$ and Eq. (3.3a) yields

$$B_l^{(-1)}=B_l^{(-2)}=D_l^{(-1)}=D_l^{(-2)}=0.\tag{3.4}$$

For $A_l^{(-1)}$, $A_l^{(-2)}$, $C_l^{(-1)}$ and $C_l^{(-2)}$ to have any nontrivial solutions the determinant of the matrix in Eq. (3.3b) has to be zero. The condition on ω for the vanishing determinant is

$$\omega=\pm i\frac{3Ml(l+1)}{8Q^2}\left[\left(1+\frac{4Q^2(l-1)(l+2)}{9M^2}\right)^{1/2}\pm1\right].\tag{3.5}$$

To treat the case where $\omega=0$, instead of obtaining the connection equations valid when $\omega=0$, we proceed in the following simpler way. First we set $\phi_0^B=\phi_2^B=0$ without

loss of generality (see Appendix). Then the following perturbation equations of Bianchi identities[7]:

$$(\delta^* + 4\beta)\psi_0^B - (D - 4\rho)\psi_1^B = (3\psi_2 - 2\phi_{11})\kappa^B, \tag{3.6}$$

$$(\Delta - 4\gamma + \mu)\psi_0^B - (\delta - 2\beta)\psi_1^B = (3\psi_2 + 2\phi_{11})\sigma^B, \tag{3.7}$$

yields $\kappa^B = \sigma^B = 0$ when $\psi_0^B = \psi_1^B = 0$. Furthermore, by performing appropriate infinitesimal tetrad and coordinate transformations we set

$$l^{\mu B} = 0 \text{ and } \epsilon^B = 0. \tag{3.8}$$

Imposing these conditions we obtain the following perturbations of Maxwell's equations,[8] NP equations,[7] and Bianchi identities:

$$(D - 2\rho)\phi_1^B = 2\phi_1\rho^B, \tag{3.9a}$$

$$\delta\phi_1^B = \frac{Q}{r^3}\xi^2 + 2\phi_1\tau^B, \tag{3.9b}$$

$$\delta^*\phi_1^B = \frac{Q}{r^3}\xi^{2*} - 2\phi_1\pi^B, \tag{3.9c}$$

$$(\Delta + 2\mu)\phi_1^B = \frac{Q}{r^3}X^2 - \frac{Q}{r^2}\mu^B, \tag{3.9d}$$

$$3(D - 3\rho)\psi_2^B - 4\phi_1(D + \rho)(\phi_1^B + \phi_1^{*B}) = 3(3\psi_2 + 2\phi_{11})\rho^B - 4\phi_{11}(\rho^B + \rho^{*B}), \tag{3.10a}$$

$$3\delta\psi_2^B + 4\phi_1\delta(\phi_1^B + \phi_1^{*B}) = \frac{-9Mr + 16Q^2}{r^5}\xi^2 + 3(3\psi_2 - 2\phi_{11})\tau^B + 4\phi_{11}(\tau^B - \pi^{*B}), \tag{3.10b}$$

$$3\delta^*\psi_2^B - 3(D - 2\rho)\psi_3^B + 4\phi_1\delta^*(\phi_1^B + \phi_1^{*B}) = \frac{-9Mr + 16Q^2}{r^5}\xi^{2*} - 3(3\psi_2 - 2\phi_{11})\pi^B + 4\phi_{11}(\tau^{*B} - \pi^B), \tag{3.10c}$$

$$(D - 2\rho)\rho^B = 0, \tag{3.11}$$

where $X^\mu = n^{\mu B}$ and $\xi^\mu = m^{\mu B}$.

Integration of Eq. (3.11) yields

$$\rho^B = \frac{a^0 + iA^0}{r^2}, \tag{3.12}$$

where a^0 and A^0 are real and independent of r. Henceforth the superscript "0" represents real functions which are independent of r and infinitesimal. We make $a^0 = 0$ by an infinitesimal coordinate transformation $r \to r + a^0(u, \theta, \phi)$ which does not affect Eq. (3.8). Then, integrate Eqs. (3.9a) and (3.10a) to obtain

$$\phi_1^B = -\frac{iQA^0}{r^3} + \frac{b^0 + iB^0}{r^2}, \tag{3.13}$$

$$\psi_2^B = \frac{3Mr - 2Q^2}{r^5}iA^0 + \frac{4Qb^0}{r^4} + \frac{c^0 + iC^0}{r^3}. \tag{3.14}$$

Subtracting the complex conjugate of Eq. (3.9b) from Eq. (3.9c), we obtain

$$\tau^{*B} + \pi^B = i\sqrt{2}\left(\frac{\partial}{\partial\theta} - \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)\left(\frac{A^0}{r^2} - \frac{B^0}{r}\right). \tag{3.15}$$

The equation for ψ_3^B, obtained by subtracting the complex conjugate of Eq. (3.10b) from Eq. (3.10c), is

$$(D - 2\rho)\psi_3^B = \delta^*(\psi_2^B - \psi_2^{*B}) + (3\psi_2 - 2\phi_{11})(\tau^{*B} + \pi^B). \tag{3.16}$$

Now integrate Eq. (3.16) with the use of Eqs. (3.14) and (3.15) to get

$$\psi_3^B = \frac{-r^2}{Q}\chi_{-1}^B = i\sqrt{2}\left(\frac{\partial}{\partial\theta} - \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)\left(\frac{QB^0}{r^4} - \frac{(3M/2)B^0 + C^0}{r^3}\right) + \frac{d^0 + iD^0}{r^2}. \tag{3.17}$$

This result indicates that the radial part of $\psi_3^B = \exp(-i\omega u)\exp(im\phi)\,{}_{-1}Y_l^m(\theta)T_l(r)$ has to be of the form

$$T_l(r) = \frac{-\Delta}{Q}\exp(-2i\omega r')R_l^{(-1)}(r) = \frac{\alpha_1}{r^2} + \frac{\alpha_2}{r^3} + \frac{\alpha_3}{r^4}. \tag{3.18}$$

Notice that the restriction $\omega = 0$ has not been made yet.

When $l = 1$, $T_1(r)$ has to satisfy the following equation obtained from Eq. (2.3):

$$\left[\Delta(3Mr - 2Q^2)\frac{d^2}{dr^2} + \left(2i\omega r^2(3Mr - 2Q^2) + \frac{2\Delta(9Mr - 8Q^2)}{r}\right)\times\frac{d}{dr} + 12i\omega r(Mr - Q^2) + 12Mr - (36M^2 + 16Q^2) + \frac{50MQ^2}{r} - \frac{16Q^4}{r^2}\right]T_1 = 0 \tag{3.19}$$

Putting $T_1(r)$ in the form of Eq. (3.18) into Eq. (3.19) one can see that unless $\omega = \pm i(3M/2Q^2)$, $\alpha_1 = \alpha_2 = \alpha_3 = 0$. When $\omega = i(3M/2Q^2)$, $T_1 = \alpha_1 r^{-2}$ and when $\omega = -i(3M/2Q^2)$, $T_1 = \alpha_1[r^{-2} - (4Q^2/3M)r^{-3} + (4Q^4/9M^2)r^{-4}]$. These conditions on ω are included in Eq. (3.5) obtained by a different procedure.

When $l > 2$ and $\omega = 0$, the decoupled equations for $T_l(r)$, obtained by using the method mentioned in Ref. 5, has the following asymptotic form at infinity ($r \to \infty$):

$$\left(\frac{d^4}{dr^4} + \frac{18}{r}\frac{d^3}{dr^3} + \frac{96 - 2l(l+1)}{r^2}\frac{d^2}{dr^2} + \frac{156 - 14l(l+1)}{r^3}\frac{d}{dr} + \frac{72 - 18l(l+1) + l^2(l+1)^2}{r^4}\right)T_l(r) \approx 0 \tag{3.20}$$

and hence the asymptotic solutions

$$T_l \approx r^{l-4},\ r^{l-5},\ r^{-l-5},\ r^{-l-6}, \tag{3.21}$$

T_l in Eq. (3.18) does not have any of the above asymptotic forms so that

$$\alpha_1 = \alpha_2 = \alpha_3 = 0.$$

Thus we see that if $\psi_0^B = \chi_1^B = 0$, there can exist a nontrivial solution for χ_{-1}^B and ψ_4^B only for ω satisfying Eq. (3.5). For $l = 1$, we have seen only $\omega = \pm i(3M/2Q^2)$ obtains. The pair of solutions

$$\omega = \pm i\frac{3Ml(l+1)}{8Q^2}\left[\left(1 + \frac{4Q^2(l-1)(l+2)}{9M^2}\right)^{1/2} - 1\right]$$

agree with the time dependence of the algebraically special perturbations of the Reissner–Nordström metric noted by Couch and Newman.[9] The case $\psi_0^B = \chi_1^B = 0$ is

algebraically special. Thus it appears that for $l \geq 2$ there exists no solution for the other set of ω.

In any case, since these solutions have an imaginary ω, the cubic polynomial dependence as r given by Eq. (3.18) implies ψ_3 will blow up exponentially either at the horizon ($u \to -\infty$) or spacial infinity ($u \to \infty$). Thus these solutions are physically unacceptable.

IV. PERTURBATIONS WITH $\psi_0^B = \psi_1^B = \psi_3^B = \psi_4^B = \phi_0^B = \phi_2^B = 0$

With $\psi_0^B = \psi_1^B = \psi_3^B = \psi_4^B = \phi_0^B = \phi_2^B = 0$ and condition (3.8) one obtains the following perturbation of the NP equations, Maxwell's equation, and Bianchi identities:

$$3(\Delta + 3\mu)\psi_2^B - 4\phi_1(\Delta - \mu)(\phi_1^B + \phi_1^{*B}) = \frac{-9Mr + 8Q^2}{r^5} X^2 - 3(3\psi_2 + 2\phi_{11})\mu^B + 4\phi_{11}(\mu^B + \mu^{*B}), \tag{4.1a}$$

$$\nu^B = 0, \tag{4.1b}$$

$$\lambda^B = 0, \tag{4.1c}$$

$$(D - \rho)\tau^B = \rho\pi^{*B}, \tag{4.2a}$$

$$(\delta - 2\beta)\tau^B = \rho\lambda^{*B}, \tag{4.2b}$$

$$(D - \rho)\alpha^B = \alpha\rho^B + \rho\pi^B, \tag{4.2c}$$

$$(D - \rho)\beta^B = \beta\rho^{*B}, \tag{4.2d}$$

$$D\gamma^B = \beta(\tau^{*B} + \pi^B - \tau^B - \pi^{*B}) + \psi_2^B + 2\phi_1(\phi_1^B + \phi_1^{*B}), \tag{4.2e}$$

$$(D - \rho)\mu^B - (\delta + 2\beta)\pi^B = \mu\rho^{*B} + \psi_2^B, \tag{4.2f}$$

$$\frac{\xi^2}{r^2} = -\delta\rho^B + \rho(\alpha^{*B} + \beta^B), \tag{4.2g}$$

$$\frac{\cot\theta}{\sqrt{2}r^2}(\xi^2 + \xi^{2*}) + \frac{\xi^3 + \xi^{3*}}{2\sqrt{2}r\sin^2\theta} = (\delta^* + 2\beta)\beta^B - (\delta + 2\beta)\alpha^B + \mu\rho^B + \rho\mu^B + \gamma(\rho^B - \rho^{*B}) - \beta(\alpha^B + \alpha^{*B} - \beta^B - \beta^{*B}) - \psi_2^B + 2\phi_1(\phi_1^B + \phi_1^{*B}), \tag{4.2h}$$

$$\frac{r^2 - 4Mr + 3Q^2}{2r^4}\xi^2 = (\delta^* + 4\beta)\lambda^{*B} - \delta\mu^{*B} - \mu(\alpha^{*B} + \beta^B), \tag{4.2i}$$

$$\frac{r^2 - 4Mr + 3Q^2}{2r^4} X^2 = (\delta + 2\beta)\mu^B - (\Delta + 2\mu + 2\gamma)\mu^B - \mu(\gamma^B + \gamma^{*B}). \tag{4.2j}$$

By applying the first order perturbations of the NP commutators on u, r, θ, and ϕ, one obtains:

$$\frac{\partial X^1}{\partial r} = 0, \tag{4.3a}$$

$$\frac{\partial X^2}{\partial r} = -\gamma^B - \gamma^{*B}, \tag{4.3b}$$

$$\frac{\partial X^3}{\partial r} = \frac{1}{\sqrt{2}r}(\tau^B + \tau^{*B} + \pi^B + \pi^{*B}), \tag{4.3c}$$

$$\frac{\partial X^4}{\partial r} = \frac{i}{\sqrt{2}r\sin\theta}(\pi^B - \pi^{*B} - \tau^B + \tau^{*B}), \tag{4.3d}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(r\xi^1) = 0, \tag{4.4a}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(r\xi^2) = -\alpha^{*B} - \beta^B + \pi^{*B}, \tag{4.4b}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(r\xi^3) = \frac{1}{\sqrt{2}r}\rho^{*B}, \tag{4.4c}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(r\xi^4) = \frac{1}{\sqrt{2}r\sin\theta}\rho^{*B}, \tag{4.4d}$$

$$\delta X^1 - (\Delta + \mu)\xi^1 = 0, \tag{4.5a}$$

$$\delta X^2 - (\Delta + \mu)\xi^2 = \frac{Mr - Q^2}{r^3}\xi^2 + \frac{\Delta}{2r^2}(\alpha^{*B} + \beta^B - \tau^B), \tag{4.5b}$$

$$\delta X^3 - (\Delta + \mu)\xi^3 = \frac{-1}{\sqrt{2}r^2}X^2 + \frac{1}{\sqrt{2}r}(\mu^B - \gamma^B + \gamma^{*B} + \lambda^{*B}), \tag{4.5c}$$

$$\delta X^4 - (\Delta + \mu)\xi^4 = \frac{i}{\sqrt{2}r\sin\theta}\left(-\frac{X^2}{r} + \mu^B - \gamma^B + \gamma^{*B} - \lambda^{*B}\right), \tag{4.5d}$$

$$(\delta^* + 2\beta)\xi^1 - (\delta + 2\beta)\xi^{1*} = \rho^{*B} - \rho^B, \tag{4.6a}$$

$$(\delta^* + 2\beta)\xi^2 - (\delta + 2\beta)\xi^{2*} = \mu^{*B} - \mu^B + \frac{\Delta}{2r^2}(\rho^B - \rho^{*B}), \tag{4.6b}$$

$$(\delta^* + 2\beta)\xi^3 - (\delta + 2\beta)\xi^{3*} = \frac{1}{\sqrt{2}r}(\alpha^B - \alpha^{*B} + \beta^B - \beta^{*B}), \tag{4.6c}$$

$$(\delta^* + 2\beta)\xi^4 - (\delta + 2\beta)\xi^{4*} = \frac{i}{\sqrt{2}r\sin\theta}\left(\frac{\xi^2 + \xi^{2*}}{r} + \cot\theta(\xi^3 + \xi^{3*}) + \alpha^B + \alpha^{*B} - \beta^B - \beta^{*B}\right). \tag{4.6d}$$

From Eq. (3.17) with $\psi_3^B = 0$ one obtains

$$\frac{\partial B^0}{\partial\theta} = \frac{\partial B^0}{\partial\phi} = \frac{\partial C^0}{\partial\theta} = \frac{\partial C^0}{\partial\phi} = 0. \tag{4.7}$$

Subtraction of the complex conjugate of Eq. (3.9d) from Eq. (3.9d) yields

$$\mu^B - \mu^{*B} = \frac{i\Delta A^0}{r^4} + \frac{2i}{r}\frac{\partial A^0}{\partial u} - \frac{2i}{Q}\frac{\partial B^0}{\partial u}. \tag{4.8}$$

Similarly, subtraction of the complex conjugate of Eq. (4.1a) from Eq. (4.1a) gives

$$\mu^B - \mu^{*B} = \frac{i\Delta A^0}{r^4} + \frac{2i}{r}\frac{\partial A^0}{\partial u} + \frac{2i}{3Mr - 4Q^2} \times\left(\frac{2Q^2}{r}\frac{\partial A^0}{\partial u} + r\frac{\partial C^0}{\partial u}\right). \tag{4.9}$$

Comparing Eqs. (4.8) and (4.9) one concludes

$$\frac{\partial A^0}{\partial u} = \frac{\partial B^0}{\partial u} = \frac{\partial C^0}{\partial u} = 0. \tag{4.10}$$

This, together with Eq. (4.7), indicates

$$B^0, C^0 = \text{const.} \tag{4.11}$$

Integration of Eq. (4.2a) for τ^B using Eqs. (3.15) and (4.11) yields

$$\tau^B = \frac{-i}{\sqrt{2}r^2}\left(\frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)A^0 + w^0 + iW^0. \tag{4.12}$$

Put Eq. (4.12) into Eq. (4.2b) with $\lambda^{*B} = 0$ to get

$$A^0 = A - a\cos\theta, \tag{4.13}$$

where A and a are real constants. Since τ^B is invariant under any infinitesimal tetrad or coordinate transformation, the boundary condition that the space be

asymptotically flat implies $w^0 = W^0 = 0$. Hence

$$\tau^B = \frac{-ia\sin\theta}{\sqrt{2}r^2}, \quad \pi^B = \frac{ia\sin\theta}{\sqrt{2}r^2}. \tag{4.14}$$

Now integrate Eqs. (4.2c)–(4.2f) successively using the results obtained up to the previous step of each integration to obtain:

$$\alpha^B = \frac{\cot\theta}{2\sqrt{2}r^2}(iA - ia\cos\theta) + \frac{ia\sin\theta}{\sqrt{2}r^2} + \frac{e^0 + iE^0}{r}, \tag{4.15}$$

$$\beta^B = \frac{\cot\theta}{2\sqrt{2}r^2}(iA - ia\cos\theta) + \frac{f^0 + iF^0}{r}, \tag{4.16}$$

$$\gamma^B = \frac{\Delta}{2r^4}(iA - ia\cos\theta) - \frac{2Qb^0}{r^3} - \frac{c^0 + i(A + C^0)}{2r^2} + g^0 + iG^0, \tag{4.17}$$

$$\mu^B = \frac{\Delta}{2r^4}(iA - ia\cos\theta) - \frac{2Qb^0}{r^3} - \frac{c^0 + i(A + C^0)}{r^2} + \frac{h^0 + iH^0}{r}. \tag{4.18}$$

From Eqs. (4.8) and (4.10) one can see that the imaginary part of μ^B is the first term of the right-hand side of Eq. (4.18) so that

$$C^0 = -A, \quad H^0 = 0. \tag{4.19}$$

Also, by a combined transformation $l \to l - h^0$ and $X^2 \to X^2 + rh^0$ one can make

$$h^0 = 0 \tag{4.20}$$

without affecting Eqs. (3.8) and the condition $a^0 = 0$ in Eq. (3.12).

Perturbations of the tetrad components can be obtained by integrating Eqs. (4.3a)–(4.4d),

$$\begin{aligned} &X^1 = k^0, \\ &X^2 = -\frac{2Qb^0}{r^2} - \frac{c^0}{r} + l^0 - 2rg^0, \\ &X^3 = n^0, \quad X^4 = \frac{a}{r^2} + \rho^0, \end{aligned} \tag{4.21}$$

$$\begin{aligned} &\xi^1 = \frac{r^0 + iH^0}{r}, \\ &\xi^2 = \frac{s^0 + iS^0}{r} - (e^0 + f^0) + i(E^0 - F^0), \\ &\xi^3 = \frac{iA - ia\cos\theta}{\sqrt{2}r^2} + \frac{l^0 + iT^0}{r}, \\ &\xi^4 = \frac{-A + a\cos\theta}{\sqrt{2}r^2\sin\theta} + \frac{v^0 + iV^0}{r}. \end{aligned} \tag{4.22}$$

We make use of a coordinate transformation $(X^1 \to X^1 - \int^u k^0\,du,\ X^2 \to X^2 - \int^u n^0\,du,\ X^3 \to X^3 - \int^u \rho^0\,du)$ to set

$$k^0 = n^0 = \rho^0 = 0. \tag{4.23}$$

Another tetrad transformation, $m \to (1 - \sqrt{2}iT^0)m$, makes

$$T^0 = 0. \tag{4.24}$$

Putting X^1 and ξ^1 into Eq. (4.5a) shows

$$\frac{\partial}{\partial r} r^0 = \frac{\partial}{\partial r} R^0 = 0 \tag{4.25}$$

which enables us to set (by $X^1 \to X^1 - \sqrt{2}\int^\theta r^0\,d\theta$)

$$r^0 = 0 \tag{4.26}$$

without affecting Eq. (4.23).

Now solve Eq. (4.6a) to obtain

$$R^0 = \sqrt{2}A\cot\theta + \frac{a\sin\theta}{\sqrt{2}} + \frac{R}{\sin\theta}, \tag{4.27}$$

where R is a real constant and can be removed by $X^1 \to X^1 - \sqrt{2}R\phi$.

Now solving Eqs. (3.9b), (3.9d), (3.10b), (4.1a), (4.2i) (4.2j), (4.5c), and (4.5d) simultaneously, using the results obtained so far, one finds

$$\frac{\partial b^0}{\partial u} = \frac{\partial b^0}{\partial\theta} = \frac{\partial b^0}{\partial\phi} = \frac{\partial c^0}{\partial u} = \frac{\partial c^0}{\partial\theta} = \frac{\partial c^0}{\partial\phi} = 0, \tag{4.28}$$

$$g^0 = G^0 = l^0 = 0, \tag{4.29}$$

$$e^0 = -f^0, \quad E^0 = F^0, \tag{4.30}$$

$$\frac{\partial l^0}{\partial u} = \frac{\partial v^0}{\partial u} = \frac{\partial V^0}{\partial u} = 0. \tag{4.31}$$

Equation (4.31) guarantees that we can perform the combined transformation $[X^3 \to X^3 + \eta^3,\ X^4 \to X^4 + \eta^4,\ m \to (1 + iH)m;\ \partial\eta^3/\partial\theta = -\sqrt{2}l^0,\ H = -(1/\sin\theta)(\partial\eta^3/\partial\phi),\ \partial\eta^4/\partial\phi = -\sqrt{2}\sin\theta\, V^0 - \cot\theta\eta^3]$ to remove l^0 and V^0 without affecting any of the previous restrictions. With these results, we solve Eqs. (4.6c), (4.6d), and (4.2h) simultaneously to obtain

$$E^0 = 0, \quad v^0 = \frac{v_1(\phi)}{\sin^2\theta} + v_2(\theta), \quad e^0 = \frac{-1}{2\sin^2\theta}\frac{\partial v_1}{\partial\phi}. \tag{4.32}$$

Finally, $v_2(\theta)$ can be eliminated by $X^4 \to X^4 + \eta^4(\theta)$ with $\partial\eta^4(\theta)/\partial\theta = -\sqrt{2}v_2(\theta)$ and $v_1(\theta)$ also is eliminated by a combined transformation

$$X^3 \to X^3 + \eta^3(\phi), \quad X^4 \to X^4 + \cot\theta\eta^4(\phi), \quad m \to \left(1 + i\frac{H(\phi)}{\sin(\theta)}\right)m,$$

with

$$\eta^4 = -H - \sqrt{2}v_1, \quad \eta^3 = -\eta^4$$

and

$$\frac{\partial H}{\partial\phi} + H + \sqrt{2}\frac{\partial v_1}{\partial\phi} = 0.$$

Notice that none of the transformations so far affect the preceding restrictions.

A summary of the results follows:

$$\begin{aligned} &\psi_2^B = \frac{3Mr - 2Q^2}{r^5}(iA - ia\cos\theta) + \frac{4Qb^0}{r^4} + \frac{c^0 - iA}{r^3}, \\ &\phi_1^B = -\frac{iQ}{r^3}(iA - ia\cos\theta) + \frac{b^0 + iB^0}{r^2}, \\ &l^{\mu B} = 0, \\ &n^{\mu B} = \left(-\frac{2Qb^0}{r^2} - \frac{c^0}{r}\right)\delta_2^\mu + \frac{a}{r^2}\delta_4^\mu, \\ &m^{\mu B} = \frac{1}{\sqrt{2}r}\Big(ia\sin\theta + 2iA\cot\theta)\delta_1^\mu + \frac{iA - ia\cos\theta}{r}\delta_3^\mu \\ &\qquad + \frac{-A + a\cos\theta}{r\sin\theta}\delta_4^\mu\Big), \end{aligned} \tag{4.33}$$

where A, a, b^0, c^0, and B^0 are real constants.

The perturbations generated by b^0 and c^0 correspond to infinitesimal changes in the charge and mass parameters, respectively and the perturbation by a is obtained by linearizing the Kerr–Newman solution about the Reissner–Nordström solution. The perturbation generated by B^0 represents a magnetic monopole of the black hole which is physically unacceptable. The A perturbation is the one obtained by linearizing the Kerr–NUT solution[10,11] (generalized to the charged case) about the Reissner–Nordström solution keeping the angular-momentum parameter zero. This solution is singular along the negative z axis. Thus, excluding the above two perturbations, we conclude that with $\chi_1^B = 0$, the only well-behaved perturbations are perturbations to other Reissner–Nordström solutions or to Kerr–Newman solutions.

V. DISCUSSION

We have shown that the nontrivial part of a perturbation of a Reissner–Nordström black hole is completely determined by specifying χ_1^B. That is χ_1^B determines the perturbation up to a change in the mass, charge, or angular momentum of a black hole.

Of course the interesting problem of performing a similar analysis for the Kerr–Newman black hole remains. Though the coupled perturbation equations are known[5] their apparent lack of separability is a barrier to analysis.

It should also be noted that Moncrief[12] has obtained rather simpler perturbation equations for the Reissner–Nordström black hole by use of a Hamiltonian variational principle. It appears that his variables are not simply related to those used in this paper except in the asymptotic region.

APPENDIX

The tetrad system is determined only up to the 6-parameter group of the following infinitesimal transformations[13]:

(i)

$$l \to l, \quad n \to n + dm^* + d^*m, \quad m \to m + dl, \tag{A1}$$

(ii)

$$l \to l + em^* + e^*m, \quad n \to n, \quad m \to m + en, \tag{A2}$$

(iii)

$$l \to (1+\Lambda)l, \quad n \to (1-\Lambda)l, \quad m \to (1+iH)m, \tag{A3}$$

where d and e are complex and Λ and H are real.

We also have the freedom of infinitesimal coordinate transformations

$$X^\mu \to X^\mu + \eta^\mu. \tag{A4}$$

Under the combined transformations of Eqs. (A1)–(A4),

$$\begin{aligned} &\psi_0^B \to \psi_0^B, \quad \psi_4^B \to \psi_4^B, \\ &\psi_1^B \to \psi_1^B + 3e\psi_2, \quad \psi_3^B \to \psi_3^B + 3d^*\psi_2, \\ &\phi_0^B \to \phi_0^B + 2e\phi_1, \quad \phi_2^B \to \phi_2^B + 2d^*\phi_1. \end{aligned} \tag{A5}$$

The above relations show that ψ_0^B, ψ_4^B, χ_1^B ($\equiv 3\psi_2\phi_0^B - 2\phi_1\psi_1^B$), and χ_{-1}^B ($\equiv 3\psi_2\phi_2^B - 2\phi_1\psi_3^B$) are invariant. In addition, one can always eliminate ϕ_0^B and ϕ_2^B by choosing

$$e = -\phi_0^B/2\phi_1 \quad \text{and} \quad d^* = -\phi_2^B/2\phi_1. \tag{A6}$$

Under the remaining transformations [Eqs. (A3) and (A4)], the various quantities transform as following:

$$\begin{aligned} &\epsilon^B \to \epsilon^B + \frac{1}{2}\frac{\partial}{\partial r}(\Lambda - iH), \\ &\rho^B \to \rho^B - \eta^2/r^2 - \Lambda/r, \\ &\mu^B \to \mu^B - \frac{r^2 - 4Mr + 3Q^2}{2r^4}\eta^2 + \frac{\Delta}{2r^3}\Lambda, \\ &l^{\mu B} \to l^{\mu B} + \frac{\partial \eta^\mu}{\partial r} + \Lambda\delta_2^\mu, \\ &n^{\mu B} \to n^{\mu B} + \frac{\partial \eta^\mu}{\partial u} - \frac{\Delta}{2r^2}\frac{\partial \eta^\mu}{\partial r} - \Lambda\delta_1^\mu + \left(\frac{\Delta}{2r^2} + \frac{Mr - Q^2}{r^3}\eta^2\right)\delta_2^\mu, \\ &m^{\mu B} \to m^{\mu B} + \frac{1}{\sqrt{2}r}\left(\frac{\partial \eta^\mu}{\partial \theta} + \frac{i}{\sin\theta}\frac{\partial \eta^\mu}{\partial \phi}\right) + \frac{1}{\sqrt{2}r}\left(\frac{\eta^2}{r} + iH\right)\delta_3^\mu \\ &\qquad + \frac{i}{\sqrt{2}r\sin\theta}\left(\frac{\eta^2}{r} + \cot\theta\eta^3 + iH\right)\delta_4^\mu. \end{aligned} \tag{A7}$$

[1] R.M. Wald, J. Math. Phys. 14, 1453 (1973).
[2] The superscript B represents perturbations and the quantities without any superscript are unperturbed ones except in some obviously different cases.
[3] R. Ruffini and A. Treves, Astrophys. Lett. 13, 109 (1973).
[4] R.M. Wald, Phys. Rev. D 10, 1680 (1974).
[5] C.H. Lee, J. Math. Phys. 17, 1226 (1976).
[6] S.K. Bose, J. Math. Phys. 16, 772 (1975).
[7] F.A.E. Pirani, *Lectures on General Relativity, Brandeis Summer Institute in Theoretical Physics* (Prentice-Hall, Englewood Cliffs, New Jersey, 1964).
[8] E. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962).
[9] W.E. Couch and E.T. Newman, J. Math. Phys. 14, 285 (1973).
[10] M. Demianski and E. Newman, Bull. Acad. Pol. Sci. Ser. Sci. Math. Astron. Phys. 14, 653 (1966).
[11] W. Kennersley, J. Math. Phys. 10, 1195 (1969).
[12] V. Moncrief, Phys. Rev. D 9, 2707 (1974); 10, 1057 (1974); 12, 1526 (1975).
[13] A. Janis and E. Newman, J. Math. Phys. 6, 902 (1965).

Zeros of the grand canonical partition function: Generalization of a result of Lee and Yang

M. E. Baur and R. M. Sinder

*Department of Chemistry, University of California, Los Angeles, California 90024**
(Received 18 June 1976)

A simple sufficiency condition for the zeros of a polynomial of grand partition function form to lie entirely on the unit circle in the complex fugacity (z) plane is rigorously proven. The condition has two parts: the canonical partition function $Q_n(M)$ is symmetric, $Q_n(M) = Q_{M-n}(M)$, and is bounded above by the binomial coefficient $\binom{M}{n}$. This represents a generalization of the condition given by Lee and Yang in the context of the Ising model and the proof is independent of theirs. Necessity of the condition is trivially proven.

The analysis of the grand canonical partition function (GPF) in terms of the limiting distribution of zeros of finite volume approximants by Yang and Lee[1-3] is widely accepted as providing a rigorous and definitive framework for the formal characterization of phase transitions. Unfortunately, practical application of the Yang–Lee methodology to realistic model systems has proven very difficult; these authors themselves provided one of the few such applications in their treatment of the Ising model,[4] showing rigorously that the GPF zeros for this case must lie on the unit circle in the complex z plane for an appropriately defined fugacity z.

The Lee–Yang proof for the Ising model zeros distribution is not entirely transparent and admits of no evident generalization or extension. By proceeding in a different fashion, drawing upon "polynomial" methods,[5] we have found it possible to advance somewhat beyond their work in this regard. In particular we here present a sufficiency proof for the zeros of a polynomial of GPF type to lie entirely on the unit circle which includes the Lee–Yang Ising model result as a special case, but is somewhat more general and makes contact with the properties of the familiar binomial expansion in suggestive fashion. We state our result in the form of a

Theorem: Any GPF polynomial $\mathcal{Q}_M = \sum_{n=0}^{M} Q_n(M) z^n$ for which

(1) $Q_n(M) = Q_{M-n}(M)$ (1)

(2) $Q_n(M) \leq \binom{M}{n}$, all n, where $\binom{M}{n}$ is the binomial coefficient, has all zeros on the unit circle, $|z| = 1$.

The symmetry property (1) is, of course, also required for the Lee–Yang proof and appears to limit the result to lattice models. Property (2) is a generalization of the Lee–Yang requirement that Q_n be representable as a product of $\binom{M}{n}$ and factors $x_{\alpha\beta}$ with all $x_{\alpha\beta} \leq 1$.

Proof: We use two lemmas from the literature[6,7]:

Lemma 1: Any circle C which encloses all the zeros of a polynomial $f(z)$ of order M also encloses all the zeros of all its derivatives $f^{(k)}(z)$, $1 \leq k \leq M-1$.

Lemma 2: A necessary and sufficient condition that all the zeros of $f(z) = \sum_0^M Q_n z^n$ have modulus $|z| = 1$ is that

$$Q_0^* Q_n = Q_M Q_{M-n}^*, \tag{2}$$

where (*) denotes complex conjugate, and that $f^{(1)}(z)$ have all its zeros in $|z| \leq 1$.

Evidently a GPF with $Q_0 = Q_M = 1$ which satisfies the symmetry condition (1) then satisfies (2). It remains to examine the circumstances under which the required condition on the zeros of $\mathcal{Q}^{(1)}$ is satisfied. We establish a *sufficient* condition on $\mathcal{Q}$ namely that *all* $\mathcal{Q}^{(k)}$, $1 \leq k \leq M-1$, have all zeros in $|z| \leq 1$. Consider first

$$\begin{aligned}\mathcal{Q}^{(M-1)} &= (M-1)!\, Q_{M-1} + M!\, z Q_M \\ &= M!\,[(1/M)\, Q_{M-1} + z].\end{aligned} \tag{3}$$

The condition that $\mathcal{Q}^{(M-1)}$ have its zeros in $|z| \leq 1$ is evidently that

$$Q_{M-1} = Q_1 \leq M.$$

For $\mathcal{Q}^{(M-2)}$ we have

$$\begin{aligned}\mathcal{Q}^{(M-2)} &= (M-2)!\, Q_{M-2} + (M-1)!\, Q_{M-1} z + \frac{M!}{2} z^2 \\ &= \frac{M!}{2}\left[\frac{2}{(M)(M-1)}\, Q_{M-2} + \frac{2}{M}\, Q_{M-1} z + z^2\right].\end{aligned} \tag{4}$$

It is easily verified that, with $Q_{M-1} \leq M$, one must have $Q_{M-2} \leq \frac{1}{2} M(M-1)$ for $\mathcal{Q}^{(M-2)}$ to have all its zeros in $|z| \leq 1$. This is an absolute bound on Q_{M-2}, independent of any statement about $\mathcal{Q}^{(M-1)}$; it is easily verified that there is no choice of Q_{M-1}, Q_{M-2} such that the zero of $\mathcal{Q}^{(M-1)}$ is in $|z| > 1$ but both zeros of $\mathcal{Q}^{(M-2)}$ are in $|z| \leq 1$.

Proceeding in this way, one sees that the permissible upper bounds for the Q_{m-j} in order that $\mathcal{Q}^{(k)}$ have no zeros outside $|z| \leq 1$ are just the binomial coefficients:

$$Q_{M-j} \leq \binom{M}{M-j}.$$

The general expression for $\mathcal{Q}^{(k)}$ is

$$\begin{aligned}\mathcal{Q}^{(k)} = \frac{M!}{(M-k)!}\Bigg[&z^{M-k} + \frac{(M-k)!}{(M-k-1)!}\,\frac{(M-1)!}{M!}\, Q_{M-1} z^{M-k-1}\Bigg] \\ &+ \cdots + \frac{(M-k)!}{(M-k-j)!}\,\frac{(M-j)!}{M!}\, Q_{M-j} z^{M-k-j} \\ &+ \cdots + \frac{(M-k)!\,k!}{M!}\, Q_k \quad , \qquad 0 \leq j \leq m-k.\end{aligned} \tag{5}$$

Assume now that each (real, positive) Q_{M-j} in $\mathcal{Q}^{(k)}$ is bounded by the binomial coefficient $\binom{M}{M-j}$ and that $\mathcal{Q}^{(k)}$

has all zeros in $|z| \leq 1$. Consider $Q^{(k-1)}$. We find

$$Q^{(k-1)} = \frac{1}{M-k+1}\left[zQ^{(k)} + \sum_{j=0}^{M-k} \frac{(M-j-1)!\,(j+1)}{(M-k-j)!}\, Q_{M-j-1} z^{M-k-j}\right]$$

$$= \frac{1}{M-k+1}\,[zQ^{(k)} + R(k;k-1)], \tag{6}$$

where $R(k;k-1)$ can be written as

$$R(k;k-1) = \frac{M!}{(M-k)!}\left[\frac{1}{M}\, Q_{M-1} z^{M-k} + \cdots + \frac{(M-k)!}{(M-k-j)!}\,\frac{(M-j)!}{M!} \times \left(\frac{j+1}{M-j}\right) Q_{M-j-1} z^{M-k-j} + \frac{(M-k+1)!\,(k-1)!}{M!}\, Q_{k-1}\right]. \tag{7}$$

Now, if we write $Q^{(k)}$ as $\sum_0^{M-k} C_j^{(k)} z^{M-k-j}$, then we have

$$R(k;k-1) = \sum_0^{M-k} C_j^{(k)} \left(\frac{j+1}{M-j}\,\frac{Q_{M-j-1}}{Q_{M-j}}\right) z^{M-k-j}. \tag{8}$$

Therefore, if the condition

$$\frac{Q_{M-j-1}}{Q_{M-j}} \leq \frac{M-j}{j+1}, \quad 0 \leq j \leq M-1 \tag{9}$$

holds, and if further

$$Q_{k-1} \leq \frac{M!}{(M-k+1)!\,(k-1)!},$$

then $R(k;k-1)$ is a polynomial of order $M-k$ meeting the same condition on its coefficients as does $Q^{(k)}$; by hypothesis, then, $R(k;k-1)$ also has all its zeros in $|z| \leq 1$. Thus $Q^{(k-1)}$ is the sum of two polynomials, both of which have all zeros in this domain. If now all the relations (9) are equalities, $R(k;k-1)$ and $Q^{(k)}$ are identical, and it is evident that $Q^{(k-1)}$ has all its zeros in $|z| \leq 1$. This merely implies that Q is the binomial expansion. In the more general case, we make use of:

Lemma 3: *Rouche's theorem*[8]: If $P(z)$ and $S(z)$ are functions analytic interior to a simple closed Jordan curve C, are continuous on C, and

$$|P(z)| < |S(z)| \quad \text{on } C,$$

then the function $F(z) = P(z) + S(z)$ has the same number of zeros interior to C as does $S(z)$. The positions of the zeros of $F(z)$ interior to C are, of course, not in general the same as those of $S(z)$. Now consider $Q^{(k-1)} = (M-k+1)^{-1}[zQ^{(k)} + R(k;k+1)]$ from the standpoint of Rouche's theorem. Evidently neither $Q^{(k)}$ nor $R(k;k+1)$ has zeros for $|z| = 1+\delta$, δ arbitrarily small. Moreover,

$$|Q^{(k)}| < |zQ^{(k)}| = |z|\,|Q^{(k)}|, \quad |z| = 1+\delta. \tag{10}$$

Hence it suffices to prove that $|R(k;k+1)| < |Q^{(k)}|$ in order to apply Rouche's theorem. Now, it is possible that all coefficients

$$C_j^{(k)}\left(\frac{j+1}{M-j}\,\frac{Q_{M-j-1}}{Q_{n-j}}\right)$$

in $R(k;k+1)$ are less than the coefficients $C_j^{(k)}$ of $Q^{(k)}$. If this is true, then

$$|Q^{(k)}| \geq |N(k)|\,|R(k;k+1)| > |R(k;k+1)|,$$

with $N(k)$ some real constant > 1 and application of Rouche's theorem immediately yields that $Q^{(k-1)}$ has the same number of zeros in the domain $|z| < 1+\delta$ as does $Q^{(k)}$. Because of the condition

$$\frac{Q_{M-j-1}}{Q_{M-j}} \leq \frac{M-j}{j+1},$$

no coefficient of $R(k;k+1)$ can be greater than the corresponding coefficient of $Q^{(k)}$. It is possible that a set of coefficients of $R(k;k+1)$ may be equal to those of $Q^{(k)}$ if for the first j indices the equality is realized in the above condition. If any coefficient, say the $(j+1)$th, of $R(k;k+1)$ is less than that of $Q^{(k)}$, then all succeeding coefficients must also be less. At worst, then, a finite set of coefficients of $R(k;k+1)$ is equal to those of $Q^{(k)}$. But in such a case we can divide $R(k;k+1)$:

$$R(k;k+1) = R_1 + R_2 = \sum_{j=0}^{M-k} (C_j^{(k)} - \epsilon)\left(\frac{j+1}{M-j}\,\frac{Q_{M-j-1}}{Q_{M-j}}\right) z^{M-k-j} + \epsilon \sum_{j=0}^{M-k}\left(\frac{j+1}{M-j}\,\frac{Q_{M-j-1}}{Q_{M-j}}\right) z^{M-k-j}, \tag{11}$$

where $0 < \epsilon < 1$, and then apply Rouche's theorem successively to $Q^{(k)} + R_1$ and $Q^{(k)} + R_1 + R_2$. This then proves that the zeros of each $Q^{(k-1)}$ are in the domain $|z| < 1+\delta$; hence in $|z| \leq 1$, if $Q^{(k)}$ has this property and if the coefficients of Q are bounded above by those of the binomial expansion. We have explicitly shown that $Q^{(M-1)}$ and $Q^{(M-2)}$ have all zeros in $|z| \leq 1$. Hence by induction, all $Q^{(k)}$ have all zeros in $|z| \leq 1$. Therefore, by Lemma 2 above, all zeros of Q have modulus unity.

The *necessity* of conditions (1) and (2) is easily established. A distribution of zeros confined to the unit circle implies invariance of Q under the transformation $z \to 1/z$, which requires the symmetry condition (1). With zeros confined to the unit circle, and assuming Q real for real z, we have

$$Q = \prod_j (z - \exp(i\theta_j))\,(z - \exp(-i\theta_j))^{n_j}, \quad \sum_j j n_j = M,$$

where n_j is the multiplicity of the zero of argument θ_j. Thus

$$Q = \prod_j (z^2 - 2z\cos\theta_j + 1)^{n_j}, \quad \sum_j j n_j = M,$$

and it is obvious by direct expansion that the coefficient of an arbitrary term z^k is less than or equal to that of z^k in the expansion of

$$(z^2 + 2z + 1)^{M/2}.$$

REMARKS

The original proof of Lee and Yang[4] applies to the case of the Ising model-lattice gas with binary interaction of arbitrary range. The present proof removes the binary interaction limitation, insofar as condition (2) is still satisfied.

The extreme limits correspond to the binomial expansion case, $Q_n(M) = \binom{M}{n}$ for which there is a single zero of multiplicity M at $z = -1$, and the constant coefficient case $Q_n(M) = 1$, $n \leq M$, $Q_n(M) = 0$, $n > M$, for which $Q = (1 - z^{M+1})/(1 - z)$ and the zeros are spread uniformly over the unit circle with every point a limit point. One sees qualitatively therefore that the correlation effect due to site—site interaction, in reducing the variation in Q_n below that of the binomial series, forces the zeros to spread along the unit circle away from the negative real axis. It is possible that extension of the considerations presented here will prove useful in giving this observation a more detailed form.

*Contribution No. 3215.

[1]C.N. Yang and T.D. Lee, Phys. Rev. **87**, 404 (1952).
[2]K. Huang, *Statistical Mechanics* (Wiley, New York, 1963), Chap. 15.
[3]H.L. Frisch and J.L. Lebowitz, *The Equilibrium Theory of Classical Fluids* (Benjamin, New York, 1964).
[4]T.D. Lee and C.N. Yang, Phys. Rev. **87**, 410 (1952).
[5]M. Marden, *The Geometry of the Zeros of a Polynomial in a Complex Variable*, Mathematical Surveys No. III (American Mathematical Society, Providence, R.I., 1949).
[6]Ref. 5, p. 161.
[7]I. Schur, J. Reine Angew. Math. **147**, 205 (1917); **148**, 122 (1918).
[8]E.C. Titchmarsh, *The Theory of Functions* (Oxford U.P., Oxford, 1939), 2nd ed., p. 116.

Intrinsic geometry of Killing trajectories

E. Honig

Scarborough College and Department of Physics, University of Toronto, West Hill, Ontario, Canada
(Received 5 December 1975; revised manuscript received 30 June 1976)

We use the Frenet-Serret formalism to study the intrinsic geometry of Killing trajectories that are admitted by an arbitrary n-dimensional Riemannian space. The intrinsic quantities associated with these curves, i.e. their curvatures, are found to be constants of the motion that can be evaluated in terms of Hankel determinants. The results are then applied to curves in real quantum mechanics.

I. INTRODUCTION

The advantages of adopting a coordinate free geometric description of a curve instead of a parametric representation has long been recognized by mathematicians but only recently by physicists. Besides divorcing the curve from a coordinate system, it also frees it from the accident of where it starts and the initial direction that it takes. Instead, the curve is completely characterized by a unique set of invariants or scalars that are intrinsic to it. The elegance of this formalism, as well as its revealing aspects, argue for its greater use in physics. This is especially true when the evolution of a physical system is described by a trajectory, be it in ordinary 3-space, space—time, configuration or phase space, etc.[1]

The general approach in analyzing an arbitrary but sufficiently smooth curve in an arbitrary n-dimensional Riemannian space is as follows. The classical Frenet—Serret formalism[2] (essentially a Gram—Schmidt orthogonalization process) is used to generate an orthonormal n-leg or frame of vectors along the curve. These vectors obey the Frenet—Serret equations and they in turn define a set of $(n-1)$ intrinsic scalars or invariants of the curve that uniquely describe it.

In this paper we confine ourselves to an invariant geometrical description of a special but important family of curves, namely the Killing congruence, whose very existence corresponds to special symmetries of the Riemannian space. This problem has previously been considered in four-dimensional curved space—times.[1c] There the analysis shed light on the physics of rotating and nonrotating black holes. Here the discussion is generalized to Killing congruences admitted by arbitrary n-dimensional Riemannian spaces with positive definite metrics. The invariants or curvatures associated with these curves are shown to be constants of the motion that can be evaluated in terms of Hankel determinants.[3] The results are then applied to real quantum mechanics.[4] Specifically, in the application we consider nonrelativistic quantum systems having time independent Hamiltonians and discrete energy spectra. If such a system has an n-dimensional spectrum, its state vector is normalized to unity and is confined to an n-dimensional subspace of Hilbert space. We consider rather, the motion of the $2n$-dimensional real vector formed from the real and imaginary parts of the state vector. This real vector traces out a path on a $(2n-1)$-dimensional hypersphere embedded in a $2n$-dimensional Euclidean space. In particular this curve is a Killing trajectory of the hypersphere, or equivalently a Killing trajectory corresponding to a subgroup of the orthogonal transformations of the Euclidean space.

II. PRELIMINARY REMARKS

We preface our discussion with a few well-known remarks that we shall need. Consider an arbitrary n-dimensional Riemannian manifold that admits one or more Killing vector fields; such a vector field ξ satisfies the Killing equation

$$\xi_{\alpha;\beta}+\xi_{\beta;\alpha}=0, \quad \alpha,\beta=1,\ldots,n, \tag{1}$$

where the semicolon indicates covariant differentiation. The maximal number of such fields that the space can admit is $n(n+1)/2$, and this occurs only for spaces of constant curvature. The converse is also true and, in particular, if the space is Euclidean, the Killing equations are equivalent to

$$\xi_\alpha = F_{\alpha\beta}x^\beta + k_\alpha, \tag{2}$$

where $F_{\alpha\beta}=-F_{\beta\alpha}=\text{const}$ and $k_\alpha=\text{const}$.

The first term on the right corresponds to $n(n-1)/2$ independent orthogonal transformations while the second term corresponds to n independent translations.

We shall concern ourselves here with an invariant geometrical description of these trajectories. To this end we first review briefly some of the intrinsic geometrical properties associated with an arbitrary but sufficiently smooth curve on an arbitrary Riemannian manifold. In particular, let the curve Γ be defined through $x^\mu=x^\mu(s)$, $\mu=1,2,\ldots,n$ where s is the arc length and $x^\mu(s)$ is of class C^{n+1}. The unit tangent vector to Γ is $e^\mu_{[1]}\equiv dx^\mu/ds\equiv x^{(1)\mu}$. Superscripts flanked by round brackets will indicate the order of differentiation; subscripts flanked by square brackets will be used as identification labels for a set of orthonormal vectors (to be specified), the first of which is the tangent vector $e_{[1]}$ above. By repeatedly taking the absolute derivative of $x^{(1)}$ at any point along Γ, we can generate the sequence of vectors $x^{(1)},x^{(2)},\ldots,x^{(i)}$, where

$$x^{(i)\alpha}=x^{(i-1)\alpha}{}_{;\beta}x^{(1)\beta}, \quad i=2,3,\cdots. \tag{3}$$

[When the order of differentiation is small, we will on occasion indicate the order by means of dots. Thus $x^{(2)\alpha}\equiv\dot{x}^{(1)\alpha}\equiv\ddot{x}^\alpha$.] If the first n of these vectors are linearly independent, we can use the Gram—Schmidt orthogonalization process to construct the orthonormal frame $e^\alpha_{[i]}$ $(i=1,2,\ldots,n)$, where $e^\alpha_{[i]}e_{[j]\alpha}=\delta_{ij}$ and

$e^{\alpha}_{[1]} = x^{(1)\alpha} = dx^{\alpha}/ds$. Repeated indices sum from 1 to n and δ_{ij} is the Kronecker delta. These vectors obey the Frenet–Serret equations:

$$\begin{bmatrix} \dot{e}^{\alpha}_{[1]} \\ \dot{e}^{\alpha}_{[2]} \\ \cdot \\ \cdot \\ \cdot \\ \dot{e}^{\alpha}_{[n]} \end{bmatrix} = \begin{bmatrix} 0 & \kappa_1 & 0 & 0 & 0 & \cdot \\ -\kappa_1 & 0 & \kappa_2 & 0 & \cdot & \cdot \\ 0 & -\kappa_2 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \circ & 0 & \kappa_{n-1} \\ \cdot & \cdot & \cdot & 0 & -\kappa_{n-1} & 0 \end{bmatrix} \begin{bmatrix} e^{\alpha}_{[1]} \\ e^{\alpha}_{[2]} \\ \cdot \\ \cdot \\ \cdot \\ e^{\alpha}_{[n]} \end{bmatrix}, \tag{4}$$

where the scalar coefficients or curvatures $\kappa_i(s) = \dot{e}^{\alpha}_{[i]} e_{[i+1]\alpha}$ form a complete set of invariants of the curve in Riemannian space. The curvatures can be expressed explicitly in terms of the Gramian determinants G_i of the vectors $x^{(j)\alpha}$, viz.,[5]

$$\kappa_i = \sqrt{G_{i-1}G_{i+1}}/G_i, \quad i = 1, \ldots, n-1, \tag{5}$$

where

$$G_i \equiv \begin{vmatrix} x^{(1)\mu}x^{(1)}_{\mu} & x^{(1)\mu}x^{(2)}_{\mu} & \cdots & x^{(1)\mu}x^{(i)}_{\mu} \\ x^{(2)\mu}x^{(1)}_{\mu} & x^{(2)\mu}x^{(2)}_{\mu} & \cdots & x^{(2)\mu}x^{(i)}_{\mu} \\ \cdot & \circ & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x^{(i)\mu}x^{(1)}_{\mu} & x^{(i)\mu}x^{(2)}_{\mu} & \cdots & x^{(i)\mu}x^{(i)}_{\mu} \end{vmatrix}. \tag{6}$$

All of the G_i's are positive for $1 \leq i \leq n$. Furthermore, $G_0 \equiv 1$. A knowledge of the curvatures as functions of arc length together with the initial n-leg completely characterizes the curve.

III. GENERAL RELATIONS

Let ξ^{α} be a Killing vector admitted by an n-dimensional Riemannian space, and let the corresponding Killing trajectory be denoted by $x^{\alpha} = x^{\alpha}(s)$, where s is the arc length parameter. The unit tangent vector to the curve can then be written as

$$e^{\alpha}_{[1]} = x^{(1)\alpha} \equiv e^{\phi}\xi^{\alpha}, \quad (1 = e^{\alpha}_{[1]}e_{[1]\alpha} = e^{2\phi}\xi^{\alpha}\xi_{\alpha}). \tag{7}$$

That ϕ is constant follows from differentiating $e^{-2\phi} = \xi^{\alpha}\xi_{\alpha}$ along the curve and then making use of Eq. (1). Again, by taking the absolute derivative of both sides of Eq. (7), we obtain

$$x^{(2)\alpha} = e^{\phi}\xi^{\alpha}{}_{;\beta}x^{(1)\beta}. \tag{8}$$

As in the restricted case of a four-dimensional spacetime,[1c] we define

$$F_{\alpha\beta} \equiv e^{\phi}\xi_{\alpha;\beta} = -F_{\beta\alpha} \tag{9}$$

and note that

$$\dot{F}_{\alpha\beta} = F_{\alpha\beta;\lambda}x^{(1)\lambda} = e^{2\phi}\xi_{\alpha;\beta;\lambda}\xi^{\lambda} = e^{2\phi}R_{\alpha\beta\lambda\delta}\xi^{\lambda}\xi^{\delta} = 0, \tag{10}$$

where $R_{\alpha\beta\lambda\delta}$ is the Riemann tensor. Equation (8) can then be written as

$$x^{(2)\alpha} = F^{\alpha}{}_{\beta}x^{(1)\beta}, \quad F^{\alpha}{}_{\beta} = e^{\phi}\xi^{\alpha}{}_{;\beta} = -F_{\beta}{}^{\alpha} = \text{const}. \tag{11}$$

The above relation, though first noted for spacetimes, is in fact quite independent of the dimension of the space or of the signature of the metric.

From Eq. (11) it follows that

$$x^{(i)\alpha} = F^{\alpha}{}_{\beta}x^{(i-1)\beta} = (F^{i-1})^{\alpha}{}_{\lambda}x^{(1)\lambda}, \quad i = 2, \ldots, n, \tag{12}$$

where

$$(F^{i-1})^{\alpha}{}_{\lambda} = F^{\alpha}{}_{\beta}F^{\beta}{}_{\gamma}\cdots F^{\rho}{}_{\lambda}, \tag{13}$$

$F^{\mu}{}_{\nu}$ being repeated $i-1$ times in Eq. (13). We are now in a position to evaluate all possible inner products between the vectors $x^{(i)\alpha}$ and hence, via Eqs. (5) and (6), all of the curvatures of the curve. In particular, by using Eqs. (12) and (13) together with the antisymmetry property of $F_{\alpha\beta}$, we find that

$$\begin{aligned} x^{(i)\alpha}x^{(j)}_{\alpha} &= (-1)^{i-1}(F^{i+j-2})_{\mu\nu}x^{(1)\mu}x^{(1)\nu} \\ &= (-1)^{j-1}(F^{i+j-2})_{\mu\nu}x^{(1)\mu}x^{(1)\nu}. \end{aligned} \tag{14}$$

Since $(F^{i+j-2})_{\mu\nu} = (-1)^{i+j-2}(F^{i+j-2})_{\nu\mu}$, we see from Eq. (14) that $x^{(i)\alpha}x^{(j)}{}_{\alpha}$ is nonzero or zero as $i+j$ is an even or odd integer. If we define

$$\lambda_k \equiv (F^{2k-2})_{\mu\nu}x^{(1)\mu}x^{(1)\nu}, \tag{15}$$

then it follows that

$$x^{(i)\alpha}x^{(j)}{}_{\alpha} = \begin{cases} (-1)^{i-1}\lambda_k = (-1)^{j-1}\lambda_k \\ \quad \text{for all } i,j,k = 1, \ldots, n \ni i+j = 2k \\ 0 \\ \quad \text{for all } i,j,k = 1,2,\ldots,n \ni i+j = 2k+1. \end{cases} \tag{16}$$

We shall return to Eqs. (15) and (16) for the explicit evaluation of the curvatures, but first we note that all of these invariants are constants of the motion, i.e.,

$$\kappa_i = \text{const}, \quad i = 1, \ldots, n-1 \tag{17}$$

provided that $\kappa_{i-1} \neq 0$. [Once we encounter a vanishing curvature, the Gram–Schmidt process ends and the ensuing curvatures can no longer be defined via Eq. (5). They are usually set equal to zero.] To verify Eq. (17), we need only generalize Gluck's observation[6] made in reference to a specific curve in a four-dimensional Euclidean space R^4. For our case, i.e., for any Killing trajectory in an arbitrary n-dimensional Riemannian space V^n, it reads as follows: "Given any two points on this curve, there is an isometry of V^n [7] onto itself which takes the curve onto itself and takes the one point onto the other. Thus the various curvatures are the same at all points of the curve." [That the curvatures are constant can also be seen if we differentiate λ_k along the trajectory. Equations (9) and (11) guarantee that $\dot{\lambda}_k = 0$, and so it follows from Eq. (16) that $x^{(i)\alpha}x^{(j)}{}_{\alpha}$ is constant along the curve for all $i,j = 1,2,\ldots,n$. The defining equations, (5) and (6), then yield Eq. (17).]

Finally we note that each of the Frenet vectors obeys the same equation as the first. Again the proof is a generalization from curved spacetime. Remembering that $x^{(1)\alpha} = e^{\alpha}_{[1]}$, we can combine Eqs. (4) and (11) as follows:

$$\kappa_1 e^{\alpha}_{[2]} = \dot{e}^{\alpha}_{[1]} = F^{\alpha}{}_{\beta}e^{\beta}_{[1]}. \tag{18}$$

By differentiating Eq. (18) and making use of Eqs. (17), (10), and (4), we find that for $\kappa_1 \neq 0$

$$\mathring{e}^{\alpha}_{[2]} = F^{\alpha}{}_{\beta} e^{\beta}_{[2]}. \tag{19}$$

If this procedure is continued we find, in general, that

$$\mathring{e}^{\alpha}_{[i]} = F^{\alpha}{}_{\beta} e^{\beta}_{[i]}. \tag{20}$$

IV. EXPLICIT EVALUATION OF THE CURVATURES

We are now in a position to evaluate the Gramian determinants G_i of the vectors $x^{(j)\alpha}$. It is particularly useful to look at a specific case, say G_6, as this will give us insight into the general structure of G_i. From Eqs. (6) and (16) we obtain

$$G_6 = \begin{vmatrix} \lambda_1 & 0 & \lambda_2 & 0 & \lambda_3 & 0 \\ 0 & -\lambda_2 & 0 & -\lambda_3 & 0 & -\lambda_4 \\ \lambda_2 & 0 & \lambda_3 & 0 & \lambda_4 & 0 \\ 0 & -\lambda_3 & 0 & -\lambda_4 & 0 & -\lambda_5 \\ \lambda_3 & 0 & \lambda_4 & 0 & \lambda_5 & 0 \\ 0 & -\lambda_4 & 0 & -\lambda_5 & 0 & -\lambda_6 \end{vmatrix}. \tag{21}$$

By successively interchanging three rows and three columns it is simple to verify that G_6 reduces to

$$G_6 = \begin{vmatrix} \lambda_1 & \lambda_2 & \lambda_3 & 0 & 0 & 0 \\ \lambda_2 & \lambda_3 & \lambda_4 & 0 & 0 & 0 \\ \lambda_3 & \lambda_4 & \lambda_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_2 & -\lambda_3 & -\lambda_4 \\ 0 & 0 & 0 & -\lambda_3 & -\lambda_4 & -\lambda_5 \\ 0 & 0 & 0 & -\lambda_4 & -\lambda_5 & -\lambda_6 \end{vmatrix} \tag{22}$$

or

$$G_6 = (-1)^3 \begin{vmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_2 & \lambda_3 & \lambda_4 \\ \lambda_3 & \lambda_4 & \lambda_5 \end{vmatrix} \begin{vmatrix} \lambda_2 & \lambda_3 & \lambda_4 \\ \lambda_3 & \lambda_4 & \lambda_5 \\ \lambda_4 & \lambda_5 & \lambda_6 \end{vmatrix}. \tag{23}$$

The same operations applied to G_5 yield

$$G_5 = (-1)^2 \begin{vmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_2 & \lambda_3 & \lambda_4 \\ \lambda_3 & \lambda_4 & \lambda_5 \end{vmatrix} \begin{vmatrix} \lambda_2 & \lambda_3 \\ \lambda_3 & \lambda_4 \end{vmatrix}. \tag{24}$$

Now we define

$$D_{1,j} \equiv \begin{vmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \cdots & \lambda_j \\ \lambda_2 & \lambda_3 & \cdot & & \\ \lambda_3 & \cdot & & & \\ \vdots & & & & \\ \lambda_j & \cdot & \cdot & \cdots & \lambda_{2j-1} \end{vmatrix}, \quad j \geq 1, \tag{25}$$

and

$$D_{2,j} \equiv \begin{vmatrix} \lambda_2 & \lambda_3 & \lambda_4 & \cdots & \lambda_j \\ \lambda_3 & \lambda_4 & \cdot & & \\ \lambda_4 & \cdot & & & \\ \vdots & & & & \\ \lambda_j & \cdot & \cdot & \cdots & \lambda_{2j-2} \end{vmatrix}, \quad j \geq 2. \tag{26}$$

We further stipulate that $D_{2,j} = D_{1,j-1} \equiv 1$ for $j = 1$. Then it is not hard to show that in general

$$G_{2j} = (-1)^j D_{1,j} D_{2,j+1} \tag{27}$$

while

$$G_{2j-1} = (-1)^{j-1} D_{1,j} D_{2,j}. \tag{28}$$

Note that $G_1 = D_{1,1} = \lambda_1 = x^{(1)\mu} x^{(1)}_{\mu} = 1$ and that we have already defined $G_0 = 1$. $D_{1,j}$ and $D_{2,j}$ as defined by Eqs. (25) and (26) are Hankel determinants.[3] They are particularly important in the theory of rational polynomials.[8]

It follows immediately from Eqs. (5), (27), and (28) that

$$\kappa^2_{2j} = -D_{1,j+1} D_{2,j} / D_{1,j} D_{2,j+1} \tag{29}$$

while

$$\kappa^2_{2j-1} = -D_{1,j-1} D_{2,j+1} / D_{1,j} D_{2,j}. \tag{30}$$

By explicitly expressing each of the κ_i's in terms of the previous ones we arrive, after some manipulation, at

$$\kappa^2_{2i} = \frac{D_{1,i+1}}{\kappa^2_{2i-1} \prod_{j=1}^{i-1} [\kappa_{2j-1} \kappa_{2j}]^{2(i-j+1)}} \tag{31}$$

and

$$\kappa^2_{2i-1} = \frac{(-1)^i D_{2,i+1}}{\prod_{j=1}^{i-1} \kappa^{2(i-j+1)}_{2j-1} \kappa^{2(i-j)}_{2j}}. \tag{32}$$

Simpler relationships can be obtained by noting that Eq. (5) implies that

$$\prod_{j=1}^{i} \kappa^2_j = \frac{G_{i+1}}{G_i} \tag{33}$$

and this together with Eqs. (27) and (28) yields

$$\prod_{j=1}^{2i} \kappa^2_j = \frac{D_{1,i+1}}{D_{1,i}}, \tag{34}$$

$$\prod_{j=1}^{2i-1} \kappa^2_j = -\frac{D_{2,i+1}}{D_{2,i}}. \tag{35}$$

Alternatively, we can also use the same equations to show that

$$\prod_{j=1}^{i} \kappa^2_{2j} = (-1)^i \frac{D_{1,i+1}}{D_{2,i+1}}, \tag{36}$$

$$\prod_{j=1}^{i} \kappa^2_{2j-1} = (-1)^i \frac{D_{2,i+1}}{D_{1,i}}. \tag{37}$$

Equations (36) and (37) in turn imply Eqs. (34) and (35).

V. APPLICATIONS TO REAL QUANTUM MECHANICS

One example of a Killing trajectory representing the evolution of a physical system is the world line of a charged particle in a homogeneous electromagnetic field in flat spacetime.[1c] Here we shall consider another example, which, interestingly enough, comes from quantum theory.

In the standard formulation of quantum mechanics the state vector corresponding to a physical system moves in complex Hilbert space. However, it has been shown[9] that the pure states of a general quantum system can always be represented in a one-to-one manner by rays of a Hilbert space that is defined over the fields of real, complex or quaternion numbers. The relation between real, complex and quaternion quantum mechanics has been discussed in some detail in the literature.[4]

Here we consider an arbitrary nonrelativistic quantum system having a time independent Hamiltonian and a discrete energy spectrum. The time evolution of the state vector describing this system in complex Hilbert space is, of course, given by the Schrödinger equation, viz.,

$$H|\psi\rangle = i\frac{\partial|\psi\rangle}{\partial t}. \tag{38}$$

(We have set $\hbar = 1$ for convenience.) If the system has an n-dimensional spectrum, then we can, in the usual manner, write

$$|\psi\rangle = \sum_{i=1}^{n} c_i|i\rangle, \tag{39}$$

where $\{|i\rangle; i = 1, \ldots, n\}$ consists of a fixed (i.e., time-independent), orthonormal and complete set of basis vectors while the c_i's are time dependent complex coefficients. The normalization condition on $|\psi\rangle$ reduces to $1 = \langle\psi|\psi\rangle = \sum_{i=1}^{n} c_i^* c_i$. It follows from Eq. (39) that the Schrödinger equation (38) can be recast as

$$\frac{dc}{dt} = -iHc, \tag{40}$$

where c is a column vector whose elements are c_j $(j = 1, \ldots, n)$ while H is the matrix representation of the Hermitian Hamiltonian operator. In particular its elements are given by $H_{ij} \equiv \langle i|H|j\rangle$. Now we define

$$c \equiv a + ib, \quad c^* = a - ib, \tag{41}$$

where a and b are real column vectors whose elements correspond to the real and imaginary parts of c, i.e., $c_j = a_j + ib_j$. By adding and subtracting Eq. (40) and its complex conjugate and then substituting the relations (41) into the resulting equations, we obtain

$$\frac{da}{dt} = -\frac{i}{2}(H - H^*)a + \tfrac{1}{2}(H + H^*)b, \tag{42}$$

$$\frac{db}{dt} = -\tfrac{1}{2}(H + H^*)a - \frac{i}{2}(H - H^*)b. \tag{43}$$

If we further define the real $2n \times 1$ column vector x as

$$x = \begin{pmatrix} a \\ b \end{pmatrix}, \tag{44}$$

the real, constant, symmetric $n \times n$ matrix S as

$$S \equiv \tfrac{1}{2}(H + H^*) = S^T, \tag{45}$$

and the real, constant antisymmetric $n \times n$ matrix A as

$$A \equiv -\frac{i}{2}(H - H^*) = -A^T, \tag{46}$$

then Eqs. (42) and (43) can be written in compact form, viz.,

$$\frac{dx}{dt} = \overline{F}x, \tag{47}$$

where

$$\overline{F} \equiv \begin{bmatrix} A & S \\ -S & A \end{bmatrix} = -\overline{F}^T \tag{48}$$

is a $2n \times 2n$ real, constant, antisymmetric matrix.

The normalization condition on $|\psi\rangle$ immediately translates into

$$x^T x = \sum_{i=1}^{n} (a_i^2 + b_i^2) = \sum_{i=1}^{n} c_i^* c_i = 1 \tag{49}$$

so that the state of the system is now represented by a real $2n$-dimensional vector, one of whose ends is fixed, while the other traces out a path on a $(2n-1)$-dimensional unit hypersphere embedded in a $2n$-dimensional Euclidean space. The motion is governed by Eq. (47), i.e., by the real quantum mechanical equivalent of the ordinary Schrödinger equation.

We note at this point that dx/dt is a tangent vector to the curve though it is not a unit vector. By comparing Eqs. (47) and (48) with Eq. (2) and keeping in mind the discussion dealing with the latter, we see that the real Schrödinger equation is, in effect, a Killing equation corresponding to orthogonal transformations in a $2n$-dimensional Euclidean space. In other words, the curve that develops with time is a Killing trajectory corresponding to orthogonal transformations of the Euclidean space, or equivalently it is a Killing trajectory of the hypersphere.

That the dynamics involved do not give rise to the translations of flat space is to be expected. Rather, we encounter an n^2 parameter subgroup of the rotation group. In order to see this, we consider the following. Orthogonal transformations in a $2n$-dimensional Euclidean space are generated via equations of the form

$$\frac{dx}{dt} = \mathcal{G}x \tag{50}$$

where $\mathcal{G} = -\mathcal{G}^T = \text{const}$ is a generator of the full $[2n(2n-1)/2]$-parameter rotation group. Since we are dealing with even-dimensional spaces, $\mathcal{G}$ can always be expressed in matrix form as

$$\mathcal{G} = \begin{bmatrix} A & K \\ -K^T & B \end{bmatrix}, \tag{51}$$

where $A = -A^T$ and $B = -B^T$ are constant, antisymmetric $n \times n$ matrices and K is an arbitrary constant $n \times n$ matrix. [Note that A and B each contribute $n(n-1)/2$ parameters while K accounts for n^2 parameters.] The

generators satisfy a Lie algebra, and it is quite straightforward to verify that

$$[\mathcal{G}_1, \mathcal{G}_2] \equiv \mathcal{G}_1\mathcal{G}_2 - \mathcal{G}_2\mathcal{G}_1 = \mathcal{G}_3, \tag{52}$$

where each of the $\mathcal{G}_i$'s, $i=1,2,3$, is of the form given by Eq. (51).

Now $\overline{F}$ as defined through Eq. (48) is certainly a generator of the full rotation group since it is a special example of $\mathcal{G}$. On the other hand, it is again a simple task to show that

$$[\overline{F}_1, \overline{F}_2] = \overline{F}_3; \tag{53}$$

that is, the $\overline{F}$ generators satisfy their own Lie algebra. It follows at once that the orthogonal transformations corresponding to Eqs. (47) and (48) form an n^2 parameter subgroup of the full $2n(2n-1)/2$ parameter rotation group. [$S = S^T$ and $A = -A^T$ have respectively $n(n+1)/2$ and $n(n-1)/2$ independent parameters so that $\overline{F}$ as well as $H = S + iA$ has n^2 independent elements.]

The curvatures of the "quantum Killing trajectories" can best be evaluated if we recast Eq. (47) in the form of Eq. (11). To this end we first replace the time parameter in Eq. (47) by an arc length parameter. Thus the unit tangent vector to the curve, $x^{(1)}(s) = \dot{x} = (dx/dt)\,dt/ds$, satisfies

$$x^{(1)} = Fx, \tag{54}$$

where

$$F \equiv \alpha\overline{F} \equiv \frac{dt}{ds}\,\overline{F} \tag{55}$$

and

$$1 = x^{(1)T}x^{(1)} = x^T F^T F x = -\alpha^2 x^T \overline{F}^2 x. \tag{56}$$

Since $\overline{F} = -\overline{F}^T = \text{const}$, it follows readily from Eqs. (54) and (55) that $\alpha = 1/\sqrt{-x^T\overline{F}^2 x}$ is constant. Successive differentiation of both sides of Eq. (54) with respect to s leads first to $x^{(2)} = Fx^{(1)}$ or Eq. (11) and then, in general, to Eq. (12).

From Eqs. (15), (54), (55), and (56) we obtain

$$\lambda_k \equiv x^{(1)T} F^{2k-2} x^{(1)} = -\alpha^{2k} x^T \overline{F}^{2k} x$$

$$= (-1)^{k+1}\,\frac{x^T\overline{F}^{2k}x}{[x^T\overline{F}^2 x]^k} = \text{const}. \tag{57}$$

Since λ_k is a ratio of scalars, it can be evaluated in any frame and in particular in the representation where H is diagonal. Then $\overline{F}$ simplifies to

$$\overline{F} = \left[\begin{array}{c|c} \bigcirc & E_1 \cdot \cdot \cdot E_n \\ \hline -E_1 \cdot \cdot \cdot -E_n & \bigcirc \end{array}\right], \tag{58}$$

where E_i is the ith eigenvalue of H corresponding to an eigenvector whose weight here is given by the coefficient $c_i = a_i + ib_i$. It is simple to see that $\overline{F}^2$, and hence $\overline{F}^{2k}$, is a diagonal matrix and that

$$\lambda_k = (-1)^{k+1}\sum_{i=1}^{n}(a_i^2 + b_i^2)\,E_i^{2k} \Big/ \left[\sum_{j=1}^{n}(a_j^2 + b_j^2)\,E_j^2\right]^k, \tag{59}$$

where $\sum_{i=1}^{n}(a_i^2 + b_i^2) = 1$. We can replace the a_i's and b_i's with their initial values a_{i0} and b_{i0} because in the energy representation $c_i = c_{i0}e^{-iE_i t}$ and this in turn implies that $(a_i^2 + b_i^2) = (a_{i0}^2 + b_{i0}^2)$.

A few observations are in order here. First $\lambda_1 = 1$ as expected. Second, if our system is prepared in either an energy eigenstate or a single set of degenerate energy eigenstates or if the system is completely degenerate, then $\lambda_k = (-1)^{k+1}$. Now, in general, we have from Eqs. (29) and (30) that $\kappa_1^2 = -\lambda_2/\lambda_1$ while $\kappa_2^2 = \lambda_2 - \lambda_1\lambda_3/\lambda_2$. Thus for any of the above cases $\kappa_1^2 = 1$ and $\kappa_2^2 = 0$. This corresponds to a curve that not only lies on the unit hypersphere, but whose radius of curvature is itself one. The trajectory is then a geodesic or "great circle" of the unit hypersphere. Again, the question as to whether the general trajectories are closed or not is best answered by considering the energy representation where $c_i = c_{i0}e^{-iE_i t}$. It is evident that c, and hence x, returns to its original position if and only if the energy eigenstates satisfy $\sum_{i=1}^{n} n_i E_i = m$ where $\{n_i; i=1,\ldots,n\}$ and m are integers.

Finally, if the system initially has equal probability of being in any one of its energy eigenstates, i.e., if $c_{j0}^* c_{j0} = a_{j0}^2 + b_{j0}^2 = 1/n$ for $j=1,\ldots,n$, then it follows directly that

$$\lambda_k = (-n)^{k-1}\sum_{i=1}^{n} E_i^{2k} \Big/ \left[\sum_{j=1}^{n} E_j^2\right]^k. \tag{60}$$

A simple example of the foregoing is any two state system for which $c_j = c_{j0}e^{-iE_j t} = (1/\sqrt{2})\,e^{-i(E_j t + \theta_j)}$, $j=1,2$, where θ_j is an arbitrary constant phase. The curve is described by

$$x = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos(E_1 t + \theta_1) \\ \cos(E_2 t + \theta_2) \\ -\sin(E_1 t + \theta_1) \\ -\sin(E_2 t + \theta_2) \end{bmatrix}. \tag{61}$$

A straightforward application of Eqs. (30), (29), and (60) leads to

$$\kappa_1 = \frac{[2(E_1^4 + E_2^4)]^{1/2}}{E_1^2 + E_2^2}, \quad \kappa_2 = \frac{E_1 E_2\,|E_1^2 - E_2^2|}{E_1^2 + E_2^2}\left(\frac{2}{E_1^4 + E_2^4}\right)^{1/2},$$

$$\kappa_3 = E_1 E_2\left(\frac{2}{E_1^4 + E_2^4}\right)^{1/2}. \tag{62}$$

Aside from a normalization factor of $\sqrt{2}$, Eq. (62) coincides with the results that Gluck obtains in a particular demonstration of his algorithm for computing curvatures in a Euclidean space.[6] Note that if $E_1 = E_2$, then $\kappa_1 = 1$ while $\kappa_2 = 0$. However, the expression for κ_3 is not valid in this case.

ACKNOWLEDGMENTS

It is a pleasure to thank Dr. T. Callahan, Dr. H.C. Corben, Dr. M. Melvin, and Dr. J. P. Wilker for stim-

ulating discussions. To Dr. E. L. Schucking I owe the suggestion of studying quantum mechanics from an intrinsic geometrical viewpoint. Finally, I wish to thank the National Research Council of Canada for financial assistance and Dr. H. C. Corben and Dr. R. C. Roeder for their general forbearance during the period of time spent working on this paper.

[1]The intrinsic geometrical approach of describing a curve, i.e., the Frenet–Serret formalism, has been used in physics by: (a) J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960) in refrerence to a number of problems such as the bouncing photon, falling apple, ballistic suicide problem; (b) F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, Mass., 1965), pp. 173ff., in reference to ordinary 3-space trajectories of relativistic charged particles in electromagnetic fields; (c) E. Honig, E. L. Schucking, and C. V. Vishveshwara, "Motion of charged particles in homogeneous electromagnetic fields," J. Math. Phys. **15**, 774 (1974). We also note two recent articles by J. Krause, Int. J. Theor. Phys. **11**, 273 (1974) and **12**, 35 (1975).

[2]D. Laugwitz, *Differential and Riemannian Geometry* (Academic, New York, 1965), pp. 120ff.

[3]F. R. Gantmacher, *Theory of Matrices, Vol. I* (Chelsea, New York, 1959, 1960). See also N. Bourbaki, *Éléments de mathématique, Fasc. XI, Algèbre* (Hermann, Paris, 1967), Chap. 4, pp. 66–67. In the older literature, Hankel determinants were sometimes referred to as orthosymmetrical determinants. See, for example, R. F. Scott, *Theory of Determinants and Their Applications*, revised by G. B. Mathews (Cambridge U. P., Cambridge, 1904), p. 99. In the more recent literature, they are also referred to as persymmetric determinants. See A. C. Aitken, *Determinants and Matrices* (Oliver & Boyd, Edinburgh & London; Interscience, New York, 1958), pp. 121ff.

[4]E. C. G. Stueckelberg, Helv. Physica Acta **33**, 727 (1960); E. C. G. Stueckelberg and M. Guenin, Helv. Physica Acta **34**, 621 (1961); E. C. G. Stueckelberg, M. Guenin, C. Piron, and H. Ruegg, Helv. Physica Acta **34**, 675 (1961); E. C. G. Stueckelberg and M. Guenin, Helv. Physica Acta, **35**, 673 (1962); D. Finkelstein, J. Jauch, S. Schiminovich, and D. Speiser, J. Math. Phys. **3**, 207 (1962).

[5]W. Blaschke, Math. Z. **6**, 94 (1920); D. Laugwitz, Ref. 2, pp. 91, 122.

[6]H. Gluck, Amer. Math. Monthly **73**, 699 (1966).

[7]We have replaced R^4 by V^n in this quote.

[8]See, for example, N. Bourbaki, Ref. 3.

[9]D. Finkelstein, J. Jauch, S. Schiminovich, and D. Speiser, Ref. 4.

Radial integrals with finite energy loss for Dirac–Coulomb functions

Krishan Sud, L. E. Wright, and D. S. Onley

Ohio University, Athens, Ohio 45701
(Received 7 June 1976)

Analytic results for radial integrals over products of Dirac–Coulomb functions and the radial part of the electromagnetic Green's function are expressed in terms of a matrix generalization of the gamma function. This matrix gamma function has many useful properties, including a recurrence relation similar to that of the gamma function, and provides a compact easily manipulated method of evaluating the Dirac–Coulomb radial integrals. These results can be used to calculate the virtual and real photon spectra associated with electron scattering from the nucleus.

I. INTRODUCTION

Exact radial integrals of Dirac—Coulomb functions for the case of zero energy loss were given by Reynolds, Onley, and Biedenharn[1] some years ago. These results led to distorted wave calculations of inelastic electron scattering from low-lying discrete levels in the nucleus where the energy transfer to the nucleus could be neglected in comparison with the incident electron energy.[2] These calculations were extended to include energy loss by means of numerical integration,[3] and this calculation has been of major importance in the analysis of inelastic electron scattering from discrete levels during the past ten years. Numerical techniques for calculating the radial integrals, however, are too time consuming and not sufficiently precise for many purposes.

A number of investigators have also given analytic results for the finite energy loss radial integrals.[4–6] Rozics and Johnson[4] showed that the radial integrals can be written in terms of a Lauricella hypergeometric function, a triple infinite series, and gave analytic continuations of the Lauricella function which can be used for low energy electrons. The authors of Refs. 5 and 6 write the integrals in terms of a finite series of Appell F_2 functions which are doubly infinite series. They give analytic continuations which allow one to calculate the radial matrix elements at higher energies, but which have difficulty in evaluating the matrix elements corresponding to the emission of higher multipole photons. In this paper we will make use of the Appell functions of Ref. 6, but we will stress the recurrence relations among the various radial integrals by use of matrix techniques which will permit the evaluation of the higher multipole matrix elements.

The Dirac—Coulomb radial functions obey coupled first-order differential equations. Prewitt and Wright[7] were able to make explicit use of the differential equations to evaluate the zero energy loss integrals for electron scattering from magnetic and electric nuclear moments from some radius R to infinity in terms of an asymptotic matrix series. This allows one to restrict the numerical integration to the nuclear volume. Onley[8] has generalized this concept and investigated the matrix solutions of first-order differential equations, and integrals over those solutions.

In this paper, we will make use of these techniques to obtain compact expressions for the Dirac—Coulomb functions, products of Dirac—Coulomb functions, integrals over products of Dirac—Coulomb functions, and the radial part of the electromagnetic Green's function for finite energy transfer. By use of these compact easily manipulated results, we can reduce the number of basic integrals that are required to a minimum. The basic integrals that are required will be evaluated by various analytic continuation techniques, including the ones given by Gargaro and Onley.[6]

In Sec. II we will discuss the matrix form of the power and asymptotic series expansions of the Dirac—Coulomb functions. The results here correspond to standard series expansions for Whittaker functions, but the matrix techniques provide a compact method of carrying out these expansions, and of transforming from one expansion to another. Furthermore, we obtain a single series (albeit with matrix coefficients) for products of Whittaker functions.

In Sec. III we will define and discuss the point integrals over the Dirac—Coulomb functions. Point integrals are radial integrals of products of Dirac—Coulomb wavefunctions over the full range of the radial coordinate with suitable subtractions being made for the integrals of irregular functions. These can be related to a generalized matrix gamma function whose elements are in the form of the Appell function[9] F_2 for the power series expansions, and the Appell function[9] F_3 for the asymptotic series expansions. We also give an expansion of the Appell F_2 function in terms of four of Appell's F_3 functions in addition to the analytic continuations of the Appell F_2 function given by Gargaro and Onley.[6] Either of these continuations permits the evaluation of the matrix gamma function.

II. DIRAC-COULOMB RADIAL FUNCTIONS

The Dirac equation with a spherically symmetric Coulomb potential can be separated into radial and angular functions in the standard way.[10] The radial Dirac—Coulomb functions for a lepton of energy E and rest mass m satisfy the following coupled first-order equation if we take the nucleus to be a point charge:

$$\frac{dU(r)}{dr} = \left(\frac{A}{r} - B\right) U(r), \tag{1}$$

where the constant 2×2 matrices A and B corresponding to the standard form of this equation are given by,

$$A=\begin{pmatrix}-\kappa & \alpha Z\\ -\alpha Z & \kappa\end{pmatrix} \quad \text{and} \quad B=\begin{pmatrix}0 & -(E+m)\\ (E-m) & 0\end{pmatrix},$$

where κ is the eigenvalue of the Dirac operator $\underset{\sim}{K}$ $=\beta(\sigma\cdot\mathbf{L}+1)$. The normalized solutions for this particular choice for A and B can be written as

$$U(r)=\begin{pmatrix}v_\kappa^R(r) & v_\kappa^I(r)\\ u_\kappa^R(r) & u_\kappa^I(r)\end{pmatrix}, \tag{2}$$

where the regular radial functions v_κ^R and u_κ^R are given explicitly in terms of the Whittaker function by

$$\begin{Bmatrix}v_\kappa^R(r)\\ u_\kappa^R(r)\end{Bmatrix}= \begin{pmatrix}1\\ -\left(\frac{E-m}{E+m}\right)^{1/2}\end{pmatrix}\frac{\exp(\pi\eta/2)}{p\sqrt{2pr}}\,\frac{|\Gamma(\gamma+i\eta)|}{\Gamma(2\gamma+1)} \times\begin{Bmatrix}\mathrm{Re}\\ \mathrm{Im}\end{Bmatrix}(\gamma+i\eta)\exp[-i(\gamma+1/2)\pi/2] \times\exp(i\eta_\kappa)M_{-1/2-i\eta,\gamma}(2ipr).$$

The parameters are $\gamma=(\kappa^2-\alpha^2Z^2)^{1/2}$, $\eta=\alpha ZE/p$, $p=(E^2-m^2)^{1/2}$, and η_κ, the Coulomb phase shift, given explicitly by

$$\eta_\kappa(\gamma)=-\frac{\pi}{2}\,\frac{1+S_\kappa}{2}-\frac{1}{2}\arctan\left(\frac{\eta(\kappa+\gamma m/E)}{\kappa\gamma-\eta^2m/E}\right)$$

with $\beta=\alpha Zm/p$ and $S_\kappa=\kappa/|\kappa|$. A second pair of solutions (irregular solutions labeled by I) are obtained by changing the sign of γ everywhere.

The matrix power series solution to Eq. (1) is most easily obtained in a representation when A is diagonal which we distinguish here by the label (A). Thus, $U^{(A)}(r)=CU(r)$ with $B^{(A)}=CBC^{-1}$, $A^{(A)}=CBC^{-1}$, where the transformation matrix C is given by

$$C=\frac{\gamma-i\eta}{2\gamma(\gamma+i\eta)}\begin{pmatrix}\frac{-i}{\sqrt{E+m}}\left(1-\frac{\kappa-i\beta}{\gamma+i\eta}\right) & \frac{-1}{\sqrt{E+m}}\left(1+\frac{\kappa-i\beta}{\gamma+i\eta}\right)\\ \frac{i}{\sqrt{E+m}}\left(1+\frac{\kappa-i\beta}{\gamma-i\eta}\right) & \frac{1}{\sqrt{E+m}}\left(1-\frac{\kappa-i\beta}{\gamma-i\eta}\right)\end{pmatrix}. \tag{3}$$

The transformed A and B matrices are

$$A^{(A)}=\begin{pmatrix}\gamma & 0\\ 0 & -\gamma\end{pmatrix}, \qquad B^{(A)}=\frac{ip}{\gamma}\begin{pmatrix}-i\eta & \gamma-i\eta\\ \gamma+i\eta & i\eta\end{pmatrix}.$$

The matrix solution to Eq. (1) with A diagonal has been given by Onley[8] and is

$$U_0^{(A)}(r)=\sum_{n=0}^{\infty}V_n r^{n+A}N,$$
$$\{V_n\}_{ij}=\frac{\{-B^{(A)}V_{n-1}\}_{ij}}{a_j-a_i+n}, \quad V_0=I, \tag{4}$$

where $A=\mathrm{diag}(a_1,a_2,\cdots)$. The notation r^A means $\mathrm{diag}(r^{a_1},r^{a_2},\cdots)$; more generally if A and B are diagonal matrices, the notation B^A means $\mathrm{diag}(b_1^{a_1},b_2^{a_2},\cdots)$. The subscript 0 on a function is used to distinguish a power series expansion from an asymptotic expansion which carried a subscript ∞. The elements of $U_0^{(A)}(r)$ can be identified in terms of the well-known Whittaker function as follows:

$$U_0^{(A)}(r)=\begin{pmatrix}M_{-i\eta,\gamma-1/2}(2ipr) & \frac{\gamma-i\eta}{2\gamma(2\gamma-1)}M_{-i\eta,-\gamma+1/2}(2ipr)\\ \frac{-(\gamma+i\eta)}{2\gamma(2\gamma+1)}M_{-i\eta,\gamma+1/2}(2ipr) & M_{-i\eta,-\gamma-1/2}(2ipr)\end{pmatrix}\begin{pmatrix}N(\gamma) & 0\\ 0 & N(-\gamma)\end{pmatrix}, \tag{5}$$

where

$$N(\gamma)=-\frac{\exp[i\eta_\kappa(\gamma)]\,i^{1-\gamma}\exp(\pi\eta/2)(\gamma-i\eta)}{p\sqrt{E+m}\,(\gamma+i\eta)}\,\frac{|\Gamma(\gamma+i\eta)|}{\Gamma(2\gamma+1)}.$$

The matrix power series corresponds to the standard power series solutions for point regular and irregular Dirac–Coulomb functions, but here one calculates the four functions in one series simultaneously.

It is also easy to calculate products of Coulomb functions by increasing the dimensions of our matrix equations and solutions. The solution given in Eq. (4) is not confined to the particular matrices A and B given here; it is valid for any pair of finite matrices of the same size. This can be particularly useful in dealing with products of solutions of the type denoted here by $U(r)$. Suppose we have two such matrix equations distinguished by the suffix $i=1,2$,

$$\frac{d}{dr}U_i(r)=\left(\frac{A_i}{r}-B_i\right)U_i(r).$$

The direct product of such solutions written $W(r)=U_2(r)\otimes U_1(r)$ is a 4×4 matrix, each element of which would be a product of two Whittaker functions corresponding to different parameters. Furthermore, W is a solution of an equation of the same form as Eq. (1) with,

$$A=A_2\otimes I+I\otimes A_1, \quad B=B_2\otimes I+I\otimes B_1. \tag{6}$$

Another relation which we will find useful is

$$r^a\exp(-br)\,W(A,B;r)=W(A+aI,B+bI;r), \tag{7}$$

which is true for any solution with A and B $n\times n$ matrices and a,b scalars [the expression $r^a\exp(-br)$ is simply the solution of Eq. (1) in the case where all the arguments are scalar]. Thus, the integrand of many integrals of physical interest can be generated by a single matrix series.

To assist us in obtaining asymptotic series expansions for the Dirac–Coulomb functions, we examine the power series expansion in a B-diagonal form. Firstly we consider, $U_0^{(B_1)} = GU_0^{(A)}$, where

$$G = \begin{pmatrix} 1 & -\dfrac{\gamma - i\eta}{\gamma + i\eta} \\ 1 & 1 \end{pmatrix}, \quad A^{(B_1)} = \begin{pmatrix} i\eta & \gamma - i\eta \\ \gamma + i\eta & -i\eta \end{pmatrix}, \quad B^{(B_1)} = \begin{pmatrix} -ip & 0 \\ 0 & ip \end{pmatrix},$$

and the normalized solutions are again in terms of Whittaker functions

$$U_0^{(B_1)} = \frac{1}{\sqrt{2ipr}} \begin{pmatrix} M_{-1/2-i\eta,\gamma}(2ipr) & -\left(\dfrac{\gamma - i\eta}{\gamma + i\eta}\right) M_{-1/2-i\eta,-\gamma}(2ipr) \\ M_{1/2-i\eta,\gamma}(2ipr) & M_{1/2-i\eta,-\gamma}(2ipr) \end{pmatrix} \begin{pmatrix} N(\gamma) & 0 \\ 0 & N(-\gamma) \end{pmatrix}. \tag{8}$$

As before, the first column contains the regular Coulomb functions and the second column contains the irregular Coulomb functions. For reference, we note that we could also obtain $U_0^{(B_1)}$ from $U(r)$ by the transformation $U_0^{(B_1)} = DU(r)$, where

$$D = \frac{-1}{(\gamma + i\eta)^2} \begin{pmatrix} \dfrac{i(\gamma - i\eta)}{\sqrt{E+m}} & \dfrac{\gamma - i\eta}{\sqrt{E-m}} \\ \dfrac{-i(\kappa - i\beta)}{\sqrt{E+m}} & \dfrac{\kappa - i\beta}{\sqrt{E-m}} \end{pmatrix}. \tag{9}$$

There is an additional B-diagonal power series solution which is sometimes of interest. It is $U_0^{(B_2)} = HU$, where

$$H = \begin{pmatrix} \dfrac{i}{\sqrt{E+m}} & \dfrac{1}{\sqrt{E-m}} \\ \dfrac{-i}{\sqrt{E+m}} & \dfrac{1}{\sqrt{E-m}} \end{pmatrix}, \quad A^{(B_2)} = \begin{pmatrix} i\eta & \kappa - i\beta \\ \kappa + i\beta & -i\eta \end{pmatrix},$$

and $B^{(B_2)} = B^{(B_1)}$. Since U is real, it follows from the form of H, that $U_0^{(B_2)} = K(U_0^{(B_2)})^*$, where $K = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ interchanges rows.

The asymptotic matrix series solution to Eq. (1) is most easily obtained in B-diagonal form. Onley[8] has shown that it is

$$U_\infty^{(B)} = \sum_{n=0}^{\infty} D_n r^{\bar{A}-n} e^{-Br}, \tag{10}$$

where $\bar{A}$ denotes the diagonal elements of A, and if $\tilde{\bar{A}} = A - \bar{A}$, then for $i \neq j$

$$\{D_n\}_{ij} = \frac{\{\tilde{\bar{A}} + n - 1)D_{n-1}\}_{ij} - (A_{ii} - A_{jj})\{D_{n-1}\}_{ij}}{b_i - b_j},$$

and for $i = j$

$$\{D_n\}_{ii} = \frac{1}{n} \sum_{k \neq i} A_{ik}(D_n)_{ki},$$

where b_i are the diagonal elements of B and $D_0 = I$. The elements of $U_\infty^{(B_1)}$ can be identified explicitly in terms of Whittaker functions of the second kind by

$$U_\infty^{(B_1)} = r^{-1/2} \begin{pmatrix} (-2ip)^{-1/2-i\eta} W_{1/2+i\eta,\gamma}(-2ipr) & -(2ip)^{-1/2+i\eta}(\gamma - i\eta) W_{-1/2-i\eta,\gamma}(2ipr) \\ -(-2ip)^{-1/2-i\eta}(\gamma + i\eta) W_{-1/2+i\eta,\gamma}(-2ip) & (2ip)^{-1/2+i\eta} W_{1/2-i\eta,\gamma}(2ipr) \end{pmatrix}.$$

Note that this solution has the interesting symmetry property that $U_\infty^{(B_1)} = KU_\infty^{(B_1)*}K$. Just as in the power series case, products of functions are given by the same asymptotic matrix series for higher dimensional A and B matrices.

Since $U_0^{(B_1)}$ and $U_\infty^{(B_1)}$ both represent the general solution to the differential equation, we must have $U_0^{(B_1)}N^{-1} = U_\infty^{(B_1)}T$, where T is a constant matrix and we have removed the normalization matrix of Eq. (8) for convenience. We can find T for the Dirac–Coulomb functions by using the relationships between the power series and asymptotic series Whittaker functions.[11] We find

$$T = \begin{pmatrix} \dfrac{\Gamma(2\gamma+1)(2ip)^{i\eta}}{\Gamma(\gamma+1+i\eta)} & \dfrac{\Gamma(-2\gamma+1)(2ip)^{i\eta}}{(\gamma+i\eta)\,\Gamma(-\gamma+i\eta)} \\ \dfrac{\Gamma(2\gamma+1)(-2ip)^{-i\eta}}{\Gamma(1+\gamma-i\eta)} \exp(i\pi\gamma) & \dfrac{\Gamma(-2\gamma+1)(-2ip)^{-i\eta}}{\Gamma(1-\gamma-i\eta)} \exp(-i\pi\gamma) \end{pmatrix}. \tag{11}$$

This result provides a compact expression for obtaining the regular Whittaker functions in terms of the asymptotic functions.

III. INTEGRALS OVER DIRAC-COULOMB FUNCTIONS

We now turn to the problem of integrals over products of Dirac–Coulomb wavefunctions and the radial part of the electromagnetic Green's function, which is $j_L(wr)$ or $h_L^{(1)}(wr)$ corresponding to emission of real or virtual photons, respectively.[5] We will discuss this in general terms before turning to the explicit problem of electron scattering.

Following Onley,[8] we define an integral operator $S(A,B;R)$ by

$$S(A,B;R)\,W(A,B;R)=\int_R^\infty W(A,B;r)\,dr, \tag{12}$$

where $W(A,B;r)$ is a solution of the first-order matrix differential equation Eq. (1), and we confine ourselves to to cases where the integral is convergent at the upper limit. The operator $S(A,B;r)$ satisfies the inhomogeneous equation

$$\frac{dS}{dr}+S\left(\frac{A}{r}-B\right)=-I, \tag{13}$$

and has the particular power series solution

$$P(A,B;r)=-(A+1)^{-1}r-(A+1)^{-1}B(A+2)^{-1}r^2 -(A+1)^{-1}B(A+2)^{-1}B(A+3)^{-1}r^3\cdots \tag{14}$$

as long as $A+n$ is not singular. The solution to the homogeneous equation obtained from Eq. (13) by removing the term I is $W(A,B;r)^{-1}$, thus the general solution for S is of the form

$$S(A,B;r)=\Gamma W(A,B;r)^{-1}+P(A,B;r), \tag{15}$$

where Γ is a constant matrix. Combining Eqs. (12) and (15), we can write

$$\Gamma=\int_R^\infty W(A,B;r)\,dr-P(A,B;R)W(A,B;R). \tag{16}$$

In Eq. (16), the left-hand side does not depend on R although this variable appears explicitly on the right. In fact R can take any value for which the functions on the right are defined. In particular we take the limit $R\to 0$ and define the subtracted integral as follows:

$$\int_{(0)}^\infty W(A,B;r)\,dr=\lim_{R\to 0}\left[\int_R^\infty W(A,B;r)\,dr -P(A,B;R)W(A,B;R)\right]. \tag{17}$$

Clearly if W is a regular function of r at the origin, the product PW on the right side of Eq. (17) vanishes in the limit, and the integral is simply the conventional integral with lower limit zero. If W is irregular, however, Eq. (17) still gives a finite result since the functions on the right are defined for any nonzero R and the combination is independent of R. This process is related to the subtracted integral definitions used with scalar functions—compare the definition of the conventional gamma function for negative arguments given in Ref. 12 on p. 243. The choice of the letter Γ for the definition of Eq. (16) is not entirely arbitrary. If we display the matrix arguments A and B, on which the function depends, as follows:

$$\Gamma(A+1,B)=\int_{(0)}^\infty W(A,B;r)\,dr, \tag{18}$$

this function is related to the conventional gamma function with a matrix argument. In fact, when B is the unit matrix $\Gamma(A,I)=\Gamma(A)$. The recurrence relation for the gamma function also holds in a modified form, namely

$$A\Gamma(A,B)=B\Gamma(A+1,B), \tag{19}$$

and we will have occasion to use this later. Relation (19) is readily obtained by integration by parts and use of Eq. (7).

In the case of electron scattering from a finite nucleus we require integrals over products of solutions of the the Dirac equation for an electron moving in the electric potential of a distributed charge. Such solutions and their integrals we suppose are necessarily found by standard numerical methods of integration starting at the origin (where the wavefunctions are regular) and continued to an approximate radius R beyond which the density of nuclear material is negligible. The integrals, however, do not cut off at this point but rather assume the standard form discussed here, the component wavefunctions being simply solutions in the field of a point charge. In general we may expect to calculate an integrating operator S, such as that defined in Eq. (12), which when operating on the wavefunctions at the nuclear boundary radius R, produces the integral of the operand from that point to infinity. In general the difficulty in evaluating such an operator is in finding a suitable expression for the function Γ.

The radial integrals relevant to the electron scattering problem can be obtained from linear combinations of the elements of

$$RI=\int_{(0)}^\infty U(2)^*\otimes U(1)\begin{Bmatrix}h_L^1(wr)\\ j_L(wr)\end{Bmatrix}dr, \tag{20}$$

where L is related to the photon multipolarity and the spherical Hankel function corresponds to emission of a virtual photon of energy w and the spherical Bessel function corresponds to the emission of a real photon of energy w. If we transform the electron wavefunction $U(1)(U(2))$ corresponding to the initial (final) state to the standard representation given in Eq. (2), then we can obtain the Bessel integral from the Hankel integral by taking the real part of RI, since in the standard representation the wavefunctions are real. We further note that the spherical Hankel function can be written as an exponential times a finite series expansion in inverse powers of the radial coordinate r,

$$h_L^{(1)}(wr)=\exp(iwr)\sum_{n=1}^{L+1}a_n(L)r^{-n}, \tag{21}$$

where

$$a_n(L)=\frac{2\Gamma(L+n)i^{n-L+2}}{\Gamma(n)\,\Gamma(2+L-n)(2w)^n}.$$

This allows us to use the general properties of the solutions to Eq. (1), given in Eq. (7), to write the integrals

RI in terms of a finite series of matrix gamma functions with the A matrix differing by integers. Furthermore, use of the recurrence relation for the matrix gamma function given in Eq. (19) allows us to write the integral RI as

$$RI = X\Gamma(A, B - iw), \tag{22}$$

where A and B are 4×4 matrices arising from the direct products of the initial and final electron wavefunctions, and X is a 4×4 matrix which depends on A, B, and L, the multipolarity of the photon emitted. The matrix X is given explicity by

$$X = a_1(L)I + \sum_{n=2}^{L+1} a_n(L) \prod_{m=1}^{n-1} \{(A-m)^{-1}B\}, \tag{23}$$

where the successive terms in the matrix product multiply from the left.

Thus, radial integrals for all multipoles L can be generated from one basic matrix gamma function, $\Gamma(A, B - iw)$. This result will allow considerable savings in computation time, particularly for the case of bremsstrahlung where many multipoles contribute. Before discussing how we evaluate this basic matrix gamma function, we note that we have checked the recurrence scheme used above on the computer and when recurring in the direction outlined above it works extremely well.

We can obtain an expression for $\Gamma(A, B - iw)$ by using the power series solution in A-diagonal form for the wavefunctions. Explicitly, we require

$$\Gamma_0(A^{(A)}, B^{(A)} - iw) = \int_{(0)}^{\infty} U_0^{(A)}(2)^* \otimes U_0^{(A)}(1) \frac{\exp(iwr)}{r}\, dr.$$

Defining

$$W(A^{(A)} - 1, B^{(A)}) = U_0^{(A)}(2)^* \otimes U_0^{(A)}(1) \frac{1}{r},$$

and integrating the power series solution for W given in Eq. (4) term by term, we obtain

$$\Gamma_0(A^{(A)}, B^{(A)} - iw) = \sum_{n=0}^{\infty} V_n \Gamma(n + A)(-iw)^{-A-n}, \tag{24}$$

where w is considered to have a small positive imaginary part to ensure convergence of the individual integrals. The coefficients V_n are given by Eq. (4) with $B = B^{(A)}$. Unfortunately this matrix series is not convergent for any set of values of physical interest and we do not know any analytic continuation techniques for matrix series.

If we consider Γ_0 in the B-diagonal form,

$$\Gamma_0(A^{(B_1)}, B^{(B_1)} - iw) = \int_{(0)}^{\infty} U_0^{(B_1)}(2)^* \otimes U_0^{(B_1)}(1) \frac{\exp(iwr)}{r}\, dr, \tag{25}$$

we can integrate each individual element of Γ_0. Apart from the normalization matrix N, these elements are given by

$$I_0 = k_1^{(b_1-1)/2} k_2^{(b_2-1)/2} \int_{(0)}^{\infty} r^{\alpha-1} \exp(-\Delta r) \times {}_1F_1(a_2, b_2, k_2 r)\ {}_1F_1(a_1, b_1, k_1 r)\, dr, \tag{26}$$

where $\Delta = i(p_1 - p_2 - w)$, $k_1 = 2ip_1$, and $k_2 = -2ip_2$.

The parameters α, b_1, b_2 are constant for each column of Γ_0. They are given explicitly by the following: column 1, $\alpha = \gamma_1 + \gamma_2$, $b_1 = 2\gamma_1 + 1$, $b_2 = 2\gamma_2 + 1$; column 2, $\alpha = -\gamma_1 + \gamma_2$, $b_1 = -2\gamma_1 + 1$, $b_2 = 2\gamma_2 + 1$; column 3, $\alpha = \gamma_1 - \gamma_2$, $b_1 = 2\gamma_1 + 1$, $b_2 = -2\gamma_2 + 1$; column 4, $\alpha = -\gamma_1 - \gamma_2$, $b_1 = -2\gamma_1 + 1$, $b_2 = -2\gamma_2 + 1$. The parameters a_1 and a_2 vary within the columns in a regular pattern. For the elements, i, of the first column of Γ_0, they (1) $a_1 = \gamma_1 + i\eta_1 + 1$, $a_2 = \gamma_2 - i\eta_2 + 1$; (2) $a_1 = \gamma_1 + i\eta_1$, $a_2 = \gamma_2 - i\eta_2 + 1$; (3) $a_1 = \gamma_1 + i\eta_1 + 1$, $a_2 = \gamma_2 - i\eta_2$; (4) $a_1 = \gamma_1 + i\eta_1$, $a_2 = \gamma_2 - i\eta_2$. The parameters a_1 and a_2 for the elements of the remaining columns are given by the same expressions with the signs of γ_1 and γ_2 changed as in columns 2, 3, and 4 above. By giving Δ a small positive real part ϵ to insure convergence at infinity (e.g., let $p_1 = p_1 - i\epsilon$), and using our subtracted definition of the integral when necessary, we can expand the hypergeometric series and integrate term by term to obtain

$$I_0 = \frac{\Gamma(\alpha)}{\Delta^{\alpha}} k_1^{(b_1-1)/2} k_2^{(b_2-1)/2} F_2(\alpha, a_2, b_2, a_1, b_1; x, y), \tag{27}$$

where $x = k_2/\Delta$, $y = k_1/\Delta$, and F_2 is the double hypergeometric series of Appell. Explicitly,

$$F_2(\alpha, \beta, \beta', \gamma, \gamma'; x, y) = \sum_{m,n} \frac{(\alpha)_{m+n}(\beta)_m(\beta')_n}{(\gamma)_m(\gamma')_n m!\, n!} x^m y^n,$$

which is absolutely convergent for $|x| + |y| < 1$. For the physical variables of interest this series does not converge, but has been analytically continued by Gargaro and Onley[6] to convergent series and used to evaluate electron radial integrals for point charge wavefunctions which would correspond to the first column of the matrix Γ_0. We note here that the same analytic continuations can be used for evaluating the integrals (in the subtracted sense) in the second, third, and fourth columns of Γ_0 which contain the irregular wavefunctions. This analytic continuation is given explicitly by $F_2(\alpha, \beta, \beta', \gamma, \gamma'; x, y) = Q_1 + Q_2 + Q_3$, where the three double series are

$$Q_1 = \frac{\Gamma(\gamma')\Gamma(\beta' - \alpha)}{\Gamma(\beta')\Gamma(\gamma' - \alpha)} (-y)^{-\alpha} \times \sum_{m,n} \frac{(\alpha)_m(1 - \gamma' + \alpha)_m(\alpha + m)_n(1 - \gamma' + \alpha + m)_n}{(1 - \beta' + \alpha)_m(1 - \beta' + \alpha + m)_n(\gamma)_n\, m!\, n!} \times \left(\frac{1}{y}\right)^m \left(\frac{-x}{y}\right)^n,$$

$$Q_2 = \frac{\Gamma(\gamma')\Gamma(\alpha - \beta')\Gamma(\gamma)\Gamma(\beta - \alpha + \beta')}{\Gamma(\alpha)\Gamma(\gamma' - \beta')\Gamma(\beta)\Gamma(\gamma - \alpha + \beta')} \left(\frac{x}{y}\right)^{\beta'} (-x)^{-\alpha} \times \sum_{m,n} \frac{(\beta')_n(1 - \gamma' + \beta')_n(\alpha - \beta')_m(\alpha + 1 - \gamma - \beta')_m(\beta - \alpha + \beta' - m)_n}{(1 - \alpha + \beta' - m)_n(\gamma + \beta' - \alpha - m)_n(1 + \alpha - \beta - \beta')_m\, m!\, n!} \times \left(\frac{1}{x}\right)^m \left(\frac{-x}{y}\right)^n, \tag{28}$$

$$Q_3 = \frac{\Gamma(\gamma')\Gamma(\gamma)\Gamma(\alpha - \beta' - \beta)}{\Gamma(\alpha)\Gamma(\gamma' - \beta')\Gamma(\gamma - \beta)} (-y)^{-\beta'} (-x)^{-\beta} \times F_3(\beta, \beta', 1 - \gamma + \beta, 1 - \gamma' + \beta', 1 + \beta + \beta' - \alpha;\ 1/x,\ 1/y),$$

where Q_3 is written in terms of Appell's hypergeometric function F_3 which is absolutely convergent for $|1/x| < 1$

and $|1/y|<1$. The series Q_1 and Q_2 are absolutely convergent for $|1/y|+|x/y|<1$. These convergence conditions are met for the physical values of x and y in the electron scattering problem.

An additional analytic continuation of Appell's F_2 function, and hence of Γ_0, can be obtained by use of the asymptotic series expansions for the wavefunctions given in Eq. (9). Defining

$$\Gamma_\infty(A, B-iw)=\int_{(0)}^{\infty} U_\infty^{(B_1)}(2)^* \otimes U_\infty^{(B_1)}(1)\frac{\exp(iwr)}{r}\,dr \tag{29}$$

and substituting the matrix series expansion given in Eq. (10), we can integrate term by term if the integral converges at infinity (i.e., the elements of B all contain positive real parts) to obtain

$$\Gamma_\infty(A, B-iw)=\sum_{n=0}^{\infty} D_n\,\Gamma(\bar{A}-n)(B-iw)^{-\bar{A}+n}, \tag{30}$$

where D_n is given in terms of $(A-1)$ and $(B-iw)$ following Eq. (10). Since B and $\bar{A}$ are both diagonal $B^{\bar{A}}$ is defined [see note following Eq. (4)].

In order to investigate the convergence of this matrix series, first given in Ref. (8), we examine the individual elements of Γ_∞. A typical element is of the form

$$I_\infty=\int_{(0)}^{\infty}\exp(-\Delta r)\,r^{\gamma'-1}\;{}_2F_0(a_2,b_2,1/k_2r)\;{}_2F_0(a_1,b_1,1/k_1r)\,dr \tag{31}$$

where k_1, k_2, and Δ are different for each column. The values for the four columns are: $\Delta_1=-i(p_1-p_2+w)$, $(k_1)_1=2ip_1$, $(k_2)_1=-2ip_2$; $\Delta_2=i(p_1+p_2-w)$, $(k_1)_2=-2ip_1$, $(k_2)_2=-2ip_2$; $\Delta_3=-i(p_1+p_2+w)$, $(k_1)_3=2ip_1$, $(k_2)_3=2ip_2$; and $\Delta_4=i(p_1-p_2-w)$, $(k_1)_4=-2ip_1$, $(k_2)_4=2ip_2$; where the subscript refers to the column of Γ_∞. The other parameters $a_1b_1a_2b_2\gamma'$ in Eq. (31) vary with the elements within a column in Γ_∞ but their magnitude has no special significance for the present argument as long as γ' is not a negative integer. In order to have the integrals convergent at infinity, it is necessary to choose $p_1=p_1+i\epsilon$ for the first and third columns, and $p_1=p_1-i\epsilon$ for the second and fourth columns. Doing so we can integrate, in the subtracted sense, the asymptotic expansions of the wavefunctions in Eq. (31) term by term to obtain

$$I_\infty=\frac{\Gamma(\gamma')}{\Delta_i^{\gamma'}}\,F_3(a_1,a_2,b_1,b_2,1-\gamma';x_i,y_i), \tag{32}$$

where the Appell function F_3 is given explicitly by

$$F_3(a_1,a_2,b_1,b_2,\gamma;x,y)=\sum_{m,n}\frac{(a_1)_m(a_2)_n(b_1)_m(b_2)_n}{(\gamma)_{m+n}\,m!\,n!}\,x^m y^n \tag{33}$$

and is absolutely convergent for $|x|<1$, $|y|<1$. The variables in Eq. (32) are $x_i=-\Delta_i/(k_1)_i$, $y_i=-\Delta_i/(k_2)_i$, where i labels the column of Γ_∞. This 4×4 collection of F_3 functions corresponds exactly to the matrix series for Γ_∞ given in Eq. (30). Thus we have obtained a single matrix series expansion which could be used to calculate Appell's F_3 function in lieu of Eq. (33). When we examine the convergence of the individual elements of Γ_∞ we find that some columns are convergent, but that one or two columns are not convergent depending on the value of w, the energy lost by the electron.

Fortunately, the matrix series, Eq. (30), can be used to calculate any column of Γ_∞ independently, while the remaining columns can be evaluated by other means. We have derived the required analytic continuations for the F_3 functions which will permit the evaluation of all the nonconvergent columns,[13] but will not give them here.

Given that we can calculate $\Gamma_\infty(A^{(B_1)}, B^{(B_1)}-iw)$, we require the first column of $\Gamma_0(A^{(B_1)}, B^{(B_1)}-iw)$ for the point nucleus case. For the case of the wavefunctions we gave a relation between the power series expansion and the asymptotic series expansion in Eq. (11) which can be used to relate direct products of wavefunctions times scalar factors. This relation, however, *cannot* be used to relate the integrands of Γ_0 and Γ_∞ given in Eqs. (25) and (29) since it does not hold when the necessary convergence factors, ϵ, are introduced to make the integrals convergent at infinity. We have, however, been able to find an analogous transformation relating Γ_0 and Γ_∞ by investigating the relationship between Appell's F_2 and F_3 functions. Appell's F_2 function can be written in terms of four of Appell's F_3 functions,[13]

$$\begin{aligned}
&\frac{1}{\Gamma(b_1)\,\Gamma(b_2)\,\Gamma(\alpha-a_1-a_2)}\,F_2(1+a_1+a_2-\alpha,a_1,a_2,b_1,b_2;x,y)\\
&\quad=AF_3\left(a_1,a_2,a_1+1-b_1,a_2+1-b_2,\alpha;\frac{1}{x},\frac{1}{y}\right)\\
&\qquad+BF_3\left(a_1,b_2-a_2,a_1+1-b_1,1-a_2,\right.\\
&\qquad\left.b_2-2a_2+\alpha;\frac{1-y}{x},\frac{y-1}{y}\right)\\
&\qquad+CF_3\left(b_1-a_1,a_2,1-a_1,a_2+1-b_2,\right.\\
&\qquad\left.b_1-2a_1+\alpha;\frac{x-1}{x},\frac{1-x}{y}\right)\\
&\qquad+DF_3\left(b_1-a_1,b_2-a_2,1-a_1,1-a_2,\right.\\
&\qquad\left.b_1+b_2-2a_1-2a_2+\alpha;\frac{x+y-1}{x},\frac{x+y-1}{y}\right),
\end{aligned}$$

where

$$A=x^{-a_1}y^{-a_2}\,\frac{1}{\Gamma(b_1-a_1)\,\Gamma(b_2-a_2)\,\Gamma(\alpha)},$$

$$B=(-y)^{a_2-b_2}x^{-a_1}\,\frac{(1-y)^{\alpha-2a_2+b_2-1}}{\Gamma(b_1-a_1)\,\Gamma(a_2)\,\Gamma(b_2-2a_2+\alpha)},$$

$$C=y^{-a_2}(-x)^{a_1-b_1}\,\frac{(1-x)^{\alpha-2a_1+b_1-1}}{\Gamma(a_1)\,\Gamma(b_2-a_2)\,\Gamma(b_1-2a_1+\alpha)},$$

$$D=(-x)^{a_1-b_1}(-y)^{a_2-b_2}\,\frac{(1-x-y)^{\alpha+b_1+b_2-2a_1-2a_2}}{\Gamma(a_1)\,\Gamma(a_2)\,\Gamma(b_1+b_2-2a_1-2a_2+\alpha)}.$$

Each of the Appell F_2 functions which appear in Γ_0 can be expressed as a sum of four F_3 functions which appear in the same row of Γ_∞. This allows us to write

$$\Gamma_0=\Gamma_\infty T' N, \tag{35}$$

where the normalization matrix N is given by the direct product of the incident and final wavefunction normalization matrix given in Eq. (5), and the transformation

matrix T' can be written in terms of the electron variables in the following form:

$$T' = P_2(M_2 \otimes M_1)P_1, \tag{36}$$

where the diagonal matrices P_2 and P_1 have elements

$$P_2 = \mathrm{diag}[\sin\pi(i\eta_2 - i\eta_1),\ \sin\pi(i\eta_1 + i\eta_2),$$
$$-\sin\pi(i\eta_1 + i\eta_2),\ \sin\pi(i\eta_1 - i\eta_2)],$$
$$P_1 = \mathrm{diag}\left[\frac{-\Gamma(2\gamma_2+1)\Gamma(2\gamma_1+1)}{\sin\pi(\gamma_1+\gamma_2)}, \frac{\Gamma(2\gamma_2+1)\Gamma(1-2\gamma_1)}{\sin\pi(\gamma_1-\gamma_2)},\right.$$
$$\left.\frac{\Gamma(1-2\gamma_2)\Gamma(2\gamma_1+1)}{\sin\pi(\gamma_2-\gamma_1)}, \frac{\Gamma(1-2\gamma_2)\Gamma(1-2\gamma_1)}{\sin(\gamma_1+\gamma_2)}\right],$$

and the 2×2 matrices M_2 and M_1 are given by

$$M_2 = \begin{pmatrix} \dfrac{\exp(-i\pi\gamma_2)(2ip_2)^{-i\eta_2}}{\Gamma(1+\gamma_2-i\eta_2)} & \dfrac{\exp(i\pi\gamma_2)(2ip_2)^{-i\eta_2}}{\Gamma(-\gamma_2-i\eta_2)(\gamma_2-i\eta_2)} \\ \dfrac{(-2ip_2)^{i\eta_2}}{\Gamma(1+\gamma_2+i\eta_2)} & \dfrac{(-2ip_2)^{i\eta_2}}{\Gamma(1-\gamma_2+i\eta_2)} \end{pmatrix},$$

$$M_1 = \begin{pmatrix} \dfrac{\exp(-i\pi\gamma_1)(-2ip_1)^{i\eta_1}}{\Gamma(1+\gamma_1+i\eta_1)} & \dfrac{\exp(i\pi\gamma_1)(-2ip_1)^{i\eta_1}}{\Gamma(-\gamma_1+i\eta_1)(\gamma_1+i\eta_1)} \\ \dfrac{(2ip_1)^{-i\eta_1}}{\Gamma(1+\gamma_1-i\eta_1)} & \dfrac{(2ip_1)^{-i\eta_1}}{\Gamma(1-\gamma_1-i\eta_1)} \end{pmatrix}.$$

This result permits an alternate method of calculating the basic integral $\Gamma_0(A, B-iw)$.

The results given in Eqs. (25) and (35) permit the evaluation of Γ_0 in B-diagonal form. To obtain Γ_0 in the standard representation of the wavefunction, see Eq. (2), we simply make use of the matrix D given in Eq. (9) which transforms the wavefunctions from the standard representation to the B-diagonal representation to write

$$\Gamma(A, B-iw) = (D_2^{*-1} \otimes D_1^{-1})\,\Gamma_0(A^{(B_1)}, B^{(B_1)} - iw), \tag{37}$$

where the labels 1 and 2 refer to initial and final states of the electron, and the matrices A, B, $A^{(B_1)}$ and, $B^{(B_1)}$ and 4×4 matrices generated from the 2×2 matrices in the standard or B-diagonal representation by means of Eq. (6), and $\Gamma_0(A^{(B_1)}, B^{(B_1)} - iw)$ is given explicitly in terms of Appell's F_2 function in Eqs. (25) and (27).

IV. CONCLUSIONS

The results presented in Sec. III allows one to calculate the radial integrals required in the analysis of the real and virtual photon spectrum associated with high energy electron scattering from the nucleus in terms of a single matrix function $\Gamma_0(A, B-iw)$ for each value of γ_1, γ_2, and w for any multipole L. The matrix function Γ_0 can be evaluated in terms of analytic continuations of Appell's F_2 series, or in terms of a matrix series coupled with analytic continuations of Appell's F_3 series. For the case of the point nucleus, Appell's F_2 series is simpler to use since all sixteen elements of Γ_∞ are required to calculate just the first column of Γ_0. For the case of finite nuclei where all sixteen elements of Γ_0 are required, the use of Γ_∞ may be preferable.

To summarize, the radial integrals from R to infinity involving the standard form of the radial functions given in Eq. (2), both regular and irregular, are given by the following matrix:

$$\int_R^\infty U_2(r)^* \otimes U_1(r)\frac{\exp(iwr)}{r}\,dr$$
$$= \Gamma_0(A, B-iw) - P(A-1, B-iw; R)$$
$$\times U_2(R)^* \otimes U_1(R)\frac{\exp(iwR)}{R},$$

where expressions for Γ_0 and P are given in Eqs. (37) and (14).

As a test of the practicality of the techniques given above, we have evaluated the point radial integrals needed for calculating the virtual photon spectra for 100 MeV electrons with multipole contributions up to $L=5$. This requires only the evaluation of the first column of Γ_0 which we carried out by using the analytic continuation of Appell's F_2 series given in Eq. (28). Our results agree with the previous results of Gargaro and Onley[6] for the lower multipoles which they calculated and, when the nuclear charge was put equal to unity, reproduced the plane wave Born approximation results to within 1%. Furthermore, the use of the matrix recursion techniques reduced the calculation time, as compared to Gargaro and Onley, by at least a factor of 2. We intend to extend these calculations to the problem of calculating the radiation tail accompanying electron scattering, where many multipoles contribute, and, to include the finite nuclear size effects.

[1] J.T. Reynolds, D.S. Onley, and L.C. Biedenharn, J. Math. Phys. 5, 411 (1964).
[2] T.A. Griffy, D.S. Onley, J.T. Reynolds, and L.C. Biedenharn, Phys. Rev. 128, 833 (1962).
[3] S.T. Tuan, L.E. Wright, and D.S. Onley, Nucl. Instrum. Methods 60, 70 (1968).
[4] J.D. Rozics and W.R. Johnson, Phys. Rev. 135, B56 (1964).
[5] W.W. Gargaro and D.S. Onley, J. Math. Phys. 11, 1191 (1970).
[6] W.W. Gargaro and D.S. Onley, Phys. Rev. C 4, 1032 (1971).
[7] J.F. Prewitt and L.E. Wright, Phys. Rev. C 9, 2033 (1974).
[8] D.S. Onley, *Nuclear Structure Studies Using Electron Scattering and Photoreaction* (Tohoku University, Sendai, Japan, 1972).
[9] P. Appell and J. Kampé de Fériet, *Fonctions Hypergéométriques et Hyperspheriques* (Gauthier-Villars, Paris, 1926).
[10] M.E. Rose, *Relativistic Electron Theory* (Wiley, New York, 1961).
[11] L.J. Slater, *Confluent Hypergeometric Functions* (Cambridge U.P., Cambridge, 1960).
[12] E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis* (Cambridge, U.P., Cambridge, 1927).
[13] Krishan Sud and L.E. Wright, J. Math. Phys. 17, 1719 (1976).

The frequency dependent electrical conductivity for disordered alloys: Application of an abstract Hilbert space generalization of Feenberg's perturbation theory*

James G. Berryman† and Samuel P. Bowen‡

Department of Physics, University of Wisconsin, Madison, Wisconsin 53706
(Received 18 July 1975; revised manuscript received 20 April 1976)

An abstract Hilbert space with a particularly convenient scalar product is introduced to permit a generalization of Feenberg's rearrangement method of perturbation theory to be applied to thermal Green's function calculations. This method has the advantage of treating averages (either thermal or configuration) rigorously from the start. Explicit calculations are done for the frequency dependent electrical conductivity for alloys with diagonal disorder at zero temperature. Three practical approaches are discussed: (1) the Gram–Schmidt orthogonalization procedure, (2) a trick which depends on the Hermitian character of the polarization operator, and (3) a general procedure for using nonorthogonal basis vectors to expand the Feenberg formulas. To second order in the scattering strength, a new expression for the conductivity is found which is valid for all frequencies. This expression agrees with earlier perturbation theory results when the frequency is very small or very large.

I. INTRODUCTION

In an earlier paper[1] (hereafter referred to as I), one of the present authors generalized Feenberg's perturbation theory[2–4] by applying it to the calculation of thermal Green's functions and Kubo formulas. This abstract Hilbert space generalized Feenberg (AHGF) method gives an immediate expression for a self-energy which is tractable, possesses a satisfactory thermodynamic limit, and has a rigorous foundation. Motivation for the method was found in an analysis of the properties of the moment expansion of the thermal Green's functions. A number of brief applications were discussed to demonstrate the breadth and tractability of the method.

The present paper gives a more detailed account of the application of AHGF to the calculation of the frequency dependent conductivity for disordered alloys. This work was first discussed in Ref. 5 which also includes numerical computations of the frequency dependent conductivity using a different generalization of Freenberg's perturbation method. The latter computations will be published elsewhere.

Previous work on the frequency dependent conductivity has been done by several authors. The high frequency conductivity has been studied by Yamada,[6] Lonke and Ron,[7] Velicky and Levin,[8] and Sen.[9] Yamada[6] treats the very low and very high frequency cases as a second-order expansion in the impurity potential. Lonke and Ron[7] calculate the conductivity for a system of free electrons with dilute impurities. Velicky and Levin[8] and Sen[9] study the conductivity using the coherent potential approximation.[10,11] Numerous authors have studied the low frequency and dc conductivity. Much of the earlier work is subsumed by the coherent potential approximation to the conductivity.[12] Further references to the earlier literature may be found in Refs. 5, 12, and 13.

Section II introduces the disordered alloy Hamiltonian, the Green's functions, and conductivity. Section III discusses the alternative choices for scalar products to be used in Sec. IV. The AHGF method is described and motivated in Sec. IV. The canonical form for the conductivity is found and comparison with the classical Drude formula is also included in Sec. IV. Section V gives a very simple application of AHGF and the Gram–Schmidt orthogonalization procedure to the conductivity when the bandwidth is small compared to the scattering strength. Section VI derives a new result valid for all frequencies and correct to $O(\delta^2)$ in the scattering strength. The possibility of expanding the Feenberg formulas in terms of nonorthogonal basis vectors is examined in Sec. VII. The resulting formalism is applied to give the general formulas for AHGF in the Bloch representation in Sec. VIII. Evaluation of the inverse optical effective mass and of more general expectation values required in the analysis are given in Appendices A and B.

II. THE DISORDERED ALLOY CONDUCTIVITY

The disordered alloy Hamiltonian in second quantized form is

$$H = H_0 + \tilde{H} \tag{2.1}$$

with

$$H_0 = \sum_{ijn\sigma} t_{ij}^{(n)} c_{in\sigma}^{\dagger} c_{jn\sigma} \tag{2.2}$$

and

$$\tilde{H} = \sum_{in\sigma} \epsilon_i^{(n)} c_{in\sigma}^{\dagger} c_{in\sigma}, \tag{2.3}$$

where n is the band index, $t_{ij}^{(n)}$ is the transfer (or hopping) matrix element from site i to j—hopping between bands (hybridization) is excluded, $\epsilon_i^{(n)}$ is the energy of an electron at site i, and $c_{in\sigma}^{\dagger}$ ($c_{in\sigma}$) creates (annihilates) an electron with spin σ and band index n at site i. The fermion operators obey the anticommutation relations $\{c_{in\sigma}^{\dagger}, c_{i'n'\sigma'}\} = \delta_{i,i'}\delta_{n,n'}\delta_{\sigma,\sigma'}$ where $\delta_{\alpha,\beta}$ is the Kronecker delta. It is frequently assumed that the state created at site i is the Wannier state associated with the Bloch states of the pure crystal. Consideration can generally be limited to a single band unless one is interested in optical transitions between bands. Furthermore, the

Hamiltonian does not alter the spin so one may ignore the spin index by introducing a factor of 2 in the formulas for physically measurable quantities. The band and spin indices will be suppressed in the future.

Analysis is simplified further if we restrict study to the three-dimensional simple cubic lattice structure with nearest neighbor (n. n.) distance a and n. n. transfer only. Thus, $t_{ij}=0$ unless i and j are n. n. ($t_{ii}=0$ also). In a pure crystal t_{ij} and ϵ_i are each constant independent of the site index.

The most commonly studied model in the literature is that of the binary disordered alloy with diagonal disorder. In this case t_{ij} takes on the same value as that for a pure crystal of A or B (assumed same for both) while ϵ_i equals either ϵ_A or ϵ_B depending on which type of atom occupies site i. Concentrations of A and B are x and y respectively ($x+y=1$). For simplicity we restrict discussion to this case in the work that follows.

When t_{ij} does not depend on the distribution of impurities, the Hamiltonian can be usefully transformed to a wave-vector representation by making the definitions

$$c_{\mathbf{k}}^{\dagger}=N^{-1/2}\sum_i \exp(i\mathbf{k}\cdot\mathbf{R}_i)c_i^{\dagger}. \tag{2.4}$$

Upon inverting the Fourier transform we find

$$c_i^{\dagger}=N^{-1/2}\sum_{\mathbf{k}} \exp(-i\mathbf{k}\cdot\mathbf{R}_i)c_{\mathbf{k}}^{\dagger} \tag{2.5}$$

and H becomes (suppressing the boldface roman vector symbolism)

$$H=\sum_k \epsilon_k c_k^{\dagger}c_k+\sum_{k,q} U(q)c_k^{\dagger}c_{k+q}, \tag{2.6}$$

where

$$U(q)=N^{-1}\sum_i \epsilon_i \exp(iq\cdot R_i), \tag{2.7}$$

$$s(k)=\tfrac{1}{3}(\cos k_x a+\cos k_y a+\cos k_z a), \tag{2.8}$$

$$\epsilon_k=Ws(k) \quad \text{where } W=6t. \tag{2.9}$$

Our studies will be centered around properties of Zubarev-type retarded thermal Green's functions for general operators A and B[14]:

$$G(z)=\langle\langle A;B\rangle\rangle_z=\int_{-\infty}^{\infty}G(t)\exp(izt)\,dt$$
$$=-i\int_0^{\infty}dt\exp(izt)\langle[A(t),B(0)]_{\pm}\rangle, \tag{2.10}$$

$\mathrm{Im}z>0$.

Inverting the transform one finds

$$G(t)=\langle\langle A(t);B(0)\rangle\rangle=\int_{-\infty}^{\infty}\frac{dz}{2\pi}G(z)\exp(-izt)$$
$$=-i\theta(t)\langle[A(t),B(0)]_{\pm}\rangle. \tag{2.11}$$

The double bracket convention is defined by (2.10) and (2.11). The choice of Fourier transform convention and multiplicative constants is uniquely determined by requiring $G(z)$ for single-particle operators to have the same matrix elements as the resolvent operator $(z-H)^{-1}$. $G(z)$ also satisfies the equation of motion (EOM)

$$z\langle\langle A;B\rangle\rangle_z=\langle[A(0),B(0)]_{\pm}\rangle+\langle\langle[A,H];B\rangle\rangle_z$$
$$=\langle[A(0),B(0)]_{\pm}\rangle-\langle\langle A;[B,H]\rangle\rangle_z. \tag{2.12}$$

In these equations $\langle 0\rangle=\mathrm{Tr}\rho 0$, where ρ is the density matrix $\rho=Q^{-1}\exp[-\beta(H-\mu N)]$ for the grand canonical ensemble and $\mathrm{Tr}\rho=1$ determines the partition function Q. $A(t)$ is the Heisenberg operator

$$A(t)=\exp(iHt)A(0)\exp(-iHt), \tag{2.13}$$

Tr is the trace operation, $\beta=(kT)^{-1}$ where T is absolute temperature, and $\theta(t)$ is the Heaviside function. Units are chosen so that $\hbar=1$. And finally $[\ ,\ ]_{\pm}$ is either the commutator or anticommutator depending on whether the operators A and B are bosonlike or fermionlike operators. See Zubarev[14] for a discussion of the proper choice of sign.

Including the configuration average in the definition, we have

$$\langle\langle A;B^{\dagger}\rangle\rangle_{z+i\eta}=-i\int_0^{\infty}dt\exp(izt-\eta t)\langle\mathrm{Tr}\rho[A(t),B^{\dagger}(0)]_{\pm}\rangle_c. \tag{2.14}$$

If we define the operator L by

$$L0=[0,H], \tag{2.15}$$

then it is easy to show that

$$A(t)=\exp(-iLt)A. \tag{2.16}$$

Putting (2.16) into (2.14) we have

$$\langle\langle A;B^{\dagger}\rangle\rangle_z=-i\int_0^{\infty}dt\exp(izt)\langle\mathrm{Tr}\rho[\exp(-iLt)A,B^{\dagger}]_{\pm}\rangle_c$$
$$=\langle\mathrm{Tr}\rho[(z-L)^{-1}A,B^{\dagger}]_{\pm}\rangle_c$$
$$=\langle\mathrm{Tr}\rho[A,(z+L)^{-1}B^{\dagger}]_{\pm}\rangle_c. \tag{2.17}$$

In order to express the conductivity as a Zubarev–Green function we need to define appropriate operators A and B. The relevant operators are the polarization and current operators. The polarization operator is

$$\mathbf{P}=e\sum_i \mathbf{R}_i n_i=e\sum_{ikq}\mathbf{R}_i\exp(iq\cdot R_i)c_k^{\dagger}c_{k+q}. \tag{2.18}$$

The current operator may be calculated from the polarization operator as

$$J_{\mu}=-iLP_{\mu}=ie\sum_{ij}t_{ij}(R_i-R_j)_{\mu}c_i^{\dagger}c_j. \tag{2.19}$$

Given the operators (2.18) and (2.19), we can calculate the frequency dependent conductivity per unit volume according to the formula

$$\sigma_{\mu\nu}(\omega)=-(1/\Omega)\langle\langle J_{\mu};P_{\nu}\rangle\rangle_{\omega+i\eta}. \tag{2.20}$$

III. THE SCALAR PRODUCTS

In this section we present the scalar products needed for the AHGF discussed in Sec. IV. Two different scalar products are needed: one for fermionlike operators (when the anticommutator appears in the Green's function) and one for bosonlike operators (when the commutator appears). The fermion scalar product is easily chosen. However, for the boson operators two choices present themselves and one is free to choose the most convenient.

A scalar product must satisfy three conditions:

$$(A,B)=(B,A)^*, \tag{3.1}$$

$$(A,A)>0 \quad \text{for } A\neq 0, \tag{3.2}$$

$$\left(\sum_i \alpha_i A_i, B\right) = \sum_i \alpha_i^* (A_i, B), \tag{3.3}$$

and

$$\left(A, \sum_i \beta_i B_i\right) = \sum_i \beta_i (A, B_i). \tag{3.4}$$

In addition we require the operator L to be Hermitian,

$$(A, LB) = (LA, B). \tag{3.5}$$

(1) The fermion scalar product has been discussed by Lonke,[15]

$$(A, B) = \mathrm{Tr}\rho\{A^\dagger, B\} = \langle\{A^\dagger, B\}\rangle. \tag{3.6}$$

The single parenthesis notation will be used to denote the fermion scalar product. That conditions (3.1) and (3.3)–(3.5) are satisfied is easily checked. Condition (3.2) follows from

$$(A, A) = \frac{1}{Q}\sum_{l,m}[\exp(-\beta E_l) + \exp(-\beta E_m)]\,|\langle m|A|l\rangle|^2 \geq 0. \tag{3.7}$$

In Eq. (3.7) and throughout this section E_l is the lth eigenvalue of the operator $H - \mu N$ and Q is the grand canonical partition function.

(2) One possible choice of boson scalar product has been discussed by Mori[16]:

$$(A, B)_M = \frac{1}{\beta}\int_0^\infty d\lambda \langle \exp(\lambda H) B \exp(-\lambda H) A^\dagger\rangle. \tag{3.8}$$

The single parenthesis with M subscript notation will be used to denote Mori's boson scalar product. Again (3.1) and (3.3)–(3.5) are easily checked. (3.2) follows because

$$(A, A)_M = \frac{1}{\beta Q}\sum_{lm}\frac{\exp(-\beta E_m) - \exp(-\beta E_l)}{E_l - E_m}\,|\langle m|A^\dagger|l\rangle|^2 \geq 0. \tag{3.9}$$

(3) A second possible choice of boson scalar product has been found,

$$((A, B)) = \langle[[B, H], A^\dagger]\rangle = \langle[LB, A^\dagger]\rangle. \tag{3.10}$$

For reasons that should be apparent, we will call (3.8) the integral scalar product and (3.10) the differential scalar product. The double parenthesis notation will be used to denote the differential scalar product. As before, (3.1) and (3.3)–(3.5) are easily proven. (3.2) follows from

$$((A, A)) = \frac{1}{Q}\sum_{lm}[\exp(-\beta E_l) - \exp(-\beta E_m)] \times (E_m - E_l)\,|\langle m|A|l\rangle|^2 \geq 0. \tag{3.11}$$

The scalar products (3.8) and (3.10) are related in general by

$$((A, B)) = \mathrm{Tr}\rho[A^\dagger, LB] = \beta(\dot{B}^\dagger, \dot{A}^\dagger)_M, \tag{3.12}$$

where $\dot{0} = (d/dt)0$. Notice that, whereas (3.6) and (3.8) give dimensionless scalars when A and B are dimensionless, (3.10) has dimensions of energy ($\hbar = 1$).

An important example of the use of these scalar products is in the evaluation of the optical effective mass (see Appendix A):

$$\mathrm{Tr}\rho[P_\nu, LP_\mu] = ((P_\nu, P_\mu)) = \beta(J_\mu, J_\nu)_M. \tag{3.13}$$

Either of these scalar products may be used in a particular application. We will find the differential scalar product to be the most convenient choice in the present work. This scalar product recommends itself because a direct evaluation in terms of thermal expectation values requires fewer operator transformations than is required when using the integral scalar product. The reason for this is that the exponential operators in the integral scalar product must be transformed away unless we are able to diagonalize the Hamiltonian exactly. Since the differential scalar product is expressed in terms of commutators of the Hamiltonian, these can be evaluated directly in any convenient representation—diagonal or nondiagonal. This fact proves to be a considerable advantage in practical calculations.

Finally we note that the differential scalar product gives zero norm to any operator that commutes with the Hamiltonian. Hence, the norm of an operator as defined by (3.11) is a quantitative measure of the deviation of the operator from its invariant part (the part that commutes with H). The integral scalar product does not automatically have this property. However, Mori has accomplished the same result by restricting the operators in the Hilbert space to have invariant parts set equal to zero. We believe the present approach to be somewhat simpler.

(4) In order to apply the scalar products to the disorder problem, we wish to include the configuration average as well as the thermal average in the scalar product. This generalization may be included with essentially no additional work by defining

$$(A, B) = \langle \mathrm{Tr}\rho\{B, A^\dagger\}\rangle_c \tag{3.14}$$

for fermion operators and

$$((A, B)) = \langle \mathrm{Tr}\rho[LB, A^\dagger]\rangle_c \tag{3.15}$$

for boson operators. Since the configuration average does not cause any spurious interference between configurations in calculating condition (3.2), we find that (3.1)–(3.5) are automatically satisfied for (3.14) and (3.15) since they are satisfied for (3.6) and (3.10). The definitions (3.14) and (3.15) are the ones used in Sec. IV.

IV. AHGF FOR THE CONDUCTIVITY

The resolventlike form of (2.17) leads one to consider the possibility that a thermal Green's function may be viewed in general as a matrix element of $(z \pm L)^{-1}$ in an abstract Hilbert space whose scalar product is temperature dependent. Zwanzig[17] has discussed the operator $(z - L)^{-1}$ and Mori[16] has introduced a Hilbert space similar to the ones to be discussed here. In Sec. III, we discussed the possible choices of scalar product used in fixing our Hilbert space. The scalar products we choose are

$$(B, A) = \langle \mathrm{Tr}\rho\{A, B^\dagger\}\rangle_c \tag{4.1}$$

for fermionlike operators[17] and

$$((B, A)) = \langle \mathrm{Tr}\rho[LA, B^\dagger]\rangle_c \tag{4.2}$$

for bosonlike operators.[1]

Using these definitions we find the Green's functions can be written in convenient and illustrative form as

$$\langle\langle A; B^\dagger\rangle\rangle_z = (B, (z - L)^{-1}A) \tag{4.3}$$

for fermion operators and

$$\langle\langle \dot{a}; B^\dagger\rangle\rangle_z = -i((B, (z - L)^{-1}A)) \tag{4.4}$$

for boson operators where $\dot{0} = d0/dt$. The reader might be concerned that not every Green's function involving bosonlike operators can be expressed in scalar product form (4.4) because of the required time derivative on the left-hand side. However, this question is quickly resolved by noting that a single application of the EOM (2.3) gives

$$\langle\langle A; B^\dagger\rangle\rangle_z = \frac{1}{z}\langle \mathrm{Tr}\rho[A, B^\dagger]\rangle_c + \frac{1}{z}((B, (z - L)^{-1}A)). \tag{4.5}$$

Thus, by using Eqs. (4.3) and (4.5), we can express any Green's function as a matrix element of a resolventlike operator in the abstract Hilbert spaces determined by the scalar products (4.1) and (4.2).

As discussed in Sec. III, the operator L is Hermitian in the Hilbert spaces defined by (4.1) and (4.2). In fact we used this condition to determine the proper choices of scalar product. That L is Hermitian will be useful to us in later work when we attempt to evaluate matrix elements of L in these Hilbert spaces.

Once we establish a Hilbert space, an Hermitian operator L, and a resolvent $(z - L)^{-1}$, I shows that AHGF may be applied to the problem of calculating matrix elements of the resolvent. Matrix elements between different operators A, B as in (4.3) are known as off-diagonal matrix elements of the resolvent; the matrix element of A with itself is a diagonal element. I shows how both diagonal and off-diagonal matrix elements may be expressed using the Feenberg formulas.[2]

Although we only require the diagonal elements in most of our later work, we include both expressions here for completeness. We propose to form the vectors of our Hilbert space from the set of operators A_i so that the A_i's are orthonormal and complete in the Hilbert space of interest. (It is possible to remove the orthogonality requirement under conditions to be discussed in Sec. VII.) Under these conditions I shows that the diagonal element of the resolvent becomes

$$(A_n, (z - L)^{-1}A_n) = (z - \xi_n(z))^{-1}, \tag{4.6}$$

where

$$\xi_{ni\cdots pq} = L_{qq} + \sum_{r\neq ni\cdots pq} \frac{L_{qr}L_{rq}}{z - \xi_{ni\cdots pqr}} + \sum_{\substack{r\neq ni\cdots pq \\ s\neq ni\cdots pqr}} \frac{L_{qr}L_{rs}L_{sq}}{(z - \xi_{ni\cdots pqr})(z - \xi_{ni\cdots pqrs})} + \cdots \tag{4.7}$$

and

$$L_{mn} = (A_m, LA_n). \tag{4.8}$$

For the off-diagonal type of matrix elements, I shows that

$$(A_m, (z - L)^{-1}A_n) = (z - \xi_m)^{-1}\left(L_{mn} + \sum_{r\neq m,n} \frac{L_{mr}L_{rn}}{z - \xi_{mnr}} + \sum_{\substack{r\neq m,n \\ s\neq m,n,r}} \frac{L_{mr}L_{rs}L_{sn}}{(z - \xi_{mnr})(z - \xi_{mnrs})} + \cdots\right) \times (z - \xi_{mn})^{-1}. \tag{4.9}$$

These formulas may be derived in a fashion similar to that of Feshbach[4] who obtained the Feenberg formulas directly from the eigenvalue problem using a method of successive approximations.

Another argument using a matrix analogy may be summarized as follows: If we were dealing with a finite matrix, we would say that Feenberg perturbation theory could be used to obtain an algebraic expression for the Cramer's rule formula for the elements of the inverse of a matrix. The philosophy used then would be to assume we have a finite matrix, obtain the Feenberg formula, then allow the matrix to become arbitrarily large. The formulas (4.6) and (4.9) result. Similar ideas have been used by Masson[18] while establishing the rigorous convergence properties of an approximation scheme using finite dimensional subspaces of a Hilbert space to approximate a general Hamiltonian.

To illustrate these ideas consider the electron creation (annihilation) operators $a_m^\dagger(a_m)$ for the states $m = 1, \ldots, N$. We have in mind now a problem without the additional complication of disorder. For example, an alloy in a particular configuration could be discussed in the following terms using the scalar products (3.6) and (3.10). Then for the fermion case we find

$$(a_m, a_n) = \delta_{mn}, \quad (a_m^\dagger, a_n) = 0 \tag{4.10}$$

so that these operators are automatically orthonormal and the AHGF may be applied straightforwardly to single-particle Green's functions. Suppose further that these operators diagonalize the Hamiltonian

$$H = \sum E_l a_l^\dagger a_l. \tag{4.11}$$

Then choosing our two-particle operators as $A_{lm} = a_l^\dagger a_m$ we find

$$((A_{lm}, A_{pq})) = (E_l - E_m)(\langle n_m\rangle - \langle n_l\rangle)\delta_{lp}\delta_{mq}. \tag{4.12}$$

A_{lm}'s which commute with H have zero norm and are excluded from the Hilbert space. The remaining A_{lm}'s are mutually orthogonal and can all be normalized.

In order to apply AHGF to the conductivity, we need to express the conductivity as a matrix element of a resolvent operator. This is accomplished by first noting that Eq. (2.20) may be rewritten as

$$\sigma_{\mu\nu}(\omega) = \frac{i}{\Omega}((P_\nu, (\omega + i\eta - L)^{-1}P_\mu)), \tag{4.13}$$

where we have used definition (4.2). For a simple cubic crystal we may restrict discussion to the case $\mu = \nu$. In this case (4.13) is the diagonal element of a resolvent and we may apply (4.6). Since P_μ is not a normalized vector, we must be careful to account for its normalization. Also note that Ω is the total crystal volume. We

find that

$$\sigma(\omega)=\frac{i}{\Omega}\frac{\langle\langle P_\mu, P_\mu\rangle\rangle}{\omega-\xi_P(\omega)}. \tag{4.14}$$

Equation (4.14) is an exact result independent of the precise definitions of either P or H. For example, the Hamiltonian may be generalized to include many-body effects (e.g., electron–electron interactions via a disordered Hubbard Hamiltonian[19]) without changing (4.14). In a practical calculation, the AHGF formulas become more difficult to evaluate, the more general the Hamiltonian. However, the principles remain unchanged. We consider (4.14) to be the canonical form for the conductivity. ξ_P serves the same role in the two-particle theory that the self-energy serves in the single-particle theory. Equation (4.14) is as important to the theory of the conductivity as

$$G(E)=[E-\Sigma(E)]^{-1}$$

is to the theory of the one electron Green's function. It should be the starting point of most approximation schemes.

Using the result of Appendix A to express the numerator of (4.14) in terms of H_0, we find

$$\sigma(\omega)=-\frac{ia^2e^2}{3\Omega}\frac{\langle \mathrm{Tr}\rho H_0\rangle_c}{\omega+i\eta-\xi_P(\omega)}. \tag{4.15}$$

This formula is convenient in many cases. For example, it is easy to see that the classical Drude formula is a special case of (4.15). To illustrate we assume a simple exponential decay for

$$\langle \mathrm{Tr}\rho[J_\mu(t), P_\mu(0)]\rangle_c=i\frac{a^2e^2}{3}\langle \mathrm{Tr}\rho H_0\rangle_c\exp(-2\Gamma t) \tag{4.16}$$

and obtain from (4.7)

$$\sigma(\omega)=-\frac{ia^2e^2}{3\Omega}\frac{\langle \mathrm{Tr}\rho H_0\rangle_c}{\omega+2i\Gamma} \tag{4.17}$$

which is formally equivalent to the classical Drude formula.[8] We see that (4.15) reduces to this form when the two-particle self-energy $\xi_P(\omega+i\eta)=-2i\Gamma$ is independent of ω.

Depending on the particular limit of the conductivity being studied, three different methods of using AHGF may be employed. The necessity of using different methods for different problems is imposed by requirements of orthogonality on the basis vectors in the abstract Hilbert space. Three possibilities have been encountered in practice: (1) orthogonal basis vectors may be found through Gram–Schmidt orthogonalization; (2) nonorthogonality of the basis vectors is negligible to the order of approximation being studied; (3) a set of nonorthogonal basis vectors must be employed to obtain results correct to the desired approximation. Each of these cases will be discussed further in the following paragraphs and examples will be given in Secs. V–VIII.

It is always possible to obtain an orthonormal set of vectors from an arbitrary set by using the Gram–Schmidt orthogonalization procedure. In a finite time, we can only obtain a finite number of vectors this way. In the limit of small bandwidth, this restriction is acceptable because only three basis vectors are required to span the Hilbert space for the conductivity calculation. In more general problems, a technical difficulty arises because expanding the Feenberg formulas in terms of the Gram–Schmidt basis results in a continued fraction. In order to obtain an imaginary part for ξ_P, we must be able to sum this continued fraction or obtain an appropriate terminating function for it. Summing the continued fraction requires knowledge of all the basis vectors of the Hilbert space and this knowledge requires considerable effort to obtain in most cases (if it is possible at all). A physically motivated terminating function may possibly be found but this is also a difficult problem to solve in many cases.

Under certain special conditions, nonorthogonality of the basis vectors may not be important so an alternative method may be used. Suppose we wish to calculate $\xi_{A_n}(\omega)$, where $A_n=A_n^\dagger$ is an Hermitian operator. Then we notice that if we were to orthogonalize A_m with respect to A_n according to

$$A_m'=A_m-S_{mn}A_n, \tag{4.18}$$

where S_{mn} is the overlap matrix element, then we would find

$$\langle\langle A_n, LA_m'\rangle\rangle=\langle\langle A_n, LA_m\rangle\rangle \tag{4.19}$$

because $\langle\langle A_n, LA_n\rangle\rangle=0$ when A_n is Hermitian. Thus, if we are doing an approximate calculation and only retain the first nonzero term of Eq. (4.7) for $\xi_{A_n}(\omega)$, we may safely ignore the nonorthogonality. This fact saves us much work in Sec. VI where $A_n=P_\mu$.

If our problem is such that the Gram–Schmidt procedure is not satisfactory and the nonorthogonality is not negligible, then we must either orthogonalize the vectors by some clever trick or use the nonorthogonal vectors themselves in our calculation. In general, the rigorous application of Eq. (4.7) to two-particle Green's functions requires exactly the same amount of work in choosing an orthonormal set of operators as is required to diagonalize the Hamiltonian. This is obviously out of the question. It would be most convenient if we could apply AHGF to a set of nonorthogonal vectors. A procedure for doing this is discussed in Sec. VII. Having established a procedure for using the nonorthogonal basis vectors, a number of alternatives will still remain. We can expand P_μ in terms of Bloch functions or Wannier functions. The Bloch representation is most convenient in the weak scattering strength limit and the Wannier representation is convenient for the small bandwidth limit. We will derive the general formulas for the Bloch representation in Sec. VIII.

V. SMALL BANDWIDTH LIMIT

The Feenberg formulas can be explicitly evaluated in the limiting cases of small bandwidth or small scattering strength. This evaluation is convenient and useful for two reasons: (1) The operator expectation values can be evaluated approximately in these limits. This is done in Appendix B. (2) Exact results are known for both of these limiting cases (see Ref. 5).

Since the two-particle self-energy in the small bandwidth limit has only a trivial imaginary part (proportional to $\mathrm{Im}\omega$), it is satisfactory to use the Gram–

Schmidt orthogonalization procedure on the vectors of our abstract Hilbert space.

Recalling expression (4.13) for the conductivity it is clear that the first vector of the sequence must be P. Using the full Hamiltonian, we can generate a second vector according to

$$LP_\mu = iJ_\mu. \tag{5.1}$$

It is a trivial consequence of the scalar product (4.2) that P and J are orthogonal. Normalizing these vectors according to

$$p_\mu \equiv \frac{P_\mu}{\|P_\mu\|} \quad \text{and} \quad j_\mu \equiv \frac{J_\mu}{\|J_\mu\|} \tag{5.2}$$

we can rewrite (5.1) as

$$Lp_\mu = i\frac{\|J_\mu\|}{\|P_\mu\|} j_\mu. \tag{5.3}$$

Then by using the Hermitian character of L, we can write

$$LJ_\mu = -i\frac{\|J_\mu\|^2}{\|P_\mu\|^2} P_\mu + Q_\mu \tag{5.4}$$

which defines Q_μ. (5.4) guarantees that Q is orthogonal to P by Gram–Schmidt construction and to J because

$$((J_\mu, LJ_\mu)) = ((J_\mu, Q_\mu)) = 0. \tag{5.5}$$

Continuing the sequence one more step we have

$$LQ_\mu = (L^2 - \|J_\mu\|^2/\|P_\mu\|^2)J. \tag{5.6}$$

In the small bandwidth limit it is essential to use the full Hamiltonian (2.1) in evaluating (5.1). However, we can use the truncated Hamiltonian

$$H = \sum_i \epsilon_i n_i \tag{5.7}$$

for calculating (5.4) and (5.6). This simplifies the analysis greatly. Using the notation from Sec. II and defining $\delta = \epsilon_B - \epsilon_A$, we find

$$((P_\mu, P_\mu)) = -\frac{2e^2t^2a^2ZN}{3}\left(x^2f'(\epsilon_A) + y^2f'(\epsilon_B) + 2xy\frac{f(\epsilon_A) - f(\epsilon_B)}{\delta}\right), \tag{5.8}$$

$$((J_\mu, J_\mu)) = -\frac{4e^2t^2a^2ZN}{3}xy\delta[f(\epsilon_A) - f(\epsilon_B)], \tag{5.9}$$

$$((J_\mu, L^{2n}J_\mu)) = \delta^{2n}((J_\mu, J_\mu)), \tag{5.10}$$

$$((J_\mu, LQ_\mu)) = \alpha\delta^2((J_\mu, J_\mu)), \tag{5.11}$$

and

$$((LQ_\mu, LQ_\mu)) = \alpha^2\delta^4((J_\mu, J_\mu)). \tag{5.12}$$

γ and α are defined by

$$2\gamma\delta^2 = \|J_\mu\|^2/\|P_\mu\|^2 \tag{5.13}$$

and $\alpha = 1 - 2\gamma$. Combining (5.11) and (5.12) clearly implies

$$LQ_\mu/\|LQ_\mu\| = j_\mu \tag{5.14}$$

so the series truncates at this order.

Given these results we can calculate the Feenberg formulas straightforwardly. They give

$$\xi_P(\omega) = 2\gamma\delta^2/[\omega - \xi_{P,J}(\omega)] \tag{5.15}$$

and

$$\xi_{P,J}(\omega) = \alpha\delta^2/\omega. \tag{5.16}$$

Finally, we find that

$$\xi_P(\omega) = 2\gamma\delta^2/(\omega^2 - \alpha\delta^2) \tag{5.17}$$

which is identical with the exact result obtained from a straightforward perturbation calculation correct to $O(t^2)$.

Notice that $\xi_P(0) = 0$, $\xi_P(\delta) = \delta$, and $\xi_P(-\delta) = -\delta$, so (5.17) correctly determines the poles of the conductivity. These poles correspond to transitions from A to A or B to B-type neighbors ($\omega = 0$), from A to B-type neighbors ($\omega = \delta$), and from B to A-type neighbors ($\omega = -\delta$). The residues of the poles correctly account for the fact that in order to have a transition from site l to site m, site l must be occupied with probability $f(\epsilon_l)$ and site m unoccupied with probability $[1 - f(\epsilon_m)]$. The probability that both sites are either A type or B type is x^2 or y^2, respectively and the probability that one is A type with the other B type is xy. These configuration effects are properly accounted for by the residues of the poles. See Ref. 5 for more details.

Thus the AHGF reproduces the exact result for the small bandwidth limit. We could improve this result by retaining the next higher order contributions in W to ξ_P. This will not be done here.

VI. SMALL SCATTERING STRENGTH LIMIT

We develop the AHGF formulas for $\sigma(\omega)$ in the small scattering strength limit in this section. We find an expression for $\sigma(\omega)$ correct to $O(\delta^2)$.

The conductivity is given by (4.14), where

$$\xi_{P_\mu}(\omega) = \sum_{k,q} \frac{|L_{P_\mu, B_{k,k+q;q}}|^2}{\omega - \xi_{P_\mu, B_{k,k+q;q}}(\omega)} + \cdots, \tag{6.1}$$

$$B'_{k,k+q;q} = U(q)c_k^\dagger c_{k+q}, \tag{6.2}$$

and B is B' divided by the norm of B'.

Since P_μ is an Hermitian operator, we may ignore the nonorthogonality of P_μ and the B's according to the argument giving Eq. (4.19). This procedure is valid in the small scattering strength limit because we need only keep the term in (6.1) which is explicitly shown in this limit.

The matrix elements are given by

$$L_{P_\mu, B_{k,k+q;q}} = \frac{((B'_{k,k+q;q}, LP_\mu))}{\|P_\mu\|\,\|B'_{k,k+q;q}\|}. \tag{6.3}$$

We can evaluate these vector norms and scalar products using the results of Appendix B. We find

$$\|B'_{k,k+q;q}\|^2 = xy\delta^2N^{-1}(\epsilon_k - \epsilon_{k+q})[f(\epsilon_{k+q}) - f(\epsilon_k)] + O(\delta^4) \tag{6.4}$$

and

$$((B'_{k,k+q;q}, LP_\mu)) = iexy\delta^2N^{-1}(v_k^u - v_{k+q}^u)[f(\epsilon_{k+q}) - f(\epsilon_k)] + O(\delta^3). \tag{6.5}$$

Furthermore, we can write

$$\xi_{P_\mu, B_{k,k+q;q}}(\omega) = \epsilon_{k+q} - \epsilon_k + O(\delta^2). \tag{6.6}$$

Substituting all these expressions into (6.1), our final expression for the self-energy is

$$\xi_P(\omega) = -2xy\delta^2 m^* \Delta(\omega), \tag{6.7}$$

where

$$(m^*)^{-1} = \sum_k (v_k^u)^2[-f'(\epsilon_k)] \tag{6.8}$$

and

$$\Delta(\omega) = \omega N^{-1} \sum_{k,q} \frac{(v_k^u)^2[f(\epsilon_{k+q}) - f(\epsilon_k)]/(\epsilon_{k+q} - \epsilon_k)}{\omega^2 - (\epsilon_{k+q} - \epsilon_k)^2}. \tag{6.9}$$

We have used the time reversal symmetry

$$\epsilon_k = \epsilon_{-k} \quad \text{and} \quad v_k^u = -v_{-k}^u \tag{6.10}$$

in simplifying (6.7). The conductivity may be obtained by substituting (6.7) into (4.14).

Checking (6.7) in the $\omega \to 0$ limit we find

$$\mathrm{Im}\,\xi_{P_\mu}(0) = -2\pi xy\delta^2 \rho(\epsilon_F) \tag{6.11}$$

when $T = 0$. Thus the standard Boltzmann result is recovered. (Here F stands for Fermi surface.)

Equation (6.7) should also be valid for high frequencies. Previous work on the high frequency conductivity has been discussed in the Introduction. Yamada's work[6] corresponds closely with our own. It is not difficult to show that when Yamada's Eq. (30) is specialized to the binary alloy and then configuration averaged, the result is in complete agreement with (6.7) at high frequencies. See Ref. 5 for more details.

VII. NONORTHOGONAL BASIS VECTORS

Similar problems of expressing formulas in terms of nonorthogonal basis vectors have been studied in other contexts by Roth[20] and by Halpern and Bergmann.[21,22]

Let $\{A_i\}$ be any complete normalized but not necessarily orthogonal set of vectors in the Hilbert space defined by either (4.1) or (4.2). If A is an eigenfunction of L with eigenvalue z, then

$$A = \sum \alpha_i A_i \tag{7.1}$$

is a representation (not necessarily unique) of A in terms of $\{A_i\}$. The coefficients α_i satisfy the system of linear homogeneous equations

$$(z - R_{ij})\alpha_i = \sum_{j \neq i} R_{ij}\alpha_j, \tag{7.2}$$

where

$$R_{ij}(z) = L_{ij} - zS_{ij}(1 - \delta_{ij}) \tag{7.3}$$

and the overlap integral is defined by

$$S_{ij} = \begin{cases} (A_i, A_j) \\ ((A_i, A_j)) \end{cases} \tag{7.4}$$

for scalar products (4.1) and (4.2), respectively. When the A_i's are orthogonal, (7.2) reduces to

$$(z - L_{ii})\alpha_i = \sum_{j \neq i} L_{ij}\alpha_j. \tag{7.5}$$

If we specify an arbitrary coefficient n as $\alpha_n = 1$ (thus normalizing all other α_i's with respect to α_n), then the system (7.5) is equivalent to the one solved in general by Feenberg. The only difference between (7.5) and (7.2) is that the off-diagonal elements with $S_{ij} \neq 0$ depend linearly on z. Feshbach's derivation of the Feenberg formulas is still valid for (7.2). Hence, the only alteration required in the formulas (4.7) and (4.9) is to replace L_{ij} everywhere by R_{ij}. This result follows immediately by applying the matrix analogy argument discussed following (4.9).

We conclude that any Green's function can be expressed as a Feenberg series by expanding the Feenberg formulas in terms of a complete set of nonorthogonal vectors. This fact is important for practical applications of the method.

VIII. THE BLOCH REPRESENTATION

If we wish to obtain a systematic approximation for the conductivity correct to higher than second order in the scattering strength, we must apply the full AHGF method using nonorthogonal vectors. Such a calculation is necessarily tedious but can be carried through in a reasonable amount of time. For reasons of space, we will not work such an example here. Instead we will exhibit the general AHGF formulas in the Bloch representation. These formulas will then be the starting point for a better calculation of the conductivity in the future.

We define

$$A'_{kk'} = c_k^\dagger c_{k'}. \tag{8.1}$$

The prime on A indicates the vector is unnormalized. Then using the Hamiltonian in the form (2.6) we find

$$LA'_{k,k'} = (\epsilon_{k'} - \epsilon_k)A'_{k,k'} + \sum_{q'} (B'_{k,k'+q';q'} - B'_{k-q',k';q'}), \tag{8.2}$$

where we define

$$B'_{k_1,k_2;q} = U(q)c_{k_1}^\dagger c_{k_2}. \tag{8.3}$$

Similarly,

$$LB'_{k,k';q'} = (\epsilon_{k'} - \epsilon_k)B'_{k,k';q'} + \sum_{q''} (D'_{k,k'+q'';q',q''} - D'_{k-q'',k';q',q''}), \tag{8.4}$$

where

$$D'_{k_1,k_2;q_1,q_2} = U(q_1)U(q_2)c_{k_1}^\dagger c_{k_2}. \tag{8.5}$$

The pattern should be apparent.

The next step in the Feenberg procedure is to normalize the vectors and determine their orthogonality properties. To save writing throughout the remainder of this section we understand $\langle 0 \rangle = \langle \mathrm{Tr}\rho 0 \rangle_c$. Then

$$\begin{aligned}((A'_{k_1,k_2}, A'_{k_3,k_4})) = {} & (\epsilon_{k_3} - \epsilon_{k_4})(\langle n_{k_4} \rangle - \langle n_{k_3} \rangle)\delta_{k_2,k_4}\delta_{k_1,k_3} \\ & + \delta_{k_2-k_4,k_1-k_3}[\langle U(k_1 - k_3)c_{k_3}^\dagger c_{k_1} \rangle \\ & + \langle U(k_4 - k_2)c_{k_2}^\dagger c_{k_4} \rangle] - \delta_{k_3,k_1}\delta_{k_4,k_2} \\ & \times \sum_q [\langle U(q)c_{k_4}^\dagger c_{k_4+q} \rangle + \langle U(q)c_{k_3-q}^\dagger c_{k_3} \rangle],\end{aligned} \tag{8.6}$$

where (8.6) has been simplified using the expectation values (B7)–(B9). Similarly,

$$\begin{aligned}&((B'_{k_1,k_2;q_1}, B'_{k_3,k_4;q_2})) \\ &\quad = (\epsilon_{k_3} - \epsilon_{k_4})[\langle U(-q_1)U(q_2)c_{k_2}^\dagger c_{k_4} \rangle \delta_{k_1,k_3}\end{aligned}$$

$$
\begin{aligned}
&- \langle U(-q_1)U(q_2)c^\dagger_{k_3}c_{k_1}\rangle \delta_{k_2,k_4}] \\
&+ \sum_q [\langle U(-q_1)U(q_2)U(q)\, c^\dagger_{k_3}c_{k_1}\rangle \delta_{k_2,k_4+q} \\
&+ \langle U(-q_1)U(q_2)U(q)c^\dagger_{k_2}c_{k_4}\rangle \delta_{k_1,k_3-q} \\
&- \langle U(-q_1)U(q_2)U(q)c^\dagger_{k_2}c_{k_4+q}\rangle \delta_{k_1,k_3} \\
&- \langle U(-q_1)U(q_2)U(q)c^\dagger_{k_3-q}c_{k_1}\rangle \delta_{k_2,k_4}] \qquad (8.7)
\end{aligned}
$$

and

$$
\begin{aligned}
&((B'_{k_1,k_2;q_1}, A'_{k_3,k_4})) \\
&= (\epsilon_{k_3} - \epsilon_{k_4})[\langle U(-q_1)c^\dagger_{k_2}c_{k_4}\rangle \delta_{k_1,k_3} \\
&- \langle U(-q_1)c^\dagger_{k_3}c_{k_1}\rangle \delta_{k_2,k_4}] + \sum_q [\langle U(-q_1)U(q)c^\dagger_{k_3}c_{k_1}\rangle \\
&\times \delta_{k_2,k_4+q} + \langle U(-q_1)U(q)c^\dagger_{k_2}c_{k_4}\rangle \delta_{k_1,k_3-q} \\
&- \langle U(-q_1)U(q)c^\dagger_{k_2}c_{k_4+q}\rangle \delta_{k_1,k_3} \\
&- \langle U(-q_1)U(q)c^\dagger_{k_3-q}c_{k_1}\rangle \delta_{k_2,k_4}]. \qquad (8.8)
\end{aligned}
$$

Although these general expressions are quite complicated, the algorithm for calculating them is very simple. For particular values of the subscripts, the expressions may be greatly simplified using (B7)–(B9).

If we define the norm of a vector A as

$$\|A\|^2 = ((A, A)), \qquad (8.9)$$

we can define the normalized vectors

$$A_{k,k'} = A'_{k,k'}/\|A'_{k,k'}\| \qquad (8.10)$$

and similarly for the B's, D's, etc. Then we can rewrite (8.2) as

$$
\begin{aligned}
LA_{k,k'} = (\epsilon_{k'} - \epsilon_k)A_{k,k'} + \sum_q [&L_{A_{k,k'},B_{k,k'+q;q}} B_{k,k'+q;q} \\
&+ L_{A_{k,k'},B_{k-q,k';q}} B_{k-q,k';q}], \qquad (8.11)
\end{aligned}
$$

where

$$L_{A_{k,k'},B_{k,k'+q;q}} = \|B'_{k,k'+q;q}\|/\|A'_{k,k'}\|, \qquad (8.12)$$

$$L_{A_{k,k'},B_{k-q,k';q}} = \|B'_{k-q,k';q}\|/\|A'_{k,k'}\|. \qquad (8.13)$$

The corresponding expression for $LB_{k,k';q}$ may be obtained by analogy.

IX. SUMMARY AND CONCLUSIONS

We introduce scalar products with associated Hilbert spaces for single- and two-particle operators and show how these may be used to generalize Feenberg perturbation theory in thermal Green's function applications. Three methods of using the AHGF are demonstrated in calculations of the disordered alloy conductivity.

First, the Gram–Schmidt orthogonalization procedure is used in the small bandwidth (strong scattering) limit. The Gram–Schmidt procedure results in a continued fraction which truncates in this example but may not truncate in general. The real part of the conductivity is a finite sum of delta functions when the continued fraction does truncate. In a typical physical application, one is interested in finding a conductivity whose real part is a continuous function of the frequency (not a sum of delta functions). Therefore, general application of the Gram–Schmidt procedure requires some physically motivated truncation of the continued fraction. In some cases, clever truncations (such as that of Onodera and Toyozawa[23]) no doubt exist.

Second, under special circumstances one may obtain a valid approximation for the conductivity without treating the nonorthogonality of the vectors in the Feenberg series. The requirements are that (1) the starting vector (e.g., P_μ for the conductivity) must be Hermitian and (2) only the first nontrivial term of the Feenberg series for the self-energy is needed. This technique is exploited to calculate a closed expression (6.7) for $\sigma(\omega)$ correct to $O(\delta^2)$ for all values of ω. The authors believe this to be a new result.

Third, under more general conditions it is shown that nonorthogonal basis vectors may be used to expand the Feenberg formulas. To demonstrate the type of calculations required, the first few AHGF formulas are written down in the Bloch representation. This technique can be used systematically to obtain expressions for the conductivity valid to any order desired although considerable effort may be required to do so.

In conclusion, we wish to point out that the AHGF method is *not* the same as the usual Feenberg method in the case of the one electron Green's function with disorder.[24,25] The usual Feenberg–Watson approach[26] is used strictly in the site representation and the arguments suggesting the behavior of the most probable Green's function form the final step of the calculation. The AHGF method treats the configuration averaged Green's function rigorously from the start and is not restricted to any particular representation. In fact, one important reason for studying the AHGF approach is that it permits the configuration average and the thermal average to be treated equally.

It has been shown elsewhere[5] that the AHGF method can be used to derive the CPA for single-particle Green's functions. An outstanding problem at present is to show that the method can also be used to derive the CPA for the transport coefficients.[12] This demonstration must await further developments in the theory.

APPENDIX A

A quantity proportional to the inverse of the optical effective mass of the conducting electrons is

$$
\begin{aligned}
\langle [[P_\mu, H], P_\mu]\rangle &= -e^2 \sum_{ij} t_{ij}(R_i - R_j)^2_\mu \langle c^\dagger_i c_j\rangle \\
&= -\frac{a^2e^2}{3} \sum_{ij} t_{ij}\langle c^\dagger_i c_j\rangle \\
&= -\frac{a^2e^2}{3}\langle H_0\rangle. \qquad (A1)
\end{aligned}
$$

The last two equalities come from averaging over the three equivalent directions in the s.c. crystal. H_0 is defined in (2.2). The formula so far assumes nothing concerning the disorder. If we assume only diagonal disorder, we find

$$
\begin{aligned}
\langle [[P_\mu, H], P_\mu]\rangle &= e^2 \sum_k \frac{\partial^2 \epsilon_k}{\partial k_\mu^2}\langle n_k\rangle \\
&= e^2 \sum_k (v^\mu_k)^2 \left(-\frac{\partial}{\partial \epsilon_k}\langle n_k\rangle\right)
\end{aligned}
$$

$$= -\frac{a^2e^2}{3}\sum_k \epsilon_k\langle n_k\rangle$$

$$= \frac{e^2}{3}\sum_k v_k^2\left(-\frac{\partial}{\partial\epsilon_k}\langle n_k\rangle\right). \tag{A2}$$

APPENDIX B

In this appendix we collect various results of expectation value calculations. To save space we make the convention throughout this appendix that $\langle 0\rangle = \langle \mathrm{Tr}\rho 0\rangle_c$.

1. Expectation values of U(q)

We assume the origin of the energy is chosen so that the average alloy potential energy vanishes. In other words, for

$$U(q) = N^{-1}\sum_i \epsilon_i \exp(iq\cdot R_i), \tag{B1}$$

we have

$$U(0) = \frac{1}{N}\sum_i \epsilon_i = \bar{\epsilon} = 0. \tag{B2}$$

Hence, we find immediately that

$$\langle U(q)\rangle = \bar{\epsilon}\delta_{q,0} = 0. \tag{B3}$$

Thus if any of the q arguments vanish in the expressions (B4)–(B6) which follow, then the expression vanishes identically. For example, $\langle U(0)U(0)\rangle = 0$. In all other cases we find

$$\langle U(q)U(q')\rangle = \left(\frac{xy\delta^2}{N}\right)\delta_{q+q',0}, \tag{B4}$$

$$\langle U(q)U(q')U(q'')\rangle = \left(\frac{xy(x-y)\delta^3}{N^2}\right)\delta_{q+q'+q'',0}, \tag{B5}$$

and

$$\langle U(q)U(q')U(q'')U(q''')\rangle$$

$$= \frac{xy(1-6xy)\delta^4}{N^3}\,\delta_{q+q'+q''+q''',0} + \frac{x^2y^2\delta^4}{N^2}$$

$$\times(\delta_{q+q',0}\delta_{q''+q''',0} + \delta_{q+q'',0}\delta_{q'+q''',0} + \delta_{q+q''',0}\delta_{q'+q'',0}), \tag{B6}$$

where $q^{(i)} \neq 0$.

2. Expectation values in Bloch representation

Because of the translational invariance of the configuration average crystal, we can also simplify certain expectation values involving electron annihilation and creation operators:

$$\langle c_k^\dagger c_{k'}\rangle = \langle n_k\rangle\,\delta_{k,k'}, \tag{B7}$$

$$\langle U(q)c_k^\dagger c_{k'}\rangle = \langle U(q)c_k^\dagger c_{k+q}\rangle\delta_{k',k+q} \quad \text{for } q\neq 0, \tag{B8}$$

$$\langle U(q_1)\cdots U(q_n)c_k^\dagger c_{k'}\rangle = \langle U(q_1)\cdots U(q_n)c_k^\dagger c_{k+q_1+\cdots+q_n}\rangle \times\delta_{k',k+q_1+\cdots+q_n}. \tag{B9}$$

These results are proven very simply by expanding in a Wannier basis and noting that the resulting expectation values can only depend on the difference between the site indices. This fact always causes one index to be free to sum over. This gives rise to the Kronecker deltas in (B7)–(B9).

3. Expectation values in Wannier representation

In the Wannier basis we find

$$\langle \epsilon_n c_l^\dagger c_m\rangle = \langle \epsilon_n n_n\rangle\,\delta_{l,m}\delta_{l,n} + \langle \epsilon_l c_l^\dagger c_m\rangle\delta_{n,l}(1-\delta_{l,m}) + \langle \epsilon_m c_l^\dagger c_m\rangle\delta_{n,m}(1-\delta_{l,m}) \tag{B10}$$

and

$$\langle \epsilon_n\epsilon_p c_l^\dagger c_m\rangle = \langle \epsilon_n^2 n_n\rangle\,\delta_{n,p}\delta_{n,l}\delta_{n,m} + \langle \epsilon_l\epsilon_m c_l^\dagger c_m\rangle(\delta_{p,l}\delta_{n,m} + \delta_{n,l}\delta_{p,m})(1-\delta_{n,p}) + \langle \epsilon_n^2\rangle\langle c_l^\dagger c_m\rangle\delta_{n,p}(1-\delta_{l,n})(1-\delta_{m,n}). \tag{B11}$$

4. Expectation values for small bandwidth

When the transfer matrix element t_{ij} is very small compared to the scattering strength in Hamiltonian (2.3), we can evaluate various operator expectation values approximately to lowest order in t_{ij}. If the expectation value has a zero order contribution independent of t, we will generally neglect $O(t)$ in the formulas. Note that $f(E) = [1+\exp\beta(E-\mu)]^{-1}$ is the Fermi distribution function and $f' = df/dE$,

$$\langle c_l^\dagger c_m\rangle \approx [xf(\epsilon_A) + yf(\epsilon_B)]\delta_{lm} + t_{lm}\left[x^2f'(\epsilon_A) + y^2f'(\epsilon_B) + 2xy\frac{f(\epsilon_B)-f(\epsilon_A)}{\epsilon_B-\epsilon_A}\right], \tag{B12}$$

$$\langle \epsilon_n c_l^\dagger c_m\rangle \approx [x\epsilon_A f(\epsilon_A) + y\epsilon_B f(\epsilon_B)]\delta_{lm}\delta_{ln} + t_{lm}\left(x^2\epsilon_A f'(\epsilon_A) + y^2\epsilon_B f'(\epsilon_B) + xy(\epsilon_A+\epsilon_B) \times\frac{f(\epsilon_B)-f(\epsilon_A)}{\epsilon_B-\epsilon_A}\right)(\delta_{nm}+\delta_{nl}), \tag{B13}$$

$$\langle \epsilon_n\epsilon_p c_l^\dagger c_m\rangle \approx [x\epsilon_A^2 f(\epsilon_A) + y\epsilon_B^2 f(\epsilon_B)]\delta_{np}\delta_{nl}\delta_{nm} + \cdots, \tag{B14}$$

and

$$\langle(\epsilon_i-\epsilon_j)^m(n_j-n_i)\rangle \approx [1-(-1)^m]xy\delta^m[f(\epsilon_A)-f(\epsilon_B)]. \tag{B15}$$

All other expectation values required in our work can be obtained straightforwardly from those given in (B12)–(B15).

5. Expectation values for small scattering strength

When the scattering strength δ is much smaller than the bandwidth, we can obtain all the operator expectation values we need from the formula

$$\langle U(q_1)U(q_2)\cdots U(q_n)c_{k_1}^\dagger c_{k_2}\rangle = \langle U(q_1)U(q_2)\cdots U(q_n)\rangle f(\epsilon_{k_1})\delta_{k_1,k_2} + \langle U(q_1)U(q_2)\cdots U(q_n)U(k_1-k_2)\rangle \times\frac{f(\epsilon_{k_1})-f(\epsilon_{k_2})}{\epsilon_{k_1}-\epsilon_{k_2}}(1-\delta_{k_1,k_2}). \tag{B16}$$

We must use the results of (B3)–(B6) in evaluating the expectation values of the scattering factors. Note that when $\epsilon_{k_1} = \epsilon_{k_2}$ for $k_1 \neq k_2$ we use the convention that

$$\frac{f(\epsilon_{k_1})-f(\epsilon_{k_2})}{\epsilon_{k_1}-\epsilon_{k_2}} = f'(\epsilon_{k_1}).$$

*Supported in part by the Wisconsin Alumni Research Foundation.

†Present address: Continental Oil Company, Ponca City, Oklahoma 74601.

†Present address: Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061.

[1]S.P. Bowen, J. Math. Phys. **16**, 620 (1975).
[2]E. Feenberg, Phys. Rev. **74**, 206 (1948).
[3]P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. II., Chap. 9.
[4]H. Feshbach, Phys. Rev. **74**, 1548 (1948).
[5]J.G. Berryman, Ph.D. thesis (University of Wisconsin, 1975) (unpublished).
[6]K. Yamada, Progr. Theor. Phys. **28**, 299 (1962).
[7]A. Lonke and A. Ron, in *Optical Properties and Electronic Structure of Metals and Alloys*, edited by F. Abeles (North-Holland, Amsterdam, 1966), p. 613.
[8]B. Velicky and K. Levin, Phys. Rev. B **2**, 938 (1970).
[9]P.N. Sen, Phys. Rev. B **8**, 5613 (1973).
[10]P. Soven, Phys. Rev. **156**, 809 (1967).
[11]B. Velicky, S. Kirkpatrick, and H. Ehrenreich, Phys. Rev. **175**, 747 (1968).
[12]B. Velicky, Phys. Rev. **184**, 614 (1969).
[13]R.J. Elliott, J.A. Krumhansl, and P.L. Leath, Rev. Mod. Phys. **46**, 465 (1974).
[14]D.N. Zubarev, Usp. Fiz. Nauk, **71**, 71 (1960) [Sov. Phys. Usp. **3**, 320 (1960)].
[15]A. Lonke, J. Math. Phys. **12**, 2422 (1971).
[16]H. Mori, Progr. Theor. Phys. **33**, 423 (1965).
[17]R. Zwanzig, in *Lectures in Theoretical Physics*, edited by W.E. Brittin, B.E. Brittin, B.W. Downs, and J. Downs (Interscience, New York, 1961), Vol. III.
[18]P. Masson, in *The Padé Approximant in Theoretical Physics*, edited by G.A. Baker, Jr. and J.L. Gammel (Academic, New York, 1970), p. 197.
[19]H. Fukuyama and H. Ehrenreich, Phys. Rev. B **7**, 3266 (1973).
[20]L.M. Roth, Phys. Rev. Lett. **28**, 1570 (1972).
[21]V. Halpern and A. Bergmann, J. Phys. C **5**, 1953 (1972).
[22]A. Bergmann and V. Halpern, J. Phys. C **7**, 289 (1974).
[23]Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. **24**, 341 (1968).
[24]P.W. Anderson, Phys. Rev. **109**, 1492 (1958).
[25]R. Abou-Chacra, P.W. Anderson, and D.J. Thouless, J. Phys. C **6**, 1734 (1973).
[26]K.M. Watson, Phys. Rev. **105**, 1388 (1957).

On a solution of the Einstein–Maxwell equations admitting a nonsingular electromagnetic field*

Raymond G. McLenaghan and Nessim Tariq

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
(Received 20 April 1976)

Solutions of the Einstein–Maxwell equations are investigated for which the electromagnetic field is nonsingular and weakly parallelly propagated along its principal null congruences. It is shown that a subclass of these solutions admits an invertible two-dimensional Abelian group of motions. A weaker characterization of a recently found solution is thereby obtained.

1. INTRODUCTION

Recently investigations[1–4] of a special class of solutions of the Einstein—Maxwell field equations admitting a nonsingular electromagnetic field have been made. In the present paper we consider space—times[5] V_4 satisfying the following conditions which are weaker than those considered in the paper of McLenaghan and Tariq[6] (this reference will be denoted by MT):

I. The Einstein—Maxwell equations are satisfied; that is,

$$R_{ab} = F_{ac}F_{b}{}^{c} - \tfrac{1}{4}g_{ab}F_{cd}F^{cd}, \tag{1.1}$$

$$F_{ab}{}^{;b} = 0, \quad F_{[ab;c]} = 0. \tag{1.2}$$

II. The electromagnetic field tensor is nonsingular which is equivalent to the existence of a null tetrad $(l_a, m_a, \bar{m}_a, n_a)$ such that

$$\overset{+}{F}_{ab} = \phi(l_{[a}n_{b]} - m_{[a}\bar{m}_{b]}). \tag{1.3}$$

III. The field tensor $\overset{+}{F}_{ab}$ is "weakly" parallelly propagated along its principal null congruences. By this we mean that

$$\overset{+}{F}_{ab;c}l^{c} = f\overset{+}{F}_{ab}, \tag{1.4}$$

$$\overset{+}{F}_{ab;c}n^{c} = g\overset{+}{F}_{ab}. \tag{1.5}$$

IV. The Weyl and Ricci tensors satisfy

$$2C_{abcd}l^{a}n^{b}(l^{c}n^{d} - m^{c}\bar{m}^{d}) + R_{ab}(l^{a}n^{b} + m^{a}\bar{m}^{b}) = 0. \tag{1.6}$$

The following remarks may be made concerning the above conditions:

(a) The null tetrad is defined by F_{ab} up to the transformation

$$l'_a = e^{a}l_a, \quad m'_a = e^{ib}m_a, \quad n'_a = e^{-a}n_a, \tag{1.7}$$

where a and b are arbitrary scalar functions of the coordinates. The conditions I—IV are invariant under these transformations.

(b) It is shown in MT that conditions III and IV are equivalent to the existence of a null tetrad, in the family of tetrads defined by F_{ab}, which is parallelly propagated along the congruences defined by l^a and n^b. In fact, condition IV is the integrability condition for the existence of such a tetrad.

(c) The condition III is satisfied, for example, by the Coulomb field in Minkowski space.

(d) Condition III but not condition IV is satisfied by the Reisner—Nordström solution.

In MT it is shown that the conditions I—IV together with the condition

$$l^{a}{}_{;a} = 0 \tag{1.8}$$

imply the existence of a global system of coordinates (t, x, y, v) for which the metric and electromagnetic field have the following form:

$$ds^2 = dt^2 - dv^2 - 2kx\,dt\,dy + (k^2x^2 - e^{kv})\,dy^2 - e^{-kv}\,dx^2, \tag{1.9}$$

$$\overset{+}{F} = (1/\sqrt{2})k\exp(ikv)(dv \wedge dt + kx\,dy \wedge dv + i\,dx \wedge dy), \tag{1.10}$$

where k is any positive real number.

Some of the properties of the above solution are the following:

(a) The space—time is locally homogeneous possessing a simply transitive group of motions G_4. The Killing vectors are

$$X_1 = \frac{\partial}{\partial t}, \quad X_2 = \frac{\partial}{\partial y}, \quad X_3 = ky\frac{\partial}{\partial t} + \frac{\partial}{\partial x}, \quad X_4 = \frac{\partial}{\partial v} + \tfrac{1}{2}kx\frac{\partial}{\partial x} - \tfrac{1}{2}ky\frac{\partial}{\partial y}. \tag{1.11}$$

We note that X_1 is timelike and not hypersurface orthogonal which implies that the space—time is stationary and not static. The Killing vectors X_1 and X_2 define an Abelian subgroup of G_4 and furthermore the transformation $t' = -t$ and $y' = -y$ is an isometry. Thus the solution admits an invertible two-parameter Abelian group of motions and is an example of a space—time of Petrov type I possessing a Riemannian—Maxwellian invertible structure.[7]

(b) The electric and magnetic fields for an observer with 4-velocity u^a are defined respectively by

$$E_a = F_{ab}u^{b}, \quad H_a = -\,{}^{*}F_{ab}u^{b}.$$

For the solution under consideration these fields are collinear for the stationary observer since

$$E_a = \sqrt{2}\,k\cos kv\,\delta^{3}_{a}, \quad H_a = \sqrt{2}\,k\sin kv\,\delta^{3}_{a}. \tag{1.12}$$

In this paper we study the conditions I—IV without the condition (1.8). The main results are given in the fol-

lowing theorems:

Theorem 1: The conditions I—IV imply that there exists in the family of null tetrads defined by F_{ab} a null tetrad $(l_a, m_a, \bar{m}_a, n_a)$ such that optical scalars of the principal null congruences defined by the vectors l^a and n^a satisfy either

$$\rho = \mu, \quad \sigma = \lambda \tag{1.13}$$

or

$$\rho = -\mu, \quad \sigma = -\lambda. \tag{1.14}$$

Theorem 2: In the case $\rho = \mu$, $\sigma = \lambda$ of Theorem 1 there exists a system of coordinates (t, x, y, v) for which the metric of space—time and the electromagnetic field are given by the expressions (1.9) and (1.10) respectively.

These results are proved using the Newman—Penrose[8] (NP) spin coefficient formalism together with the equivalent vectorial formalism of Debever[9,10] described in MT.

2. PROOF OF THEOREM 1

If we insert the canonical form for $\overset{+}{F}_{ab}$ into (1.4) and contract successively with $l^a m^b$ and $\bar{m}^a n^b$, we obtain the conditions

$$\kappa = \pi = 0. \tag{2.1}$$

By a similar procedure we obtain from (1.5) the conditions

$$\nu = \tau = 0. \tag{2.2}$$

The only further consequences of contractions on (1.4) and (1.5) are

$$f = D\ln\phi, \tag{2.3}$$

$$g = \Delta\ln\phi. \tag{2.4}$$

It is easy to verify that the conditions (2.1) and (2.2) are invariant under the tetrad transformations (1.7). These transformations induce the following transformations on the nonzero tetrad components of the Weyl tensor:

$$\Psi_0' = e^{2w}\Psi_0, \quad \Psi_2' = \Psi_2, \quad \Psi_4' = e^{-2w}\psi_4. \tag{2.5}$$

where

$$w = a + ib \tag{2.6}$$

is an arbitrary complex function of four real variables. Thus we are able to set

$$\Psi_0 = \Psi_4. \tag{2.7}$$

In order to preserve (2.7) the remaining tetrad freedom (1.7) is restricted as follows:

$$e^{4w} = 1. \tag{2.8}$$

As a consequence of Eqs. (2.1), (2.2), and (2.7) together with the hypotheses I—IV the commutation relations[11] are

$$[\Delta, D] = (\gamma + \bar{\gamma})D + (\epsilon + \bar{\epsilon})\Delta, \tag{2.9}$$

$$[\delta, D] = (\bar{\alpha} + \beta)D - \sigma\bar{\delta} - (\bar{\rho} + \epsilon - \bar{\epsilon})\delta, \tag{2.10}$$

$$[\delta, \Delta] = -(\bar{\alpha} + \beta)\Delta + \bar{\lambda}\bar{\delta} + (\mu - \gamma + \bar{\gamma})\delta, \tag{2.11}$$

$$[\bar{\delta}, \delta] = (\bar{\mu} - \mu) + (\bar{\rho} - \rho)\Delta - (\bar{\alpha} - \beta)\bar{\delta} - (\bar{\beta} - \alpha)\delta, \tag{2.12}$$

the Ricci identities are

$$D\rho = \rho^2 + \sigma\bar{\sigma} + (\epsilon + \bar{\epsilon})\rho, \tag{2.13}$$

$$D\sigma = (\rho + \bar{\rho})\sigma + (3\epsilon - \bar{\epsilon})\sigma + \Psi_0, \tag{2.14}$$

$$D\alpha - \bar{\delta}\epsilon = (\rho + \bar{\epsilon} - 2\epsilon)\alpha + \beta\bar{\sigma} - \bar{\beta}\epsilon, \tag{2.15}$$

$$D\beta - \delta\epsilon = \alpha\sigma + (\bar{\rho} - \bar{\epsilon})\beta - \bar{\alpha}\epsilon, \tag{2.16}$$

$$D\gamma - \Delta\epsilon = -(\epsilon + \bar{\epsilon})\gamma - (\gamma + \bar{\gamma})\epsilon, \tag{2.17}$$

$$D\lambda = \rho\lambda + \bar{\sigma}\mu - (3\epsilon - \bar{\epsilon})\lambda, \tag{2.18}$$

$$D\mu = \bar{\rho}\mu + \sigma\lambda - (\epsilon + \bar{\epsilon})\mu + \Psi_2, \tag{2.19}$$

$$\Delta\lambda = -(\mu + \bar{\mu})\lambda - (3\gamma - \bar{\gamma})\lambda - \Psi_0, \tag{2.20}$$

$$\delta\rho - \bar{\delta}\sigma = \rho(\bar{\alpha} + \beta) - \sigma(3\alpha - \bar{\beta}), \tag{2.21}$$

$$\delta\alpha - \bar{\delta}\beta = \rho\mu - \sigma\lambda + \alpha\bar{\alpha} + \beta\bar{\beta} - 2\alpha\beta + \gamma(\rho - \bar{\rho}) + \epsilon(\mu - \bar{\mu}) - 2\Psi_2, \tag{2.22}$$

$$\delta\lambda - \bar{\delta}\mu = \mu(\alpha + \bar{\beta}) + \lambda(\bar{\alpha} - 3\beta), \tag{2.23}$$

$$\Delta\mu = -\mu^2 - \lambda\bar{\lambda} - (\gamma + \bar{\gamma})\mu, \tag{2.24}$$

$$\delta\gamma - \Delta\beta = -(\bar{\alpha} + \beta)\gamma - \beta(\gamma - \bar{\gamma} - \mu) + \alpha\bar{\lambda}, \tag{2.25}$$

$$\Delta\sigma = -\mu\sigma - \bar{\lambda}\rho + (3\gamma - \bar{\gamma})\sigma, \tag{2.26}$$

$$\Delta\rho = -\rho\bar{\mu} - \sigma\lambda + (\gamma + \bar{\gamma})\rho - \Psi_2, \tag{2.27}$$

$$\Delta\alpha - \bar{\delta}\gamma = -\beta\lambda + (\bar{\gamma} - \bar{\mu})\alpha + \bar{\beta}\gamma, \tag{2.28}$$

the Bianchi identities are

$$D\Psi_0 = (\rho - 4\epsilon)\Psi_0 - \lambda\Psi_2, \quad \Delta\Psi_0 = (4\gamma - \mu)\Psi_0 + \sigma\Psi_2, \tag{2.29a}$$

$$\delta\Psi_0 = 4\beta\Psi_0, \quad \bar{\delta}\Psi_0 = 4\alpha\Psi_0, \tag{2.29b}$$

$$D\Psi_2 = 2(\rho + \bar{\rho})\Psi_2, \quad \Delta\Psi_2 = -2(\mu + \bar{\mu})\Psi_2,$$

$$\delta\Psi_2 = \bar{\delta}\Psi_2 = 0, \tag{2.30}$$

$$(\rho + 2\bar{\rho})\sigma = (\mu + 2\bar{\mu})\lambda, \quad \lambda\Psi_0 = -(\rho + 2\bar{\rho})\Psi_2, \tag{2.31}$$

and the Maxwell's equations are

$$D\phi = 2\rho\phi, \quad \Delta\phi = -2\mu\phi, \quad \delta\phi = \bar{\delta}\phi = 0. \tag{2.32}$$

By applying the commutation relations (2.9)—(2.12) to Maxwell's equations (2.32) we obtain

$$\delta\rho = \rho(\bar{\alpha} + \beta), \quad \bar{\delta}\rho = \rho(\alpha + \bar{\beta}), \tag{2.33}$$

$$\delta\mu = -\mu(\bar{\alpha} + \beta), \quad \bar{\delta}\mu = -\mu(\alpha + \bar{\beta}), \tag{2.34}$$

$$\rho\bar{\mu} = \bar{\rho}\mu. \tag{2.35}$$

Then with the help of equations (2.33)—(2.34) and the Ricci identities (2.21) and (2.23) we obtain

$$\bar{\delta}\sigma = \sigma(3\alpha - \bar{\beta}), \quad \delta\lambda = \lambda(\bar{\alpha} - 3\beta). \tag{2.36}$$

To proceed further, we need expressions for $\delta\sigma$ and $\bar{\delta}\lambda$. These are obtained by applying the commutator $[\delta, D]$ to ρ and the commutator $[\delta, \Delta]$ to μ:

$$\delta\sigma = \sigma(3\beta - \bar{\alpha}), \quad \bar{\delta}\lambda = \lambda(\bar{\beta} - 3\alpha). \tag{2.37}$$

If we now operate with δ and $\bar{\delta}$ on the relationship given by Eq. (2.31) and use Eqs. (2.33), (2.34), (2.36), and (2.37), we find that

$$\alpha = \beta = 0. \tag{2.38}$$

By operating on (2.35) with D and Δ and using (2.13), (2.19), (2.24), (2.27) and (2.31), we obtain respectively the expressions

$$(\rho-\bar{\rho})(\Psi_2+2\rho\bar{\mu})+(\bar{\mu}-\mu)\sigma\bar{\sigma}+\rho\bar{\delta}\bar{\lambda}-\bar{\rho}\sigma\lambda=0, \tag{2.39}$$

$$(\mu-\bar{\mu})(\Psi_2+2\bar{\rho}\mu)+(\bar{\rho}-\rho)\lambda\bar{\lambda}+\mu\bar{\sigma}\bar{\lambda}-\bar{\mu}\sigma\lambda=0. \tag{2.40}$$

When ρ is not real the equations (2.39) and (2.40) immediately yield the following tetrad invariant expression:

$$\rho^2\lambda\bar{\lambda}=\mu^2\sigma\bar{\sigma}. \tag{2.41}$$

If ρ is real (2.22) and (2.31) yield

$$2\Psi_2=\rho\mu-\sigma\lambda, \tag{2.42a}$$

$$\lambda\Psi_0+3\rho\Psi_2=0, \tag{2.42b}$$

$$\rho\sigma-\mu\lambda=0. \tag{2.42c}$$

By operating on (2.42b) with Δ and using (2.20), (2.27), (2.29a), (2.30), (2.42a), and (2.42c) we obtain

$$\sigma\lambda+3\rho\mu=0. \tag{2.43}$$

By operating on (2.43) with D and using (2.13), (2.14), (2.18), and (2.19) we obtain the expression

$$3\rho^2=\sigma\bar{\sigma}. \tag{2.44}$$

Similarly, operating on (2.43) with Δ and using the pertinent Ricci identities we obtain

$$3\mu^2=\lambda\bar{\lambda}. \tag{2.45}$$

From (2.44) and (2.45) it follows that the relationship (2.41) is also true when ρ is real.

Next we observe that the three expressions (2.31), (2.35), and (2.41) relating the spin coefficients ρ, μ, σ, and λ imply that if ρ is either complex or imaginary, we obtain

$$\rho=\mu, \quad \sigma=\lambda \tag{2.46}$$

or

$$\rho=-\mu, \quad \sigma=-\lambda. \tag{2.47}$$

For the case ρ real (2.42c) implies

$$\rho^2\sigma\bar{\sigma}=\mu^2\lambda\bar{\lambda} \tag{2.48}$$

on comparing (2.48) with (2.41) we find (2.46) or (2.47).

This completes the proof of Theorem 1.

3. PROOF OF THEOREM 2

From Eq. (2.38) we have $\alpha=\beta=0$. Thus it follows that the result of operating on the quantities ρ, σ, ϵ, ϕ, Ψ_0, Ψ_2, with δ or $\bar{\delta}$ is always zero. A further consequence of this is that for ρ not real the commutator (2.12) becomes

$$D+\Delta=0. \tag{3.1}$$

Applying (3.1) to Ψ_0 and using (2.29a), we obtain

$$\gamma=\epsilon; \tag{3.2}$$

similarly applying (3.1) to ϵ, we have

$$D\epsilon+\Delta\epsilon=0. \tag{3.3}$$

Comparing (3.3) with the Ricci identity (2.17) and using (3.2), we obtain the following equation:

$$D\epsilon=-(\epsilon+\bar{\epsilon})\epsilon. \tag{3.4}$$

We next show that (3.2) also holds in the case ρ real. To see this, we first note from Eq. (2.22) that

$$\sigma^2=\bar{\sigma}^2. \tag{3.5}$$

By operating on (3.5) with D and using (2.18) we obtain

$$\epsilon=\bar{\epsilon}. \tag{3.6}$$

Similarly operating on (3.5) with Δ and using (2.26), we have

$$\gamma=\bar{\gamma}. \tag{3.7}$$

Finally we recall Eq. (2.44) and operate on it with D and Δ to obtain the following expressions:

$$\sigma^2=9\rho^2+12\rho\epsilon, \quad \sigma^2=9\rho^2+12\rho\gamma. \tag{3.8}$$

Hence (3.8) immediately implies

$$\gamma=\epsilon. \tag{3.9}$$

Consequently, for ρ real (3.1) and (3.4) also hold.

Before proceeding further with the proof of Theorem 2 we summarize the problem. The hypotheses of Theorem 2 are equivalent to the following set of equations:

$$\mu=\rho, \quad \lambda=\sigma, \quad \gamma=\epsilon,$$
$$\kappa=\nu=\pi=\tau=\alpha=\beta=0, \tag{3.10}$$

$$\Psi_4=\Psi_0, \quad \Psi_2+\Phi_{11}=0,$$
$$\Psi_1=\Psi_3=\Phi_{ab}=0 \ \text{ unless } a=b=1, \tag{3.11}$$

$$D+\Delta=0, \quad \delta=\bar{\delta}=0, \tag{3.12}$$

$$D\epsilon=-(\epsilon+\bar{\epsilon})\epsilon, \tag{3.13}$$

$$D\rho=\rho^2+\sigma\bar{\sigma}+\rho(\epsilon+\bar{\epsilon}), \tag{3.14}$$

$$D\sigma=\rho(\sigma+\bar{\sigma})-3\epsilon\sigma+\bar{\epsilon}\sigma, \tag{3.15}$$

$$D\phi=\rho\phi, \tag{3.16}$$

$$D\Psi_0=(\rho-4\epsilon)\Psi_0-\sigma\Psi_2, \quad D\Psi_2=2(\rho+\bar{\rho})\Psi_2, \tag{3.17}$$

$$\sigma\Psi_0=-(\rho+2\bar{\rho})\Psi_2, \quad 2\Psi_2=\rho^2-\sigma^2+2\epsilon(\rho-\bar{\rho}), \tag{3.18}$$

$$\rho^2-\sigma^2+2\sigma\bar{\sigma}-2\rho\bar{\rho}+2\rho\epsilon+2\bar{\rho}\epsilon+4\rho\bar{\epsilon}=0, \tag{3.19}$$

$$\sigma^2(12\epsilon+4\bar{\rho}-4\bar{\epsilon}+\rho)+4\epsilon\bar{\rho}^2-\rho^3-2\rho\sigma\bar{\sigma}$$
$$-2\rho^2\epsilon-2\rho\bar{\rho}(\epsilon+\rho)=0. \tag{3.20}$$

The proof of Theorem 2 splits into three cases: (i) ρ real, (ii) ρ complex, (iii) ρ imaginary. In Secs. 4, 5, and 6 we consider separately these three cases.

4. THE CASE ρ REAL

From Eqs. (2.44), (3.8), and (3.19) we deduce

$$\rho+\epsilon=0. \tag{4.1}$$

Since $\sigma^2=\bar{\sigma}^2$ from (3.5) we must consider two possibilities: (i) $\sigma=\bar{\sigma}$ and (ii) $\sigma=-\sigma$. In the first case we find from (2.44) and (3.8) that

$$\rho+2\epsilon=0. \tag{4.2}$$

However, (4.1) and (4.2) imply $\rho=\epsilon=0$, which is impossible.

In the second case (3.19) yields

$$3\sigma^2 = 8\epsilon\rho - \rho^2. \tag{4.3}$$

By eliminating σ^2 between (4.3) and (3.20) we obtain

$$4\epsilon^2 + 3\epsilon\rho - \rho^2 = 0. \tag{4.4}$$

Operating with D on (4.4) and using (3.13), (3.14), and (4.3) yields

$$\rho^3 - 2\epsilon\rho^2 + 3\epsilon^2\rho + 2\epsilon^3 = 0. \tag{4.5}$$

It follows from (4.4) and (4.5) that

$$\epsilon(\rho^2 + 7\rho\epsilon + 2\epsilon^2) = 0. \tag{4.6}$$

Thus either $\epsilon = 0$, which is impossible, or

$$\rho^2 + 7\rho\epsilon + 2\epsilon^2 = 0 \tag{4.7}$$

However, Eqs. (4.4) and (4.7) also lead to an impossibility.

5. THE CASE ρ COMPLEX

Recently Debever[12] has proved the following theorem:

Theorem 3: Suppose there exists a null frame such that

$$ds^2 = 2(\theta^0\theta^3 - \theta^1\theta^2) \tag{5.1}$$

and

$$\overset{*}{F} = \phi(\theta^0 \wedge \theta^3 - \theta^1 \wedge \theta^2). \tag{5.2}$$

If the vacuum Maxwell's equations

$$d\overset{*}{F} = 0,$$

are satisfied and if the spin coefficients satisfy the following conditions,[13]

$$\kappa = \nu, \quad \pi = \tau, \quad \rho = \mu, \quad \epsilon = \gamma, \quad \alpha = \beta, \tag{5.3}$$

then the space–time admits a two-dimensional Abelian invertible isometry group with timelike two-dimensional orbits. In other words, there exists a system of coordinates (t, z, x, y) in which the Killing vectors and the metric have the form respectively

$$\frac{\partial}{\partial t}, \quad \frac{\partial}{\partial z}, \tag{5.4}$$

$$ds^2 = (Ldt + Mdz)^2 - (Ndt + Pdz)^2 - S^2\left(\frac{dx^2}{X^2} + \frac{dy^2}{Y^2}\right), \tag{5.5}$$

where L, M, P, N, S are functions only of x and y. The functions $X = X(x)$ and $Y = Y(y)$ are introduced here purely for mathematical convenience. The appropriate null frame is[14]

$$\theta^0 = l_a dx^a = \frac{1}{\sqrt{2}}\left[Ldt + Mdz + S\frac{dy}{Y}\right], \tag{5.6}$$

$$\theta^1 = m_a dx^a = \frac{1}{\sqrt{2}}\left[Ndt + Pdz - iS\frac{dx}{X}\right], \tag{5.7}$$

$$\theta^2 = \bar{m}_a dx^a = \bar{\theta}^1, \tag{5.8}$$

$$\theta^3 = n_a dx^a = \frac{1}{\sqrt{2}}\left[Ldt + Mdz - S\frac{dy}{Y}\right], \tag{5.9}$$

In view of (1.2), (1.3), (3.10), and (3.11) it is clear that the hypothesis of Debever's theorem is satisfied in the situation under consideration. Thus it remains only to solve the Einstein–Maxwell equations (1.1) with the metric (5.5) for the condition IV.

The spin coefficients are

$$\kappa = \nu = \frac{1}{2\sqrt{2}}\left[\frac{Y}{S}(mL_y - lM_y) + i\frac{X}{S}\left(pL_x - nM_x - \frac{S_x}{S}\right)\right], \tag{5.10}$$

$$\pi = \tau = \frac{1}{2\sqrt{2}}\left[\frac{Y}{S}(pN_y - nP_y) + \frac{iX}{S}\left(pL_x - nM_y + \frac{S_x}{S}\right)\right], \tag{5.11}$$

$$\sigma = \lambda = \frac{1}{2\sqrt{2}}\left[\frac{Y}{S}\left(lP_y - mN_y - \frac{S_y}{S}\right) + \frac{iX}{S}(nP_x - pN_x)\right], \tag{5.12}$$

$$\rho = \mu = \frac{1}{2\sqrt{2}}\left[\frac{Y}{S}\left(lP_y - mN_y + \frac{S_y}{S}\right) + \frac{iX}{S}(lM_x - mL_x)\right], \tag{5.13}$$

$$\epsilon = \gamma = \frac{1}{2\sqrt{2}}\left[\frac{Y}{S}(nM_y - pL_y) + \frac{iX}{S}(nP_x - pN_x + lM_x - mL_x)\right], \tag{5.14}$$

$$\alpha = \beta = \frac{1}{2\sqrt{2}}\left[\frac{Y}{2S}(mL_y - lm_y + pN_y - nP_y) + \frac{iX}{S}(mN_x - lP_x)\right], \tag{5.15}$$

where $m = M/(LP - MN)$ etc.

The vanishing of the spin coefficients κ, π, and α produces the following simplification:

$$L = L(x, y), \quad M = AL, \quad N = C(B - A)^{-1}L^{-1}, \tag{5.16}$$

$$P = BN, \quad S = S(y), \quad M_x = BL_x, \tag{5.17}$$

where A and B are functions of x and C is a function of y.

The differential operators are given by

$$D = (1/\sqrt{2})(p\partial_t - n\partial_z - YS^{-1}\partial_y), \tag{5.18}$$

$$\delta = -(1/\sqrt{2})(m\partial_t - l\partial_z + iXS^{-1}\partial_x), \tag{5.19}$$

$$\Delta = (1/\sqrt{2})(p\partial_t - n\partial_z + YS^{-1}\partial_y). \tag{5.20}$$

However, since all the functions occuring in (5.5) involve atmost only the variables x and y, they simplify to

$$D = -\Delta = -(1/\sqrt{2})YS^{-1}\partial_y, \quad \delta = -i(1/\sqrt{2})XS^{-1}\partial_x. \tag{5.21}$$

Furthermore, since S is only a function of y, we therefore choose the free function Y such that

$$Y = S. \tag{5.22}$$

In view of (5.22), (5.21) becomes

$$D = -\Delta = -(1/\sqrt{2})\partial_y, \quad \delta = -i(1/\sqrt{2})XS^{-1}\partial_x. \tag{5.23}$$

Using (3.12) and integrating (3.13), we obtain

$$\epsilon = k(1 + ie)(y + d)^{-1}, \tag{5.24}$$

where $k = -1/2\sqrt{2}$ and e and d are arbitrary constants. Comparing (5.14) and (5.24), we find that the metric function L must be of the form

$$L = F(y + d), \tag{5.25}$$

where F is a function of x only.

Next it is found expedient to choose the free function X such that

$$X = B - A. \tag{5.26}$$

Using (5.26) and the fact that by (3.12)

$$(\mathrm{Im}\rho)_x = 0, \tag{5.27}$$

we have

$$(F^2)_{xx} = 0. \tag{5.28}$$

Integration yields

$$F^2 = ax + b, \tag{5.29}$$

where a, b are arbitrary constants. Similarly we have the condition

$$(\mathrm{Im}\sigma)_x = 0, \tag{5.30}$$

and consequently after integration we find that

$$B - A = (ax + b)(cx + f)^{-1/2}, \tag{5.31}$$

where c, f are arbitrary constants. Finally the integration of the differential equation in (5.17) yields the following expression for A [15]:

$$A = a(bc - af)^{-1}(ax + b)^{-1/2}(cx + f)^{1/2} + g, \tag{5.32}$$

where g is an arbitrary constant.

If we denote the y-dependent part of N by N_0, we find from (5.12), (5.13), (5.14) the equation

$$cN_0^2 - 4eN_0S - a(y + d)^2 = 0. \tag{5.33}$$

The expressions for the spin coefficients ρ and σ read respectively

$$\rho = -k[(\ln N_0 S)_y + i(\tfrac{1}{2})a(y + d)N_0^{-1}S^{-1}], \tag{5.34}$$

$$\sigma = -k[(\ln(N_0 S^{-1}))_y - i(\tfrac{1}{2})cN_0 S^{-1}(y + d)^{-1}]. \tag{5.35}$$

From Eqs. (3.11), (3.12), it is found that N_0 satisfies the following three ordinary differential equations:

$$aN_0'^2Y^4 - 2aN_0N_0'Y^3 + (cN_0^2N_0'^2 - 4aN_0^2 - ae^2N_0^2)Y^2 + 4cN_0^3N_0'Y - 3ce^2N_0^4 = 0, \tag{5.36}$$

$$(cN_0^3N_0'' + ae^2N_0^2)Y^2 - aN_0'^2Y^4 + 2aN_0N_0'Y^3 - cN_0^3N_0'Y + ce^2N_0^4 = 0, \tag{5.37}$$

$$(2aN_0N_0'' - 2aN_0'^2)Y^4 + 6aN_0N_0'Y^3 - 4aN_0^2Y^2 - 4ce^2N_0^4 = 0, \tag{5.38}$$

where Y denotes $y + d$ and prime indicates differentiation with respect to y.

Addition of (5.36) and (5.37) yields an Euler differential equation

$$Y^2(N_0^2)'' + 3Y(N_0^2)' - 4e^2(N_0^2) = (8a/c)Y^2. \tag{5.39}$$

If $e^2 - 2 \neq 0$, the general solution of (5.39) is

$$N_0^2 = c_1Y^{m_1} + c_2Y^{m_2} + 2a(2 - e^2)^{-1}c^{-1}Y^2, \tag{5.40}$$

where c_1, c_2 are arbitrary constants and $m_1 = -1 + (1 + 4e^2)^{1/2}$, $m_2 = -1 - (1 + 4e^2)^{1/2}$. For $e^2 - 2 = 0$, the general solution to (5.39) has the form

$$N_0^2 = c_1Y^2 + c_2Y^{-4} + (4a/3c)Y^2 \ln Y, \tag{5.41}$$

where c_1, c_2 are arbitrary constants. However, neither of the solutions (5.40) or (5.41) satisfy the differential equation (5.38).

Thus there is no solution to the conditions I–IV for ρ complex in the case $\rho = \mu$, $\sigma = \lambda$.

6. THE CASE ρ IMAGINARY

From equations (3.19) and (3.20) it follows that

$$\sigma^2 = (\bar{\rho}\rho^2 - \bar{\epsilon}\rho^2)(\bar{\rho} + 3\epsilon - \bar{\epsilon})^{-1}. \tag{6.1}$$

Also from Eqs. (3.14), (3.19) we obtain

$$\sigma^2 = \rho^2 + 4\rho\epsilon. \tag{6.2}$$

Since $\rho + \bar{\rho} = 0$, we deduce from (6.1) and (6.2) that

$$\rho[\rho(\epsilon - \bar{\epsilon}) + \bar{\epsilon}(3\epsilon - \bar{\epsilon})] = 0. \tag{6.3}$$

This in turn implies that for a nontrivial situation we must have

$$\epsilon + \bar{\epsilon} = 0. \tag{6.4}$$

Substituting (6.4) into (6.3) and (6.2), we find that

$$\rho = 2\epsilon, \quad \sigma = \pm i\rho. \tag{6.5}$$

The ambiguity in the sign of σ may now be removed by using the remaining tetrad freedom given by (2.8). Hence we choose the tetrad so that

$$\sigma = -i\rho. \tag{6.6}$$

In view of (6.4) the differential equation (3.13) has solution

$$\epsilon = ik, \tag{6.7}$$

where k is an arbitrary positive constant.

The set of differential equations obtained on equating the values of the spin coefficient given above with their values given by (5.15), (5.16), (5.17) is

$$\begin{aligned} &mL_y - lN_y = 0, \quad pN_y - nP_y = 0, \quad lP_x - mN_x = 0, \\ &pL_x - nM_x + S_x/S = 0, \quad pL_x - nM_x - S_x/S = 0, \\ &nP_x - pN_x = 0, \quad lP_y - mN_y + S_y/S = 0, \\ &nM_y - pL_y = 0, \quad (Y/S)(lP_y - mN_y - S_y/S) = 4\sqrt{2}\,k, \\ &(X/S)(lM_x - mL_x) = 4\sqrt{2}\,k. \end{aligned} \tag{6.8}$$

Integration of (6.8) in the manner of Sec. 5 yields

$$\begin{aligned} &L = 2kx + l, \qquad M = 2kbx + m, \quad N = n\exp(2\sqrt{2}\,ky), \\ &P = bN, \quad S = s\exp(-2\sqrt{2}\,ky), \quad Y = S, \quad X = 2\sqrt{2}\,ns, \end{aligned} \tag{6.9}$$

when b, l, m, n, s are arbitrary constants. The metric thus becomes

$$ds^2 = [(2kx + l)\,dt + (2kbx + m)\,dz]^2 - n^2\exp(4\sqrt{2}\,ky)(dt + bdz)^2 - dy^2 - \exp(-4\sqrt{2}\,ky)\,dx^2. \tag{6.10}$$

By means of the substitutions

$$\begin{aligned} &t = -(n^{-1}y' + b(m - lb)^{-1}(t'), \quad x = 2\sqrt{2}\,nx' - \tfrac{1}{2}l/k, \\ &y = v, \quad z = (m - lb)^{-1}(t'), \qquad k = 4\sqrt{2}\,k', \end{aligned} \tag{6.11}$$

and on dropping the primes we recover the metric (1.9).

The integration of Maxwell's equations yields the expression (1.10).

The case $\rho = -\mu$ and $\sigma = -\lambda$ of Theorem 1 will be discussed in a subsequent paper.

ACKNOWLEDGMENT

The authors wish to thank Professor R. Debever for many stimulating discussions.

*This work was supported in part by a grant from the National Research Council of Canada.

[1]N. Tariq and B.O.J. Tupper, Tensor N.S. **28**, 83 (1974).
[2]N. Tariq and B.O.J. Tupper, Gen. Relativ. Gravit. **6**, 777 (1975).
[3]R.G. McLenaghan and N. Tariq, J. Math. Phys. **16**, 2306 (1975).
[4]B.O.J. Tupper, Gen. Relativ. Gravit., to appear.
[5]The conventions of this paper are the same as those of Ref. 3.
[6]See Ref. 3.
[7]These properties were pointed out to us by Prof. R. Debever (private communication).
[8]E.T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[9]R. Debever, Cah. Phys. 168–169, 303 (1964).
[10]R. Debever, Bull. Cl. Sci. Acad. Roy. Belg. **60**, 998 (1974).
[11]See Ref. 8.
[12]R. Debever, communication at the Conference on Differential Geometrical Methods in Mathematical Physics, Bonn, July 1975.
[13]These conditions are satisfied if and only if there exists a tetrad transformation $\theta^0 \to -\theta^3 \to -\theta^0$, $\theta^1 \to -\theta^2$, $\theta^2 \to -\theta^1$, which leaves the spin coefficients κ, ρ, σ, τ, ϵ, α invariant.
[14]We note that the metric (5.5) is invariant under the finite isometry $t \to -t$, $z \to -z$ but the nonsingular electromagnetic field gives rise to an involution since it implies that $F^* \to -F^*$.
[15]If $bc = af$, we obtain an integral of the form $\int dx/x(ax+b)^{3/2}$ which is not difficult to integrate. However, since we never need anywhere the explicit forms of A or B, we do not elaborate this calculation further.

A note on certain multiple integrals

M. L. Mehta*

Japan Atomic Energy Research Institute, Tokai-mura, Nakagan, Ibaraki-ken, Japan
(Received 30 June 1976)

An analytical expression is given for the multiple integral

$\int\cdots\int d\mu(x_{n+1})\, d\mu(x_{n+2}) \cdots d\mu(x_N)\Pi_{1\le j<k\le N}|x_j - x_k|^\beta$,

where $0 \le n \le N$, $\beta = 1,2$, or 4 and the positive measure $d\mu(x)$ is such that all its moments exist, $\int d\mu(x)x^j < \infty$, $j = 0,1,2,\cdots$. The case $d\mu(x) = dx$, for $-1 \le x \le 1$, and $d\mu(x) = 0$, for $|x| > 1$, is given as an example. In the limit $N \to \infty$ the correlation functions of this example, the so-called Legendre ensembles, coincide with those of the circular or the Gaussian ensembles of random matrices.

1. INTRODUCTION

Recently Thien Vo-Dai and J.R. Derome studied the so called Legendre ensenbles,[1] defined by the joint probability density

$$Q_{N\beta}(x_1, x_2, \ldots, x_N) = \text{const} \prod_{1\le j<k\le N} |x_j - x_k|^\beta, \tag{1.1}$$

where $\beta = 1$, 2, or 4 and $|x_j| \le 1$. They found that in the limit $N \to \infty$ the spacing distribution for these ensembles is identical with that for the Gaussian or circular ensembles.[1,2] They also calculated the two-level correlation functions, which amounts to integrating $Q_{N\beta}$ over all the variables except 2, and found that in the limit $N \to \infty$ the result is again identical with that for Gaussian or circular ensembles. For this purpose they used an old method,[2] which is tedious and becomes prohibitive for higher correlations. The new powerful method[3,4] seems not to be widely noticed or appreciated for the apparent reason that it was never explained in enough detail. We hereby try to remedy this defect.

The method is explained in Secs. 2—5. It is hard to avoid some repetitions. We then apply the method to the particular case of Legendre ensembles in Sec. 6. In the limit $N \to \infty$, correlation functions of all orders (and not only of order two) are identical to those for the Gaussian or circular ensembles. Thus the local statistical properties seem to be quite insensitive to details of the matrix ensemble considered.

The subject matter of Secs. 2—5 forms a part of the planned publication Elements of Matrix Theory, where motivation for various manipulations and a fuller account can be found.

2. ALTERNANTS AND CONFLUENT ALTERNANTS

Let $p_1(x), p_2(x), \ldots, p_N(x)$ be a set of N linearly independent polynomials, the degree of $p_j(x)$ being less than or equal to $N-1$ for $j = 1, 2\ldots, N$. Then

$$\det[p_j(x_k)]_{j,k=1,2,\ldots,N} = \text{const} \prod_{1\le j<k\le N} (x_j - x_k), \tag{2.1}$$

as can be seen from the fact that the left-hand side vanishes on setting $x_j = x_k$. Similarly if $\{p_j(x)\}$, $j = 1, 2, \ldots, 2N$ are $2N$ linearly independent polynomials, the degree of each of them being this time less than or equal to $2N-1$ and $\{p_j'(x)\}$ their derivatives, then the $2N \times 2N$ confluent alternant $[p_j(x_k)\, p_j'(x_k)]$, $j = 1, 2, \ldots, 2N$, $k = 1, 2, \ldots, N$, has the determinant[5]

$$\det[p_j(x_k)\ p_j'(x_k)] = \text{const} \prod_{j<k} (x_j - x_k)^4. \tag{2.2}$$

The constants in Eqs. (2.1) and (2.2) will depend on the choice of $p_j(x)$. In particular, if

$$p_j(x) = \sum_{k=1}^{N} u_{jk} x^{k-1},$$

then

$$\det[p_j(x_k)] = \det U \det[x_j^{k-1}] = \det U \prod_{j<k} (x_k - x_j).$$

Similar remarks apply for the confluent alternant.

We will write $Q_{N\beta}(x_1, x_2, \ldots, x_N)$ as an alternant or a confluent alternant, choosing $p_j(x)$ conveniently.

3. DETERMINANTS OF MATRICES WITH QUATERNION ELEMENTS

A quaternion[6] is a linear combination of four units

$$q = q_0 + \mathbf{q}\cdot\mathbf{e} = q_0 + q_1 e_1 + q_2 e_2 + q_3 e_3, \tag{3.1}$$

q_0, q_1, q_2, q_3 are real or complex numbers; the four basic units satisfy the multiplication laws

$$\begin{aligned} &1 \cdot 1 = 1, \quad 1 \cdot e_j = e_j \cdot 1 = e_j, \quad j = 1, 2, 3, \\ &e_1^2 = e_2^2 = e_3^2 = e_1 e_2 e_3 = -1, \end{aligned} \tag{3.2}$$

and multiplication is associative. The scalar part of q is q_0. Any quaternion q has the dual

$$\bar{q} = q_0 - \mathbf{q}\cdot\mathbf{e} \tag{3.3}$$

A matrix Q with quaternion elements q_{ij} has the dual matrix $\bar{Q} = [\bar{q}_{ji}]$. One may sometimes think of the quaternion units as 2×2 complex matrices (related to the so-called Pauli matrices in physics)

$$\begin{aligned} &1 \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad e_1 \to \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \\ &e_2 \to \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}, \quad e_3 \to \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}. \end{aligned} \tag{3.4}$$

Among the various possibilities we choose the following definition of a determinant,[3,7] denoting it as Tdet. For

self-dual Q (i.e., when $Q=\bar{Q}$),

$$\mathrm{Tdet}Q=\sum_P(-1)^{N-l}\prod_1^l(q_{ab}q_{bc}\cdots q_{da})_0, \tag{3.5}$$

where P is any permutation of the indices $(1, 2, \ldots, N)$ consisting of l exclusive cycles of the form

$$(a\to b\to c\to\cdots\to d\to a), \tag{3.6}$$

and $(-1)^{N-l}$ is the parity of P. The subscript 0 means that we take the scalar part of the product over each cycle. We will need two theorems.

Theorem 3.1: In a self-dual matrix A, $A=\bar{A}$, replace its quaternion elements by their 2×2 matrix representatives, Eq. (3.4), getting a matrix $C(A)$ of twice the size with complex elements. Then

$$\det C(A)=(\mathrm{Tdet}A)^2. \tag{3.7}$$

For a proof see Ref. 3.

Theorem 3.2: Consider N real or complex variables $x_1, x_2, \ldots, x_N$. Let the $N\times N$ matrix A_N with quaternion elements depend on these variables as follows:

$$a_{jk}=f(x_j, x_k)=\bar{a}_{kj}, \tag{3.8}$$

i.e., a_{jk} depends only on x_j and x_k. The matrix A_N is self-dual, $A_N=\bar{A}_N$. Moreover, let

$$\int f(x, x)d\mu(x)=c, \tag{3.9}$$

$$\int f(x, y)f(y, z)\,d\mu(y)=f(x, z)+g(x, z), \tag{3.10}$$

$$g(x, y)=\lambda f(x, y)-f(x, y)\lambda, \tag{3.11}$$

where $d\mu$ is a suitable measure, c a constant scalar, i.e., $c=\bar{c}$, and λ a constant quaternion. Then

$$\int \mathrm{Tdet}A_N\,d\mu(x_N)=(c-N+1)\mathrm{Tdet}A_{N-1}, \tag{3.12}$$

where A_{N-1} is the $(N-1)\times(N-1)$ matrix obtained by removing the row and column containing x_N.

For a proof see Ref. 3 or Ref. 9.

4. ORTHOGONAL AND SKEW-ORTHOGONAL POLYNOMIALS

A linear form of two quantities f and g will be called their product and will be denoted as (f, g). This product may have some symmetry. Let $c_{j,k}=c_{k,j}=(x^j, x^k)_s$ be a symmetric product and $\alpha_{j,k}=-\alpha_{k,j}=(x^j, x^k)_a$ be an antisymmetric product of x^j and x^k. Examples are

$$c_{j,k}=c_{j+k}=\int x^{j+k}\,d\mu(x), \tag{4.1}$$

$$\alpha_{j,k}=(k-j)\int x^{j+k-1}\,d\mu(x), \tag{4.2}$$

or

$$\alpha_{j,k}=\int\int x^j y^k\,\mathcal{E}(y-x)\,d\mu(x)\,d\mu(y), \tag{4.3}$$

where $\mathcal{E}(x)=x/|x|$, for $x\neq0$, and $\mathcal{E}(0)=0$. The symmetric or antisymmetric product of two polynomials is then defined by the linearity of the product.

Let $p_j(x)$ be a polynomial of precise degree j. Any monomial x^k can be uniquely expressed as a linear combination of the $p_j(x)$, for $j\leq k$. Hence, if $(p_j(x), p_k(x))_s=\delta_{jk}$ for every j and k, then $(p_j(x), x^k)_s=0$, for $0\leq k<j$ and $(p_j(x), x^j)\neq0$. Therefore we can write $p_j(x)$ as a determinant. (See for example, Ref. 8.) Similarly, if $(p_j(x), p_k(x))_a=Z_{jk}$ where Z is the canonical antisymmetric matrix,

$$Z_{jk}=-Z_{kj}=1, \quad \text{if } j=k-1=\text{even},$$
$$=0, \quad \text{otherwise},$$

i.e., written as a direct sum,

$$Z=\begin{bmatrix}0&1\\-1&0\end{bmatrix}\dotplus\begin{bmatrix}0&1\\-1&0\end{bmatrix}\dotplus\cdots\dotplus\begin{bmatrix}0&1\\-1&0\end{bmatrix}, \tag{4.4}$$

then $(p_{2j}(x), x^k)_a=(p_{2j+1}(x), x^k)_a=0$, for $0\leq k<2j$, and $(p_{2j+1}(x), x^{2j})_a\neq0$. Therefore, one may again write $p_j(x)$ as a determinant.[9]

If $c_{j,k}=(x^j, x^k)_s=0$, for $j+k$ odd, then each orthogonal polynomial has a definite parity $p_j(-x)=(-1)^jp_j(x)$. Similarly, if $\alpha_{j,k}=(x^j, x^k)_a=0$, for $j+k$ even, then each skew-orthogonal polynomial has a definite parity,[9]

$$p_j(-x)=(-1)^jp_j(x).$$

5. THE PRODUCT OF DIFFERENCES AS A Tdet

To evaluate the multiple integral

$$F_{N\beta}(x_1, \ldots, x_n)=\text{const}\int\cdots\int\prod_{1\leq j<k\leq N}|x_j-x_k|^\beta$$
$$\times d\mu(x_{n+1})\,d\mu(x_{n+2})\cdots d\mu(x_N), \tag{5.1}$$

we express the above integrand as a $\mathrm{Tdet}A_N$, where A_N is an $N\times N$ matrix with real or quaternion elements. Theorem 3.2 applied several times will then give the above integral as a Tdet of an $n\times n$ matrix for any n lying between 0 and N, $0\leq n\leq N$.

A. The case $\beta=2$

Every real symmetric matrix can be diagonalized by a real orthogonal matrix. Let $M=[c_{i+j-2}]$, $i, j=1, 2, \ldots, N$, where $c_i=\int x^i\,d\mu(x)$. This M is symmetric, so let $M=U^TBU$, where U is real orthogonal and B is diagonal, $U=[u_{ij}]$, $B=[b_i\delta_{ij}]$. Choose

$$p_i(x)=b_i^{-1/2}\sum_{j=1}^N u_{ij}x^{j-1}, \quad i=1, 2, \ldots, N. \tag{5.2}$$

The polynomials $\{p_i(x)\}$ are orthonormal,

$$\int p_i(x)p_j(x)\,d\mu(x)=(b_ib_j)^{-1/2}(UMU^T)_{ij}=\delta_{ij}. \tag{5.3}$$

Let

$$f_2(x, y)=\sum_{i=1}^N p_i(x)p_i(y). \tag{5.4}$$

This f_2 satisfies Eqs. (3.9) and (3.10) with $c=N$ and $g=0$. Also

$$\det[f_2(x_i, x_j)]=\det\left[\sum_{k=1}^N p_k(x_i)p_k(x_j)\right]_{i,j=1,2,\ldots,N}$$
$$=\{\det[p_i(x_j)]\}^2=(\det M)^{-1}\prod_{1\leq i<j\leq N}(x_i-x_j)^2. \tag{5.5}$$

B. The case $\beta = 4$

Every real antisymmetric matrix can be quasidiagonalized by a real orthogonal matrix.[9] Let then $M = [\alpha_{j,k}] = U^T B Z U$, where $\alpha_{j,k}$ are given by Eq. (4.2), $UU^T = I$, $B = [b_i \delta_{ij}]$, and Z is given by Eq. (4.4), all the matrices being of order $2N \times 2N$. Choose with the elements u_{ij} of U,

$$p_i(x) = b_i^{-1/2} \sum_{j=1}^{2N} u_{ij} x^{j-1}, \quad i = 1, 2, \ldots, 2N. \tag{5.6}$$

The polynomials $\{p_i(x)\}$ are skew-orthonormal,

$$\int \left\{ p_i(x) \frac{d}{dx} p_j(x) - p_j(x) \frac{d}{dx} p_i(x) \right\} d\mu(x) = Z_{ij}, \tag{5.7}$$

where the antisymmetric matrix Z is given by Eq. (4.4),

$$z_{ij} = -z_{ji} = 1, \quad \text{if } i = j + 1 = \text{odd},$$
$$= 0, \quad \text{otherwise}. \tag{5.8}$$

Let us denote the derivative of $p_i(x)$ by $q_i(x)$,

$$q_i(x) = \frac{d}{dx} p_i(x). \tag{5.9}$$

Define

$$S(x, y) = \sum Z_{ij} p_i(x) q_j(y), \tag{5.10}$$

$$S^+(x, y) = S(y, x) = -\sum Z_{ij} q_i(x) p_j(y), \tag{5.11}$$

$$D(x, y) = -\sum Z_{ij} p_i(x) p_j(y), \tag{5.12}$$

$$I(x, y) = \sum Z_{ij} q_i(x) q_j(y), \tag{5.13}$$

all the sums running from 1 to $2N$. Put

$$f_4(x, y) = \begin{bmatrix} S(x, y) & D(x, y) \\ I(x, y) & S^+(x, y) \end{bmatrix}. \tag{5.14}$$

This quaternion f_4 satisfies Eqs. (3.9) and (3.10) with $c = N$ and $g = 0$. Also from Theorem 3.1

$$\{\mathrm{Tdet}[f_4(x_i, x_j)]\}^2 = \det[C(f_4(x_i, x_j))],$$

and

$$C(f_4(x_i, x_j)) = \begin{bmatrix} p_i(x_j) \\ q_i(x_j) \end{bmatrix} Z [q_i(x_j) \quad -p_i(x_j)], \tag{5.15}$$

so that

$$\det[C(f_4)] = \{\det[p_i(x_j) \quad q_i(x_j)]\}^2$$
$$= (\det M)^{-1} \prod_{1 \le i < j \le N} (x_i - x_j)^8. \tag{5.16}$$

In the last step we have used Eq. (2.2).

C. The case $\beta = 1$, N even

Let $\alpha_{j,k}$ be given by Eq. (4.3). Quasidiagonalize the $N \times N$ antisymmetric real matrix $M = [\alpha_{j,k}] = U^T B Z U$, where U is real orthogonal, B is diagonal, and Z is given by Eq. (4.4). With the matrix elements u_{ij} of U choose

$$p_i(x) = b_i^{-1/2} \sum_{j=1}^{N} u_{ij} x^{j-1}, \quad i = 1, 2, \ldots, N. \tag{5.17}$$

The polynomials $\{p_i(x)\}$ are skew-orthonormal,

$$\int \int p_i(x) p_j(y) \mathcal{E}(y - x)\, d\mu(x)\, d\mu(y) = Z_{ij}. \tag{5.18}$$

Let

$$q_j(x) = \int p_j(y) \mathcal{E}(y - x)\, d\mu(y). \tag{5.19}$$

Let $S(x, y)$, $S^+(x, y)$, $D(x, y)$, and $I(x, y)$ be again given by Eqs. (5.10)–(5.13), the sums running this time from 1 to N. Put

$$f_1(x, y) = \begin{bmatrix} S(x, y) & D(x, y) \\ J(x, y) & S^+(x, y) \end{bmatrix},$$

where

$$J(x, y) = I(x, y) + \mathcal{E}(x - y).$$

This f_1 satisfies Eqs. (3.9)–(3.11) with $c = N$ and

$$\lambda = \tfrac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Moreover,

$$\{\mathrm{Tdet}[f_1(x_i, x_j)]\}^2 = \det[C(f_1(x_i, x_j))],$$

and, to evaluate $\det[C(f_1)]$, we observe that the matrix

$$A \equiv \begin{bmatrix} S(x_i, x_j) & D(x_i, x_j) \\ I(x_i, x_j) & S^+(x_i, x_j) \end{bmatrix}$$
$$= \begin{bmatrix} p_i(x_j) \\ q_i(x_j) \end{bmatrix} Z [q_i(x_j) \quad -p_i(x_j)]$$

has rank N. Hence only N rows of A are linearly independent. The N rows $[I(x_i, x_j) \; S^+(x_i, x_j)]$ being linear combinations of the first N rows, the determinant of $C(f_1)$ will not change if we subtract the last N rows of A from the corresponding ones of $C(f_1)$. Thus

$$\det[C(f_1(x_i, x_j))]$$
$$= \det[\mathcal{E}(x_i - x_j)] \cdot \det[D(x_i, x_j)].$$

The first determinant on the right-hand side is $+1$, while the second one is

$$\det[D(x_i, x_j)] = (\det M)^{-1} \prod_{1 \le i < j \le N} (x_i - x_j)^2.$$

D. The case $\beta = 1$, N odd

With $\alpha_{j,k}$ given by Eq. (4.3), the $N \times N$ matrix $[\alpha_{j,k}]$ can again be quasidiagonalized with a real orthogonal matrix U. As N is odd, $[\alpha_{j,k}]$ is singular,

$$U[\alpha]U^T = BZ + 0,$$

the diagonal B and canonical antisymmetric Z are $(N-1) \times (N-1)$ matrices. Choose

$$p_i(x) = b_i^{-1/2} \sum_{j=1}^{N} u_{ij} x^{j-1}, \quad i = 1, 2, \ldots, N-1,$$

$$p_N(x) = b_N^{-1} \sum_{j=1}^{N} u_{Nj} x^{j-1},$$

where we have introduced the number

$$b_N = \sum_{j=1}^{N} u_{Nj} \int x^{j-1}\, d\mu(x) \neq 0.$$

The polynomials $\{p_i(x)\}$, $i = 1, 2, \ldots, N-1$, are skew-orthonormal, Eq. (5.18). Also

$$\int\int p_i(x) p_N(y)\, \mathcal{E}(y-x)\, d\mu(x)\, d\mu(y) = 0,$$

and

$$\int p_i(x)\, d\mu(x) = \delta_{iN} = \begin{cases} 1, & \text{if } i = N, \\ 0, & \text{if } i < N. \end{cases}$$

Let $q_i(x)$ be given by Eq. (5.19) and S, S^*, D, and I by Eqs. (5.10)–(5.13), the sums now running from 1 to $N-1$. Put

$$f_{1\eta}(x,y) = \begin{bmatrix} S(x,y) + b_N^{-1} p_N(x) & D(x,y) + \eta p_N(x) p_N(y) \\ J(x,y) - b_N^{-1}(q_N(x) - q_N(y)) & S^*(x,y) + b_N^{-1} p_N(y) \end{bmatrix},$$

where $J(x,y) = I(x,y) + \epsilon(x-y)$, and η is a small positive number which will ultimately be taken to be zero. This $f_{1,\eta=0}(x,y)$ satisfies Eqs. (3.9) to (3.11). The $\mathrm{Tdet}[f_{1\eta}(x_i, x_j)]$ can again be calculated by subtracting from the last N rows of $C(f_{1\eta})$ the corresponding rows of

$$A_\eta = \begin{bmatrix} S(x_i,x_j) + b_N^{-1} p_N(x_i) & D(x_i,x_j) + \eta p_N(x_i) p_N(x_j) \\ I(x_i,x_j) + 1/\eta - b_N^{-1}(q_N(x_i) - q_N(x_j)) & S^*(x_i,x_j) + b_N^{-1} p_N(x_j) \end{bmatrix}.$$

Finally take the limit $\eta \to 0$.

6. AN EXAMPLE: LEGENDRE ENSEMBLES

For the particular case

$$\int d\mu(x) \to \int_{-1}^{1} dx, \tag{6.1}$$

the above procedure gives the following results. In what follows $\{P_j(x)\}$ denote Legendre polynomials[8]:

$$P_j(x) = 2^{-j} \sum_{m=0}^{[j/2]} (-1)^m \binom{j}{m} \binom{2j-2m}{j} x^{j-2m}, \tag{6.2}$$

$$\int_{-1}^{1} P_j(x) P_k(x)\, dx = (j+\tfrac{1}{2})^{-1} \delta_{jk}, \tag{6.3}$$

while $P_{-1}(x) = 0$ by convention. If $Q_{N\beta}$ in Eq. (1.1) is interpreted as a probability density while its integral $F_{N\beta}$, Eq. (5.1), as a correlation function, then the overall constants should be properly fixed. Here we will not care about this particular detail.

A. Case $\beta = 2$

$$C_{j,k} = C_{j+k}, \quad C_j = \begin{cases} 0, & \text{if } j \text{ is odd}, \\ (j+\tfrac{1}{2})^{-1}, & \text{if } j \text{ is even}; \end{cases} \tag{6.4}$$

$$p_j(x) = (j-\tfrac{1}{2})^{1/2} P_{j-1}(x), \quad j = 1, 2, \ldots, N; \tag{6.5}$$

$$f_2(x,y) = \sum_{j=0}^{N-1} (j+\tfrac{1}{2}) P_j(x) P_j(y)$$

$$= \tfrac{1}{2} N \{P_N(x) P_{N-1}(y) - P_{N-1}(x) P_N(y)\}/(x-y). \tag{6.6}$$

Let $xN = \pi\xi$, $yN = \pi\eta$. We take the limits $N \to \infty$, $x, y \to 0$, keeping ξ and η finite. Then

$$P_{2m}(\pi\xi/2m) \to (-1)^m (\pi m)^{-1/2} \cos\pi\xi,$$
$$P_{2m+1}(\pi\xi/2m) \to (-1)^m (\pi m)^{-1/2} \sin\pi\xi. \tag{6.7}$$

Hence

$$\lim_{N\to\infty} \frac{\pi}{N} f_2\left(\frac{\pi\xi}{N}, \frac{\pi\eta}{N}\right) = S(\xi-\eta) \equiv \sigma_2(r), \tag{6.8}$$

where

$$S(r) = \frac{\sin\pi r}{\pi r}, \quad r = |\xi - \eta|. \tag{6.9}$$

B. Case $\beta = 4$

$$\alpha_{j,k} = (k-j) \int_{-1}^{1} x^{j+k-1}\, dx$$
$$= \begin{cases} 0, & \text{if } j+k \text{ is even}, \\ 2(k-j)/(k+j), & \text{if } j+k \text{ is odd}; \end{cases} \tag{6.10}$$

$$p_{2j-1}(x) = (j-\tfrac{3}{4})^{1/2} \{1 + P_{2j-2}(x)\},$$

$$p_{2j}(x) = (j-\tfrac{3}{4})^{1/2} \int_0^x P_{2j-2}(y)\, dy$$

$$= \tfrac{1}{4}(j-\tfrac{3}{4})^{-1/2} \{P_{2j-1}(x) - P_{2j+1}(x)\}, \tag{6.11}$$

for $j = 1, 2, \ldots, N$;

$$S(x,y) = \sum_{j=1}^{N} \{p_{2j-1}(x) p'_{2j}(y) - p_{2j}(x) p'_{2j-1}(y)\},$$

where $p'_j(x) = (d/dx) p_j(x)$. After some algebra,

$$S(x,y) = \tfrac{1}{2} \sum_{j=0}^{2N-1} (j+\tfrac{1}{2}) P_j(x) P_j(y) + \tfrac{1}{4} P'_{2N-1}(y)$$
$$- \tfrac{1}{4} P_{2N-1}(x) P'_{2N}(y). \tag{6.12}$$

On taking the limit $N \to \infty$, while keeping $xN = \pi\xi$ and $yN = \pi\eta$ finite, the first sum on the right-hand side of the above equation gives N/π. $\frac{1}{2} S(2|\xi-\eta|)$, where $S(r)$ is given by Eq. (6.9); the other two terms in (6.12) tend to zero. The limits of $D(x,y)$ and $I(x,y)$ are

$$D(x,y) = \int_0^y S(x,z)\, dz - \tfrac{1}{4} P_{2N-1}(y) + \tfrac{1}{4} P_{2N-2}(0) \int_0^y dz\, \frac{P_{2N-1}(z)}{z}$$
$$\to \tfrac{1}{4} \int_0^{2r} S(r')\, dr',$$

$$I(x,y) = \frac{d}{dx} S(x,y)$$
$$\to \left(\frac{N}{\pi}\right)^2 \frac{1}{2} \frac{d}{dr} S(2r), \quad r = |\xi - \eta|.$$

In the limit $f_4(x,y)$ is replaced with

$$\sigma_4(\xi,\eta) = \tfrac{1}{2} \begin{bmatrix} S(2r) & D(2r) \\ \mathcal{I}(2r) & S(2r) \end{bmatrix}, \quad r = |\xi - \eta|, \tag{6.13}$$

where

$$D(r) = \frac{d}{dr} S(r), \quad \mathcal{I}(r) = \int_0^r S(r')\, dr', \tag{6.14}$$

and $S(r)$ is given by Eq. (6.9).

C. Case $\beta = 1$

In this case

$$\alpha_{j,k} = \int\int_{-1}^{1} x^j y^k \mathcal{E}(y-x)\,dx\,dy,$$

$$\alpha_{2j,2k} = \alpha_{2j+1,\,2k+1} = 0, \tag{6.15}$$

$$\alpha_{2j,2k+1} = -\alpha_{2k+1,2j} = 4\{(2j+1)(2j+2k+3)\}^{-1};$$

$$p_{2j-1}(x) = (j-\tfrac{1}{4})^{1/2}\frac{d}{dx}P_{2j-1}(x),$$

$$p_{2j}(x) = (j-\tfrac{1}{4})^{1/2}P_{2j-1}(x), \tag{6.16}$$

for $j = 1, 2, \ldots, [\frac{1}{2}N]$. After some algebra we get for N even,

$$S(x,y) = \sum_{j=0}^{N-1}(j+\tfrac{1}{2})P_j(x)P_j(y) - \tfrac{1}{2}P'_{N-1}(x)P_N(y) - \frac{N}{\pi}\mathcal{S}(r), \tag{6.17}$$

$$D(x,y) = -\frac{1}{2}\frac{d}{dy}S(x,y) + \left(\frac{N}{\pi}\right)^2 \tfrac{1}{2}\mathcal{D}(r), \tag{6.18}$$

$$I(x,y) = \int_{-1}^{1} S(z,y)\,\mathcal{E}(z-x)\,dz - 2\mathcal{I}(r), \tag{6.19}$$

where $\mathcal{S}(r)$, $\mathcal{D}(r)$, and $\mathcal{I}(r)$ are given by Eqs. (6.9) and (6.14). For N odd, there are some extra terms which tend to zero as $N \to \infty$, $x \to 0$, $y \to 0$. The $f_1(x,y)$ is replaced with

$$\sigma_1(\xi,\eta) = \begin{bmatrix} \mathcal{S}(r) & \mathcal{D}(r) \\ \mathcal{I}(r) - \frac{1}{2}\mathcal{E}(r) & \mathcal{S}(r) \end{bmatrix}, \tag{6.20}$$

$$r = |\xi - \eta|.$$

7. CONCLUSION

Since the $\sigma_\beta(\xi,\eta)$, Eqs. (6.8), (6.13), and (6.20), for the Legendre ensembles are identical to those for the Gaussian and circular ensembles, the correlation functions are also identical to all orders; hence all local properties are identical.

*On leave from C.N.R.S., France and C.E.N. de Saclay, 91190 Gif-sur-Yvette, France.

[1]H.S. Leff, J. Math. Phys. 5, 763 (1964); D. Fox and P.B. Kahn, Phys. Rev. 134, B1151 (1964); Thien Vo-Dai and J.R. Derome, Nuovo Cimento 30, B 239 (1975); M.L. Mehta, *Random Matrices* (Academic, New York, 1967).
[2]F.J. Dyson, J. Math. Phys. 3, 166 (1962).
[3]F.J. Dyson, Comm. Math. Phys. 19, 235 (1970).
[4]M.L. Mehta, Comm. Math. Phys. 20, 245 (1971); F.J. Dyson, J. Math. Phys. 13, 96 (1971).
[5]M.L. Mehta, Ref. 1, Appendix A.7.
[6]W.R. Hamilton, *Elements of Quaternions* (reprinted) (Chelsa, New York, 1965).
[7]E.H. Moore, *General Analysis I*, Memorial of the American Mathematical Society (Am. Math. Soc., Philadelphia, 1935); M.L. Mehta, J. Math. Phys. Sc. 8, 559 (1974).
[8]H. Bateman, *Higher Transcendental Functions*, edited by A. Erdelyi (McGraw-Hill, New York, 1953), Vol. 2, Chap. 10.
[9]M.L. Mehta, Elements of Matrix Theory (to be published by Hindustan Publishing Corp., Delhi).

Some basic properties of Killing spinors

Shahen Hacyan

Instituto de Astronomía, Universidad Nacional Autónoma de México, México 20, D. F., Mexico

Jerzy Plebański*

Centro de Investigaciones y de Estudios Avanzados del IPN, México 14, D. F., Mexico
(Received 21 May 1976)

The concept of Killing spinor is analyzed in a general way by using the spinorial formalism. It is shown, among other things, that higher derivatives of Killing spinors can be expressed in terms of lower order derivatives. Conformal Killing vectors are studied in some detail in the light of spinorial analysis: Classical results are formulated in terms of spinors. A theorem on Lie derivatives of Debever–Penrose vectors is proved, and it is shown that conformal motion in vacuum with zero cosmological constant must be homothetic, unless the conformal tensor vanishes or is of type N. Our results are valid for either real or complex space–time manifolds.

1. INTRODUCTION

Killing vectors and Killing tensors have proved to be useful tools in the study of Riemannian geometries. Killing vectors describe the symmetries of a V_n.[1] For instance, Debney and Kerr[2,3] have analyzed the properties of Killing vectors in the light of the null tetrad formalism, as developed by Debney, Kerr, and Schild.[4] As for Killing tensors—particularly rank two Killing tensors—they have been studied by many authors,[5-10] mainly in relation with the problem of integrating geodesic equations by separation of variables in the Hamilton–Jacobi equation.

On the other hand, since the spinorial formulation has played such an important role in general relativity,[11] it is natural to introduce the concept of a Killing spinor.[8] The aim of this work is to present, in a very general way, basic properties of Killing spinors, and then to study some particular cases which deserve special attention. Our work will make full use of the spinorial formalism.

In another context, the idea of using "complex extensions" of general relativity has attracted much attention in recent years.[12,13] This complexification takes a remarkably natural form when spinorial analysis is introduced. Recently, one of us (Plebański[14]) has shown that, in a complex V_4, spinors can be defined which transform according to the $SL(2,\mathbb{C})\times\overline{SL}(2,\mathbb{C})$ group. The implication of this is that objects with dotted spinorial indices are no longer related to those with undotted indices, as it would be the case in a real V_4. For instance, the spinorial images of the complex Weyl tensor are C_{ABCD} and $C_{\dot A\dot B\dot C\dot D}$, but, in general, $C_{\dot A\dot B\dot C\dot D}\neq(C_{ABCD})^*$ if the V_4 is complex.

Thus, loosely speaking, the general procedure for complexifying a real V_4 is simply to "forget" (!) that dotted spinors are related to undotted ones by a complex conjugation. The results we present in this work take this fact into account: They are valid either in a real or in a complex V_4. This will be mentioned explicitly whenever necessary. We hope that these results will be useful in further studies of complex and real V_4's.

Consider a symmetric spinor $K^{\dot B_1\cdots\dot B_{2q}}_{A_1\cdots A_{2p}}=K^{(\dot B_1\cdots\dot B_{2q})}_{(A_1\cdots A_{2p})}$ which transforms as the irreducible $D(p,q)$ representation of the Lorentz group, and which satisfies the equation

$$\nabla^{(\dot B}_{(A}K^{\dot B_1\cdots\dot B_{2q})}_{A_1\cdots A_{2p})}=0 \tag{1.1}$$

in a Riemannian V_4. Such a spinor is a Killing spinor, Eq. (1.1) being the spinorial analog of the equation for a conformal Killing tensor $K_{b_1\cdots b_m}=K_{(b_1\cdots b_m)}$,

$$\nabla_{(a}K_{b_1\cdots b_m)}=g_{(ab_1}\chi_{b_2\cdots b_m)}. \tag{1.2}$$

In flat space, Eq. (1.1) is the twistor equation (Penrose[15]). In a curved space it will give some very strong conditions on the type of spacetime. For instance, Sommers[10] has shown that any vacuum metric which admits a nontrivial $D(p,0)$ Killing spinor must be flat, plane fronted gravitational wave or type D.

We shall investigate further properties of Killing spinors. We will be particularly interested in the important case of a $D(\frac{1}{2},\frac{1}{2})$ Killing spinor, which satisfies the equation

$$\nabla^{(\dot A}_{(A}K^{\dot B)}_{B)}=0. \tag{1.3}$$

This equation is the "trace-free" part of the conformal Killing equation,

$$\nabla_{(a}K_{b)}=\chi g_{ab}. \tag{1.4}$$

Since conformal invariance seems to play an important role in physics (see, e.g., Penrose,[15] Plebański[16]), we will consider conformal Killing vectors ($\chi=0$), normal Killing vectors ($\chi=0$) being obvious subcases.

2. GENERAL CONSIDERATIONS

It can be shown (Sommers[10]) that the $(m+1)$th derivative of a rank m Killing tensor can always be expressed in terms of its lower order derivatives. We will prove a similar theorem for Killing spinors (KS). Following Sommers,[10] it is enough to notice that the rth covariant derivative of any tensor or spinor can be expressed as

$$\nabla_{a_1}\cdots\nabla_{a_r}K_{b_1\cdots b_m}=\nabla_{(a_1}\cdots\nabla_{a_r)}K_{b_1\cdots b_m}+[\text{terms containing } (r-2)\text{th or lower order derivatives of } K_{b_1\cdots b_m}]. \tag{2.1}$$

This is due to the commutation relations between covariant derivatives. Now,

$$N_r := (r+1)[(r+1) + r(r-1)/6] \tag{2.2}$$

is the number of algebraically independent components of a fully symmetric rank r tensor in V_4. Thus the rth derivative of an irreducible $D(p, q)$ spinor is given by $N_r(2p+1)(2q+1)$ linearly independent terms containing rth derivatives, plus terms with $(r-2)$th or lower order derivatives. On the other hand, Eq. (1.1) gives $N_{r-1}(2p+2)(2q+2)$ linearly independent relations among the rth order derivatives, since

$$\nabla_{(a_1}\cdots\nabla_{a_{r-1})}\nabla^{(\dot{B}}_{(A}K^{\dot{B}_1\cdots\dot{B}_q)}_{A_1\cdots A_{2p})} = 0. \tag{2.3}$$

Thus, if

$$\frac{N_{r-1}}{N_r} \geqslant \frac{(2p+1)(2q+1)}{(2p+2)(2q+2)} \tag{2.4}$$

one can obtain the rth derivative of the KS in terms of lower order derivatives[17] [plus, in general, some integrability conditions; i.e., Eqs. (3.2) or (3.16) below].

We will now study the integrability condition of Eq. (1.1). The commutation relations for covariant derivatives of (in particular) symmetric spinors is given by[18]:

$$dx^\mu \wedge dx^\nu \nabla_\mu \nabla_\nu \Psi^{\dot{B}_1\cdots\dot{B}_{2q}}_{A_1\cdots A_{2p}} = 2q\Psi^{\dot{N}}_{A_1\cdots A_{2p}}{}^{(\dot{B}_1\cdots\dot{B}_{2q-1}}R^{\dot{B}_{2q})}{}_{\dot{N}}$$
$$- 2p\Psi^{\dot{B}_1\cdots\dot{B}_{2q}}_{\dot{N}(A_1\cdots A_{2p-1}}R^{N}{}_{A_{2p})}, \tag{2.5}$$

where $R^A{}_B$ is the two-form associated to the Riemann tensor. According to the Cartan structure equations written in spinorial notation, one has

$$R^A{}_B = d\Gamma^A{}_B + \Gamma^A{}_N \wedge \Gamma^N{}_B, \tag{2.6}$$

where $\Gamma^A{}_B$ are one-forms associated to the Ricci rotation coefficients (we will use exactly the same notation and conventions as in Ref. 14). More explicitly, Eq. (2.5) gives

$$\nabla_{(C}{}^{\dot{N}}\nabla_{D)\dot{N}}\Psi^{\dot{B}_1\cdots\dot{B}_{2q}}_{A_1\cdots A_{2p}} = 4qC_{CD\dot{N}}{}^{(\dot{B}_1}\Psi^{\dot{B}_2\cdots\dot{B}_{2q})\dot{N}}_{A_1\cdots A_{2p}}$$
$$+ 4pC^N{}_{CD(A_1}\Psi^{\dot{B}_1\cdots\dot{B}_{2q}}_{A_2\cdots A_{2p})N}$$
$$+ p\frac{R}{3}\delta^N_{(C}\epsilon_{D)(A_1}\Psi^{\dot{B}_1\cdots\dot{B}_{2q}}_{A_2\cdots A_{2p})N}, \tag{2.7a}$$

$$\nabla^{N(\dot{C}}\nabla_N{}^{\dot{D})}\Psi^{\dot{B}_1\cdots\dot{B}_{2q}}_{A_1\cdots A_{2p}} = -4qC^{\dot{C}\dot{D}(\dot{B}_1}{}_{\dot{N}}\Psi^{\dot{B}_2\cdots\dot{B}_{2q})\dot{N}}_{A_1\cdots A_{2p}}$$
$$- 4pC^N{}_{(A_1}{}^{\dot{C}\dot{D}}\Psi^{\dot{B}_1\cdots\dot{B}_{2q}}_{A_2\cdots A_{2p})N}$$
$$+ q\frac{R}{3}\delta^{(\dot{C}}_{\dot{N}}\epsilon^{\dot{D})(\dot{B}_1}\Psi^{\dot{B}_2\cdots\dot{B}_{2q})\dot{N}}_{A_1\cdots A_{2p}}, \tag{2.7b}$$

where C_{ABCD} and $C_{\dot{A}\dot{B}\dot{C}\dot{D}}$ are the spinorial images of the Weyl tensor, $C_{AB\dot{C}\dot{D}}$ is the spinorial image of the traceless Ricci tensor, and R is the trace of the Ricci tensor. If the considered V_4 is *real*, then $C_{\dot{A}\dot{B}\dot{C}\dot{D}} = (C_{ABCD})^*$, $C_{AB\dot{C}\dot{D}} = (C_{CD\dot{A}\dot{B}})^*$ and $R = R^*$, otherwise these last relations do not hold (for a detailed discussion of this point, see Ref. 14).

Now, the fundamental equation (1.1) implies that the covariant derivative of a KS has the irreducible form

$$\nabla_A{}^{\dot{B}}K^{\dot{B}_1\cdots\dot{B}_{2q}}_{A_1\cdots A_{2p}} = \epsilon_{A(A_1}l^{\dot{B}\dot{B}_1\cdots\dot{B}_{2q}}_{A_2\cdots A_{2p})} + \epsilon^{\dot{B}(\dot{B}_1}\bar{l}^{\dot{B}_2\cdots\dot{B}_{2q})}_{AA_1\cdots A_{2p}}$$
$$+ \epsilon_{A(A_1}\epsilon^{\dot{B}(\dot{B}_1}j^{\dot{B}_2\cdots\dot{B}_{2q})}_{A_2\cdots A_{2p})} \tag{2.8}$$

where[19]:

$$l^{\dot{B}\dot{B}_1\cdots\dot{B}_{2q}}_{A_1\cdots A_{2p-1}} = \frac{2p}{2p+1}\nabla^{N(\dot{B}}K^{\dot{B}_1\cdots\dot{B}_{2q})}_{A_1\cdots A_{2p-1}N}, \tag{2.9a}$$

$$\bar{l}^{\dot{B}_1\cdots\dot{B}_{2q-1}}_{A_1A_1\cdots A_{2p}} = -\frac{2q}{2q+1}\nabla_{(A|\dot{N}|}K^{\dot{B}_1\cdots\dot{B}_{2q-1}\dot{N}}_{A_1\cdots A_{2p})}, \tag{2.9b}$$

$$j^{\dot{B}_1\cdots\dot{B}_{2q-1}}_{A_1\cdots A_{2p-1}} = -\frac{2q}{2p+1}\frac{2q}{2q+1}\nabla^N{}_{\dot{N}}K^{\dot{B}_1\cdots\dot{B}_{2q-1}\dot{N}}_{A_1\cdots A_{2p-1}N}. \tag{2.9c}$$

3. SPECIAL CASES

Consider first a $D(p, 0)$ KS. According to (2.8) one has

$$\nabla_A{}^{\dot{B}}K_{A_1\cdots A_{2p}} = \epsilon_{A(A_1}l^{\dot{B}\dot{B}_1\cdots\dot{B}_{2q}}_{A_2\cdots A_{2p})}. \tag{3.1}$$

Substituting this in the formula (2.7a), and symmetrizing over indices $CDA_1\cdots A_{2p}$, one obtains

$$K_{N(B_1\cdots B_{2p-1}}C^N{}_{B_{2p}CD)} = 0. \tag{3.2}$$

Substitution of (3.1) in (2.7b) gives, after proper symmetrization of indices,

$$\nabla^{(\dot{P}}_{(B_1}l^{\dot{Q})}_{B_2\cdots B_{2q})} = 4pK_{N(B_1\cdots B_{2q-1}}C_{B_{2p})}{}^{N\dot{P}\dot{Q}} \tag{3.3}$$

Thus, in empty space, $l^{\dot{Q}}_{B_1\cdots B_{2p-1}}$ itself is a KS of type $D(p-\frac{1}{2}, \frac{1}{2})$. In a complex V_4, Eq. (3.2) has a counterpart

$$K_{\dot{N}(\dot{B}_1\cdots\dot{B}_{2p-1}}C^{\dot{N}}_{\dot{B}_{2p}\dot{C}\dot{D})} = 0 \tag{3.4}$$

which, for a real V_4, would be just the complex conjugate of (3.3). A similar result applies for relation (3.3). It is interesting that, in a complex V_4, this KS's $K_{B_1\cdots B_{2p}}$ and $K_{\dot{B}_1\cdots\dot{B}_{2p}}$ are independent objects: that is, one has a "heavenly" and a "hellish" KS (in the terminology of Ref. 14). In fact, a spacetime can admit a "heavenly" KS without necessarily admitting a "hellish" KS.

Consider, for instance, a $D(\frac{1}{2}, 0)$ KS which satisfies the equation

$$\nabla^A_{(A}K_{B)} = 0 \tag{3.5}$$

and consequently

$$\nabla_A{}^{\dot{A}}K_B = \epsilon_{AB}l^{\dot{B}}. \tag{3.6}$$

The integrability conditions give, after some algebraic manipulations, the following relations:

$$K_N C^N{}_{ABC} = 0, \tag{3.7a}$$

$$\nabla_A{}^{\dot{A}}l^{\dot{B}} = 2C^N{}_A{}^{\dot{A}\dot{B}}K_N + \frac{R}{12}\epsilon^{\dot{A}\dot{B}}K_A. \tag{3.7b}$$

Equation (3.7a) implies that the spacetime is of type $[N]\otimes$ [something] or $[-]\otimes$ [something], and (3.7b) implies that in vacuum, l^A itself is a Killing spinor. Equation (3.7b) permits one to express the second derivative of the KS in terms of the KS itself.

Consider, now, KS's of type $D(1, 0)$. One has

$$\nabla^{\dot{A}}_{(A}K_{BC)} = 0 \tag{3.8}$$

from which

$$\nabla_A{}^{\dot{A}} K_{BC} = \epsilon_{A(B} l_{C)}{}^{\dot{A}}. \tag{3.9}$$

The integrability conditions give

$$C^N{}_{(ABC} K_{D)N} = 0, \tag{3.10a}$$

$$\nabla_A{}^{\dot{A}} l_B{}^{\dot{B}} = 4C^N{}_{(A\dot{A}\dot{B}} K_{B)N} + \epsilon_{AB}\left(C^{\dot{N}\dot{S}\dot{A}\dot{B}} K_{NS} + \frac{R}{6} K^{\dot{A}\dot{B}}\right) + \epsilon^{\dot{A}\dot{B}}\left(C^{NS}{}_{AB} K_{NS} + \frac{R}{6} K_{AB}\right). \tag{3.10b}$$

Thus, in vacuum, $l_A{}^{\dot{A}}$ is a $D(\frac{1}{2}, \frac{1}{2})$ KS. Furthermore, (3.10b) permits one to calculate all the second derivatives of K_{AB} in terms of K_{AB} itself. For further properties of type $D(\frac{1}{2}, 0)$ and $D(1, 0)$ KSs in a real V_4 see Walker and Penrose,[8] and Sommers.[10]

We now consider the important case of $D(\frac{1}{2}, \frac{1}{2})$ Killing spinors, which are related to conformal Killing vectors (CKV). From (2.8) one has

$$\nabla_{(A}{}^{(\dot{A}} K_{B)}{}^{\dot{B})} = 0, \tag{3.11}$$

$$\nabla_A{}^{\dot{A}} K_B{}^{\dot{B}} = \epsilon_{AB} l^{\dot{A}\dot{B}} + \epsilon^{\dot{A}\dot{B}} l_{AB} - 2\epsilon_{AB}\epsilon^{\dot{A}\dot{B}}\chi. \tag{3.12}$$

Notice that $l^{\dot{A}\dot{B}} = (l^{AB})^*$ and that χ is the conformal factor appearing in (1.4). Substituting (3.12) in the integrability conditions (2.7a) and (2.7b), after proper symmetrizations and contractions of indices and using the relation

$$\nabla_A{}^{\dot{A}} l_{BC} = \nabla_{(A}{}^{\dot{A}} l_{BC)} + \tfrac{2}{3}\epsilon_{A(B}\nabla^{N\dot{A}} l_{C)N} \tag{3.13}$$

one obtains, after some algebraic manipulations,

$$\nabla_A{}^{\dot{A}} l_{BC} = -2C^N{}_{ABC} K_N^{\dot{A}} - 2K^{\dot{N}}_{(A} C_{BC)\dot{N}}{}^{\dot{A}} - \tfrac{4}{3}\epsilon_{A(B} C_{C)}{}^{N\dot{N}\dot{A}} K_{N\dot{N}} + \frac{R}{6}\epsilon_{A(B} K_{C)}^{\dot{A}} - 2\epsilon_{A(B}\nabla_{C)}{}^{\dot{A}}\chi. \tag{3.14}$$

If $\chi = 0$, one would obtain an equation expressing second derivatives of the Killing vector in terms of the Killing vector itself (a well known result![1]). The next step is to use the integrability conditions of (3.14) by substituting it again in (2.7a) and (2.7b). Using the Bianchi identities in spinorial form,

$$\nabla^{\dot{S}}_{\dot{D}} C_{SABC} + \nabla_{(A}{}^{\dot{S}} C_{BC)\dot{S}\dot{D}} = 0, \tag{3.15a}$$

$$\nabla^{R\dot{S}} C_{AR\dot{B}\dot{S}} + \tfrac{1}{8}\nabla_{A\dot{B}} R = 0, \tag{3.15b}$$

one obtains, after some algebraic acrobatics, the following equations:

$$4C^N{}_{(ABC} l_{D)N} - 4C_{ABCD}\chi + K_{N\dot{N}}\nabla^{N\dot{N}} C_{ABCD} = 0, \tag{3.16}$$

$$\nabla_{(A}{}^{(\dot{A}}\nabla_{B)}{}^{\dot{B})}\chi = K^S_S \nabla_{(S}{}^{(\dot{S}} C_{AB)}{}^{\dot{A}\dot{B})} - \tfrac{2}{3}K^{S(\dot{A}}\nabla_{(A|\dot{N}|} C_{BS)}{}^{\dot{B})\dot{N}} - \tfrac{2}{3}K_{\dot{S}(A}\nabla^{N(\dot{A}} C_{B)N}{}^{\dot{B}\dot{S})} + \tfrac{1}{18}K^{(\dot{A}}_{(A}\nabla_{B)}{}^{\dot{B})} R - 2C^N{}_{(A}{}^{\dot{A}\dot{B}} l_{B)N} + 2C_{AB\dot{S}}{}^{(\dot{A}} l^{\dot{B})\dot{S}} + 4\chi C_{AB}{}^{\dot{A}\dot{B}} \tag{3.17a}$$

$$\nabla_{N\dot{N}}\nabla^{N\dot{N}}\chi = -\tfrac{1}{6}K^{\dot{N}}_N \nabla^N{}_{\dot{N}} R - \tfrac{2}{3}\chi R. \tag{3.17b}$$

These last two equations permits one to calculate the second derivatives of χ in terms of the Killing spinor and χ only, by using the formula

$$\nabla_A{}^{\dot{A}}\nabla_B{}^{\dot{B}}\chi = \nabla_{(A}{}^{(\dot{A}}\nabla_{B)}{}^{\dot{B})}\chi + \tfrac{1}{4}\epsilon_{AB}\epsilon^{\dot{A}\dot{B}}\nabla^{N\dot{N}}\nabla_{N\dot{N}}\chi. \tag{3.18}$$

Summing up: for any CKV, the second and higher order derivatives of K_a and χ can always be expressed as linear combinations of K_a, $\nabla_a K_b$, χ, and $\nabla_a\chi$. In other words, a CKV can be described by a set of 15 parameters (10 if $\chi = 0$). This is a well known result which can be obtained by classical tensorial methods.[1] Furthermore, this result is in agreement with inequality (2.4) obtained above: Third order and higher derivatives of $K^{\dot{A}}{}_A$ can be expressed in terms of lower derivatives (remember that $\chi = -\frac{1}{8}\nabla^{N\dot{N}} K_{N\dot{N}}$). As for the relation (3.16), it is the spinorial form of the equation for the Lie derivative of the Weyl tensor along a CKV. Equations for the Lie derivatives of the Riemann tensor and its covariant derivatives entirely determine the number of possible Killing vectors[20,21] (this result is also valid in complex Riemannian spaces, as a simple examination shows).

We will now examine the concept of Lie derivative in the light of spinorial analysis. By definition, the Lie derivative of a tensor $T_{\alpha_1\cdots\alpha_p}$ along a vector K_α is

$$\underset{K}{\pounds} T_{\alpha_1\cdots\alpha_p} = T_{\mu\alpha_2\cdots\alpha_p} K^\mu{}_{;\alpha_1} + \cdots + T_{\alpha_1\cdots\alpha_{p-1}\mu} K^\mu{}_{;\alpha_p} + T_{\alpha_1\cdots\alpha_p;\mu} K^\mu. \tag{3.19}$$

In particular, if K^α is a CKV, then

$$\underset{K}{\pounds} g_{\alpha\beta} = 2\chi g_{\alpha\beta}. \tag{3.20}$$

When (3.19) is applied to the vectors of a tetrad, one obtains

$$\underset{K}{\pounds} e^a_\alpha = \Lambda^a{}_n e^n_\alpha, \tag{3.21}$$

where

$$\Lambda^a{}_b = K^a{}_{;b} - K^n \Gamma^a{}_{bn}. \tag{3.22}$$

This follows easily from the definition of the Ricci rotation coefficients

$$\Gamma^a{}_{bc} := -e^a{}_{\alpha;\beta} e^\alpha{}_b e^\beta{}_c. \tag{3.23}$$

The concept of Lie derivative along a CKV can be introduced in a straightforward manner (e.g., Geroch[22]). From the definition of the Pauli matrices and the null tetrad[14]

$$g^{A\dot{A}} := \sqrt{2}\begin{pmatrix} e^4 & e^2 \\ e^1 & -e^3 \end{pmatrix}, \tag{3.24}$$

one can find a formula similar to (3.21),

$$dx^\alpha \underset{K}{\pounds} g_\alpha^{A\dot{A}} = \Lambda^A{}_N g^{N\dot{A}} + \Lambda^{\dot{A}}{}_{\dot{N}} g^{A\dot{N}} + \chi g^{A\dot{A}}, \tag{3.25}$$

where

$$\Lambda^A{}_B = -\tfrac{1}{2} l^A{}_B - \Gamma^A{}_{Bn} K^n. \tag{3.26}$$

From this it can be seen that Eq. (3.16) can be rewritten as

$$4\Lambda^N{}_{(A} C_{BCD)N} - 2\chi C_{ABCD} + K^\mu \partial_\mu C_{ABCD} = 0. \tag{3.27}$$

We can now prove the following theorem.

Theorem: The Lie derivative along a CKV of a Debever–Penrose[11] vector is proportional to it.

That is, if $e^3{}_\mu$ is a Debever–Penrose vector and Eq. (3.20) is satisfied, then

$$\underset{K}{\pounds} e^3_\alpha = f e^3_\alpha \quad (f \text{ is some function}). \tag{3.28}$$

The proof is very simple: If $e^3{}_\mu$ is a single DP vector, then $C_{1111}=0$ and $C_{1112}\neq 0$, and (3.27), for indices $ABCD\to 1111$, implies $\Lambda_{11}=0$. If $e^3{}_\alpha$ is a double DP vector, then $C_{1111}=C_{1112}=0$ and $C_{1122}\neq 0$; Eq. (3.27), for indices $ABCD\to 1112$, again implies $\Lambda_{11}=0$, and so on for triple and quadruple DP vectors: one always finds that $\Lambda_{11}=0$. But, from Eq. (3.25), this is precisely the condition for $e^3{}_\alpha$ to have its Lie derivative parallel to itself. This completes the proof of the theorem.[23] A simple reexamination of the proof given above shows that it is also valid in complex V_4. All one has to do is to use the dotted counterpart of Eq. (3.27) and the generalization of the concept of a Debever–Penrose vector to complex V_4.[14] The theorem has an obvious geometrical interpretation: Since conformal curvature is invariant along conformal Killing vector fields, the principal null directions it defines must also be preserved.

We will now consider the special case of a CKV in a vacuum spacetime, that is $C_{AB\dot A\dot B}=0$ and $R=-4\lambda$ (cosmological constant). The equations obtained above take the simpler form,

$$\nabla_A{}^{\dot A} l_{BC} = -2C^N{}_{ABC}K^{\dot A}_N - \tfrac{2}{3}\lambda\epsilon_{A(B}K^{\dot A}_{C)} - 2\epsilon_{A(B}\nabla_{C)}{}^{\dot A}\chi, \quad (3.29)$$

$$\nabla_A{}^{\dot A}\nabla_B{}^{\dot B}\chi = \tfrac{2}{3}\epsilon_{AB}\epsilon^{\dot A\dot B}\lambda\chi, \quad (3.30)$$

$$4C^N{}_{(ABC}l_{D)N} - 4\chi C_{ABCD} + K_{N\dot N}\nabla^{N\dot N}C_{ABCD} = 0. \quad (3.31)$$

In particular, the integrability condition of (3.30) implies

$$C^N{}_{ABC}\nabla^{\dot A}{}_N\chi = 0. \quad (3.32)$$

Thus, if $\nabla^{\dot A}{}_A\chi\neq 0$, one has two possibilities:

(i) $\det(\nabla_N{}^{\dot A}\chi)\neq 0$, in which case $C_{ABCD}=0$,

(ii) $\det(\nabla^{\dot A}{}_N\chi)=0$, then $\nabla^{\dot A}{}_N\chi=\phi\alpha_N\alpha^{\dot A}$ and, according to (3.32), α_N is a quadruple Penrose spinor and the spacetime is of type N. Thus, the existence of a nonconstant conformal factor χ in vacuum implies that the spacetime is Minkowskian, de Sitter or type N.[24] Moreover, if $\lambda\neq 0$ then χ is necessarily zero except for the de Sitter and type N cases. Notice that χ satisfies an eikonal equation if $C_{ABCD}\neq 0$.

Next, we consider the case when a Maxwell field is present. The Einstein–Maxwell equations without sources are

$$C_{AB\dot A\dot B} = -8f_{AB}f_{\dot A\dot B}, \quad \nabla^{N\dot A}f_{NB}=0, \quad R=-4\lambda, \quad (3.33)$$

where f_{AB} is the $D(1,0)$ spinor associated to the electromagnetic field. In this particular case, Eqs. (3.17a) and (3.17b) somewhat simplify,

$$\frac{1}{8\pi}\nabla_{(A}{}^{(\dot A}\nabla_{B)}{}^{\dot B)}\chi = K^S_S\nabla^{\dot S}{}_S[f_{AB}f^{\dot A\dot B}] - 2f^{\dot A\dot B}f^{\dot N}{}_{(A}l_{B)N} + 2f_{AB}f_{\dot S}{}^{(\dot A}l^{\dot B)\dot S} + 4\chi f_{AB}f^{\dot A\dot B}, \quad (3.34a)$$

$$\nabla_{N\dot N}\nabla^{N\dot N}\chi = \tfrac{8}{3}\chi\lambda. \quad (3.34b)$$

The formulas given above are the spinorial versions of some fundamental classical results on Killing vectors (see, e.g., Einsenhart[1,20] and Petrov[21]). As for Killing vectors in spaces where an electromagnetic field is present, they have been of fundamental importance in the works of Ernst[25] and Kinnersley,[26] who found a general procedure for generating new solutions. We hope that our Eqs. (3.34) will be useful in further investigations of this point.

We note that this paper is to be considered as the fourth of the series of articles which study the "analytic continuation" of general relativity (the first is Ref. 14; then follow Refs. 27 and 28) and will be followed by subsequent publications, where the present rather formal results will be applied as useful technical tools in the study of the dynamics of complex relativity.

*On leave of absence from University of Warsaw, Warsaw, Poland.

[1] L. P. Eisenhart, *Riemannian Geometry* (Princeton U. P., Princeton, N.J., 1964).
[2] R. P. Kerr and G. C. Debney, Jr., J. Math. Phys. **11**, 2807 (1970).
[3] G. C. Debney, Jr., J. Math. Phys. **12**, 1088 (1971).
[4] G. C. Debney, Jr., R. P. Kerr, and A. Schild, J. Math. Phys. **10**, 1842 (1969).
[5] B. Carter, Phys. Rev. **174**, 1559 (1968).
[6] P. Havas, J. Math. Phys. **16**, 2476 (1975).
[7] C. P. Boyer, E. G. Kalnins, and W. Miller, Jr., Nagoya Math. J. **60**, 35 (1976).
[8] M. Walker and R. Penrose, Commun. Math. Phys. **18**, 265 (1970).
[9] I. Hauser and R. J. Malhiot, J. Math. Phys. **15**, 816 (1974); **16**, 150 (1975).
[10] P. Sommers, Ph.D. thesis, The University of Texas at Austin, (1973).
[11] R. Penrose, Ann. Phys. (N.Y.) 10, 171 (1960).
[12] A. Trautman, Proc. R. Soc. **A270**, 326 (1962) (a private communication of I. Robinson is acknowledged).
[13] E. T. Newman in Tel Aviv GR7 Lecture (1974), Bergmann's Symposium, Syracuse (1975), and Math. Phys. Symposium, Mexico (1976); R. Penrose, "The Nonlinear Graviton," First Award Gravity Foundation Essay (1975).
[14] J. Plebański, J. Math. Phys. **16**, 2395 (1975).
[15] R. Penrose, in *Quantum Gravity. An Oxford Symposium*, edited by C. J. Isham, R. Penrose, and D. W. Sciama (Clarendon, Oxford, 1975).
[16] J. Plebański, *Conformal Equivalent Riemannian Spaces* (Centro de Investigaciones y Estudios Avanzados del IPN, México, 1965).
[17] A basic property of Killing vectors is that their second or higher derivatives can be expressed in terms of the Killing vector and its first derivatives. Since a generalization of this property applies for KSs, the name of "Killing spinors" seems to be well justified, and we will use it without quotation marks.
[18] J. Plebański, *Spinors, Tetrads and Forms* (Centro de Investigaciones y Estudios Avanzados del IPN, México, 1975). Our convention for normalizing Pauli matrices is $g^{N\dot N}_a g^b_{N\dot N} = -2\delta^b_a$; thus, for instance, $K^{A\dot A} = g^{A\dot A}_n K^n$ and $K^n = -\frac{1}{2}g^a_{N\dot N}K^{N\dot N}$. Also, in our convention: $\epsilon_{AB}\epsilon^{AB}=2$, $\tau^A=\tau_N\epsilon^{NA}$, $\tau_A=\epsilon_{AN}\tau^N$. This follows the notation of Ref. 14.
[19] A symbol like $(AB\,|\,C\,|\,D)$, for instance, means that symmetrization is over all indices except C (this with obvious gengeneralizations).
[20] L. P. Eisenhart, *Continuous Groups of Transformation* (Princeton U. P., Princeton, N.J., 1933).
[21] A. Z. Petrov, *Einstein Spaces* (Pergamon, Oxford, 1969).
[22] Mentioned by Sommers.[10]
[23] Petrov gives a proof of this theorem in Chapter 4 of his book[21] for the particular case of a vacuum type III metric.
[24] If the space is complex, one has type $[-]\otimes[-]$, $[-]\otimes N$, and $N\otimes N$ in the notation of Ref. 14.
[25] F. J. Ernst, Phys. Rev. **168**, 1415 (1968); J. Math. Phys. **15**, 1409 (1974).
[26] W. Kinnerley, J. Math. Phys. **14**, 651 (1973).
[27] J. F. Plebański and S. Hacyan, J. Math. Phys. **16**, 2403 (1975).
[28] J. D. Finley, III and J. F. Plebański, J. Math. Phys. **17**, 585 (1976).

The intrinsic spinorial structure of hyperheavens

J. D. Finley, III*
University of New Mexico, Albuquerque, New Mexico 87131

J. F. Plebański†
Centro de Investigación y de Estudios Avanzados del I. P. N., AP 14-740, México 14, D. F., Mexico
(Received 30 June 1976)

Following Plebański and Robinson, complex V_4's which admit a congruence of totally null surfaces are shown to have coordinates which, in pairs, have a spinor structure which generates the usual spinor structure of the 2-forms over the space. This structure allows Einstein's vacuum equations to fracture into three triples and a singlet, which allow for easy reduction of the entire set to one nonlinear partial differential equation needed for consistency. An inhomogeneous $\widetilde{GL}(2,C)$ group of coordinate transformations, constrained to leave the tetrad form invariant, is constructed and used to simplify the equations and clarify the geometrical meaning of the parameters introduced during the integration process.

1. INTRODUCTION

The analytic continuation of general relativity, with emphasis on complex Einstein-flat four-dimensional spaces endowed with purely self-dual conformal curvature (heavens, according to the Newman–Penrose terminology[1]), was studied in a previous series of papers.[2–6] The present work is the second of a new series of studies, being an extension along lines initiated by Plebański and Robinson.[7] They found that if the conformal curvature is algebraically degenerate from (at least) one side—that is, e.g., the self-dual part of the conformal curvature tensor is algebraically degenerate—then all quantities on the manifold can be written in terms of a canonical tetrad such that Einstein's equations in vacuum may be integrated up to the very end, reducing the problem of such a (complex, minimally degenerate) gravitational field to the solution of one partial differential equation of second order, involving only quadratic nonlinearities. (This equation contains as a very special case the heavenly equation of Ref. 2, and is therefore called the hyperheavenly equation.) It is to be noted that the existence of this tetrad is already guaranteed by the totally null surface which Ref. 3 guarantees us that such a space must have. [These important null (extremal) 2-surfaces have also been called twistor surfaces by Flaherty.[8]]

We use the notation and techniques of this series of papers, particularly Ref. 2. (For real tetradial formalism one could also see Ref. 9.) The spinorial ideas used relative to the structure of the coordinates in these spaces were first developed for heavenly spaces in Ref. 6. In Sec. 2 we write the metrics of both cases of Ref. 7 in a uniform spinorially-based notation, which is chosen to emphasize as much as possible the existence of the congruence of totally null surfaces, thus making it easier to see how the "crystal structure" of Einstein's vacuum equations may be split apart. We then determine the general form of the connections and the Riemann tensor, and define a group of automorphisms of the space which preserve the forms of the tetrad, as natural changes in the (arbitrary) parametrization of the various null surfaces in the congruence. We also show the natural relationship of the intrinsic spinor structure of the coordinates to the usual spinor spaces defined over the manifold.

In Sec. 3 we show how the complete set of Einstein vacuum field equations "fractures" into three triples corresponding to helicity[10] +1, 0, and −1, and a singlet equation, which, when solved in the proper order are very amenable. Then, having reduced the problem to the solution of the hyperheavenly equation for W, a function of all the coordinates, we may recast the components of the conformal curvature into their (simple and plausible) form as functions of W and the integration parameters which appear.

In Sec. 4 we show how one easily makes a choice of conformal factor to get to the case I of Ref. 7. At each step in both Secs. 3 and 4, we show how one may take advantage of the particular group of automorphisms so as to both simplify the resultant equations and emphasize the geometrical content of the parametrization. In appendices we give some of the more complicated details of the calculations used to reduce Einstein's equations, as well as to show how the results should be (trivially) modified if one wants to include a nonzero cosmological constant.[11]

2. INTRINSIC SPINOR FORMALISM

We start off by insisting only that our (complex) spacetime admit a congruence of totally null surfaces. These surfaces can be described by the (closed) 2-form

$$\Sigma \equiv du \wedge dv, \tag{2.1}$$

where u and v are (functionally) independent quantities constant on each surface, and the expansion 1-form

$$\theta \equiv \theta_a dx^a \equiv \tfrac{1}{2}(u^{ia}{}_{;a}\,dv - v^{ia}{}_{;a}\,du). \tag{2.2}$$

Since Σ is totally null, it is clear that du and dv are null and orthogonal. As was shown in Ref. 7, the existence of this congruence is sufficient to allow us to choose a (null) tetrad:

$$\begin{aligned} &e^1 = \phi^{-2}\,du, \quad e^2 = dx + P\,du + R\,dv, \\ &e^3 = \phi^{-2}\,dv, \quad e^4 = dy + R\,du + Q\,dv, \end{aligned} \tag{2.3a}$$

$$\longleftrightarrow$$

$$\begin{aligned} &\partial_1 = \phi^2(\partial_u - P\partial_x - R\partial_y), \quad \partial_2 = \partial_x \\ &\partial_3 = \phi^2(\partial_v - R\partial_x - Q\partial_y), \quad \partial_4 = \partial_y, \end{aligned} \tag{2.3b}$$

where we use coordinates especially adapted to our

congruence, u, v being constant on each surface, along with some independent variables x, y, which are longitudinal variables along a given surface, while ϕ, P, Q, and R are sufficiently smooth functions of these coordinates, and the metric is given by

$$g = 2e^1 \underset{s}{\otimes} e^2 + 2e^3 \underset{s}{\otimes} e^4 \in \Lambda^1 \otimes \Lambda^1, \tag{2.4}$$

with $\underset{s}{\otimes}$ as the symmetrized tensor product: $A \underset{s}{\otimes} B \equiv \frac{1}{2}(A \otimes B + B \otimes A)$.

Since we have started out with a real V_4 and complexified it, we may always look upon it as some sort of formal direct sum of two two-dimensional complex spaces in a fairly simple way.[12] However, we intend to show that, because of the existence of a totally null congruence in this space, this correspondence may be thought of as indicating important underlying structures of such spaces. To this end, then, we proceed to look upon the coordinates, in pairs, as forming spinors:

$$q_{\dot{A}} \equiv \begin{pmatrix} v \\ u \end{pmatrix}, \quad p^{\dot{B}} \equiv \begin{pmatrix} -y \\ -x \end{pmatrix}, \tag{2.5}$$

where, being spinors, the indices of $q_{\dot{A}}$ and $p^{\dot{B}}$ are manipulated in the usual way:

$$q^{\dot{A}} \equiv \epsilon^{\dot{B}\dot{A}} q_{\dot{B}}, \quad q_{\dot{A}} \equiv \epsilon_{\dot{A}\dot{B}} q^{\dot{B}}, \tag{2.6}$$

and similarly for $p^{\dot{B}}$, as well as being subject to transformations from spinor-type groups such as SL(2, C). We will say much more, later, about the transformation properties of these coordinate spinors. It should also be pointed out that the spinor indices are dotted because of the fact that our tetrad puts us in a space where it is the left-conformal curvature tensor (the *self-dual* part) which is algebraically degenerate instead of the anti-self-dual part. (The two situations are, of course, isomorphic, and this choice is made to follow the notation of Ref. 7.) When identification of these spinor indices is made with the usual spinor indices in the spinor form of general relativity, it will be found that these were the correct identifications of type of indices.

We may then discuss the (complex) cotangent (and tangent) spaces at each point on the manifold, where the pairing of the four coordinates into two spinors suggests that we, instead, actually look at the two-dimensional spinor cotangent and tangent spaces. We proceed by rewriting the tetrad as

$$e_{\dot{A}} \equiv \phi^{-2}\, dq_{\dot{A}} = \begin{pmatrix} e^3 \\ e^1 \end{pmatrix}, \quad E^{\dot{A}} \equiv \begin{pmatrix} e^4 \\ e^2 \end{pmatrix}, \tag{2.7}$$

with the metric given by

$$g = 2\phi^{-2} E^{\dot{A}} \underset{s}{\otimes} dq_{\dot{A}}. \tag{2.8}$$

These equations allow us to define a very useful symmetric object

$$Q^{\dot{A}\dot{B}} \equiv \begin{pmatrix} Q & R \\ R & P \end{pmatrix}, \tag{2.9a}$$

such that

$$E^{\dot{A}} = -\, dp^{\dot{A}} + Q^{\dot{A}\dot{B}}\, dq_{\dot{B}}, \tag{2.9b}$$

which we will call a rotation spinor of type (0, 1). It is given this name because: (a) Since it is a symmetric traceless second-rank spinor—i.e., type (0, 1)—it generates in the usual way a second-rank anti-self-dual skew tensor which then can be taken as the infinitesimal generator for some (complex) rotation which tells us something about the lack of complete transversality of the coordinates $p^{\dot{A}}$ to $q_{\dot{B}}$; (b) under the appropriate inhomogeneous $\widetilde{\mathrm{GL}}(2, C)$ transformations it will be shown to transform not as a true spinor, but with inhomogeneous terms. By rewriting Eq. (2.8) as

$$g = -\, 2\, dp^{\dot{A}} \underset{s}{\otimes} dq_{\dot{A}} + 2Q^{\dot{A}\dot{B}}\, dq_{\dot{A}} \underset{s}{\otimes} dq_{\dot{B}}, \tag{2.10}$$

we see that $Q^{\dot{A}\dot{B}}$ is that which distinguishes the metric from being trivially flat, and, as such, determines the character of the metric as a double Kerr–Schild metric.[13] In particular, if we define

$$\Delta \equiv \tfrac{1}{2}\phi^{-2} Q^{\dot{A}\dot{B}} Q_{\dot{A}\dot{B}}, \tag{2.11}$$

where the factor ϕ^{-2} has been inserted for later convenience only, it is the vanishing of Δ which causes degeneration to an ordinary Kerr—Schild metric.

We also exhibit explicitly a basis for the (tangent) spaces dual to the (cotangent) spaces spanned by $e_{\dot{A}}$ and $E^{\dot{A}}$. These are

$$-\, \partial_{\dot{A}} \equiv -\frac{\partial}{\partial p^{\dot{A}}} = \begin{pmatrix} \partial_y \\ \partial_x \end{pmatrix} \quad \text{—dual to } E^{\dot{A}}, \tag{2.12a}$$

and

$$\vec{\partial}^{\dot{A}} \equiv \begin{pmatrix} \partial_3 \\ \partial_1 \end{pmatrix} = \phi^2 \left(\frac{\partial}{\partial q_{\dot{A}}} + Q^{\dot{A}\dot{B}} \partial_{\dot{B}} \right) \quad \text{—dual to } e_{\dot{A}}. \tag{2.12b}$$

By using the phraseology that these are dual, one to another, we mean, among other things, that one may write out the exterior derivative, in this basis, as

$$d = E^{\dot{A}} \partial_{\dot{A}} + \phi^{-2}\, dq_{\dot{A}} \vec{\partial}^{\dot{A}}, \quad \text{on a scalar.} \tag{2.13}$$

At this point we should point out that, because of the usual chain rules for calculus and the desire to have $\partial^{\dot{A}}$ mean $\partial/\partial p_{\dot{A}}$, where $p_{\dot{A}} = \epsilon_{\dot{A}\dot{B}} p^{\dot{B}}$, it is necessary to raise and lower spinor indices in the tangent spaces oppositely to those in the cotangent space:

$$\partial^{\dot{A}} = \epsilon^{\dot{A}\dot{B}} \partial_{\dot{B}}, \quad \partial_{\dot{A}} = \epsilon_{\dot{B}\dot{A}} \partial^{\dot{B}}, \tag{2.14}$$

where, of course, the same rules follow for $\partial^{\dot{A}}$ and $\partial/\partial q_{\dot{A}}$, as well. One could also see the necessity for this reversal by noting that it ensures that there is a well-defined action of the cotangent spinors on the tangent ones. That is, if $\partial_{\dot{A}}$ is a tangent spinor and $L^{\dot{A}}$ a cotangent one, then $+L^{\dot{A}}\partial_{\dot{A}} = +L_{\dot{A}}\partial^{\dot{A}}$, instead of the more usual situation with only one spinor space in which, e.g., $p^{\dot{A}} q_{\dot{A}} = -\, p_{\dot{A}} q^{\dot{A}}$.

We now list the bases for self-dual and anti-self-dual 2-forms in their usual[2,8,9] spinor form:

$$\begin{aligned} S^{11} &\equiv 2e^4 \wedge e^2 = E^{\dot{A}} \wedge E_{\dot{A}} \\ S^{12} &\equiv e^1 \wedge e^2 + e^3 \wedge e^4 = \phi^{-2}\, dq_{\dot{A}} \wedge E^{\dot{A}}, \\ S^{22} &\equiv 2e^3 \wedge e^1 = \phi^{-4}\, dq^{\dot{A}} \wedge dq_{\dot{A}} \end{aligned} \tag{2.15a}$$

$$S^{\dot{A}\dot{B}} \equiv \begin{pmatrix} 2e^4 \wedge e^1, & -e^1 \wedge e^2 + e^3 \wedge e^4 \\ -e^1 \wedge e^2 + e^3 \wedge e^4, & 2e^3 \wedge e^2 \end{pmatrix} = 2\phi^{-2}\, dq^{(\dot{A}} \wedge E^{\dot{B})}. \tag{2.15b}$$

We may then calculate the commutation coefficients and

the connections in the standard way, with the results

$$\Gamma_{11} \equiv -\omega_{42} = \phi^{-3} J^{\dot{A}}\, dq_{\dot{A}},$$

$$\Gamma_{12} \equiv -\tfrac{1}{2}(\omega_{12}+\omega_{34}) = \tfrac{1}{2}[\partial_{\dot{A}} Q^{\dot{A}\dot{B}} - \phi^{-3}\partial^{\dot{B}}\phi]\, dq_{\dot{B}} + \tfrac{3}{2}\phi^{-1} J_{\dot{A}} E^{\dot{A}},$$

$$\Gamma_{22} \equiv -\omega_{31} = (\partial_{\dot{A}} Q^{\dot{A}\dot{B}})\, dq_{\dot{B}} - \phi^{-1}(\partial_{\dot{A}}\phi) E^{\dot{A}}, \tag{2.16a}$$

$$\Gamma_{\dot{A}\dot{B}} \equiv \begin{pmatrix} -\omega_{41}, & \tfrac{1}{2}(\omega_{12}-\omega_{34}) \\ \tfrac{1}{2}(\omega_{12}-\omega_{34}), & -\omega_{32} \end{pmatrix} = (\partial_{(\dot{A}} Q_{\dot{B})\dot{C}})\, dq^{\dot{C}} + \phi^{-1} J_{(\dot{A}} E_{\dot{B})} - \phi^{-3}(\partial_{(\dot{A}}\phi)\, dq_{\dot{B})}. \tag{2.16b}$$

We have used ω_{ab} for the Riemannian connection on our (four-complex-dimensional) manifold and Γ_{AB}, $\Gamma_{\dot{A}\dot{B}}$ for the usual induced spinor connections. (The symbol $J_{\dot{A}}$ is an abbreviation for $\partial_{\dot{A}}\phi$ which is a quantity which will reappear many times.) In both these sets of equations one sees a connection between the (usual) spinor indices on the 2-form bases and the connections, and the spinor indices inherited from the component spinors in this particular tetradial representation. That is, the anti-self-dual (hellish) quantities are such that the indices have a 1—1 correspondence, while the self-dual (heavenly) quantities are without free spinor indices in this basis, and therefore are scalars under homogeneous SL(2, C) transformations.

The relations given so far are moderately complicated, but it must be remembered that they are completely general. They will simplify greatly after appropriate gauging and solution of Einstein's equations are accomplished. Now, by calculating Cartan's second curvature form and comparing coefficients as usual (see, e.g., Refs. 2 or 7) we may determine the spinor forms of the curvature quantities:

$$\tfrac{1}{2}C^{(5)} = C_{1111} = 0,$$

$$\tfrac{1}{2}C^{(4)} = C_{1112} = 0,$$

$$\tfrac{1}{2}C^{(3)} = C_{1122} = -\tfrac{1}{6}\phi^2 \partial_{\dot{A}}\partial_{\dot{B}} Q^{\dot{A}\dot{B}}, \tag{2.17a}$$

$$\tfrac{1}{2}C^{(2)} = C_{1222} = -\tfrac{1}{2}\phi^5 \partial^{\dot{A}}\phi^{-3}\partial^{\dot{B}} Q_{\dot{A}\dot{B}} - \tfrac{1}{2}\phi^2 \partial_{\dot{A}}\phi^{-3}\partial^{\dot{A}}\phi,$$

$$\tfrac{1}{2}C^{(1)} = C_{2222} = -\phi^4 \partial^{\dot{A}}\phi^{-2}\partial^{\dot{B}} Q_{\dot{A}\dot{B}} + \phi^4 (\partial^{\dot{A}} Q_{\dot{A}\dot{B}})(\partial_{\dot{C}} Q^{\dot{B}\dot{C}}),$$

$$C_{\dot{A}\dot{B}\dot{C}\dot{D}} = -\phi^2 \partial_{(\dot{A}}\partial_{\dot{B}} Q_{\dot{C}\dot{D})}, \tag{2.17b}$$

which determine the conformal curvature, and the quantities which determine the Einstein tensor

$$-\tfrac{1}{2}R = -\tfrac{1}{2}R^{ab} g_{ab} = +\phi^2 \partial_{\dot{A}}\partial_{\dot{B}} Q^{\dot{A}\dot{B}} - 6\phi^3 \partial_{\dot{A}}\phi^{-4}\partial^{\dot{A}}\phi, \tag{2.18a}$$

$$C_{11\dot{A}\dot{B}} = -\phi^{-1}\partial_{\dot{A}}\partial_{\dot{B}}\phi,$$

$$C_{12\dot{A}\dot{B}} = -\tfrac{1}{2}\phi^2 \partial_{(\dot{A}}\partial^{\dot{C}} Q_{\dot{B})\dot{C}} - \phi\partial_{(\dot{A}}\phi^{-2}\partial_{\dot{B})}\phi, \tag{2.18b}$$

$$C_{22\dot{A}\dot{B}} = -\phi^5 \partial_{(\dot{A}}\phi^{-3}\partial^{\dot{C}} Q_{\dot{B})\dot{C}} + \phi(\partial^{\dot{C}}\phi)\partial_{\dot{C}} Q_{\dot{A}\dot{B}} - \phi\partial_{(\dot{A}}\phi^{-2}\partial_{\dot{B})}\phi.$$

Having, now, general formulas at our disposal, we want to investigate to what extent gauge freedom can help in simplifying the equations. In particular, there is reason to believe that no essential physics lies in a particular parametrization of our totally null congruence—the variables $q_{\dot{A}}$. Therefore, we wish now to consider a transformation to new parameters $q'_{\dot{R}} = q'_{\dot{R}}(q_{\dot{A}})$, only restricting the transformation to be invertible. In addition, since this particular tetrad form is known to be useful, we will determine new longitudinal coordinates $p'^{\dot{S}}$ in such a way as to be able to have the form of the new tetrad in terms of the new variables the same as the form of the old tetrad in terms of the old variables. The purpose of this will be to determine the transformation properties of the various quantities entering into the equations so that we may determine the underlying geometrical significance of the various "constants of integration" entering into the solution of Einstein's equations in such a way as to both simplify the procedure and simplify the interpretation.

We therefore set

$$T_{\dot{R}}^{\ \dot{A}} \equiv \frac{\partial}{\partial q_{\dot{A}}} q'_{\dot{R}}, \quad e'_{\dot{R}} = \phi'^{-2}\, dq'_{\dot{R}} = \rho T_{\dot{R}}^{\ \dot{A}} e_{\dot{A}}, \tag{2.19}$$

where $\rho \equiv (\phi'/\phi)^{-2}$ is a function only of $q_{\dot{A}}$, and

$$E'^{\dot{S}} = \rho^{-1}(T_{\dot{A}}^{-1\,\dot{S}} E^{\dot{A}} - \eta e'^{\dot{S}}), \tag{2.20}$$

where the latter line is required by invariance of the metric and η is so far arbitrary. Form invariance requires that

$$E'^{\dot{S}} = -dp'^{\dot{S}} + Q'^{\dot{S}\dot{T}}\, dq'_{\dot{T}}. \tag{2.21}$$

From Eqs. (2.19) and (2.20) we find that

$$p'^{\dot{R}} = \rho^{-1} T_{\dot{A}}^{-1\,\dot{R}} p^{\dot{A}} + \sigma'^{\dot{R}},$$

$$2\eta\phi^{-2} = T p_{\dot{A}}(\partial\rho^{-1}/\partial q_{\dot{A}}) - \epsilon_{\dot{R}\dot{S}}(\partial\sigma'^{\dot{R}}/\partial q_{\dot{S}}), \tag{2.22}$$

$$Q'^{\dot{R}\dot{S}} = \rho^{-1} T_{\dot{A}}^{-1\,\dot{R}} T_{\dot{B}}^{-1\,\dot{S}} Q^{\dot{A}\dot{B}} + \tfrac{1}{2} T_{\dot{A}}^{-1(\dot{S}}\partial p^{\dot{R})}/\partial q_{\dot{A}},$$

where $\sigma'^{\dot{R}}$ are functions of $q^{\dot{A}}$ only and T is the determinant of the transformation matrix $T_{\dot{R}}^{\ \dot{A}}$. Therefore, we see that the longitudinal variables must go into $q_{\dot{A}}$-dependent linear combinations of themselves modulo possible inhomogeneous terms, which tell us that it is not particularly important which surface of the congruence is labeled which.

Continuing with straightforward calculations, one finds that

$$S'^{RS} = L^R{}_A L^S{}_B S^{AB}, \quad L^R{}_A \equiv \begin{pmatrix} \rho^{-1}T^{-1/2}, & \eta T^{1/2} \\ 0, & \rho T^{1/2} \end{pmatrix} \tag{2.23a}$$

and

$$S'^{\dot{R}\dot{S}} = T T_{\dot{A}}^{-1\,\dot{R}} T_{\dot{B}}^{-1\,\dot{S}} S^{\dot{A}\dot{B}}. \tag{2.23b}$$

Remembering that $T_{\dot{R}}^{\ \dot{A}}$ is an element of $\widetilde{\mathrm{GL}}(2, C)$, we see that the anti-self-dual basis forms transform effectively under $\widetilde{\mathrm{SL}}(2, C)$, while the self-dual ones transform by a simple $\widetilde{\mathrm{SL}}(2, C)$ representation of the differences between our inhomogeneous $\widetilde{\mathrm{GL}}(2, C)$ and SL(2, C). (We use the tilde over quantities without indices to indicate that they refer to notions in the anti-self-dual subspace as opposed to the self-dual one. See Ref. 2 for a more general discussion of the full transformation group.) Analogously one also finds

$$C'_{\dot{R}\dot{S}\dot{T}\dot{U}} = T^{-2} T_{\dot{R}}^{\ \dot{A}} T_{\dot{S}}^{\ \dot{B}} T_{\dot{T}}^{\ \dot{C}} T_{\dot{U}}^{\ \dot{D}} C_{\dot{A}\dot{B}\dot{C}\dot{D}},$$

$$C'_{RSTU} = L_R{}^A L_S{}^B L_T{}^C L_U{}^D C_{ABCD}. \tag{2.24}$$

Among other things which will be useful later, these equations say that $C'^{(3)} = C^{(3)}$.

3. SOLUTION OF EINSTEIN'S EQUATIONS IN VACUUM

Referring to Eqs. (2.18), we can immediately integrate the first triple of equations, which asserts the

existence of a spinor $J_{\dot{A}}$ and a quantity κ, both functions of $q_{\dot{B}}$ only, such that

$$\phi = J_{\dot{A}} p^{\dot{A}} + \kappa. \tag{3.1}$$

(Note that the choice $J_{\dot{1}} = J_{\dot{2}} = -1$, $\kappa = 0$ was made in the discussion of what was called Case II in Ref. 7, while $J_{\dot{1}} = J_{\dot{2}} = 0$, $\kappa = 1$ was made in the discussion there of Case I.) By looking at the transformation equations [Eqs. (2.22)] and setting $\phi' = J'_{\dot{R}} p'^{\dot{R}} + \kappa'$, we find that

$$J'_{\dot{R}} = \rho^{1/2} T_{\dot{R}}{}^{\dot{A}} J_{\dot{A}}, \quad \kappa' = \rho^{-1/2}\kappa - \rho^{1/2} T_{\dot{R}}{}^{\dot{A}} J_{\dot{A}} \sigma'^{\dot{R}}. \tag{3.2}$$

It is clear that, if so desired, one may then utilize the freedom embodied in $T_{\dot{R}}{}^{\dot{A}}$, ρ, and $\sigma'^{\dot{R}}$ to choose $J_{\dot{A}}$ in various specific ways. Because it facilitates greatly the solution of the remaining equations, we use this freedom to *choose $J_{\dot{A}}$ and κ to be constant*. This clearly puts restrictions on future transformations that one may wish to do. However, there is in fact still considerable freedom.

We now introduce a partition of the identity relative to our distinguished constant spinor $J_{\dot{A}}$. That is, we introduce a spinor $K_{\dot{A}}$ such that

$$\delta^{\dot{A}}{}_{\dot{B}} = (1/\tau)(K^{\dot{A}} J_{\dot{B}} - J^{\dot{A}} K_{\dot{B}}), \quad \tau \equiv K^{\dot{A}} J_{\dot{A}}. \tag{3.3}$$

A convenient choice is $K_{\dot{1}} = J_{\dot{1}}$, $K_{\dot{2}} = -J_{\dot{2}}$. We may use this pair of (basis) spinors, thus naturally defined, when needed, to decompose the spinor coordinates into "scalar" coordinates.

The next equation to attack is the one for the curvature scalar. Here we set $R = 0$, but see Appendix B where we show how to include a possible cosmological constant λ. Using the fact that $J_{\dot{A}}$ and κ are constants, we find that

$$\phi^5 \partial_{\dot{A}} \partial_{\dot{B}} \phi^{-3} Q^{\dot{A}\dot{B}} = -\tfrac{1}{2} R = 0. \tag{3.4}$$

From this we infer the existence (see Appendix A for more details) of a spinor $A^{\dot{B}}$ such that

$$Q^{\dot{A}\dot{B}} = \phi^3 \partial^{(\dot{A}} A^{\dot{B})}. \tag{3.5}$$

When this form is inserted into the second triple of Einstein's equations, they take the form

$$\begin{aligned} &\partial_{\dot{A}} \partial_{\dot{B}} \Lambda_0 = 0, \\ &\Lambda_0 = \phi^{-1} \partial_{\dot{C}} \phi^2 A^{\dot{C}}, \end{aligned} \tag{3.6}$$

which imply the existence of $H_{\dot{A}}$ and ξ, functions of $q_{\dot{C}}$ only, such that

$$\phi^{-1} \partial_{\dot{C}} \phi^2 A^{\dot{C}} = \Lambda_0 = 2 H_{\dot{A}} p^{\dot{A}} + 2\xi. \tag{3.7}$$

We may now integrate this equation to find $A^{\dot{C}}$. However, first we note that Einstein's vacuum equations separate into three triples and a singlet. The $C_{11\dot{A}\dot{B}}$ triple is of the form $\partial_{\dot{A}} \partial_{\dot{B}} \Lambda_+ = 0$ for $\Lambda_+ = \phi$. By use of the scalar equation and the Λ_+ triple, we are able to rearrange the second triple to read $\partial_{\dot{A}} \partial_{\dot{B}} \Lambda_0 = 0$. We will find, in addition, that this is sufficient information to allow us to rearrange the third triple to the form $\partial_{\dot{A}} \partial_{\dot{B}} \Lambda_-$, where Λ_- is yet to be defined. It is this particularly simple end form which makes this particular approach successful and which shows up the basic structure of the equations.

Secondly, before integrating Eq. (3.7), we calculate from Eqs. (2.17a) and (3.5), that

$$C^{(3)} = -2 J^{\dot{A}} H_{\dot{A}} \phi^3 \equiv -2\mu\phi^3. \tag{3.8}$$

Since $C^{(5)} = C^{(4)} = 0$, $C^{(3)}$ is an invariant, which suggests that there is particular intrinsic geometric importance to this component of $H_{\dot{A}}$. For simplicity, therefore, we regauge $A^{\dot{B}}$ so as to eliminate the other terms on the right-hand side of Eq. (3.7). In particular, let

$$A^{\dot{A}} \to A^{\dot{A}} + (1/2\tau) H_{\dot{B}} K^{\dot{B}} p^{\dot{A}} + b^{\dot{A}},$$

where

$$J_{\dot{A}} b^{\dot{A}} - (\kappa/2\tau) H_{\dot{B}} K^{\dot{B}} - \xi = 0, \tag{3.9}$$

which in no way changes the value of $Q^{\dot{A}\dot{B}}$ defined by Eq. (3.5) but reduces Eq. (3.7) to a simpler form

$$\tfrac{1}{2} \phi^{-1} \partial_{\dot{A}} \phi^2 A^{\dot{A}} = \mu K^{\dot{A}} p_{\dot{A}} / \tau, \tag{3.10a}$$

which has the solution

$$A^{\dot{B}} = -\phi^{-2} \partial^{\dot{B}} W + (\mu/\tau^2) K^{\dot{B}} K^{\dot{C}} p_{\dot{C}}, \tag{3.10b}$$

where W is an arbitrary sufficiently smooth function of all the coordinates, which plays a role completely analogous to that of the key function in heavenly spaces.[2] Inserting this into Eq. (3.5), we find that

$$Q^{\dot{A}\dot{B}} = -\partial^{(\dot{A}} \phi^4 \partial^{\dot{B})} \phi^{-3} W + (\mu/\tau^2) \phi^3 K^{\dot{A}} K^{\dot{B}}. \tag{3.11}$$

It is worth noting here that Eqs. (2.24) guarantee the invariance of $C^{(3)}$ under our group of transformations, which implies that

$$-2\mu' \phi'^3 = C'^{(3)} = -2\mu\phi^3$$

or

$$\mu' = \rho^{3/2} \mu. \tag{3.12}$$

We may therefore choose ρ so as to give μ whatever *constant* value we might want. (That constant is essentially just a scale factor.) We will not take advantage of this at this time, but it is a very useful thing to do when dealing with the metric for other calculations, such as trajectories or Killing vectors.

We may now proceed to the consideration of the last triple of equations. They can be rearranged, by virtue of our current information, to take the form

$$\partial_{\dot{A}} \partial_{\dot{B}} \Lambda_- = 0, \tag{3.13}$$

where

$$\begin{aligned} \Lambda_- \equiv{}& \Delta + (\phi^{-1} J_{\dot{A}} \partial^{\dot{A}} W)^2 + \phi^{-1} \partial^{\dot{A}} \frac{\partial W}{\partial q^{\dot{A}}} + \mu \phi^4 \partial_\phi \phi^{-3} W \\ &+ \frac{\psi}{2\tau^2} [\psi J^{\dot{C}} - (\phi + \kappa) K^C] \frac{\partial \mu}{\partial q^{\dot{C}}} \\ ={}& \tfrac{1}{2} \phi^4 (\partial^{\dot{A}} \phi^{-2} \partial^{\dot{B}} W)(\partial_{\dot{A}} \phi^{-2} \partial_{\dot{B}} W) + \phi^{-1} \partial^{\dot{A}} \frac{\partial W}{\partial q^{\dot{A}}} \\ &- \mu \phi^4 \partial_\phi \phi^{-1} \partial_\phi \phi^{-1} W + \frac{\psi}{2\tau^2} [\psi J^C - (\phi + \kappa) K^C] \frac{\partial \mu}{\partial q^{\dot{C}}}, \end{aligned} \tag{3.14}$$

where ψ is just an abbreviation for $K^{\dot{A}} p_{\dot{A}}$, and the equations for Λ_- have been written to separate out the (messy) terms containing $\partial\mu/\partial q^{\dot{C}}$ which would vanish if the gauge were chosen so that μ were constant. But Eqs. (3.13) just tell us of $N_{\dot{A}}$ and γ, functions of $q_{\dot{C}}$,

such that

$$\Lambda_- = N_{\dot{A}} p^{\dot{A}} + \gamma, \tag{3.15}$$

which, when Eq. (3.14) is inserted, is simply the (hyperheavenly) equation of Ref. 7, which is the only constraint on W. (In the special case when μ, $N_{\dot{A}}$, and γ vanish and $\phi = 1$, it becomes the second heavenly equation of Ref. 2 in the notation of Ref. 6.) This is a single, second order differential equation for W, linear except for the first term, whose solutions determine all minimally degenerate complex V_4's which are Einstein-flat.

We may insert Eq. (3.15) into the forms for the curvature quantities, and find that the self-dual ones simplify considerably as a result:

$$C_{\dot{A}\dot{B}\dot{C}\dot{D}} = -\phi^2 \partial_{(\dot{A}} \partial_{\dot{B}} Q_{\dot{C}\dot{D})}$$

$$= \phi^3 \partial_{\dot{A}} \partial_{\dot{B}} \partial_{\dot{C}} \partial_{\dot{D}} W - 6\frac{\mu}{\tau^2}\phi^3 J_{(\dot{A}} J_{\dot{B}} K_{\dot{C}}, K_{\dot{D})} \tag{3.16a}$$

$$C^{(5)} = 0 = C^{(4)}, \quad C^{(3)} = 2\mu\phi^3,$$

$$C^{(2)} = 2\phi^5 \left[N_{\dot{A}} J^{\dot{A}} - \left(p^{\dot{A}} + \frac{\kappa}{2\tau} K^{\dot{A}} \right) \frac{\partial \mu}{\partial q^{\dot{A}}} \right],$$

$$C^{(1)} = 2\phi^7 \left\{ \phi \left(\phi^2 \frac{\mu}{\tau} K^{\dot{A}} - \frac{\partial}{\partial q_{\dot{A}}} \right) \left(N_{\dot{A}} + \frac{1}{2\tau} p_{\dot{A}} K^{\dot{R}} \frac{\partial \mu}{\partial q^{\dot{R}}} \right) \right.$$

$$+ J^{\dot{B}} \frac{\partial}{\partial q^{\dot{B}}} (N_{\dot{A}} p^{\dot{A}} + \gamma + 3\mu W)$$

$$- \frac{\psi}{2\tau} J^{\dot{C}} p^{\dot{D}} \frac{\partial^2 \mu}{\partial q^{\dot{C}} \partial q^{\dot{D}}} + \left[2 N_{\dot{A}} J^{\dot{A}} J_{\dot{B}} - \left(p_{\dot{B}} J^{\dot{C}} + J_{\dot{B}} p^{\dot{C}} \right.\right.$$

$$\left.\left. + \frac{\kappa}{\tau} K_{\dot{B}} J^{\dot{C}} \right) \frac{\partial \mu}{\partial q^{\dot{C}}} \right] \partial^{\dot{B}} W \Big\}. \tag{3.16b}$$

Again, one notices that the equations simplify greatly if μ is chosen constant. We should also note that one may regauge Λ_-, if desired, to modify $N_{\dot{A}}$ and γ. In particular, we note that the form of the metric, as determined by $Q^{\dot{A}\dot{B}}$ via Eqs. (3.11), is unchanged by the gauge $W \to W + \alpha + \delta\phi^3$, whereas $\Lambda_- \to \Lambda_- - 3\mu\alpha - 3\phi J^{\dot{A}} \partial\delta/\partial q^{\dot{A}}$, where α and δ are functions of $q_{\dot{C}}$ only. However, $\Lambda_- = N_{\dot{A}} p^{\dot{A}} + \gamma = (N_{\dot{A}} K^{\dot{A}} J_{\dot{B}} p^{\dot{B}} - N_{\dot{A}} J^{\dot{A}} K_{\dot{B}} p^{\dot{B}})/\tau + \gamma = N_{\dot{A}} K^{\dot{A}} (\phi/\tau) - N_{\dot{A}} J^{\dot{A}} (K_{\dot{B}} p^{\dot{B}}/\tau) + \gamma - \kappa N_{\dot{A}} K^{\dot{A}}/\tau$. Therefore, we see that we could choose δ so as to eliminate the $N_{\dot{A}} K^{\dot{A}} \phi/\tau$ term. (A portion of this freedom was utilized in Ref. 7.) By looking at the equations for $C^{(1)}$ and $C^{(2)}$, we see that, of the quantities $N_{\dot{A}}, \gamma$, only $N_{\dot{A}} J^{\dot{A}}$ appears in $C^{(2)}$—and only that for constant μ—which tells us why it cannot be gauged away. Then in $C^{(1)}$, which is rather complicated, our regauging argument tells us that γ has geometrical significance only when $\mu \equiv 0$, in which case we have

$$C^{(1)} \to \phi^7 \left[2 N_{\dot{A}} J^{\dot{A}} J_{\dot{B}} \partial^{\dot{B}} W + (J^{\dot{B}} p^{\dot{A}} + \phi \epsilon^{\dot{B}\dot{A}}) \frac{\partial}{\partial q^{\dot{B}}} N_{\dot{A}} + J^{\dot{A}} \frac{\partial \gamma}{\partial q^{\dot{A}}} \right],$$

$$\text{for } \mu = 0. \tag{3.17}$$

(We will see, when discussing Case I hyperheavenly metrics, that $J^{\dot{A}} \partial\gamma/\partial q^{\dot{A}}$ plays a large role in distinguishing the two cases.)

4. TRANSITION TO $\phi = 1$

In Ref. 7 the solution of Einstein's vacuum equation is divided into two cases: I and II. Their Case I has the conformal function $\phi = 1$. The structure in this work has been set up with the idea of making the transition between the two as easy as possible. Since, in the previous section, $J_{\dot{A}}$ has been an arbitrary constant spinor, we may pass to the limit in which it vanishes. We will see that it is necessary to regauge slightly the canonical function W in order to keep all quantities finite, but it is a very straight forward process. It is worthwhile, first, to point out the geometrical distinction between the two cases.[7] The expansion form θ distinguishes the two cases. The vanishing of θ determines that the totally null surfaces in our space are *plane*, as well as geodesic, of course. Using Eq. (2.2), we easily calculate that

$$\theta = i * [d \ln\phi \wedge \Sigma] = \phi(\phi_y du - \phi_x dv) = -\phi K_{\dot{A}} dq^{\dot{A}} \tag{4.1}$$

so that it becomes quite clear that one gets to these very special spaces, where there is a congruence of totally null (geodesic) plane surfaces, by allowing ϕ to become constant.

In particular, then, we want to allow $J_{\dot{A}}$ to vanish, and $\kappa \to 1$. In order to do this we allow

$$W = \Theta + \mu\kappa\phi\psi^2/2\tau^2. \tag{4.2}$$

Note that as $J_{\dot{A}}$ goes to zero, so do $K_{\dot{A}}$, μ, and $\psi \equiv K^{\dot{A}} p_{\dot{A}}$, so that τ vanishes quadratically, and the above additional term, which diverges, has been chosen so that the new potential Θ and all metric quantities will be well-behaved in this limit. It would have been possible, of course, to have used Θ as a generating potential in all of our equations, from the start. We have chosen W and Θ, however, because they give the simplest form to $Q^{\dot{A}\dot{B}}$ (the metric) and to Λ_- for Cases II and I, respectively, while one easily goes between them via Eq. (4.2). Reevaluating, then, all relevant quantities in terms of Θ instead of W, we may easily make the transition to Case I, where we give names to certain limits:

$$\begin{pmatrix} -g \\ -f \end{pmatrix} \equiv F^{\dot{A}} \equiv \lim_{J_{\dot{C}} \to 0} \left(\frac{\mu K^{\dot{A}}}{\tau} \right),$$

$$\tfrac{1}{2} \begin{pmatrix} -g/f, & 1 \\ -1, & f/g \end{pmatrix} \equiv U^{\dot{A}\dot{B}} \equiv \lim_{J_{\dot{C}} \to 0} \left(\frac{K^{\dot{A}} J^{\dot{B}}}{\tau} \right). \tag{4.3}$$

With $J_{\dot{A}} = 0$, then we obtain

$$A^{\dot{B}} = -\partial^{\dot{B}}\Theta - p_{\dot{A}} p_{\dot{C}} U^{\dot{C}(\dot{A}} F^{\dot{B})},$$

$$Q^{\dot{A}\dot{B}} = -\partial^{\dot{A}}\partial^{\dot{B}}\Theta - F^{\dot{A}} U^{\dot{B}\dot{C}} p_{\dot{C}},$$

$$\Lambda_+ = 0, \quad \Lambda_0 = 2F^{\dot{A}} p_{\dot{A}}, \tag{4.4}$$

$$\Lambda_- = \tfrac{1}{2}(\partial^{\dot{A}}\partial^{\dot{B}}\Theta)(\partial_{\dot{A}}\partial_{\dot{B}}\Theta) + \partial^{\dot{A}} \frac{\partial\Theta}{\partial q^{\dot{A}}} + F^{\dot{A}} \partial_{\dot{A}}\Theta + F_{\dot{A}} p_{\dot{B}} U^{\dot{C}\dot{B}} \partial^{\dot{A}} \partial_{\dot{C}}\Theta$$

$$+ \tfrac{1}{2} p^{\dot{A}} p^{\dot{B}} U_{\dot{A}\dot{B}} \frac{\partial F^{\dot{C}}}{\partial q^{\dot{C}}} = N_{\dot{A}} p^{\dot{A}} + \gamma.$$

All of the connection and curvature quantities are easily calculated from their earlier expression:

$$\Gamma_{11} = 0, \quad \Gamma_{12} = \tfrac{1}{2} F^{\dot{B}} dq_{\dot{B}}, \quad \Gamma_{22} = \left(N^{\dot{B}} + \tfrac{1}{2} p^{\dot{B}} \frac{\partial F^{\dot{C}}}{\partial q^{\dot{C}}} \right) dq_{\dot{B}},$$

$$\Gamma_{\dot{A}\dot{B}}^{\dot{C}}=(\partial_{\dot{A}}\partial_{\dot{B}}\partial_{\dot{C}}\Theta)\,dq^{\dot{C}}+\tfrac{1}{3}F_{(\dot{A}}\,dq_{\dot{B})},$$

$$C^{(5)}=C^{(4)}=C^{(3)}=0,\quad C^{(2)}=-\frac{\partial F^{\dot{A}}}{\partial q_{\dot{A}}},$$

$$C^{(1)}=2\left(F^{\dot{A}}-\frac{\partial}{\partial q_{\dot{A}}}\right)\left(N_{\dot{A}}+\tfrac{1}{2}p_{\dot{A}}\frac{\partial F^{\dot{B}}}{\partial q^{\dot{B}}}\right),$$

$$C_{\dot{A}\dot{B}\dot{C}\dot{D}}=\partial_{\dot{A}}\partial_{\dot{B}}\partial_{\dot{C}}\partial_{\dot{D}}\Theta. \tag{4.5}$$

It is worth pointing out two special facts about the gauge group in this case. Since we have let $J_{\dot{A}}$ vanish, there is no longer any constraint on the group discussed in Sec. 2, to maintain $J_{\dot{A}}$ constant. Much more freedom is allowed than previously: the entire $\widetilde{\mathrm{GL}}(2,C)$ group, but with $\rho=1$. By using the limiting value of Eq. (2.22), we find that such a form-invariant transformation generates a new connection spinor $F'^{\dot{R}}$ such that

$$F'^{\dot{R}}=T_{\dot{A}}^{\dot{R}}F^{\dot{A}}-T^{-1}\frac{\partial}{\partial q_{\dot{R}}}T. \tag{4.6}$$

There is clearly considerable freedom to regauge $F'^{\dot{R}}$, in this way. One may, for example, choose $f'=g'$, which would make the form of $U^{\dot{A}\dot{B}}$ much simpler.

There is, also, another way to simplify the equations in this case; if

$$\Theta\to\Theta-\tfrac{1}{6}L_{\dot{A}\dot{B}\dot{C}}p^{\dot{A}}p^{\dot{B}}p^{\dot{C}}, \tag{4.7}$$

where $L_{\dot{A}\dot{B}\dot{C}}$ is totally symmetric, then the form of $Q^{\dot{A}\dot{B}}$ is changed, but the role of $F^{\dot{A}}$ is not altered in the connections and conformal curvature. Therefore, one may use $L_{\dot{A}\dot{B}\dot{C}}$ to "simplify" the form of $Q^{\dot{A}\dot{B}}$ or, more importantly, of Λ_-. One can also use this freedom to "explain" the occurrence of the somewhat unnatural-looking $U^{\dot{A}\dot{B}}$ in this expression. That is, it just corresponds to a particular choice of $L_{\dot{A}\dot{B}\dot{C}}$ generated by the convenient form we had found for the more general form of the metric, with conformal factor. By insisting on not having terms in Λ_- which are other than first order in $F^{\dot{A}}$ (or zeroth order), we find that a very useful choice for $L_{\dot{A}\dot{B}\dot{C}}$ exists:

$$L_{\dot{1}\dot{1}\dot{1}}=-\tfrac{1}{2}(g-f^2/g),\quad L_{\dot{2}\dot{2}\dot{2}}=-\tfrac{1}{2}(f-g^2/f),$$
$$L_{\dot{1}\dot{1}\dot{2}}=0=L_{\dot{1}\dot{2}\dot{2}}, \tag{4.8}$$

which generates a choice made in Ref. 7 and has the effect of causing

$$Q^{\dot{A}\dot{B}}=-\partial^{\dot{A}}\partial^{\dot{B}}\Theta'+\tfrac{1}{2}\begin{pmatrix}gy+fx, & gx+fy\\ gx+fy, & gy+fx\end{pmatrix}, \tag{4.9}$$

and making minor changes in Λ_- which will not be given here explicitly, but are easily performed via Eq. (4.7).

5. CONCLUSIONS

Our results may be summarized by pointing out that we have shown that Einstein's vacuum equations split into three triples and a singlet which, when in the presence of the minimum possible amount of algebraic degeneracy for a complex V_4, are amenable to reconstruction in the form $\partial_{\dot{A}}\partial_{\dot{B}}\Lambda_\alpha=0$, $\alpha=+,\ 0,\ -$, introducing three spinors and three scalars (along with the cosmological constant as a scalar, from the tenth equation) as constants of integration—that is, functions of $q_{\dot{C}}$ only. We have utilized a group of coordinate and tetrad transformations which leave form-invariant the tetrad, and thereby the metric, which has the interpretation of a reparametrization of the surfaces in the congruence of totally null surfaces which the space possesses by virtue of the (minimal) algebraic degeneracy of its conformal curvature tensor. This group is originally an inhomogeneous version of $\widetilde{\mathrm{GL}}(2,C)$ times a simple representation of $\mathrm{SL}(2,C)$ generated by the inhomogeneities where, of course, $\widetilde{\mathrm{GL}}(2,C)$ acts on the dotted indices and $\mathrm{SL}(2,C)$ on the undotted ones. During the derivation of the general solution it was deemed desirable to restrict somewhat the available group parameters by restraining ϕ to be a function of $p^{\dot{A}}$ only. However, we point out that, once the solution is found, we may always retrieve the more general situation by the action of the group once again.

We have also shown how to generate both cases of Ref. 7 from a uniform formalism, and how to incorporate a nonzero cosmological constant into the equations. It is to be hoped that this is only a start toward a complete investigation of many of the properties of these spaces, whose real cross sections certainly include most of the solutions currently known. Work is progressing on study of the Killing spinors which they possess, as well as other properties.

APPENDIX A

We want here to outline, in somewhat more detail, the calculations leading to the derivation of the form of the hyperheavenly equation given in Eqs. (3.14) and (3.15). We first note a few lemmas which are easily proven or well known, but perhaps unfamiliar in spinor language. First, if $X^{ABC\cdots D}$ is a spinor such that $\partial_A X^{ABC\cdots D}=0$, then there exists, locally, $Y^{B\cdots D}$ such that $X^{AB\cdots D}=\partial^A Y^{B\cdots D}$. Secondly, if $X^{A\cdots D}$ is totally symmetric, iteration of the argument gives (see also Sec. 3 of Ref. 4) a scalar H such that $X^{A\cdots D}=\partial^A\cdots\partial^D H$. Third, any quantity, say Z, can be written in the form $Z=\partial_A B^A$, where B^A is not uniquely determined by this requirement. We also note that the equation $\partial_A\partial_B L_{C\cdots D}=0$ implies the existence of $M_{BC\cdots D}$ and $N_{C\cdots D}$, functions of q_E only, such that $L_{C\cdots D}=p^B M_{BC\cdots D}+N_{C\cdots D}$, i.e., it must, of course, just be linear in p^B. We need a final statement about the equation

$$\phi^n\partial_{\dot{A}}X^{\dot{A}}=L_{\dot{A}}p^{\dot{A}}+\ell. \tag{A1}$$

As may be checked by insertion, the solution is given by

$$X^{\dot{A}}=\partial^{\dot{A}}\Lambda-\frac{\phi^{1-n}}{\tau(n-1)}\left[L_{\dot{B}}p^{\dot{B}}+\ell+\frac{\phi}{\tau(n-2)}K^{\dot{B}}L_{\dot{B}}\right]K^{\dot{A}}, \tag{A2}$$

where $n\neq 1,2$.

From Eq. (3.4), we easily infer the existence of a scalar G such that $\partial_{\dot{B}}\phi^{-3}Q^{\dot{A}\dot{B}}=\partial^{\dot{A}}G$. Now we pick some spinor $A^{\dot{B}}$ such that $G=\frac{1}{2}\partial_{\dot{B}}A^{\dot{B}}$, from which we have that $\partial_{\dot{B}}(\phi^{-3}Q^{\dot{A}\dot{B}}-\partial^{\dot{A}}A^{\dot{B}})=0$. Since $Q^{\dot{A}\dot{B}}$ is symmetric and $\partial_{\dot{B}}\partial^{\dot{B}}\equiv 0$, this is more conveniently written as $\partial_{\dot{B}}(\phi^{-3}Q^{\dot{A}\dot{B}}-\partial^{(\dot{A}}A^{\dot{B})})=0$. Since the quantity in parentheses is symmetric, we easily have that $\phi^{-3}Q^{\dot{A}\dot{B}}-\partial^{(\dot{A}}A^{\dot{B})}=\partial^{\dot{A}}\partial^{\dot{B}}H$, but by regauging $A^{\dot{B}}\to A^{\dot{B}}+\partial^{\dot{B}}H$, not changing

the equation defining $A^{\dot{B}}$, we see that there is always an $A^{\dot{A}}$ such that $Q^{\dot{A}\dot{B}}=\phi^3\partial^{(\dot{A}}A^{\dot{B})}$—Eq. (3.5). Therefore, the second triple of equations is

$$0=C_{12\dot{A}\dot{B}}=-\tfrac{1}{2}\phi^2\,\partial_{(\dot{A}}\partial^{\dot{C}}Q_{\dot{B})\dot{C}}-\phi\partial_{(\dot{A}}\phi^{-2}\mathcal{D}_{\dot{B})}\phi$$
$$=-\tfrac{1}{2}\phi^4\,\partial^{\dot{C}}\phi^{-2}\partial_{(\dot{A}}Q_{\dot{B})\dot{C}}$$
$$=\tfrac{1}{4}\phi^4(\partial_{\dot{A}}\partial^{\dot{C}}\phi\partial_{(\dot{B}}A_{\dot{C})}+\partial_{\dot{B}}\partial^{\dot{C}}\phi\partial_{(\dot{A}}A_{\dot{C})}+2J_{(\dot{A}}\partial^{\dot{C}}\partial_{\dot{B})}A_{\dot{C}})$$
$$=\tfrac{1}{4}\phi^4(\partial_{\dot{A}}\partial_{\dot{B}}\partial^{\dot{C}}\phi A_{\dot{C}}+\partial_{\dot{B}}\partial^{\dot{C}}J_{[\dot{A}}A_{\dot{C}]}+\partial_{\dot{A}}\partial^{\dot{C}}J_{[\dot{B}}A_{\dot{C}]})$$
$$=\tfrac{1}{4}\phi^4\partial_{\dot{A}}\partial_{\dot{B}}\phi^{-1}\partial^{\dot{C}}\phi^2A_{\dot{C}},$$

where some simple steps have been omitted. Therefore, we have

$$\phi^{-1}\partial^{\dot{C}}\phi^2A_{\dot{C}}=N_{\dot{A}}p^{\dot{A}}+\xi. \tag{A3}$$

Using Eq. (A2) for $n=-1$, we acquire the result

$$A^{\dot{C}}=\phi^{-2}\partial^{\dot{C}}W+(H_{\dot{B}}p^{\dot{B}}+\xi-\phi H_{\dot{B}}K^{\dot{B}}/3\tau)K^{\dot{C}}/2\tau.$$

However, if, as pointed out in Eqs. (3.9), we first regauge the right-hand side of Eq. (A3), we clearly get just Eq. (3.10), from which Eq. (3.11) follows directly.

We must now rearrange the third triple, which is the most complicated task so far. Using the constancy of $J_{\dot{A}}$, which implies $\mathcal{D}^{\dot{C}}\phi=\phi^2Q^{\dot{C}\dot{D}}J_D$, and Eq. (2.18b) gives

$$C_{22\dot{A}\dot{B}}=-\phi^5\partial_{(\dot{A}}\phi^{-3}\mathcal{D}^{\dot{C}}Q_{\dot{B})\dot{C}}+\phi^3Q^{\dot{C}\dot{D}}J_{\dot{D}}\partial_{\dot{C}}Q_{\dot{A}\dot{B}}-\phi\mathcal{D}_{(\dot{A}}Q_{\dot{B})\dot{C}}J^{\dot{C}}.$$

The last two terms can be manipulated to give

$$\phi^3(Q_{\dot{C}[\dot{A}}\partial^{\dot{C}}Q_{\dot{D}]\dot{B}}+Q_{\dot{C}[\dot{B}}\partial^{\dot{C}}Q_{\dot{D}]\dot{A}})J^{\dot{D}}-\phi^3J^{\dot{C}}\partial Q_{\dot{C}(\dot{B}}/\partial q^{\dot{A})}$$
$$=\phi J_{(\dot{A}}\partial^{\dot{R}}Q_{\dot{B})\dot{R}}-\phi^3J^{\dot{C}}\partial Q_{\dot{A}\dot{B}}/\partial q^{\dot{C}},$$

which implies

$$C_{22\dot{A}\dot{B}}=-\phi^6\partial_{(\dot{A}}[\phi^{-4}\partial^{\dot{C}}Q_{\dot{B})\dot{C}}-2J^{\dot{C}}\partial A_{\dot{B})}/\partial q^{\dot{C}}]$$
$$\equiv-\phi^6\partial_{(\dot{A}}X_{\dot{B})}. \tag{A4}$$

The spinor $X_{\dot{B}}$ is defined by this equation only to within a linear term in $p^{\dot{A}}$, so we may take

$$X_{\dot{B}}=\phi^{-2}Q^{\dot{C}\dot{D}}\partial_{\dot{D}}Q_{\dot{B}\dot{C}}+\phi^{-2}\partial Q_{\dot{B}\dot{C}}/\partial q_{\dot{C}}-J^{\dot{C}}\partial A_{\dot{B}}/\partial q^{\dot{C}}$$
$$+\alpha p_{\dot{B}}, \tag{A5}$$

where α will be determined conveniently later. We define the first term on the right-hand side of this equation to be $Z_{\dot{B}}$ and return to it shortly. Using Eq. (3.5) the rest of the terms on the right-hand side, excluding $\alpha p_{\dot{B}}$, can be written as

$$-\phi\partial_{(\dot{B}}\partial A_{\dot{C})}/\partial q_{\dot{C}}-J^{\dot{C}}\partial A_{\dot{B}}/\partial q^{\dot{C}}$$
$$=-\partial_{\dot{B}}\phi\partial A_{\dot{C}}/\partial q_{\dot{C}}+\phi\partial_{[\dot{B}}\partial A_{\dot{C}]}/\partial q_{\dot{C}},$$
$$-J_{[\dot{C}}\partial A_{\dot{B}]}/\partial q_{\dot{C}}$$
$$=-\partial_{\dot{B}}\phi\partial A_{\dot{C}}/\partial q_{\dot{C}}+\tfrac{1}{2}\partial(\phi\partial_{\dot{C}}A^{\dot{C}}+2J_{\dot{C}}A^{\dot{C}})/\partial q^{\dot{B}}$$
$$=\partial_{\dot{B}}(-\phi\partial A_{\dot{C}}/\partial q_{\dot{C}})-(K_{\dot{A}}p^{\dot{A}}\partial\mu/\partial q^{\dot{B}})/\tau,$$

where the last equality requires the use of Eq. (3.7) regauged by Eq. (3.9), i.e., $\phi^{-1}\partial_{\dot{C}}\phi^2A^{\dot{C}}=2\mu\psi/\tau$. The trick is now to choose $\alpha=-(K^{\dot{A}}\partial\mu/\partial q^{\dot{A}})/2\tau$, so that the entire expression becomes a gradient; that is, with this choice of α, all terms on the right-hand side of Eq. (A4) except the first take the form $\partial_{\dot{B}}[-\phi\partial A_{\dot{C}}/\partial q_{\dot{C}}+(\psi p^{\dot{C}}\partial\mu/\partial q^{\dot{C}})/2\tau+\eta_{\dot{C}}p^{\dot{C}}]$.

Now the other term, named $Z_{\dot{B}}=\partial_{\dot{B}}\Delta-Q_{\dot{B}\dot{C}}\partial_{\dot{S}}\phi^{-2}Q^{\dot{C}\dot{S}}$, where Δ is the function defined by Eq. (2.11). For the last term in $Z_{\dot{B}}$, then, we calculate $\partial_{\dot{B}}\phi^{-2}Q^{\dot{B}\dot{A}}=2\phi^{-3}J^{\dot{A}}J_{\dot{B}}\partial^{\dot{B}}W+\mu K^{\dot{A}}/\tau$, and rewrite $Q_{\dot{B}\dot{C}}$ in the convenient form

$$\phi^{-3}Q_{\dot{B}\dot{C}}=-\phi^{-1}\partial_{\dot{B}}\partial_{\dot{C}}\phi^{-1}W+2\phi^{-4}J_{\dot{B}}J_{\dot{C}}W+\mu K_{\dot{B}}K_{\dot{C}}/\tau^2.$$

It is then easy to show that

$$-Q_{\dot{B}\dot{C}}\partial_{\dot{S}}\phi^{-2}Q^{\dot{C}\dot{S}}=\partial_{\dot{B}}(\phi^{-1}J_{\dot{A}}\partial^{\dot{A}}W)^2+\mu\partial_{\dot{B}}(\phi^4K^{\dot{C}}\partial_{\dot{C}}\phi^{-3}W)/\tau,$$

from which the first form of Λ_- in Eq. (3.14) follows if we insert our expression for A_B and collect terms.

APPENDIX B

In order to include a nonzero cosmological constant, we must solve

$$\phi^5\partial_{\dot{A}}\partial_{\dot{B}}\phi^{-3}Q^{\dot{A}\dot{B}}=2\lambda \tag{B1}$$

from Eq. (3.4). Using Eq. (A2) for $n=5$, we easily find that $\partial_{\dot{B}}\phi^{-3}Q^{\dot{A}\dot{B}}=\partial^{\dot{A}}\Lambda-(\lambda\phi^{-4}K^{\dot{A}})/2\tau$, from which we obtain

$$Q^{\dot{A}\dot{B}}=\phi^3\partial^{(\dot{A}}A^{\dot{B})}+(\lambda K^{\dot{A}}K^{\dot{B}})/6\tau^2 \tag{B2}$$

with $A^{\dot{A}}$ independent of λ, so that the solution to the second triple, in terms of $A^{\dot{A}}$, is just as before. However, note that this is not actually the most general form for $Q^{\dot{A}\dot{B}}$ which satisfies Eq. (B1), however, but rather

$$Q^{\dot{A}\dot{B}}=\phi^3\partial^{(\dot{A}}A^{\dot{B})}+\frac{\lambda/\tau^2}{6\alpha-4\beta+\gamma}$$
$$\left[\alpha K^{\dot{A}}K^{\dot{B}}+2\beta\frac{\psi}{\phi}K^{(\dot{A}}J^{\dot{B})}+\gamma\frac{\psi^2}{\phi^2}J^{\dot{A}}J^{\dot{B}}\right]. \tag{B3}$$

where α, β, and γ are arbitrary functions of q_C such that $6\alpha-4\beta+\gamma$ does not vanish. The choice given above is just $\beta=0=\gamma$ and is the simplest choice. However, for purposes of making the transition to the case $\phi=1$, other choices such as $\alpha=\beta=0$ are best. We continue now, in the general case with just Eq. (B2). Then, noting that λ is, of course, a constant, we proceed to integrate the third triple. The quantity $X_{\dot{B}}$ in Eq. (A4) has exactly the same form as before, but now the $Q^{\dot{A}\dot{B}}$ within it contains λ terms. We easily follow through the same derivation as in Appendix A, with the result that we must add a single additional (λ-dependent) term to Λ_-; i.e., to the hyperheavenly equation

$$\Lambda_-(\text{with }\lambda)=\Lambda_-(\text{without }\lambda)+\tfrac{1}{6}\lambda\phi^3\partial_\phi\phi^{-4}\partial_\phi W. \tag{B4}$$

If, on the other hand we want to "add" λ to the solution with $\phi=1$, we take Eq. (B3) with $\alpha=\beta=0$ and let $J_{\dot{A}}$ vanish, along with all the other concommitant quantities which must vanish. Then a convenient regauging is performed, and we obtain the appropriate formulas for Case I ($\phi=1$):

$$Q^{\dot{A}\dot{B}}\to-\partial^{\dot{A}}\partial^{\dot{B}}\Theta+\tfrac{2}{3}F^{(\dot{A}}p^{\dot{B})}+\tfrac{1}{3}\lambda p^{\dot{A}}p^{\dot{B}}, \tag{B5}$$

$$\Lambda_-\to\tfrac{1}{2}(\partial^{\dot{A}}\partial^{\dot{B}}\Theta)(\partial_{\dot{A}}\partial_{\dot{B}}\Theta)+\partial^{\dot{A}}\partial\Theta/\partial q^{\dot{A}}+F^{\dot{A}}\partial_{\dot{A}}\Theta-\tfrac{2}{3}F^{\dot{A}}p^{\dot{B}}\partial_{\dot{A}}\partial_{\dot{B}}\Theta$$
$$+\tfrac{1}{18}(F^{\dot{A}}p_{\dot{A}})^2+\tfrac{1}{6}p^{\dot{A}}p^{\dot{B}}\partial F_{\dot{B}}/\partial q^{\dot{A}}-\lambda(p^{\dot{A}}\partial_{\dot{A}}\Theta-\Theta), \tag{B6}$$

where, in the language of Sec. 4, another choice of $L_{\dot{A}\dot{B}\dot{C}}$ has been made relative to the form of Λ_- derived there which, although it has the disadvantage of generat-

ing quadratic terms in $F^{\dot{A}}$, has the advantages that (1) it makes the additional λ-dependent terms take on their simplest possible form and (2) it does not need any ad hoc spinors, such as $U^{\dot{A}\dot{B}}$, for its description. Another way of looking at this last comment is just that this, in Case I, is the "natural" gauge for F^{A}, from which all the others may be generated by an (arbitrary) assumption of a particular form for $L_{\dot{A}\dot{B}\dot{C}}$. Starting from the form given in Eqs. (4.4) for $Q^{\dot{A}\dot{B}}$ and Λ_{α} in terms of $U^{\dot{A}\dot{B}}$, we may think of having gotten there from this natural gauge by a spinor $L^{\dot{A}\dot{B}\dot{C}}$ [see Eq. (4.7)] of the form

$$L^{\dot{1}\dot{1}\dot{1}}=-\tfrac{1}{2}f^{2}/g, \quad L^{\dot{1}\dot{1}\dot{2}}=\tfrac{1}{6}f,$$
$$L^{\dot{1}\dot{2}\dot{2}}=\tfrac{1}{6}g, \quad L^{\dot{2}\dot{2}\dot{2}}=-\tfrac{1}{2}g^{2}/f.$$

*Work supported in part by the Fomento Educacional, A.C., México 5, D.F., Mexico, and by the Centro de Investigacion y de Estudios Avanzados del Instituto Politecnico Nacional, México 14, D.F., Mexico.

†On leave of absence from the University of Warsaw, Warsaw, Poland.

[1]R. Penrose, First Award Gravity Foundation Essay, "The Non-linear Graviton" (1975); E.T. Newman, Tel Aviv GR7 Lecture (1974) and a lecture at the Enrico Fermi Summer School of Varenna, 1975.

[2]J.F. Plebański, J. Math. Phys. **16**, 2395 (1975).

[3]J.F. Plebański and S. Hacyan, J. Math. Phys. **16**, 2403 (1975).

[4]J.D. Finley, III and J.F. Plebański, J. Math. Phys. **17**, 585 (1976).

[5]S. Hacyan and J.F. Plebański, submitted to J. Math. Phys.

[6]C.P. Boyer and J.F. Plebański, submitted to J. Math. Phys.

[7]J.F. Plebański and I. Robinson, Phys. Rev. Lett. **37**, 493 (1976). A detailed discussion is in the Proceedings of the Symposium on Asymptotic Structure of Space Time, University of Cincinnati, Ohio, June, 1976, to be published by Plenum Press.

[8]E.J. Flaherty, "Complex Strings and Asymptotic Twistors," in the Proceedings of the International Symposium on Mathematical Physics, Mexico City, January 1976, p. 713.

[9]G.C. Debney, R.P. Kerr, and A. Schild, J. Math. Phys. **10**, 1842 (1969), or J.F. Plebański, "Spinors, Tetrads and Forms," unpublished monograph from Centro de IEA del IPN, México 14, D.F.

[10]For more on the helicity structure of the representation of the Lorentz group and its relevance to the structure of the Riemann tensor see, e.g., R. Debever, Acad. Roy. Belg. Cl. Sci. Bull. Ser. 5, **60**, 998 (1974).

[11]See also A. Garcia, J.F. Plebański, and I. Robinson, preprint, on the addition of λ to hyperheavens.

[12]For a more rigorous point of view, see E.J. Flaherty, *Hermitian and Kählerian Geometry in Relativity* (Springer-Verlag, Berlin, 1976). The spaces considered here are (related to) complex extensions of what he calls 1GSF spacetimes on pp. 206ff.

[13]J.F. Plebański and A. Schild, to be published in Nuovo Cimento.

Lie series and invariant functions for analytic symplectic maps*

Alex J. Dragt and John M. Finn†

Center for Theoretical Physics, Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742
(Received 3 March 1976)

Symplectic maps (canonical transformations) are treated from the Lie algebraic point of view using Lie series and Lie algebraic techniques. It is shown that under very general conditions an analytic symplectic map can be written as a product of Lie transformations. Under certain conditions this product of Lie transformations can be combined to form a single Lie transformation by means of the Campbell–Baker–Hausdorff theorem. This result leads to invariant functions and generalizes to several variables a classic result of Birkhoff for the case of two variables. It also provides a new approach since the connection between symplectic maps, Lie algebras, invariant functions, and Birkhoff's work has not been previously recognized and exploited. It is expected that the results obtained will be applicable to the normal form problem in Hamiltonian mechanics, the use of the Poincaré section map in stability analysis, and the behavior of magnetic field lines in a toroidal plasma device.

1. INTRODUCTION AND NOTATION

The purpose of this paper is to discuss canonical transformations from the Lie algebraic point of view using Lie series and Lie algebraic techniques. The study of canonical transformations or maps is important for several reasons: As is well known, canonical transformations preserve Hamilton's equations of motion.[1–3] In this context, they can be used systematically to bring a Hamiltonian to a simpler form from which the solutions to the equations of motion can be more easily discovered.[4–6] Secondly, the canonical coordinates $p(t)$, $q(t)$ at time t for any Hamiltonian system are related to their values p_0, q_0 at time $t=t_0$ by a canonical transformation.[1–3] In addition, the Poincaré section map used to investigate stability behavior is canonical.[7,8] Finally, the behavior of magnetic field lines in a toroidal plasma device can be characterized by a canonical map.[9] We expect that our results will have application to all these areas.

The most commonly used method of describing canonical transformations is by means of transformation functions of mixed variables.[1–3] As has been discussed by Deprit and others, this method has certain drawbacks which can be overcome by the use of Lie series.[10–12] In this paper we will employ a variant of the Lie series approach.

The remainder of this section and Theorem 1 of the next section are devoted to a review of well-known material concerning Lie series and to a development of notation.[3,5,10–12] Our purpose is to make the material of this paper relatively self-contained.

We shall be working with a phase space consisting of the $2n$ variables $(q_1 \cdots q_n, p_1 \cdots p_n)$. The *Lie product* of any two functions f and g of the phase space variables will be defined by the Poisson bracket operation,

$$[f,g]=\sum_i \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}-\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}. \tag{1.1}$$

The set of all functions defined on phase space has an obvious linear vector space structure since it is closed under addition and scalar multiplication. Also the "multiplication" rule (1.1) satisfies all the requirements for a Lie product including the Jacobi condition

$$[f,[g,h]]+[g,[h,f]]+[h,[f,g]]=0. \tag{1.2}$$

Consequently, *functions on phase space may be viewed as elements in a Lie algebra.* We remind the reader that the equations of motion generated by a Hamiltonian H can themselves be written in terms of Lie products,

$$\dot{q}_i=[q_i,H], \quad \dot{p}_i=[p_i,H]. \tag{1.3}$$

A canonical transformation to new variables $Q(q,p)$, $P(q,p)$ is defined to be any transformation satisfying

$$\begin{aligned}&[Q_i,Q_j]=[q_i,q_j]=0,\\ &[P_i,P_j]=[p_i,p_j]=0,\\ &[Q_i,P_j]=[q_i,p_j]=\delta_{ij}.\end{aligned} \tag{1.4}$$

That is, *canonical transformations are those transformations which preserve the Lie algebraic structure.*

At this point it is notationally convenient to collect the two sets of n variables q, p into a combined set of $2n$ variables $z_1 \cdots z_{2n}$ by the rule

$$z_i=q_i, \quad z_{n+i}=p_i, \quad i=1,\ldots,n. \tag{1.5}$$

In terms of the z's the fundamental Poisson bracket rules (1.4) become

$$[z_i,z_j]=J_{ij}, \tag{1.6}$$

where J denotes the antisymmetric $2n\times 2n$ matrix

$$J=\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}. \tag{1.7}$$

Here each entry in J is an $n\times n$ block. We note that J has the properties

$$\begin{aligned}&\tilde{J}=-J,\\ &J^2=-I,\\ &\det J=I.\end{aligned} \tag{1.8}$$

The general Lie product (1.1) is given in terms of the z's by the relation

$$[f,g]=\sum_{k,l}\left(\frac{\partial f}{\partial z_k}\right)J_{kl}\left(\frac{\partial g}{\partial z_l}\right). \tag{1.9}$$

Suppose we introduce new variables $\bar{z}(z)$ and require that the transformation be canonical. Combining (1.4) and (1.9) we find

$$J_{ij}=[\bar{z}_i,\bar{z}_j]=\sum_{kl}\left(\frac{\partial \bar{z}_i}{\partial z_k}\right)J_{kl}\left(\frac{\partial \bar{z}_j}{\partial z_l}\right). \tag{1.10}$$

Let M be the Jacobian matrix for the transformation going from z to $\bar{z}$,

$$M_{ik}(z)=\frac{\partial \bar{z}_i}{\partial z_k}. \tag{1.11}$$

Employing M, we find that (1.10) can be written in the compact form

$$MJ\tilde{M}=J. \tag{1.12}$$

This is just the condition that the matrix M must satisfy in order to belong to the symplectic group in $2n$ dimensions. We conclude that the necessary and sufficient condition for a transformation to be canonical is that its Jacobian matrix be symplectic.[8,12,13] For this reason a canonical transformation is often called a symplectic map.

In this paper we will study analytic symplectic maps. They are canonical transformations given by convergent power series. We write these power series as

$$\bar{z}_i=\sum_{|\sigma|>0}a_i(\sigma)z^\sigma. \tag{1.13}$$

Here σ denotes a collection of exponents $(\sigma_1\cdots\sigma_{2n})$ and

$$|\sigma|=\sum_1^{2n}\sigma_i,\quad z^\sigma=z_1^{\sigma_1}z_2^{\sigma_2}\cdots z_{2n}^{\sigma_{2n}}. \tag{1.14}$$

Note that in the sum (1.13) we have purposely excluded constant terms by requiring $|\sigma|>0$. We do this to eliminate a possible nuisance later on and because we are not interested in transformations which simply translate the origin in phase space.

More specifically, our purpose is to study the relation between transformations of the form (1.13) and Lie series. Lie series and Lie transformations will be defined in the next section. There we will also see that under certain conditions the transformation (1.13) can be written as a product of Lie transformations. In Sec. 3 we will apply our results to several symplectic maps studied previously by other authors. Section 4 is devoted to the development and application of various Lie algebraic tools including the Campbell–Baker–Hausdorff formula. In Sec. 5 we apply the Campbell–Baker–Hausdorff formula to produce invariant functions for the map (1.13). An invariant function is a function f with the property $f(\bar{z})=f(z)$. The existence and form of an invariant function enables one to study the effect of applying the map (1.13) many times in succession. We will learn that the determination of invariant functions is closely related to the determination of integrals of motion in Hamiltonian mechanics. In particular, invariant functions tell us a great deal about the underlying map just as integrals of motion characterize trajectories in mechanics. Our results are summarized in a final section.

2. LIE SERIES AND TRANSFORMATIONS

For the remainder of this paper we adopt the notational convention that lower case letters f,g, etc., denote functions and capital letters F, G, etc., denote operators.

Let f be a specified function on phase space, and let e be any function. We associate with f the linear differential operator F by the rule

$$Fe=[f,e]. \tag{2.1}$$

For example, if $f=z_1$, then $F=\partial/\partial z_{n+1}$. We shall call F the *Lie operator* associated with f.

In general, Lie operators do not commute. Let F and G be the Lie operators associated with the functions f and g. We will denote their commutator by $\{F,G\}$,

$$\{F,G\}=FG-GF. \tag{2.2}$$

Suppose h is the function defined by

$$h=[f,g]. \tag{2.3}$$

We find, using the Jacobi relation (1.2),

$$\begin{aligned}\{F,G\}e&=[f,[g,e]]-[g,[f,e]]\\&=[[f,g],e]=He,\end{aligned} \tag{2.4}$$

where H is the Lie operator associated with h. Since e is any arbitrary function we may rewrite (2.4) as

$$H=\{F,G\}. \tag{2.5}$$

Comparing (2.5) and (2.3), we see that Lie operators form a Lie algebra under commutation which is homomorphic to the Poisson bracket Lie algebra of the underlying functions.[14] In particular, we are guaranteed that the commutator of two Lie operators is again a Lie operator. This fact will be important in Secs. 4 and 5.

We next consider infinite operator power series, called *Lie series*, of the form $\sum_0^\infty a_n F^n$ with the convention $F^0=I$. Of particular interest is the exponential series $\exp(F)$ defined as expected by

$$\exp(F)=\sum_0^\infty F^n/n!. \tag{2.6}$$

We shall call $\exp(F)$ the *Lie transformation associated* with f and *generated* by F.

Lie transformations have two remarkable properties: Suppose d and e are any two functions. Then we find[10]

$$\exp(F)(de)=(\exp(F)\,d)(\exp(F)\,e) \tag{2.7}$$

and

$$\exp(F)[d,e]=[\exp(F)\,d,\exp(F)\,e]. \tag{2.8}$$

These results follow from the properties of the exponential series and the relations

$$F(de)=(Fd)\,e+d(Fe), \tag{2.9}$$

$$F[d,e]=[Fd,e]+[d,Fe]. \tag{2.10}$$

That is, F is a *derivation* with respect to both ordinary and Poisson bracket multiplication.[15]

We are ready to explore the relation between symplectic maps and Lie transformations. The first result is immediate:

Theorem 1: If $\exp(F)$ is the Lie transformation associated with the analytic function f, then the infinite series given by

$$\bar{z}_i = \exp(F)\, z_i \tag{2.11}$$

is, providing it converges, an analytic symplectic map.

Proof: We simply use (2.8) and (2.6) to find

$$[\bar{z}_i, \bar{z}_j] = [\exp(F)\, z_i, \exp(F)\, z_j]$$
$$= \exp(F)[z_i, z_j] = \exp(F)\, J_{ij} = J_{ij}. \tag{2.12}$$

The converse result is somewhat more difficult to state and to prove. We shall first state the result, and then work up to its proof in stages.

Theorem 2: Suppose one is given an analytic symplectic map in the form (1.13). Let $M(0)$ denote the matrix defined by (1.11) with all $z_i = 0$. Assume that $M(0)$ is joined to the identity matrix by a continuous one parameter subgroup of symplectic matrices. Or equivalently, assume that $M(0)$ can be written in the form

$$M(0) = \exp(JS), \tag{2.13}$$

where S is a symmetric matrix. Then there exist homogeneous polynomials g_2, g_3, etc., of degree 2, 3, etc., and associated operators G_2, G_3, etc., such that the map (1.13) can be written in the infinite product form

$$\bar{z}_i = [\exp(G_2)\exp(G_3)\cdots]\, z_i. \tag{2.14}$$

The proof of this result is most easily accomplished by a series of lemmas.

Lemma 1: *A* set of $2n$ functions $f_1 \cdots f_{2n}$ satisfying

$$[z_i, f_j] = [z_j, f_i] \tag{2.15}$$

exists if and only if there is a function g such that

$$f_i = [g, z_i] = G\, z_i. \tag{2.16}$$

Proof: First suppose that each f_i is given by (2.16). Then we quickly verify (2.15): We find

$$[z_i, f_j] - [z_j, f_i] = [z_i, [g, z_j]] - [z_j, [g, z_i]]$$
$$= -[g, [z_i, z_j]] = -[g, J_{ij}] = 0.$$

Now suppose (2.15) is true. We introduce auxiliary variables z^* by the rule

$$z_i^* = \sum_j J_{ij} z_j. \tag{2.17}$$

Because $\tilde{J}J = I$, we can immediately write the inverse relation

$$z_k = \sum_i J_{ik} z_i^*. \tag{2.18}$$

Let f be any function. We find

$$[z_i, f] = \sum_{jk} \left(\frac{\partial z_i}{\partial z_j}\right) J_{jk} \left(\frac{\partial f}{\partial z_k}\right)$$
$$= \sum_k J_{ik} \left(\frac{\partial f}{\partial z_k}\right) = \sum_k \left(\frac{\partial f}{\partial z_k}\right)\left(\frac{\partial z_k}{\partial z_i^*}\right) = \left(\frac{\partial f}{\partial z_i^*}\right). \tag{2.19}$$

Because of this relation, the hypothesis (2.15) implies

$$\frac{\partial f_i}{\partial z_j^*} = \frac{\partial f_j}{\partial z_i^*}$$

which means that $\sum_i f_i\, dz_i^*$ is an exact differential. Therefore the function g given by the path integral

$$g = -\int^{z^*} \sum_i f_i\, dz_i'^* \tag{2.20}$$

is well defined, and satisfies $[g, z_i] = -(\partial g/\partial z_i^*) = f_i$. Using (2.17) we obtain the explicit formula

$$g(z) = -\int^{z} \sum_{ij} f_i J_{ij}\, dz_j'. \tag{2.21}$$

Lemma 2: Let g_s be a homogeneous polynomial of degree s. That is, we have

$$g_s(z) = \sum_{|\sigma|=s} b(\sigma)\, z^\sigma \tag{2.22}$$

for some set of coefficients. Also, let $\mathcal{P}_s$ denote the set of all homogeneous polynomials of degree s. Then, since the Poisson bracket operation involves multiplication and two differentiations, we have for any two homogeneous polynomials g_r, g_s the relation

$$[g_r, g_s] \in \mathcal{P}_{r+s-2}. \tag{2.23}$$

Lemma 3: A necessary and sufficient condition for a symplectic matrix N to lie on a continuous one parameter symplectic subgroup joined to the identity is that there exist a symmetric matrix S such that

$$N = \exp(JS). \tag{2.24}$$

Proof: Suppose N is a matrix of the form (2.24). Then we find by direct computation that N is symplectic,

$$N\tilde{J}\tilde{N} = \exp(JS)\, J \exp(JS)^{\sim} = \exp(JS)\, J \exp(S\tilde{J})$$
$$= \exp(JS)\, J \exp(-SJ)\, J^{-1} J = \exp(JS)\exp(-JS)\, J = J. \tag{2.25}$$

A similar result holds for the matrix $N(\tau)$ defined by

$$N(\tau) = \exp(\tau JS) \tag{2.26}$$

where τ is a parameter. It follows that N lies on a continuous one parameter subgroup joined to the identity. Now assume the converse, namely that N does lie on a one parameter subgroup. Without loss of generality we assume that the group is parameterized in such a way that

$$N(\tau_1 + \tau_2) = N(\tau_1) N(\tau_2), \tag{2.27a}$$
$$N(0) = I, \tag{2.27b}$$
$$N(1) = N. \tag{2.27c}$$

Now differentiate (2.27a) with respect to τ_1 and then set $\tau_1 = 0$ and $\tau_2 = \tau$ to obtain the result

$$N'(\tau) = N'(0)\, N(\tau). \tag{2.28a}$$

This equation with the initial condition (2.27b) has the unique solution

$$N(\tau) = \exp[\tau N'(0)]. \tag{2.28b}$$

Let us write $N'(0) = JS$ where S is an undetermined matrix. Next suppose τ is small. Then

$$N(\tau) = \exp(\tau JS) \cong I + \tau JS. \tag{2.29a}$$

Enforcing the symplectic condition (1.12) gives

$$(I + \tau JS)\, J (I + \tau \tilde{S}\tilde{J}) \cong J. \tag{2.29b}$$

Consequently, equating powers of τ, we have

$$JSJ + J\tilde{S}\tilde{J} = 0. \tag{2.29c}$$

Finally, use of (1.8) implies the expected conclusion

$$S = \tilde{S}. \tag{2.30}$$

Now set $\tau = 1$. The result is a matrix written in the form (2.24).

Cautionary remark: Not every symplectic matrix can be written in the form (2.24). A counter example in the 2×2 case is the matrix given by

$$N = \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix}. \tag{2.31}$$

Lemma 4: Suppose that $M(0)$ is joined to the identity by a continuous one parameter subgroup. Then there exists a second degree homogeneous polynomial g_2 such that

$$(\exp G_2)\, z_i = \sum_j M_{ij}(0)\, z_j. \tag{2.32}$$

Proof: According to Lemma 3 we may write $M(0)$ in the form

$$M(0) = \exp(JS). \tag{2.33}$$

We define g_2 by the expression

$$g_2 = -\tfrac{1}{2}\sum_{i,k} S_{ik} z_i z_k, \tag{2.34}$$

and find

$$G_2 z_i = [g_2, z_i] = \sum_j (JS)_{ij} z_j. \tag{2.35}$$

The desired result (2.32) follows immediately by exponentiation.

Lemma 5: Let $r(>1)$ denote a "remainder" series consisting of terms higher than first degree. Then, under the conditions of Theorem 2,

$$\exp(-G_2)\, \bar{z}_i = z_i + r(>1). \tag{2.36}$$

Proof: From (1.13) we have

$$\exp(-G_2)\, \bar{z}_i = \sum_{|\sigma|=1} a_i(\sigma) \exp(-G_2)\, z^\sigma + r(>1). \tag{2.37}$$

Since the first term on the right is of first order, we can also equivalently write

$$\exp(-G_2)\, \bar{z}_i = \sum_j M_{ij}(0) \exp(-G_2)\, z_j + r(>1), \tag{2.38}$$

but from (2.32) we conclude

$$\exp(-G_2)\, z_j = \sum_k (M^{-1})_{jk} z_k. \tag{2.39}$$

Combining (2.38) and (2.39) completes the proof.

Lemma 6: There exist polynomials g_3, g_4, etc., such that when $\exp(-G_3)$, $\exp(-G_4)$, etc., are consecutively applied to (2.36), the order of the remainder term can be made arbitrarily large.

Proof: We shall find g_3. The higher order g's are found in the same fashion. Let us decompose the remainder term $r(>1)$ in (2.36) into a second degree term $f_i(2;z)$ plus a higher order remainder $r(>2)$,

$$\exp(-G_2)\, \bar{z}_i = z_i + f_i(2;z) + r(>2). \tag{2.40}$$

Now form the Poisson bracket of (2.40) with the analogous expression having the index set equal to j. Using (1.10), (2.8), and (2.23) we find

$$J_{ij} = [z_i + f_i(2) + r(>2),\ z_j + f_j(2) + r(>2)],$$

or

$$J_{ij} = J_{ij} + [z_i, f_j(2)] + [f_i(2), z_j] + r(>1). \tag{2.41}$$

Equating like powers of z, we get

$$[z_i, f_j(2)] + [f_i(2), z_j] = 0. \tag{2.42}$$

It follows from Lemma 1 that there is a function g_3 such that

$$f_i(2;z) = G_3 z_i. \tag{2.43}$$

In fact, g_3 can be found explicitly from (2.21), and is clearly homogeneous of degree 3. Using (2.43), we rewrite (2.40) as

$$\exp(-G_2)\, \bar{z}_i = z_i + G_3 z_i + r(>2). \tag{2.44}$$

Finally, we apply $\exp(-G_3)$ to both sides of (2.44). The result is

$$\exp(-G_3) \exp(-G_2)\, \bar{z}_i = z_i + r(>2). \tag{2.45}$$

We have all the necessary machinery to complete the proof of Theorem 2. Comparing (2.36) and (2.45), we find that we have been able to raise the order of the remainder term by 1. As stated in the last lemma, it is easy to see that the process can be repeated at will. That is, there exist further homogeneous polynomials $g_4, g_5, \ldots, g_s$ such that

$$\exp(-G_s) \cdots \exp(-G_3) \exp(-G_2)\, \bar{z}_i = z_i + r(>s-1) \tag{2.46}$$

for any value of s. Inverting the left-hand side of (2.46), we obtain the result

$$\bar{z}_i = \exp(G_2) \cdots \exp(G_s)\, z_i + r(>s-1). \tag{2.47}$$

Now let $s \to \infty$. Then, if the remainder tends to zero, we obtain the advertised result (2.14). Otherwise the result is true only formally. In the latter case the infinite product is also divergent.

We close this section with the remark that it is often more convenient to have a product representation, usually with different G's, in the opposite order,

$$\bar{z}_i = \exp(G_s'') \cdots \exp(G_3'') \exp(G_2)\, z_i + r(>s-1). \tag{2.48}$$

We will show in Sec. 4 that this is always possible providing (2.47) holds, and vice versa.

3. EXAMPLES

In this section we will apply the results of Sec. 2 to some maps studied previously in the literature by other authors.

The first examples are *Cremona* maps. They are symplectic maps for which the power series (1.13) terminates.[16] The simplest nontrivial Cremona map terminates at the second power. In the easiest case of two dimensions, where symplectic maps are merely area preserving maps, a suitable linear transformation brings the quadratic Cremona map into the form[17]

$$\bar{z}_1 = \lambda[z_1 + (z_1 - z_2)^2], \quad \bar{z}_2 = \lambda^{-1}[z_2 + (z_1 - z_2)^2] \tag{3.1}$$

if the matrix $M(0)$ for the original map has the real distinct positive eigenvalues λ and λ^{-1}. (Note that for a symplectic matrix M, one always has $\det M = 1$, and hence the eigenvalues must be reciprocals in the 2×2 case.) If the eigenvalues are $\exp(\pm i\alpha)$, i.e., if they lie on the unit circle, then the quadratic Cremona map can be brought to the form[17–19]

$$\begin{aligned} \bar{z}_1 &= z_1 \cos\alpha + z_2 \sin\alpha + z_2^2 \cos\alpha, \\ \bar{z}_2 &= -z_1 \sin\alpha + z_2 \cos\alpha - z_2^2 \sin\alpha. \end{aligned} \tag{3.2}$$

There are also other possibilities for the eigenvalues of $M(0)$ which are of less interest.

Let us apply our formalism. In the case (3.1) we have

$$M(0) = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix} = \exp \begin{pmatrix} \log\lambda & 0 \\ 0 & -\log\lambda \end{pmatrix}. \tag{3.3}$$

Therefore, using (2.33) and (2.34) we find

$$g_2 = -(\log\lambda) z_1 z_2. \tag{3.4}$$

Correspondingly, we have for G_2 the expression

$$G_2 = (\log\lambda)\left(z_1 \frac{\partial}{\partial z_1} - z_2 \frac{\partial}{\partial z_2}\right). \tag{3.5}$$

Next we compute the $f_i(2;z)$ following (2.40). We find

$$\begin{aligned} \exp(-G_2)\bar{z}_1 &= \exp(-G_2)\{\lambda[z_1 + (z_1 - z_2)^2]\} \\ &= \lambda \exp(-G_2) z_1 + \lambda[\exp(-G_2) z_1 - \exp(-G_2) z_2]^2 \\ &= z_1 + \lambda[\lambda^{-1} z_1 - \lambda z_2]^2. \end{aligned} \tag{3.6}$$

Thus, we get for $f_1(2;z)$ the expression

$$f_1(2;z) = \lambda^{-1} z_1^2 - 2\lambda z_1 z_2 + \lambda^3 z_2^2. \tag{3.7}$$

Similarly, we find for $f_2(2;z)$ the result

$$f_2(2;z) = \lambda^{-3} z_1^2 - 2\lambda^{-1} z_1 z_2 + \lambda z_2^2. \tag{3.8}$$

We are ready to apply (2.21) to find g_3. The line integral is most easily evaluated along the path $z_i' = \tau z_i$ with the parameter τ ranging from zero to one. If the f_i are homogeneous polynomials, the integral can be evaluated immediately in the general case to give

$$g_{s+1}(z) = -(s+1)^{-1} \sum_{i,j} f_i(s;z) J_{ij} z_j. \tag{3.9}$$

In particular, for g_3 we find the result

$$g_3(z) = (\lambda^{-1} z_1 - \lambda z_2)^3/3. \tag{3.10}$$

It follows that

$$G_3 = (\lambda^{-1} z_1 - \lambda z_2)^2 \left(\lambda^{-1} \frac{\partial}{\partial z_2} + \lambda \frac{\partial}{\partial z_1}\right). \tag{3.11}$$

We must now continue on to compute the higher order remainder terms following (2.45). The calculation is simplified by the observation that

$$G_3^2 z_i = 0 \tag{3.12}$$

and hence

$$\exp(-G_3) z_1 = z_1 - \lambda(\lambda^{-1} z_1 - \lambda z_2)^2, \tag{3.13a}$$

$$\exp(-G_3) z_2 = z_2 - \lambda^{-1}(\lambda^{-1} z_1 - \lambda z_2)^2. \tag{3.13b}$$

Also, we have

$$G_3(\lambda^{-1} z_1 - \lambda z_2)^2 = 0 \tag{3.14}$$

and hence

$$\exp(-G_3)(\lambda^{-1} z_1 - \lambda z_2)^2 = (\lambda^{-1} z_1 - \lambda z_2)^2. \tag{3.15}$$

We are ready. We find, using (3.6), (3.13), and (3.15) the result

$$\exp(-G_3)\exp(-G_2)\bar{z}_1 = \exp(-G_3)[z_1 + \lambda(\lambda^{-1} z_1 - \lambda z_2)^2] = z_1. \tag{3.16}$$

That is, the remainder term vanishes! The same is true for $\bar{z}_2$. Thus in this case, the higher order Lie operators G_4, G_5, etc., are all zero. We conclude that for the two-dimensional quadratic Cremona map in the case (3.1) we have

$$\exp(-G_3)\exp(-G_2)\bar{z}_i = z_i, \tag{3.17}$$

and hence

$$\bar{z}_i = \exp(G_2)\exp(G_3) z_i. \tag{3.18}$$

The calculation in the case (3.2) can also be carried out with equal ease. The result is

$$g_2(z) = -(\alpha/2)(z_1^2 + z_2^2), \tag{3.19}$$

$$G_2 = -\alpha\left(z_1 \frac{\partial}{\partial z_2} - z_2 \frac{\partial}{\partial z_1}\right), \tag{3.20}$$

$$g_3(z) = -(z_1 \sin\alpha + z_2 \cos\alpha)^3/3, \tag{3.21}$$

$$G_3 = -(z_1 \sin\alpha + z_2 \cos\alpha)^2 \left(\sin\alpha \frac{\partial}{\partial z_2} - \cos\alpha \frac{\partial}{\partial z_1}\right). \tag{3.22}$$

The higher order Lie operators again vanish, and Eq. (3.18) is exact.

Another symplectic map which has received considerable study is the ninth order Cremona map in two dimensions given implicitly by the relations[20,21]

$$\bar{z}_1 = z_1 + a z_2 - a z_2^3, \quad \bar{z}_2 = z_2 - a\bar{z}_1 + a\bar{z}_1^3 \tag{3.23}$$

and explicitly by

$$\begin{aligned} \bar{z}_1 &= z_1 + a z_2 - a z_2^3, \\ \bar{z}_2 &= z_2 - a(z_1 + a z_2 - a z_2^3) + a(z_1 + a z_2 - a z_2^3)^3. \end{aligned} \tag{3.24}$$

Here a is a parameter. Due to algebraic complications, we have not attempted to express this map in the form (2.14) although we have verified that $M(0)$ does lie on a continuous one parameter subgroup connected to the identity providing a is small enough. This task seems better suited to digital computers programmed to perform algebraic operations. However, we have discovered that the map can be written in the form

$$\bar{z}_i = \exp(F_2)\exp(F_4)\exp(G_2)\exp(G_4) z_i \tag{3.25}$$

with

$$\begin{aligned} f_2 &= -(a/2) z_2^2, \quad f_4 = (a/4) z_2^4, \\ g_2 &= -(a/2) z_1^2, \quad g_4 = (a/4) z_1^4. \end{aligned} \tag{3.26}$$

We will see in the next section that, should it be desirable, there are standard Lie algebraic manipulations which can be used to bring (3.25) into the form (2.14).

As mentioned by Moser,[22] two-dimensional Cremona mappings can be expressed as repeated products of linear transformations and shear mappings of the form

$$\bar{z}_1 = z_1 + h(z_2), \quad \bar{z}_2 = z_2. \tag{3.27}$$

The shear mapping (3.27) can be expressed as a Lie series $\exp(F)$, where F is the Lie operator associated with the function f given by

$$f(z_1, z_2) = -\int^{z_2} h(z')\,dz'. \tag{3.28}$$

Also, linear maps connected to the identity can be expressed as Lie transformations as seen earlier. It follows that quite generally Cremona maps can be expressed as products of Lie series. This factorization may be distinct from that of (2.14) since the functions given by (3.28) need not be homogeneous polynomials.

In addition, some area preserving maps which are more general than Cremona maps can be factored in a similar way. For example, the mapping $\mathcal{T}$ given by[23]

$$\mathcal{T}: \begin{cases} \bar{z}_1 = z_1 + z_2, \\ \bar{z}_2 = z_1 - \epsilon \sin(z_1 + z_2) \end{cases} \tag{3.29}$$

can be expressed as a product of two shear mappings. We have

$$\mathcal{T} = \mathcal{RS}, \tag{3.30}$$

where $\mathcal{R}$ and $\mathcal{S}$ denote the mappings

$$\mathcal{R}: \begin{cases} \bar{z}_1 = z_1, \\ \bar{z}_2 = \bar{z}_2 - \epsilon \sin z_1, \end{cases} \tag{3.31}$$

$$\mathcal{S}: \begin{cases} \bar{z}_1 = z_1 + z_2, \\ \bar{z}_2 = z_2. \end{cases} \tag{3.32}$$

The mapping $\mathcal{T}$ therefore has the representation

$$\bar{z}_i = \exp(F)\exp(G)\, z_i, \tag{3.33}$$

where F and G are the Lie operators corresponding to the functions f and g given by

$$f(z_1, z_2) = \epsilon \cos z_1, \quad g(z_1, z_2) = -z_2^2/2. \tag{3.34}$$

Similar results hold for the mapping

$$\bar{z}_1 = z_1 + \epsilon \sin z_2, \quad \bar{z}_2 = z_1 + z_2 + \epsilon \sin z_2 \tag{3.35}$$

studied by Froeschlé.[18] It can be written in the product form $\mathcal{SR}$ provided the roles of z_1 and z_2 are interchanged.

Finally, we close this section with a brief study of the four-dimensional map given by[24]

$$\begin{aligned} \bar{z}_1 &= z_1 + a_1 \sin(z_1 + z_3) + b \sin(z_1 + z_2 + z_3 + z_4), \\ \bar{z}_2 &= z_2 + a_2 \sin(z_2 + z_4) + b \sin(z_1 + z_2 + z_3 + z_4), \\ \bar{z}_3 &= z_1 + z_3, \\ \bar{z}_4 &= z_2 + z_4. \end{aligned} \tag{3.36}$$

A routine calculation shows that $M(z)$ given by (1.11) is indeed a symplectic matrix so that (3.36) is a symplectic map. In particular, $M(0)$ is given by the matrix

$$M(0) = \begin{pmatrix} (1 + a_1 + b) & b & (a_1 + b) & b \\ b & (1 + a_2 + b) & b & (a_2 + b) \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}. \tag{3.37}$$

We next inquire whether $M(0)$ lies on a continuous one parameter subgroup connected to the identity. We have not been able to treat the general case. However, we have been able to verify this condition if a_1, a_2, and b are sufficiently small. Details are given in Appendix A. Therefore Theorem 2 applies, and with sufficient effort the polynomials g_2, g_3, etc., can be computed.

4. LIE ALGEBRAIC TOOLS

The content of Theorem 2 is that under rather general conditions a symplectic map can be written as a product of Lie transformations. Earlier, from Eqs. (2.2)—(2.5), we found that the Lie operators which generate Lie transformations also form a Lie algebra under commutation. The purpose of this section is to review some Lie algebraic tools which will enable us to manipulate the various Lie operators appearing in products of Lie transformations.

We begin by introducing yet another Lie algebra, the *adjoint* Lie algebra.[14] Let F be a given Lie operator and E an arbitrary Lie operator. We associate with F an operator $\hat{F}$ (which acts on Lie operators) by the rule

$$\hat{F}E = \{F, E\}. \tag{4.1}$$

Here, as before, $\{\ ,\ \}$ denotes commutation. Next, let F and G be any two Lie operators. We define a Lie operator H by the rule

$$H = \{F, G\}. \tag{4.2}$$

Then we find

$$\begin{aligned} \{\hat{F}, \hat{G}\}E &= (\hat{F}\hat{G} - \hat{G}\hat{F})E \\ &= \{F, \{G, E\}\} - \{G, \{F, E\}\} \\ &= \{\{F, G\}, E\} = \hat{H}E. \end{aligned} \tag{4.3}$$

Here we have used the Jacobi identity

$$\{E, \{F, G\}\} + \{F, \{G, E\}\} + \{G, \{E, F\}\} = 0 \tag{4.4}$$

which always holds for commutators. Since E is arbitrary, we may rewrite (4.3) as

$$\hat{H} = \{\hat{F}, \hat{G}\}. \tag{4.5}$$

Comparison of (4.5) and (4.2) shows that the adjoint Lie algebra is homomorphic to the parent Lie algebra of Lie operators.

Our discussion should have a familiar ring. It parallels, in fact, the discussion surrounding Eqs. (2.1)—(2.5). Reviewing these equations, we see that the commutator Lie algebra of Lie operators is actually the adjoint Lie algebra of the underlying Poisson bracket Lie algebra. And, consequently, the "adjoint" we have been discussing is really the "adjoint-adjoint" of the basic Poisson bracket Lie algebra.

We now have the necessary notation to state the simplest theorem about the rearrangement of Lie transformations,

Theorem 3: Let A and B be Lie operators. Then

$$[\exp(A)]\,B[\exp(-A)] = (\exp\hat{A})\,B \tag{4.6}$$

and

$$\exp(A)\exp(B)\exp(-A) = \exp[(\exp\hat{A})\,B]. \tag{4.7}$$

Proof: Let τ be a parameter, and define $C(\tau)$ by the equation

$$C(\tau) = [\exp(\tau A)]\,B[\exp(-\tau A)]. \tag{4.8}$$

Then we have the relation

$$C(0) = B. \tag{4.9}$$

Further, we find by differentiation

$$\frac{dC}{d\tau} = AC - CA = \hat{A}C. \tag{4.10}$$

The solution to this differential equation with the boundary condition (4.9) is

$$C(\tau) = \exp(\tau\hat{A})\,B. \tag{4.11}$$

Now set $\tau = 1$ to obtain (4.6). We next observe that for any two operators B and C we have

$$\hat{A}(BC) = \{A, BC\} = \{A, B\}C + B\{A, C\} = (\hat{A}B)\,C + B\hat{A}C. \tag{4.12}$$

It follows that $\hat{A}$ acts as a derivation on products, and in analogy to (2.7) and (2.9) we obtain

$$(\exp\hat{A})(BC) = ((\exp\hat{A})\,B)((\exp\hat{A})\,C). \tag{4.13}$$

This result is consistent with (4.6) and the observation that

$$[\exp(A)](BC)[\exp(-A)] = [\exp(A)]\,B[\exp(-A)][\exp(A)]\,C[\exp(-A)]. \tag{4.14}$$

We conclude that

$$[\exp(A)]\,B^n[\exp(-A)] = [(\exp\hat{A})\,B]^n \tag{4.15}$$

for any power n. The desired result (4.7) now follows directly term by term.

As an application of Theorem 3, let us consider the product $\exp(G_2)\exp(G_3)$ which occurs in Eq. (3.18). We have

$$\exp(G_2)\exp(G_3) = \exp(G_2)\exp(G_3)\exp(-G_2)\exp(G_2) = \exp[(\exp\hat{G}_2)\,G_3]\exp(G_2). \tag{4.16}$$

Let us define G_3' by the expression

$$G_3' = (\exp\hat{G}_2)\,G_3. \tag{4.17}$$

Is there a polynomial g_3' which has G_3' as its associated Lie operator? We know there must be from the remarks following Eq. (2.5). Using the homomorphisms between the Lie algebras involved, we obtain

$$g_3' = (\exp G_2)\,g_3. \tag{4.18}$$

Consequently, for the first example of Sec. 3 we calculate that

$$g_3' = (z_1 - z_2)^3/3 \tag{4.19}$$

and hence

$$G_3' = (z_1 - z_2)^2\left(\frac{\partial}{\partial z_2} + \frac{\partial}{\partial z_1}\right). \tag{4.20}$$

We have shown that

$$\exp(G_2)\exp(G_3) = \exp(G_3')\exp(G_2) \tag{4.21}$$

with G_3' given by (4.20).

We are now ready to move on to a far deeper result generally known as the Campbell–Baker–Hausdorff (CBH) formula. In its usual mathematical setting it provides the connecting link between Lie algebras and Lie groups.[14] We will use it to reexpress the product of two Lie transformations as a single Lie transformation, or more generally as a method of combining exponents.

Theorem 4: Let A and B be any two operators, and let α and β be parameters. Then we can formally write

$$\exp(\alpha A)\exp(\beta B) = \exp(C) \tag{4.22}$$

with

$$C = \alpha A + \beta B + (\alpha\beta/2)\{A, B\} + (\alpha^2\beta/12)\{A, \{A, B\}\} + (\alpha\beta^2/12)\{B, \{B, A\}\} + \cdots. \tag{4.23}$$

The series for C may or may not converge depending on the properties of αA and βB. The really remarkable fact is that the right-hand side of (4.23) involves *only* Lie products. Thus, all we need to evaluate (4.23) is a knowledge of the Lie algebra generated by A and B, and we are guaranteed that C is contained in this Lie algebra. The general form of all the coefficients in the series is not known.[25] However, the series can be formally summed to all orders in α and the first few orders in β. Through first order in β we have

$$C = \alpha A + \beta\alpha\hat{A}[1 - \exp(-\alpha\hat{A})]^{-1}B + O(\beta^2). \tag{4.24}$$

A proof of these results and an expression for the quadratic term in β are given in Appendix B.

As a simple example of the use of the CBH formula, we will derive (2.48) starting from (2.47). A more complicated example of its use will be given in the next section. Beginning with (2.47), we write

$$\exp(G_2)\cdots\exp(G_s) = \exp(G_2)\cdots\exp(G_s)\exp(-G_2)\exp(G_2) = \exp(G_3')\cdots\exp(G_s')\exp(G_2), \tag{4.25}$$

where

$$G_\tau' = \exp(\hat{G}_2)\,G_\tau. \tag{4.26}$$

Note that as in our earlier example, G_τ' will be the Lie operator associated with the function g_τ' given by

$$g_\tau' = (\exp G_2)\,g_\tau, \tag{4.27}$$

and that the degree of g_τ' is as indicated because of (2.23). Next we repeatedly use the CBH formula (4.23) to combine the various operators G_τ' to obtain an expression of the form

$$\exp(G_3')\cdots\exp(G_s') = \exp(H_3 + \cdots + H_s + \cdots). \tag{4.28}$$

Observe that because of (2.23), only a finite number of terms in the series (4.23) are required in the calculation of each H_r. Again using the CBH formula, we may write

$$\exp(H_3 + \cdots) = \exp(H_3 + \cdots)\exp(-H_3)\exp(H_3)$$
$$= \exp(H_4' + \cdots)\exp(H_3). \tag{4.29}$$

This process can be repeated again and again to get

$$\exp(H_3 + \cdots) = \exp(H_{s+1}'' + \cdots)\exp(H_s')\cdots\exp(H_3). \tag{4.30}$$

Combining (4.25), (4.28), and (4.29), we find

$$\exp(G_2)\cdots\exp(G_s)\,z_i$$
$$= \exp(H_s')\cdots\exp(H_3)\exp(G_2)\,z_i + r(> s-1). \tag{4.31}$$

Consequently, an expression of the form (2.47) implies an expression of the form (2.48). The converse can be proven analogously.

5. CONSTRUCTION OF INVARIANT FUNCTIONS

In the study of a symplectic transformation $\mathcal{T}$ arising from either a Poincaré surface of section or from following the field lines in a toroidal plasma device, one is interested in studying the result of applying the map many times in succession. That is, we are interested in studying $\mathcal{T}^n z$ for large n. This study is simplified if one can construct invariant functions f with the property

$$f(\bar{z}) = f(z), \tag{5.1}$$

where

$$\bar{z} = \mathcal{T} z. \tag{5.2}$$

For if such functions can be found, one knows that the points generated by $\mathcal{T}^n z$, for various n must all lie on a surface of constant f. The more invariant functions one can find, the more one can say about the map and its powers. The situation is quite analogous to the role played by integrals of motion for Hamiltonian systems or magnetic surfaces in a toroidal plasma. We will see shortly that the analogy is more than coincidental.

The problem of constructing invariant functions in the case of symplectic transformations in *two* variables was first considered in detail by Birkhoff.[26] We shall begin this section by proving some simple lemmas which will enable us to restate his result in our language. We will then show how the same results can be obtained *for any number of variables* from the CBH formula.

Lemma 7: Consider a one parameter family of symplectic maps. That is, we write

$$\bar{z}_i(s) = g_i(z, s) \tag{5.3}$$

with the understanding that the new variables $\bar{z}(s)$ are symplectically related to the original variables z for every value of the parameter s. Then there exists a function h, which we shall call the *generating function*, such that

$$\left.\frac{\partial \bar{z}_i(s)}{\partial s}\right|_z = [h(\bar{z}, s), \bar{z}_i]. \tag{5.4}$$

Proof: Since the functions g_i are viewed as given, we have by direct calculation

$$\frac{\partial \bar{z}_i(s)}{\partial s} = \frac{\partial g_i(z, s)}{\partial s}. \tag{5.5}$$

Next invert the transformation (5.3) to solve for the z's in terms of $\bar{z}$'s, and substitute this result into the right-hand side of (5.5) to obtain expressions of the form

$$\frac{\partial \bar{z}_i(s)}{\partial s} = f_i(\bar{z}, s). \tag{5.6}$$

From Taylor's theorem we have

$$\bar{z}_i(s+\epsilon) = \bar{z}_i(s) + \epsilon f_i(\bar{z}, s) + O(\epsilon^2). \tag{5.7}$$

Take the Poisson bracket of (5.7) with the analogous expression having the index set equal to j. The result is

$$[\bar{z}_i(s+\epsilon), \bar{z}_j(s+\epsilon)] = [\bar{z}_i(s), \bar{z}_j(s)] + \epsilon\{[\bar{z}_i(s), f_j(\bar{z}, s)] + [f_i(\bar{z}, s), \bar{z}_j(s)]\} + O(\epsilon). \tag{5.8}$$

Using the first part of (1.10) and equating powers of ϵ, we find

$$[\bar{z}_i, f_j(\bar{z}, s)]_z + [\bar{z}_j, f_i(\bar{z}, s)]_z = 0. \tag{5.9}$$

Here we have written the subscript z to emphasize that the Poisson bracket is taken with respect to the variables z. However, as is well known, the Poisson bracket can also be taken with respect to the variables $\bar{z}$. For let u and v be any two functions. Then by the chain rule and (1.12) we have

$$[u(\bar{z}), v(\bar{z})]_z = \sum_{ij}\left(\frac{\partial u}{\partial z_i}\right) J_{ij}\left(\frac{\partial v}{\partial z_i}\right)$$
$$= \sum_{ijkl}\left(\frac{\partial u}{\partial \bar{z}_k}\right)\left(\frac{\partial \bar{z}_k}{\partial z_i}\right) J_{ij}\left(\frac{\partial v}{\partial \bar{z}_l}\right)\left(\frac{\partial \bar{z}_l}{\partial z_j}\right)$$
$$= \sum_{k,l}\left(\frac{\partial u}{\partial \bar{z}_i}\right) M_{ki} J_{ij} \tilde{M}_{jl}\left(\frac{\partial v}{\partial \bar{z}_l}\right)$$
$$= \sum_{k,l}\left(\frac{\partial u}{\partial \bar{z}_k}\right) J_{kl}\left(\frac{\partial v}{\partial \bar{z}_l}\right) = [u(\bar{z}), v(\bar{z})]_{\bar{z}}. \tag{5.10}$$

Thus we may also write (5.9) in the form

$$[\bar{z}_i, f_j(\bar{z}, s)]_{\bar{z}} + [\bar{z}_j, f_i(\bar{z}, s)]_{\bar{z}} = 0. \tag{5.11}$$

The existence of the advertised generating function $h(\bar{z}, s)$ now follows from Lemma 1.

Lemma 8: Suppose the one parameter family in Lemma 7 is also a one parameter *group*. Without loss of generality, we may assume that the parameterization is selected in such a way that it satisfies

$$\bar{z}_i(0) = z_i \tag{5.12}$$

and is additive,

$$\bar{z}_i(s_1 + s_2) = g_i(\bar{z}(s_1), s_2). \tag{5.13}$$

Then the generating function h is *independent* of s.

Proof: Partially differentiate (5.13) with respect to s_2 and then set s_2 equal to zero. The result is a relation of the form (5.6) with

$$f_i(\bar{z}) = \left.\frac{\partial g_i}{\partial s_2}\right|_{s_2=0}. \tag{5.14}$$

Note, however, that in this case f_i is independent of s. It follows from the remainder of Lemma 7 that the generating function h is independent of s.

Lemma 9: If the generating function is independent of s, the differential equation

$$\frac{\partial \bar{z}_i(s)}{\partial s} = [h(\bar{z}), \bar{z}_i] \tag{5.15}$$

with the initial condition (5.12) has the unique solution

$$\bar{z}_i(s) = \exp(sH)\, z_i\,. \tag{5.16}$$

We note that apart from a sign, (5.15) is analogous to Hamilton's equations of motion (1.3).

Proof: Evidently (5.12) is satisfied. Now differentiate (5.16) with respect to s to get

$$\frac{\partial \bar{z}_i}{\partial s} = \exp(sH)\, Hz_i = \exp(sH)[h(z), z_i]$$
$$= [\exp(sH)\, h(z),\ \exp(sH)\, z_i]. \tag{5.17}$$

Here we have used (2.8). Also, from (2.9) and (5.16) it follows that

$$\exp(sH)\, z^\sigma = \bar{z}(s)^\sigma. \tag{5.18}$$

Consequently, since polynomials are dense in the set of functions, we must have

$$\exp(sH)\, u(z) = u(\bar{z}) \tag{5.19}$$

for any function u. Employing (5.16) and (5.19) in (5.17), we see that the differential equation (5.15) is satisfied.

Lemma 10: The function $h(z)$ is an invariant function for the transformation (5.16).

Proof: By (5.19) we have

$$h(\bar{z}) = \exp(sH)\, h(z) = h(z) + s[h, h] + \cdots = h(z), \tag{5.20}$$

since all the Poisson brackets are zero.

We are now ready to appreciate the result of Birkhoff, which we summarize in the next theorem.

Theorem 5: Denote by $\bar{z}_i(k)$ the result of applying the transformation (1.13) k times in succession. We also adopt the convention (5.12). Evidently $\bar{z}_i(k)$ has an expansion similar to (1.13). We write

$$\bar{z}_i(k) = \sum_{|\sigma|>0} a_i(k, \sigma)\, z^\sigma \tag{5.21}$$

with the explicit recognition that the coefficients a_i will depend on k. Then, in the case of two variables and provided the eigenvalues of $M(0)$ satisfy certain conditions, the dependence of the coefficients a_i on k can be *extended* from *integer* values to *all* real values by analytic interpolation in such a way that the series (5.21) "behaves" as a one parameter group, as in Lemma 8, with k playing the role of a continuous parameter. We use the word "behaves" advisedly, because the series may not be convergent for nonintegral k even though $a_i(k, \sigma)$ is well defined. That is, the group property may hold only as a formal relation between power series.

The direct verification of Birkhoff's theorem is beyond the purpose of this paper. However, we point out that his theorem, with the aid of Lemmas 9 and 10, produces a formal power series for an invariant function h. If the series is convergent, it yields a true invariant function; otherwise the result is a series which formally satisfies (5.20) term by term.

We will now show that the same results can be derived almost immediately for the case of any number of variables with the aid of the CBH formula. We assume the conditions of Theorem 3 are satisfied, and begin by repeatedly applying the CBH formula to the terms $\exp(G_3)\exp(G_4)\exp(G_5)\cdots$ appearing in (2.14). In view of (2.23), we can combine the exponents into one grand exponent to find

$$\exp(G_3)\exp(G_4)\cdots = \exp(G_3' + G_4' + \cdots), \tag{5.22}$$

and each term G_r' can be written as a sum of a finite number of commutators. Next we try to combine the result (5.22) with $\exp(G_2)$ to find

$$\exp(G_2)\exp(G_3)\exp(G_4)\cdots = \exp(G_2)\exp(G_3' + G_4' + \cdots)$$
$$= \exp(H_2 + H_3 + H_4 + \cdots). \tag{5.23}$$

This last step is somewhat more problematical since, because of (2.23), we must now sum infinite series involving arbitrarily many commutators of G_2 to find each of the terms $H_3, H_4, \cdots$. We will study this matter somewhat further in a moment. Assuming that the various series converge, we can formally write

$$\bar{z}_i = \exp(H)\, z_i \tag{5.24}$$

with

$$H = H_2 + H_3 + H_4 + \cdots. \tag{5.25}$$

Furthermore, because of the relations (2.3) and (2.5), we know that there must be functions $h_2, h_3, \cdots$ corresponding to the operators H_2, H_3, etc. Thus, H is a Lie operator corresponding to h given by

$$h = h_2 + h_3 + h_4 + \cdots. \tag{5.26}$$

Finally, from Lemma 10 (with $s = 1$) we conclude that the function $h(z)$ constructed in this manner will be an invariant function of the transformation (2.14).

It is easy to verify that the invariant function h we have obtained from the CBH formula is the same as would be found by Birkhoff's method. Let us apply the transformation (5.24) twice in succession. We write

$$\bar{z}_i = \exp(H_z)\, z_i, \tag{5.27a}$$

$$\bar{\bar{z}}_i = \exp(H_{\bar{z}})\, \bar{z}_i, \tag{5.27b}$$

and use subscripts to indicate exactly which variables occur in the various Poisson brackets. Expanding (5.27b) we can write

$$\bar{\bar{z}}_i = \exp(H_{\bar{z}})\, \bar{z}_i = \bar{z}_i + [h(\bar{z}), \bar{z}_i]_{\bar{z}} + \cdots$$
$$= \bar{z}_i + [h(\bar{z}), \bar{z}_i]_z + \cdots = \bar{z}_i + [h(z), \bar{z}_i]_z + \cdots$$
$$= \exp(H_z)\, \bar{z}_i\,. \tag{5.28}$$

Here we have used (5.10) and the fact that h is an invariant function. Now substitute (5.28) into (5.27a) to get

$$\bar{\bar{z}}_i = \exp(H_z)\exp(H_z)\, z_i = \exp(2H_z)\, z_i. \tag{5.29}$$

Employing the notation of Theorem 5, it is clear that (5.29) generalizes to

$$\bar{z}_i(k) = \exp(kH)\, z_i. \tag{5.30}$$

Finally, (5.30) can be extended from integer values to all real values simply by replacing k with s to give (5.16).

There is an interpretation of the result (5.30) which is worth emphasizing: We have already remarked that (5.15) is analogous to Hamilton's equations of motion. We now see that if Birkhoff's theorem, or equivalently the use of the CBH formula, is applicable, then the transformation (1.13) can be viewed as the result of integrating Hamilton's equations of motion for the time independent Hamiltonian $(-h)$ from the initial "time" $s=0$ to the "time" $s=1$. Subsequent iterations of the map are obtained by integrating on to successive integer values.

So far, we have not discussed the convergence of the various procedures we have employed. This question is very difficult, and much remains to be learned. A theorem of Moser[27] can be used in the simplest case of two variables z_1, z_2 if $M(0)$ can be brought to the form (3.3). In our language, he shows in this case by indirect methods that if $\mathcal{T}$ is the symplectic transformation in question, then there exists another symplectic transformation $\mathcal{U}$ of the form

$$\mathcal{U} = \exp(F_3)\exp(F_4)\exp(F_5)\cdots \tag{5.31}$$

such that

$$\mathcal{U}^{-1}\mathcal{T}\mathcal{U} = \exp\left(\sum_1^\infty \alpha_n G_{2n}\right) \tag{5.32}$$

with

$$g_2 = z_1 z_2, \quad g_{2n} = (g_2)^n, \quad \alpha_1 = -\log\lambda. \tag{5.33}$$

Both the infinite product (5.31) and the infinite series in (5.32) converge. By undoing the transformation $\mathcal{U}$, one finds the desired result

$$\mathcal{T} = \exp H \tag{5.34}$$

with

$$H = \mathcal{U}\sum_1^\infty \alpha_n G_{2n}\mathcal{U}^{-1}. \tag{5.35}$$

Thus, there are nontrivial classes of problems for which our methods (and also Birkhoff's) succeed.

By contrast, a theorem of Moser's on Cremona maps[22] can be used to infer that the CBH series diverges for the example (3.2). The method of proof is again indirect. However, direct examination of the CBH series shows that it repeatedly contains terms of the form

$$[1-\exp(-\hat{G}_2)]^{-1}F_n, \tag{5.36}$$

and we will see that these terms can cause problems.

Rather than examining (5.36), it is convenient to use the homomorphism between Lie algebras and their adjoints to work instead with the expression

$$[1-\exp(-G_2)]^{-1}f_n. \tag{5.37}$$

We next observe from (2.23) that G_2 maps $\mathcal{P}_n$ into itself, and hence the action of G_2 on each $\mathcal{P}_n$ can be represented by a matrix. Let v_1 be a polynomial of first degree which is an eigenvector of G_2. We write the eigenvalue as $(-\log\lambda)$ so that we have

$$G_2 v_1 = (-\log\lambda)\, v_1 \tag{5.38}$$

and

$$\exp(-G_2)\, v_1 = \lambda v_1. \tag{5.39}$$

Now suppose f_n is a polynomial of the form

$$f_n = (v_1)^n. \tag{5.40}$$

Then we find

$$\exp(-G_2)f_n = \exp(-G_2)(v_1)^n = \lambda^n (v_1)^n = \lambda^n f_n. \tag{5.41}$$

Here we have used (2.7). In this case we find for (5.37) the result

$$[1-\exp(-G_2)]^{-1}f_n = (1-\lambda^n)^{-1}f_n. \tag{5.42}$$

Suppose λ lies on the unit circle as it does for the example (3.2). Then the expression $(1-\lambda^n)^{-1}$ either is infinite for some n [if $\alpha = i\log\lambda$ is a rational multiple of 2π], or becomes arbitrarily large with increasing n [if α is an irrational multiple of 2π]. What we are observing here is a manifestation of the classic problem of "small denominators" which has been known to celestial mechanicians in connection both with perturbation theory and mapping problems since the time of Poincaré.[28,29] We see that it may also occur in the Lie algebraic approach in such a way as to spoil the convergence of the CBH series, and that this problem can potentially occur if any of the eigenvalues of $M(0)$ lie on the unit circle.

There is one last topic we wish to discuss. We have used the CBH series to obtain an invariant function h. In the case of sympletic transformations in two variables, a single invariant function suffices to characterize the map, and all other invariant functions are simple functions of h. However, in the case of four or more variables, e.g., (3.27), there may be additional invariant functions beyond h.

In view of (5.19), f will be an invariant function if it satisfies the relation

$$Hf = [h, f] = 0. \tag{5.43}$$

Consequently, the problem of finding further invariant functions is equivalent to the classical mechanics problem of finding integrals of motion for a system having $(-h)$ as a Hamiltonian. By analogy to classical mechanics, we expect to be able to find at most $2n-1$ functionally independent invariant functions including h itself.

There is as yet no fully developed algorithm for finding integrals of motion for any specified Hamiltonian h. However, there is a germ for such an algorithm in Birkhoff's procedure of attempting to bring Hamiltonians to a normal form.[4,5] In our notation, one attempts in this procedure to find polynomials g_3, g_4, etc., such that the Hamiltonian h' given by

$$h' = \cdots\exp(G_4)\exp(G_3)h \tag{5.44}$$

has a particularly simple form. If the form is simple enough, one can read off the integrals of motion direct-

ly. This method has been applied successfully by Gustavson and others[5,6] to the case of several variables provided h_2 has the form

$$h_2 = \sum_i \alpha_i z_i^2 \tag{5.45}$$

with all $\alpha_i > 0$. The case with some $\alpha_i = 0$ can also be treated.[17] An analysis which we intend to publish later shows that what is essential to this whole procedure is a detailed treatment of the range and null spaces of the operator H_2.

Now suppose that f' is an integral of h',

$$[h', f'] = 0. \tag{5.46}$$

We define f by the rule

$$f = \exp(-G_3)\exp(-G_4)\cdots f'. \tag{5.47}$$

Then we find

$$\begin{aligned}[h,f] &= [\exp(-G_3)\exp(-G_4)\cdots h', \exp(-G_3)\exp(-G_4)\cdots f'] \\ &= \exp(-G_3)\exp(-G_4)\cdots [h', f'] = 0,\end{aligned} \tag{5.48}$$

which shows that f is an integral. Usually f' can be taken to be a second degree polynomial. However the series f given by (5.47) will generally contain an infinite number of terms and may not converge. In the latter case, (5.43) is only satisfied term by term, and f is only a formal series. We expect the case of divergent series to be the most common. This is because if the series were to converge, it would produce an analytic global integral for the Hamiltonian h. However, most Hamiltonians do not possess global analytic integrals.[7,29]

6. CONCLUDING SUMMARY

In Sec. 2 it was shown that the Lie transformation associated with an analytic function produces an analytic symplectic map, and that conversely, under certain general conditions, an analytic symplectic map can be written as a product of Lie transformations. Section 3 treated several examples of analytic symplectic maps that had been studied previously by other authors. The discussion then turned in Sec. 4 to a further development of Lie algebraic tools and culminated with the Campbell–Baker–Haudsdorff formula. Next, after some preliminary background work, it was shown in Sec. 5 that the CBH formula can be used to formally combine a product of Lie transformations into a single Lie transformation, and that in so doing one obtains a generalization of Birkhoff's theorem for the construction of invariant functions. Thus, the existence of invariant functions is intimately related to the convergence of the CBH formula, and vice versa. Finally, in the case of symplectic maps involving more than two variables, the construction of additional invariant functions was shown to be analogous to the construction of integrals of motion in Hamiltonian dynamics.

ACKNOWLEDGMENTS

The first author is indebted to the Institut des Hautes Etudes Scientifiques where this work was begun. He wishes to thank Professor N. Kuiper and Professor L. Michel for their fine hospitality.

APPENDIX A

The purpose of this appendix is to demonstrate that M as given by (3.37) lies on a one parameter subgroup connected to the identity. We begin by observing that M can be written as a product of two symplectic matrices N and R,

$$M = NR, \tag{A1}$$

where

$$N = \begin{pmatrix} 1 & 0 & a_1+b & b \\ 0 & 1 & b & a_2+b \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{A2}$$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}. \tag{A3}$$

Each of these matrices can be written in exponential form, and use of Lemma 3 reveals that they lie on one parameter symplectic subgroups continuously connected to the identity,

$$N = \exp\begin{pmatrix} 0 & Q \\ 0 & 0 \end{pmatrix}, \tag{A4}$$

$$R = \exp\begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix}, \tag{A5}$$

Here each block is a 2×2 matrix, and Q denotes the matrix

$$Q = \begin{pmatrix} a_1+b & b \\ b & a_2+b \end{pmatrix}. \tag{A6}$$

The next step is to try to combine the two exponents by using the CBH formula. For this purpose we note that the two exponents occurring in (A4) and (A5) can be written in the respective forms

$$\log N = Q \otimes (\sigma_1 + i\sigma_2)/2 \tag{A7}$$

$$\log R = I \otimes (\sigma_1 - i\sigma_2)/2. \tag{A8}$$

Here the symbols σ_j denote the Pauli matrices,[1] and "$\otimes$" indicates that we have taken a tensor product.[30] For example,

$$Q \otimes (\sigma_1 + i\sigma_2)/2 = Q \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & Q \\ 0 & 0 \end{pmatrix}. \tag{A9}$$

It is easily verified that the tensor product operation obeys the multiplication rules

$$(A \otimes \sigma_j)(B \otimes \sigma_k) = AB \otimes \sigma_j\sigma_k \tag{A10}$$

and the addition rules

$$(A+B) \otimes \sigma_j = A \otimes \sigma_j + B \otimes \sigma_j, \tag{A11}$$

$$A \otimes (\alpha\sigma_j + \beta\sigma_k) = \alpha A \otimes \sigma_j + \beta A \otimes \sigma_k. \tag{A12}$$

Consequently, from the CBH formula we conclude that $\log M$ must be given by an expression of the form

$$\log M = \log NR = f \otimes \sigma_1 + g \otimes \sigma_2 + h \otimes \sigma_3, \tag{A13}$$

where f, g, and h are power series in the matrix Q. This is because I and powers of Q all commute, and the Pauli matrices are closed under commutation. Furthermore f, g, and h must be the same series that occur in the expression

$$\exp[\alpha(\sigma_1 + i\sigma_2)/2]\exp[(\sigma_1 - i\sigma_2)/2]$$
$$= \exp[f(\alpha)\sigma_1 + g(\alpha)\sigma_2 + h(\alpha)\sigma_3], \tag{A14}$$

where α is a parameter.

It remains to be shown that the series f, g, and h converge for Q sufficiently small. Or equivalently, we must show that $f(\alpha)$, $g(\alpha)$, and $h(\alpha)$ are all analytic and have nonzero radii of convergence in the complex variable α. A short calculation for the group SL(2, C) gives the multiplication rule

$$\exp(\mathbf{n}_1 \cdot \boldsymbol{\sigma})\exp(\mathbf{n}_2 \cdot \boldsymbol{\sigma}) = \exp(\mathbf{n}_3 \cdot \boldsymbol{\sigma}), \tag{A15}$$

with $\mathbf{n}_3$ given by the formulas

$$\boldsymbol{\tau}_3 = (\boldsymbol{\tau}_1 + \boldsymbol{\tau}_2 + i\boldsymbol{\tau}_1 \times \boldsymbol{\tau}_2)/(1 + \boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2), \tag{A16}$$

$$\boldsymbol{\tau}_j = \mathbf{n}_j(\tanh\sqrt{\mathbf{n}_j \cdot \mathbf{n}_j}/\sqrt{\mathbf{n}_j \cdot \mathbf{n}_j})$$
$$= \mathbf{n}_j(1 - \tfrac{1}{3}\mathbf{n}_j \cdot \mathbf{n}_j + \cdots). \tag{A17}$$

For the case in question, Eq. (A14), we have

$$\mathbf{n}_1 = \alpha(\hat{e}_1 + i\hat{e}_2)/2, \quad \mathbf{n}_2 = (\hat{e}_1 - i\hat{e}_2)/2. \tag{A18}$$

Inserting this information into (A16) and (A17) gives the result

$$\boldsymbol{\tau}_3 = [\hat{e}_1(1+\alpha)/2 + i\hat{e}_2(\alpha - 1)/2 + \hat{e}_3\alpha/2]/[1 + \alpha/2]. \tag{A19}$$

Now we need to solve for $\mathbf{n}_3$. Combining (A19) and (A17) we find

$$\boldsymbol{\tau}_3 \cdot \boldsymbol{\tau}_3 = \alpha(1 + \alpha/4)/(1 + \alpha/2)^2 = (\tanh\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3})^2$$
$$= \mathbf{n}_3 \cdot \mathbf{n}_3 - \tfrac{2}{3}(\mathbf{n}_3 \cdot \mathbf{n}_3)^2 + \cdots. \tag{A20}$$

Consequently,

$$(\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3}/\tanh\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3}) = 1 + (\alpha/3) + \cdots. \tag{A21}$$

Finally, from the relation

$$\mathbf{n}_3 = \boldsymbol{\tau}_3(\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3}/\tanh\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3}) \tag{A22}$$

we find

$$f(\alpha) = (1/2) + (5\alpha/12) + \cdots, \tag{A23}$$
$$g(\alpha) = (-i/2) + (7i\alpha/12) + \cdots,$$
$$h(\alpha) = (\alpha/2) - (\alpha^2/12) + \cdots.$$

It is clear from (A19)–(A22) that the series for f, g, and h all have nonzero radii of convergence.

APPENDIX B

The purpose of this appendix is to prove Theorem 4. Writing (4.22) a bit more explicitly, we have

$$\exp[C(\alpha, \beta)] = \exp(\alpha A)\exp(\beta B). \tag{B1}$$

The first result we will need is that C obeys the differential equation

$$\frac{\partial C}{\partial \beta} = \hat{C}[1 - \exp(-\hat{C})]^{-1}B. \tag{B2}$$

To see how this comes about, let us differentiate both sides of (B1) with respect to β. The derivative of the right-hand side is easily computed,

$$\frac{\partial}{\partial \beta}\exp(\alpha A)\exp(\beta B) = \exp(\alpha A)\exp(\beta B)B = \exp(C)B. \tag{B3}$$

Computation of the derivative of the left-hand side requires more work. We find through first order in $\delta\beta$ that

$$\exp[C(\alpha, \beta + \delta\beta)] = \exp\left(C(\alpha, \beta) + \delta\beta\,\frac{\partial}{\partial\beta}C(\alpha, \beta)\right)$$
$$= \sum_0^\infty (1/n!)\left(C(\alpha, \beta) + \delta\beta\,\frac{\partial}{\partial\beta}C(\alpha, \beta)\right)^n. \tag{B4}$$

Now expand the power series and retain zero and first order terms. The result through first order is

$$\exp[C(\alpha, \beta + \delta\beta)]$$
$$= \exp(C) + \delta\beta\sum_{n=1}^{\infty}\sum_{m=0}^{n-1}(1/n!)C^m\left(\frac{\partial C}{\partial\beta}\right)C^{n-m-1}. \tag{B5}$$

Here we have paid careful attention to the possibility that C and $\partial C/\partial\beta$ may not commute. From (B5) we conclude

$$\left(\frac{\partial}{\partial\beta}\right)\exp(C) = \sum_{n=1}^{\infty}\sum_{m=0}^{n-1}\frac{1}{n!}\,C^m\left(\frac{\partial C}{\partial\beta}\right)C^{n-m-1}. \tag{B6}$$

Next change the order of summation in (B6) to obtain

$$\left(\frac{\partial}{\partial\beta}\right)\exp(C) = \sum_{m=0}^{\infty}\sum_{l=0}^{\infty}[1/(l+m+1)!]C^m\left(\frac{\partial C}{\partial\beta}\right)C^l. \tag{B7}$$

It is a remarkable fact that the series (B7) has an integral representation,

$$\sum_{m=0}^{\infty}\sum_{l=0}^{\infty}[1/(l+m+1)!]\,C^m\left(\frac{\partial C}{\partial\beta}\right)C^l$$
$$= \int_0^1 d\gamma\exp[(1-\gamma)C]\left(\frac{\partial C}{\partial\beta}\right)\exp(\gamma C). \tag{B8}$$

This is easily verified by expanding out the two exponentials and integrating term by term.

Only a few more steps are required. We write

$$\int_0^1 d\gamma\exp[(1-\gamma)C]\left(\frac{\partial C}{\partial\beta}\right)\exp(\gamma C)$$
$$= \exp(C)\int_0^1 d\gamma\exp(-\gamma C)\left(\frac{\partial C}{\partial\beta}\right)\exp(\gamma C)$$
$$= \exp(C)\int_0^1 d\gamma\exp(-\gamma\hat{C})\left(\frac{\partial C}{\partial\beta}\right)$$
$$= \exp(C)\{[1 - \exp(-\hat{C})]/\hat{C}\}\left(\frac{\partial C}{\partial\beta}\right). \tag{B9}$$

Here we have used (4.6). Also, the last integration over γ was performed by expanding $\exp(-\gamma\hat{C})$ in a power series, integrating term by term, and then resuming the result. We conclude that

$$\left(\frac{\partial}{\partial\beta}\right)\exp(C) = \exp(C)\{[1 - \exp(-\hat{C})]/\hat{C}\}\left(\frac{\partial C}{\partial\beta}\right). \tag{B10}$$

We note for future use that Eqs. (B4)–(B10) hold quite generally, and make no use of any special form C might have.

The differentiation of both sides has been completed. Comparing (B3) and (B10) and cancelling the common factor $\exp(C)$, we find

$$\{[1-\exp(-\hat{C})]/\hat{C}\}\ \frac{\partial C}{\partial \beta} = B. \tag{B11}$$

This expression, when solved for $\partial C/\partial\beta$, gives the advertised result (B2).

The proof of Eq. (4.22) now follows immediately. Make the expansion

$$C(\alpha,\beta)=\sum \alpha^m \beta^n C_{mn} \tag{B12}$$

and substitute it into (B2) with the observation that

$$C(\alpha,0)=\alpha C_{10}=\alpha A, \tag{B13}$$

$$C(0,\beta)=\beta C_{01}=\beta B. \tag{B14}$$

A comparison of coefficients of like powers of $\alpha^m\beta^n$ gives the series (4.23).

To prove Eq. (4.24) we write a Taylor expansion of C with respect to β,

$$C(\beta)=C(0)+\beta C'(0)+(\beta^2/2)C''(0)+\cdots. \tag{B15}$$

Here we have suppressed the dependence of C on α for notational convenience. The quantity $C(0)$ is already known from (B13), and $C'(0)$ can be found from (B2) with β set equal to zero. The result is

$$C'(0)=\alpha\hat{A}[1-\exp(-\alpha\hat{A})]^{-1}B. \tag{B16}$$

Insertion of (B13) and (B16) into (B15) gives the desired result (4.24).

The computation of successively higher derivatives becomes increasingly more complicated. To find $C''(0)$, we write equation (B11) in the form

$$\mathcal{P}(\beta)C'(\beta)=B, \tag{B17}$$

where $\mathcal{P}(\beta)$ denotes the operator

$$\mathcal{P}(\beta)=[1-\exp(-\hat{C})]/\hat{C}=\int_0^1 d\nu \exp(-\nu\hat{C}). \tag{B18}$$

Next we differentiate both sides of (B17) with respect to β to find

$$\mathcal{P}'C'+\mathcal{P}C''=0, \tag{B19}$$

and hence

$$C''(0)=-\mathcal{P}^{-1}(0)\mathcal{P}'(0)C'(0). \tag{B20}$$

From the integral representation (B18) we find

$$\mathcal{P}'(\beta)=-\int_0^1 d\nu \exp(-\nu\hat{C})\int_0^1 d\gamma \exp(\gamma\nu\hat{C}) \times \nu\hat{C}'(\beta)\exp(-\gamma\nu\hat{C}). \tag{B21}$$

Here we have also used (B4)–(B8) and part of (B9) to differentiate $\exp(-\nu\hat{C})$. It follows that

$$-\mathcal{P}'(0)C'(0)=\int_0^1\int_0^1 \nu\,d\nu\,d\gamma \exp(-\nu\alpha\hat{A})\exp(\gamma\nu\alpha\hat{A}) \times\hat{C}'(0)\exp(-\gamma\nu\alpha\hat{A})C'(0). \tag{B22}$$

We also observe that

$$\hat{C}'(0)\exp(-\gamma\nu\alpha\hat{A})C'(0) = \{C'(0),\exp(-\gamma\nu\alpha\hat{A})C'(0)\} = \exp(-\gamma\nu\alpha\hat{A})\{\exp(\gamma\nu\alpha\hat{A})C'(0),C'(0)\}. \tag{B23}$$

In obtaining the last expression we have used a result analogous to (2.8). By combining (B18), (B20), (B22), and (B23) we find the final result

$$C''(0)=\alpha\hat{A}[1-\exp(-\alpha\hat{A})]^{-1}\int_0^1\int_0^1 \nu\,d\nu\,d\gamma\exp(-\nu\alpha\hat{A}) \times\{\exp(\gamma\nu\alpha\hat{A})C'(0),C'(0)\}. \tag{B24}$$

*Supported in part by the National Science Foundation under Grant No. GP-41822X.

†Present address: Laboratory of Plasma Studies, Cornell University, Ithaca, New York 14853.

[1]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading Massachusetts, 1950).
[2]E. Saletan and A. Cromer, *Theoretical Mechanics* (Wiley, New York, 1971).
[3]E.C.G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective* (Wiley, New York, 1974).
[4]G.D. Birkhoff, *Dynamical Systems*, Am. Math. Soc. No. 9 (Am. Math. Soc., New York, 1927).
[5]A. Deprit, J. Henrard, J. Price, and A. Rom, Celest. Mech. 1, 222 (1969).
[6]F. Gustavson, Astron. J. 71, 670–86 (1966).
[7]A.J. Dragt and J.M. Finn, J. Geophys. Res. 81, 2327–40 (1976).
[8]V.I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).
[9]J.M. Finn, Nucl. Fusion 15, 845 (1975); N. Filonenko, R. Sagdeev, G. Zaslavski, Nucl. Fusion 7, 253 (1967); M.N. Rosenbluth, R.Z. Sagdeev, J.B. Taylor, and G.M. Zaslavski, Nucl. Fusion 6, 297 (1966); D.W. Kerst, J. Nucl. Energy Part C 4, 253–62 (1962).
[10]A. Deprit, Celest. Mech. 1, 12 (1969).
[11]A.H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973) E. Leimanis, *The General Problem of the Motion of Coupled Rigid Bodies About a Fixed Point* (Springer, New York, 1965), p. 121.
[12]G.E.O. Giacaglia, *Perturbation Methods in Non-Linear Systems* (Springer-Verlag, Berlin, 1972).
[13]R. Abraham, *Foundations of Mechanics* (Benjamin, New York, 1967).
[14]M. Hausner and J.T. Schwartz, *Lie Groups; Lie Algebras* (Gordon and Breach, New York, 1968).
[15]S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[16]W. Engel, Math. Ann. 130, 11 (1955); 136, 319 (1958).
[17]J.M. Finn, University of Maryland Thesis (unpublished 1974).
[18]C. Froeschle, Astron. Astrophys. 9, 15 (1970).
[19]M. Henon, Q. Appl. Math. 27, 291 (1969); J. Roels and M. Henon, Bull. Astron. Ser. 3 2, 267 (1967); M. Henon and C. Heiles, Astron. J. 69, 73 (1964).
[20]J.H. Bartlett, *Classical and Modern Mechanics* (University of Alabama Press, Birmingham, Alabama, 1975); B.Z. Jenkins and J.H. Bartlett, Celest. Mech. 5, 407 (1972).
[21]J. Greene, J. Math. Phys. 9, 760 (1968).
[22]J. Moser, Bol. Soc. Mat. Mex. 5, 176 (1960).
[23]F. Rannou, Astron. Astrophys. 31, 289 (1974).
[24]C. Froeschle, Astron. Astrophys. 9, 15 (1970); 16, 172 (1972); Astrophys. Space Sci. 14, 110 (1971); C. Froeschle and J.P. Scheidecker, Astron. Astrophys. 22, 431 (1973).
[25]N. Bourbaki, *Elements de mathématiques, Fasc.* 37: *Groupes et algebres de Lie* (Hermann, Paris, 1972), Chapters 2 and 3.
[26]G.D. Birkhoff, Acta Math. 43, 1 (1920).
[27]J. Moser, Commun. Pure Appl. Math. 9, 673 (1956).
[28]H. Poincaré, *Les methods nouvelles de la mechanique celeste*, 1892 (Reprinted by Dover, New York, 1957).
[29]J. Moser, *Stable and Random Motions in Dynamical Systems* (Princeton U.P., Princeton, N.J., 1973).
[30]P. Halmos, *Finite Dimensional Vector Spaces* (Van Nostrand, Princeton, N.J., 1958).

Monopoles, vortices and the geometry of the Yang–Mills bundles*

Z. F. Ezawa

Laboratoire de Physique Theorique et Particules Elementaires,† Universite de Paris XI, 91405 Orsay, France

H. C. Tze

Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305
(Received 16 March 1976)

A topological classification of monopoles and vortices is formulated in terms of fibre bundles. The distinction between Dirac and 't Hooft monopoles is made in the light of the energy finiteness problem. Finite-length vortices with Dirac monopoles at the end points are also discussed.

I. INTRODUCTION AND CONCLUSION

Originally introduced to formulate and solve *global* topological problems, fibre bundles[1] have provided us with an ideal language for discussing relativity, invariance, and gauge transformations.[2] They may well become the standard mathematical baggage for particle physicists. In this paper, we apply the topology of fibre bundles to classify monopoles and vortices in non-Abelian gauge theories. The bundle formulation allows a compact, unified treatment of the 't Hooft–Polyakov monopoles,[3,4] the Dirac monopoles,[5,6] and the Nielsen–Olesen vortices.[7,8]

Here we summarize the results of our analysis. The 't Hooft and Dirac monopoles are classified by $\pi_1(H)$ and $\pi_1(G)$, respectively, where G is the global symmetry group of the physical system and H is an isotropy subgroup of G. In spontaneously broken gauge the theories, 't Hooft monopoles have finite energy while Dirac monopoles in general have infinite energy, possibly corresponding to unobservable quarks. Dirac monopoles are accompanied by Nielsen–Olesen vortices, the latter being also characterized by $\pi_1(G)$. A system of n Dirac monopoles of the same type has finite energy if n is the dimension of $\pi_1(G)$. This implies that in theories with $G=\mathrm{SU}(3)/Z_3$ a three-monopole system is physically realizable as well as monopole–antimonopole system. They would be field-theoretical candidates for extended baryons and mesons.[9,6]

II. THE YANG-MILLS BUNDLES AND MONOPOLES

In a geometrical approach to local gauge invariance, a fibre bundle can be uniquely associated with a physical system of fields, as we now explain. The base space x is the $(3+1)$-dimensional space–time manifold covered by a set of coordinate patches $\{V_i\}$. With each point $x\in V_i$, we associate a set of fields $\{\Phi_1, \Phi_2, \ldots\}$, which we will make cross-sections of the fibre bundle. All the realizable sets in the system make up a field manifold Y. The field Φ is defined with respect to a certain holomorphic representation $T[G]$ of a Lie group G, the symmetry group of the system. The Yang–Mills potential A_μ is the connection which assumes value in the Lie algebra of G. Let G be n-dimensional. A set of n continuous mappings $\{X_k\}$ from V_i is chosen to form a base in the Lie algebra of G. G acts on it through the adjoint representation $\mathrm{Ad}[G]$. Under the base transformation

$$S\colon \{X_k\}\to\{Y_k\},\quad Y_k=\mathrm{Ad}[S]X_k,\quad S(x)\in G, \tag{1a}$$

the fields and the connection transform as

$$S\colon \Phi(x)\to\Phi'(x)=T[S(x)]\Phi(x), \tag{1b}$$

$$A_\mu(x)\to A'_\mu(x)=\mathrm{Ad}[S(x)]A_\mu(x)+(i/e)\partial_\mu S(x)\cdot S^{-1}(x), \tag{1c}$$

respectively. We also refer to (1) as a gauge transformation.

Consider two overlapping coordinate patches V_i, $V_1\cap V_2\neq\emptyset$. The fields are defined in each patch as $\Phi_i(x_i)$, $x_i\in V_i$. At $x\in V_1\cap V_2$, $\Phi_1(x)$ and $\Phi_2(x)$ are related by (1b) if they represent the same physical object. Similarly, the connection is related by (1c). In this way, the fields as well as the connection are seamed all over the space–time manifold X.

As an example of the field system with a connection, let us consider the Lagrangian:

$$\mathcal{L}=\mathrm{Tr}F_{\mu\nu}F_{\mu\nu}+\sum_{k=1}^{n}|\nabla_\mu\Phi_k|^2+U(\Phi_1,\cdots,\Phi_n), \tag{2a}$$

$$F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu+ie[A_\mu, A_\nu], \tag{2b}$$

$$\nabla_\mu\Phi_k=\partial_\mu\Phi_k+ieT(A_\mu)\Phi_k, \tag{2c}$$

with $T(A_\mu)$ being the representation of A_μ. The equations of motion are derived from (2). By the foregoing discussion they are only valid in each coordinate patch V_i. In the overlap domain $V_1\cap V_2\neq\emptyset$, the fields are seamed with the coordinate transformation (1). The set of Higgs field solutions $\{\Phi_1,\ldots,\Phi_n\}$ to the Lagrangian system makes up the field manifold Y. Y crucially depends on the potential U and the boundary conditions. Though it is useful to have in mind a model Lagrangian, our arguments in this paper are entirely independent of it.

The system is specified by (i) the space–time manifold X, (ii) the field manifold Y, (iii) the topological group G and (iv) the gauge transformation (1). The representation $T[G]$ acts on Y effectively. G is a covering group of $T[G]$. In the following we identify G with $T[G]$ by requiring that G acts on Y effectively. By this we mean that such an element $g\in G$ is only the unit element as satisfies $T(g)u=u$ for all $u\in Y$. According to the existence theorem[1] a fibre bundle is constructed, uniquely up to isomorphism, with X as the base space,

Y the fibre and G the structure group. The field Φ is nothing but a cross section of the fibre bundle.

For simplicity we assume that G is connected. Then, there exist the universal covering group G^* and a subgroup C of the center Z^* of G^* so that the isomorphism

$$G^*/C = G, \quad C \subset Z^* \tag{3}$$

follows.[10] By definition G^* is simply connected. Next, we define the isotropy group $H(u)$ at a point $u \in Y$ by $H(u) = \{h \in G;\ T(h)u = u\}$. By the equivalence relation that $u \sim v$ if and only if $G/H(u) = G/H(v)$, all the elements of Y are grouped into the equivalence classes $\{Y_i\}$. Then, by fixing $u \in Y_i$ arbitrarily, the homeomorphism

$$Y_i = G/H(u), \quad u \in Y_i, \tag{4}$$

is proved.[10] G acts on Y_i transitively.

We explain this. Since G is the symmetry group, if $u(x)$ is a solution then $T[S(x)]u(x)$ is a solution. But it is not true that any solution $v(x)$ is obtainable from $u(x)$ by a gauge transformation. We define an homogeneous space $Y(u)$ by $Y(u) = \{v \in Y;\ v = T(S)u,\ \forall S \in G\}$. the fibre Y is the union of all such nonoverlapping homogeneous spaces Y_i. Y is the set of all possible solutions to the system, with the boundary condition picking out one of the Y_i's. By taking an element u from Y_i, the topological structure of Y_i is determined by (4). In the spontaneously symmetry broken theory, $H(u)$ is the unbroken symmetry at the point u. Especially, if the symmetry is completely broken, $H(u) = \{e\}$ for all $u \in Y$, and the fibre bundle is the principal bundle.[1]

In field theory, we consider an object occupying a finite domain in X. Either it arises from the dynamics or must be introduced explicitly by hand. The extended 't Hooft monopole[3] is an example of the former and the point-like Dirac monopole[5] is one of the latter. The object interacts with the fields and we seek to categorize it through this interaction.[11]

We take an element $u(x_0) \in Y$. We perform the equivalent transport of $u(x_0)$ along a loop l. At a point $x \in l$, it gives

$$u_l(x) = T[g_l(x, x_0)]u(x_0), \tag{5a}$$

$$g_l(x, x_0) = P \exp(-ie \int_{x_0}^{x} dx_\mu A_\mu) \in G, \tag{5b}$$

P being the ordering operator along the loop l. In (5b) we have assumed that the loop l exists in a single coordinate patch, but the modification in the general case is straightforward with the use of the gauge transformation.[12]

The operator $g_l(x, x_0)$ draws a curve l^* in G as x moves from x_0 along the loop l in X. It is to be noticed that the curve l^* is not necessarily a loop in G. Now we sweep the loop l over a 2-sphere S at a fixed time in X. The corresponding curve l^* traces out a surface $g(S)$ in G. Its boundary $\partial g(S)$ is a loop in G and traced by $g_l(x_0, x_0)$. We denote by R the subset of G which is swept by all possible boundaries $\partial g(S)$. Later we shall show how to determine R. In this way, with any sphere S in X, we can associate a member of the second relative homotopy group $\pi_2(G, R, e)$. Accordingly, if two spheres S_1 and S_2 are mapped into the same member, we can find a continuous gauge transformation so that the connections are the same at the corresponding points on the spheres S_1 and S_2. During this transformation, the energy of the system changes only continuously. We are thus led to the following definition.

Definition: Two spheres S_1 and S_2 are said to enclose the same type of monopoles if they are mapped into the same homotopy class of $\pi_2(G, R, e)$.

From this definition we get the ensuing theorem.

Theorem: Monopoles are classified by the fundamental homotopy group $\pi_1(R)$.

Proof: There is an *exact* homotopy sequence[1]: $\pi_2(G) \to \pi_2(G, R) \xrightarrow{\partial} \pi_1(R)$. By the construction of R, the boundary operator ∂ is a surjection. Since G is a compact Lie group, $\pi_2(G) = 0$. Hence, $\pi_2(G, R, e) = \pi_1(R)$ by the homeomorphism ∂.

Thus far we have yet made no analysis of the boundary conditions on the fields. Of utmost importance to our work is the *energy finiteness condition*. From the Lagrangian (2) it is seen that the kinetic energy contribution of the field Φ_k at infinity is negligible if and only if the condition

$$\nabla_\mu \Phi_k = 0 \tag{6}$$

is obeyed asymptotically.[9] In general, the energy finiteness condition for a field is that the connection becomes flat with respect to the field at infinity.

If the connection is flat, then we obtain $\Phi_k(x_0) = T[g_l(x_0, x_0)]\Phi_k(x_0)$ with (5b). This expression is in fact the integrated form of (6). Since $\partial g(S)$ is the trace of $g_l(x_0, x_0)$, we have

$$T[\partial g(S)]\Phi_k(x_0) = \Phi_k(x_0). \tag{7}$$

(7) turns out to determine R, the aggregate of all boundaries $\partial g(S)$. The types of monopoles are in fact determined by this energy finiteness condition (6).

The condition (6) is purely a physical requirement. In classical electrodynamics we treat a charged test particle moving in the electromagnetic field. If the field represents such a test particle, there is no reason to put (6) for the field. Yet, the energy of the system is finite since the field represents a single particle. On the contrary, if we treat a spontaneously broken gauge theory, the vacuum is a medium, for instance, the condensed phase of the Higgs fields. The system has finite energy if and only if (6) holds asymptotically for all fields that participate in making the vacuum. Depending on the physical requirements, we have to impose the appropriate energy finiteness condition.

At present, we are most interested in finding extended objects in spontaneously broken gauge theories. Thus, we require (6) for all the fields eventually. However, there are two different categories. (A) It is satisfied for a single monopole, or (B) it is not satisfied by a single monopole. In the case (A), any monopole system has finite energy and is realizable. In the case (B), only certain sets of monopoles are realizable. To the categories (A) and (B) correspond the 't Hooft monople and the Dirac monopole, respectively.

A. The 't Hooft monopoles

We require all fields Φ_k to satisfy the condition (6) asymptotically. Because the asymptotic behaviors of $\{\Phi_k\}$ determine the solution uniquely, the fibre over x at infinity is enough to characterize the object. In the Lagrangian theory, the set $\{\Phi_k\}$ which gives the absolute minimum of the potential constitutes the classical vacuum. The set of all classical vacua Y is the fibre. Y is decomposed into the union of the homogeneous spaces $\{Y_i\}$.

Since (7) holds all the fields we deduce $\partial g(S) \in H(u)$, $H(u)$ being the isotropy group at $u \in Y_i$. This is necessary and sufficient for finite energy of the system. Hence, we have

Corollary A: 't Hooft monopoles are classified by $\pi_1(H(u))$ for each homogeneous space Y_i where $u \in Y_i$.

The isotropy group $H(u)$ is the unbroken symmetry at a point u. $\pi_1(H(u))$ does not depend on u but only on the homogeneous space Y_i that contains u. If the system is subject to no additional boundary conditions, it allows all types of monopoles given by $\pi_1(H_i)$ for all homogeneous spaces Y_i of Y.

Here we recall the classification of 't Hooft monopole proposed in Ref. 4. The equivalent transport maps a sphere S into a sphere $g(S)u(x_0)$ in $Y(u)$ because of (7). Therefore, the solutions are also classified by $\pi_2(Y(u))$. Due to (3) and (4) we get $\pi_2(Y_i) = \pi_1(H_i)$, and the two classifications are equivalent. However, this result is only applicable to 't Hooft monopoles. A few illustrations are in order.

A trivial example is the case where the symmetry is completely broken. The bundle is the principal bundle. There are no 't Hooft monopoles. Physically, this is also clear since there is no long range gauge field remaining.

The next simplest example is the case where the symmetry is completely broken except for a U(1) subgroup $H_{\rm em}$ which we identify with "electromagnetism."[3] The monopoles are classified by $\pi_1(H_{\rm em}) = Z_\infty$, the additive group of integers. The topological spectrum is the same as the well-known Abelian U(1) monopoles. In gauge theories with $G = {\rm SU}(N)/Z_N$ we introduce a sufficient number of fields in the adjoint representation until the only unbroken symmetry is given by the generator

$$\lambda_{\rm em} = (1/N){\rm diag}(1, 1, \ldots, 1-N) \tag{8}$$

regarded as an element of the universal covering group $G^* = {\rm SU}(N)$. As the loop $l = \partial\hat{S}$ covers the sphere S, $g_l(x_0, x_0)$ moves in G as $\exp[-ie\Phi(\hat{S})\lambda_{\rm em}]$, with the flux $\Phi(\hat{S})\lambda_{\rm em} = \oint_{\partial\hat{S}} dx_\mu A_\mu$. $\partial g(S)$ traces a loop if and only if $\exp[-ie\Phi(S)\lambda_{\rm em}] = 1$. By regarding $g_l(x_0, x_0)$ as an element of $G^* = {\rm SU}(N)$, this is equivalent to requiring $\exp[-ie\Phi(S)\lambda_{\rm em}] \in Z_N$, with $\lambda_{\rm em}$ explicitly given by (8). Using the relation $\exp(i2\pi m\lambda_{\rm em}) = \exp(i2\pi m/N)$, we get $\Phi(S) = (2\pi/e)m$. The monopole is quantized in Dirac units. In general, the space–time manifold cannot be covered by a single coordinate patch. But, *this is possible* for the 't Hooft monopole bearing m-Dirac flux units if $m = 0 \pmod N$, since $g_l(x_0, x_0)$ traces out a loop not only in $G = {\rm SU}(N)/Z_N$ but also in $G^* = {\rm SU}(N)$. The monopole solution found by 't Hooft[3] is just such an example for $G = {\rm SU}(2)/Z_2 = {\rm O}(3)$. Finally, we notice that by a gauge transfromation we can make $A_\mu = A_\mu^{\rm em}\lambda_{\rm em}$ locally at infinity. In this gauge $g_l(x, x_0)$ is entirely in $H_{\rm em}$. Even if $m = 0 \pmod N$, we need more than one coordinate patch. The system looks as if there are only the electromagnetic field and the magnetic monopoles. This is the Abelian gauge obtained by a singular gauge transformation in Ref. 13 for the 't Hooft model.

B. The Dirac monopoles

The Dirac monopole is by definition any monopole that is not 't Hooft's. This is too general to get any result. Practically the most important case is the one where no fields are subject to the condition (6). We call the corresponding monopole a Dirac monopole in the narrow sense. We consider this case only. There are no constraints on $\partial g(S)$ and we get $R = G$. Hence we conclude

Corollary B: Dirac monopoles in the narrow sense are classified by $\pi_1(G) = C$.

Here we have used (3) which relates G to the universal covering group. In electrodynamics $G = {\rm U}(1)$ and $\pi_1(G) = Z_\infty$; there exists an infinite variety of Abelian monopoles. In the $G = {\rm SU}(N)$ case, $\pi_1(G) = 0$ and there are no monopoles. This corresponds to introducing the fields, say, in the fundamental representation. In the $G = {\rm SU}(N)/Z_N$ case, $\pi_1(G) = Z_N$ and the monopole strength is defined only by modulo N.

In the presence of Dirac monopoles, we always need more than one coordinate patch. This is in practice very inconvenient since we have to solve the equations in each coordinate patch and seam the solution across the patches. However, as is known,[5] there is an equivalent description where the connection is allowed to be *singular* along the Dirac strings. For instance, in the ${\rm SU}(N)/Z_N$ theory, we make the substitution[6]

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + ie[A_\mu, A_\nu] - G^*_{\mu\nu}\lambda_{\rm em} \tag{9}$$

with (8). $G^*_{\mu\nu}$ represents the string singularities originating from the monopoles.[5] Then a single coordinate patch A_μ is sufficient for any Dirac monopole system with this singular gauge choice. In (9) we have introduced the strings in the $\lambda_{\rm em}$ direction because Z_N is a subgroup generated by $\lambda_{\rm em}$.

In this description the phase factor (5b) picks up the factor $P\exp(-ie\oint dx_\mu A_\mu)$, the loop being *infinitesimal* around the singular point, as the loop l passes across the point. This factor belongs to C, regarding as an element in the covering group G^*. Because G^* is simply connected, the types of the strings are given by the discrete elements of C in accordance with the corollary B. In the ${\rm SU}(N)/Z_N$ case, there is a gauge in which $\oint A_\mu(x)dx_\mu = (2\pi/e)m\lambda_{\rm em}$. The monopoles are quantized in Dirac units.

We have seen that we can apply the same topological analysis to 't Hooft and Dirac monopoles. But this does not mean that these monopoles are of similar natures. 't Hooft monopoles are dynamically created in a spontaneously broken vacuum, and hence the masses and other parameters are to be calculated. On the other

hand, Dirac monopoles are introduced by hand: The masses, spins and other internal degrees of freedom are free parameters we can assign arbitrarily.

III. THE YANG-MILLS BUNDLES AND VORTICES

The system of a single Dirac monopole has infinite energy in a spontaneously broken vacuum since the energy finiteness condition (6) is not satisfied asymptotically. Without violating the analysis of Sec. II B, we can require (6) outside of a *tube* starting at the monopole and ending at infinity. The object that is confined in the tube is a Dirac monopole accompanied by the vortex. It is easy to see that, far enough away from the point monopole, its accompanying vortex is indistinguishable from the Nielsen—Olesen vortex.[7] This point has been discussed in ample detail in a recent work.[8]

Here we define the Nielsen—Olesen vortices in our terminology. We take the same fibre bundle, the fibre Y being the aggregate of all classical vacua. Any loop l around the tube is mapped into a path $g_l(x_0, x_0)$ in G by the equivalent transport (5). Since the connection is flat outside the tube, we get $T[g_l(x_0, x_0)]u = u$. The path g_l connects the unit element e and an element of the isotropy subgroup $H(u)$. Just as in the definition for topologically distinct monopoles, we are led to the following classification.

Definition: Two loops l_1 and l_2 are said to enclose the same type of vortices if they are mapped into the same homotopy class of $\pi_1(G, H_i, e)$.

From this we deduce:

Theorem: Vortices are formed and classified by $\pi_1(G)$ if and only if H_i is embedded in a simply connected domain of G.

Proof: There is an exact sequence[1]: $\pi_1(H_i) \xrightarrow{j^*} \pi_1(G) \to \pi_1(G, H_i) \to \pi_0(H_i) = 0$. Here $\pi_1(G) = \pi_1(G, H_i)$ if and only if j^* sends $\pi_1(H_i)$ to the neutral element of $\pi_1(G)$.

The classification is the same as that of the Dirac monopoles. Nielsen—Olesen vortices are such objects that terminate at Dirac monopoles.[8] The isotropy group $H(u)$ is the unbroken symmetry at a point $u \in Y$. A sufficient condition for the vortex formation is to break the symmetry completely, i.e., $H(u) = \{e\}$. In the U(1) theory,[7] one Higgs field is enough for this purpose. In the $SU(2)/Z_2$ theory,[7] two Higgs fields are introduced in the regular representation. In the $SU(3)/Z_3$ theory, we take two Higgs fields $u = \{\Phi_1, \Phi_2\}$ in the $(\lambda_4 - \lambda_5)$ or $(\lambda_6 - \lambda_7)$ plane. They are arbitrarily fixed to minimize the potential; hence $u \in Y$. The homogeneous space $Y(u)$ is homeomorphic to $SU(3)/Z_3$ and $H(u) = \{e\}$. Vortices, characterized by $\pi_1(SU(3)/Z_3) = Z_3$, will be generated along the λ_8 axis.

A single Dirac monopole cannot exist in the spontaneously broken vacuum because the energy is infinite. By the modulo property of their strength, a system of Dirac monopoles has finite energy if their net strength is zero. This implies that in the $SU(N)/Z_N$ theory we can confine N monopoles of the same type in a finite domain outside of which (6) holds. These monopoles are tied together by finite-length vortices. Especially in the $SU(3)/Z_3$ case, *this gives a baryonic vortex string as well as a mesonic one*. This remarkable feature was pointed out by Mandelstam,[9] and recently analysed in details in Refs. 6 and 8. However, this system of Dirac monopoles confined in a domain has no monopole as a whole system. It would decay into pure vacuum unless the monopole has additional degrees of freedom, like quark quantum numbers. But this problem will be the subject of another paper.

ACKNOWLEDGMENTS

We wish to warmly thank Holger Nielsen, Jeffrey Goldstone, and Marvin Weinstein for helpful suggestions at the early phase of this work.

*Work supported in part by the Energy Research and Development Administration.

†Laboratoire associé au Centre National de la Recherche Scientifique.

[1]There are many books on fibre bundles. See, e.g., N. Steenrod, *The Topology of Fibre Bundles* (Princeton U. P., Princeton, N.J., 1951); S.T. Hu, *Homotopy Theory* (Academic, New York, 1959). For connection, see, e.g: S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. 1 (Interscience, New York, 1963).

[2]A. Trautman, Rep. Math. Phys. **1**, 27 (1970); L.N. Chang, K.I. Macrae and F. Mansouri, Phys. Rev. D **13**, 235 (1976).

[3]G. 't Hofft, Nucl. Phys. B **79**, 276 (1974); A.M. Polyakov, Zh. Eksp. Teor. Fiz. Pis. Red. **20**, 430 (1974) [JETP Lett. **20**, 194 (1974)].

[4]Yu.S. Tyupkin, V.A. Fateev, and A.S. Shvarts, Zh. Eksp. Teor. Fiz. Pis. Red. **21**, 91 (1975) [JETP Lett. **21**, 42 (1975)]; M.I. Monastyrsky and A.M. Perelomov, *ibid.* **21**, 194 (1975) [*ibid.* **21**, 43 (1975)].

[5]P.A.M. Dirac, Phys. Rev. **74**, 817 (1948); Proc. R. Soc. (Lond.) A **133**, 60 (1931).

[6]Z.F. Ezawa and H.C. Tze, Nucl. Phys. B **100**, 1 (1975).

[7]H.B. Nielsen and P. Olesen, Nucl. Phys. B **61**, 45 (1973).

[8]H.C. Tze and Z.F. Ezawa, "Global Signatures of Gauge Invariance: Vortices and Monopoles," SLAC PUB-1735, April 1976 to appear in Phys. Rev. D. For the Abelian case, see Z.F. Ezawa and H.C. Tze, Nucl. Phys. B **98**, 264 (1975).

[9]S. Mandelstam, Phys. Lett. B **53**, 476 (1975).

[10]L. Pontrjagin, *Topological Group* (Princeton U.P., Princeton, N.J., 1939).

[11]E. Lubkin, Ann. Phys. (N.Y.) **23**, 233 (1963).

[12]T.T. Wu and C.N. Yang, Phys. Rev. D **12**, 3845 (1975) and Stony Brook Preprint ITP-SB 76-5.

[13]J. Arafune, P.G.O. Freund, and C.J. Goebel, J. Math. Phys. **16**, 433 (1975).

On the stationary axisymmetric Einstein–Maxwell field equations

Roberto Catenacci* and Joaquin Diaz Alonso†

Groupe d'Astrophysique Relativiste, Observatoire de Meudon, 92 Meudon, France
(Received 15 March 1976; revised manuscript received 27 May 1976)

We show the existence of a formal identity between Einstein's and Ernst's stationary axisymmetric gravitational field equations and the Einstein–Maxwell and the Ernst equations for the electrostatic and magnetostatic axisymmetric cases. Our equations are invariant under very simple internal symmetry groups, and one of them appears to be new. We also obtain a method for associating two stationary axisymmetric vacuum solutions with every electrostatic known.

1. INTRODUCTION

Although it has been recently proved[1] that the Kerr metric represents the most general topologically spherical stationary axistymmetric black hole solution of the vacuum Einstein equations and the various no hair theorems[2] provide a fairly convincing proof that the Kerr–Newmann metric is the most general "physically well behaved" black hole solution of the coupled Einstein–Maxwell stationary axisymmetric equations, it is worthwhile to investigate the problem of finding exact solutions of Einstein–Maxwell equations, having in mind to obtain solutions which may be applied to more complicated situations. Moreover, we think that this problem has an intrinsic theorical interest due to the nonlinear nature of the equations involved and the relative scarcety of exact solutions.

In this paper we present a formalism which can be used for generating families of exact solutions of stationary axisymmetric vacuum Einstein field equations and of the static axistymmetric Einstein–Maxwell equations.

Our method generalizes a previous work,[3] and it is essentially the working out of the consequences of an interesting formal analogy between the Einstein–Maxwell and the Ernst equations.[4] In Sec. 2 we give a short review of the Ernst approach, and we present the general equations in a form of vanishing divergences. In Sec. 3 we consider some particular cases and we show explicitly the above mentioned formal analogy. By introducing maps between the sets of solutions, we show how this analogy can be used for generating solutions.

In the last section we study the internal symmetry groups of the equations, and we find a geometrical origin for the Kinnersley groups[5] (in the static case). We present also a new internal symmetry groups of Einstein–Maxwell static axisymmetric equations with can be used (eventually combined with the other knowns groups and the transformations presented in Sec. 3) for generating families of solutions.

2. THE ERNST METHOD AND THE GENERAL EQUATIONS

The most general invertible stationary axisymmetric electrovacuum line element can be written (see, e.g., the review by Carter[6]) in the Papapetrou system of coordinates as:

$$ds^2 = f^{-1}\{e^{2\gamma}(d\rho^2 + dz^2) + \rho^2 d\phi^2\} - f(dt + \omega d\phi)^2, \tag{2.1}$$

where f, ω, and γ are functions of ρ and z only. The metric function γ is not independent and can be obtained from f and ω.[7]

The independent Einstein-Maxwell equations can be written as

$$\begin{aligned}
&\nabla\cdot\left(\frac{f}{\rho^2}(\nabla A_3 - \omega\nabla A_4)\right) = 0,\\
&\nabla\cdot\left(\frac{\nabla A_4}{f} + \frac{f\omega}{\rho^2}(\nabla A_3 - \omega\nabla A_4)\right) = 0,\\
&\nabla\cdot\left(\frac{f^2}{\rho^2}\nabla\omega + 4\frac{f}{\rho^2}(\nabla A_3 - \omega\nabla A_4)\right) = 0,\\
&\nabla\cdot\left\{\frac{\nabla f}{f} + \frac{f^2}{\rho^2}\omega\nabla\omega - 2\frac{f}{\rho^2}(A_3 + \omega A_4)(\nabla A_3 - \omega\nabla A_4)\right.\\
&\qquad\left. - \frac{\nabla A_4^2}{f}\right\} = 0,
\end{aligned} \tag{2.2}$$

where A_4 and A_3 are the usual t and φ component of the electromagnetic potential, and these are functions of ρ and z only. The differential operators are defined with respect to the flat three-dimensional metric:

$$dl^2 = d\rho^2 + dz^2 + \rho^2 d\phi^2.$$

The form (2.2) of the equations can be obtained from the form presented by Ernst[4] by some algebraic manipulations where we have utilized the relation

$$\nabla^2(\log\rho) = 0,$$

which holds in the pure vacuum and pure electromagnetic case. Ernst[4] introduces two auxiliary potentials A_3' and φ_{Ernst}, related to A_3 and ω by

$$\rho^{-1}f(\nabla A_3 - \omega\nabla A_4) = \hat{n} \times A_3',$$

$$\frac{f^2}{\rho^2}\nabla\omega - 2A_4\frac{\hat{n} \times \nabla A_3'}{\rho} + 2A_3'\frac{\hat{n} \times \nabla A_4}{\rho} = \frac{\hat{n} \times \nabla\varphi_{\text{Ernst}}}{\rho}$$

(where $\hat{n}$ is a unit vector in the azimuthal direction), and he defines two complex potentials $\mathscr{E}$ and ψ:

$$\psi = A_4 + iA_3',$$

$$\mathscr{E} = f - \psi\psi^* + i\varphi_{\text{Ernst}}.$$

In terms of the potentials $\mathscr{E}$ and ψ, the Einstein–Maxwell equations can be written as

$$\begin{aligned}\{\text{Re}\mathscr{E} + |\psi|^2\}\nabla^2\mathscr{E} &= \nabla\mathscr{E}\cdot(\nabla\mathscr{E} + 2\psi^*\nabla\psi),\\ \{\text{Re}\mathscr{E} + |\psi|^2\}\nabla^2\psi &= \nabla\psi\cdot(\nabla\mathscr{E} + 2\psi^*\nabla\psi).\end{aligned} \tag{2.3}$$

By introducing a potential φ, related to the Ernst one by

$$\varphi_{\text{Ernst}} = \varphi + 2A_4A_3'$$

after some algebraic manipulations, one can recast the system (2.3) in the form

$$\begin{aligned}&\nabla\cdot\left(\frac{\nabla A_3'}{f} - \frac{A_4}{f_2}(\nabla\varphi + 4A_4\nabla A_3')\right) = 0,\\ &\nabla\cdot\left(\frac{\nabla A_4}{f} + \frac{A_3'}{f_2}(\nabla\varphi + 4A_4\nabla A_3')\right) = 0,\\ &\nabla\cdot\left(\frac{1}{f_2}(\nabla\varphi + 4A_4\nabla A_3')\right) = 0,\\ &\nabla\cdot\left(\frac{\nabla f}{f} + \frac{(\varphi + 2A_4A_3')}{f^2}(\nabla\varphi + 4A_4\nabla A_3') - \frac{\nabla(A_4^2 + A_3'^2)}{f}\right) = 0,\end{aligned} \tag{2.4}$$

which is analog to the (2.2) form of Einstein–Maxwell equations. This formal "analogy" can be converted into a formal "identity" in several restricted cases as we shall show in the following section.

3. PARTICULAR CASES

From the Einstein-Maxwell (2.2) and the Ernst (2.4) general equations, we can obtain various restricted (but important) cases.

First of all, the pure vacuum stationary case: we take $A_4 = 0$, $A_3 = 0$ in the Eqs. (2.2) and $A_4 = 0$, $A_3' = 0$ in the Eqs. (2.4); we obtain respectively[3]

$$\begin{aligned}&\nabla\cdot\left(\frac{f^2}{\rho^2}\nabla\left(\frac{\rho^2}{f^2} - \omega^2\right)\right) = 0,\\ &\nabla\cdot\left(\frac{f^2}{\rho^2}\nabla\omega\right) = 0\end{aligned} \qquad \text{(system E1)} \tag{3.1}$$

and

$$\begin{aligned}&\nabla\cdot\left(\frac{1}{f^2}\nabla(f^2 + \varphi^2)\right) = 0,\\ &\nabla\cdot\left(\frac{1}{f^2}\nabla\varphi\right) = 0\end{aligned} \qquad \text{(system E2).} \tag{3.2}$$

The purely electrostatic case is obtained by taking $\omega = 0$, $A_3 = 0$ and $\varphi = 0$, $A_3' = 0$ respectively in Eqs. (2. 2) and (2. 4). Both systems reduce to

$$\begin{aligned}&\nabla\cdot\left(\frac{1}{f}\nabla(f - A_4^2)\right) = 0,\\ &\nabla\cdot\left(\frac{1}{f}\nabla A_4\right) = 0\end{aligned} \qquad \text{(system E3).} \tag{3.3}$$

In Eqs. (3.3) one can easily introduce [due to the fact that $\nabla\cdot\{f^{-1}\nabla A_4\} = 0$] an auxiliary electrostatic potential A'_4 by

$$\nabla A_4 = f\rho^{-1}\hat{n} \times \nabla A_4'.$$

Equations (3.3) become then

$$\begin{aligned}&\nabla\cdot\left(\frac{f}{\rho^2}\nabla\left(\frac{\rho^2}{f^2} + A_4'^2\right)\right) = 0,\\ &\nabla\cdot\left(\frac{f}{\rho^2}\nabla A_4'\right) = 0\end{aligned} \qquad \text{(system E4).} \tag{3.4}$$

We can consider also the magnetostatic case: $\varphi = 0$, $A_4 = 0$ in (2.4) and $\omega = 0$, $A_4 = 0$ in (2.2). The general equations reduce, respectively, to

$$\begin{aligned}&\nabla\cdot\left(\frac{1}{f}\nabla(f - A_3'^2)\right) = 0,\\ &\nabla\cdot\left(\frac{1}{f}\nabla A_3'\right) = 0\end{aligned} \qquad \text{(system E5)} \tag{3.5}$$

and

$$\begin{aligned}&\nabla\cdot\left[\frac{f}{\rho^2}\nabla\left(\frac{\rho^2}{f} + A_3^2\right)\right] = 0,\\ &\nabla\cdot\left(\frac{f}{\rho^2}\nabla A_3\right) = 0\end{aligned} \qquad \text{(system E6).} \tag{3.6}$$

We are left then with six systems of equations formally identical; in fact, all of them can be put in the standard form:

$$\begin{aligned}\frac{\mathscr{E}_K^+ + \mathscr{E}_K^-}{2}\nabla^2\mathscr{E}_K^+ &= \nabla\mathscr{E}_K^+\cdot\nabla\mathscr{E}_K^+,\\ \frac{\mathscr{E}_K^- + \mathscr{E}_K^+}{2}\nabla^2\mathscr{E}_K^- &= \nabla\mathscr{E}_K^-\cdot\nabla\mathscr{E}_K^-,\end{aligned} \tag{3.7}$$

where $K = 1, 2, 3, 4, 5, 6$ and (with an obvious meaning of the index)

$$\begin{aligned}&\mathscr{E}_1^\pm = \rho/f \pm \omega, && \mathscr{E}_2^\pm = f \pm i\varphi,\\ &\mathscr{E}_3^\pm = f^{1/2} \pm A_4, && \mathscr{E}_4^\pm = \rho/f^{1/2} \pm iA_4',\\ &\mathscr{E}_5^\pm = f^{1/2} \pm A_3', && \mathscr{E}_6^\pm = \rho/f^{1/2} \pm iA_3.\end{aligned}$$

This formal identity can be used for generating automatically solutions of the various systems when we know a solution of one of them.

More precisely one can define the following maps between the sets $\{EK\}_{K=1,2,3,4,5,6}$ of the solutions of the systems EK:

$$\begin{array}{ccc} \{E1\} & \xrightarrow[B_{2,1}]{A_{2,1}} & \{E2\} \\ {\scriptstyle B_{1,3}}\downarrow & & \uparrow{\scriptstyle B_{4,2}} \\ \{E3\} & \xrightarrow[B_{4,3}]{A_{4,3}} & \{E4\} \\ {\scriptstyle B_{3,5}}\downarrow & & \uparrow{\scriptstyle B_{6,4}} \\ \{E5\} & \xrightarrow[B_{6,5}]{A_{6,5}} & \{E6\} \end{array}$$

where the maps $B_{i,j}$ are defined by

$$\mathscr{E}_i^{\pm} \xrightarrow{B_{i,j}} \mathscr{E}_j^{\pm} = \mathscr{E}_i^{\pm}, \qquad i,j = 1,2,3,4,5,6.$$

The metric functions (or the electromagnétic potentials) associated with the solutions generated by some of the $B_{i,j}$ are in general complex, but, in several cases, it is possible to arrange for them to be real by parameter adjustments. The $A_{i+1,i}$, $i = 1, 3, 5$ are defined by

$$(f,\varphi) \xrightarrow[\substack{\tilde{f}=f \\ \nabla\tilde{\omega}=\rho f^{-2}\hat{n}\times\nabla\varphi}]{A_{2,1}} (\tilde{f},\tilde{\omega}),$$

$$(f,A_4') \xrightarrow[\substack{\tilde{f}=f \\ \nabla\tilde{A}_4=f\rho^{-1}\hat{n}\times\nabla A_4'}]{A_{4,3}} (\tilde{f},\tilde{A}_4),$$

$$(f,A_3) \xrightarrow[\substack{\tilde{f}=f \\ \nabla\tilde{A}'_3=f\rho^{-1}\hat{n}\times\nabla A_3}]{A_{6,5}} (\tilde{f},\tilde{A}_3')$$

(modulo a trivial indeterminacy, one can define the converse applications).

The most important relations satisfied by these maps are

$$\left.\begin{aligned} &(B_{i,K})^{-1} = B_{K,i} \\ &B_{i,K}\circ B_{K,l} = B_{i,l} \end{aligned}\right\}, \quad i,K,l = 1,2,3,4,5,6, \tag{3.8a, 3.8b}$$

$$\left.\begin{aligned} &A_{i+1,i}\circ B_{i,i-1} = B_{i+1,i}\circ A_{i,i+1} \\ &A_{i+1,i} = B_{l,i}\circ A_{l+1,l}\circ B_{i+1,l-1} \end{aligned}\right\}, \quad i,l = 1,3,5. \tag{3.8c, 3.8d}$$

The relation (3.8c) rules out the possibility of generating infinite chains of solutions by repeated applications of the transformations $A_{i+1,i}\circ B_{i+1}$.

The maps $B_{5,3}$ and $B_{6,4}$ correspond to the well-known invariance of the energy–momentum tensor of an electromagnetic field (in vacuum) under the changement of the electric into the magnetic field and vice versa.

The map $B_{3,2}$ was utilized by Esposito and Witten[8] for generating infinite chains of solutions; its existence; has been also pointed out by Misra *et al.*[9] in a different context.

4. INTERNAL SYMMETRY GROUPS

An important consequence of the fact that we have all the equations in the standard form (3.7) is that one can easily find an internal symmetry group.

As is well known, the Einstein vacuum stationary axisymmetric equations are invariant under the group of linear transformations in the two-space generated by the two killing vectors $\partial/\partial t$ and $\partial/\partial\varphi$. If we compute the laws of transformation of the potentials $\mathscr{E}_1^{\pm}$ under an action of this group, we can find automatically an internal symmetry groups of the others systems of equations. We have

$$\mathscr{E}_K^{\pm} \xrightarrow{\mathscr{G}_K(a,b,c,d)} \tilde{\mathscr{E}}_K^{\pm} = \frac{a\mathscr{E}_K^{\pm} \pm b}{\pm c\mathscr{E}_K^{\pm} + d}, \tag{4.1}$$

$K = 1, 3, 5, \quad ad+bc > 0, \quad a, d, b, c$ real,

and

$$\mathscr{E}_{K+1}^{\pm} \xrightarrow{\mathscr{G}_{K+1}(a,b,c,d)} \tilde{\mathscr{E}}_{K+1}^{\pm} = \frac{a\mathscr{E}_{K+1}^{\pm} \pm ib}{\pm ic\mathscr{E}_K^{\pm} + d}, \tag{4.2}$$

$K = 1, 3, 5, \quad ad - bc > 0, \quad a, d, b, c$ real.

$\tilde{\mathscr{E}}_K^{\pm}(\tilde{\mathscr{E}}_{K+1}^{\pm})$ belongs to $\{EK\}$ ($\{EK+1\}$) whenever $\mathscr{E}_K^{\pm}(E_{K+1}^{\pm})$ belongs to $\{EK\}$ ($\{EK+1\}$). By using the procedure presented in a paper by Kloster *et al.*,[10] one can see that the invariance groups obtained are the most general for which

$$\frac{\partial\tilde{\mathscr{E}}_i^{+}}{\partial\mathscr{E}_i^{-}} = \frac{\partial\tilde{\mathscr{E}}_i^{-}}{\partial\mathscr{E}_i^{+}} = 0.$$

From the explicit form of the transformation laws of the functions

$$\psi_K = \frac{\mathscr{E}_K^{+} + \mathscr{E}_K^{-}}{2}, \qquad \phi_K = \frac{\varepsilon_K^{+} - \varepsilon_K^{-}}{2},$$

$$\psi_{K+1} = \frac{\mathscr{E}_{K+1}^{+} + \mathscr{E}_{K+1}^{-}}{2}, \qquad \phi_{K+1} = \frac{\varepsilon_{K+1}^{+} - \varepsilon_{K+1}^{-}}{2i},$$

$K = 1, 3, 5.$

One can verify that (formally) it is possible to take a, b, c, d complex. One has the interesting relations

$$g_k(a,b,c,d) = B_{1,k}\cdot g_1(a,b,c,d)\cdot B_{k,1}, \quad |k-1| = 0, 2, 4, \tag{4.3}$$

and

$$g_k(a,b,c,d) = B_{1,k}\cdot g_1(a,ib,ic,d)\cdot B_{k,1}, \quad |k-1| = 1, 3, 5 \tag{4.4}$$

The group g_2 is the well-known Geroch–Kinnersley group.[10,11] One can easily see that the groups g_3 and g_5 are respectively the pure electrostatic and the pure magnetostatic part of the Kinnersley group.[5]

The groups g_4 and g_6 appear to be new. As an example, if we apply a transformation belonging to the g_4 group to the Minkowski space–time, we obtain a space–time with

$$f = (1+\alpha^2\rho^2)^2, \qquad \omega = 0$$

and an electric field $F = -2\alpha(dz\wedge dt)$, where α is a real constant.

The application of the groups g_4 and g_6 gives directly some of the metrics recently presented by Ernst.[12,13]

This is due to the fact that in the relation (4.4), the transformation B_{36} can be formally interpreted as an imaginary transformation of the type

$$dt' = \pm i \cdot d\phi,$$

$$d\phi' = \pm i \cdot dt.$$

The transformation B_{45} also is of the same nature. We want to also point out that by combining the groups g_j with the transformations $B_{i,k}$ and $A_{i+1,i}$ one can generate infinite chains of solutions (see, e.g., Ref. 8).

ACKNOWLEDGMENTS

The authors would like to thank Dr. B. Carter and Professor B. Bertotti for useful discussions. One of us (R.C.) acknowledges the hospitality of the Groupe d'Astrophysique Relativiste at Meudon Observatory. We thank also the referee for useful remarks.

*Permanent address: Istituto di Fisica Teorica, Università di Pavia, Italy.
†Fellow of European Space Agency.

[1]D.C. Robinson, Phys. Rev. Lett. **34,** 905 (1975).
[2]B.Carter, "The vacuum black hole uniqueness theorem and its conceivable generalisations," presented at the Marcel Grossman Meeting, Trieste, July 1975, Preprint.
[3]W. Kinnersley, "Recent Progress in Exact Solutions," in Proc. of the 7th International Conference on Gravitation and General Relativity.
[4]F.J. Ernst, Phys. Rev. **168,** 1415 (1968).
[5]W. Kinnersley, J. Math. Phys. **14,** 651 (1973).
[6]B. Carter, article on "Les astres occlus," edited by Dewitt and Dewitt, Les Houches (1972).
[7]A. Papapetrou, Ann. Inst. H. Poincaré A **4,** 83 (1966).
[8]P. Esposito and L. Witten, General Relativity Gravitation **6,** 387 (1975).
[9]R.M. Misra *et al.*, Phys. Rev. D **7,** 1587 (1973).
[10]S. Kloster *et al.*, J. Math. Phys. **15,** 1096 (1974).
[11]R.P. Geroch, J. Math. Phys. **12,** 918 (1971).
[12]F.J. Ernst, J. Math. Phys. **17,** 54 (1976).
[13]F.J. Ernst, J. Math. Phys. **17,** 515 (1976).