A NOTE ON THE APPLICATION OF SCHWINGER'S VARIATIONAL PRINCIPLE TO DIRAC'S EQUATION OF THE ELECTRON*

By H. E. MOSES (*Institute for Mathematics and Mechanics, New York University*)

Schwinger's variational principle has been used for a wide variety of problems involving wave motion in which it is desired to find the amplitude of a scattered wave in terms of the incident wave.** Schwinger's method makes use of the fact that the amplitude of the scattered wave satisfies a variational principle. We shall indicate this variational principle briefly.

Let us consider a vector space. We may define two different inner products, the Hermitian and symmetric inner product, in this vector space. The Hermitian inner product (a,b) of two vectors a and b is defined by the condition that

$$(a, b) = (b, a)^* \tag{1}$$

where the asterisk indicates the complex conjugate. The symmetric inner product is defined by the condition

$$(a, b) = (b, a). \tag{2}$$

Let us consider a vector space with a Hermitian inner product and consider a pair of equations

$$\begin{cases} a = Ky, \\ a' = K'y', \end{cases} \tag{3}$$

where K and K' are Hermitian adjoint operators which by definition satisfy the condition

$$(K'u, v) = (u, Kv) \tag{4}$$

for any two vectors u and v. If $K' = K$, K is a Hermitian operator. It can be shown from (3) that

$$(a', y) = (y', a) \tag{5}$$

which is called the *reciprocity theorem.* Let us define a number λ by

$$\lambda = \frac{1}{(y', a)} = \frac{1}{(a', y)}, \tag{6}$$

and the functional $\lambda\{v,v'\}$ by

$$\lambda\{v, v'\} = \frac{(v', Kv)}{(v', a)(a', v)} = \frac{(K'v', v)}{(v', a)(a', v)} \tag{7}$$

so that

$$\lambda\{y, y'\} = \lambda. \tag{8}$$

---

*Received March 25, 1952.

**See, for example, Schwinger's unpublished nuclear physics notes, or the lectures of N. Marcuvitz in the notes "Recent Developments in the Theory of Wave Propagation", Inst. for Math. and Mech., N.Y.U., 1949-50. The problem of the present paper is treated abstractly in the first set of notes. The point of view of the present note is close to that of the latter set of notes.

It can be shown that $\lambda\{v,v'\}$ is stationary for independent variations of v and v' about the values y and y' respectively and that, therefore, from (8), the stationary value of $\lambda\{v,v'\}$ is λ.

A similar statement holds if the symmetric inner product is used. In this case K' is said to be the symmetric adjoint operator of K if K' and K are related by (4).

In Dirac's theory of the electron, the elements of the vector space are functions $f(x,\gamma)$ of the coordinates denoted collectively by the vector x, and of a variable γ which is restricted to four values which may be taken as 1, 2, 3, 4. These functions are called spinor components. The Hermitian inner product of $f(x,\gamma)$ and $g(x,\gamma)$ is given by

$$\sum_{\gamma=1}^{4} \int f(x, \gamma)^* g(x, \gamma)\, dx$$

and the symmetric inner product by

$$\sum_{\gamma=1}^{4} \int f(x, \gamma) g(x, \gamma)\, dx.$$

Dirac's wave equation for the electron in an electromagnetic field is

$$\frac{i\partial\psi(x, \gamma; t)}{\partial t} = H\psi(x, \gamma; t) \tag{9}$$

where H is an operator which operates on x and γ and is given by

$$H = H_0 + q \tag{10}$$

with

$$\left.\begin{aligned} H_0 &= \sum_{j=1}^{3} i\alpha_j \frac{\partial}{\partial x_j} - m\beta \\ q &= \sum_{j=1}^{3} \alpha_j A_j(x) + e\phi(x) \end{aligned}\right\}. \tag{11}$$

Here α_j and β are Hermitian operators which operate with respect to the variable only. They satisfy the following commutation relations

$$\left.\begin{aligned} \alpha_j\alpha_i + \alpha_i\alpha_j &= 2\delta_{ij}I, \\ \beta\alpha_j + \alpha_j\beta &= 0, \\ (\beta)^2 &= I, \end{aligned}\right\} \tag{12}$$

where I is the identity operator.

The operators α_j , β can be expressed as integral operators with kernels $\alpha_j(\gamma,\gamma')$, $\beta(\gamma,\gamma')$ which are the well-known Dirac matrices.

We have taken $h/2\pi = c = 1$. The mass of the electron is m and its charge is e. The functions $A_i(x)$ and $\phi(x)$ are the vector and scalar potentials of an electromagnetic field and are taken as real and are assumed to vanish if $| x | > r_0$, for some r_0.

We shall look for solutions of equation (9) which can be written as

$$\psi(x, \gamma; t) = e^{-iEt}\chi(x, \gamma; E) \tag{13}$$

so that equation (9) leads to

$$H\,\chi(x,\gamma;E) = E\chi(x,\gamma;E). \tag{14}$$

We shall write

$$\chi = \chi_{in} + \chi_{sc}\,, \tag{15}$$

and require that χ_{in} be a solution of

$$H_0\chi_{in}(x,\gamma;E) = E\chi_{in}(x,\gamma;E) \tag{16}$$

where H_0 is given by (11). Suitable solutions are the "spinor plane wave" solutions which have the form

$$\chi_{in}(x,\gamma;E) = \chi(\gamma;E,\eta,\tau)\,\frac{e^{i|k|(\eta x)}}{(2\pi)^{3/2}} \tag{17}$$

where η is a unit vector which specifies the direction of propagation, (ηx) is the inner product of the vectors x and η, and $|k|$ is the absolute value of the momentum vector and is given by the relation

$$|k|^2 = E^2 - m^2. \tag{18}$$

Here τ is a variable which is restricted to two values which may be taken as $+1$ and -1. The significance of τ is that it represents the component of the spin in the direction of the momentum.

By substituting (31) into the equation $H_0\,\chi_{in} = E\chi_{in}$ it is seen that the functions $\chi(\gamma;E,\eta,\tau) \equiv \chi(\gamma;E)$ satisfy the following equation

$$\left\{E + |k|\sum_{j=1}^{3}\alpha_j\eta_j + \beta m\right\}\chi(\gamma;E,\eta,\tau) = 0. \tag{19}$$

For the purpose of the present note, it is not necessary to give an explicit form for the functions $\chi(\gamma;E)$; these functions can be found in textbooks. However, it will be useful to indicate the orthogonality properties of the functions. These orthogonality relations are

$$\left.\begin{aligned}\sum_{\gamma}\chi(\gamma;E,\eta,\tau)^*\chi(\gamma;E,\eta,\tau) &= \delta_{\tau\tau'}\\ \sum_{\gamma}\chi(\gamma;-E,\eta,\tau)^*\chi(\gamma;E,\eta,\tau') &= 0,\end{aligned}\right\} \tag{20}$$

also

$$\sum_{\tau}\chi(\gamma;E,\eta,\tau)\chi(\gamma';E,\eta,\tau) + \sum_{\tau}\chi(\gamma;-E,\eta,\tau)\chi(\gamma';-E,\eta,\tau)^* = \delta(\gamma,\gamma'). \tag{21}$$

From (14), (15) and (16) it is seen that χ_{sc} satisfies

$$[E - H_0]\chi_{sc}(x,\gamma;E) = q\chi(x,\gamma;E) \tag{22}$$

the solution of which can be written in terms of influence function $g(x,\gamma;x',\gamma')$,

$$\chi_{sc}(x,\gamma;E) = \sum_{\gamma'}\int g(x,\gamma;x',\gamma')q\chi(x',\gamma';E) \tag{23}$$

where

$$[E - H_0]g(x, \gamma; x', \gamma') = \delta(x - x')\ \delta(\gamma, \gamma'). \tag{24}$$

In (24) the operator $[E - H_0]$ operates on the variables x, γ. Since the physically interesting problem is that for which χ_{sc} is an outgoing wave, we shall take an outgoing wave solution for (24). It will now be shown how this solution which we donote by g is obtained.

From the commutation rules (12) for α_i, β one has

$$[E + H_0][E - H_0] = [E - H_0][E + H_0] = (E^2 - m^2 + \nabla^2) = (|k|^2 + \nabla^2). \tag{25}$$

Consider now a solution of the differential equation

$$(E^2 - m^2 + \nabla_x^2)s(x, x') \equiv (|k|^2 + \nabla_x^2)s(x, x')$$
$$= \delta(x - x'). \tag{26}$$

(The subscript x on the operator ∇^2 indicates that the differentiations are to be carried out on the variable x rather than x'.)

Any solution $s(x,x')$ of equation (26) can be used to form a solution of equation (24). From (25) and (26) one has

$$[E - H_0][E + H_0]s(x, x')\ \delta(\gamma, \gamma') = \delta(x - x')\ \delta(\gamma, \gamma'). \tag{27}$$

Hence a solution $g(x, \gamma; x', \gamma')$ of (24) is

$$g(x, \gamma; x', \gamma') = [E + H_0]s(x, x')\ \delta(\gamma, \gamma'). \tag{28}$$

This method of obtaining influence functions for the Dirac operator is a well-known one. A solution $s_r(x,x')$ of (26) which leads to an outgoing wave is

$$s_r = -\frac{e^{i|k||x-x'|}}{4\pi\,|x - x'|}. \tag{29}$$

The influence function g_r obtained using s_r is in explicit form

$$g_r(x, \gamma; x', \gamma') = -\left\{\frac{-[|k|\,|x - x'| + i]}{|x - x'|^2} \sum_{i=1}^{3} \alpha_i(\gamma, \gamma')(x_i - x_i') - m\beta(\gamma, \gamma') + E\ \delta(\gamma, \gamma')\right\} \frac{e^{i|k||x-x'|}}{4\pi\,|x - x'|}. \tag{30}$$

Here $\alpha_i(\gamma, \gamma')$, $\beta(\gamma, \gamma')$ are the matrices which represent the operators α, β. The function g_r represents an outgoing wave because the time factor is e^{-iEt} (see (13)).

For large values of $|x|$, the expression for χ_{sc} becomes

$$\chi_{sc}(x, \gamma; E) = -\frac{e^{i|k||x|}}{4\pi\,|x|} \sum_{\gamma'} \left\{\left[E\ \delta(\gamma, \gamma') - |k| \sum_{i=1}^{3} \alpha_i(\gamma, \gamma')\eta_{1i} - m\beta(\gamma, \gamma')\right] \cdot \int e^{-i|k|(x', \eta_1)} q\chi(x', \gamma'; E)\ dx'\right\}. \tag{31}$$

In the above expressions and in those following the result of operating with an operator such as q on a function $f(x, \gamma)$ will be a function of x, γ which will be denoted by $qf(x, \gamma)$.

Here η_1 is a unit vector with components η_{1i} defined by

$$x = |x| \eta_1 \tag{32}$$

We should like to re-write (31) so that the amplitudes of the spherical waves are inner products in order that we may ultimately use the variational principle to obtain these amplitudes. This line of thought motivates our use of the identity

$$\left[E\, \delta(\gamma, \gamma') - |k| \sum_{i=1}^{3} \alpha_i(\gamma, \gamma')\eta_i - m\beta(\gamma, \gamma')\right] = 2E \sum_{\tau} \chi(\gamma; E, \eta, \tau)\chi(\gamma'; E, \eta, \tau)^* \tag{33}$$

Though this identity is fundamental in our treatment, we shall not prove it, since it follows directly from (19) and (21).

If the incident wave part of the function $\chi(x, \gamma; E)$ has the direction η' and the value of τ is τ', we shall denote $\chi(x, \gamma; E)$ by $\chi(x, \gamma; E, \eta', \tau')$. We proceed to define the "spinor spherical wave" $\theta(x, \gamma; E, \eta, \tau)$ by

$$\theta(x, \gamma; E, \eta, \tau) = -\frac{Ee^{i|k||x|}}{2\pi |x|} \chi(\gamma; E, \eta, \tau). \tag{34}$$

The spinor spherical wave is analogous to the spinor plane wave (17). Furthermore, we define $T(E; \eta, \tau; \eta', \tau')$ by

$$\begin{aligned} T(E; \eta, \tau; \eta', \tau') &= \sum_{\gamma'} \int \chi(\gamma'; E, \eta, \tau)^* e^{-i|k|(x', \eta)} q\chi(x', \gamma'; E, \eta', \tau')\, dx' \\ &= \sum_{\gamma'} \int [q^*\chi(\gamma'; E, \eta, \tau)^*] e^{-i|k|(x', \eta)} \chi(x', \gamma'; E, \eta', \tau')\, dx'. \end{aligned} \tag{35}$$

The expression for χ_{sc} takes the form for large $|x|$,

$$\chi_{sc}(x, \gamma; E) = \sum_{\tau} \theta(x, \gamma; E, \eta_1, \tau) T(E; \eta_1, \tau; \eta', \tau'). \tag{36}$$

The scattered wave may be regarded as the sum of two spinor spherical waves, each being characterized by a different value of τ. The function $T(E; \eta_1, \tau_1; \eta_2, \tau_2)$ may be regarded as the amplitude of the spinor spherical wave in the direction η_1 and with $\tau = \tau_1$ when the incident spinor plane wave has as its direction of propagation η_2 and its value of τ is τ_2. We have therefore obtained χ_{sc} in a form where $T(E; \eta, \tau; \eta', \tau')$ is an inner product to which, as will now be shown, a variational principle can be applied.

In order to show how the variational principle discussed abstractly earlier may be used to find the amplitudes $T(E; \eta, \tau; \eta', \tau')$ we need only show what quantities are to be identified with the quantities appearing in equation (3).

Let us first consider $T(E; \eta, \tau; \eta', \tau')$ as being the symmetric inner product of two vectors. We identify $\chi(x, \gamma; E, \eta', \tau')$ with the vector y of equation, (3). We see that the vector a' is to be identified with $q^* \chi(\gamma; E, \eta, \tau)^* e^{-i|k|} (x, \eta)$. From (15) and (23) we construct the equation corresponding to the first of equations (3).

$$q\chi(\gamma; E, \eta', \tau')e^{i|k|(x, \eta')} = q\chi(x, \gamma; E, \eta', \tau') - q \sum_{\gamma'} \int g_r(x, \gamma; x', \gamma') q\chi(x', \gamma'; E, \eta', \tau')\, dx'. \tag{37}$$

The integral equation corresponding to the second equation (3) will also be given, since it is not difficult to obtain symmetric adjoint operator K'. Accordingly, the integral equation for Λ which is the function corresponding to the vector y' is

$$q^*\chi'(\gamma; E, \eta, \tau)^* e^{-i|k|(x,\eta)} = q^*\Lambda(x; \gamma; E, \eta, \tau) - q^* \sum_{\gamma'} \int g_s(x, \gamma; x', \gamma') q^* \Lambda(x', \gamma', E, \eta, \tau)\, dx' \tag{38}$$

where

$$g_s(x, \gamma; x', \gamma') = g_r(x', \gamma'; x, \gamma). \tag{39}$$

Just as the integral equation (37) for $\chi(x, \gamma; E, \eta, \tau)$ was derived from the differential equation

$$H\chi(x, \gamma; E, \eta, \tau) = E\chi(x, \gamma; E, \eta, \tau) \tag{40}$$

together with appropriate boundary conditions, the integral equation (38) for $\Lambda(x, \gamma; E; \eta, \tau)$ can be derived from a differential equation with certain boundary conditions. The differential equation satisfied by Λ is

$$H^*\Lambda(x, \gamma; E, \eta, \tau) = E\Lambda(x, \gamma; E, \eta, \tau) \tag{41}$$

so that Λ is an eigenfunction of H^*. The boundary condition on Λ is that it is to be expressed as the sum of an incident wave Λ_{in} and a scattered wave Λ_{sc} such that Λ_{in} satisfies

$$H_0^*\Lambda_{in}(x, \gamma; E, \eta, \tau) = E\Lambda_{in}(x, \gamma; E, \eta, \tau) \tag{42}$$

and Λ_{sc} behaves like the spherical wave $(e^{ik}|x|)/|x|$ for large values of $|x|$.

We take as a suitable solution Λ_{in} of (42) the function χ_{in}^* where χ_{in} is given by (17.) The function Λ_{sc} as can be shown from (40), satisfies the equation

$$\Lambda_{sc}(x, \gamma; E) = \sum_{\gamma'} \int g_s(x, \gamma; x', \gamma') q^* \Lambda(x', \gamma'; E)\, dx' \tag{43}$$

where the inverse operator $[E - H_0^*]^{-1}$ is represented by an integral operator with the kernel g_s.

The functional corresponding to $\lambda(v', v)$ whose extremal value is $1/[T(E; \eta, \tau; \eta', \tau')]$ is written as $1/T\{v', v\}$ and is given by

$$\frac{1}{T}\{v', v\} = \frac{\sum_{\gamma} \int v'(x, \gamma)\{qv(x, \gamma) - q\sum_{\gamma'} \int g_r(x, \gamma; x', \gamma') qv(x', \gamma')\, dx'\}\, dx}{\left(\sum_{\gamma} \int v'(x, \gamma) q\chi(\gamma; E, \eta', \tau') e^{i|k|(x,\eta')}\, dx\right)\left(\sum_{\gamma} \int q^*\chi(\gamma; E, \eta, \tau)^* e^{-i|k|(x,\eta)} v(x, \gamma)\, dx\right)} \tag{44}$$

where v' and v are the trial functions which approximate $\Lambda(x, \gamma; E, \eta, \tau)$ and $\chi(x, \gamma; E, \eta', \tau')$ respectively.

Having considered the case where $T(E; \eta, \tau; \eta', \tau')$ is a symmetric inner product, we shall now discuss the case in which this amplitude is considered a Hermitian inner product of the two vectors $q\ \chi(\gamma; E, \eta, \tau)\ e^{i}\ |\ k\ |\ (\eta, x)$ and $\chi(x, \gamma; E, \eta', \tau')$. The identification of the vectors a and y and of the operator K are as before in the case of the symmetric inner product. The vector a' is now identified with $q\ \chi(\gamma; E, \eta, \tau)e^{i|k|(\eta,x)}$. The vector y' is identified with the function $\Omega(x, \gamma; E, \eta, \tau)$ which satisfies the following integral equation which corresponds to the second of equations (3).

$$q\chi(\gamma'; E, \eta, \tau)e^{i|k|(\eta,x)} = q\Omega(x, \gamma; E, \eta, \tau) - q \sum_{\gamma'} \int g_t(x, \gamma; x', \gamma')q\Omega(x', \gamma'; E, \eta, \tau)\, dx' \tag{45}$$

where

$$g_t(x, \gamma; x', \gamma') = g_r(x', \gamma'; x, \gamma)^* = -\left\{\frac{[|\ k\ |\ |\ x - x'\ | - i]}{|\ x - x'\ |^2} \sum_{i=1}^{3} \alpha_i(\gamma, \gamma')(x_i - x'_i) - m\beta(\gamma, \gamma') + E\ \delta(\gamma, \gamma')\right\} \frac{e^{-i|k||x-x'|}}{4\pi\ |\ x - x'\ |}. \tag{46}$$

It can be shown that $\Omega(x, \gamma; E, \eta, \tau)$ as a solution of a differential equation with suitable boundary conditions. The function $\Omega(x, \gamma; E, \eta, \tau)$ satisfies the differential equation

$$H\Omega(x, \gamma; E, \eta, \tau) = E\Omega(x, \gamma; E, \eta, \tau) \tag{47}$$

so that Ω, like χ, is an eigenfunction of H. However, the function Ω has different boundary conditions than χ. The boundary condition on Ω is that it should be expressed as the sum of an incident part Ω_{in} and a "concentrating" part Ω_{con}. The function Ω_{in} satisfies

$$H_0\Omega_{in}(x, \gamma; E, \eta, \tau) = E\Omega_{in}(x, \gamma; E, \eta, \tau) \tag{48}$$

and is taken to be χ_{in} as given by (17). The function Ω_{con} is specified by the condition that Ω_{con} is to behave like an *inwardly* moving spherical wave for large values of $|\ x\ |$. Using the boundary condition that Ω_{con} represent an inward moving spherical wave for large values of $|\ x\ |$, the inverse operator $[E - H_0]^{-1}$ can be expressed as an integral operator while the function g_t is the solution of equation (24) expressed in the form (28) when the solution s of (26) is taken to be

$$s_t = -\frac{e^{-i|k||x-x'|}}{4\pi\ |\ x - x'\ |}$$

instead of s_r given by (29).

In the case of the Hermitian inner product the functional $1/T(v',v)$ is given by

$$\frac{1}{T}(v', v) = \frac{\sum_{\gamma} \int v'(x, \gamma)^* \{qv(x, \gamma) - q \sum_{\gamma'} \int g_r(x, \gamma; x', \gamma')qv(x', \gamma')\, dx'\}\, dx}{\left(\sum_{\gamma} \int v'(x, \gamma)^* q\chi(\gamma; E, \eta', \tau')e^{i|k|(x,\eta')}\, dx\right)\left(\sum_{\gamma} \int q^*\chi(\gamma; E, \eta, \tau)^* e^{-i|k|(x,\eta)}v(x, \gamma)\, dx\right)} \tag{49}$$

where v' and v approximate

$$\Omega(x, \gamma; E, \eta, \tau) \qquad \text{and} \qquad \chi(x, \gamma; E, \eta', \tau'),$$

respectively.

Acknowledgments

An extended version of this note appeared as a report prepared under Navy Contract N6ori-201 Task Order No. 1 by the Institute for Mathematics and Mechanics of New York University.

The author wishes to express his thanks to Professors K. O. Friedrichs and J. B. Keller for their criticisms.

A FORM OF NEWTON'S METHOD WITH CUBIC CONVERGENCE*

By W. M. STONE (*Boeing Airplane Co. and Oregon State College*)

For obtaining an approximation to a root of a transcendental equation $f(x) = 0$ Newton's method may well be unsatisfactory because it is only quadratically convergent, thus requiring considerable interpolation if the functions involved are scantily tabulated. On the other hand, the formulas of Stewart [1], Hamilton [2], Bodewig [3], and others, which offer cubic or higher convergence have the serious drawback of requiring the evaluation of second or higher derivatives. Formula (4) below, based on the generalized Taylor expansion of Hummel and Seebeck [4], offers cubic convergence in terms of $f(x)$ and $f'(x)$ evaluated at points on each side of the root.

Taking $n = m$ in the Hummel-Seebeck expansion,

$$f(x) = f(a) + (x - a)\frac{f'(a) + f'(x)}{2} + (x - a)^2 \frac{f''(a) - f''(x)}{12} + \cdots , \tag{1}$$

we obtain two approximations to a root,

$$x - a = \frac{-2f(a)}{f'(a) + f'(x)} \qquad \text{and} \qquad x - b = \frac{-2f(b)}{f'(b) + f'(x)}. \tag{2}$$

We choose a and b so $f(a)$ and $f(b)$ are opposite in sign, $f'(a)$ and $f'(b)$ same sign. Elimination of $f'(x)$ in equations (2) yields

$$\frac{f(b)}{x - b} - \frac{f(a)}{x - a} + \frac{f'(b) - f'(a)}{2} = 0 \tag{3}$$

or, by an obvious procedure,

$$x = \frac{b + a}{2} - \frac{f(b) - f(a)}{f'(b) - f'(a)} \pm \left\{\left[\frac{b - a}{2} + \frac{f(b) - f(a)}{f'(b) - f'(a)}\right]^2 - \frac{2(b - a)f(b)}{f'(b) - f'(a)}\right\}^{1/2}, \tag{4}$$

where choice of the ambiguous sign is quite obvious.

*Received April 22, 1952.