Perturbation Methods for Dirac Radial Equations*

G. BREIT AND G. E. BROWN
Yale University, New Haven, Connecticut
(Received July 18, 1949)

Two methods of calculating perturbation effects for the Dirac electron equation are discussed. The two radial functions enter symmetrically in the main contributions to the effects calculated. The first method is related to the method of the Riccati equation for ordinary second-order differential equations but differs from it through the use of the ratio of the two radial functions rather than that of the logarithmic derivative of one of them. The second method has some similarity to Milne's phase amplitude method. The methods described have been helpful in the two preceding papers.

I. INTRODUCTION[1,2]

THE Schrödinger single particle wave equation can often be conveniently treated by introducing the logarithmic derivative of the radial wave function. Perturbation calculations can at times be made very conveniently by this method. Useful results for nuclear scattering theory[3,4] and for the JWKB method have been obtained in this manner. The general plan is as follows. The linear second-order equation is transformed into a non-linear first-order equation of the Riccati type, the dependent variable being the logarithmic derivative. Perturbation calculations can then be made conveniently. A special advantage of this method is that an estimate of the effect of an inaccuracy can be made by a simple quadrature. For the Dirac equation in a central field one deals with two simultaneous radial equations which connect the two radial functions. One of these functions can be eliminated and there results an ordinary second-order differential equation which can be dealt with in the same general way as in the Schrödinger case. Such a procedure has been used[5] and is feasible. It has certain disadvantages, however. In the first place the second-order differential equation for a Dirac radial function has to be reduced to a standard form by a transformation. Secondly the physically important radial functions are lost track of in the manipulations. It is clear on the other hand that the Rayleigh-Schrödinger perturbation calculus can be applied to this problem, that in general the perturbation Hamiltonian can then be represented by two-row square matrices with operators for elements and that the two radial functions enter explicitly in the calculations. It is a disadvantage of the method of elimination of one of the radial functions that expressions obtainable from it are not simply related to those of the Rayleigh-Schrödinger perturbation calculus. It will be shown below that equations of a simpler and more practical form can be obtained by a different choice of the dependent variable. Two methods will be described. In the first the variable $y=f/g$ will be used. Here f, g are the two radial functions in a rather usual choice of the Dirac matrices which will become directly obvious through the formulas below. The employment of the variable y yields convenient results whenever the change in y caused by the perturbation, which is here called δy, does not become too large in the region of distances r for which the calculation is made. Experience with a similar equation for the Schrödinger case[4] has shown that analytic approximations can be improved by employing step by step numerical constructions for taking into account non-linear terms in δy. The method requires special consideration close to a node of g, a case which arises often and leads to an infinite δy. In order to have available a treatment in which infinities do not occur a second perturbation method has been worked out and is reported on also. This is suggested by Milne's phase amplitude method for the Schrödinger equation.[6] The amplitude A is here $r(f^2+g^2)^{\frac{1}{2}}$ and the phase φ is introduced by $\tan\varphi=f/g$. The direct application of Milne's method would bring in variables different from those used here. The equations reported on have been found useful in the two preceding papers.[1,2]

The feature which is especially convenient in both phase amplitude methods is the replacement of oscillatory functions by the monotonic φ and the smoothly varying A.

II. THE LOGARITHMIC DERIVATIVE METHOD

The radial equations can be written in the form

$$\{\|\epsilon_{ij}\|(d/idr)+\|P_{ij}\|\}\chi=0, \quad (i=1,2;j=1,2) \quad (1)$$

where χ is a column matrix with two rows. The two matrix elements of χ are denoted by

$$\chi_1=F=rf, \quad \chi_2=G=rg \quad (1.1)$$

and the normalization integral is

$$\int_0^\infty (\chi_1^2+\chi_2^2)dr=1. \quad (1.2)$$

* Assisted by the Joint Program of the ONR and the AEC.
[1] The methods described in this paper have been helpful in Breit, Brown, and Arfken, Phys. Rev. 76, 1299 (1949).
[2] The methods described in this paper have been helpful in G. E. Brown and G. B. Arfken, Phys. Rev. 76, 1305 (1949).
[3] Breit, Thaxton, and Eisenbud, Phys. Rev. 55, 1018 (1939).
[4] The possibilities of step by step numerical constructions of perturbation effects by means of the Riccati equation have been successfully tried out by M. H. Hull at the suggestion of one of the writers in connection with the calculation of Coulomb functions.
[5] G. Breit and R. E. Meyerott, Phys. Rev. 72, 1023 (1947).
[6] W. E. Milne, Phys. Rev. 35, 863 (1930).

The relation of the other symbols in Eq. (1) to the scalar potential A_0, the electronic charge and mass $-e$, m is

$$\|P_{ij}\| = (E+eA_0+mc^2\rho_3)/\hbar c+ik\epsilon\rho_3/r, \quad (1.3)$$

in the notation of Dirac.[7] The quantum number k is the characteristic value of

$$\rho_3[(\mathbf{L}\sigma)+1],$$

while

$$\epsilon=\|\epsilon_{ij}\|=\rho_1(\sigma\mathbf{r})/r. \quad (1.4)$$

The matrices ρ_3 and ϵ/i are taken in the representation

$$\rho_3=\begin{pmatrix}1, & 0\\ 0, & -1\end{pmatrix}, \quad \epsilon/i=\begin{pmatrix}0, & -1\\ 1, & 0\end{pmatrix}. \quad (1.5)$$

Employing the abbreviation $p_0=(E+eA_0)/c$ the matrix elements of $\|P_{ij}\|$ are

$$P_{11}=(p_0+mc)/\hbar, \quad P_{12}=P_{21}=-k/r, \quad P_{22}=(p_0-mc)/\hbar. \quad (1.6)$$

The definitions made above describe the connection with standard theory. Changes in A_0 are seen to affect only the diagonal elements P_{11}, P_{22}. For the treatment of problems in which perturbation effects of the central potential are needed it suffices to consider the special case of a diagonal perturbation matrix

$$\|\delta P_{ij}\|=\|P_{ij}\|-\|P_{ij}{}^0\| \quad (1.7)$$

with c numbers as the perturbing elements.

In the problem of nuclear mass motion[1,2,5] and in other questions[8] there occur more general changes of the δP_{ij}. The calculations below will be made in part without any restriction on the δP_{ij} and in part with the specialization of all the δP_{ij} being c numbers. In the latter case the restrictions $\delta P_{12}=\delta P_{21}$, $P_{ij}=P_{ij}{}^*$ will be used. The functions F, G will be supposed real throughout.

Extensions of the general formulas to complex P_{ij} are obvious. The simplicity of reductions in the special case, just mentioned, seems to be lost, however, when extensions to complex χ_i are made. It may be noted that if the equations are not reducible to real χ_i, $y=F/G$ contains two functions rather than one and the application of the methods under discussion is necessarily more involved. It will be recalled that if the P_{ij} are c numbers then the Hermitian requirement for $\|P_{ij}\|$ is

$$\int_0^\infty \Sigma P_{ij}{}^*\chi_j{}^*\chi_i dr=\int_0^\infty \Sigma P_{ij}\chi_i{}^*\chi_j dr$$

and this can be satisfied for arbitrary χ_i, χ_j only if $P_{ij}{}^*=P_{ji}$. This condition reduces to $P_{ij}=P_{ji}$ if all the quantities are real. Only symmetric matrices $\|P_{ij}\|$ are, therefore, considered below. For differential operators P_{12} the requirement of symmetry cannot be used. An example is the Hermitian operator $(\epsilon/i)(d/dr)$.

It will now be supposed that there is available a solution of Eq. (1) for some set of values of the P_{ij} given by Eqs. (1.3), (1.6) with a suitable choice of the scalar potential. The quantities corresponding to this solution will be referred to as unperturbed quantities. The unperturbed matrix will be denoted by

$$\|P_{ij}{}^0\|\equiv\|p_{ij}\|. \quad (2)$$

The perturbed function satisfies Eq. (1) with a different set of matrix elements. The relation between the two sets of elements will be taken as in Eq. (1.7). Introducing

$$y=F/G, \quad y_0=F_0/G_0, \quad \delta y=y-y_0 \quad (2.1)$$

with the convention of denoting unperturbed quantities by subscripts or superscripts 0, one obtains

$$d(\delta y)/dr\equiv -G^{-2}\Sigma\chi_i(\delta P_{ij})\chi_j +\Sigma p_{ij}(G_0{}^{-2}\chi_i{}^0\chi_j{}^0-G^{-2}\chi_i\chi_j). \quad (2.2)$$

Here and below it will be understood that the summations are performed over both i and j. By straightforward calculation one can remove first-order effects from the right side of the above formula with the result

$$d(G_0{}^2\delta y)/dr=-(G_0{}^2/G^2)\Sigma\chi_i(\delta P_{ij})\chi_j-G_0{}^2p_{11}(\delta y)^2. \quad (3)$$

The last term on the right side is of the second order in the small quantity δy. The main effects are contained in the first term. To the first order of small quantities one has the approximation

$$d(G_0{}^2\delta y)/dr\cong-\Sigma\chi_i{}^0(\delta P_{ij})\chi_j{}^0. \quad (3.1)$$

If the boundary conditions are such that $G_0{}^2\delta y$ vanishes at $r=0$ and $r=\infty$ this formula implies with neglect of higher order effects,

$$\int_0^\infty \Sigma\chi_i{}^0(\delta P_{ij})\chi_j{}^0 dr=0 \quad (3.2)$$

which gives the Rayleigh-Schrödinger first-order perturbation formula for the energy. Account has to be taken of the fact that, in general, Eq. (3.2) cannot be satisfied unless E is changed from its unperturbed value.

By integration of Eq. (3.1) one obtains δy to the first order within an interval $r_1<r<r_2$ provided δy is known either at r_1 or r_2 and provided the nodes of G and G_0 do not interfere with the calculation. Having δy to the first order one also has first-order corrections to F and G because Eq. (1) gives

$$d(\delta G)/dr=[\delta(P_{11}y+P_{12})]G+(p_{11}y_0+p_{12})\delta G \quad (3.3)$$

so that $\delta G=G-G_0$ is determined for known $\delta(P_{11}y+P_{12})$ and G as the solution of an ordinary linear inhomogeneous differential equation. Similarly δF is determined by Eq. (1). An iteration procedure for the solution of Eq.

[7] P. A. M. Dirac, Proc. Roy. Soc. **A117**, 611 (1928).

[8] The method outlined proved useful in unpublished calculations made by one of the authors (G.B.) on the distortion of the electronic wave functions produced by the nuclear magnetic moment.

(1) is thus possible, in the general case. It has practical value only if the convergence is rapid.

In the special case of $P_{12}=P_{21}$ and of c numbers for the P_{ij} the consideration of δF and δG can be avoided. One finds in this case

$$-d(G_0^2\delta y)/dr=G_0^2(y^2\delta P_{11}+2y\delta P_{12}+\delta P_{22}) +G_0^2 p_{11}(\delta y)^2=\Sigma\chi_i^0(\delta P_{ij})\chi_j^0+2G_0(F_0\delta P_{11} +G_0\delta P_{12})\delta y+G_0^2 P_{11}(\delta y)^2. \quad (3.4)$$

In the latter form there are only known functions multiplying the unknown δy. The second and third terms are of second order of small quantities. The term $P_{11}(\delta y)^2$ contains the third-order effect $(\delta P_{11})(\delta y)^2$, since $P_{11}=p_{11}+\delta P_{11}$. The error committed by neglecting them can be readily estimated and either an iteration procedure or a numerical step by step construction is relatively simple. In this case the knowledge of y gives G and F directly from Eq. (1) because

$$dG/Gdr=P_{11}y+P_{12} \quad (3.5)$$

so that G is obtainable by a quadrature.

If the δP_{ij} are differential operators it is also possible to perform reductions similar to that of Eq. (3.4). The differentiations on the χ_i which occur in Eq. (3) bring in either dG/dr which is expressible as $P_{11}yG+P_{12}G$, which in turn contains dG/dr, dy/dr, y, G. The relation thus obtained can be solved for dG/dr resulting in an expression for dG/Gdr in y and dy/dr. Since $dF/Fdr=dy/ydr+dG/Gdr$ one is also able to deal with dF/Fdr. The iteration via F and G discussed in connection with Eq. (3.3) is thus not a necessary step and is only one of several possibilities.[9]

III. THE PHASE-AMPLITUDE METHOD

In order to avoid special considerations at the poles of y equations somewhat similar to those of the phase amplitude method of Milne have been worked out. The defining relations are

$$F=A\mathbf{s}, \quad G=A\mathbf{c}, \quad \mathbf{s}=\sin\varphi, \quad \mathbf{c}=\cos\varphi. \quad (4)$$

The quantities φ and A will be referred to as phase and amplitude respectively. The equations on A and φ can be put in various forms having their respective advantages. The forms listed below have been decided on with the view of classifying effects according to different orders of smallness. The relations between A and φ can be put in the form

$$\begin{aligned} -Ad\varphi/dr&=\Sigma u_i P_{ij}u_j A,\\ dA/dr&\equiv\Sigma u_i R_{ij}u_j A,\end{aligned} \quad (4.1)$$

with

$$u_1\equiv\mathbf{s}, \quad u_2=\mathbf{c} \quad (4.2)$$

and

$$\|R_{ij}\|=\begin{pmatrix}-P_{21}, & -P_{22}\\ P_{11}, & P_{12}\end{pmatrix}. \quad (4.3)$$

Adhering to the previously introduced meaning of superscript and subscript 0 as well as the symbol δ for difference between perturbed and unperturbed quantities one finds

$$-d(\delta\varphi)/dr=A^{-2}\Sigma\chi_i(\delta P_{ij})\chi_j +2Q_2'^0\delta_1-Q_1'^0\delta_2-Q_2'^0\delta_3, \quad (4.4)$$

where

$$\delta_1=\delta\varphi, \quad \delta_2=1-\cos2\delta_1, \quad \delta_3=2\delta_1-\sin2\delta_1, \quad (4.5)$$

$$Q_1'=Q_1\mathbf{C}_0+Q_2\mathbf{S}_0, \quad Q_2'=-Q_1\mathbf{S}_0+Q_2\mathbf{C}_0, \quad (4.6)$$

$$\mathbf{C}=\cos2\varphi, \quad \mathbf{S}=\sin2\varphi \quad (4.7)$$

and the superscript ^{0} on the Q' indicates that the quantities Q in Eq. (4.6) take the values Q^0. The Q are as follows

$$Q_1=(P_{22}-P_{11})/2, \quad Q_2=(P_{12}+P_{21})/2. \quad (4.8)$$

The notation

$$P=(P_{11}+P_{22})/2 \quad (4.9)$$

will also be used.

The quantities Q_i' are connected with the Q_i by an orthogonal transformation. The latter has coefficients which are determined by the known value φ_0 of φ in the zeroth approximation. The value of φ_0 does not change during the improvement of values of $\delta\varphi$. The orthogonal transformation can be conveniently remembered as corresponding to a rotation through an angle $2\varphi_0$. Removal of first-order terms on the right side of Eq. (4.4) gives

$$-A_0^{-2}d(A_0^2\delta\varphi)/dr =A^{-2}\Sigma\chi_i(\delta P_{ij})\chi_j-Q_1'^0\delta_2-Q_2'^0\delta_3 \quad (5)$$

which is analogous to Eq. (3). The quantities δ_2, δ_3 are of second and third order respectively. The first term on the right side of Eq. (5) is seen to be the main one for small perturbations. It does not become infinite at a node of G and is more convenient in this respect than the first term on the right side of Eq. (3). Iteration procedures for Eq. (5) are similar to those for Eq. (3). The special case which was considered for Eq. (3.4) leads to the following form

$$-A_0^{-2}d(A_0^2\delta\varphi)/dr=\delta P+\delta Q_1' +2(\delta Q_2')\delta_1-Q_1'\delta_2-Q_2'\delta_3. \quad (5.1)$$

On the right side of this equation the quantities Q' are completely known and the δ_j are definitely determined by $\delta\varphi$. For a sufficiently small $\delta\varphi$ the δ_j can be approximated by one or two terms of the power series in δ_1, which represent them. Iteration or step by step constructions are again practical.

It is clear from Eq. (4.1) that in a calculation of δA the matrix $\|R_{ij}\|$ plays a part similar to that of P in the calculation of $\delta\varphi$. Replacement of the P_{ij} by the R_{ij}

[9] Another possibility for treating higher order effects is the expansion of δy in a power series similar to that discussed by Breit, Thaxton, and Eisenbud (see reference 3) for the Schrödinger equation. Coefficients of the terms in this series are given by relations similar to Eqs. (9.1) of reference 3.

changes P into $(P_{12}-P_{21})/2$, Q_1 into Q_2, Q_2 into $-Q_1$, Q_1' into Q_2', Q_2' into $-Q_1'$. These relations will be helpful to those who wish to verify the calculations. It is found that

$$d(\delta \ln A)/dr = A^{-2}\Sigma\chi_i(\delta R_{ij})\chi_j - 2Q_1'^0\delta_1 - Q_2'^0\delta_2 + Q_1'^0\delta_3 \quad (6)$$

and for the special case of c numbers for the P_{ij} the corresponding formula is

$$d(\delta \ln A)/dr = \delta Q_2' - 2Q_1'\delta_1 - Q_2'\delta_2 + Q_1'\delta_3. \quad (6.1)$$

The calculation of changes in the amplitude A is thus reducible to the evaluation of simple expressions.

PHYSICAL REVIEW VOLUME 76, NUMBER 9 NOVEMBER 1, 1949

Electron-Nuclear Potential Fields from Hyperfine Structure

M. F. CRAWFORD AND A. L. SCHAWLOW
McLennan Laboratory, University of Toronto, Toronto, Canada

(Received July 26, 1949)

Theoretical predictions for the effects of nuclear size on hyperfine structure are compared with experimental data. New data show that isotope shifts of *ns* levels in Tl III are proportional to $\psi^2(0)$ as predicted. The fractional change in nuclear radius for the addition of two neutrons, determined from *s* electron shifts in Hg II, Tl III, Pb IV, is the same for the three elements. Köhler's results for Tl II are consistent with those for Tl III when allowance is made for the mutual screening of the valence electrons.

Corrections for the approximation of the perturbation method and for screening of inner electrons by the valence electron are considered and found to be appreciable. Even after applying the corrections, the fractional change in nuclear radius is smaller than expected if the charge is uniformly distributed throughout the nucleus and the volume is proportional to the mass.

Values of nuclear magnetic moments of thallium deduced from hyperfine structure measurements are compared with those measured by radiofrequency induction and found to be 15 percent lower. This discrepancy is removed by the correction for finite nuclear size assuming either a uniform charge distribution or a concentration of charge toward the surface of the nucleus. Thus both the isotope shift and the magnetic effect indicate that the electron-nuclear potential is consistent with a charge distribution of this form and that the non-electrical forces between electrons and nuclei are relatively small.

1. INTRODUCTION

A RELATIVISTIC theory of the effects of finite nuclear size, with a consequent departure from a Coulomb field, on the atomic energy levels has been formulated by Rosenthal and Breit[1,2] and Racah.[3] The first two authors have also considered the effect of a finite nucleus on the interaction between electrons and the nuclear magnetic moment. The first effect leads to an isotope shift, and the second to a correction for the nuclear magnetic moment deduced from hyperfine structure splittings. Previous comparisons[2,4,5] of the theory with experimental data on isotope shifts agreed only as to order of magnitude; in general, the theoretical and observed values differed by a factor of the order of three. In many of these cases the comparison was complicated by the mutual screening of several electrons outside closed shells and by inter-configuration perturbations. It is therefore desirable to consider unperturbed levels arising from one-electron configurations, preferably those of penetrating *s* electrons. Such configurations occur in the Hg II, Tl III, Pb IV sequence and their isotope shifts are analyzed here.

The analysis shows that significant information about electron-nuclear potential fields can be obtained from the absolute magnitude of the shifts. The corrections for the approximations in the perturbation method and for the screening of electrons in completed shells are found to be important.

Hitherto the effect of nuclear radius on the magnetic interaction between electrons and the nuclear spin has not been confirmed experimentally. Evidence confirming this effect is presented. Both the isotope shifts and the magnetic effect are more consistent with a uniform charge distribution in the nucleus than with a well-type potential for the electron-nuclear interaction.

2. ISOTOPE SHIFT

The isotope shift for a single *s* electron as derived by Rosenthal and Breit[1] using the perturbation method is

$$\Delta\delta W = \frac{4\pi R a_H^3 \psi^2(0)}{Z} \frac{1+\rho}{[\Gamma(2\rho+1)]^2} y_0^{2\rho} \frac{\Delta y_0}{y_0} \cdot B, \quad (1)$$

where R is the Rydberg constant,

$a_H = h^2/4\pi^2 me^2$ is the radius of the first Bohr orbit for hydrogen,
Z is the nuclear charge,
$\psi^2(0)$ is the square of the non-relativistic atomic wave function at the center of the nucleus,
$\rho = (1-Z^2\alpha^2)^{\frac{1}{2}}$,
$\alpha = 2\pi e^2/hc$ is the fine structure constant
$y_0 = 2Zr_0/a_H$ where r_0 is the radius of the nucleus,

[1] J. E. Rosenthal and G. Breit, Phys. Rev. **41**, 459 (1932).
[2] G. Breit, Phys. Rev. **42**, 348 (1932).
[3] G. Racah, Nature **129**, 723 (1932).
[4] P. Köhler, Zeits. f. Physik **113**, 306 (1939).
[5] S. Mrozowski, Phys. Rev. **61**, 605 (1942).