and from this, one sees easily that the cross section for deuteron formation without emission of radiation is of the order

$$\sigma \sim \frac{1}{\alpha^2}\frac{\hbar\alpha}{Mv_0}\left(\frac{MUb^2}{\hbar^2}\right)^2\frac{1}{(\alpha b)^7}, \qquad (21)$$

again obeying a $1/v_0$ law, but of negligible importance compared to the magnetic dipole capture process, unless there is a stable singlet level, extremely close to zero potential.

I wish to use this opportunity for expression of my gratitude to Professor J. R. Oppenheimer and Dr. Robert Serber for their generous interest and many valuable suggestions during the course of this work.

FEBRUARY 1, 1937 PHYSICAL REVIEW VOLUME 51

The Effect of Nuclear Motion in the Dirac Equation

Irving S. Lowen
New York University, Washington Square East, New York, N. Y.
(Received December 1, 1936)

Corrections to the Dirac equation, relativistic to the order v^2/c^2, due to the motion of the nucleus, for the case of an electron in the field of a heavy particle of mass M, are derived with the aid of the Breit two-body relativistic interaction, to the order m/M, m being the mass of the electron. A calculation of the value of the correction for a $1s$ electron gives a result in agreement with that obtained from the Schrödinger treatment to the order m/M.

1.

The Dirac equation is limited in its application to such problems as the treatment of the energy levels of hydrogen, by the fact that it is valid only for the case of one body moving in an external field. As a consequence, the application to the energy levels of a single electron moving around a nucleus is justified only to the extent that the electron may be considered to move in a fixed field of force while a discussion of the effect of the nuclear motion on the energy levels by means of this equation is not available. A discussion of the "Mitbewegung" has been given by Bechert and Meixner[1] making use of the relativistic two-body interaction of Breit[2] in which both the nucleus and electron are treated in the Pauli approximation. The original treatment of Breit for the two-body interaction, however, suggests an approach in which the effect of the nuclear motion introduces small perturbing terms in the Dirac equation. In the discussion below these terms are evaluated to the order v^2/c^2 and $(m/M)^2$ m, M, being the mass of the electron and heavy particle, respectively. The correction due to the terms in $(m/M)^2$ turns out to be smaller than the hyperfine structure term by a factor m/M and may therefore be discarded as unobservable. We are therefore left with the equations (11), (12) in which the terms may be taken to represent the effect of the mitbewegung to the order m/M and v^2/c^2.

Our procedure in obtaining the desired corrections is to eliminate from the two-body equation of Breit, the "small" components of the wave function by expressing them in terms of the "large" components in the usual way in which the Dirac equation is reduced in the Pauli approximation. The large and small components mentioned refer of course only to the dependence on the nuclear coordinates.

2.

As our starting point, we consider the equation[2]

$$\left\{\frac{E}{c}+\frac{Ze^2}{cr}+\frac{e}{c}V+\boldsymbol{\alpha}_1\cdot\mathbf{p}_1+\boldsymbol{\alpha}_2\cdot\mathbf{p}_2+\beta_1Mc+\beta_2mc -\frac{Ze^2}{2c}\left(\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r}+\frac{\boldsymbol{\alpha}_1\cdot\mathbf{r}\,\boldsymbol{\alpha}_2\cdot\mathbf{r}}{r^3}\right)\right\}\Psi=0, \qquad (1)$$

[1] Bechert and Meixner, Ann. d. Physik **22**, 525 (1935).
[2] G. Breit, Phys. Rev. **34**, 553 (1929).

where the subscript 1 refers to heavy particle of mass M charge Ze, subscript 2 referring to electron; $\mathbf{r}$ vector distance between the two particles pointing from 1 to 2. The α's are the usual Dirac's α's while the wave function Ψ has sixteen components, and $V=V_1-ZV_2$, V_1 and V_2 being potentials at particles one and two, respectively, due to the external field. Disregarding the last term temporarily, (1) becomes, putting

$$p_0=\frac{E}{c}+\frac{Ze^2}{cr}+\frac{e}{c}V$$

$$\{p_0+\boldsymbol{\alpha}_1\cdot\mathbf{p}_1+\beta_1Mc+\boldsymbol{\alpha}_2\cdot\mathbf{p}_2+\beta_2mc\}\Psi=0, \quad (2)$$

setting

$$\Psi=\begin{pmatrix}\psi_a\\ \psi_b\end{pmatrix},$$

where ψ_a consists of the "small" components and ψ_b of the "large" components of the wave functions, referring to the heavy particle, (2) breaks up into

$$(p_0+Mc+\boldsymbol{\alpha}_2\cdot\mathbf{p}_2+\beta_2mc)\psi_a-\boldsymbol{\sigma}_1\cdot\mathbf{p}_1\psi_b=0. \quad (3)$$

$$(p_0-Mc+\boldsymbol{\alpha}_2\cdot\mathbf{p}_2+\beta_2mc)\psi_b-\boldsymbol{\sigma}_1\cdot\mathbf{p}_1\psi_a=0, \quad (4)$$

where the σ's are the negatives of Pauli's and both ψ_a and ψ_b are four component functions.

From (3), we obtain to the order $(m/M)^2$ and $(v/c)^2$

$$\psi_a=-\frac{1}{4M^2c^2}(\boldsymbol{\alpha}_2\cdot\mathbf{p}_2+\beta_2mc)\boldsymbol{\sigma}_1\cdot\mathbf{p}_1\psi_b+(p_0+Mc)^{-1}\boldsymbol{\sigma}_1\cdot\mathbf{p}_1\psi_b.$$

Substituting this in (4), we obtain

$$\Big\{p_0-Mc+\boldsymbol{\alpha}_2\cdot\mathbf{p}_2+\beta_2mc+\frac{1}{4M^2c^2}(\boldsymbol{\alpha}_2\cdot\mathbf{p}_2+\beta_2mc)\boldsymbol{\sigma}_1\cdot\mathbf{p}_1\boldsymbol{\sigma}_1\cdot\mathbf{p}_1-\boldsymbol{\sigma}_1\cdot\mathbf{p}_1(p_0+Mc)^{-1}\boldsymbol{\sigma}_1\cdot\mathbf{p}_1\Big\}\psi_b=0. \quad (5)$$

The last term may be simplified as follows:

$$-\boldsymbol{\sigma}_1\cdot\mathbf{p}_1(p_0+Mc)^{-1}\boldsymbol{\sigma}_1\cdot\mathbf{p}_1=-\frac{1}{2Mc}\left(p_1^2+\frac{Ze\hbar}{c}\boldsymbol{\sigma}_1\cdot\mathbf{H}_1\right)+\frac{p_0-Mc}{4M^2c^2}\left(p_1^2+\frac{Ze\hbar}{c}\boldsymbol{\sigma}_1\cdot\mathbf{H}_1\right)-\frac{Zei\hbar}{4M^2c^3}(\mathbf{E}_1\cdot\mathbf{p}_1-i\boldsymbol{\sigma}_1\cdot\mathbf{E}_1\times\mathbf{p}_1),$$

$\mathbf{E}_1$ and $\mathbf{H}_1$ being the electric and magnetic intensities, use having been made of the following:

$$(p_0+Mc)^{-1}=\frac{1}{2Mc}\left(1+\frac{p_0-Mc}{2Mc}\right)^{-1}=\frac{1}{2Mc}\left(1-\frac{p_0-Mc}{2Mc}\right) \quad (6)$$

to terms of the order $(v/c)^2$, and of the result

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B})=\mathbf{A}\cdot\mathbf{B}-i\boldsymbol{\sigma}_1\cdot\mathbf{A}\times\mathbf{B}, \quad (7)$$

where $\mathbf{A}$, $\mathbf{B}$ are two vectors commuting with $\boldsymbol{\sigma}$.

Substituting in (5) and solving for p_0-Mc we obtain

$$\Big\{p_0-Mc+\boldsymbol{\alpha}_2\cdot\mathbf{p}_2+\beta_2mc-\frac{1}{2Mc}\left(p_1^2+\frac{Ze\hbar}{c}\boldsymbol{\sigma}_1\cdot\mathbf{H}_1\right)+\frac{p_1^4}{8M^3c^3}-\frac{Zei\hbar}{4M^2c^3}(\mathbf{E}_1\cdot\mathbf{p}_1-i\boldsymbol{\sigma}_1\cdot\mathbf{E}_1\times\mathbf{p}_1)\Big\}\psi_b=0. \quad (8)$$

We must now reduce the term

$$-\frac{Ze^2}{2c}\left[\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r}+\frac{\boldsymbol{\alpha}_1\cdot\mathbf{r}\,\boldsymbol{\alpha}_2\cdot\mathbf{r}}{r^3}\right]=-\frac{Ze^2}{2c}X$$

in a similar manner. Following a procedure of Breit's[3] instead of eliminating the small components from this term by the above procedure, it is convenient to calculate its average value and replace the small components in the integral for the average by their values in terms of the large. For this we require to calculate the integral of

[3] G. Breit, Phys. Rev. **39**, 616 (1932).

$$\overbrace{\psi_a^*\psi_b^*}\left[\frac{\alpha_1\cdot\alpha_2}{r}+\frac{(\alpha_1\cdot\mathbf{r})(\alpha_2\cdot\mathbf{r})}{r^3}\right]\begin{pmatrix}\psi_a\\ \psi_b\end{pmatrix}=$$

$$\overbrace{\psi_a^*\psi_b^*}\left[-\frac{\sigma_1\cdot\alpha_2}{r}-\frac{(\sigma_1\cdot\mathbf{r})(\alpha_2\cdot\mathbf{r})}{r^3}\right]\begin{pmatrix}\psi_b\\ \psi_a\end{pmatrix}=$$

$$-[\psi_a^*Q\psi_b+\psi_b^*Q\psi_a],$$

where $$Q=\frac{\sigma_1\cdot\alpha_2}{r}+\frac{(\sigma_1\cdot\mathbf{r})(\alpha_2\cdot\mathbf{r})}{r^3},$$

over the coordinate space of both particles.

Making use of (4a) this becomes

$$\psi_b\left\{\frac{1}{4M^2c^2}[\sigma_1\cdot\mathbf{p}_1(\alpha_2\cdot\mathbf{p}_2+\beta_2mc)Q+Q(\alpha_2\cdot\mathbf{p}_2+\beta_2mc)\sigma_1\cdot\mathbf{p}_1]-\sigma_1\cdot\mathbf{p}_1(p_0+Mc)^{-1}Q+Q(p_0+Mc)^{-1}\sigma_1\cdot\mathbf{p}_1\right\}\psi_b.$$

The first bracket may be simply transformed by exchanging $\alpha_2\cdot\mathbf{p}_2+\beta_2mc$ with $\sigma_1\cdot\mathbf{p}_1$ so that in the first term it finally operates on ψ_b^* and in the second on ψ_b. Furthermore, one can substitute in the first bracket to a sufficient degree of approximation

$$(\alpha_2\cdot\mathbf{p}_2+\beta_2mc)\psi_b=-(p_0-Mc)\psi_b.$$

Thus we have for $\bar{X}$ at length, omitting the integral sign

$$\bar{X}=-\frac{1}{2Mc}\psi_b^*(\sigma_1\cdot\mathbf{p}_1Q+Q\sigma_1\cdot\mathbf{p}_1)\psi_b+\frac{Zeih}{4M^2c^3}\psi_b^*(Q\sigma_1\cdot\mathbf{E}_1-\sigma_1\cdot\mathbf{E}_1Q)\psi_b,\quad(9)$$

where use has been made of (6).

Using (7) the first bracket is transformed as follows:

$$-\frac{1}{2Mc}\psi_b^*[\sigma_1\cdot\mathbf{p}_1Q+Q\sigma_1\cdot\mathbf{p}_1]\psi_b=\frac{1}{Mc}\psi_b^*\left[\frac{\alpha_2\cdot\mathbf{p}_1}{r}+h\sigma_1\cdot\frac{\mathbf{r}}{r^3}\times\alpha_2-\alpha_2\cdot\mathbf{r}\frac{\mathbf{r}}{r^3}\cdot\mathbf{p}_1+\frac{ih}{2}\left\langle\nabla\cdot\frac{\mathbf{r}}{r^3}\right\rangle\alpha_2\cdot\mathbf{r}\right]\psi_b.\quad(10)$$

The term containing $\nabla\cdot\mathbf{r}/r^3$ is zero even for s states on account of the factor $\alpha_2\cdot\mathbf{r}$ and may therefore be disregarded. This may be verified by a direct calculation for the special case of a Coulomb field which is the application with which we shall be concerned.

By a similar procedure, the second bracket of (9) can be reduced and when combined with (10) gives at length

$$X=-\frac{1}{Mc}\left[\frac{\alpha_2}{r}\cdot\mathbf{p}_1-h\sigma_1\cdot\frac{\mathbf{r}}{r^3}\times\alpha_2+\alpha_2\cdot\mathbf{r}\frac{\mathbf{r}}{r^3}\cdot\mathbf{p}_1\right]+\frac{Zeh}{M^2c^3}\left[\sigma_1\cdot\mathbf{E}_1\times\frac{\alpha_2}{r}+\frac{\sigma_1\cdot\mathbf{E}_1\times\mathbf{r}}{r^3}\alpha_2\cdot\mathbf{r}\right].\quad(10a)$$

Consequently the complete Hamiltonian may be written, in ascending order of magnitude in m/M

$$H=H_a+H_b+H_c,$$

where

$$\frac{H_a}{c}=-\left(\frac{Ze^2}{cr}+\alpha_2\cdot\mathbf{p}_2+\beta_2mc+\frac{e}{c}V\right),$$

$$\frac{H_b}{c}=\frac{1}{2Mc}\left(p_1^2+\frac{Zeh}{c}\sigma_1\cdot\mathbf{H}_1\right)-\frac{Ze^2}{2Mc^2}\left(\frac{\alpha_2}{r}\cdot\mathbf{p}_1-h\sigma_1\cdot\frac{\mathbf{r}}{r^3}\times\alpha_2+\frac{\alpha_2\cdot\mathbf{r}\mathbf{r}\cdot\mathbf{p}_1}{r^3}\right),$$

$$\frac{H_c}{c}=\frac{Zeih}{4M^2c^3}(\mathbf{E}_1\cdot\mathbf{p}_1-i\sigma_1\cdot\mathbf{E}_1\times\mathbf{p}_1)+\frac{Z^2e^3h}{2M^2c^4}\left(\sigma_1\cdot\mathbf{E}_1\times\frac{\alpha_2}{r}+\frac{\sigma_1\cdot\mathbf{E}_1\times\mathbf{r}}{r^3}\alpha_2\cdot\mathbf{r}\right).$$

These results may be checked against those of Bechert and Meixner by eliminating the small components of the electron wave function in H_a and H_b in the usual way. We reduce then to the results 5, 6, 7 of their paper.

3.

We apply the preceding result to the calculation of the energy levels of a single electron in a Coulomb field. We may gain some idea of the relative importance of the various terms by comparing them with the term of H_b

$$\frac{Ze^2}{2Mc^2}\hbar\boldsymbol{\sigma}_1\cdot\frac{\mathbf{r}}{r^3}\times\boldsymbol{\alpha}_2=-\frac{e}{c}\boldsymbol{\mu}\cdot\frac{\mathbf{r}}{r^3}\times\boldsymbol{\alpha}_2,$$

$\boldsymbol{\mu}$ being the magnetic moment of the heavy particle, to which is due the hyperfine structure. Simple considerations show that the terms in the second bracket of H_c are of the order $(m/M)\alpha^2Z^2$ while those in the first are $(m/M)\alpha Z$ in terms of the h.f.s. term. The contribution of H_c may therefore be dropped in applications to the derivation of the term values of a single electron in a central field. The terms in H_b on the other hand are either of the order of magnitude of the hyperfine structure term or differ from it by a factor of at most $1/\alpha Z$.

For application to s terms, these considerations require a slight refinement since H_c then becomes infinite. The more exact calculations result in the replacement of the factors M^2c^2 in the denominators of the terms in H_c by $(p_0+Mc)^2$. Since the chief contribution of these terms to the energy comes from a region $r=e^2/Mc^2\sim10^{-16}$ cm their meaning is doubtful unless we cut the wave functions off at about r equal to 10^{-13} cm. In this case also the terms in H_c may be dropped as being negligible. We take as the terms of our Hamiltonian for the one electron case in the Coulomb field of the nucleus assuming zero external magnetic field

$$\frac{H_a}{c}=-\boldsymbol{\alpha}\cdot\mathbf{p}-\beta mc-\frac{Ze^2}{rc}-\frac{e}{c}V,$$

$$\frac{H_b}{c}=\frac{p^2}{2Mc}-\frac{e}{c}\boldsymbol{\mu}\cdot\frac{\mathbf{r}}{r^3}\times\boldsymbol{\alpha}+\frac{Ze^2}{2Mc^2}\left(\frac{\boldsymbol{\alpha}}{r}\cdot\mathbf{p}+\boldsymbol{\alpha}\cdot\mathbf{r}\frac{\mathbf{r}}{r^3}\cdot\mathbf{p}\right),$$

where the substitution $\mathbf{p}=\mathbf{p}_1=-\mathbf{p}_2$ has been made. The subscripts have been omitted on the right since all the terms refer now to the electron.

We calculate the contribution of H_b to the energy for an s state first transforming to polar coordinates. In these coordinates we have[4] in the absence of an external field

$$\frac{H_a}{c}=-\epsilon p_r-i\epsilon\rho_3\frac{kh}{r}-\rho_3 mc-\frac{Ze^2}{rc},\tag{11}$$

$$\frac{H_b}{c}=\frac{1}{2Mc}\left(p_r^2-\frac{kh^2}{r^2}\rho_3+\frac{k^2h^2}{r^2}\right)+\frac{Ze^2}{2Mc^2}\left(\frac{2\epsilon}{r}p_r+i\epsilon\rho_3\frac{kh}{r^2}+\frac{i\epsilon h}{r^2}\right)\tag{12}$$

with the omission of the h.f.s. term, and the same definition of k as in Dirac. Thus for $s_{\frac{1}{2}}$, $p_{\frac{1}{2}}$, $p_{\frac{3}{2}}$, $p_{\frac{3}{2}}$ states, k has values -1, 1, -2, respectively.

Use has been made here of the relations

$$r\epsilon=\boldsymbol{\alpha}\cdot\mathbf{r},\tag{13a}$$

$$\mathbf{r}\cdot\mathbf{p}=rp_r+ih,\tag{13b}$$

$$\boldsymbol{\alpha}\cdot\mathbf{p}=\epsilon p_r+i\epsilon\rho_3(kh/r),\tag{13c}$$

where $\epsilon=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}$, $\rho_3=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$.

Finally by squaring (13c) we obtain

$$p^2=p_r^2-\frac{kh^2}{r^2}\rho_3+\frac{k^2h^2}{r^2}.$$

In terms of radial functions, the average value of H_b is given by

$$\Delta W=\frac{1}{2Mc^2}\int_0^\infty\left\{[E^2-(eA_0)^2+m^2c^4](\phi_1^2+\phi_2^2)+2Emc^2(\phi_1^2-\phi_2^2)+2Ze^2c\frac{kh}{r^2}\phi_1\phi_2\right\}dr,\tag{14a}$$

$$=\frac{h}{2Mc}\int_0^\infty\left\{(E-eA_0)\left(\phi_1\frac{d\phi_2}{dr}-\phi_2\frac{d\phi_1}{dr}\right)+2E\frac{k}{r}\phi_1\phi_2\right\}dr,\tag{14b}$$

where ϕ_1 and ϕ_2 satisfy the equations

$$\frac{d\phi_1}{dr}-\frac{k}{r}\phi_1+\frac{1}{hc}(E+eA_0-mc^2)\phi_2=0,$$

$$\frac{d\phi_2}{dr}+\frac{k}{r}\phi_2-\frac{1}{hc}(E+eA_0+mc^2)\phi_1=0,$$

where $eA_0=Ze^2/r$.

ϕ_1 and ϕ_2 are subject to the normalizing condition

$$\int_0^\infty(\phi_1^2+\phi_2^2)dr=1.$$

[4] Dirac, *Quantum Mechanics*, second edition, p. 266.

As a check on the calculations, it is convenient to use

$$\left(\frac{p^2}{2M}\right)=\frac{1}{2Mc^2}\int_0^\infty[(E+eA_0+mc^2)^2\phi_1^2+(E+eA_0-mc^2)^2\phi_2^2]dr.$$

The first form (14) is convenient if one does not desire derivatives in the expression, the second is more compact and has the additional advantage that for 1*s* states

$$\phi_1\frac{d\phi_2}{dr}-\phi_2\frac{d\phi_1}{dr}=0.$$

For 1*s* states these expressions give

$$\Delta W=\frac{m}{M}\frac{mc^2}{2}\alpha^2Z^2.$$

This relativistic result is to be compared with what one might expect from naive considerations: $\Delta W=(m/M)mc^2(1-(1-\alpha^2Z^2)^{\frac{1}{2}})$. In the case of the former one has for $Z=90$, a correction $\Delta W=0.215(m/M)mc^2$ while the latter gives $.245\ (m/M)\ mc^2$. The strict relativistic treatment therefore gives an appreciable effect.

We close this section by giving the Mitbewegung corrections for the case of two electrons revolving around a nucleus. If we refer to the electrons by the subscripts 1 and 2, we have in place of H_b above, for two electrons, omitting the h.f.s. term

$$\frac{H_b'}{c}=\frac{(\mathbf{p}_1+\mathbf{p}_2)^2}{2Mc}+\frac{Ze^2}{2Mc^2}\left[\left(\frac{\boldsymbol{\alpha}_1}{r_1}\right)\cdot(\mathbf{p}_1+\mathbf{p}_2)+\left(\frac{\boldsymbol{\alpha}_1\cdot\mathbf{r}_1}{r_1^3}\right)\mathbf{r}_1\cdot(\mathbf{p}_1+\mathbf{p}_2)\right]+\frac{Ze^2}{2Mc^2}\left[\left(\frac{\boldsymbol{\alpha}_2}{r_2}\right)\cdot(\mathbf{p}_1+\mathbf{p}_2)+\left(\frac{\boldsymbol{\alpha}_2\cdot\mathbf{r}_2}{r_2^3}\right)\mathbf{r}_2\cdot(\mathbf{p}_1+\mathbf{p}_2)\right],$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are vector distances pointing from the nucleus to the electrons 1 and 2, respectively. From this the generalization to any number of electrons may be simply made.

4.

Since in the derivation, Eq. (1) from which we start presupposes the existence of an equation of Dirac type for both particles and since present data on nuclear moments seems to indicate that the heavy particles are not described by such an equation, the question naturally arises as to the extent to which our results are limited by this condition. The answer to this question may be based on two circumstances. In the first place, the effect of the magnetic moment of the heavy particle is represented in our result by the term in (10*a*)

$$\frac{e}{c}\boldsymbol{\alpha}_2\cdot\frac{\boldsymbol{\mu}\times\mathbf{r}}{r^3},\quad\text{where}\quad\boldsymbol{\mu}=-\frac{Ze\hbar}{2Mc}\boldsymbol{\sigma},$$

which is *formally* the same as that obtained from the single electron equation for a nuclear magnetic moment $\boldsymbol{\mu}$ so that by giving $\boldsymbol{\mu}$ appropriate values the h.f.s. is correctly given, insofar as it be considered to arise from a magnetic doublet at the heavy particle, by this calculation even though the heavy particle does not obey a Diracian equation. Presumably, then, the other terms which do not even involve the magnetic moment of the heavy particle are correct.

The other justification for believing our result to be correct to the stated degree of approximation is due to the fact that (1) may be derived in the case of two electrons by means of the quantum electrodynamics.[3] This derivation should hold also in case one of the particles is moving sufficiently slowly to be treated in the Pauli approximation in which case the applicability of the Dirac formalism to this particle would not then be a necessary prerequisite to the validity of the resulting interaction.

In conclusion it is a pleasure to acknowledge my indebtedness to Professor G. Breit for most generous advice and discussions. I should also like to express my thanks to the physics department of the University of Wisconsin for its kind hospitality during the past summer when part of this work was done.