Effect of Nuclear Motion of the Hyperfine Structure of the Ground Term of Hydrogen

G. BREIT AND R. E. MEYEROTT
Yale University, New Haven, Connecticut*

(Received August 19, 1947)

The correction factor for the effect of nuclear motion on the hyperfine structure of hydrogen is discussed. It is found that this factor can be represented by $(1+m/M)^{-3}$ to within terms of order $(m/M)\alpha^2 \log\alpha$, where m, M are, respectively, the masses of the electron and nucleus while α is the fine structure constant. It is assumed that the Coulomb potential is that of a point charge for distances greater than $r_0=e^2/mc^2$ and that for distances smaller than r_0 it is of the order mc^2. This assumption makes it possible to treat the problem by means of existing theories. First-order perturbation theory for the effect of the nuclear magnetic field is employed. The reasons for doing the work are explained in the introduction. The calculations for the part of the proton's magnetic moment following from Dirac's equation are described in Section II. Section III is concerned with the effect of the part of the magnetic moment of the proton which is not accounted for by Dirac's equation and is referred to as the Pauli part. The deuteron is also discussed in Section III.

I. INTRODUCTION

THE recent results[1] on the hyperfine structure of hydrogen make it desirable to make sure of the correction for the nuclear mass motion to the theoretically expected formula for the hyperfine structure splitting. Fermi's result[2] for the energy splitting of s terms of single-electron spectra for a stationary nucleus is proportional to the product of the magnetic moments of the electron and the nucleus and contains besides, as a factor, the square of the non-relativistic Schrödinger function at the stationary nucleus $\psi_S^2(0)$. The latter quantity has the dimensions of the cube of a reciprocal length. Together with the product of the magnetic moments it gives a dimensionally correct combination for the expression of an energy. The quantity $\psi_S^2(0)$ is for hydrogen $1/\pi a_H^3 n^3$, where a_H is the Bohr radius. Its appearance suggests that to within a constant factor it replaces the mean of the reciprocal electron-nucleus distance and that for a nucleus of finite rather than infinite mass the formula for the energy splitting has to be modified by the introduction of the factor

$$(1+m/M)^{-3},$$

which represents the reciprocal of the cube of the ratio of the spatial extension of the wave function for finite and infinite mass. Here m, M stand, respectively, for the masses of the electron and the nucleus. It is in fact well-known that for the non-relativistic quantum-mechanical two-body problem the unnormalized wave function is obtainable from that of the one-body problem of mass m with $M=\infty$ by stretching a graph of the latter in such a way as to increase all interparticle distances in the ratio $1+m/M$, which is the reciprocal of the corresponding ratio of reduced masses.

The above simple arguments make it plausible to assume that the inverse cube of the reduced mass ratio represents the correction factor to Fermi's formula apart from typically relativistic corrections.[3] A partial confirmation of this view has been obtained previously in connection with estimates of the magnetic energy contribution to the binding energy of the deuteron and is referred to below in connection with Eq. (6.1) of the present paper. It was found that the non-relativistic limit for this energy is obtained by employing Fermi's formula with the interpretation of $\psi_S^2(0)$ as the square of the non-relativistic Schrödinger function for the relative motion of proton and neutron. By the non-relativistic limit one means here the asymptotic form in the limit $c=\infty$, where c is the velocity of light.

The arguments given above are not as con-

* Assisted by Contract N6ori-44, Task Order XVI of the Office of Naval Research.

[1] J. E. Nafe, E. B. Nelson, and I. I. Rabi, Phys. Rev. **71**, 914 (1947).

[2] E. Fermi, Zeits. f. Physik **60**, 320 (1930).

[3] G. Breit, Phys. Rev. **35**, 1447 (1930).

vincing, however, as is desirable. The accuracy of the experiments at Columbia is very high, and there is a need for a careful analysis of the effects neglected in the inverse cube of reduced mass-correction factor. The arguments for the employment of this factor have, in fact, a number of defects. In the first place the mean of the inverse cube of the interparticle distance diverges in the non-relativistic limit for s terms. The simple argument based on the similarity transformation which has been given above is, therefore, not correct even though there is no doubt that it has something to do with the true state of things. Secondly, the considerations concerning the magnetic energy of proton and neutron in the deuteron have to do only with the special case of particles of equal masses. Besides, they have been carried out on the assumption that the interaction energy between the proton and neutron is of a non-singular type. Characteristic effects of the inverse-square field of force might have been overlooked. Thirdly, the order of magnitude of what is neglected in the $(1+m/M)^{-3}$ correction factor has not been previously estimated.

The calculations reported on below are reassuring concerning the validity of the simple correction factor. They show that the effects neglected are of the order $(m/M)\alpha^2\log\alpha$, where α is the fine structure constant. In obtaining this result it was assumed that the nucleus has a finite radius of the order e^2/mc^2, where e is the electronic charge. For distances smaller than this amount the inverse-square law potential was supposed to be inapplicable. The calculations would have failed if the inverse-square law of force were supposed to apply at all distances. Some of the integrals entering the answer would have diverged logarithmically. The starting point for the calculation would also be an uncertain one if the assumption of a finite radius were not made. By making the potential finite rather than infinite at small distances the velocity of the nucleus is kept small, and the discussion by means of a Hamiltonian of known type is possible.

II. EFFECT FOR PROTON OBEYING DIRAC'S EQUATION

A. General Assumptions

The wave equation will be taken to be

$$(p_0+\boldsymbol{\alpha}_e\mathbf{p}_e+\boldsymbol{\alpha}_M\mathbf{p}_M+\beta_e mc+\beta_M Mc-Y/c)\psi=0, \quad (1)$$

where

ψ = 16 component wave function ψ_{ab}; $(a, b)=1, 2, 3, 4$.

$p_0=-\hbar\partial/ic\partial t+e^2/cr$.

$Y=(e^2/2)[(\boldsymbol{\alpha}_e\boldsymbol{\alpha}_M)/r+(\boldsymbol{\alpha}_e\mathbf{r})(\boldsymbol{\alpha}_M\mathbf{r})/r^3]$.

$\boldsymbol{\alpha}_e=(\alpha_{ex}, \alpha_{ey}, \alpha_{ez})$ = vector having for components direct products of the first three Dirac matrices for the electron index, a, and a unit matrix for the proton index b; the original representation of Dirac is used here so that the parts of α_{ex}, α_{ey} referring to the subscript a have, respectively, elements 1, 1, 1, 1 and $-i$, i, $-i$, i on the diagonal perpendicular to the principal diagonal starting at the upper right-hand corner.

$\boldsymbol{\alpha}_M$ = vector having for components the first three Dirac matrices for proton index b; conventions followed for $\boldsymbol{\alpha}_M$ are similar to those for $\boldsymbol{\alpha}_e$, reversing a with b and the words proton and electron.

β_e = matrix diagonal in b and operating on a as Dirac's α_4; the representation is such that the only non-vanishing matrix elements of α_4 are on the principal diagonal and have the values 1, 1, -1, -1 starting at the upper left-hand corner.

β_M = matrix diagonal in a, operating on b as Dirac's α_4; the conventions regarding representation are like those for β_e.

m = electron mass.

M = proton mass.

e = electronic charge—a positive number.

c = velocity of light.

$\mathbf{p}_e=(\hbar/i)(\partial/\partial x_e, \partial/\partial y_e, \partial/\partial z_e)$.

$\mathbf{p}_M=(\hbar/i)(\partial/\partial x_M, \partial/\partial y_M, \partial/\partial z_M)$.

$\mathbf{r}_e=(x_e, y_e, z_e)$ = electron coordinates.

$\mathbf{r}_M=(x_M, y_M, z_M)$ = proton coordinates.

r = distance between nucleus and electron.

It is convenient to list at this stage the meaning of some of the symbols which are introduced later in this paper as well.

α = fine structure constant $e^2/\hbar c$.

ψ_S = non-relativistic Schrödinger function.

$\psi_S(0)$ = value of ψ_S for $r=0$.

$\mu_0=e\hbar/2mc$ = electronic Bohr magneton.

$\mu_{0M}=e\hbar/2Mc$ = nuclear Bohr magneton.

$\mathbf{J}\hbar$ = angular momentum of relative motion.

$\rho=(1-\alpha^2)^{\frac{1}{2}}$.

R.P. stands for the real part of quantity to the right of symbol R.P.

Equation (1) is only approximate. Arguments previously[4] discussed show that it is a good equation to within the order v^2/c^2, where v stands for the velocity of either particle. The velocity of the electron is, of course, large when it is close to the proton and in the corresponding region of configuration space the validity of Eq. (1) may be questioned. It is clear, however, that the error introduced by this effect can be made to be vanishingly small by making M very large, because the effect of high electron velocity is represented in a relativistically consistent and supposedly correct manner by Dirac's equation for an electron subject to the action of an electromagnetic field.

The term in Y represents the combined effect of the magnetic interaction between the particles and the correction to the electrostatic interaction owing to the finiteness of the velocity of light. It will be omitted at first, and its effect will be taken into account later by means of a first-order perturbation calculation. Omitting the term in $-Y/c$, one finds from Eq. (1) that

$$\begin{aligned}(p_0+Mc+\boldsymbol{\alpha}_e\mathbf{p}_e+\beta_e mc)\Phi+(\boldsymbol{\sigma}_M\mathbf{p}_M)\Psi=0,\\(p_0-Mc+\boldsymbol{\alpha}_e\mathbf{p}_e+\beta_e mc)\Psi+(\boldsymbol{\sigma}_M\mathbf{p}_M)\Phi=0,\end{aligned}\tag{1.1}$$

where Ψ and Φ are 8-component wave functions obtained from ψ by restricting the proton index b in ψ_{ab} to the values 3, 4, in the case of Ψ, and to the values 1, 2 in the case of Φ. In the notation employed from Eq. (1.1) on the rows corresponding to $b=1, 2$ are arranged to have the same position and order in Φ as the rows corresponding to $b=3, 4$ have in Ψ. The vectors $\boldsymbol{\sigma}_M$ have for components matrices which are direct products of unit matrices in the electron index a and Pauli matrices in the proton index. The operator $\mathbf{p}_e+\mathbf{p}_M$ represents the total momentum. It commutes with the Hamiltonian which corresponds to Eq. (1) whether one neglects Y or not. It is, therefore, possible to confine the discussion to such ψ that

$$(\mathbf{p}_e+\mathbf{p}_M)\psi=0.\tag{1.2}$$

It follows from the last equation that every ψ_{ab} is a function only of the relative coordinates

$$x=x_e-x_M,\quad y=y_e-y_M,\quad z=z_e-z_M,\tag{1.3}$$

and is independent of the centroid coordinates $(m\mathbf{r}_e+M\mathbf{r}_M)/(m+M)$. Consequently, one has also

$$(\mathbf{p}_e+\mathbf{p}_M)\Psi=0,\quad(\mathbf{p}_e+\mathbf{p}_M)\Phi=0.\tag{1.4}$$

Because of Eq. (1.3), the relative momentum operator,

$$\mathbf{p}=(\hbar/i)(\partial/\partial x,\ \partial/\partial y,\ \partial/\partial z),\tag{1.5}$$

has the property

$$\mathbf{p}\Psi=\mathbf{p}_e\Psi,\quad\mathbf{p}\Phi=\mathbf{p}_e\Phi.\tag{1.6}$$

For states of definite total energy E one has

$$cp_0=E+e^2/r=(M+m)c^2+\epsilon+e^2/r,\tag{1.7}$$

where ϵ is the energy of the system in excess of the sum of the rest mass energies of the two particles. In view of Eq. (1.6), one may replace $\mathbf{p}_e$ in Eq. (1.1) by $\mathbf{p}$, and, similarly, on account of Eq. (1.4) one may replace $\mathbf{p}_M$ in Eq. (1.1) by $-\mathbf{p}$. From now on the subscript e will be dropped in α_e and β_e. The first line in Eq. (1.1) is used next to express Φ in terms of Ψ. Substitution of the result into the second line gives

$$(p_0-Mc+\boldsymbol{\alpha}\mathbf{p}+\beta mc-A/c)\Psi=0,\tag{2}$$

where

$$A=c(\boldsymbol{\sigma}_M\mathbf{p})(p_0+Mc+\boldsymbol{\alpha}\mathbf{p}+\beta mc)^{-1}(\boldsymbol{\sigma}_M\mathbf{p}).\tag{2.1}$$

Denoting by the sign † the conjugate transposed of a matrix one finds also

$$\Phi^\dagger\Phi=\Psi^\dagger(\boldsymbol{\sigma}_M\mathbf{p})\times(p_0+Mc+\boldsymbol{\alpha}\mathbf{p}+\beta mc)^{-2}(\boldsymbol{\sigma}_M\mathbf{p})\Psi.\tag{2.2}$$

Since $p_0\Psi\sim Mc\Psi$, one finds that through most of the configuration space

$$\Phi^\dagger\Phi\cong\Psi^\dagger(p^2/4M^2c^2)\Psi\cong\alpha^2(m/2M)\Psi^\dagger\Psi,\tag{2.3}$$

where α is the fine-structure constant. The factor multiplying $\Psi^\dagger\Psi$ in Eq. (2.3) is of the order of 1.5×10^{-8}. The effect of Φ on the normalization integral may be neglected therefore. Calculations with Eq. (2) may be carried on accordingly without reference to Φ. The

[4] G. Breit, Phys. Rev. **34**, 553 (1929); Phys. Rev. **36**, 383 (1930); Phys. Rev. **39**, 616 (1932).

operator A of Eq. (2.1) is expanded by means of

$$(p_0+Mc+\boldsymbol{\alpha}\mathbf{p}+\beta mc)^{-1}$$
$$=(p_0+Mc)^{-1}-(p_0+Mc)^{-1}(\boldsymbol{\alpha}\mathbf{p}+\beta mc)$$
$$\times(p_0+Mc)^{-1}+(p_0+Mc)^{-1}$$
$$\times[(\boldsymbol{\alpha}\mathbf{p}+\beta mc)(p_0+Mc)^{-1}]^2+\cdots. \quad (2.4)$$

In view of the fact that $(\boldsymbol{\alpha}\mathbf{p}+\beta mc)\Psi$ is of the order $mc^2\alpha^2\Psi$, the second term in the above expansion is of the order $m\alpha^2/M$ of the first, and the third term is of the order $m\alpha^2/M$ of the second. The contributions to the energy arising from the first term are of the order of the non-relativistic effect of the nuclear motion, those arising from the second term are still of interest for the hyperfine structure being of the order

$$(e^2/2a_H)(m/M)^2\alpha^2\sim 0.5\times 10^{-5}\ \text{cm}^{-1}.$$

The third term is of the order 10^{-13} cm^{-1} and may be neglected. In view of the fact that the second term in Eq. (2.4) is on the limit of what is of interest it will be taken account of at the end of the calculation, and it will be neglected at first. The consideration of the first and main term in Eq. (2.4) has to be carried out in more detail, however. One has

$$c(\boldsymbol{\sigma}_M\mathbf{p})(p_0+Mc)^{-1}(\boldsymbol{\sigma}_M\mathbf{p})=c(p_0+Mc)^{-1}p^2$$
$$-\hbar^2(e^2/r^3)(p_0+Mc)^{-2}(r\partial/\partial r)+A_1, \quad (2.5)$$

where

$$A_1=(\hbar e^2/r^3)(p_0+Mc)^{-2}[\mathbf{r}\times\mathbf{p}]\cdot\boldsymbol{\sigma}_M. \quad (2.6)$$

The term A_1 is the analog of the Thomas term. If the Thomas terms were appreciable, complications would arise. An estimate of the effect of A_1 will be made, therefore, at this stage. The estimate may be carried out in the approximation of neglecting nuclear motion. One finds that in the state in which the proton spin and the electronic angular momentum are parallel, the expectation value of A_1 is

$$\langle A_1\rangle\cong(\pi m/3M)\alpha^2(e\hbar/mc)(e\hbar/Mc)\psi_S^2(0). \quad (2.7)$$

The intermediate steps leading to this formula are straightforward and are omitted. The contribution to the hyperfine-structure separation arising from the operator A_1 is 4 times the value of $\langle A_1\rangle$ given in Eq. (2.7) and is thus

$$(16\pi m/3M)\alpha^2\mu_0\mu_{0M}\psi_S^2(0).$$

This is $m\alpha^2/2M\cong 1.5\times 10^{-8}$ times the hyperfine structure splitting

$$(32\pi/3)\mu_0\mu_{0M}\psi_S^2(0),$$

which is expected according to Fermi's formula. The effect of A_1 is seen to be negligible, and it will be omitted below.

By means of Eqs. (2), (2.1), (2.5), and with the omission of A_1 as well as of the second term in Eq. (2.4), one obtains

$$[p_0-Mc+\boldsymbol{\alpha}\mathbf{p}+\beta mc-(p_0+Mc)^{-1}p^2$$
$$+\hbar^2(e^2/cr^2)(p_0+Mc)^{-2}\partial/\partial r]\Psi=0. \quad (3)$$

A special solution is found by arranging for Ψ to be the direct product of the function for index b, which corresponds to $\begin{pmatrix}1\\0\end{pmatrix}$ upward orientation of the proton spin and of a column matrix in the index a having elements depending on angles in the same way as for Dirac's equation in a central field for upward orientation of the electronic angular momentum. The non-vanishing components of Ψ are then

$$\Psi_{1,1}=izf/r, \quad \Psi_{2,1}=i(x+iy)f/r, \quad \Psi_{3,1}=g.$$

Substitution into Eq. (3) shows that it suffices to satisfy the radial equations

$$\begin{aligned}&(p_0-Mc-mc)g\\&\quad+\frac{\hbar^2 d}{r^2dr}\left(\frac{r^2}{p_0+Mc}\frac{dg}{dr}\right)+\frac{\hbar}{r^2}\frac{d(r^2f)}{dr}=0,\\&(p_0-Mc+mc)f\\&\quad+\frac{\hbar^2}{r^2}\frac{d}{dr}\left(\frac{r^2}{p_0+Mc}\frac{df}{dr}\right)-\frac{\hbar dg}{dr}=0.\end{aligned} \quad (3.1)$$

The angles and spin orientations have been eliminated at this stage and only the radial motion has been left. The possibility of separating variables in this manner is not unexpected. It is, in fact, readily verified that the Hamiltonian

$$H_0=-e^2/r-c(\boldsymbol{\alpha}_e\mathbf{p}_e)$$
$$-c(\boldsymbol{\alpha}_M\mathbf{p}_M)-\beta_e mc^2-\beta_M Mc^2 \quad (3.15)$$

commutes with the total angular-momentum components. Thus, for instance,

$$[H_0, xp_y-yp_x+\hbar(\sigma_{ez}+\sigma_{Mz})/2]\psi=0 \quad (3.16)$$

for any ψ satisfying Eq. (1.2). In this equation σ_{ez}, σ_{Mz} are the z components of the Dirac 4-row matrix vectors $\boldsymbol{\sigma}_e$, $\boldsymbol{\sigma}_M$, referring to the electron and proton. The original Eq. (1), with Y omitted, can be satisfied, therefore, by a wave function for which the total momentum is zero and for which also the z component as well as the square of the total angular momentum can have eigenvalues. It will be noted that the orbital angular momentum entering the problem is that of relative motion and that the elimination of Φ does not spoil the validity of the conservation of angular momentum.

If M is made to approach ∞ in Eq. (3.1), and if p_0 is expressed in terms of ϵ by means of Eq. (1.7), the radial equations become identical with corresponding radial equations for Dirac's electron in a central field, and aside from questions connected with the polarization of a vacuum they are the correct equations for an electron in the field of a point charge within the limitations of present day theory.

It has been shown by Bechert and Meixner,[5] and by Lowen,[6] that the effect of nuclear motion on the energy is not changed by relativistic effects within the first-order correction terms of order m/M. It is to be expected, therefore, that the main effect of m/M on the wave function is that of changing the linear scale, since this is its effect for the non-relativistic Schrödinger wave function ψ_S.

The Schrödinger radial equation is obtained by replacing the quantity p_0+Mc occurring in the first line of Eq. (3.1) by $2Mc$ expressing f as approximately $(\hbar/2mc)dg/dr$ by means of the second line and substituting this value of f into the first line. One obtains in this way

$$\left(\epsilon_S+\frac{e^2}{r}\right)\psi_S+\frac{\hbar^2}{2}\left(\frac{1}{m}+\frac{1}{M}\right)\frac{d(r^2d\psi_S/dr)}{r^2dr}=0, \quad (3.2)$$

which is the non-relativistic Schrödinger equation. The largest error in this reduction is introduced by the replacement of $(p_0-Mc+mc)f$ by $2mcf$, when f is solved for in terms of g. The fractional error introduced into f is of the order $(\epsilon+e^2/r)/(2mc^2)$, and this is of the order $\alpha^2/4$ for large r. At the turning point of the Bohr orbit the error vanishes. For sufficiently small r the fractional error becomes large, but for $r\sim a_H/10$ the fractional error in f is still of the order 10^{-4}. For smaller distances the fractional error in f introduced by the approximation used here increases, and one cannot claim that this approximation is good enough on the grounds of numerical smallness of the effect everywhere. On the other hand, it is readily verified that if $M=\infty$ the approximation made in approximating f by $(\hbar/2mc)dg/dr$ is harmless and can be taken care of by the correction factor to hyperfine structure splitting having the value

$$[2(1-\alpha^2)-(1-\alpha^2)^{\frac{1}{2}}]^{-1}\cong 1+3\alpha^2/2, \quad (3.3)$$

which has to be taken into account for large atomic numbers[3] but is negligible for hydrogen. This correction factor arises as the quotient of the exact and approximate value of

$$\int_0^\infty fgdr. \quad (3.4)$$

The approximate value of the integral being obtained by replacement of g by ψ_S and of f by $(\hbar/2mc)d\psi_S/dr$. The order of magnitude of the correction factor is what one would estimate from the fact that the average value of e^2/r is -2ϵ and that $\epsilon\sim -mc^2\alpha^2/2$. The value of the correction factor cannot be correctly estimated in this manner, however, because the function g is only approximately represented by ψ_S. Aside from the normalization factor, this difference consists only in the presence of an extra factor,

$$r^{\rho-1}\cong r^{-\alpha^2/2}; \quad \rho=(1-\alpha^2)^{\frac{1}{2}},$$

in g, which results again in effects of order α^2.

The smallness of the difference between the correction factor Eq. (3.3) and unity has just been explained for $M=\infty$. The corresponding correction factor for $M\neq\infty$ will be studied next. It will be found in connection with Eqs. (5.5) that the radial integrals needed for the evaluation of the expectation value of Y are of the form listed in Eq. (3.4). The examination of the

[5] K. Bechert and J. Meixner, Ann. d. Physik **22**, 525 (1935).
[6] I. S. Lowen, Phys. Rev. **51**, 190 (1937).

radial equations will be made first, and the reduction of the expectation value to radial integrals will follow.

B. Orientation Regarding Radial Functions by the Ritz Method

Expressing energy in units $mc^2\alpha^2$, and length in $a_H=\hbar^2/me^2=$ Bohr radius, one obtains in place of Eq. (3.1)

$$\alpha\left(\epsilon+\frac{1}{r}\right)g+\frac{d(r^2f)}{r^2dr}+\alpha\frac{d}{r^2dr}\left[\frac{r^2dg/dr}{2/\beta+1+\alpha^2(\epsilon+1/r)}\right]=0,$$
$$\left[1+(\alpha^2/2)\left(\epsilon+\frac{1}{r}\right)\right]f-\frac{\alpha dg}{2dr}+\frac{\alpha^2 d}{2r^2dr}\left[\frac{r^2df/dr}{2/\beta+1+\alpha^2(\epsilon+1/r)}\right]=0, \quad (3.5)$$

where

$$\beta=m/M\ll 1. \quad (3.6)$$

These equations can be derived from the variational equation

$$\delta\int_0^\infty\{\tfrac{1}{2}[1+(\alpha^2/2)(\epsilon+1/r)]f^2-(\alpha/2)fdg/dr+(\alpha^2/4)(\epsilon+1/r)g^2-(\alpha^2/4)[(df/dr)^2+(dg/dr)^2]/[(2/\beta+1+\alpha^2(\epsilon+1/r)]\}r^2dr=0. \quad (3.7)$$

The smallness of $\alpha^2\epsilon$, in comparison with $2/\beta$, allows the replacement of ϵ by ϵ_S in the last term and the restatement of the variational equation for the ground state in the form

$$-\alpha^2\epsilon\int_0^\infty(f^2+g^2)r^2dr=\int_0^\infty\{(2+\alpha^2/r)f^2+\alpha^2g^2/r-2\alpha fdg/dr-\alpha^2[(df/dr)^2+(dg/dr)^2]/[(2/\beta)+1+\alpha^2(\epsilon_S+1/r)]\}r^2dr, \quad (3.8)$$

$$(\epsilon+\alpha^{-2})^2=\text{minimum}. \quad (3.9)$$

Minimizing the energy gives $\epsilon=-\infty$. This is a consequence of the existence of negative energy states. The problem is stated, therefore, in the form of minimizing the square of the energy.

The minimum of $(\epsilon+\alpha^{-2})^2$ for $M=\infty$ is $(1-\alpha^2)\alpha^{-1}$ and corresponds to the normal state of hydrogen.

By means of the variational equation the best values of the parameters A, σ, γ will now be determined for the trial functions

$$g=r^\sigma e^{-\gamma r},\quad f=Ar^\sigma e^{-\gamma r}, \quad (4)$$

by the Ritz method. One has

$$\int_0^\infty g^2r^2dr=A^{-2}\int_0^\infty f^2r^2dr=(2\gamma)^{-2\sigma-3}\Gamma(2\sigma+3), \quad (4.1)$$

and, similarly, the other needed integrals are readily obtained. One finds

$$1+\alpha^2\epsilon=(1-A^2)/(1+A^2)-\gamma[\alpha^2+2\alpha A/(1+A^2)]/(\sigma+1)+\alpha^2\beta\gamma^2/(2+\beta)(2\sigma+1), \quad (4.2)$$

where the approximation

$$(2/\beta)+1+\alpha^2(\epsilon_S+1/r)\cong(2/\beta)+1 \quad (4.21)$$

has been made in the last term of the integral on the right side of Eq. (3.8). Minimizing the square of the energy with respect to γ, A and σ in accordance with Eq. (3.9), one obtains, respectively,

$$\alpha^2+2\alpha A/(1+A^2)=2\alpha^2\beta\gamma(\sigma+1)/(2+\beta)(2\sigma+1), \quad (4.22)$$

$$2A+\alpha\gamma(1-A^2)/(1+\sigma)=0, \quad (4.23)$$

$$\alpha^2+2\alpha A/(1+A^2)=2\alpha^2\beta\gamma(\sigma+1)^2/(2+\beta)(2\sigma+1)^2. \quad (4.24)$$

From Eqs. (4.22) and (4.24) it follows that $\sigma=0$, and the solution of the remaining two independent relations gives

$$A=-\alpha\gamma/2+\alpha^3/8(1+\beta)^3+\cdots, \quad (4.3)$$

$$\gamma=1/(1+\beta)+\alpha^2/2+\cdots, \quad (4.31)$$

where the quantities A, γ are being represented in descending powers of $1+\beta$. It is seen that if A and γ are approximated by the first term in

the expansions on the right side of Eqs. (4.3), (4.31), the error introduced into the radial integral in Eq. (3.4) is of the order $\alpha^2\beta$ which may be neglected. It is natural that the adjustment of σ to the best value gives $\sigma=0$ rather than $-\alpha^2/2$ because the Ritz method is not sensitive to errors of the wave function close to $r=0$. It is this region that is responsible for the occurrence of the exponent $-\alpha^2/2$. On the other hand, the presence of the factor $\exp[-(\alpha^2/2)\ln r]$ in the wave function produces an error of the order of 10^{-4} in the density as r changes by a factor $2.72^2=7.4$. It is, therefore, natural to expect that the integral of Eq. (3.4), as well as the variational principle expressions for the energy, are not sensitive to the difference between $\sigma=0$ and $\sigma=-\alpha^2/2$. That the integral is not sensitive to this difference has been shown in connection with Eq. (3.3).

In view of the lack of sensitivity of the variational principle to σ, the question arises regarding the error introduced by the failure of Eqs. (4.3), (4.31) to reproduce correctly the wave function close to $r=0$ for $M=\infty$. An estimate of the error can be made by assigning to σ the value $-\alpha^2/2$ and adjusting γ and A by the Ritz method by means of Eqs. (4.22), (4.23). It is then found that A, γ have values A_σ, γ_σ given by

$$A_\sigma=-(\alpha\gamma/2)[1-\sigma-\alpha^2/4(1+\beta)^2+\cdots], \quad (4.32)$$

$$\gamma_\sigma=1/(1+\beta)-\beta\alpha^2/(1+\beta)^2+\cdots. \quad (4.33)$$

The second term of the formula for γ_σ is negligible compared with the first. The term $-\sigma$ on the right side of Eq. (4.32) changes A_σ by one part in 30,000 and is the largest correction. From the point of view of over-all adjustment by the Ritz method, with the three parameters which have been used above, the radial integral which matters for hyperfine structure is seen to be stable to within fractional errors of order α^2.

The answer just obtained for the radial integral by the variational method leads to a correction factor involving the cube of the ratio of the electronic and reduced mass ratios. Considerations will be given next, however, which show that the form of the wave function assumed in Eq. (4) is not sufficiently flexible and that a more accurate solution modifies the result for the wave function without affecting, however, the radial integrals that are important for hyperfine structure.

For larger r the last term in the second Eq. (3.5) can be neglected, and the equation just mentioned yields then an expression for f in terms of g which, when substituted into the first of the two Eqs. (3.5), gives a second-order equation in one variable g. In this one can replace the denominator $2/\beta+1+\alpha^2(\epsilon+1/r)$ by $2/\beta$ for $r>\alpha^2$ and even for smaller r. Replacing g by

$$G=rg, \quad (4.4)$$

one has

$$\frac{d^2G}{dr^2}+\frac{P'(r)}{P(r)}\frac{dG}{dr}+\left[\frac{\epsilon+1/r}{P(r)}-\frac{P'(r)}{rP(r)}\right]G=0, \quad (4.5)$$

where

$$P(r)=1/(A+\alpha^2/r)+1/(B+\alpha^2/r), \quad (4.51)$$

$$P'(r)=\alpha^2r^{-2}[(A+\alpha^2/r)^{-2}+(B+\alpha^2/r)^{-2}], \quad (4.52)$$

$$A=2+\alpha^2\epsilon, \quad B=2/\beta+\alpha^2\epsilon. \quad (4.53)$$

Equation (4.5) can be discussed by standard methods. The term in the first derivative can be removed by the transformation

$$G=v/[P(r)]^{\frac{1}{2}}, \quad (4.54)$$

and it is found from the differential equation satisfied by v that the asymptotic form of v is a constant multiple of

$$r^n\exp(-\gamma r), \quad (4.55)$$

where

$$\gamma^2\cong-2\epsilon(1+\beta)^{-1}(1+\alpha^2\epsilon/2), \quad (4.56)$$

and

$$n\cong[(1+\beta)^{-1}+\alpha^2\epsilon]/\gamma. \quad (4.57)$$

Equations (4.54), (4.55), (4.56), and (4.57) determine the asymptotic value of the logarithmic derivative of g for large r.

C. Perturbation Calculation of the Radial Functions

The calculation of g will next be made by determining first its logarithmic derivative. The latter will be found by means of a perturbation calculation in which the radial equations for $\beta=0$ serve as the reference point. The logarithmic derivative

$$y=dG/Gdr \quad (4.6)$$

is expressed in terms of

$$y_0 = dG_0/G_0 dr \tag{4.61}$$

as

$$y = y_0 + \delta y. \tag{4.62}$$

It is found from the second-order differential equation by first converting to the Riccati type of equation in y and then introducing the integrating factor corresponding to the terms linear in δy that

$$\frac{d}{dr}(G_0^2 P \delta y) + G_0^2 P\{\delta[(\epsilon+1/r)/P - P'/rP] + y_0\delta(P'/P) + (\delta y)^2\} = 0, \tag{4.63}$$

where the following conventions are employed: (a) all quantities with the exception of ϵ are functions of r, (b) the quantity P' is the derivative of P with respect to r in agreement with the notation of Eq. (4.52), (c) the symbol δ before a parenthesis or a bracket indicates that one subtracts from the quantity in parentheses or brackets the value of the quantity for $\beta = 0$. Since according to Eqs. (4.56), (4.57) the quantity δy is finite at $r = \infty$, and since G_0^2 vanishes exponentially at $r = \infty$, one can neglect the value of $G_0^2 P \delta y$ at $r = \infty$. Integration of Eq. (4.63) yields, therefore,

$$G_0^2 P \delta y = \int_r^\infty G_0^2 P\{\delta[(\epsilon+1/r)/P - P'/rP] + y_0\delta(P'/P) + (\delta y)^2\} dr. \tag{4.64}$$

The determination of the effect of β has been now put into the form of obtaining a solution of an integral equation which is suitable for iteration. The process of joining smoothly the branch of g extending to $r = \infty$ with the branch starting at $r = 0$ involves an adjustment of $\delta\epsilon$. The joining of the two branches of the curve will be made at $r = \alpha^2$, which is the value of r in units of Bohr radii, which corresponds to the classical electron radius e^2/mc^2.

The right side of Eq. (4.64) contains under the integral sign the quantity δy, which has to be determined. The practicability of employing the above equation depends on the fact that the first-order result which is obtained by neglecting $(\delta y)^2$ in the integrand yields a small value of δy if the adjustment of ϵ is made so as to secure smooth joining of the two branches of the wave function.

Neglecting $(\delta y)^2$ under the integral sign of Eq. (4.64) for the present, one obtains

$$\delta y \cong (A_0 + \alpha^2/r) r^{-2\rho} e^{2r} \int_r^\infty r^{2\rho} e^{-2r} \times \{\delta\epsilon + (\beta/2)[1 - (2/r) - \alpha^2(r^{-2} - r^{-1})] + \alpha^2\beta/[2r^2(A_0 + \alpha^2/r)]\} dr. \tag{4.65}$$

In addition to the approximation, consisting of neglecting quadratic effects, which has already been mentioned, the following approximations have been made at this stage: (a) A term in $y_0\alpha^2\beta^2/r^2$ has been omitted since it occurs side by side with a term in $y_0\beta\alpha^2/r^2$; (b) The quantity $\beta\alpha^2(1-\alpha^2)^{\frac{1}{2}}r^{-2}$ has been replaced by $\beta\alpha^2 r^{-2}$; (c) The energy ϵ_0 has been expressed as $-\frac{1}{2} - \alpha^2/8$. By means of Eq. (4.65) one finds by straightforward calculation

$$\delta y \cong (1 + \alpha^2/2r)\{(\delta\epsilon - \beta/2)[1/2r^2 + 1/r + \alpha^2 e^{2r} Ei(-2r)/2r^2 + 1 - (\alpha^2/2)(1/r + 3/2r^2)] + \beta[1 - \alpha^2/4r^2 + \alpha^4 e^{2r} Ei(-2r)/4r^2]\}, \tag{4.66}$$

where

$$-Ei(-x) = \int_x^\infty (e^{-x}/x) dx. \tag{4.67}$$

In this calculation terms of higher order than $\alpha^2\beta$ have been dropped, with the exception of those containing $\alpha^4\beta Ei/r^2$, since it is not clear without further consideration that their effect is negligible. In the evaluation of the integral of $(2 + \alpha^2/r)^{-1}\exp(-2r)$ the factor $\exp(\alpha^2)$ multiplying the Ei has been dropped. This term is multiplied by $\alpha^4\beta$. Inclusion of the factor $\exp(\alpha^2)$ would amount to taking into account effects of order $\alpha^6\beta$ and it is not done for this reason.

Equations (1), (1.1) are not valid at very small r. If the interparticle distance becomes small enough, and if the Coulomb force is assumed to act at small distances, then the velocity of both the proton and the electron becomes comparable to the velocity of light, and there is no reason for expecting Eq. (1) to apply under such conditions. Equations (1.1), which are obtained from Eq. (1) by omitting the term in Y, are also not directly applicable at $r = 0$. If

one had to rely on Eq. (1) altogether, it would be impossible to obtain a solution with any certainty.

It is questionable, however, whether the Coulomb law is valid at distances much smaller than e^2/mc^2. There is a definite indication of a modification of the inverse-square law of force at this distance for two protons, and whether the origin of the specific proton-proton force at short distances be of the nature of a meson or some other type of field, it would not be surprising if the force between a proton and an electron should also turn out to be modified at distances comparable to those at which proton-proton interactions set in. In view of the likelihood of the absence of a strong increase in the velocity of the proton at distances smaller than the electronic radius and the gain in definiteness of discussion which results through the removal of the problem of finding a suitable Hamiltonian for the description of the motion of two rapidly moving point charges, *it will now be assumed* that at distances smaller than e^2/mc^2 the potential energy $-e^2/r$ is modified so as to be approximately constant. For simplicity it will be taken to have a constant value $-mc^2$ for $0<r<e^2/mc^2$. The conclusions do not depend, however, on whether one takes the potential energy to be exactly constant in this range of values of r or not.

Inasmuch as the last term in the second Eq. (3.5) is of the order $\alpha^2\beta$ of the first, it will be neglected in the calculation that will be made now. Since $f\ll g$, the effect on the term in $d(r^2f)/r^2dr$ in the first of the two Eqs. (3.5) is small. One obtains in this manner, on elimination of f, a simplified special case of Eq. (3.5)

$$d^2G/dr^2+K^2G=0, \tag{4.68}$$

with

$$K^2=(1+\alpha^2\epsilon)\alpha^{-2}[1/(3+\alpha^2\epsilon)+(\beta/2)(1+\beta+\alpha^2\beta\epsilon/2)^{-1}]^{-1}, \tag{4.69}$$

so that

$$ry=Kr\cot(Kr)=1-K^2r^2/3+\cdots. \tag{4.7}$$

Substituting K^2 and evaluating for $r=\alpha^2$, one has

$$(ry)_{r=\alpha^2}\cong 1-\alpha^2+3\alpha^2\beta/2. \tag{4.71}$$

On the other hand, the solution for the ground state corresponding to $\beta=0$ is, apart from a normalizing factor common to f and g,

$$f_0=-(1-\rho)^{\frac{1}{2}}r^{\rho-1}e^{-r};\qquad g_0=(1+\rho)^{\frac{1}{2}}r^{\rho-1}e^{-r}, \tag{4.72}$$

where ρ is as explained in the list of notations following Eq. (1). It follows that

$$y_0=\rho/r-1,\ (ry_0)_{r=\alpha^2}=(1-\alpha^2)^{\frac{1}{2}}-\alpha^2=1-3\alpha^2/2. \tag{4.73}$$

Combining Eqs. (4.71) and (4.73), one obtains

$$(r\delta y)_{r=\alpha^2}=\alpha^2/2+3\alpha^2\beta/2. \tag{4.74}$$

In this formula the first term $\alpha^2/2$ has to do with the modification introduced by changing the inverse first-power potential to a constant in the range $0<r<\alpha^2$. It has nothing to do with the effect of the mass of the proton on δy. Its inclusion in the calculation would correspond to making a combined estimate of the effect of the proton's size and of the effect of nuclear motion. In order to keep the two effects separate from each other, only the part of δy containing β will be kept at this stage. By separating the effects in this manner one obtains the effect of changing β from 0 to its experimental value. A slight error is introduced at this point through the employment of the functions given by Eqs. (4.72) rather than functions corresponding to

$$(r\delta y)_{r=\alpha^2,\,\beta=0}=\alpha^2/2, \tag{4.75}$$

which correspond to infinite proton mass and a constant potential energy in $0<r<\alpha^2$. The difference between these functions and the functions of Eq. (4.72) is very slight, however. The right side of Eq. (4.66) is equated, therefore, to $3\beta/2$, which is the part of δy corresponding to the term containing β on the right side of Eq. (4.74). Calculation gives then

$$\delta\epsilon-\frac{\beta}{2}=\frac{(1/4\alpha^2)-Ei(-2\alpha^2)(\exp 2\alpha^2)/4}{(1/2\alpha^4)+(1/4\alpha^2)+Ei(-2\alpha^2)(\exp 2\alpha^2)/2\alpha^2+1/2}\beta \cong(\alpha^2\beta/2)[1-2\alpha^2Ei(-2\alpha^2)-\alpha^2/2], \tag{4.8}$$

and the value of the first-order effect δy corresponding to this value of $\delta\epsilon-(\beta/2)$ is obtained

from Eq. (4.66) as

$$\delta y \cong \beta\{1+\alpha^2/2+\alpha^2/r + (\alpha^4/2r^2)[e^{2r}Ei(-2r)-Ei(-2\alpha^2)] - (\alpha^4/r)Ei(-2\alpha^2)\}. \quad (4.81)$$

The factor β, which is present in all terms of the above formula, makes the result small, and the quadratic term in the integrand of Eq. (4.64) is negligible on account of the presence of the factor β^2. This is the result of making δy small at $r=\alpha^2$.

The integral which matters for the hyperfine structure splitting is

$$\int_0^\infty fg dr. \quad (4.82)$$

The effect of δy on this integral can be broken down into: (a) The presence of the factor $\exp(\int^r(\delta y)dr)$ in both f and g, which results in a net factor $\exp(2\int^r(\delta y)dr)$, (b) an extra term arising in f because it is expressible in terms of dg/dr, so that the exponential just referred to brings in an extra term on differentiation. The correction factor to the integrand resulting from both causes is

$$[1+(g_0 dr/dg_0)\delta y]\exp\left[2\int^r(\delta y)dr\right],$$

which on substituting g_0 becomes

$$[1-(1+\alpha^2/2r)^{-1}\delta y]\exp\left(2\int^r \delta y dr\right). \quad (4.83)$$

This correction factor does not include the change in the normalization constant. It will be simpler to take this change into account later. It will first be explained that in Eq. (4.81) the only important term inside the brace is 1. The other terms give effects of the order $\alpha^2\beta$. For $r=\alpha^2$ the largest additional term inside the brace is that containing α^2/r. In the exponent of Eq. (4.83) it brings in a factor, $2\alpha^2\beta \log r$, which is entirely negligible in a larger range of values of r than $\alpha^2<r<10^3$. For the factor multiplying the exponential in Eq. (4.83) the term in α^2/r has the effect of multiplying the integrand by $1-\beta\alpha^2$, which is a negligible effect. At $r=\alpha^2$ the effect is greater, but an estimate of it shows that the integral of Eq. (4.82) is changed by the fractional amout $-\alpha^2\beta \ln 2\alpha^2$, which is also negligible. The term in square brackets containing $Ei(-2r)$ vanishes at $r=\alpha^2$. The square bracket changes slowly at $r=\alpha^2$, because at small r the dependence of the Ei-function on its argument is logarithmic. The effect of the square bracket is thus also negligible.

It is thus seen that the effect of β on the integral of Eq. (4.82) is to multiply it by the factor

$$\text{correction factor} = \frac{\int(1-\beta)\exp(2\beta r)f_0 g_0 dr}{\int f_0 g_0 dr} \times \frac{\int(f_0^2+g_0^2)r^2 dr}{\int(f_0^2+g_0^2)\exp(2\beta r)r^2 dr}, \quad (4.84)$$

all integrals being taken from 0 to ∞. To within first-order terms in β this is simply $(1+\beta)^{-3}$. The justification for employing the integral of Eq. (4.82) is given in the next section.

D. The Hyperfine Structure

The expectation value $\langle Y\rangle$, of the operator Y of Eq. (1), will now be computed. One has from the first Eq. (1.1) with the aid of Eqs. (1.6) and (2.4)

$$\Phi = [(p_0+Mc)^{-1}-(p_0+Mc)^{-1}(\boldsymbol{\alpha}\mathbf{p}+\beta mc) \times (p_0+Mc)^{-1}](\boldsymbol{\sigma}_M\mathbf{p})\Psi, \quad (5)$$

where the subscript e is dropped and c.g.s. units are used. The operator cp_0 is

$$cp_0 = (M+m)c^2+\epsilon+e^2/r. \quad (5.1)$$

The first term in Eq. (5) contributed the main part of the expectation value. The second term contributes a correction of order β. Relationship to familiar expressions is secured by means of the approximation

$$cp_0 \cong Mc^2. \quad (5.1')$$

This approximation is not good enough for the evaluation of contributions arising from the first

term on the right side of Eq. (5), but suffices for the calculation of effects arising from the second term. The integral representing $\langle Y\rangle$ contains Φ linearly, and the approximation made in Eq. (5.1′) is taken care of for the first term of Eq. (5) by the correction factor

$$2M/(2M+m)=1/(1+\beta/2). \tag{5.2}$$

The quantity $\epsilon+e^2/r$ in Eq. (5.1) contributes terms of the order α^2 in comparison with mc^2. The correction factor of Eq. (5.2) is, therefore, good enough. The second term of Eq. (5) can be taken into account by noting that commuting $\boldsymbol{\alpha}\mathbf{p}+\beta mc$ with $(p_0+Mc)^{-1}$ brings in terms of still higher order and

$$(\boldsymbol{\alpha}\mathbf{p}+\beta mc)\Psi\cong -mc\Psi.$$

The correction factor brought in by the second term of Eq. (5) is thus

$$1-(-mc)(2Mc)/(2Mc)^2=1+\beta/2. \tag{5.2′}$$

Combining this correction factor with that of Eq. (5.2) one obtains a correction factor of unity to within terms of order $\alpha^2\beta$. It is, therefore, good enough to calculate with the simple approximation of Eq. (5.1′), which corresponds to

$$\Phi\cong(\boldsymbol{\sigma}_M\mathbf{p})\Psi/(2Mc). \tag{5.3}$$

One obtains in this manner

$$\langle 2Y\rangle=(e^2/2Mc)\int\Psi^\dagger[(\boldsymbol{\alpha}\boldsymbol{\sigma}_M)/r + (\boldsymbol{\alpha}\mathbf{r})(\boldsymbol{\sigma}_M\mathbf{r})/r^3](\boldsymbol{\sigma}_M\mathbf{p})\Psi+\text{comp. conj.}\}d\tau,$$

where comp. conj. stands for the complex conjugate of the immediately preceding expression. Linearizing in $\boldsymbol{\sigma}_M$ one obtains

$$\langle 2Y\rangle=(e^2/2Mc)\int\{\Psi^\dagger[r^{-1}(\boldsymbol{\alpha}\mathbf{p})+ir^{-1}\boldsymbol{\sigma}_M[\boldsymbol{\alpha}\times\mathbf{p}] + r^{-3}(\boldsymbol{\alpha}\mathbf{r})(\mathbf{r}\mathbf{p}+i[\mathbf{r}\times\mathbf{p}]\boldsymbol{\sigma}_M)]\Psi +\text{comp. conj.}\}d\tau. \tag{5.4}$$

The value of this quantity is calculated for the state in which the z component of $\boldsymbol{\sigma}_M/2$, which represents the nuclear spin, has the characteristic value $\frac{1}{2}$ and in which the projection of the electronic angular momentum on the z axis is also $\frac{1}{2}$. This state is one of the three magnetic levels with hyperfine-structure quantum number $f=1$. It suffices to calculate $\langle Y\rangle$ for this state because according to well-known relations it has the same value for the other magnetic levels of $f=1$; the value of $\langle Y\rangle$ for $f=0$ is known to be (-3) times the value for $f=1$. Straightforward substitution and calculation gives

$$\langle Y\rangle=-(8\pi/3)(e^2h/Mc)\int_0^\infty fg\,dr. \tag{5.5}$$

The normalization is supposed to be

$$4\pi\int_0^\infty(f^2+g^2)r^2dr=1. \tag{5.5′}$$

The integral of Eq. (4.82) is seen to occur in Eq. (5.5). The normalization integrals of Eq. (4.84) are also seen to be in agreement with Eq. (5.5′). It remains to evaluate the right side of Eq. (4.84).

The factor $(1-\beta)\exp(2\beta r)$ of the first integral in the numerator of Eq. (4.84) just suffices to change it into the first integral in the denominator by a change of variable

$$r'=r/(1+\beta), \tag{5.5″}$$

which corresponds to the non-relativistic effect of expansion of the wave-function space in the ratio $(1+\beta):1$. A slight inaccuracy is involved at this point. The factor $[1+(\alpha^2/2)(\epsilon+1/r)]^{-1}$ in the expression for f in terms of dg/dr does not have its $1/r$ changed into $1/r'$ by the transformation of Eq. (5.5″). The error involved in this inaccuracy is of the order $\alpha^2\beta$ and will be neglected. It may be well to recall here that the whole term $(\alpha^2/2)(\epsilon+1/r)$ could be omitted without affecting the result to more than terms of relative order α^2 and that Fermi's well-known result involving the Schrödinger $\psi^2(0)$ has been obtained in this manner. The ratio of the first integral of Eq. (4.84) in the numerator to the first integral in the denominator is thus unity to within terms of relative order $\alpha^2\beta$. The ratio of the second integral in the numerator to the second integral in the denominator can again be considered by means of the transformation of Eq. (5.5″). The factor $\exp(2\beta r)$ suffices for the change to r' in g_0^2 and within errors of relative order $\alpha^2\beta$ or β^2 in f_0^2. The conversion of r^2dr in the denominator brings about the appearance of

the factor $(1+\beta)^3$ in the denominator, which amounts to $1/(1+\beta)^3$ for the correction factor.

Another way of explaining the result just arrived at is to note that the factors $(1-\beta)$ and $\exp(2\beta r)$ in the first integral in the numerator of Eq. (4.84) suffice to make the integrand $(\alpha/4)dg^2/dr$ with an error of relative order $\alpha^2\beta$, while the factor $\exp(2\beta r)$ makes the second integrand of the second integral in the denominator in the same approximation as previously g^2r^2. Similarly, the integrands of the first integral in the denominator and the second integral in the numerator are f_0g_0 and g_0^2 respectively. The right side of Eq. (4.84) is thus $g^2(0)/g_0^2(0)$. The difference between $g^2(r)$ and $g_0^2(r)$, aside from normalization, is that

$$g(r)=\text{const.}g_0(r/(1+\beta)).$$

Since the normalization of g^2 is obtained by division by

$$4\pi\int g^2(r)r^2dr=4\pi(1+\beta)^3\int g_0^2(r)r^2dr,$$

the correction factor is $(1+\beta)^{-3}$.

In both explanations f_0^2/g_0^2 has been neglected in the normalization integral. The error introduced by doing so is represented according to Eq. (3.5) by the factor

$$1-(\alpha^2/4)\int(dg/dr)^2r^2dr\Big/\int g^2r^2dr\cong 1-\alpha^2/4.$$

This factor is compensated for, however, by the fact that

$$f_0\cong-(\alpha/2)(1+\alpha^2/4)g_0.$$

The remaining inaccuracy is thus represented to relative order α^2 by the factor

$$\left[\int(r^{\rho-1}e^{-r})^2dr\Big/\int r^2(r^{\rho-1}e^{-r})^2dr\right]$$

$$\times\left[\int r^2e^{-2r}dr\Big/\int e^{-2r}dr\right]$$

$$\cong[\Gamma(2\rho-1)/\Gamma(2\rho+1)][\Gamma(3)/\Gamma(2)]$$

$$=1/(2\rho^2-\rho)\cong 1+3\alpha^2/2, \quad (5.6)$$

in agreement with the exact result[3] for $\beta=0$.

A check on the validity of the factor $(1+\beta)^{-3}$ can be obtained by going to the extreme case of two equal masses. For this case the problem has been essentially worked out previously.[7] The scheme employed a limiting process in which the potential energy $-e^2/r$ was first replaced by a function $-e^2\{1/r\}$ which was kept finite at $r=0$. The range of values of r within which the function was assumed to differ from $-e^2/r$ was made to approach zero, and the value of the function at $r=0$ was made to approach ∞. At all stages in the limiting process the Laplacian of the modified function was well defined. In the limit it approached a multiple of the Dirac δ function as below,

$$\lim\Delta(-e^2\{1/r\})=4\pi e^2\delta(\mathbf{r}). \quad (6)$$

The spin interactions between the particles were brought in by a generalization of the operator Y. The expectation value of the perturbing energy for parallel spin orientations was calculated to be the same as that of

$$H'=-(e^2\hbar^2/6M^2c^2)\Delta\{1/r\}, \quad (6.1)$$

where M is the common mass of the two particles with the understanding that the unperturbed Hamiltonian is that of the non-relativistic Schrödinger equation. The expectation value of H' is, therefore, according to Eqs. (6), (6.1)

$$\langle H'\rangle=(8\pi/3)(he/2Mc)^2\psi_S^2(0)$$

$$=(E_1-E_0)/4, \quad (6.2)$$

where ψ_S is the non-relativistic Schrödinger function for the relative motion of the two particles. The energy difference between levels having fine quantum numbers $f=1$ and 0 should be 4 times the above $\langle H'\rangle$, as is also indicated in Eq. (6.2). Since $he/2Mc$ is the magnetic moment appropriate to each particle, according to the Dirac equation which was supposed to be obeyed by each particle, this result is in agreement with Fermi's formula

$$E_{i+\frac{1}{2}}-E_{i-\frac{1}{2}}=(8\pi/3)(2+1/i)\mu\mu_0\psi_S^2(0), \quad (6.3)$$

where i is the nuclear spin and μ, μ_0 are, respectively, the magnetic moments of the nucleus and the electron. In the present case $i=\frac{1}{2}$. According

[7] G. Breit, Phys. Rev. **51**, 249 (1936), see Eqs. (17.7), (17.8). These are practically the same as the equations for spin-spin interactions of electrons referred to in reference 4. *Ibid.*, **53**, 153 (1938) pp. 159–160 discusses the relations of Casimir, Physica **3**, 936 (1936).

to Eq. (6.2) the hyperfine-structure splitting depends on the mass of one of the particles partly through μ and partly through $\psi_S^2(0)$. The latter dependence is taken care of by the transformation of Eq. (5.5″), which brings the factor $(1+\beta)^{-3}$ into the answer.

III. THE PAULI PART OF THE MOMENT. THE DEUTERON

Since the proton does not have a magnetic moment which would be expected for it according to Dirac's equation, it is necessary to take into account the additional moment of ~ 1.79 nuclear Bohr magnetons. An understanding of the origin of this addition is lacking, and the additional moment will be treated, therefore, as an intrinsic property of the proton. Pauli[8] gave an equation suitable for the representation of a particle obeying an equation of the general Dirac form but having in addition to the Bohr magneton an intrinsic moment. The excess of ~ 1.79 nuclear Bohr magnetons will be represented by a term in the Hamiltonian of the type invented by Pauli and will be referred to as the Pauli part of the proton's moment. The part of μ corresponding to it will be referred to as μ_P. The interaction with an external field is represented by an addition to the Hamiltonian

$$H_P' = \mu_P[\rho_{3M}(\mathfrak{H}\boldsymbol{\sigma}_M) - \rho_{2M}(\mathcal{E}\boldsymbol{\sigma}_M)], \qquad (7)$$

where for a discussion of a single particle ρ_2, ρ_3 are four row four column matrices introduced by Dirac in his original notation. The matrix ρ_3 is the same as β. The matrix ρ_2 is

$$\rho_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}. \qquad (7.1)$$

In Eq. (7) the matrices $\boldsymbol{\sigma}_M$ are four row matrices in the proton index identical with Dirac's vector matrix $\boldsymbol{\sigma}$. For two particles ρ_{2M} operates on the proton indices as though there were no electron and is diagonal in electron indices. For a single particle in an external field Eq. (7) gives a rigorously consistent relativistic description. If the electron's motion could be considered as preassigned one could substitute for $\mathcal{E}$ and $\mathfrak{H}$ the values for the electric and magnetic fields at the proton produced by the electron without bringing into question the propriety of employing Eq. (7). Since the electron's motion is not preassigned, however, one has to examine the errors introduced by borrowing an interaction Hamiltonian from single particle theory.

In employing Eq. (7) for two particles, the proton's equations of motion are still formally correct. The addition to the x component of the force on the proton brought about by H_P' is

$$\delta\dot{p}_{xM} = (i/\hbar)[H', p_{xM}] = -\mu_P[\rho_{3M}(\boldsymbol{\sigma}_M \partial\mathfrak{H}/\partial x_M) - \rho_{2M}(\boldsymbol{\sigma}_M \partial\mathcal{E}/\partial x_M)], \qquad (7.2)$$

corresponding to magnetic moment $-\mu_P\rho_{3M}\boldsymbol{\sigma}_M$ and electric moment $\mu_P\rho_{2M}\boldsymbol{\sigma}_M$. For a proton at rest the magnetic moment is represented by $\mu_P\boldsymbol{\sigma}_M$, and there is no electric moment. The densities of electric- and magnetic-moment distributions can be verified to transform themselves correctly under Lorentz transformations. The relation between $\dot{x}_M$ and α_M is left undisturbed by Eq. (7). From the point of view of the correspondence principle the proton's motion is, therefore, represented satisfactorily by Eq. (7).

For the consideration of the electron's motion Eq. (7) would be exact if the proton could be considered as stationary. The operator, ρ_{3M}, would then be equivalent to multiplication by -1, the operator ρ_{2M} to multiplication by 0. The interaction Hamiltonian in this limit becomes the same form as that commonly used for a stationary nucleus. It is well known that this form represents satisfactorily the electron's motion, the expressions for $\mathfrak{H}$ at the nucleus in terms of Dirac's α matrices for the electron being just right to make the magnetic vector potential of the nucleus appear in the right linear combination with electrons momentum operator. The inaccuracies of Eq. (7) for the discussion of the electron's motion have their origin, therefore, in the finiteness of the velocity of the nucleus. For a slowly moving nucleus these inaccuracies are small, because so far as the motion of the electron is concerned only the retardation of the effects of the nucleus at the electron caused by the finiteness of the velocity of light is neglected by the employment of Eq. (7). The retardation effects of nuclear motion

[8] W. Pauli, *Handbuch der Physik* (Verlagsbuchhandlung, Julius Springer, Berlin, 1933), Vol. 24/1, p. 221.

are of relative order $(m/M)^2(v^2/c^2)$, where v is the electron velocity, and are negligible since v^2/c^2 is $\sim\alpha^2$.

In terms of the electron operators employed in Eq. (1)

$$H_p'=\mu_P e[\rho_{3M}\sigma_M[\mathbf{r}\times\boldsymbol{\alpha}_e]-\rho_{2M}(\mathbf{r}\sigma_M)]/r^3. \quad (7.3)$$

The second term in the brackets requires special consideration. One finds by calculation that

$$\int\psi^\dagger\rho_{2M}(\mathbf{r}\sigma_M)r^{-3}\psi d\tau=(h/Mc)\text{R.P.}\int\Psi^\dagger r^{-3}\{(\mathbf{r}\nabla)+i\sigma_M[\mathbf{r}\times\nabla]\}\Psi d\tau, \quad (7.31)$$

where the σ_M are four row matrices in the nuclear subscript when they multiply ψ and two row matrices when they multiply Ψ. The quantity Φ has been eliminated in this calculation by means of Eq. (5.3). The integrand of the right side of Eq. (7.31) depends on the orientation of the nuclear spin only through the term in σ_M which contains also the orbital angular-momentum operator $[\mathbf{r}\times\mathbf{p}]$. Non-relativistically the latter factor is equal to zero for s terms. In the approximation of the present paper it is

$$(h/Mc)\int_0^\infty(-8\pi/3)(f^2/r)dr, \quad (7.32)$$

the expression having been evaluated for parallel spin orientations. This expression has to be multiplied by $(-\mu_P e)$ in order to give the contribution to the expectation value of H_P'. If the integration in the last formula were carried out for a value of f corresponding to a Coulomb field, one would obtain an infinite result. For a Coulomb potential modified so as to have finite value for $0<r<e^2/mc^2$ the integral is finite, and only a logarithmic term in mc^2/e^2 is brought in. The ratio f/g being of the order α, the contribution of the term under discussion to the hyperfine structure is of the order $(m/M)\alpha^2\log\alpha$ and may be neglected.

It may be noted that the effect just discussed is that of the mutual energy of the electron's electric field and of the nuclear electric dipole produced by the nuclear motion of translation as a result of the nuclear magnetic moment. In non-quantum analogy the effect is zero for s terms. It is also zero in any approximation in which L is a good quantum number.

The first term of Eq. (7.3) differs from the interaction Hamiltonian generally employed for a fixed nucleus only through the presence of the factor ρ_{3M}. The effect introduced by this factor is that of changing the sign of the terms containing Φ and is of the relative order $(m/M)^2\alpha^2$.

The effect of the electric-dipole moment induced by nuclear translation is present also for the Dirac part of the moment. Calculation shows, in fact, that the part of the particle-density formula which is spin dependent when expressed in terms of Ψ alone gives an electric-dipole contribution of the same order as the second term in brackets in Eq. (7.3). This effect is also of order $(m/M)\alpha^2\log\alpha$.

It is seen that for the Pauli part of the proton's moment the mass effect is represented by the same factor $(1+m/M)^3$ as for the Dirac part within terms of the order $(m/M)\alpha^2\log\alpha$.

All of the deuteron's moment is reasonably represented as an intrinsic one. The magnetic moment of the Dirac type does not enter the deuteron problem because the proton contained in the deuteron is moving in the field of the neutron. The circulation of the proton's charge contributes to the magnetic moment of the deuteron, but its perturbation by recoil action from the electron is negligibly small because of the much larger force to which the proton is subjected as a result of the forces acting on it within the deuteron.

A few arguments concerned with physical plausibility will be dealt with first. It will be supposed that the neutron is not interacting with the electron in any way except through the rather small effect of its magnetic moment. If the neutron's mass were negligible and if its mass were instead assigned to the proton, one could construct a substitute model for the deuteron which would behave in weak external fields in the same way as the deuteron. To obtain this result one would have to arrange for the magnetic model of the model to be the same as that of the deuteron. It is plausible to assume that such a model will be equivalent to the deuteron in its action on the electron. For the state with magnetic-quantum number 3/2 this model shows a simple relationship to the case of magnetic-

quantum number 1 which has been worked out for the case of the proton. The spins of the proton and neutron are parallel to the angular momentum of relative electronic motion. Since the neutron does not interact with the electron except through its magnetic moment, and since in the present model the neutron's mass has been transferred to the proton, the calculations made for ordinary hydrogen are directly applicable. The only changes that have to be made are: (a) the mass of the deuteron has to be substituted for the mass of the proton, (b) the magnetic moment of the neutron has to be added to the magnetic moment of the proton. The formula for the hyperfine-structure splitting expected on the basis of the above model is, therefore, the one derived by Fermi with the modification of the factor $(1+m/M_D)^{-3}$.

The applicability of the correction to deuterium can also be demonstrated along more rigorous lines. The Hamiltonian for three particles can be written down so as to describe the system of proton, neutron, and electron. The term corresponding to Y of the present paper is neglected at first. Solutions of the proton-neutron relative motion problem are introduced. These solutions have to be modified first for the effect of the momentum of the deuteron. The magnetic momentum changes slightly, and a small electric moment results. Both of these changes are insignificant as in the case of the single proton. The wave function of the whole system is expanded in products of the internal deuteron function and the function describing the relative motion of the electron with respect to the center of mass of the deuteron. Radial equations for relative motion analogous to those dealt with for the single proton are obtained. The term in Y for the proton, the Pauli part of the proton's magnetic moment, and the magnetic moment of the deuteron combine into one term corresponding to the magnetic moment of the deuteron. The result of this consideration is the same as that of the simplified model just considered.[9]

A preliminary report on the present work was made at the conference on The Foundations of Quantum Mechanics held at Shelter Island, New York on June 2–4, 1947. The conference was sponsored by the National Academy of Sciences at the suggestion of Dr. D. A. MacInnes who has organized it. It is a pleasure to acknowledge the stimulus derived from the conference for finishing this work.

[9] After this manuscript had been completed there appeared a note on the same subject by O. Halpern in Phys. Rev. **72**, 245(L) (1947). The results obtained in the present paper differ from Halpern's.