

Direct Exchange in Ferromagnets

R. STUART

Lawrence Radiation Laboratory, Livermore, California

AND

W. MARSHALL

Atomic Energy Research Establishment, Harwell, England

(Received June 6, 1960)

The direct exchange integral which occurs in the Heisenberg theory of ferromagnetism is evaluated for all internuclear spacings. We find that it is always positive, whereas Bethe originally suggested it would be positive only at large spacing and more recently it has been suggested that the integral should always be negative. However, at the observed internuclear separation the magnitude calculated is of the order of 70 times too small to explain the experimentally determined exchange constant in ferromagnetic metals, and we therefore conclude that direct exchange is not responsible for ferromagnetism in these metals.

I. INTRODUCTION

EVER since Heisenberg¹ explained how exchange effects could account for ferromagnetism, there has been controversy over the sign of the direct-exchange integral which appears in the Heisenberg theory. Using a simple Heitler-London model with one electron per atom, Heisenberg showed that the spin coupling between two nearest-neighbor atoms A and B of separation R was given by the exchange integral

$$J(A,B) = \iint d\mathbf{r}_1 d\mathbf{r}_2 B(\mathbf{r}_1)A(\mathbf{r}_2) \times \left[\frac{e^2}{R} + \frac{e^2}{r_{12}} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} \right] B(\mathbf{r}_2)A(\mathbf{r}_1), \quad (1)$$

where \mathbf{r}_1 and \mathbf{r}_2 are the coordinates of the two electrons and $A(\mathbf{r})$ and $B(\mathbf{r})$ are the wave functions (assumed real) centered one on each nucleus. The first two terms of this expression are positive, the last two are negative, and clearly all four are of the same order of magnitude. The sign of this integral is therefore not easy to estimate and, in the Heisenberg theory, ferromagnetism results only if it is positive.

It is exactly this integral which occurs in the theory of the hydrogen molecule where the wave functions A and B are the $1s$ orbitals of the hydrogen atoms. For this latter problem J is known to be negative [singlet state lowest] but Heisenberg postulated that J was positive for wave functions and lattice spacings appropriate for the ferromagnetic elements. It was left to Bethe² to give plausible arguments why J should be positive. He pointed out that the second term of (1) would dominate the others if all the overlap between A and B were concentrated in a small region away from the nuclei. He suggested that this would be the case if the wave functions A and B had (a) small amplitude at their parent nuclei, (b) angular lobes pointing towards and

overlapping one another, and (c) small radial extent compared to the internuclear spacing R .

All three conditions are satisfied by the ferromagnetic metals. In these metals the magnetic electrons are in $3d$ wave functions and near the nucleus these are small (like r^2) so (a) is satisfied. These $3d$ wave functions also have strong angular lobes which can be made to overlap one another so (b) is also satisfied. Detailed estimates by Slater³ show that the ratio of the mean radius of these wave functions compared to the internuclear separation is smaller for the three ferromagnetic metals than for any other metal. Hence condition (c) is also well satisfied. It is clear that in general it is hard to satisfy all three conditions simultaneously so it seemed not surprising that only three metals were ferromagnetic.

From general arguments Bethe was able to suggest how J varied with internuclear spacing R . At small R , he concluded J was negative because (c) was not satisfied: as R was increased he suggested that J approached zero, became positive, passed through a maximum and then decayed exponentially to zero. Bethe gave a rough sketch of this behavior and this has become known as the Bethe⁴ curve.

Subsequently the validity of Bethe's arguments was severely questioned and it was suggested by Van Vleck,⁵ by Slater⁶ himself, and by Zener,⁷ that J could never be positive but simply varied monotonically from a large negative value at small R to an exponentially small negative value at large R .

Because of this controversy an explicit calculation of J is obviously of value but the integrals involved appeared so difficult that only two such calculations have been made. The first was by Wohlfarth⁸ who concluded that J was always negative. However, for ease of calcu-

³ J. C. Slater, *Phys. Rev.* **36**, 57 (1930).

⁴ R. M. Bozorth, *Ferromagnetism* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1951), p. 444.

⁵ J. H. Van Vleck, *Revs. Modern Phys.* **25**, 220 (1953).

⁶ J. C. Slater, *Phys. Rev.* **49**, 537 (1936); *Revs. Modern Phys.* **25**, 199 (1953).

⁷ C. Zener, *Phys. Rev.* **81**, 440 (1951); **82**, 403 (1951); **83**, 299 (1951).

⁸ E. P. Wohlfarth, *Nature* **163**, 57 (1949).

¹ W. Heisenberg, *Z. Physik* **49**, 619 (1928).

² A. Sommerfeld and H. Bethe, *Handbuch der Physik*, edited by S. Flügge (Verlag Julius Springer, Berlin, 1933), Vol. 24, Part II, p. 595.

lation he replaced his $3d$ wave functions by spherically averaged functions. Because this violates condition (b), his result is not surprising. The other calculation was by Kaplan⁹ who took proper account of the angular variation of the wave functions but, because of the difficulties of the calculation, made only one accurate estimate of J and this was for the small internuclear separation of $1.5a_0$. This value was positive and equal to 0.60 ev. For larger values of R , Kaplan was unable to make any reliable estimates of J but concluded that J probably remained positive out to the observed internuclear separation of $4.7a_0$.

In this paper we describe an explicit calculation of J for all R . We do not use the Barnett-Coulson¹⁰ method of evaluation but after some preliminary algebra place the calculation directly onto an IBM 704. In the Appendix the procedure adopted is described in some detail because we believe that this direct method may often be preferable to the Barnett-Coulson approach and could be used for a wide variety of problems. In the next section we define the various integrals and give the results and these are discussed in Sec. III.

However, we should first remark that there is no absolute definition of "direct exchange"; the exchange energies in one model can occur as other energies in another model and only the total energy has an absolute meaning. We examine here what has, by convention, become known as direct exchange in the theory of ferromagnetism using a strict Heitler-London model.

II. THEORY AND RESULTS

Before discussing the precise form of the integrals involved, it is well to remark that J as given by (1) is only an approximate form for the spin-spin coupling parameter. Indeed this is evident from the appearance of the term e^2/R in (1) because the effective spin coupling cannot depend directly on the interaction energy between the nuclei. It is easy to show that if we define the effective spin coupling between orbits A and B by

$$H = -2J'(A,B)\mathbf{S}_A \cdot \mathbf{S}_B, \quad (2)$$

where \mathbf{S}_A and \mathbf{S}_B are the spins in orbits A and B , respectively, then

$$J'(A,B) = \{J(A,B) - S^2(A,B)C(A,B)\} \times \{1 - S^4(A,B)\}^{-1}, \quad (3)$$

where $C(A,B)$ is the Coulomb interaction

$$C(A,B) = \iint d\mathbf{r}_1 d\mathbf{r}_2 B^2(\mathbf{r}_2) \times \left[\frac{e^2}{R} + \frac{e^2}{r_{12}} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} \right] A^2(\mathbf{r}_1), \quad (4)$$

⁹ H. Kaplan, Phys. Rev. **85**, 1038 (1952).

¹⁰ M. P. Barnett and C. A. Coulson, Phil. Trans. Roy. Soc. London **A243**, 221 (1951).

and $S(A,B)$ is the overlap integral

$$S(A,B) = \int d\mathbf{r}_1 B(\mathbf{r}_1)A(\mathbf{r}_1). \quad (5)$$

Our results will show $J'(A,B)$ differs appreciably from $J(A,B)$ only at small R so that for all practical purposes the correction can be ignored.

The first term of (1) we call $J_1(A,B)$ and is simply $e^2 S^2(A,B)/R$; the second term is difficult to evaluate and we call it $J_2(A,B)$; the last two terms are equal and their sum $J_3(A,B)$ is $-2S(A,B)P(A,B)$ where $P(A,B)$ is defined by

$$P(A,B) = \int d\mathbf{r}_1 B(\mathbf{r}_1)A(\mathbf{r}_1)e^2/r_{1B}. \quad (6)$$

We label the various terms of $C(A,B)$ using a similar notation; thus $C_1(A,B)$ is simply e^2/R , $C_2(A,B)$ is the integral involving e^2/r_{12} , and $C_3(A,B)$ is the sum of the last two terms of (4) and given by

$$C_3(A,B) = -2 \int d\mathbf{r}_1 A^2(\mathbf{r}_1)e^2/r_{1B}. \quad (7)$$

The geometry of the problem is illustrated in Fig. 1. We imagine nucleus B placed at distance R along the z axis relative to A ; \mathbf{r}_{1A} and \mathbf{r}_{1B} are the vector distances of electron 1 from the nuclei A and B , respectively; \mathbf{r}_{2A} and \mathbf{r}_{2B} are defined similarly. For this problem the functions A and B are the $3d$ wave functions of iron and these are conveniently taken as real and given by

$$\begin{aligned} d_{z^2}(\mathbf{r}) &= (5/16\pi)^{1/2}(3 \cos^2\theta - 1)P(r)/r, \\ d_{x^2-y^2}(\mathbf{r}) &= (15/16\pi)^{1/2} \sin^2\theta(\cos^2\varphi - \sin^2\varphi)P(r)/r, \\ d_{xy}(\mathbf{r}) &= (15/4\pi)^{1/2} \sin^2\theta \cos\varphi \sin\varphi P(r)/r, \\ d_{yz}(\mathbf{r}) &= (15/4\pi)^{1/2} \cos\theta \sin\theta \sin\varphi P(r)/r, \\ d_{zx}(\mathbf{r}) &= (15/4\pi)^{1/2} \cos\theta \sin\theta \cos\varphi P(r)/r. \end{aligned} \quad (8)$$

The radial function $P(r)$ is given by Wood and Pratt¹¹ and agrees very well with that determined from neutron diffraction form factor measurements.^{12,13}

Choosing various pairs of these wave functions for A and B in (1), it is possible to construct 25 exchange integrals, eleven of which are independent. The largest overlap integral is given by choosing d_{z^2} for both A and

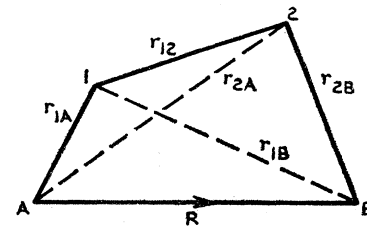
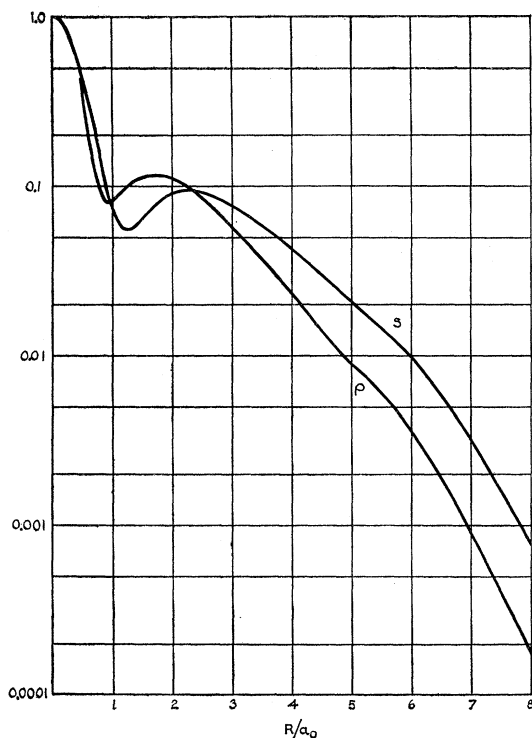


FIG. 1. The geometry of the problem.

¹¹ J. H. Wood and G. W. Pratt, Jr., Phys. Rev. **107**, 995 (1957).

¹² R. Nathans, C. G. Shull, G. Shirane, and A. Anderson, J. Phys. Chem. Solids **10**, 138 (1959).

¹³ R. J. Weiss and A. J. Freeman, J. Phys. Chem. of Solids **10**, 147 (1959).

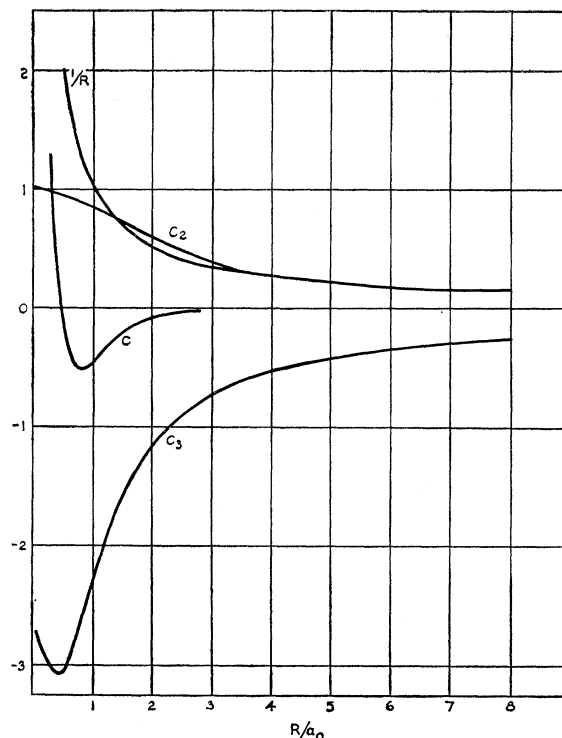
FIG. 2. $P(z^2, z^2)$ in units of e^2/a_0 , and $S(z^2, z^2)$.

B and we may confidently expect the corresponding exchange integral $J(z^2, z^2)$ to be much larger than the others; this is the integral we calculate in this paper. The only other choices for A, B which give nonzero overlap are yz, yz and zx, zx with equal exchange integrals $J(yz, yz)$ and $J(zx, zx)$, respectively, but these are certainly much smaller than $J(z^2, z^2)$. For all other choices of the pair A, B the overlap integral vanishes and the corresponding exchange coupling (1) reduces to $J_2(A, B)$ alone; this is certainly positive, for the second term of (1) is positive definite, and therefore assists ferromagnetism but the overlap between the wave functions is negligibly small everywhere so we may neglect all such contributions just as we neglect $J(yz, yz)$ and $J(zx, zx)$. [To be quite sure on this point, we have evaluated $J(z^2, zx) = J_2(z^2, zx)$ at the observed spacing of $4.7a_0$ and found it to be only 2% of $J(z^2, z^2)$.]

In calculating $J(z^2, z^2)$ it is easy to obtain $S(z^2, z^2)$, $P(z^2, z^2)$ and $C_3(z^2, z^2)$; $C_2(z^2, z^2)$ is only moderately difficult (details are given in Appendix B) but $J_2(z^2, z^2)$ reduces to a difficult five-dimensional integral which is evaluated by the method described in Appendix A. The results are shown in Figs. 2, 3, 4, and 5.

III. DISCUSSION

The first point to notice is that $J(z^2, z^2)$ and $J'(z^2, z^2)$ are positive for all R ; thus Bethe's qualitative arguments (a) and (b) are proved correct but his argument (c) [which would lead to negative $J(z^2, z^2)$ at small R] is proved incorrect. The small local minimum which

FIG. 3. $1/R$, $C_2(z^2, z^2)$, $C_3(z^2, z^2)$, and $C(z^2, z^2)$ in units of e^2/a_0 .

appears in $J(z^2, z^2)$ at $R = 1.1a_0$ is a consequence of the similar local minimum which appears in the overlap integral. It occurs at a distance where the positive and negative lobes of $A = d_{z^2}$ and $B = d_{z^2}$ overlap appreciably; as R is increased the magnitude of this negative overlap contribution falls very rapidly so that the total overlap actually increases and thus gives rise to the local minimum and maximum in $S(z^2, z^2)$ and $J(z^2, z^2)$. This minimum and maximum are of no practical significance because they occur at such small R values. At the observed interatomic spacing in Fe, $4.7a_0$, $J(z^2, z^2)$ has the value of 6.8×10^{-3} ev. We have not repeated the calculation for Ni or for Co but for these elements the variation of $J(z^2, z^2)$ with R must be qualitatively similar to that for Fe and roughly the same $J(z^2, z^2)$ value may be assumed for them.

It is very important to recognize that this calculated value for $J(z^2, z^2)$ must not be compared directly to the observed exchange coupling between atoms because we have not yet taken proper account of the probability that the two d_{z^2} orbitals are occupied simultaneously by a single electron. For simplicity let us first consider Ni metal: We assume a Van Vleck model for Ni where each atom has either a $(3d)^9$ or $(3d)^{10}$ configuration. The coupling between two nearest-neighbor $(3d)^9$ configurations we describe by the usual spin Hamiltonian

$$H = -2JS_1 \cdot S_2, \quad (9)$$

where S_1 and S_2 are of magnitude $\frac{1}{2}$ and represent the spins on atoms A and B , respectively [(9) should be

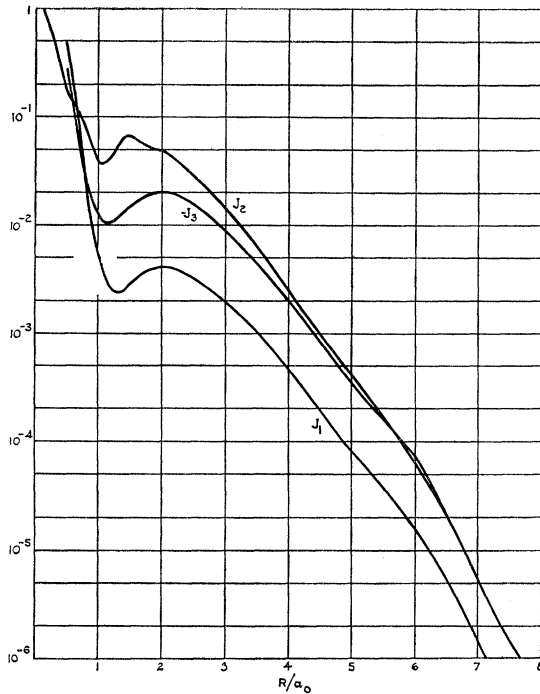


FIG. 4. $J_1(z^2, z^2)$, $J_2(z^2, z^2)$, and $-J_3(z^2, z^2)$ in units of e^2/a_0 .

clearly distinguished from (2); in the latter S_A and S_B denote the spins associated with the particular orbitals A and B , respectively]. Neutron diffraction experiments^{12,13} show that there is essentially no preference for the $3d$ "hole" to occupy any particular orbital of (8). Assuming therefore that each $3d$ orbital has a equal probability of 0.2 of being occupied by the hole, we deduce that the probability of the two z^2 orbitals being occupied simultaneously is 0.04. Hence

$$J = 0.04J(z^2, z^2). \quad (10)$$

Using a value of 6.8×10^{-3} ev for $J(z^2, z^2)$ we get 2.72×10^{-4} ev for J . This is to be compared to the experimental results of 0.020 ev from spin wave measurements and 0.010 ev from specific heat measurements.¹⁴ We therefore conclude that direct exchange fails by a factor of about 70 to account for the observed coupling.

In Fe we again use the Hamiltonian (9) but this time S_1 and S_2 have magnitude 1. Because experiment^{12,13} tells us the $3d$ orbitals are equally occupied, we assume equal probabilities for all occupation assignments. On each atom there are two holes which may be placed in any of 10 distinct pairs; four of these pairs involve the z^2 orbital so the probability of having a hole in z^2 is 0.4. The probability of simultaneous occupation of the z^2 of A and that of B is therefore 0.16 and for each such case S_A is $0.5S_1$ and S_B is $0.5S_2$; hence (10) is correct for Fe also. A similar argument for Co, assumed to involve a mixture of $(3d)^8$ and $(3d)^9$ configurations, shows that

¹⁴ J. A. Hofmann, A. Paskin, K. J. Tauer, and R. J. Weiss, *J. Phys. Chem. of Solids* **1**, 45 (1956).

(10) is again applicable. The observed value of J for Fe is 0.018 ev from spin wave measurements or 0.011 ev from specific heat measurements,¹⁴ and hence we conclude that direct exchange fails by a factor of order 70 to explain the observed coupling in all these cases.

It is interesting to remark that this explicit calculation justifies the somewhat intuitive arguments which have been put forward for why direct exchange is unimportant; but this is not for the reason suggested [that J should be negative] but because the calculated J is too small. However, we must also point out that Zener⁷ has argued that J should be negative explicitly because screening by the other $3d$ electrons of the atom is not complete. We have not allowed for this effect in our calculation and therefore cannot say if it would or would not make J negative; indeed, for our immediate purposes the question is uninteresting because we have already concluded that direct exchange cannot be responsible for ferromagnetism even if the effect is ignored.

In view of these results we believe that the true origin of ferromagnetism in these metals must be found in what Van Vleck and Slater have called "intra-atomic" exchange. This may be thought of as the effects introduced by departures from the strict Heitler-London model that we have used in this paper.

It is worth noting that the individual exchange integral $J(z^2, z^2)$ is quite large [one third of the observed

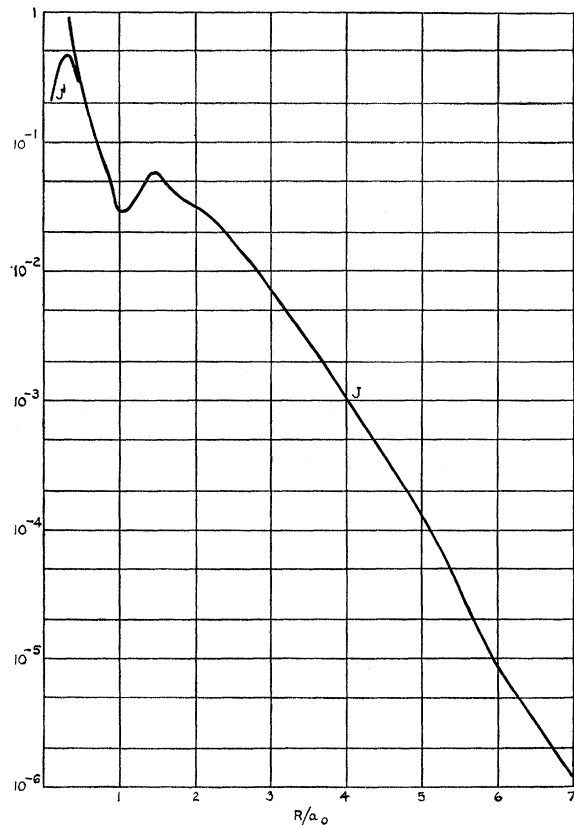


FIG. 5. $J(z^2, z^2)$ and $J'(z^2, z^2)$ in units of e^2/a_0 .

value] and it is only the small probability, 0.04, of the holes appearing simultaneously in the z^2 orbitals which renders this direct-exchange mechanism unimportant in this case. In salts the electrons may be in fixed orbitals and then, for favorable cases, direct exchange may be of importance.

APPENDIX A

Using the coordinate transformation

$$\mathbf{r} = \frac{1}{2}(\mathbf{r}_{1A} - \mathbf{r}_{2B} - \mathbf{R}_{AB}),$$

$$\mathbf{R} = \frac{1}{2}(\mathbf{r}_{1A} + \mathbf{r}_{2B}),$$

the integral $J_2(z^2, z^2)$ can be put in the following form:

$$J_2(z^2, z^2) = \frac{25e^2}{4\pi} \int_0^1 d\mu \int_0^1 dM \int_0^\pi d\phi \int_0^{R_{\max}} R^2 dR$$

$$\times \int_0^{\tau_{\max}} r dr (3r_{1B}^2 \cos^2\theta_{1B} - r_{1B}^2)$$

$$\times r_{1B}^{-3} P(r_{1B}) (3r_{2A}^2 \cos^2\theta_{2A} - r_{2A}^2) r_{2A}^{-3} P(r_{2A})$$

$$\times (3r_{1A}^2 \cos^2\theta_{1A} - r_{1A}^2) r_{1A}^{-3} P(r_{1A})$$

$$\times (3r_{2B}^2 \cos^2\theta_{2B} - r_{2B}^2) r_{2B}^{-3} P(r_{2B}),$$

where

$$r_{1A}^2 = R^2 + r^2 + a^2 + 2Rrs + 2RaM + 2ra\mu,$$

$$r_{2B}^2 = R^2 + r^2 + a^2 - 2Rrs - 2RaM + 2ra\mu,$$

$$r_{1B}^2 = R^2 + r^2 + a^2 + 2Rrs - 2RaM + 2ra\mu,$$

$$r_{2A}^2 = R^2 + r^2 + a^2 - 2Rrs + 2RaM - 2ra\mu,$$

$$r_{1A} \cos\theta_{1A} = RM + r\mu + a,$$

$$r_{2B} \cos\theta_{2B} = RM - r\mu - a,$$

$$r_{1B} \cos\theta_{1B} = RM + r\mu - a,$$

$$r_{2A} \cos\theta_{2A} = RM - r\mu + a,$$

$$s = [(1 - M^2)(1 - \mu^2)]^{\frac{1}{2}},$$

$$a = \frac{1}{2}R_{AB},$$

$$\mu = \cos\theta,$$

$$M = \cos\Theta,$$

$$\mathbf{r} = \mathbf{r}(r, \theta, \phi),$$

$$\mathbf{R} = \mathbf{R}(R, \Theta, \Phi).$$

The limits of integration R_{\max} and r_{\max} are determined by the cutoff of the tabulated function $P(r)$ at some value $r = X_0$. Thus the value of R_{\max} is found from the relation

$$(1 - M^2)R_{\max}^2 + (MR_{\max} + a)^2 = X_0^2,$$

or

$$R_{\max} = -aM + [a^2M^2 + X_0^2 - a^2]^{\frac{1}{2}}.$$

Similarly, r_{\max} is the largest value of r such that r_{1A} , r_{2B} , r_{1B} , or r_{2A} equals X_0 . We easily get

$$r_{\max} = \text{smaller of } (C_1, C_2),$$

where

$$C_1 = -|Rs + a\mu| + [(Rs + a\mu)^2 - R^2 - a^2 - 2aRM + X_0^2]^{\frac{1}{2}},$$

$$C_2 = -|Rs - a\mu| + [(Rs - a\mu)^2 - R^2 - a^2 + 2aRM + X_0^2]^{\frac{1}{2}}.$$

The expression for $J_2(z^2, z^2)$ was programmed for the IBM 704. The integrations over the variables μ , M , ϕ , and R were performed by using the Gaussian quadrature method; the trapezoidal rule was used to integrate over r . For $R_{AB} = 4a_0$ the following number of integration points was used; 4 points for μ , 5 points for M , 4 points for ϕ , 10 points for R ; and a Δr equal to 0.15 was used in the trapezoidal integration of r . The accuracy of the procedure was checked by making numerous runs with various mesh dimensions, and the above-mentioned mesh size was found to give an answer which was accurate to better than 0.1%. For this case, a running time of five minutes was required. More mesh points and a longer running time (up to thirty-five minutes) were needed for the smallest values of R_{AB} .

APPENDIX B

The integral

$$C_2(z^2, z^2) = \left(\frac{5e}{16\pi}\right)^2 \int_0^{X_0} r_{1A}^2 dr_{1A} \int_{-1}^1 d\mu_{1A}$$

$$\times \int_0^{2\pi} d\phi_{1A} \int_0^{X_0} r_{2A}^2 dr_{2A} \int_{-1}^1 d\mu_{2A} \int_0^{2\pi} d\phi_{2A}$$

$$\times (3\mu_{1A}^2 - 1)^2 r_{1A}^{-2} P^2(r_{1A})$$

$$\times (3\mu_{2B}^2 - 1)^2 r_{2B}^{-2} P^2(r_{2B}) |r_{1A} - r_{2A}|^{-1}$$

we evaluated in the following manner:

We make the expansion

$$(3\mu_{2B}^2 - 1) r_{2B}^{-2} P^2(r_{2B}) = \frac{1}{2} \sum_{n=0}^{\infty} (2n+1) P_n(\mu_{2A}) F_n(r_{2A}),$$

where

$$F_n(r_{2A}) = \int_{-1}^1 d\mu_{2B} P_n(\mu_{2A}) r_{2B}^{-2} P^2(r_{2B}) (3\mu_{2B}^2 - 1).$$

The integrand is independent of ϕ_{1A} and ϕ_{2A} except for the term $|r_{1A} - r_{2A}|^{-1}$. Consequently, that term can be replaced in the integrand by $\sum_{l=0}^{\infty} P_l(\mu_{1A}) P_l(\mu_{2A}) \times g_l(r_{1A}, r_{2A})$, where

$$g_l(r_{1A}, r_{2A}) = r_{1A}^l r_{2A}^{-l-1} \quad \text{for } r_{1A} < r_{2A},$$

$$= r_{2A}^l r_{1A}^{-l-1} \quad \text{for } r_{2A} < r_{1A}.$$

Integration over the four angular coordinates gives the form of $C_2(z^2, z^2)$ which was evaluated on the IBM 704:

$$C_2(z^2, z^2) = (5e^2/28) \int_0^{X_0} dr_{1A} P^2(r_{1A}) \int_0^{X_0} dr_{2A} r_{2A}^2$$

$$\times \left[\frac{1}{2} M g_0(r_{1A}, r_{2A}) F_0(r_{2A}) + g_2(r_{1A}, r_{2A}) F_2(r_{2A}) \right.$$

$$\left. + g_4(r_{1A}, r_{2A}) F_4(r_{2A}) \right].$$