Fine Structure of Helium*†

CHARLES SCHWARTZ
Department of Physics, University of California, Berkeley, California
(Received 6 January 1964)

Under the expectation that experiments will soon give values for the fine structure intervals of the 2^3P state of helium to an accuracy of $1/10^6$, we have undertaken a program of calculations which, it is hoped, will lead to a new determination of the fine-structure constant α. This paper gives a brief survey of the over-all program, and a detailed report of the successful completion of the first task: The construction of approximate solutions to the Schrödinger equation which lead to average values of the leading fine structure operators accurate to about one part per million.

I. INTRODUCTION

THE early studies of the fine structure of the helium atom by Breit[1] provided very important confirmation of the newly developing theory of the quantum interactions of electrons and the electromagnetic field. Over the intervening years there have been several efforts to increase the accuracy of this analysis, chiefly by the construction of successively more complicated approximations to the nonrelativistic 2-electron wave function. Breit,[1] with a 2-term trial function got values of the fine-structure intervals ν (see Fig. 1) accurate to several percent. Araki[2] and co-workers with 8 terms came within about 1%, and also considered two small corrections to the first-order theory. Traub and Foley,[3] with an 18-parameter function reduced the error to about one part in 10^3; and most recently Pekeris, Schiff, and Lifson,[4] with up to 220 terms in the trial function, came within one part in 10^4 of the final results we have attained. (These numbers refer to the larger interval ν_{01}. In the calculation of the small interval ν_{12} there is a cancellation between separate terms amounting to a full order of magnitude, and a corresponding loss of accuracy results.)

While these studies have been chiefly aimed at checking the theory, we here take another position: Assuming the current theory of quantum electrodynamics, one can combine the calculated and measured fine-structure intervals of helium in order to determine the value of the fine-structure constant α. At present, the best value of α comes from the work of Dayhoff, Triebwasser, and Lamb[5] on the fine structure $(2P_{1/2}-2P_{3/2})$ of hydrogen. That value is uncertain to about one part in 10^5, due essentially to the short lifetime of the $2p$ state, and it has not appeared feasible to improve on those measurements. However, the lifetime of the ^3P state of helium is about 2 orders of magnitude longer, and it is expected[6] that these fine structure intervals can be measured to an accuracy of one part in 10^6 or perhaps better.[7] It then becomes necessary to calculate theoretical values for the helium fine structure to $1/10^6$, and this is the task we here commence. It may be remarked that in the case of the hydrogen fine structure the theoretical formula[8] is very simple; this is not the case for helium. However, we believe that those circumstances which made the hydrogen problem simple, will have the effect of reducing the helium calculation from "impossibly difficult" to merely "difficult."

We should add that there is at present a specific need for a better value of α. Analysis[9] of the hyperfine structure of hydrogen reveals effects due to the electro-

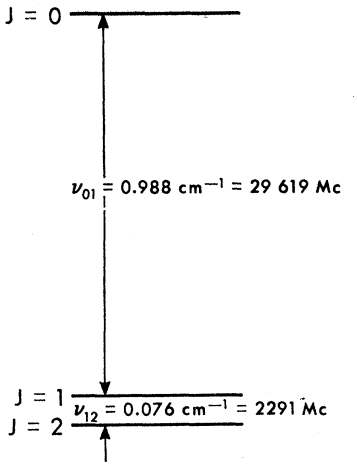

FIG. 1. Energy levels of the $1s2p^3P$ state of helium.

* This work was supported in part by the U. S. Atomic Energy Commission and the U. S. Air Force under Grant AF-AFOSR 130-63, and in part by the Advanced Research Projects Administration through the U. S. Office of Naval Research.

† A preliminary report of this work was presented at the 1963 Annual Meeting of the American Physical Society, Bull. Am. Phys. Soc. 8, 20 (1963).

[1] G. Breit, Phys. Rev. **36**, 383 (1930). References to other early studies may be found in this reference.

[2] G. Araki, M. Ohta, and K. Mano, Phys. Rev. **116**, 651 (1959), and several earlier papers referred to therein.

[3] J. Traub and H. M. Foley, Phys. Rev. **116**, 914 (1959).

[4] C. L. Pekeris, B. Schiff, and H. Lifson, Phys. Rev. **126**, 1057 (1962).

[5] E. S. Dayhoff, S. Triebwasser, and W. E. Lamb, Phys. Rev. **89**, 106 (1953).

[6] V. W. Hughes (private communication).

[7] For the best measurements to date (accurate to about $0.5/10^4$), see J. Lifsitz and R. H. Sands, Bull. Am. Phys. Soc. **6**, 424 (1961), and earlier references given therein.

[8] See, for example, H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic Press Inc., New York, 1957), p. 105.

[9] C. K. Iddings and P. M. Platzman, Phys. Rev. **115**, 919 (1959); also see D. E. Zwanziger, Bull. Am. Phys. Soc. **6**, 514 1961).

magnetic structure of the proton, but a clear interpretation of those results is spoiled by the present uncertainty in the value of α.

II. PLAN OF THE CALCULATIONS

We will now outline the several parts of the task of calculating the fine structure intervals in helium to an accuracy of about one part in 10^6. The zeroth-order problem is the nonrelativistic two-electron Hamiltonian in a fixed Coulomb field:

$$H_0=\frac{p_1^2}{2m}+\frac{p_2^2}{2m}-\frac{Ze^2}{r_1}-\frac{Ze^2}{r_2}+\frac{e^2}{r_{12}}. \tag{1}$$

The next terms in a nonrelativistic expansion are the well-known[10] fine-structure terms which we shall denote as α^2H_2. [Our reference for energy is $e^2/a_0=\alpha^2mc^2$, and we shall not note factors of Z $(=2)$ as being pertinent to the classification of the expansion terms.] We thus expect an energy level formula something like

$$E_J-E_0=\alpha^2\langle H_2\rangle_J+\alpha^4\left[\langle H_4\rangle_J+\left\langle H_2\frac{1}{E_0-H_0}H_2\right\rangle_J\right]$$
$$+(\text{terms of order }\alpha^5\text{ and smaller which we shall ignore}), \tag{2}$$

where H_4 is some higher order operator which has not yet been worked out.

There are still many terms not represented in this formula (2). The most apparent is the correction for the anomalous magnetic moment of the electron, of order α^3. This is quite simple and has already been included by some authors.[2,4] The reduced mass must be put in properly, and there is also the operator $\mathbf{p}_1\cdot\mathbf{p}_2/M$ correcting for the motion of the nucleus; this is itself spin-independent but will contribute in a second-order calculation mixed with H_2. There are then very many diagrams one could write down describing radiative corrections of many higher orders.

The most extensive work on the fine structure and Lamb shift in hydrogen has produced terms up to, but not including, the order α^7mc^2,[11] which is just to the same order that we now need. But the two-electron atom seems so much more complicated than the one-electron atom that our project might appear too difficult to attempt. A thorough relativistic theory of the two-electron atom has thus far been analyzed[12] only as far as the leading Lamb-shift terms, $\alpha^5mc^2\times(\ln\alpha+\text{const})$, and our goal is well beyond this. What encourages us to proceed is the expectation that most of the difficult terms at these higher orders are spin-independent, and thus, while affecting the absolute energy level, do not contribute to the fine structure.

As an example, consider the effect on the fine structure of some deviation from pure Coulomb potentials occurring at some small distance R. The integrals which occur are of the general form of the familiar $\langle(1/r)\times(\partial V/\partial r)\rangle\sim\langle 1/r^3\rangle$, and are taken over radial wave functions for $l>0$ orbitals. Thus the relative correction would appear to be of the order

$$\int_0^R r^2dr r^{2l-3}\Big/\int_0^a r^2dr r^{2l-3}\approx(R/a)^{2l}\leqq(R/a)^2.$$

Thus the effect of finite nuclear size would seem to be completely negligible, and the effect of vacuum polarization would appear to be of order $\alpha(\alpha)^2=\alpha^3$ on the fine structure intervals, again negligible.

The entire project may be divided into three rather well-separated parts. First, is the task of constructing a sufficiently accurate eigenfunction of H_0 so that the leading term in Eq. (2) can be determined to about $1/10^6$; this is mostly a computer problem. Second, is the task of carrying out a proper relativistic analysis and determining the operator H_4 and whatever else may appear to the required order; this will be a matter needing both formal technique and enlightened short-cutting. The third job will be the evaluation of these smaller corrections to the fine structure, and this will require a combination of modest computer effort and much algebraic detail.

The rest of this paper is concerned with the first of these tasks.

III. THE SCHRÖDINGER WAVE FUNCTION: FIRST ATTEMPT

We wish to construct a good approximation to the lowest ^3P eigenfunction of the equation

$$H_0\psi=E_0\psi. \tag{3}$$

This will be done by setting up a sequence of trial functions ψ, using the variational principle

$$\delta\langle\psi|H_0-E_0|\psi\rangle=0 \tag{4}$$

and seeing how results of interest converge as we make the trial functions larger and larger.

We started with the functions

$$\psi=\left(\frac{1-P_{12}}{4\pi\sqrt{2}}\right)\sum_{l,m,n=0}^{l+m+n\leqq\omega}C_{lmn}\mathbf{r}_1r_1^mr_2^nr_{12}^l\times e^{-(\kappa\sigma/2)r_1}e^{-(\kappa/2)r_2}, \tag{5}$$

where P_{12} exchanges coordinated r_1 and r_2, and the P-state character is represented by the vector sign. The unit of length is $a_0=\hbar^2/me^2$; and the scale parameters κ and σ were taken, by extrapolation from the results of

[10] See Ref. 8, p. 181.

[11] A. J. Layzer, J. Math. Phys. **2**, 308 (1961), has calculated terms of order $mc^2\alpha^7\ln^2\alpha$ and $mc^2\alpha^7\ln\alpha$. Only the latter contributes to the fine structure, and we hope it will not be difficult to include the term of this order in our problem.

[12] J. Sucher, Phys. Rev. **109**, 1010 (1958).

TABLE I. Result of variational calculations of $2\,^3P$ with the standard Hylleraas basis (5).

ω	Number of terms	Energy $-E(e^2/a_0)$	$C(\frac{1}{2}\alpha^2\,\mathrm{Ry})$	$D(\frac{1}{2}\alpha^2\,\mathrm{Ry})$
2	10	2.132678402	−0.07065132	−0.05404532
3	20	2.133085039	−0.06829756	−0.05379667
4	35	2.133140222	−0.06771029	−0.05394716
5	56	2.133157595	−0.06745614	−0.05400336
6	84	2.133162268	−0.06733560	−0.05402833
7	120	2.133163594	−0.06728399	−0.05403823
8	165	2.1331639812	−0.06726722	−0.05404284
9	220	2.1331641067	−0.06726320	−0.05404503
10	286	2.1331641530	−0.06726453	−0.05404627
Final result from Table II		2.1331641908	−0.06727529	−0.05404839

Traub and Foley,[3] to be 4.62 and 0.29,[13] respectively, and were not varied. The variation of the linear parameters C_{lmn} leaves us then with the numerical problem of finding the eigenvalue and eigenvector of a symmetric matrix of degree

$$N(\omega)=\tfrac{1}{6}(\omega+1)(\omega+2)(\omega+3). \tag{6}$$

Some of the details of the construction of the matrix elements and the numerical methods used will be found in the Appendix. We give in Table I the resulting eigenvalues for the series of trial functions of degree $\omega=2,\ 3,\ \cdots,\ 10$. This was done with 16-decimal arithmetic on an IBM-704 and consumed a total of some 20 h of machine time.

This work very closely parallels that of Pekeris, Schiff, and Lifson,[4] differing only in the specification of the exponential parameters, κ and σ. This makes a non-negligible difference: At 220 terms our eigenvalue is an order of magnitude better converged than theirs.[14]

Our chief interest anyway is not the energy eigenvalue, but the eigenvector, or more properly the average values of the two spin-dependent parts of H_2. These are the well-known[10] spin-orbit terms

$$H(\text{s.o.})=\left(\frac{e\hbar}{2mc}\right)^2\left(\frac{\boldsymbol{\sigma}_1+\boldsymbol{\sigma}_2}{2}\right)\cdot\left[Z\left(\frac{\mathbf{r}_1\times\mathbf{p}_1}{r_1^3}+\frac{\mathbf{r}_2\times\mathbf{p}_2}{r_2^3}\right)-\frac{3(\mathbf{r}_1-\mathbf{r}_2)\times(\mathbf{p}_1-\mathbf{p}_2)}{r_{12}^3}\right], \tag{7a}$$

and the spin-spin term

$$H(\text{s.s.})=\left(\frac{e\hbar}{2mc}\right)^2\frac{1}{r_{12}^3}\left(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2-\frac{3\boldsymbol{\sigma}_1\cdot\mathbf{r}_{12}\boldsymbol{\sigma}_2\cdot\mathbf{r}_{12}}{r_{12}^2}\right). \tag{7b}$$

[We have dropped here a term in H(s.o.) which has only off-diagonal matrix elements.] The first-order fine-structure levels are customarily given in terms of two constants C (from s.o.) and D (from s.s.):

$$\begin{aligned}\langle H_2\rangle_{J=0}&=\text{const}-2C-(10/3)D,\\ \langle H_2\rangle_{J=1}&=\text{const}-C+(5/3)D,\\ \langle H_2\rangle_{J=2}&=\text{const}+C-(1/3)D.\end{aligned} \tag{8}$$

[These values will subsequently be given in units of $(e\hbar/2mc)^2a_0^{-3}=(\alpha^2/4)(e^2/a_0)=\frac{1}{2}\alpha^2$ Ry]. These results are also shown in Table I. The accuracy of our C and D values here is hardly better than that of Pekeris *et al.*[4] (about $1/10^4$) and far short of the desired $1/10^6$. At the time when these results were obtained we attempted to extrapolate the apparent rate of convergence of the output numbers, and thus estimated that something approaching 2000 terms of the series (5) would be needed to obtain the required accuracy in C and D. It seemed that, at best, this would be an extremely expensive venture; and we thus decided that this attack had failed.

Two possible paths were then considered. We might start with the wave functions already constructed, and, by use of auxiliary variational calculations of modest size, attempt to increase the accuracy of the C and D integrals to something approaching that which is obtained directly for the energy eigenvalue.[15] However, it appeared that, due to the rather singular nature of the operators H_2, the auxiliary functions needed for this method would be of a complicated analytical nature; and so we did not make a serious attempt in this direction.

Alternatively, we could seek a better set of basis functions for the calculation of ψ. The most likely cause for the slow convergence noted above appeared to be the weak logarithmic singularity in the two-electron wave function studied by Fock.[16,17] It appeared, however, that putting into ψ the explicit logarithm term given by Fock would be a very messy job. Furthermore, it was felt that the reward would be rather slight if only the first term were accounted for; and the higher terms of Fock's expansion appear to be complicated beyond our ability to manage. We then sought, in a much more *ad hoc* manner to introduce just any covenient terms which would add flexibility to ψ especially in the region ($r_1\to 0$ and $r_2\to 0$). We thus chose to try the addition of the simple factor $(r_1+r_2)^{1/2}$ to the series (5).[18] The efficacy of this extension of the basis functions was first tested on the relatively simple calculation of the

[13] The precise values of κ and σ used were about 10^{-8} smaller than the figures given here, due to the decimal-to-binary conversion operation of the machine.

[14] For a discussion of relative convergence rates, see the article by C. Schwartz in *Methods in Computational Physics* (Academic Press Inc., New York, 1963), Vol. 2, p. 241.

[15] The method referred to is that described by C. Schwartz, Ann. Phys. (N.Y.) **6**, 170 (1959).

[16] V. A. Fock, Izv. Akad. Nauk SSSR Ser. Fiz. **18**, 161 (1954).

[17] See also the discussion and analysis in Ref. 14, p. 256.

[18] Half powers were earlier introduced into the Hylleraas expansion by H. M. Schwartz, Phys. Rev. **120**, 483 (1960).

TABLE II. Results of variational calculations of $2\,^3P$ with the extended basis (9).

Number ω	of terms	$-E(e^2/a_0)$	In units of $\frac{1}{2}\alpha^2$ Ry Cz	C_e	D
1	7	2.132580318470	0.127974449	−0.198448086	−0.053285232
2	19	2.133100285189	0.136827714	−0.203765520	−0.053888850
3	39	2.133155369966	0.138223897	−0.205167872	−0.053987664
4	69	2.133162942721	0.138649978	−0.205707289	−0.054026808
5	111	2.133164012406	0.138640520	−0.205858266	−0.054041649
6	167	2.133164164417	0.138638567	−0.205898542	−0.054046487
7	239	2.133164186301	0.138635834	−0.205908639	−0.054047869
8	329	2.133164189955	0.138636386	−0.205911292	−0.054048248
9	439	2.133164190626	0.138636755	−0.205911984	−0.054048351
Extrap-olation		2.13316419080 ±5	$C=Cz+C_e\rightarrow$	−0.067275287 ±19	−0.054048390 ±13

helium ground state (1^1S), and those very successful results have already been presented.[19]

The calculations reported in this section are not a total loss, since we can expect to use these medium-accuracy wave functions for the later calculation of the small correction terms schematically represented in Eq. (2).

IV. THE WAVE FUNCTION: SECOND ATTEMPT

The expanded basis now looks like the expression (5) with the replacement

$$C_{lmn}\rightarrow C_{lmn}+D_{lmn}(r_1+r_2)^{1/2}; \tag{9}$$

and the single term D_{000} is omitted. The scale parameter σ was left fixed as before at 0.29,[13] but after a short search we changed the value of κ to 4.0 (exactly). The matrix elements of H_0 and H_2 are more complicated with the half-power terms, but once we had learned to evaluate the integrals C and D even with the old basis, the new techniques were not very difficult (see Appendix). The results converged much more rapidly, than before, but it was found, as might be expected that as one approached more closely to the exact fitting of the function numerical accuracy became more and more critical at an alarming rate. We found it necessary to construct special programs to do all our arithmetic to an accuracy of 30, and then 52 decimals; the resulting cost in machine time was very great. The total labor took some 30 h on an IBM-7090.

In Table II are presented our results with the half-power functions up to the ninth order. We show here separately the contributions from the two parts of (7a):

$$C=C_Z+C_e. \tag{10}$$

V. RESULTS

From the numbers in Table II we now must give our best values for the fine-structure intervals along with an estimate of their probable uncertainty. This we do by attempting to extrapolate the results shown to $\omega\rightarrow\infty$; the smoothness of the sequence of computed numbers

[19] C. Schwartz, Phys. Rev. **128**, 1146 (1962).

TABLE III. Extrapolation of the calculated interval ν_{12} (in units of $\frac{1}{2}\alpha^2$ Ry).

Number of terms	Calculated $2D-2C$	Successive differences	Ratios of successive differences
7	0.0343768105		
		−0.008279	
19	0.0260979112		0.022
		−0.00018529	
39	0.0259126200		−0.800
		0.0001483866	
69	0.0260610066		1.962
		0.0002911883	
111	0.0263521949		0.257
		0.0000747812	
167	0.0264269761		0.306
		0.0000228957	
239	0.0264498718		0.151
		0.0000034446	
329	0.0264533164		0.127
		0.0000004379	
439	0.0264537543		
	Projected increment = 0.000000064 ±21		
	Extrapolation = 0.026453818 ±21		

will be used to indicate the reliability we may place on this extrapolation.

Table III shows the details of such an extrapolation for the small interval

$$\nu_{12}=\tfrac{1}{2}\alpha^2\ \text{Ry}\cdot(2D-2C). \tag{11}$$

(It does make a helpful difference on the final error estimates that we extrapolate the combined integrals, rather than combine the extrapolated integrals.) The table shows values, differences, and ratios of successive differences; attempts to extrapolate must be based on the behavior of these ratios, especially at the bottom of the sequence. It is clear from Table III that these ratios cannot be very well described in any simple analytical terms; but it is to our advantage that they are very small.[20] After some playing with these numbers, attempting various analytical and graphical fits, it was decided that a reasonable procedure would be simply to extrapolate from the last step as if we had a geometric series, and assign as our uncertainty $\frac{1}{3}$ of the added increment. That is, if Δ' is the last difference recorded and R' is the last ratio, our final answer is gotten by adding to the last computed value the quantity

$$\Delta''=\Delta'\left(\frac{R'}{1-R'}\right)(1\pm\tfrac{1}{3}). \tag{12}$$

[20] By way of contrast one may note that the corresponding ratios in the work of Pekeris *et al.* (Ref. 4) are much smoother, but also quite a bit larger. One may also notice that while their value for ν_{01} calculated with 220 terms is 0.011% larger than our final result, their extrapolated value is 0.009% too small. Their extrapolation here was thus in error by about 45%. (For ν_{12} their extrapolation erred by only about 16%) We point these numbers out intending not to chasten others, but to emphasize the need for a critical attitude toward any attempts at numerical extrapolation, and in particular our own.

We thus arrive at the following results for the intervals of the 2^3P state of helium, bare of any higher order corrections whatsoever.

$$\nu_{01}=\tfrac{1}{2}\alpha^2\ \mathrm{Ry}[0.33751721(7)] \qquad R'=0.210 \tag{13a}$$

$$\nu_{12}=\tfrac{1}{2}\alpha^2\ \mathrm{Ry}[0.02645382(2)] \qquad R'=0.127 \tag{13b}$$

$$\nu_{02}=\tfrac{1}{2}\alpha^2\ \mathrm{Ry}[0.36397101(9)] \qquad R'=0.172. \tag{13c}$$

The number in the inner parenthesis gives the uncertainty in the last figure quoted, arrived at according to (12). If our extrapolations are accepted, then we have achieved the desired goal of $1/10^6$ accuracy. If it becomes necessary, we could, without too much expense, carry the numerical work one order further and thus increase the certainty of these numbers by about a factor of 5; but at present we feel this part of the program can rest, and wait for the experimental work and the analysis of higher-order corrections to catch up.

VI. CONCLUSION

The first major goal of the plan outlined in Sec. II has now apparently been achieved, and this encourages us to proceed with the other parts of the over-all program for the redetermination of α. There are two rather apparent cautions which we wish to make. First, regarding the extrapolation and error estimation, we have tried to be both clever and objective in our analysis, but each reader will have to find his own measure of skepticism regarding the precise accuracy of our results. Secondly, while we have endeavored to check and recheck all our algebra and computer programs, the considerable complexity of the work reported here must leave open the possibility of some undiscovered error of importance. To this point we report the following experience.

Six months after we had completed the last of our computations, there was discovered an error in our computer program for multiple precision division. It turned out that the last ten decimals were sometimes treated incorrectly; but some conditions under which this error would not arise could also be precisely identified. A detailed scrutiny of all our programs then revealed the astonishing fact that this mistake did not have any effect at all on any of the matrix element computations. The error could work only in the process of solving for the eigenvectors of the matrices; but our method was, at the end, an iterative process, so that we suffered only a slower convergence (paid more dollars) as a result of the mistake, and all our final answers were unaffected.

We hope that someone, working independently of us, will repeat and check all this work; this is the best insurance we can imagine for the reliability of these important results.

One can ask for a comparison of the present best experimental and theoretical values for the fine-structure intervals. Adding to our results (13) the simple correction for the anomalous magnetic moment of the electron[2] (the only correction of relative order α) and using the value of the Rydberg for helium,

$$\tfrac{1}{2}\alpha^2\ \mathrm{Ry}=2.921374\ \mathrm{cm}^{-1}, \tag{14}$$

we get

$$\begin{aligned}\nu_{01}&=0.987837\ \mathrm{cm}^{-1},\\ \nu_{12}&=0.0765302\ \mathrm{cm}^{-1},\\ \nu_{02}&=1.064367\ \mathrm{cm}^{-1}.\end{aligned} \tag{15}$$

These may be compared with the experimental values

$$\begin{aligned}\nu_{01}&=0.98791(4)\ \mathrm{cm}^{-1},\\ \nu_{12}&=0.076423(3)\ \mathrm{cm}^{-1},\\ \nu_{02}&=1.06434(3)\ \mathrm{cm}^{-1}.\end{aligned}$$

[*Note added in proof.* The numbers quoted here have been corrected, relative to those published (Ref. 7), according to private communication from Dr. R. H. Sands. See also F. M. J. Pichanick, R. D. Swift, and V. W. Hughes, Bull. Am. Phys. Soc. **9**, 90 (1964).] The differences are about 0.0001 cm^{-1}, which is roughly what one would expect due to the neglected terms of relative order α^2. Araki[2] has considered one such higher-order term, the mixing with the nearby 2^1P state; Perkeris *et al.*[4] give for this effect a resulting downward shift of 0.000158 cm^{-1} to the $J=1$ level. The addition of this correction appears to do little more than change the signs of the discrepancies for the 0–1 and 1–2 intervals; and this indicated that all the higher order terms of Eq. (2) will have to be treated together.

ACKNOWLEDGMENTS

The author was assisted in some of the early programming by A. Miller. He also wishes to thank many staff members of the University of California Computer Center for the advice and service they have generously given him.

APPENDIX: NUMERICAL METHODS

The many matrix elements of the operators H_0 and H_2 in the basis used were evaluated in terms of pretabulated integrals of the following general type:

$$A_\Lambda(L,M,N)=\left(\frac{1+\sigma}{2\sigma}\right)^{L+M+N+(\Lambda/2)}\int\frac{dv_1}{4\pi}\int\frac{dv_2}{4\pi}e^{-r_1}e^{-r_2/\sigma}\times r_1{}^{M-2}r_2{}^{N-2}r_{12}{}^{L-2}f_\Lambda(r_1,r_2), \tag{A1}$$

$B_\Lambda(L,M,N)=$ same as (A1) but with a factor $\cos\theta_{12}$ under the integral. (A2)

L, M, N are integers greater than zero; $\sigma(0<\sigma<1)$ is a general parameter which is put $=0.29$ for the direct integrals and $=1$ for the exchange integrals; and f_Λ is a homogeneous function of degree $\frac{1}{2}\Lambda$ in the two lengths r_1 and r_2.

By use of the following identity referring to integration over the angle θ_{12} (signified by the angular brackets)

$$\langle \sin^2\theta_{12} r_{12}{}^L\rangle = \frac{-2}{(L+2)}\left\langle \cos\theta_{12}\frac{r_{12}{}^{L+2}}{r_1 r_2}\right\rangle, \qquad \text{(A3)}$$

one can easily establish the following recursion formulas:

$$\begin{aligned} A(L,M,N) &= A(L-2,M+2,N)+A(L-2,M,N+2) \\ &\quad -2B(L-2,M+1,N+1), \\ B(L,M,N) &= [(L-2)/(L+2)][B(L-2,M+2,N) \\ &\quad +B(L-2,M,N+2) \\ &\quad -2A(L-2,M+1,N+1)]. \end{aligned} \qquad \text{(A4)}$$

Our general method for building tables of integrals is by the use of such recursion formulas; but one must be careful that the particular recursive scheme chosen does not build by taking differences of nearly equal numbers (most obvious recursion schemes do suffer from this serious drawback.) For the above integrals—for large values of L, M, N in particular—the main contribution comes from large values of r_{12}; thus $\cos\theta_{12}$ is mostly negative, the integrals A and B have opposite signs, and the recursion schemes (A4) are safe.

In order to start the recursion sequence we need to evaluate A and B for $L=1$ and 2. The angular integrals are simple and we can express the results in terms of two families of two-dimensional integrals. (The subscript Λ is understood in what follows.)

$$\begin{aligned} &B(2,M,N)=0, \\ &A(2,M,N)=F(M+1,N+1)+G(N+1,M+1), \\ &B(1,M,N)=\tfrac{1}{3}F(M+2,N-1)+\tfrac{1}{3}G(N+2,M-1), \\ &A(1,M,N)=F(M+1,N)+G(N+1,M), \end{aligned} \qquad \text{(A5)}$$

where

$$\begin{aligned} F(M,N) &= [(1+\sigma)/2\sigma]^{M+N+(\Lambda/2)}\int_0^\infty dr e^{-r} r^{M-1} \\ &\quad \times\int_r^\infty ds e^{-s/\sigma} s^{N-1} f_\Lambda(r,s) \\ &= \left(\frac{1+\sigma}{2\sigma}\right)^{M+N+(\Lambda/2)}\Gamma\left(M+N+\frac{\Lambda}{2}\right) \\ &\quad \times\int_0^1 dx \frac{x^{M-1} f_\Lambda(x,1)}{[(1/\sigma)+x]^{M+N+(\Lambda/2)}}, \end{aligned} \qquad \text{(A6)}$$

and

$$\begin{aligned} G(M,N) &= \left(\frac{1+\sigma}{2\sigma}\right)^{M+N+(\Lambda/2)}\int_0^\infty dr e^{-r/\sigma} r^{M-1} \\ &\quad \times\int_r^\infty ds e^{-s} s^{N-1} f_\Lambda(s,r) \\ &= \left(\frac{1+\sigma}{2\sigma}\right)^{M+N+(\Lambda/2)}\Gamma\left(M+N+\frac{\Lambda}{2}\right) \\ &\quad \times\int_0^1 dx \frac{x^{M-1} f_\Lambda(1,x)}{[1+(x/\sigma)]^{M+N+(\Lambda/2)}}. \end{aligned}$$

We shall again tabulate these two dimensional arrays, F and G, by recursion; but all the obvious formulas we have looked at (proceeding from small to large values of M and N) are badly behaved regarding loss of accuracy from subtractions. We thus use the backward iteration formulas

$$\begin{aligned} F(M,N) &= \frac{2/(1+\sigma)}{(M+N+\frac{1}{2}\Lambda)}[\sigma F(M+1,N)+F(M,N+1)] \\ G(M,N) &= \frac{2/(1+\sigma)}{(M+N+\frac{1}{2}\Lambda)}[G(M+1,N)+\sigma G(M,N+1)]. \end{aligned} \qquad \text{(A7)}$$

These recursions start from the values of F and G along the line $M+N=$constant (about 30 in our work). We have found the following devious, but numerically safe, procedure for these evaluations. After Eq. (A6) make the change of variables

$$x=\frac{1-u}{1+u}, \qquad \text{(A8)}$$

implying

$$\begin{aligned} f_\Lambda(x,1) &= (1+u)^{-(\Lambda/2)} f_\Lambda(1-u,1+u), \\ f_\Lambda(1,x) &= (1+u)^{-(\Lambda/2)} f_\Lambda(1+u,1-u); \end{aligned} \qquad \text{(A8)}$$

then expand the resulting denominator of F in an infinite power series in

$$[\tfrac{1}{2}(1-\sigma)(1-u)],$$

and the denominator of G in terms of $\{[(1-\sigma)/(1+\sigma)]\times u\}$. Also make the binomial expansion of the resulting numerator factor $(1+u)^{N-1}$ according to the separation

$$(1+u)=2u+(1-u).$$

The resulting expressions are

$$\begin{Bmatrix} F(M,N) \\ G(M,N) \end{Bmatrix} = \sum_{J=1}^{N} \frac{(N-1)!\,2^{J-1}(M+N-J-1)!}{(N-J)!\,2^{M+N+(\Lambda/2)-1}} \times \begin{Bmatrix} \tilde{F}_J(M+N) \\ \tilde{G}_J(M+N) \end{Bmatrix}, \qquad \text{(A9)}$$

where

$$\tilde{F}_J(K)=[\tfrac{1}{2}(1+\sigma)]^{K+(\Lambda/2)}\sum_{k=0}^{\infty}\left(\frac{1-\sigma}{2}\right)^k \times\frac{(K+\frac{1}{2}\Lambda-1+k)!}{k!(J-1)!(K-1-J)!}\tilde{F}_J'(k,K), \tag{A10}$$

$$\tilde{G}_J(K)=\sum_{k=0}^{\infty}\left(\frac{1-\sigma}{1+\sigma}\right)^k\frac{(K+\frac{1}{2}\Lambda-1+k)!}{k!(J-1)!(K-1-J)!}\tilde{G}_J'(k,K),$$

and

$$\tilde{F}_J'(k,K)=\int_0^1 du f_\Lambda(1-u,1+u)u^{J-1}(1-u)^{K-1-J+k}, \tag{A11}$$

$$\tilde{G}_J'(k,K)=\int_0^1 du f_\Lambda(1+u,1-u)u^{J-1+k}(1-u)^{K-1-J}.$$

For most of the integrals of interest the function $f_\Lambda(r_1,r_2)$ is $(r_1+r_2)^{\Lambda/2}$ ($\Lambda=0$, ±1, or 2); and so in the final expressions (A11) the function f_Λ is merely a constant and these integrals are trivial. However in treating H_2 we also run into the spectral case of Eq. (A3) for $L=-2$, and this leaves us with

$$f_\Lambda(r_1,r_2)=\ln\left|\frac{r_1-r_2}{r_1+r_2}\right|\cdot(r_1+r_2)^{\Lambda/2} \tag{A12}$$

in Eq. (A6); but this reduces to $2^{\Lambda/2}\ln u$ in Eq. (A11), and these final integrals are again easy.

The infinite sums in Eq. (A10) are rapidly converging for our values of σ; and the entire procedure worked very well.

Matrices of dimension 440 in sextuple precision require over half a million words of machine storage space (an order of magnitude more than what is available in the core) and so we made extensive use of magnetic tapes. The most time-consuming operation was the matrix inversion carried out at the start of the eigenvector calculation. For the largest, dimension 438, this took 10 h of 7090 time—equivalent to about one billion simple 8-figure multiplications on this machine.

We will now describe the method used for finding the eigenvector. Given two symmetric matrices H and N ($N>0$) of order n, we separate the problem

$$\sum_{j=1}^{n}(H_{ij}-\lambda N_{ij})x_j=0,\quad i=1,2,\cdots,n \tag{A13}$$

as follows: Set $x_1=1$, and solve for x_i from

$$\sum_{j=2}^{n}(H_{ij}-\lambda N_{ij})x_j=-(H_{i1}-\lambda N_{i1}),\ i=2,3,\cdots,n \tag{A14}$$

once λ is known; but we can get a stationary estimation for λ, once an approximate vector x is known, from the Rayleigh quotient,

$$\lambda=(x,Hx)/(x,Nx). \tag{A15}$$

In this work we always have a very good guess for λ to start with, and the alternating iteration (A14), (A15) is very rapidly convergent. Most of the time goes into solving (A14); the process of solving n simultaneous linear equations

$$Ax=b \tag{A16}$$

by the direct method of elimination takes $\frac{1}{6}n^3$ operations (for a symmetric matrix A). On the other hand, iterative attacks on the solution of (A16) require only of the order of n^2 operations per cycle, but many iteration schemes do not converge at all well. The general criterion is that one have some "good" approximation to the inverse of the matrix A; then

$$x^{(l+1)}=x^{(l)}+\tilde{A}\cdot(b-Ax^{(l)}), \tag{A17}$$

where $\tilde{A}$ approximates A^{-1}, can be expected to converge "rapidly." For our problem the matrix to be inverted (A14) changes only very slightly, as our value of λ is improved, so that it is necessary to calculate *only once* an inverse matrix (by the direct method[21]), and then we use this as the kernel of the iterations (A17), until both eigenvalue and eigenvector are converged.

This process works well, but as one goes to larger and larger matrices, loss of numerical accuracy becomes a serious problem at a rate much beyond any that could be attributed to statistical phenomena. Obviously, the full matrix $(H-\lambda N)$ in Eq. (A13) is singular—for the correct λ—but one would think that the remnant matrix in Eq. (A14) is far from singular since the most important element of the basis has been removed. In actuality, however, the removal of a single trial function does not prevent the convergence of the variational sequence, it may merely slow it. Thus, as we go to larger and larger bases, the matrix in Eq. (A14) does come closer and closer to being singular, and at a painfully fast rate. We handled this problem by the brute force method of using higher precision arithmetic, but the cost was quite great. (It may be that the problem would not have been quite as bad as it did appear to us, had we caught earlier the error mentioned in Sec. VI.)

[21] After completion of this work the following, possibly more efficient, method occurred to us. The new terms one is adding to the basis at each step in the sequence of variational calculations must be numerically of decreasing importance (if the entire process is converging at all well). Then it may be sufficient to get $(\tilde{A})$ by carrying out the direct calculation on only a small submatrix representing the dominant terms, and simply using the diagonal elements for the higher parts of the matrix.