derived from Hamilton's principle if Eqs. (49) and (50) are first used to rewrite Eq. (51) in the form

$$S = \int [p_i \dot{q}_i - H(q,p,t)] dt. \tag{65}$$

With *fixed* end points t_1 and t_2, independent variations δp_i and δq_i then induce the variation

$$\delta S = \int_{t_1}^{t_2} \left(\dot{p}_i \delta q_i + \dot{q}_i \delta p_i + \frac{\partial H}{\partial q_i} \delta q_i + \frac{\partial H}{\partial p_i} \delta p_i \right) dt$$

$$= p_i \delta q_i \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(-\dot{p}_i \delta q_i + \dot{q}_i \delta p_i \right.$$

$$\left. + \frac{\partial H}{\partial q_i} \delta q_i + \frac{\partial H}{\partial p_i} \delta p_i \right) dt. \tag{66}$$

The integrated term vanishes by assumption. Equating to zero the coefficients of δp_i and δq_i then yields the set of equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{67}$$

which are Hamilton's equations of motion. Thus, both Lagrangian and Hamiltonian mechanics are implied by Hamilton's principle, provided that the variations δp_i and δq_i can be regarded as independent. The latter assumption is important because it underlies the entire approach of Hamiltonian mechanics. It is justified by the fact that the Hamiltonian form can be obtained by a direct transformation of variables from the Lagrangian form.

2.3.3   Canonical   Transformations

Transformations of the type (33) are called *point transformations* because they involve only the coordinates q_i. An advantage of Hamiltonian mechanics is that, with the q's and p's regarded as independent variables, more general types of transformation can be constructed that map the set $\{q,p\}$ to a new set $\{Q,P\}$ defined by connection equations of the form

$$Q_i = Q_i(q,p,t), \quad P_i = P_i(q,p,t). \tag{68}$$

Such a *contact transformation* is said to be *canonical* if the form of Hamilton's equations is left invariant; i.e., there exists a trans-

formed Hamiltonian function $K(Q,P,t)$ such that

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i} \tag{69}$$

are the correct equations of motion for the transformed variables. This will be true if both sets of coordinates and momenta satisfy their respective variational principles

$$\delta \int [p_i \dot{q}_i - H(q,p,t)] dt = 0, \tag{70}$$

$$\delta \int [P_i \dot{Q}_i - K(Q,P,t)] dt = 0. \tag{71}$$

The integrands need not be equal, but they can differ by at most the total time derivative of a function F, called the *generating function* for the transformation. We can thus write

$$p_i \dot{q}_i - H(q,p,t) = \lambda \left[P_i \dot{Q}_i - K(Q,P,t) + \frac{dF}{dt} \right]. \tag{72}$$

Values of the constant $\lambda \neq 1$ just correspond to a scale transformation, and so without loss of generality, we can take $\lambda = 1$. With the choice $F = F_1(q,Q,t)$, the set $\{q,Q\}$ is regarded as the $2N$ separately independent variables, and the quantity dF/dt in Eq. (72) can be replaced by

$$\frac{dF_1}{dt} = \frac{\partial F_1}{\partial q_i} \dot{q}_i + \frac{\partial F_1}{\partial Q_i} \dot{Q}_i + \frac{\partial F_1}{\partial t}. \tag{73}$$

Equating to zero the coefficients of $\dot{q}_i$ and $\dot{Q}_i$ then yields

$$p_i = \frac{\partial F_1}{\partial q_i}, \quad P_i = -\frac{\partial F_1}{\partial Q_i}, \tag{74}$$

$$K = H + \frac{\partial F_1}{\partial t}. \tag{75}$$

For example, the choice $F_1 = q_i Q_i$ interchanges the roles of the q's and p's (except for a sign change), with the result

$$p_i = Q_i, \quad P_i = -q_i, \quad K = H. \tag{76}$$

This demonstrates explicitly that the q's and

p's stand on an equal footing as independent variables in Hamiltonian mechanics—their roles can be interchanged by a canonical transformation.

Other choices for the $2N$ independent variables can be achieved by application of one or more Legendre transformations (as is done in thermodynamics to change independent variables). Of particular relevance to the Hamilton–Jacobi equation is the choice $[q,P]$. One need merely replace $F_1(q,Q)$ by $F_2(q,P) = F_1 + Q_iP_i$. Then Eq. (72) becomes

$$p_i\dot{q}_i - H(q,p,t) = -Q_i\dot{P}_i - K(Q,P,t) + \frac{dF_2}{dt}, \tag{77}$$

with

$$\frac{dF_2}{dt} = \frac{\partial F_2}{\partial q_i}\dot{q}_i + \frac{\partial F_2}{\partial P_i}\dot{P}_i + \frac{\partial F_2}{\partial t}, \tag{78}$$

and so

$$P_i = \frac{\partial F_2}{\partial q_i}, \quad Q_i = \frac{\partial F_2}{\partial P_i}, \tag{79}$$

$$K = H + \frac{\partial F_2}{\partial t}. \tag{80}$$

For example, the choice $F_2 = q_iP_i$ generates the identity transformation with $Q_i = q_i$ and $P_i = p_i$.

2.3.4 Interpretation of the Hamilton–Jacobi Equation   A comparison of Eq. (80) with Eq. (54) shows that the Hamilton–Jacobi equation can be regarded as a canonical transformation with $F_2 = S(q,P,t)$ such that the transformed Hamiltonian is $K = 0$. The transformed equations of motion are then

$$\frac{\partial K}{\partial P_i} = \dot{Q}_i = 0, \quad \frac{\partial K}{\partial Q_i} = -\dot{P}_i = 0. \tag{81}$$

Thus the transformed momenta $P_i = \alpha_i$ and the transformed coordinates $Q_i = \beta_i$ are all constants of the motion. The nature of the solution is now clear. Writing $S = S(q,\alpha,t)$, then the set of equations

$$p_i = \frac{\partial S(q,\alpha,t)}{\partial q_i}, \quad \beta_i = \frac{\partial S(q,\alpha,t)}{\partial \alpha_i} \tag{82}$$

evaluated at $t = t_1$ provides a set of $2N$ equations to determine $2N$ unknowns α_i and β_i in terms of the initial $(q_i)_1$ and $(p_i)_1$. For conservative systems, the remaining $(N + 1)$th constant of integration is the energy E of the system. For this case, the time variable is separable, and S can be written in the form

$$S(q,\alpha,t) = W(q,\alpha) - Et, \tag{83}$$

where $W(q,\alpha)$ is called *Hamilton's characteristic function*. The crucial point is that *any* complete integral to the Hamilton–Jacobi equation generates a possible dynamical motion of the system since it is the generating function for a canonical transformation. This result is called Jacobi's theorem.

As a consequence of Jacobi's theorem, any complete integral S contains within it all possible trajectories of the system as a function of the initial conditions, rather than one particular trajectory. In fact, surfaces of constant S move through configuration space like wave fronts of constant phase such that the particle trajectories follow the orthogonal set of curves. To see this, consider the example of a particle moving in a potential V, expressed in Cartesian coordinates. Equations (53) for $[q_1,q_2,q_3] = [x,y,z]$ can then be written as the single equation

$$\nabla S = \nabla W = \mathbf{p}, \tag{84}$$

which shows that the momentum $\mathbf{p} = m\mathbf{v}$ is everywhere perpendicular to surfaces of constant W, and the particle velocity is $v = |\nabla W|/m$. The Hamilton–Jacobi equation for W in this case is

$$(1/2m)(\nabla W)^2 + V = E, \tag{85}$$

so that

$$|\nabla W| = \sqrt{2m(E - V)}. \tag{86}$$

For the case of a freely falling particle, surfaces of constant W are just horizontal planes with the particle trajectories in the perpendicular direction. As time goes on, surfaces of constant S sweep through surfaces of constant W with a phase velocity given by

$$u = \frac{ds}{dt} = \frac{E}{|\nabla W|}, \tag{87}$$

where ds is a displacement in the direction normal to a surface of constant W. The above follows from the facts that the stationary phase condition $dS = 0$ corresponds to $dW = E\,dt$, and $dW = |\nabla W|ds$ is the change in W due to a displacement in the direction normal to the surface. Thus the phase velocity decreases as the particle velocity increases, just as is the case for the wave and particle pictures of light.

The above considerations in fact provide a wave picture of classical dynamics in the "geometrical optics" limit where the wavelength is infinitesimally small compared with the dimensions of the apparatus. The wave nature can then be ignored, and the trajectories determined by the principle of least action (or Fermat's principle in the case of geometrical optics).

2.4 Relativistic Generalization

For the case of a single particle acted on by forces derivable from a potential V, Hamilton's principle can be simply modified to incorporate the effects of special relativity. One need simply define

$$L = -mc^2\gamma - V, \tag{88}$$

where $\gamma = \sqrt{1 - v^2/c^2}$, and v is the velocity $|\dot{\mathbf{r}}|$ in a particularly chosen Lorentz frame. With this choice of L, Hamilton's principle and the Euler–Lagrange equations give the correct equations of motion

$$\frac{d}{dt}\left(\frac{m\dot{x}_i}{\gamma}\right) = -\frac{\partial V}{\partial x_i}. \tag{89}$$

The canonical momenta p_i and Hamiltonian H are given by

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = \frac{m\dot{x}_i}{\gamma}, \tag{90}$$

$$H = p_i\dot{x}_i - L = T + V = E, \tag{91}$$

where $T = mc^2/\gamma$ is a relativistic generalization of the kinetic energy, including the rest-mass energy mc^2. After substituting for $\dot{x}_i$, H assumes the form

$$H = \sqrt{c^2p^2 + m^2c^4} + V. \tag{92}$$

2.4.1 Inclusion of Electromagnetic Fields
In general, an electromagnetic field is derivable from a scalar potential $\phi(\mathbf{r},t)$ and a vector potential $\mathbf{A}(\mathbf{r},t)$ according to

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \tag{93}$$

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{94}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields. The equation of motion for a particle of charge q is then

$$\frac{d}{dt}\left(\frac{m\dot{x}_i}{\gamma}\right) = qE_i + \frac{q}{c}(\mathbf{v} \times \mathbf{B})_i, \tag{95}$$

which now contains a velocity-dependent force term. This equation follows from the Lagrangian

$$L = -mc^2\gamma - q\phi + \frac{q}{c}\mathbf{A}\cdot\mathbf{v}, \tag{96}$$

or its nonrelativistic counterpart with the term $-mc^2\gamma$ replaced by T. The canonical momenta are then

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = \frac{m\dot{x}_i}{\gamma} + \frac{q}{c}A_i. \tag{97}$$

A direct calculation shows that the Hamiltonian becomes [cf. Eq. (92)]

$$H = \sqrt{c^2[\mathbf{p} - (q/c)\mathbf{A}]^2 + m^2c^4} + V. \tag{98}$$

The same substitution $\mathbf{p} \to \mathbf{p} - (q/c)\mathbf{A}$ applies also in the nonrelativistic case. This simple prescription, together with $V = q\phi$, allows electromagnetic fields to be easily incorporated into the Lagrangian and Hamiltonian formulations of mechanics.

3. APPLICATIONS TO QUANTUM MECHANICS

3.1 Variational Derivation of the Schrödinger Equation

The considerations of Sec. 2.3.4 suggest that the Hamilton–Jacobi equation of classical dynamics expresses the short-wavelength

limit of an underlying wave equation, with surfaces of constant S identified as surfaces of constant phase. In fact, Eq. (85) already bears a superficial resemblance to the time-independent Schrödinger equation, but it does not yet have the form of a wave equation. Following Schrödinger, a suitable wave equation can be obtained by first making the substitution

$$W = iC \ln \Psi \Rightarrow \Psi = e^{iW/C} \qquad (99)$$

into the Hamilton–Jacobi equation

$$(1/2m)(\nabla W^* \cdot \nabla W) + V = E, \qquad (100)$$

generalized for complex W, to obtain

$$(C^2/2m)(\nabla \Psi^* \cdot \nabla \Psi) + (V - E)\Psi^* \Psi = 0. \qquad (101)$$

The left-hand side can be integrated over all space, provided that $\int \Psi^* \Psi d^3 r$ remains finite. Application of the Euler–Lagrange equations to make the integral stationary with respect to arbitrary independent variations of Ψ and Ψ^* then yields the Schrödinger equation

$$-(C^2/2m)(\nabla^2 \Psi) + (V - E)\Psi = 0, \qquad (102)$$

together with a similar equation for Ψ^*. The derivation is a simple extension of the one used to obtain Laplace's equation (31) in Sec. 1.2. The solution to Eq. (102) then determines the wave function $\Psi(\mathbf{r})$ for the system, subject to the constraint that $\int \Psi^* \Psi d^3 r$ remain finite for bound systems; i.e., that Ψ is *normalizable*. Comparison with experiment shows that one should set $C = \hbar = h/2\pi$, where h is Planck's constant. Equation (102) can then be written in the form

$$H(q,p)\Psi = E\Psi, \qquad (103)$$

where $H(q,p)$ is the Hamiltonian with the quantum-mechanical replacement $\mathbf{p} \to (\hbar/i)\nabla$. The constraint on $\int \Psi^* \Psi d^3 r$ makes this an eigenvalue problem that determines the possible energies E of the system.

If there are n interacting particles, then the term $\nabla^2 \Psi$ is to be replaced by $\Sigma_i \nabla_i^2 \Psi$, and V includes all the interaction potentials. Also, the various integrals over $d\mathbf{r}$ are replaced by multiple integrals over $d^3 r_1 d^3 r_2 \cdots d^3 r_n$.

3.2 The Rayleigh–Schrödinger Variational Principle

Consider a bound system, or one that is contained in a finite box. Under these conditions, Eq. (101) can be integrated over all space and the term $\nabla \Psi^* \cdot \nabla \Psi$ integrated by parts to obtain

$$\int \Psi^*(H - E)\Psi d^3 r = 0. \qquad (104)$$

The integrated term does not contribute under the assumed conditions because $\Psi(\mathbf{r}) \to 0$ sufficiently rapidly as $|\mathbf{r}| \to \infty$. The variational derivation of Sec. 3.1 guarantees that this integral is stationary with respect to arbitrary variations $\delta\Psi$ if Ψ satisfies the Schrödinger equation. However, the same variational condition can now be reinterpreted as the problem of making the integral $\int \Psi^* H \Psi d^3 r$ stationary, subject to the constraint that

$$\int \Psi^* \Psi d^3 r = \text{const.}, \qquad (105)$$

with E playing the role of a Lagrange undetermined multiplier. In this guise, one can say that E obtained from the Rayleigh quotient

$$E = \frac{\displaystyle\int \Psi^* H \Psi d^3 r}{\displaystyle\int \Psi^* \Psi d^3 r} \qquad (106)$$

is stationary. In fact, as discussed in the following section, E is a minimum under many circumstances.

3.3 The Rayleigh–Ritz Variational Method

The Schrödinger equation is a partial-differential equation that can be solved exactly only for certain special cases such as the Coulomb potential or the harmonic-oscillator potential. For arbitrary potentials, or for problems containing more than two bodies, the quantum-mechanical problem is no easier to solve than the corresponding classical one. In these cases, the Rayleigh–Schrödinger variational principle provides one of the most powerful methods for obtaining approximate eigenvalues E and wave functions Ψ.

Suppose one guesses by some means an approximate trial wave function Ψ_{tr} that conforms with the constraint of normalizability and approximates one of the exact solutions to

$$H\Psi_i = E_i\Psi_i, \quad i = 1, 2, \ldots. \tag{107}$$

The index i labels the spectrum of exact solutions. In general, the eigenvalue spectrum will have both discrete and continuous parts. In the latter case, summations over i include integrations over the continuous part. The crucial point is that even though the Ψ_i are not known, they form a complete basis set of functions in terms of which the trial function Ψ_{tr} can be expanded. In analogy with Fourier series, one can therefore write

$$\Psi_{\text{tr}} = \sum_{i=1}^{\infty} c_i\Psi_i, \tag{108}$$

where the c_i are the expansion coefficients. Let the eigenvalue spectrum be ordered so that $E_1 < E_2 < E_3 < \cdots$, and assume that all the Ψ_i and Ψ_{tr} are normalized to unity; i.e., using Dirac bra-ket notation for integrals,

$$\langle \Psi_{\text{tr}}|H|\Psi_{\text{tr}}\rangle \equiv \int \Psi^* H \Psi d^3r = E_{\text{tr}},$$
$$\langle \Psi_i|\Psi_j\rangle = \delta_{i,j},$$
$$\langle \Psi_i|H|\Psi_j\rangle = E_i\delta_{i,j}. \tag{109}$$

Substituting Eq. (108) into (106), and using Eqs. (109), one then obtains

$$E_{\text{tr}} = |c_1|^2E_1 + |c_2|^2E_2 + |c_3|^2E_3 + \cdots \tag{110}$$

for the corresponding trial energy. Since, by assumption, $\langle \Psi_{\text{tr}}|\Psi_{\text{tr}}\rangle = 1$, it follows that

$$\sum_{i=1}^{\infty} |c_i|^2 = 1, \tag{111}$$

and so Eq. (110) can be rewritten in the form

$$\begin{aligned}
E_{\text{tr}} = E_1 &+ |c_2|^2(E_2 - E_1) \\
&+ |c_3|^2(E_3 - E_1) + \cdots \\
&\geq E_1.
\end{aligned} \tag{112}$$

Thus E_{tr} is an *upper bound* on the lowest eigenvalue E_1 for any normalizable Ψ_{tr}.

The basic idea of variational calculations then is to write Ψ_{tr} in some arbitrarily chosen mathematical form with variational parameters (subject to normalizability and boundary conditions at the origin and infinity), and then adjust the parameters to obtain the minimum value of E_{tr}. A lower E_{tr} is guaranteed to be closer to E_1. The power of the method stems both from this and the fact that, by the Rayleigh–Schrödinger variational principle, the error term linear in $\delta\Psi = \Psi_{\text{tr}} - \Psi_1$ vanishes.

3.3.1 Algebraic Solution for Linear Variational Parameters

Suppose that Ψ_{tr} depends in some arbitrarily chosen way on a set of N variational parameters $a_1, a_2, \ldots, a_N$. [For example, in a one-dimensional case, one might choose $\Psi(r) = r^{a_1}e^{-a_2r}$ with a_1 and a_2 regarded as nonlinear variational parameters.] Then the variational condition corresponds to the system of equations

$$\frac{\partial E_{\text{tr}}}{\partial a_p} = 0, \quad p = 1, \ldots, N. \tag{113}$$

In general, this is a set of transcendental algebraic equations that cannot be solved exactly.

However, the minimization problem for the case of *linear* variational coefficients can be solved exactly by matrix diagonalization. For example, let $\{\chi_p|p = 1, \ldots, N\}$ be a finite basis set of N arbitrarily chosen functions (subject to the boundary conditions and normalizability) that need have nothing to do with the exact Ψ_i, and write Ψ_{tr} in the form

$$\Psi_{\text{tr}} = \sum_{p=1}^{N} a_p\chi_p. \tag{114}$$

Now the variational parameters a_p enter linearly, and the set of variational conditions (113) becomes exactly equivalent to the N-dimensional generalized eigenvalue problem

$$\mathbf{Ha} = \lambda\mathbf{Oa}, \tag{115}$$

where $\mathbf{a}$ is a column vector of coefficients a_p, and $\mathbf{H}$ and $\mathbf{O}$ have matrix elements $H_{pq} = \langle \chi_p|H|\chi_q\rangle$ and $O_{pq} = \langle \chi_p|\chi_q\rangle$. There are N eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$, of which the lowest is an upper bound to E_1.

Equation (115) is equivalent to the original Schrödinger equation (103) only if the basis set $\{\chi_p\}$ is complete, and in general this requires taking the limit $N \to \infty$. The significance of Eq. (115) is that it provides a computationally useful means of obtaining approximate solutions, even if the complete basis set of functions $\{\chi_p\}$ is truncated at some finite number N, and the lowest eigenvalue provides an upper bound that systematically decreases toward the exact E_1 as N is increased. As discussed in the following section, the bounds apply not just to the ground state but also to the lower-lying excited states.

3.3.2 Extension to Excited States   By the Hylleraas–Undheim–MacDonald (HUM) theorem (see Hylleraas and Undheim, 1930; MacDonald, 1933), the remaining eigenvalues $\lambda_2, \lambda_3, \ldots$ are also upper bounds to the exact energies $E_2, E_3, \ldots$, provided that the spectrum is bounded from below. The HUM theorem is a consequence of the matrix eigenvalue interleaving theorem, which states that as the dimensions of **H** and **O** are progressively increased by adding an extra row and column, the N old eigenvalues λ_p fall between the $N + 1$ new ones. Consequently, as illustrated in Fig. 3, all eigenvalues numbered from the bottom up must move inexorably downward as N is increased. Since the
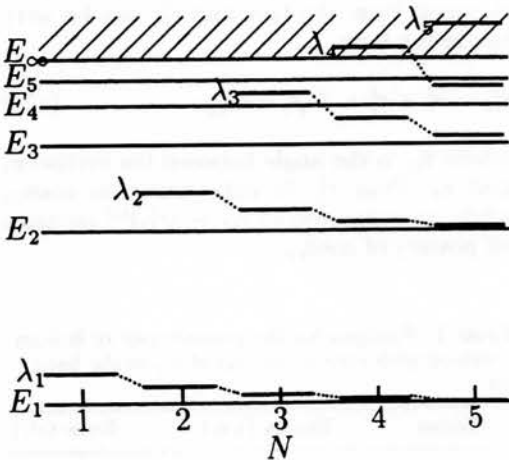
exact spectrum of bound states is obtained in the limit $N \to \infty$, no λ_p can cross the corresponding exact E_p on its way down. Thus $\lambda_p \geq E_p$ for every finite N.

If the exact Ψ_i can be formed from a linear combination of the χ_p included in the finite basis set, then the result of the variational calculation is the exact Ψ_i and E_i. Otherwise, one obtains the best variational approximation provided by the particular χ_p chosen. If the χ_p basis set becomes asymptotically complete as $p \to \infty$, then convergence to the correct answer is assured.

The HUM theorem no longer applies directly to the relativistic Dirac equation or similar problems because the spectrum is not bounded from below. However, finite basis-set methods can still be applied, provided that sufficient care is taken in their construction (see Drake and Goldman, 1988; Grant, 1996).

3.3.3 Variational Lower Bound   If the Rayleigh–Ritz method is applied to the integral

$$\int \Psi^*(H - E)(H - E_2^<)d^3r, \tag{116}$$

then the quantity

$$E^< = \frac{\langle \Psi_{tr}|H(H - E_2^<)|\Psi_{tr}\rangle}{\langle \Psi_{tr}|(H - E_2^<)|\Psi_{tr}\rangle} \tag{117}$$

is made stationary. If the quantity $E_2^<$ is chosen to be a lower bound on the energy E_2 of the first excited state, then by an argument similar to that leading to Eq. (112), $E^<$ is a lower bound on the ground-state energy E_1, called the Temple lower bound (Temple, 1928) and denoted by $E_1^<$. In fact, if Ψ_{tr} is expanded as in Eq. (108), then after some algebra, $E^<$ from Eq. (117) becomes

$$E_1^< = E_1 + \frac{\sum_{i=2}^{\infty} |c_i|^2(E_i - E_1)(E_i - E_2^<)}{E_1 - E_2^< + D}, \tag{118}$$

where

$$D = \sum_{i=2}^{\infty} |c_i|^2(E_i - E_1) = E_1^> - E_1, \tag{119}$$

FIG. 3. Diagram illustrating the HUM theorem. The λ_p, $p = 1, \ldots, N$, are the variational eigenvalues for an N-dimensional basis set, and the E_i are the exact eigenvalues of H. The highest λ_p lie in the continuous spectrum of H.

and $E_1^>$ is the variational upper bound on E_1. The numerator of the fraction in Eq. (118) is positive and the denominator is negative, provided that $E_1^< < E_2^<$, thus proving the bound. However, generally speaking, $E_1^<$ is much less accurate than $E_1^>$.

3.3.4 Illustrative Results for Helium

Application of the variational method to helium by Hylleraas (1928, 1929) played an important role in the early history of quantum mechanics because it provided the first test of the Schrödinger equation in a system more complicated than hydrogen. With its two electrons orbiting the nucleus, helium is the simplest atomic system that cannot be adequately described by the older Bohr–Sommerfeld quantum theory.

The Hamiltonian for helium (in the limit of infinite nuclear mass) is

$$H = -\sum_{i=1}^{2}\left(\frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{e^2}{r_i}\right) + \frac{e^2}{r_{12}}, \qquad (120)$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ is the electron–electron separation and Ze is the nuclear charge. The e^2/r_{12} term represents the Coulomb repulsion between the two electrons. Without this term, the Schrödinger equation would be separable, and the exact solution (including permutational symmetry) would be of the form

$$\Psi(\mathbf{r}_1,\mathbf{r}_2) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) \pm \psi_1(\mathbf{r}_2)\psi_2(\mathbf{r}_1), \qquad (121)$$

where $\psi_1(\mathbf{r})$ and $\psi_2(\mathbf{r})$ are exactly known hydrogenic wave functions, depending on the state in question. The Hartree–Fock approximation corresponds to the best variational representation that can be written in the form of a separable product with $\psi_1(\mathbf{r})$ and $\psi_2(\mathbf{r})$ regarded as arbitrary functions of r. However, even this is in error for the ground-state energy of $-2.903\,724$ a.u. [the atomic unit (a.u.) of energy is $e^2/a_0 = 27.211\,396$ eV, where a_0 is the Bohr radius] by 0.0247 a.u., or 0.673 eV. This difference, called the *correlation energy*, is much larger than typical chemical energies.

To obtain a better representation, Hylleraas suggested constructing a trial solution of the form

$$\Psi_{\text{tr}} = \sum_{i,j,k} a_{ijk} r_1^i r_2^j r_{12}^k \exp(-\alpha r_1 - \beta r_2), \qquad (122)$$

which is of the form of a hydrogenic product $\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)$, except that it contains explicit powers of r_{12} and is therefore not separable. The a_{ijk} are the linear variational parameters, and α and β are additional nonlinear parameters that can be separately varied to optimize the energy. Detailed formulas for the necessary matrix elements are given by Drake (1996). This basis set is provably complete in the limit of large $i_{\max}$, $j_{\max}$, and $k_{\max}$ (Klahn and Bingel, 1977). Typically, all combinations of powers are included in the basis set such that, for electrons with angular momenta l_1 and l_2,

$$i + j + k - l_1 - l_2 \leq \Omega, \qquad (123)$$

where Ω is an integer that is progressively increased until adequate convergence is obtained. Without further truncation, the number of terms obtained in this way is

$$N = \tfrac{1}{6}(\Omega + 1)(\Omega + 2)(\Omega + 3). \qquad (124)$$

The effect of including powers of r_{12} is dramatic and immediate. The Hartree–Fock approximation corresponds to the limit of large $i_{\max}$ and $j_{\max}$ with $k_{\max} = 0$. As shown in Table 1, an increase of $k_{\max}$ to 1 reduces the error in E to only 0.006 20 eV, thereby accounting for 99% of the correlation energy. The results in Table 1 also demonstrate that the odd powers of r_{12} are much more effective than the even powers. This can be understood from the fact that r_{12}^2 can be written in the form

$$r_{12}^2 = r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta_{12}, \qquad (125)$$

where θ_{12} is the angle between the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Thus r_{12}^2 is only linear in $\cos\theta_{12}$, while an expansion of $r_{12} = (r_{12}^2)^{1/2}$ contains all powers of $\cos\theta_{12}$.

Table 1. Energies for the ground state of helium obtained with various powers of r_{12} in the basis set.

r_{12} Terms	Energy (a.u.)	Error (eV)
No r_{12}	$-2.879\,029$	0.672
r_{12}^2	$-2.900\,503$	0.087 6
r_{12}^2, r_{12}^4	$-2.902\,752$	0.026 4
r_{12}	$-2.903\,496$	0.006 20
r_{12}, r_{12}^3	$-2.903\,700$	0.000 65
All r_{12}	$-2.903\,724$	0.000 00

Calculations of this type have been performed by many authors (see Drake, 1993a, for a review), and they have reached a high degree of sophistication. Problems typical of all variational calculations are a dramatic and progressive loss of accuracy for the more highly excited states, and numerical linear dependence in the basis set as it is enlarged. These problems can be avoided by doubling the basis set so that it contains a second set of terms with the same powers but different scale factors α and β. A complete optimization with respect to the α's and β's then leads to a natural partition of the basis set into two sectors with one representing the asymptotic form of the wave function and the other representing complex inner correlation effects. The results obtained by this method are essentially exact for all practical purposes. The nonrelativistic energies are known to better than one part in 10^{16} for the entire singly excited spectrum. A sample of results for the low-lying states is given in Table 2. The indicated convergence was obtained by progressively increasing Ω up to 17, corresponding to about 1700 terms in the doubled basis set. Table 3 shows an example of the convergence for the ground state. [For the case of S states, the basis-set sizes are smaller than indicated by Eq. (124) because terms with $i > j$ can be omitted by symmetry.] The ratios of successive differences in the last column provide a convenient method to monitor the convergence of the eigenvalue. They show that the differences themselves decrease in a fairly smooth and uniform fashion with increasing Ω. These high-precision results provide a benchmark for comparison with other less accurate

Table 3. Convergence study for the ground-state energy of helium (in atomic units). The numbers in the last column give the ratios of successive differences.

Ω	$N_{tot}(\Omega)$	$E(\Omega)$	$R(\Omega)^a$
4	44	$-2.903\ 724\ 131\ 001\ 531\ 810$	
5	67	$-2.903\ 724\ 351\ 566\ 477\ 006$	
6	98	$-2.903\ 724\ 373\ 891\ 109\ 909$	9.88
7	135	$-2.903\ 724\ 376\ 548\ 959\ 510$	8.40
8	182	$-2.903\ 724\ 376\ 960\ 412\ 587$	6.46
9	236	$-2.903\ 724\ 377\ 018\ 168\ 462$	7.12
10	302	$-2.903\ 724\ 377\ 030\ 786\ 217$	4.58
11	376	$-2.903\ 724\ 377\ 033\ 426\ 037$	4.78
12	464	$-2.903\ 724\ 377\ 033\ 966\ 492$	4.88
13	561	$-2.903\ 724\ 377\ 034\ 076\ 500$	4.91
14	674	$-2.903\ 724\ 377\ 034\ 107\ 875$	3.51
15	797	$-2.903\ 724\ 377\ 034\ 116\ 019$	3.85
16	938	$-2.903\ 724\ 377\ 034\ 118\ 518$	3.26
17	1090	$-2.903\ 724\ 377\ 034\ 119\ 239$	3.47
18	1262	$-2.903\ 724\ 377\ 034\ 119\ 479$	3.01
Extrapolation		$-2.903\ 724\ 377\ 034\ 119\ 597(15)$	

$^a R(\Omega) = [E(\Omega - 1) - E(\Omega - 2)]/[E(\Omega) - E(\Omega - 1)].$

methods of calculation such as Hartree–Fock and configuration interaction. Results for many other states are given by Drake (1993b, 1994).

A comparison of the results in Table 2 with experiment is meaningful only after corrections for finite nuclear mass, special relativity, and quantum-electrodynamic (QED) effects (such as electron self-energy and vacuum polarization) are taken into account. A detailed discussion of these corrections can be found in Drake (1993b, 1994). When they are included, the calculated transition frequencies agree to within the estimated accuracy of the QED shift. If the measurements are expressed in terms of ionization energies for the various states, then their accuracies range from ± 30 MHz ($\pm 5 \times 10^{-9}$ a.u.) for the ground state to ± 0.1 MHz ($\pm 1.5 \times 10^{-11}$ a.u.) for the higher-lying P and D states. Since the nonrelativistic energies in Table 2 are much more accurate than this, the comparison with experiment is primarily a test of higher-order contributions to the QED shift (two-electron Lamb shift), which is the dominant source of uncertainty in the calculations.

As one example, the calculated ionization energy of the $1s2s$ ^1S state is (Drake et al., 1993)

$$960\ 332\ 039.4 \pm 1 \text{ MHz}.$$

Of this total, -2808.5 ± 1 MHz comes from

Table 2. Nonrelativistic energies for several states of helium in the limit of infinite nuclear mass.

State	Energy (a.u.)
$1s^2$ ^1S	$-2.903\ 724\ 377\ 034\ 119\ 60(2)$
$1s2s$ ^1S	$-2.145\ 974\ 046\ 054\ 419(2)$
$1s2s$ ^3S	$-2.175\ 229\ 378\ 236\ 791\ 307(6)$
$1s2p$ ^1P	$-2.123\ 843\ 086\ 498\ 101\ 35(5)$
$1s2p$ ^3P	$-2.133\ 164\ 190\ 779\ 283\ 17(3)$
$1s3s$ ^1S	$-2.061\ 271\ 989\ 740\ 911(5)$
$1s3s$ ^3S	$-2.068\ 689\ 067\ 472\ 457\ 192(1)$
$1s3p$ ^1P	$-2.055\ 146\ 362\ 091\ 943\ 33(7)$
$1s3p$ ^3P	$-2.058\ 081\ 084\ 274\ 275\ 3(2)$
$1s3d$ ^1D	$-2.055\ 620\ 732\ 852\ 246\ 51(8)$
$1s3d$ ^3D	$-2.055\ 636\ 309\ 453\ 261\ 34(4)$

the calculated QED shift. For comparison, the two experimental values are

960 332 041.52 ± 0.21 MHz,
960 332 040.87 ± 0.15 MHz.

The first is obtained from an extrapolation of the $1s2s$ ^1S–$1snp$ ^1P transition frequencies to the series limit (Sansonetti and Gillaspy, 1992), and the second from the $1s2s$ ^1S–$1snd$ ^1D two-photon transition frequencies (Lichten *et al.*, 1991). Although the measurements do not quite agree with each other, taken together they determine the QED shift of the $1s2s$ ^1S state to an accuracy of about 100 parts per million and verify the calculated value to better than 0.1%. For the ground state, the calculated QED shift in the ionization energy has the much larger value $-(41\,233 \pm 35)$ MHz. This has recently been verified to an accuracy of ±45 MHz (±0.1%) from the total $1s^2$ ^1S–$1s2p$ ^1P transition frequency (Eikema *et al.*, 1996).

In summary, the results in Table 2 provide a firm foundation of nonrelativistic energies upon which higher-order corrections can be built and compared with experiment. Further improvements in the QED part of the theory remain an important challenge for the future.

3.3.5 Extensions to More Complex Systems

Fully correlated variational calculations of the type described in the previous section are difficult to extend to systems more complex than helium because of the rapid increase in the number of terms required. For an atom containing K electrons, there are K single-particle radial distances r_s and $K(K - 1)/2$ interparticle distances r_{st} for a total of $P = K(K + 1)/2$ radial coordinates. If all combinations of powers of the r_s and r_{st} are included in the basis set such that the sum of powers is $\leq \Omega$ [cf. Eq. (123)], then the generalization of Eq. (124) for the number of terms is

$$N = \frac{(\Omega + 1)(\Omega + 2) \cdots (\Omega + P)}{P!}. \qquad (126)$$

Since the time required to calculate a single eigenvector increases in proportion to N^3, the overall complexity of the calculation increases roughly in proportion to

$$\left[\frac{6(\Omega + P)!}{P!(\Omega + 3)!} \right]^3 \qquad (127)$$

relative to helium with the same Ω.

As an example, from Table 3, an accuracy of 10^{-10} a.u. for the ground state of helium requires $\Omega = 8$. A similar accuracy for lithium with $K = 3$ and $P = 6$ therefore requires about 6000 times the computer resources, and for beryllium with $K = 4$ and $P = 10$, the factor from expression (127) becomes 1.4×10^{13}.

Because of this rapid increase of complexity with the number of electrons, fully correlated calculations of spectroscopic accuracy have only been extended as far as lithium (see Yan and Drake, 1995; Yan *et al.*, 1996; and earlier references therein). The pattern of convergence for the ground state is similar to that shown in Table 3. The results up to $\Omega = 8$ yield the extrapolated nonrelativistic eigenvalue

$$E(1s^2 2s^2\, S) = -7.478\,060\,323\,10(31)\ \text{a.u.} \qquad (128)$$

The uncertainty of $\pm 3 \times 10^{-10}$ is about what one would expect from Table 3 for $\Omega = 8$.

For systems more complex than lithium, one must resort to other methods of calculation that can be extended to arbitrarily complex systems, but typically having much lower accuracy ($\pm 10^{-6}$ a.u. or more). These methods include multiconfiguration Hartree–Fock (MCHF), configuration-interaction (CI), many-body perturbation-theory, finite-element, diffusion Monte Carlo (DMC), and variational Monte Carlo (VMC) techniques. The MCHF and CI methods are similar in concept to the fully correlated variational method described in Sec. 3.3.4, except that the members of the basis set χ_p are constructed from antisymmetrized products of one-electron orbitals corresponding to definite electronic configurations. The effect is analogous to including only the *even* powers of r_{12} as shown in Table 1, and so convergence with increasing angular momentum of the individual electrons is slow. Recently, Goldman (1994) has devised a modified CI method involving extrapolation procedures to overcome this problem, at least for simple systems. For recent work on finite-element and many-body perturbation-theory meth-