

# The electronic Hamiltonian in an electromagnetic field

Trygve Helgaker

*Department of Chemistry, University of Oslo,*

*P.O.B. 1033 Blindern, N-0315 Oslo, Norway*

Poul Jørgensen and Jeppe Olsen

*Department of Chemistry, University of Aarhus, DK-8000 Århus C, Denmark*

## Contents

<b>I. Introduction</b>	5
<b>II. Classical mechanics</b>	5
A. Newtonian mechanics	5
B. Lagrangian mechanics	5
C. Hamiltonian mechanics	9
<b>III. Electromagnetic fields</b>	12
A. Maxwell's equations	12
B. The continuity equation	13
C. Scalar and vector potentials	14
D. Gauge transformations	15
E. Maxwell's equations in vacuum: electromagnetic waves	16
F. Maxwell's equations for static fields and charge distributions	17
G. Electric multipole expansions	18
H. Magnetic multipole expansions	21
I. The potential generated by a moving point-charge particle	22
<b>IV. Particle in an electromagnetic field</b>	24
A. Lagrangian mechanics	24
B. Hamiltonian mechanics	25
C. Gauge transformations	26
<b>V. Relativistic mechanics</b>	26
A. One-particle conservative systems	27
B. Relativistic particle in an electromagnetic field	28
C. Two-particle systems	30
<b>VI. The one-electron nonrelativistic Hamiltonian</b>	32
A. Quantization	32
B. The one-component one-electron Hamiltonian	34
C. The two-component one-electron Hamiltonian	35

D. Two-component wave functions and spin variables	39
E. Gauge transformation of the Schrödinger equation	41
<b>VII. The one-electron relativistic Hamiltonian</b>	<b>43</b>
A. The Dirac equation	43
B. Reduction of the Dirac equation	45
C. The first-order relativistic one-electron Hamiltonian	49
D. The two-component Pauli Hamiltonian	51
<b>VIII. The two-electron relativistic Hamiltonian</b>	<b>53</b>
A. The Breit equation	53
B. Matrix form of the Breit equation	54
C. Reduction of the Breit equation	55
D. Reduction of the two-electron Coulomb operator	58
E. Reduction of the Gaunt operator	59
F. Reduction of the retardation operator	61
<b>IX. Many-electron systems and the Breit-Pauli Hamiltonian</b>	<b>62</b>
<b>X. The molecular electronic Breit–Pauli Hamiltonian</b>	<b>64</b>
A. Nuclear electromagnetic fields	65
B. External electromagnetic fields	67
C. Purely nuclear contributions	68
D. Survey of terms in the molecular Breit–Pauli Hamiltonian	69
1. Kinetic energy	70
2. Coulomb interactions	71
3. External electric-field interactions	72
4. Zeeman interactions	72
5. Spin–orbit interactions	73
6. Spin–spin interactions	75
7. Orbit–orbit interactions	76
8. Diamagnetic interactions	76
<b>XI. Exercises</b>	<b>78</b>

A. Vector identities	81
B. Matrix direct products	82

## I. INTRODUCTION

## II. CLASSICAL MECHANICS

### A. Newtonian mechanics

In *Newtonian mechanics*, the motion of a (nonrelativistic) particle of mass  $m$  is described by *Newton's equation of motion*

$$\mathbf{F}(\mathbf{r}, \mathbf{v}, t) = m\mathbf{a} \quad (1)$$

where  $\mathbf{F}(\mathbf{r}, \mathbf{v}, t)$  is the *force* that acts on the particle and  $\mathbf{a}$  its acceleration in a Cartesian coordinate system. In general, the force is a function of the position  $\mathbf{r}$  of the particle, its velocity  $\mathbf{v}$ , and the time  $t$ . For a system of many particles, there is one set of equations for each particle, where the force may or may not depend on the other particles in the system.

The explicit form of  $\mathbf{F}(\mathbf{r}, \mathbf{v}, t)$  depends on the system studied. In a *conservative system*, it is (by definition) independent of the velocity and time. The force may then be obtained from a *potential*  $V(\mathbf{r})$  as follows:

$$\mathbf{F}(\mathbf{r}) = -\nabla V(\mathbf{r}) \quad (2)$$

In a *nonconservative system*, the force depends also on the velocity of the particle and possibly on time. In particular, in an electromagnetic field, a particle of charge  $z$  is subject to the *Lorentz force*

$$\mathbf{F}(\mathbf{r}, \mathbf{v}, t) = z(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (3)$$

where  $\mathbf{E}$  and  $\mathbf{B}$  are the *electric field strength* and the *magnetic induction*, respectively, of the electromagnetic field at the position of the particle.

### B. Lagrangian mechanics

The Newtonian formulation of mechanics in terms of forces acting on particles provides a complete description of classical mechanics in the sense that the solution of Newton's equation of motion leads to a correct description of the motion of the particles. Nevertheless, for many purposes, it is more convenient to use the *Lagrangian formulation* of classical mechanics. First, the Lagrangian formulation is more general than the Newtonian formulation in

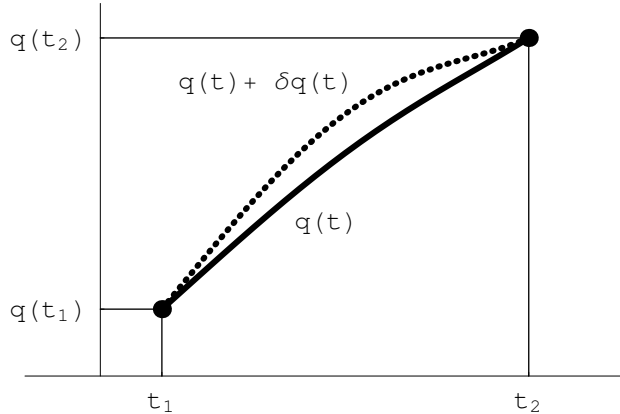


FIG. 1: Trajectories between the fixed points  $q(t_1)$  and  $q(t_2)$  in one-dimensional configuration space that it provides a unified description of particles and fields. Next, it provides a description of the system that, unlike the Newtonian formulation, is invariant to coordinate transformations and also handles constraints more naturally. Most important here, it provides the springboard to quantum mechanics.

Let us consider a classical system of  $n$  degrees of freedom—that is, a system in which the positions of all particles are uniquely specified by the values of  $n$  *generalized coordinates*  $q_i$ , which in general are functions of time  $q_i(t)$ . With each coordinate  $q_i$  in *configuration space*, we associate a *generalized velocity*  $\dot{q}_i$ , which represents the time derivative of the coordinate. The *principle of least action* (*Hamilton's principle*) then states that, for this system, there exists a *Lagrangian*  $L(\mathbf{q}, \dot{\mathbf{q}}, t)$  (with  $q_i$  and  $\dot{q}_i$  collected in vectors) such that the *action integral*

$$S = \int_{t_1}^{t_2} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) dt \quad (4)$$

is stationary  $\delta S = 0$  with respect to variations in the trajectory taken by the system as it moves from  $\mathbf{q}(t_1)$  to  $\mathbf{q}(t_2)$ . To emphasize that the time integration is for a given path  $\mathbf{q}(t)$ , this dependence has been indicated explicitly in Eq. (4). In Fig. 1, we have plotted the stationary path  $q(t)$  together with an allowed variation  $q(t) + \delta q(t)$  in a one-dimensional trajectory between the fixed end points  $q(t_1)$  and  $q(t_2)$ . Note that the trajectories are paths in the configuration space of dimension  $n$  equal to the number of coordinates  $q_i$ .

The Lagrangian is not unique since we may add to it any function that does not contribute to  $\delta S$ . In particular, consider the new Lagrangian

$$L' = L + \frac{\partial f}{\partial t} + \sum_i \dot{q}_i \frac{\partial f}{\partial q_i} \quad (5)$$

where  $f(\mathbf{q}, t)$  is an arbitrary velocity-independent function. Integration of this Lagrangian along a path gives the new action integral

$$\begin{aligned} S' &= \int_{t_1}^{t_2} \left( L + \frac{\partial f}{\partial t} + \sum_i \dot{q}_i \frac{\partial f}{\partial q_i} \right) dt = \int_{t_1}^{t_2} \left( L + \frac{df}{dt} \right) dt \\ &= S + f(\mathbf{q}_2, t_2) - f(\mathbf{q}_1, t_1) \end{aligned} \quad (6)$$

To within a constant,  $S$  and  $S'$  are therefore identical for all paths between the two fixed end points. Therefore, if  $L$  identifies the correct trajectory by  $\delta S = 0$ , then so does the transformed Lagrangian  $L'$  of Eq. (5). The velocity-independent function  $f(\mathbf{q}, t)$  is known as a *gauge function* and the associated transformation Eq. (5) is a *gauge transformation*.

Carrying out the variations in the action integral Eq. (4) and imposing the stationary condition, we obtain

$$\delta S = \sum_i \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) dt = 0 \quad (7)$$

Integrating the last term in Eq. (7) by parts and using the fact that the  $\delta q_i$  vanish at the end points, we obtain

$$\delta S = \sum_i \int_{t_1}^{t_2} \left[ \frac{\partial L}{\partial q_i} \delta q_i - \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i \right] dt = 0 \quad (8)$$

and conclude that the Lagrangian satisfies the  $n$  second-order differential equations (one for each coordinate)

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \quad (9)$$

These are *Lagrange's equations of motion*, which, unlike Newton's equations Eq. (1), preserve their form in any coordinate system. Whereas in Newtonian mechanics the force defines the system and is chosen so that Newton's equations Eq. (1) reproduce the motion of the system, in Lagrangian mechanics the Lagrangian defines the system and is determined so that Lagrange's equations Eq. (9) reproduce the motion of the system.

There exists no general prescription for constructing the Lagrangian. In the important case of a (nonrelativistic) particle in a conservative force field Eq. (2), the Lagrangian may be taken as the difference between the kinetic and potential energies

$$L(\mathbf{q}, \dot{\mathbf{q}}) = T(\mathbf{q}, \dot{\mathbf{q}}) - V(\mathbf{q}) \quad (10)$$

where the potential is related to the force as

$$F_i = - \frac{\partial V(\mathbf{q})}{\partial q_i} \quad (11)$$

In particular, for a single particle in Cartesian coordinates, the Lagrangian Eq. (10) takes the form

$$L(\mathbf{r}, \mathbf{v}) = \frac{1}{2}mv^2 - V(\mathbf{r}) \quad (12)$$

Lagrange's equations of motion Eq. (9), which in Cartesian coordinates take the form

$$\frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}} = \frac{\partial L}{\partial \mathbf{r}} \quad (13)$$

then reduce to Newton's equations Eq. (1) with the force Eq. (2):

$$\frac{d}{dt} m\mathbf{v} = -\nabla V(\mathbf{r}) \quad (14)$$

For a particle subject to the (nonconservative) Lorentz force Eq. (3), the Lagrangian takes a slightly different form, as discussed in Section IV A. For all systems of interest to us, the Lagrangian is thus readily set up, in any convenient coordinate system.

In the description of physical systems, it is important to identify conserved quantities. We here consider the *generalized momentum* conjugate to the generalized coordinate  $q_i$

$$p_i(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{\partial L}{\partial \dot{q}_i} \quad (15)$$

and the *energy function*

$$h(\mathbf{q}, \dot{\mathbf{q}}, t) = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L(\mathbf{q}, \dot{\mathbf{q}}, t) \quad (16)$$

Taking the total time derivative of the conjugate momentum and of the energy function, we obtain

$$\frac{dp_i}{dt} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \quad (17)$$

$$\begin{aligned} \frac{dh}{dt} &= \sum_i \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - \frac{dL}{dt} \\ &= \sum_i \left[ \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i - \frac{\partial L}{\partial q_i} \dot{q}_i - \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right] - \frac{\partial L}{\partial t} = -\frac{\partial L}{\partial t} \end{aligned} \quad (18)$$

where, in both cases, we have used Lagrange's equations Eq. (9) in the final step. We conclude that  $p_i$  and  $h$  are constants of motion provided the Lagrangian does not depend explicitly on  $q_i$  and  $t$ , respectively. For a conservative system of one particle in Cartesian coordinates with Lagrangian Eq. (12), the generalized momentum Eq. (15) and the energy function Eq. (16) become the *linear momentum* and the *total energy*, respectively:

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = m\mathbf{v} \quad (19)$$

$$h = \frac{\partial L}{\partial \mathbf{v}} \mathbf{v} - L = \frac{1}{2}mv^2 + V(\mathbf{r}) = T + V(\mathbf{r}) \quad (20)$$



These quantities are preserved whenever the potential is independent of  $\mathbf{r}$  and  $t$ , respectively.

### C. Hamiltonian mechanics

In Lagrangian mechanics, the particles move along trajectories that, for a system of  $n$  degrees of freedom, are determined by the  $n$  second-order differential equations Eq. (9) involving the Lagrangian  $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ . The motion of the particles is completely specified once the initial values of the  $n$  coordinates and the  $n$  velocities are given. We shall now consider the related *Hamiltonian formulation* of classical mechanics, where the  $n$  second-order differential equations are replaced by  $2n$  first-order equations. To achieve this, we shall proceed by promoting the conjugate momenta  $p_i$  to independent variables on a par with the generalized coordinates  $q_i$ .

To establish Hamiltonian mechanics, consider the differential of the energy function Eq. (16) for arbitrary changes in  $\mathbf{q}$ ,  $\dot{\mathbf{q}}$  and  $t$ , at a point on the path where Lagrange's equations Eq. (9) are satisfied:

$$\begin{aligned} dh &= d \left[ \sum_i p_i \dot{q}_i - L(\mathbf{q}, \dot{\mathbf{q}}, t) \right] = \sum_i (p_i d\dot{q}_i + \dot{q}_i dp_i - \dot{p}_i dq_i - p_i d\dot{q}_i) - \frac{\partial L}{\partial t} dt \\ &= \sum_i (-\dot{p}_i dq_i + \dot{q}_i dp_i) - \frac{\partial L}{\partial t} dt \end{aligned} \quad (21)$$

where we have used Eqs. (15) and (17). Noting that this differential depends on  $dp_i$  rather than on  $d\dot{q}_i$ , we introduce the *Hamiltonian function* as the energy function expressed as a function of  $\mathbf{q}$ ,  $\mathbf{p}$ , and  $t$  rather than as a function of  $\mathbf{q}$ ,  $\dot{\mathbf{q}}$ , and  $t$ :

$$H(\mathbf{q}, \mathbf{p}, t) = \sum_i p_i \dot{q}_i - L(\mathbf{q}, \dot{\mathbf{q}}, t) \quad (22)$$

Consider now the differential of the Hamiltonian function with respect to independent variations in  $\mathbf{q}$ ,  $\mathbf{p}$ , and  $t$ :

$$dH(\mathbf{q}, \mathbf{p}, t) = \sum_i \frac{\partial H}{\partial q_i} dq_i + \sum_i \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt \quad (23)$$

To within an arbitrary constant,  $H(\mathbf{q}, \mathbf{p}, t)$  and  $h(\mathbf{q}, \dot{\mathbf{q}}, t)$  become identical if we require the fulfillment of *Hamilton's equations of motion*

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad (24)$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (25)$$

and

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \quad (26)$$

In Hamiltonian mechanics, the  $n$  generalized coordinates  $q_i$  and their conjugate momenta  $p_i$  are considered independent variables in a  $2n$ -dimensional space called the *phase space*, connected by the  $2n$  first-order differential equations Eqs. (24) and (25). The system is then represented by points  $(\mathbf{q}, \mathbf{p})$  in phase space, on a trajectory that satisfies Hamilton's equations of motion.

As an example, we consider again a single particle in conservative force field. In Cartesian coordinates, the Lagrangian is given by Eq. (12) and the generalized momentum corresponds to the linear momentum Eq. (19). From Eq. (22), we then find that the energy function is given as

$$h(\mathbf{r}, \mathbf{v}) = \mathbf{p} \cdot \mathbf{v} - \frac{1}{2}mv^2 + V(\mathbf{r}) = \frac{1}{2}mv^2 + V(\mathbf{r}) \quad (27)$$

from which we obtain the Hamiltonian by using Eq. (19) to eliminate  $\mathbf{v}$ :

$$H(\mathbf{r}, \mathbf{p}) = \frac{p^2}{2m} + V(\mathbf{r}) \quad (28)$$

Substituting this Hamiltonian in Eqs. (24) and (25), we find that Hamilton's equations of motion in this particular case yield

$$\mathbf{v} = \frac{\mathbf{p}}{m} \quad (29)$$

$$\dot{\mathbf{p}} = -\nabla V(\mathbf{r}) \quad (30)$$

Eliminating  $\mathbf{p}$  from these equations, we recover Newton's equations of motion  $\mathbf{F} = m\mathbf{a}$  in Cartesian coordinates.

In Hamiltonian mechanics, *constants of motion* are easily identified. First, from Eq. (25), we note the conjugate momentum  $p_i$  is a constant of motion whenever the Hamiltonian function does not depend explicitly on  $q_i$ . Next, from Eqs. (18) and (26), we note that the total and partial time derivatives of the Hamiltonian are identical:

$$\frac{dH}{dt} = \frac{dh}{dt} = -\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} \quad (31)$$

The Hamiltonian is thus a constant of motion provided it does not depend explicitly on time.

Finally, let us consider the time development of a general dynamical variable  $A(\mathbf{q}, \mathbf{p}, t)$ :

$$\frac{dA}{dt} = \sum_i \left( \frac{\partial A}{\partial q_i} \dot{q}_i + \frac{\partial A}{\partial p_i} \dot{p}_i \right) + \frac{\partial A}{\partial t} = \sum_i \left( \frac{\partial A}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial H}{\partial q_i} \right) + \frac{\partial A}{\partial t} \quad (32)$$

where we have used Eqs. (24) and (25). Introducing the *Poisson bracket* of the two variables  $A$  and  $B$  as

$$\{A, B\} = \sum_i \left( \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i} \frac{\partial A}{\partial p_i} \right) \quad (33)$$

we may express the equations of motion of  $A$  in Eq. (32) succinctly as

$$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t} \quad (34)$$

Important special cases are  $\dot{q}_i = \{q_i, H\}$  and  $\dot{p}_i = \{p_i, H\}$  (i.e., Hamilton's equations) and  $dH/dt = \partial H/\partial t$  since  $\{H, H\} = 0$ . Therefore, a dynamical variable  $A(\mathbf{q}, \mathbf{p})$  that does not depend explicitly on time is a constant of motion provided its Poisson bracket with the Hamiltonian function vanishes. In classical mechanics, the Poisson brackets play a role similar to that of commutators in quantum mechanics. In particular, we note that the Poisson brackets among the conjugate coordinates and momenta are given by

$$\{q_i, q_j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{q_i, p_j\} = \delta_{ij} \quad (35)$$

Upon quantization, the dynamical variables of Hamiltonian mechanics are replaced by operators whose commutators are identical to the corresponding Poisson brackets multiplied by  $i\hbar$  such that  $i\hbar\{A, B\} \rightarrow [\hat{A}, \hat{B}]$ .

Let us summarize the procedure for establishing the Hamiltonian formulation of a classical system. We begin by choosing the generalized coordinates  $\mathbf{q}$ , in terms of which we set up the Lagrangian  $L(\mathbf{q}, \dot{\mathbf{q}}, t)$  such that the solution of Lagrange's equations Eq. (9) determines the motion of the particles. Next, we introduce the conjugate momenta  $\mathbf{p}$  according to Eq. (15) and set up the energy function  $h(\mathbf{q}, \dot{\mathbf{q}}, t)$  of Eq. (16). Finally, we invert Eq. (15) to eliminate the velocities from the energy function, obtaining the Hamiltonian function  $H(\mathbf{q}, \mathbf{p}, t)$  of Eq. (22) as a function of  $\mathbf{q}$  and  $\mathbf{p}$ . Finally, we solve Hamilton's equations of motion Eqs. (24) and (25) for the given initial conditions.

### III. ELECTROMAGNETIC FIELDS

#### A. Maxwell's equations

The electric field  $\mathbf{E}(\mathbf{r}, t)$  and magnetic induction  $\mathbf{B}(\mathbf{r}, t)$  of an electromagnetic field satisfy *Maxwell's equations* (in SI units)

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad \text{Coulomb's law} \quad (36)$$

$$\nabla \times \mathbf{B} - \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J} \quad \text{Ampère's law with Maxwell's correction} \quad (37)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (38)$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0} \quad \text{Faraday's law} \quad (39)$$

where  $\rho(\mathbf{r}, t)$  and  $\mathbf{J}(\mathbf{r}, t)$  are the *charge and current densities*, respectively, of the particles in the system, whereas  $\epsilon_0$  and  $\mu_0$  are the *electric constant* (the permittivity of vacuum) and the *magnetic constant* (the permeability of vacuum), respectively. As we shall see later, the electric and magnetic constants are related to the speed of light in vacuum  $c$  as  $\epsilon_0 \mu_0 = c^{-2}$ . When  $\rho$  and  $\mathbf{J}$  are known, Maxwell's equations may be solved for  $\mathbf{E}$  and  $\mathbf{B}$ , subject to suitable boundary conditions. Conversely, if  $\mathbf{E}$  and  $\mathbf{B}$  are known, then  $\rho$  and  $\mathbf{J}$  can be determined from Newton's equations with the Lorentz force Eq. (3).

In general, therefore, Maxwell's equations for the electromagnetic field must be solved simultaneously with Newton's equations for the particles in that field. This problem may be approached by setting up a Lagrangian for the total system of particles and field in such a manner that the solution of the Lagrange's equations is equivalent to the simultaneous solution of Maxwell's equations for the electromagnetic field and Newton's equations for the charged particles. Such a general approach is not required here, however, since we shall take the externally applied electromagnetic field to be fixed and unaffected by the presence of the particles. Consequently, we shall here restrict ourselves to setting up a Lagrangian (and subsequently the Hamiltonian) for these particles, subject to the Lorentz force. However, before such a Lagrangian is constructed in Section IV, we shall in this section consider some of the properties of the electromagnetic field itself, as obtained by solving Maxwell's equations Eq. (36)–(39).

Before considering the solution of Maxwell's equations in special cases, it is worthwhile to point out that these equations uniquely specify the electric field and the magnetic induction,

subject to suitable boundary conditions. This result follows from Helmholtz' theorem, which states that any vector field  $\mathbf{F}$  is uniquely determined by its divergence  $\nabla \cdot \mathbf{F}$  and by its curl  $\nabla \times \mathbf{F}$  in a given volume of space  $V$ , provided we also specify its component  $F_{\perp}$  normal to the surface of  $V$ . For example, Maxwell's equations uniquely determine  $\mathbf{E}$  and  $\mathbf{B}$  if these are required to vanish at infinity.

## B. The continuity equation

The inhomogeneous pair of Maxwell's equations Eqs. (36) and (37) relate the charge and current densities to the electric field and the magnetic induction. However, charge cannot be destroyed and Maxwell's equations must be consistent with this observation. Taking the time derivative of Coulomb's law Eq. (36) and the divergence of Ampère's law Eq. (37), we obtain

$$\frac{\partial \rho}{\partial t} = \epsilon_0 \nabla \cdot \frac{\partial \mathbf{E}}{\partial t} \quad (40)$$

$$\nabla \cdot \mathbf{J} = \mu_0^{-1} \nabla \cdot \nabla \times \mathbf{B} - \epsilon_0 \nabla \cdot \frac{\partial \mathbf{E}}{\partial t} \quad (41)$$

Adding these two equations together and noting that  $\nabla \cdot \nabla \times \mathbf{F} = 0$  for any vector  $\mathbf{F}$  [Eq. (A9)], we obtain the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \quad (42)$$

For a physical interpretation of the continuity equation, we integrate over a volume of space  $V$  to obtain

$$\int_V \left( \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} \right) d\mathbf{r} = 0 \quad (43)$$

and invoke Gauss' theorem, which states that, for an arbitrary vector function  $\mathbf{F}$ ,

$$\int_V \nabla \cdot \mathbf{F} d\mathbf{r} = \int_S \mathbf{n} \cdot \mathbf{F} dS \quad (44)$$

where  $S$  is the surface that encloses  $V$  and  $\mathbf{n}$  its unit outward normal. Inserting this result into Eq. (43), we obtain

$$\frac{\partial}{\partial t} \int_V \rho d\mathbf{r} + \int_S \mathbf{n} \cdot \mathbf{J} dS = 0 \quad (45)$$

which is an expression of *charge conservation*, relating the change of the charge in a volume to the flow of charge out of this volume. In static systems,  $\partial \rho / \partial t = 0$  and the continuity equation Eq. (42) then shows that the current density becomes divergenceless  $\nabla \cdot \mathbf{J} = 0$ .

### C. Scalar and vector potentials

The homogeneous pair of Maxwell's equations Eqs. (38) and (39) are independent of the charge and current densities, being mathematical identities involving only  $\mathbf{E}$  and  $\mathbf{B}$ . Since  $\nabla \cdot \nabla \times \mathbf{F} = 0$  for any vector function  $\mathbf{F}$  [Eq. (A9)] and since any *solenoidal (divergenceless) vector* may be expressed as  $\nabla \times \mathbf{F}$ , we may satisfy Eq. (38) by writing the magnetic induction  $\mathbf{B}$  as the curl of a vector  $\mathbf{A}$

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (46)$$

Substituting this expression into the second homogeneous equation Eq. (39), we obtain

$$\nabla \times \left( \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = \mathbf{0} \quad (47)$$

Next, since  $\nabla \times \nabla f = \mathbf{0}$  for any scalar function  $f$  [Eq. (A10)] and since any *irrotational (curl-free) vector* may be expressed as  $\nabla f$ , Eq. (39) is automatically satisfied by introducing the scalar function  $\phi$  and writing

$$\mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \quad (48)$$

In conclusion, the homogeneous pair of Maxwell's equations Eqs. (38) and (39) is satisfied whenever the electric field and the magnetic induction are expressed in terms of a *scalar potential*  $\phi(\mathbf{r}, t)$  and a *vector potential*  $\mathbf{A}(\mathbf{r}, t)$  according to Eqs. (46) and (48). We note that the potentials  $(\phi, \mathbf{A})$  contain four rather than six components as in  $(\mathbf{E}, \mathbf{B})$ . The scalar and vector potentials thus provide a more compact representation of the electromagnetic field than do the observable electric field strength  $\mathbf{E}$  and magnetic induction  $\mathbf{B}$ .

Substituting Eqs. (46) and (48) into the inhomogeneous Maxwell's equations Eqs. (36) and (37), we obtain using Eq. (A11) the following second-order partial differential equations for the scalar and vector potentials

$$\nabla^2 \phi + \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) = -\frac{\rho}{\epsilon_0} \quad (49)$$

$$\nabla^2 \mathbf{A} - \nabla (\nabla \cdot \mathbf{A}) - \epsilon_0 \mu_0 \frac{\partial^2}{\partial t^2} \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial}{\partial t} \nabla \phi = -\mu_0 \mathbf{J} \quad (50)$$

which must be solved with the appropriate boundary conditions imposed. Before considering the solution of Maxwell's equations for the potentials, we shall consider an important point regarding these potentials—namely, that they are not uniquely defined and may be subjected to certain transformations without affecting the physical situation.

## D. Gauge transformations

The observable electromagnetic fields  $\mathbf{E}$  and  $\mathbf{B}$  are uniquely determined by the physical system. By contrast, the potentials  $\phi$  and  $\mathbf{A}$  representing these fields are not unique. To see this, let  $\phi(\mathbf{q}, t)$  and  $\mathbf{A}(\mathbf{q}, t)$  be a pair of potentials that represent the observed fields:

$$\mathbf{E}(\phi, \mathbf{A}) = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} \quad (51)$$

$$\mathbf{B}(\phi, \mathbf{A}) = \nabla \times \mathbf{A} \quad (52)$$

Let us now generate, from these potentials, a new pair of *gauge-transformed potentials*

$$\phi' = \phi - \frac{\partial f(\mathbf{q}, t)}{\partial t} \quad (53)$$

$$\mathbf{A}' = \mathbf{A} + \nabla f(\mathbf{q}, t) \quad (54)$$

where the *gauge function*  $f(\mathbf{q}, t)$  is independent of the velocities  $\dot{\mathbf{q}}$ . Such a gauge transformation does not affect the physical fields:

$$\mathbf{E}(\phi', \mathbf{A}') = -\nabla\phi' - \frac{\partial\mathbf{A}'}{\partial t} = -\nabla\phi + \nabla\frac{\partial f}{\partial t} - \frac{\partial\mathbf{A}}{\partial t} - \frac{\partial\nabla f}{\partial t} = \mathbf{E}(\phi, \mathbf{A}) \quad (55)$$

$$\mathbf{B}(\phi', \mathbf{A}') = \nabla \times (\mathbf{A} + \nabla f) = \mathbf{B} + \nabla \times \nabla f = \mathbf{B}(\phi, \mathbf{A}) \quad (56)$$

where we have used Eq. (A10). We are therefore free to choose  $f(\mathbf{q}, t)$  such that the vector potential satisfies additional conditions. In particular, we may require its divergence to be equal to some velocity-independent function  $g(\mathbf{q}, t)$ :

$$\nabla \cdot \mathbf{A}' = g \quad (57)$$

The importance of fixing the divergence stems from Helmholtz theorem, which states that the curl  $\nabla \times \mathbf{A}' = \mathbf{B}$  and the divergence  $\nabla \cdot \mathbf{A}' = g$  together uniquely determine  $\mathbf{A}'$ , subject to suitable boundary conditions. From Eq. (54), we see that Eq. (57) implies that the gauge function satisfies *Poisson's equation*

$$\nabla^2 f = g - \nabla \cdot \mathbf{A} \quad (58)$$

Note that, even though Poisson's equation fixes the divergence of the potential, it does not determine the gauge function and hence the vector potential completely as any solution of the homogeneous *Laplace's equation*

$$\nabla^2 f_h = 0 \quad (59)$$

may be added to the gauge function in Eq. (58) without affecting the divergence and curl of the vector potential; the solution to Laplace's equation are the solid harmonic functions discussed in Section 6.4.1.

The gauge may be chosen for our convenience. In relativistic work, for example, the *Lorentz gauge* is typically used, requiring the scalar and vector potentials to satisfy the conditions

$$\nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 \frac{\partial \phi}{\partial t} = 0 \quad (60)$$

giving the following symmetric (Lorentz invariant) form of Maxwell's equations Eqs. (49) and (50):

$$\nabla^2 \phi - \epsilon_0 \mu_0 \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0} \quad (\text{Lorentz gauge}) \quad (61)$$

$$\nabla^2 \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{J} \quad (\text{Lorentz gauge}) \quad (62)$$

We shall always work in the *Coulomb gauge*, where the vector potential is divergenceless:

$$\nabla \cdot \mathbf{A} = 0 \quad (63)$$

In the Coulomb gauge, Maxwell's equations Eqs. (49) and (50) take the less symmetric form

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0} \quad (\text{Coulomb gauge}) \quad (64)$$

$$\nabla^2 \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} - \epsilon_0 \mu_0 \frac{\partial}{\partial t} \nabla \phi = -\mu_0 \mathbf{J} \quad (\text{Coulomb gauge}) \quad (65)$$

We note that the Coulomb and Lorentz gauges are identical in the static case. In the following sections, we shall discuss the solution to Maxwell's equations in the Coulomb gauge in two important simple cases—namely, in vacuum and for static fields.

### E. Maxwell's equations in vacuum: electromagnetic waves

In vacuum, all charges and currents vanish. In the Coulomb gauge, we may then write Maxwell's equations Eqs. (64) and (65) in the form

$$\nabla^2 \phi = 0 \quad (66)$$

$$\nabla^2 \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = \epsilon_0 \mu_0 \frac{\partial}{\partial t} \nabla \phi \quad (67)$$



From Eq. (66), we note that the scalar potential  $\phi$  is a solid-harmonic function, satisfying Laplace's equation. As seen from Eq. (54), we may then carry out a further transformation with the solid-harmonic gauge function  $f = t\phi$  without affecting the Coulomb gauge of the vector potential. Furthermore, we see from Eq. (53) that this transformation eliminates the scalar potential entirely. Therefore, Eq. (67) reduces to the wave equations

$$\nabla^2 \mathbf{A} = c^{-2} \frac{\partial^2 \mathbf{A}}{\partial t^2} \quad (68)$$

for waves propagating with speed  $c = 1/\sqrt{\mu_0\epsilon_0}$ . A general solution to Maxwell's equations in vacuum Eq. (68) is given by the plane waves

$$\phi(\mathbf{r}, t) = 0 \quad (69)$$

$$\mathbf{A}(\mathbf{r}, t) = \text{Re} [\mathbf{A}_0 \exp(\mathbf{i}\mathbf{k} \cdot \mathbf{r} - i\omega t)] \quad (70)$$

where  $\mathbf{A}_0$  is complex, whereas the wave vector  $\mathbf{k}$  and angular frequency  $\omega$  are real and related as

$$k = \|\mathbf{k}\| = \frac{\omega}{c} \quad (71)$$

To satisfy the Coulomb gauge-condition Eq. (63) for all  $t$ , we must impose  $\mathbf{k} \cdot \mathbf{A} = 0$  as seen by taking the divergence of Eq. (70) using Eq. (A4). The corresponding electric field and the magnetic induction are obtained from Eqs. (51) and (52) as

$$\mathbf{E}(\mathbf{r}, t) = \omega \text{Im} \mathbf{A}(\mathbf{r}, t) \quad (72)$$

$$\mathbf{B}(\mathbf{r}, t) = \mathbf{k} \times \text{Im} \mathbf{A}(\mathbf{r}, t) \quad (73)$$

In vacuum,  $\mathbf{E}$  and  $\mathbf{B}$  are thus transverse waves  $\mathbf{k} \cdot \mathbf{E} = \mathbf{k} \cdot \mathbf{B} = 0$  in phase with each other, traveling in the direction of  $\mathbf{k}$  in such a manner that  $\mathbf{k}$ ,  $\mathbf{E}$  and  $\mathbf{B} = \omega^{-1}\mathbf{k} \times \mathbf{E}$  constitute a right-handed coordinate system.

## F. Maxwell's equations for static fields and charge distributions

In electrostatics and magnetostatics, we study static fields and static charge and current distributions. Maxwell's equations Eq. (64) and (65) then become

$$\nabla^2 \phi = -\rho/\epsilon_0 \quad (\text{electrostatics}) \quad (74)$$

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J} \quad (\text{magnetostatics}) \quad (75)$$

from which the electric and magnetic fields are obtained from Eqs. (51) and (52) as

$$\mathbf{E} = -\nabla\phi \quad (\text{electrostatics}) \quad (76)$$

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (\text{magnetostatics}) \quad (77)$$

In the static limit, there is thus a complete separation of electric and magnetic phenomena and the charge and current densities become, respectively, the sources of the scalar and vector potentials Eqs. (74) and (75). Consider the following scalar and vector potentials, obtained from the charge and current densities  $\rho(\mathbf{r}')$  and  $\mathbf{J}(\mathbf{r}')$ :

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (\text{electrostatics}) \quad (78)$$

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (\text{magnetostatics}) \quad (79)$$

These potentials satisfy Eqs. (74) and (75), as readily verified by using the relation  $\nabla^2 r^{-1} = -4\pi\delta(\mathbf{r})$  of Eq. (A19). Whereas  $\phi(\mathbf{r})$  represents the electrostatic potential set up by the charge density  $\rho(\mathbf{r})$ ,  $\mathbf{A}(\mathbf{r})$  is the magnetostatic potential generated by the current density  $\mathbf{J}(\mathbf{r})$ . In Exercise 1, we show that the vector potential Eq. (79) satisfies the Coulomb gauge condition Eq. (63), as assumed in Eqs. (74) and (75).

We may now understand the reason for the term *Coulomb gauge*, introduced in Section IIID. In electrostatics, Coulomb's law Eq. (36) is equivalent to Poisson's equation Eq. (74). In the Coulomb gauge Eq. (63), we ensure that the relationship between the potential and the charge distribution is always of this form Eq. (64), even in time-dependent situations, unlike in the Lorentz gauge Eq. (61).

### G. Electric multipole expansions

The expression for the scalar potential  $\phi(\mathbf{r})$  in Eq. (78) is completely general, valid for any electrostatic charge distribution  $\rho(\mathbf{r}')$ . Often, however, the observer is far removed from the source of the potential. In such cases, the detailed form of the charge distribution matters less and we may evaluate the potential at the position of the observer in simple fashion, by means of a *multipole expansion*.

To set up the multipole expansion of the potential  $\phi(\mathbf{r})$  generated by the charge distribution  $\rho(\mathbf{r}')$ , we put the origin somewhere inside the charge distribution, using  $\mathbf{r}$  to denote the position of the observer (outside the charge density) and  $\mathbf{r}'$  to denote positions inside the

charge density. Far away from the density,  $r' \ll r$  and we may expand  $|\mathbf{r} - \mathbf{r}'|^{-1}$  in orders of  $r'/r$ :

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{r} + \frac{\mathbf{r} \cdot \mathbf{r}'}{r^3} + \frac{3(\mathbf{r} \cdot \mathbf{r}')^2 - r'^2 r^2}{2r^5} - \frac{2\pi}{3} \delta(\mathbf{r}) r'^2 + \dots \quad (80)$$

where we have used Eqs. (A16) and (A17). Inserting this expression in Eq. (78) and integrating over the primed variables, we obtain the multipole expansion of the scalar potential

$$\phi(\mathbf{r}) = \frac{q}{4\pi\epsilon_0 r} + \frac{\mathbf{p}^T \mathbf{r}}{4\pi\epsilon_0 r^3} + \frac{3\mathbf{r}^T \mathbf{M} \mathbf{r} - r^2 \text{tr} \mathbf{M}}{8\pi\epsilon_0 r^5} - \frac{\delta(\mathbf{r}) \text{tr} \mathbf{M}}{6\epsilon_0} + \dots \quad (81)$$

where we have introduced the total charge, the dipole moment, and the second moment, respectively, of the charge distribution:

$$q = \int \rho(\mathbf{r}') d\mathbf{r}' \quad (82)$$

$$\mathbf{p} = \int \mathbf{r}' \rho(\mathbf{r}') d\mathbf{r}' \quad (83)$$

$$\mathbf{M} = \int \mathbf{r}' \mathbf{r}'^T \rho(\mathbf{r}') d\mathbf{r}' \quad (84)$$

Often, the multipole expansion is expressed in terms of the quadrupole moment—that is, the traceless part of the second moment defined as

$$\Theta = \frac{3}{2} \mathbf{M} - \frac{1}{2} (\text{tr} \mathbf{M}) \mathbf{I}_3 \quad (85)$$

The third term in Eq. (81) may then be written in variety of ways

$$\frac{3\mathbf{r}^T \mathbf{M} \mathbf{r} - r^2 \text{tr} \mathbf{M}}{8\pi\epsilon_0 r^5} = \frac{\text{tr} \Theta (3\mathbf{r} \mathbf{r}^T - r^2 \mathbf{I}_3)}{12\pi\epsilon_0 r^5} = \frac{\mathbf{r}^T \Theta \mathbf{r}}{4\pi\epsilon_0 r^5} \quad (86)$$

in terms of the second moment or the quadrupole moment. The second expression, in particular, represents the interaction as the trace of the quadrupole and electric field-gradient tensors.

In the multipole expansion Eq. (81), the leading term represents the potential generated by a shapeless point-charge particle. The remaining terms represent different corrections to this simple description. The dipolar and quadrupolar terms are shape corrections, depending on the orientation of the  $\rho(\mathbf{r}')$  relative to the observer. If the total charge  $q$  is zero, the dipolar term becomes the leading term; if  $q$  and  $\mathbf{p}$  are both zero, the quadrupolar term becomes leading. Higher-order terms may be generated in the same manner but it is then more economical to expand in terms of solid-harmonic functions rather than in terms of

Cartesian moments as done here—see the discussion in Section 9.13. A particularly important multipole expansion is the expansion of the scalar potential generated by atomic nuclei, introduced in Section X A.

The contact term in Eq. (81) represents a size correction to the point-charge description, reducing the interaction at the position of the source. For an *external* observer, it does not contribute to the interaction but it is needed whenever the charge distribution  $\rho(\mathbf{r}')$  extends to the observer—for example, when an electron, moving in the potential generated by the nuclei in a molecule, comes into contact with a nucleus.

Multipole expansions are also useful for evaluating the interaction between a charge distribution  $\rho(\mathbf{r})$  and a static external potential  $\phi(\mathbf{r})$ . As shown in Section IV, the energy of a point particle of charge  $q$  in a potential  $\phi(\mathbf{r})$  is given by  $q\phi(\mathbf{r})$ . For a continuous charge distribution  $\rho(\mathbf{r})$ , therefore, the energy of interaction is given by

$$W = \int \rho(\mathbf{r})\phi(\mathbf{r}) \, d\mathbf{r} \quad (87)$$

Expanding in orders of  $\mathbf{r}$  about an origin  $\mathbf{r}_0$  inside  $\rho(\mathbf{r})$  and introducing the electric field and field gradient of the electrostatic potential

$$\mathbf{E} = -\nabla\phi(\mathbf{r}) \quad (88)$$

$$\mathbf{V} = -\nabla\nabla^T\phi(\mathbf{r}) \quad (89)$$

we find that the energy interaction Eq. (87) may be expressed as

$$W = q\phi - \mathbf{p}^T\mathbf{E} - \frac{1}{2} \text{tr} \mathbf{M}\mathbf{V} + \dots \quad (90)$$

where  $q$ ,  $\mathbf{p}$  and  $\mathbf{M}$  are the charge, dipole moment and second moment of  $\rho(\mathbf{r})$  evaluated about the origin according to Eqs. (82)–(84), respectively. Note that the potential  $\phi$ , field  $\mathbf{E}$ , and field gradient  $\mathbf{V}$  in Eq. (90) must be evaluated at the origin  $\mathbf{r}_0$  of the multipole expansion. For an external potential, its source resides outside  $\rho(\mathbf{r})$ . According to Eq. (74), the field gradient generated by the source is therefore traceless inside  $\rho(\mathbf{r})$ :

$$\text{tr} \mathbf{V} = -\text{tr}(\nabla\nabla^T\phi) = -\nabla^2\phi = 0 \quad (91)$$

The trace of the quadrupole moment then does not interact with the field gradient

$$\text{tr}(\mathbf{M}\mathbf{V}) = \text{tr}(\mathbf{M}\mathbf{V}) - \frac{1}{3}(\text{tr} \mathbf{M})(\text{tr} \mathbf{V}) = \frac{2}{3} \text{tr}(\Theta\mathbf{V}) \quad (92)$$

and we may evaluate the total energy of interaction as

$$W_{\text{ext}} = q\phi - \mathbf{p}^T \mathbf{E} - \frac{1}{3} \text{tr} \Theta \mathbf{V} + \dots \quad (93)$$

We note the characteristic manner in which the total charge interacts with the potential, the dipole moment with the electric field, the traceless quadrupole moment with the electric field gradient, and so on.

## H. Magnetic multipole expansions

When the source  $\mathbf{J}(\mathbf{r}')$  of the vector potential  $\mathbf{A}(\mathbf{r})$  is far removed from the observer at  $\mathbf{r}$  in Eq. (79), it becomes advantageous to represent  $\mathbf{A}(\mathbf{r})$  by a multipole expansion, analogous to the expansion of the scalar potential Eq. (48) in Section III G. Thus, inserting Eq. (80) in Eq. (79), we obtain to first order

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi r} \int \mathbf{J}(\mathbf{r}') d\mathbf{r}' + \frac{\mu_0}{4\pi r^3} \int (\mathbf{r} \cdot \mathbf{r}') \mathbf{J}(\mathbf{r}') d\mathbf{r}' + \dots \quad (94)$$

To simplify this expansion, we note the relation

$$\int [\nabla' f(\mathbf{r}')] \cdot \mathbf{J}(\mathbf{r}') d\mathbf{r}' = \int \nabla' \cdot [f(\mathbf{r}') \mathbf{J}(\mathbf{r}')] d\mathbf{r}' = 0 \quad (95)$$

The first identity follows from Eq. (A4) and from  $\nabla' \cdot \mathbf{J}(\mathbf{r}') = 0$ , valid for static systems as discussed in Section III B; the second identity follows from Gauss' theorem Eq. (44) since, for localized currents, the integrand vanishes for large  $\mathbf{r}'$ . Substituting first  $f = r'_i$  and then  $f = r_i r'_i r'_j$  in Eq. (95), we obtain the relations

$$\int J_i(\mathbf{r}') d\mathbf{r}' = 0 \quad (96)$$

$$\int [r_i r'_i J_j(\mathbf{r}') + r_i r'_j J_i(\mathbf{r}')] d\mathbf{r}' = 0 \quad (97)$$

The first relation tells us that the zero-order term in Eq. (94) vanishes, whereas the second relation enables us to rewrite the integral of the first-order term as

$$\int (\mathbf{r} \cdot \mathbf{r}') \mathbf{J}(\mathbf{r}') d\mathbf{r}' = \frac{1}{2} \int [(\mathbf{r} \cdot \mathbf{r}') \mathbf{J}(\mathbf{r}') - (\mathbf{r} \cdot \mathbf{J}(\mathbf{r}')) \mathbf{r}'] d\mathbf{r}' = \frac{1}{2} \int \mathbf{r} \times [\mathbf{J}(\mathbf{r}') \times \mathbf{r}'] d\mathbf{r}' \quad (98)$$

where we have used Eq. (A2) in the last step. Inserting this result into Eq. (94), we obtain the following expansion of the vector potential Eq. (79):

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \mathbf{r}}{r^3} + \dots \quad (99)$$

where the magnetic dipole moment is given by

$$\mathbf{m} = \frac{1}{2} \int [\mathbf{r}' \times \mathbf{J}(\mathbf{r}')] d\mathbf{r}' \quad (100)$$

There is no zero-order term (in agreement with the observation that magnetic monopoles do not exist) and high-order terms are seldom needed. In Section X A, we shall use Eq. (99) to represent the magnetic fields generated by the atomic nuclei.

### I. The potential generated by a moving point-charge particle

Let us now consider a point-charge particle of charge  $z$ . For such a particle, we may write the charge density and the current density in the form

$$\rho(\mathbf{r}) = z\delta(\mathbf{r}) \quad (101)$$

$$\mathbf{J}(\mathbf{r}) = z\mathbf{v}\delta(\mathbf{r}) \quad (102)$$

Here  $\mathbf{r}$  is the position of the observer relative to the particle, whereas  $\mathbf{v}$  the velocity of the particle relative to the observer. We shall assume that the particle is moving at constant velocity so that its position at a given time may be written as

$$\mathbf{r}(t) = \mathbf{r}_0 - \mathbf{v}t + \mathcal{O}(v^2/c^2) \quad (103)$$

where  $\mathbf{r}_0$  is the position at time  $t = 0$  and where the term proportional to  $v^2/c^2$  has been included from relativistic considerations. The potentials generated by such a particle cannot be written in closed form but, for small  $v/c$ , they are accurately represented by their leading terms:

$$\phi(\mathbf{r}) = \frac{z}{4\pi\epsilon_0 r} + \mathcal{O}(v^4/c^4) \quad (104)$$

$$\mathbf{A}(\mathbf{r}) = \frac{z}{8\pi\epsilon_0 c^2 r^3} [r^2\mathbf{v} + (\mathbf{v} \cdot \mathbf{r})\mathbf{r}] + \mathcal{O}(v^3/c^4) \quad (105)$$

It should be emphasized that the expansions of the scalar and vector potentials considered here are different from those considered in Sections III G and III H. There, we expanded the potentials generated stationary charge and current densities, sufficiently distant that their detailed form does not matter in the evaluation of the interaction. In the present section, we expand the potentials generated by moving point-charge particles, sufficiently slow for high-order relativistic corrections to be neglected in the evaluation of the interaction.

We shall now demonstrate that these potentials satisfy Maxwell's equations Eqs. (64) and (65) with the charge and current densities given by Eqs. (101) and (102):

$$\nabla^2 \phi(\mathbf{r}) = -\frac{z}{\epsilon_0} \delta(\mathbf{r}) \quad (106)$$

$$\nabla^2 \mathbf{A}(\mathbf{r}) - c^{-2} \frac{\partial^2}{\partial t^2} \mathbf{A}(\mathbf{r}) - c^{-2} \frac{\partial}{\partial t} \nabla \phi(\mathbf{r}) = -\mu_0 z \mathbf{v} \delta(\mathbf{r}) \quad (107)$$

and also that the Coulomb-gauge condition Eq. (63) is satisfied.

Inserting the scalar potential Eq. (104) into Eq. (106) and invoking  $\nabla^2 r^{-1} = -4\pi\delta(\mathbf{r})$  of Eq. (A19), we confirm that the first of Maxwell's equations are satisfied. Next, we rewrite the vector potential using Eq. (A20):

$$\mathbf{A}(\mathbf{r}) = \frac{z}{8\pi\epsilon_0 c^2} [(\mathbf{I}_3 \nabla^2 - \nabla \nabla^T) r] \mathbf{v} + \mathcal{O}(v^3/c^4) \quad (108)$$

The divergence  $\nabla \cdot \mathbf{A}(\mathbf{r})$  vanishes since, for any vector  $\mathbf{a}$ ,

$$\nabla^T (\mathbf{I}_3 \nabla^2 - \nabla \nabla^T) \mathbf{a} = \nabla^2 \nabla^T \mathbf{a} - \nabla^T \nabla \nabla^T \mathbf{a} = 0 \quad (109)$$

demonstrating that the vector potential satisfies the Coulomb-gauge condition. With the vector potential expressed in this manner and recalling that  $r$  depends explicitly on time Eq. (103), we obtain

$$\nabla^2 \mathbf{A}(\mathbf{r}) = \frac{\mu_0 z}{4\pi} [(\mathbf{I}_3 \nabla^2 - \nabla \nabla^T) r^{-1}] \mathbf{v} + \mathcal{O}(v^3/c^4) \quad (110)$$

$$-c^{-2} \frac{\partial^2 \mathbf{A}(\mathbf{r})}{\partial t^2} = -c^{-2} (\mathbf{v} \cdot \nabla)^2 \mathbf{A}(\mathbf{r}) = \mathcal{O}(v^3/c^4) \quad (111)$$

$$-c^{-2} \nabla \frac{\partial \phi(\mathbf{r})}{\partial t} = c^{-2} [\nabla \nabla^T \phi(\mathbf{r})] \mathbf{v} = \frac{\mu_0 z}{4\pi} (\nabla \nabla^T r^{-1}) \mathbf{v} + \mathcal{O}(v^5/c^6) \quad (112)$$

where we have used  $\nabla^2 r = 2r^{-1}$  of Eq. (A18) in Eq. (110),  $\partial f(\mathbf{r})/\partial t = -\mathbf{v} \cdot \nabla f(\mathbf{r})$  in Eqs. (111) and (112), and  $\mu_0 \epsilon_0 = c^{-2}$ . Substituting these results into Eq. (107) and using Eq. (A19), we conclude that Maxwell's equations are satisfied to the required order for a point-charge particle.

The vector potential generated by a moving point-charge particle Eq. (105) is sometimes decomposed as

$$\mathbf{A}(\mathbf{r}) = \mathbf{A}^m(\mathbf{r}) + \mathbf{A}^r(\mathbf{r}) + \mathcal{O}(v^3/c^4) \quad (113)$$

$$\mathbf{A}^m(\mathbf{r}) = \frac{z\mathbf{v}}{4\pi\epsilon_0 c^2 r} \quad (114)$$

$$\mathbf{A}^r(\mathbf{r}) = -\frac{z}{8\pi\epsilon_0 c^2 r^3} [r^2 \mathbf{v} - (\mathbf{v} \cdot \mathbf{r}) \mathbf{r}] \quad (115)$$

where the first contribution, known as the *instantaneous magnetic vector potential*. This potential may be obtained by substituting Eq. (102) into the expression for the magnetostatic potential Eq. (79). The second contribution is a *retardation correction* to the instantaneous scalar potential, arising since all electromagnetic interactions propagate with the finite velocity  $c$ .

## IV. PARTICLE IN AN ELECTROMAGNETIC FIELD

### A. Lagrangian mechanics

To describe a system containing a particle in an electromagnetic field, we must set up a Lagrangian such that Lagrange's equations of motion Eq. (9) reduce to Newton's equations of motion Eq. (1) with the Lorentz force Eq. (3). Since the Lorentz force is velocity dependent, this is not a conservative system, for which  $L = T - V$  where  $V$  is related to the force as  $F_i = -\partial V/\partial q_i$ . Rather, it belongs to a broader class of systems for which the Lagrangian may be written in terms of a velocity-dependent *generalized potential*  $U(\mathbf{q}, \dot{\mathbf{q}}, t)$  as follows:

$$L = T - U \quad (116)$$

$$F_i = -\frac{\partial U}{\partial q_i} + \frac{d}{dt} \left( \frac{\partial U}{\partial \dot{q}_i} \right) \quad (117)$$

In Exercise 2, we show that, for a particle of charge  $z$  subject to the Lorentz force Eq. (3), the generalized potential takes the form

$$U = z(\phi - \mathbf{v} \cdot \mathbf{A}) \quad (118)$$

where  $\phi$  and  $\mathbf{A}$  are the scalar and vector potentials, respectively, and  $\mathbf{v}$  the velocity of the particle. For such a particle, therefore, the Lagrangian becomes

$$L = T - z(\phi - \mathbf{v} \cdot \mathbf{A}) \quad (119)$$

To see that this Lagrangian provides a correct description of a particle in an electromagnetic field, we note that, in a Cartesian coordinate system with  $T = \frac{1}{2}mv^2$ , Lagrange's equations Eq. (9) become

$$\frac{d}{dt} \frac{\partial T}{\partial \mathbf{v}} = -\frac{\partial U}{\partial \mathbf{r}} + \frac{d}{dt} \frac{\partial U}{\partial \mathbf{v}} \quad (120)$$



Whereas the left-hand side evaluates to  $m\mathbf{a}$ , the right-hand side is the generalized force of Eq. (117) in Cartesian form and evaluates to the Lorentz force Eq. (3) as demonstrated in Exercise 2.

## B. Hamiltonian mechanics

To arrive at the Hamiltonian description of a particle subject to the Lorentz force, we restrict ourselves to a Cartesian coordinate system and introduce the conjugate momentum of Eq. (15)

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{\partial}{\partial \mathbf{v}} \left[ \frac{1}{2}mv^2 - z(\phi - \mathbf{v} \cdot \mathbf{A}) \right] = m\mathbf{v} + z\mathbf{A} \quad (121)$$

where we have used the Lagrangian of Eq. (119). Next, introducing the *kinetic momentum*  $\boldsymbol{\pi}$  as the momentum in the absence of a vector potential Eq. (19)

$$\boldsymbol{\pi} = m\mathbf{v} \quad (122)$$

we conclude that, in Cartesian coordinates, the generalized momentum in the presence of an electromagnetic field is given by

$$\mathbf{p} = \boldsymbol{\pi} + z\mathbf{A} \quad (123)$$

In the absence of an electromagnetic field, the kinetic and generalized momenta are identical in a Cartesian coordinate system.

From the Lagrangian Eq. (119) and the generalized momentum Eq. (123), we obtain the energy function for a particle in an electromagnetic field according to the prescription Eq. (16)

$$h = \mathbf{p} \cdot \mathbf{v} - L = (mv^2 + z\mathbf{v} \cdot \mathbf{A}) - \left( \frac{1}{2}mv^2 - z\phi + z\mathbf{v} \cdot \mathbf{A} \right) \quad (124)$$

which, when expressed in terms of the generalized momentum, yields the Hamiltonian function of Eq. (22):

$$H = T + z\phi = \frac{(\mathbf{p} - z\mathbf{A}) \cdot (\mathbf{p} - z\mathbf{A})}{2m} + z\phi \quad (125)$$

Note that, in the presence of an external vector potential, the Hamiltonian is not given as the sum  $T+U$  of the kinetic energy and the generalized potential but rather as  $H = T+U+z\mathbf{v} \cdot \mathbf{A}$ , as seen from Eqs. (125) and (118). In Exercise 3, it is verified that this Hamiltonian, when inserted into Hamilton's equations Eqs. (24) and (25), leads to Newton's equations for a particle subject to the Lorentz force Eq. (3).

### C. Gauge transformations

Gauge transformations affect not only the potentials but also the Lagrangian, the generalized momenta, and the Hamiltonian. For example, for a particle in an electromagnetic field, the gauge-transformed Lagrangian—that is, the Lagrangian Eq. (119) constructed from the transformed potentials Eqs. (53) and (54)—becomes

$$L' = T - z(\phi' - \mathbf{v} \cdot \mathbf{A}') = L + z \left( \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f \right) \quad (126)$$

which has the same structure as the gauge-transformed Lagrangian in Eq. (5). It is left as an exercise to show that a gauge transformation affects the generalized momenta, the kinetic momenta, and the Hamiltonian as follows:

$$\mathbf{p}' = \mathbf{p} + z \nabla f \quad (127)$$

$$\boldsymbol{\pi}' = \boldsymbol{\pi} \quad (128)$$

$$H' = H - z \frac{\partial f}{\partial t} \quad (129)$$

As shown in Section II B, gauge transformations Eq. (5) do not affect the Lagrangian equations of motion. From Hamilton's equations Eqs. (24) and (25), we obtain

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{\partial H}{\partial \mathbf{p}'} \frac{\partial \mathbf{p}'}{\partial \mathbf{p}} + \frac{\partial H}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial \mathbf{p}} = \frac{\partial H}{\partial \mathbf{p}'} = \frac{\partial H'}{\partial \mathbf{p}'} + z \frac{\partial^2 f}{\partial t \partial \mathbf{p}'} = \frac{\partial H'}{\partial \mathbf{p}'} \quad (130)$$

$$\dot{\mathbf{p}}' = \frac{\partial}{\partial t} \left( \mathbf{p} + z \frac{\partial f}{\partial \mathbf{q}} \right) = - \frac{\partial H}{\partial \mathbf{q}} + z \frac{\partial^2 f}{\partial t \partial \mathbf{q}} = - \frac{\partial}{\partial \mathbf{q}} \left( H - z \frac{\partial f}{\partial t} \right) = - \frac{\partial H'}{\partial \mathbf{q}} \quad (131)$$

where we have used Eqs. (127) and (129), confirming that Hamilton's equations are also invariant to gauge transformations.

## V. RELATIVISTIC MECHANICS

In relativistic mechanics, the equations of motion differ from those of nonrelativistic mechanics, notably for particles moving at high speeds. In Cartesian coordinates, the relativistic equations of motion are given by

$$\mathbf{F} = m \frac{d\gamma \mathbf{v}}{dt} \quad (132)$$

with the *Lorentz factor*

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (133)$$

where  $c$  is the speed of light in vacuum. When the speed of a particle  $v$  approaches zero, the Lorentz factor approaches one and the relativistic equations of motion Eq. (132) reduce to Newton's equations Eq. (1). Conversely, when the speed approaches  $c$ , the particle shows an increased resistance to acceleration until, at  $v = c$ , it can be accelerated no further.

### A. One-particle conservative systems

For a relativistic particle in a conservative force field, the Lagrangian may be taken as

$$L = -mc^2\gamma^{-1} - V \quad (134)$$

Indeed, substituting this function into Lagrange's equations of motion Eq. (13), we recover the relativistic equations of motion Eq. (132):

$$m \frac{d\gamma\mathbf{v}}{dt} = -\nabla V(\mathbf{r}) = \mathbf{F} \quad (135)$$

in the same manner that substitution of the nonrelativistic Lagrangian Eq. (12) into Eq. (13) gives the nonrelativistic equations of motion Eq. (14). Note that the same Lagrange's equations of motion Eqs. (9) and (13) are used in relativistic and nonrelativistic theories, only the Lagrangians differ.

Let us consider the conjugate momentum and the energy function in relativistic theory. The conjugate momentum is obtained from Eqs. (15) and (134) as

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = m\gamma\mathbf{v} \quad (136)$$

and differs from the corresponding nonrelativistic momentum Eq. (19) by the presence of the Lorentz factor. Note that, in terms of the relativistic conjugate momentum, the equations of motion Eq. (132) have the same form as in nonrelativistic theory—namely,  $\mathbf{F} = d\mathbf{p}/dt$ . From the general expression Eq. (16), we obtain the relativistic energy function

$$h = \mathbf{p} \cdot \mathbf{v} - L = mv^2\gamma + mc^2\gamma^{-1} + V = m(v^2 + c^2\gamma^{-2})\gamma + V = mc^2\gamma + V \quad (137)$$

For a conservative relativistic system, therefore, the total energy  $h$  and the kinetic energy  $T$  are given by

$$h = T + V, \quad (138)$$

$$T = mc^2\gamma \quad (139)$$

Expanding the Lorentz factor of  $T$  in powers of  $(v/c)^2$ , we obtain

$$T = mc^2 + \frac{mv^2}{2} + \frac{3mv^4}{8c^2} + \dots \quad (140)$$

which differs from the nonrelativistic energy  $\frac{1}{2}mv^2$  in the contributions from the particle's *rest energy*  $E_0 = mc^2$  and from higher-order terms. As the particle's speed approaches that of light, the kinetic energy becomes infinite, in agreement with our observation that no force can accelerate the particle beyond the speed of light; at low speeds, the relativistic kinetic energy becomes identical to the nonrelativistic energy except for the constant rest energy, which merely represents a shift of the energy scale.

For a Hamiltonian description of a relativistic particle, we must express the energy function Eq. (138) in terms of the conjugate momentum Eq. (136). From the expression for the Lorentz factor Eq. (133), we obtain

$$\gamma^2 = 1 + \frac{v^2\gamma^2}{c^2} = 1 + \frac{p^2}{m^2c^2} \quad (141)$$

where we have used Eq. (136). Taking the positive square root of  $\gamma^2$ , we obtain the following expression for the kinetic energy Eq. (139) in terms of the conjugate momentum:

$$T = mc^2\gamma = \sqrt{m^2c^4 + p^2c^2} \quad (142)$$

For a relativistic particle in a conservative force field, the Hamiltonian obtained from Eq. (138) is therefore given by

$$H = \sqrt{m^2c^4 + p^2c^2} + V \quad (143)$$

The reader may wish to confirm that substitution of this Hamiltonian into Hamilton's equations Eqs. (24) and (25) gives us back the relativistic equations of motion Eq. (132). It is worth noting that the relativistic Hamiltonian of a conservative system is equal to  $T + V$  (as in nonrelativistic theory), even though the Lagrangian Eq. (134) of a relativistic conservative system is not given by  $T - V$ .

## B. Relativistic particle in an electromagnetic field

The treatment of a relativistic particle in an electromagnetic field follows the nonrelativistic treatment closely. For a particle of charge  $z$  in an electromagnetic field represented

by the potentials  $\phi$  and  $\mathbf{A}$ , the relativistic Lagrangian is given by

$$L = -mc^2\gamma^{-1} - z(\phi - \mathbf{v} \cdot \mathbf{A}) \quad (144)$$

by analogy with the nonrelativistic expression Eq. (119). Substitution of this Lagrangian into Lagrange's equations of motion Eq. (13) gives

$$-\left(\frac{d}{dt}\frac{\partial}{\partial\mathbf{v}} - \frac{\partial}{\partial\mathbf{r}}\right)mc^2\gamma^{-1} = \left(\frac{d}{dt}\frac{\partial}{\partial\mathbf{v}} - \frac{\partial}{\partial\mathbf{r}}\right)z(\phi - \mathbf{v} \cdot \mathbf{A}) \quad (145)$$

Whereas the left-hand side becomes  $m\mathbf{d}\gamma\mathbf{v}/dt$  as previously obtained when deriving Eq. (136) from Eq. (134), the right-hand side becomes the Lorentz force Eq. (3) as derived in Exercise 2. Equation (145) therefore becomes

$$m\frac{d\gamma\mathbf{v}}{dt} = z(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (146)$$

which are the equations of motion for a relativistic particle subject to the Lorentz force.

To arrive at the relativistic Hamiltonian for a particle in an electromagnetic field, we first determine the conjugate momentum from Eqs. (15) and (144):

$$\mathbf{p} = \frac{\partial L}{\partial\mathbf{v}} = m\gamma\mathbf{v} + z\mathbf{A} \quad (147)$$

and then the energy function in the same manner as in Eqs. (137) and (138):

$$h = \mathbf{p} \cdot \mathbf{v} - L = m\gamma v^2 + mc^2\gamma^{-1} + z\phi = mc^2\gamma + z\phi \quad (148)$$

Introducing the kinetic momentum

$$\boldsymbol{\pi} = m\gamma\mathbf{v} = \mathbf{p} - z\mathbf{A} \quad (149)$$

and using this expression to eliminate  $\mathbf{v}$  from  $mc^2\gamma$  in the same manner as for the field-free particle in Eqs. (141) and (142), we obtain the following Hamiltonian, valid for a relativistic particle in an electromagnetic field:

$$H = \sqrt{m^2c^4 + \boldsymbol{\pi}^2c^2} + z\phi \quad (150)$$

As in nonrelativistic theory, this Hamiltonian differs from the Hamiltonian of a conservative system Eq. (143) only in the appearance of  $\boldsymbol{\pi} = \mathbf{p} - z\mathbf{A}$  rather than  $\mathbf{p}$  in the kinetic energy. Expanding the Hamiltonian in powers of  $c^{-2}$ , we obtain

$$H = mc^2 + \frac{\boldsymbol{\pi}^2}{2m} - \frac{\boldsymbol{\pi}^4}{8m^3c^2} + z\phi + \mathcal{O}\left(\frac{\boldsymbol{\pi}^6}{m^5c^4}\right) \quad (151)$$

where, in addition to the rest mass and the higher-order terms, we recognize the Hamiltonian of a nonrelativistic particle in an electromagnetic field Eq. (125).

### C. Two-particle systems

Let us consider a system of two particles, with masses  $m_1$  and  $m_2$ , charges  $z_1$  and  $z_2$ , and velocities  $\mathbf{v}_1$  and  $\mathbf{v}_2$ , respectively. Each particle moves according to the laws of relativistic mechanics, with Lagrangians

$$L_1 = -m_1 c^2 \gamma_1^{-1} - z_1 (\phi_1 - \mathbf{v}_1 \cdot \mathbf{A}_1) - z_1 [\phi_p(z_2, \mathbf{v}_2, \mathbf{r}_{21}) - \mathbf{v}_1 \cdot \mathbf{A}_p(z_2, \mathbf{v}_2, \mathbf{r}_{21})] \quad (152)$$

$$L_2 = -m_2 c^2 \gamma_2^{-1} - z_2 (\phi_2 - \mathbf{v}_2 \cdot \mathbf{A}_2) - z_2 [\phi_p(z_1, \mathbf{v}_1, \mathbf{r}_{12}) - \mathbf{v}_2 \cdot \mathbf{A}_p(z_1, \mathbf{v}_1, \mathbf{r}_{12})] \quad (153)$$

where  $\gamma_i = (1 + v_i^2/c^2)^{-1/2}$  is the Lorentz factor of particle  $i$ , whereas  $\phi_i$  and  $\mathbf{A}_i$  are the external potentials evaluated at the position of particle  $i$ . The interaction between the two particles in Eqs. (152) and (153) is described in terms of the potentials Eqs. (104) and (105) generated by a point-charge particle of charge  $z$  and velocity  $\mathbf{v}$ , measured at position  $\mathbf{r}$  relative to the particle:

$$\phi_p(z, \mathbf{v}, \mathbf{r}) = \frac{z}{4\pi\epsilon_0 r} + \mathcal{O}(v^4/c^4) \quad (154)$$

$$\mathbf{A}_p(z, \mathbf{v}, \mathbf{r}) = \frac{z}{8\pi\epsilon_0 c^2 r^3} [r^2 \mathbf{v} + (\mathbf{v} \cdot \mathbf{r}) \mathbf{r}] + \mathcal{O}(v^3/c^4) \quad (155)$$

Since the potentials of a moving particle cannot be written down in exact form, we will have to be content with approximate expressions. The expressions Eqs. (154) and (155) with the  $\mathcal{O}(v^4/c^4)$  and  $\mathcal{O}(v^3/c^4)$  terms omitted are adequate for our purposes—that is, for the description of molecular electronic systems. However, for a proper order analysis, the  $\mathcal{O}$  terms will nevertheless be retained in the following.

For a simultaneous treatment of both particles, consider the two-particle function

$$L = - \sum_{i=1}^2 [m_i c^2 \gamma_i^{-1} + z_i (\phi_i - \mathbf{v}_i \cdot \mathbf{A}_i)] - C_{12} - B_{12} \quad (156)$$

where we have included the one-particle parts of  $L_1$  and  $L_2$  of Eqs. (152) and (153) and added the interaction terms

$$C_{12} = \frac{z_1 z_2}{4\pi\epsilon_0 r_{12}} [1 + \mathcal{O}(v^4/c^4)] \quad (157)$$

$$B_{12} = - \frac{z_1 z_2}{8\pi\epsilon_0 c^2} \frac{(\mathbf{v}_1 \cdot r_{12}^2 \mathbf{v}_2) + (\mathbf{v}_1 \cdot \mathbf{r}_{12})(\mathbf{r}_{12} \cdot \mathbf{v}_2)}{r_{12}^3} [1 + \mathcal{O}(v^2/c^2)] \quad (158)$$

Since the interaction  $-C_{12} - B_{12}$  in Eq. (156) is the same as in  $L_1$  and  $L_2$ , the two-particle function  $L$  gives, for each particle, the same equations of motion as do  $L_1$  and  $L_2$  separately. We shall therefore take Eq. (156) as the Lagrangian of the combined system.

To set up the Hamiltonian for the combined system, we begin in the usual manner by determining their conjugate momenta and their energy function. Differentiating the Lagrangian Eq. (156) with respect to  $\mathbf{v}_i$ , we obtain

$$\mathbf{p}_i = m_i \gamma_i \mathbf{v}_i + z_i \mathbf{A}_i - \frac{\partial B_{12}}{\partial \mathbf{v}_i} \quad (159)$$

which differs from the conjugate momentum of a one-particle system Eq. (147) in the presence of the interaction term. Next, evaluating the energy function according to the prescription Eq. (16), we obtain

$$h = \sum_i (m_i c^2 \gamma_i + z_i \phi_i) + C_{12} - B_{12} \quad (160)$$

in the same manner as for a one-particle system Eq. (148), noting that  $B_{12} = \mathbf{v}_i \cdot (\partial B_{12} / \partial \mathbf{v}_i)$ . We now introduce the kinetic momentum as

$$\boldsymbol{\pi}_i = \mathbf{p}_i - z_i \mathbf{A}_i = m_i \gamma_i \mathbf{v}_i - \frac{\partial B_{12}}{\partial \mathbf{v}_i} \quad (161)$$

and obtain for the squared Lorentz factor [see Eq. (141)]

$$\gamma_i^2 = 1 + \frac{v_i^2 \gamma_i^2}{c^2} = 1 + \frac{\pi_i^2}{m_i^2 c^2} + \frac{2\boldsymbol{\pi}_i}{m_i^2 c^2} \cdot \frac{\partial B_{12}}{\partial \mathbf{v}_i} \quad (162)$$

where we have neglected the term quadratic in  $B_{12}$ , in accordance with the error term in Eq. (158). Taking the positive square root of this expression and expanding to lowest order in  $B_{12}$ , we find that the kinetic energy [see Eq. (139)] may be written in the form

$$m_i c^2 \gamma_i = \sqrt{m_i^2 c^4 + \pi_i^2 c^2} + \frac{\boldsymbol{\pi}_i}{m_i} \cdot \frac{\partial B_{12}}{\partial \mathbf{v}_i} \quad (163)$$

where we have omitted terms of order  $\mathcal{O}(v^4/c^4)$ . From Eq. (161), we note that the last term in this expression may be simplified as

$$\frac{\boldsymbol{\pi}_i}{m_i} \cdot \frac{\partial B_{12}}{\partial \mathbf{v}_i} = \gamma_i \mathbf{v}_i \cdot \frac{\partial B_{12}}{\partial \mathbf{v}_i} - \frac{1}{m_i} \left\| \frac{\partial B_{12}}{\partial \mathbf{v}_i} \right\|^2 = B_{12} \quad (164)$$

where we have retained the leading term in agreement with Eq. (158). Substituting first Eq. (164) into Eq. (163) and then Eq. (163) into Eq. (160), we arrive at the two-particle energy function:

$$h = \sum_{i=1}^2 \sqrt{m_i^2 c^4 + \pi_i^2 c^2} + \sum_{i=1}^2 z_i \phi_i + C_{12} + B_{12} \quad (165)$$

To obtain the Hamiltonian, it only remains to express  $\mathbf{v}_i$  in terms of  $\boldsymbol{\pi}_i$  in  $B_{12}$ . From Eq. (161), we note that

$$\mathbf{v}_i = \frac{\boldsymbol{\pi}_i}{m_i} + \mathcal{O}(v/c^2) \quad (166)$$

We may therefore replace  $\mathbf{v}_i$  by  $\boldsymbol{\pi}_i/m_i$  in  $B_{12}$  of Eq. (158) without introducing errors larger than those already present in the error term. Inserting the resulting expression for  $B_{12}$  into Eq. (165), we obtain the two-particle Hamiltonian function

$$H = \sum_{i=1}^2 \sqrt{m_i^2 c^4 + \boldsymbol{\pi}_i^2 c^2} + \sum_{i=1}^2 z_i \phi_i + \frac{z_1 z_2}{4\pi\epsilon_0} \left[ \frac{1}{r_{12}} - \frac{(\boldsymbol{\pi}_1 \cdot r_{12}^2 \boldsymbol{\pi}_2) + (\boldsymbol{\pi}_1 \cdot \mathbf{r}_{12})(\mathbf{r}_{12} \cdot \boldsymbol{\pi}_2)}{2m_1 m_2 c^2 r_{12}^3} + \mathcal{O}\left(\frac{p^4}{m^4 c^4}\right) \right] \quad (167)$$

which is correct to all orders in the one-particle part and to order  $p^4/m^4 c^4$  in the two-particle part. This Hamiltonian may be generalized to an arbitrary number of particles by assuming only pairwise interactions. In the following, we shall therefore take our *relativistic many-particle Hamiltonian function* to be of the form

$$H = \sum_i \sqrt{m_i^2 c^4 + \boldsymbol{\pi}_i^2 c^2} + \sum_i z_i \phi_i + \frac{1}{4\pi\epsilon_0} \sum_{i>j} \frac{z_i z_j}{r_{ij}} - \frac{1}{8\pi\epsilon_0 c^2} \sum_{i>j} \frac{z_i z_j}{m_i m_j} \frac{(\boldsymbol{\pi}_i \cdot r_{ij}^2 \boldsymbol{\pi}_j) + (\boldsymbol{\pi}_i \cdot \mathbf{r}_{ij})(\mathbf{r}_{ij} \cdot \boldsymbol{\pi}_j)}{r_{ij}^3} \quad (168)$$

keeping in mind the approximations made in the two-particle terms. Subtracting the rest mass term  $\sum_i m_i c^2$  and letting  $c$  go to infinity, we obtain the *nonrelativistic many-particle Hamiltonian function*

$$H^{\text{NR}} = \sum_i \frac{\boldsymbol{\pi}_i^2}{2m_i} + \sum_i z_i \phi_i + \frac{1}{4\pi\epsilon_0} \sum_{i>j} \frac{z_i z_j}{r_{ij}} \quad (169)$$

In Sections VI–IX, we shall develop the quantum-mechanical versions of these Hamiltonians, for many-electron systems.

## VI. THE ONE-ELECTRON NONRELATIVISTIC HAMILTONIAN

### A. Quantization

In classical Hamiltonian mechanics, a physical system is described by values of the generalized coordinates  $q_i$  and their conjugate momenta  $p_i$ , with one such pair for each degree



of freedom. The generalized coordinates and conjugate momenta may be obtained as functions of time by solving Hamilton's equations of motion Eqs. (24) and (25). All observables  $\Omega(\mathbf{q}, \mathbf{p})$  of the system can be expressed in terms of these coordinates and momenta.

In quantum mechanics, the generalized coordinates and momenta  $q_i$  and  $p_i$  are promoted to Hermitian operators  $\hat{Q}_i$  and  $\hat{P}_i$ , whose commutators must be equal to the elementary Poisson brackets Eq. (35) multiplied by  $i\hbar$ :

$$[\hat{Q}_i, \hat{Q}_j] = 0, \quad [\hat{P}_i, \hat{P}_j] = 0, \quad [\hat{Q}_i, \hat{P}_j] = i\hbar\delta_{ij} \quad (170)$$

In the *coordinate representation*, this is accomplished by the identifications (in Cartesian coordinates)

$$\hat{Q}_i = \hat{q}_i \rightarrow q_i \quad (171)$$

$$\hat{P}_i = \hat{p}_i \rightarrow -i\hbar \frac{\partial}{\partial q_i} \quad (172)$$

Thus, in this representation, the coordinate operators work by multiplication and the momentum operators by differentiation. Other observables  $\Omega$  are represented by Hermitian operators  $\hat{\Omega}(\hat{\mathbf{q}}, \hat{\mathbf{p}})$  obtained from the corresponding classical observables  $\Omega(\mathbf{q}, \mathbf{p})$  by replacing  $q_i$  and  $p_i$  by their quantum-mechanical counterparts  $\hat{q}_i$  and  $\hat{p}_i$ . In particular, the quantum-mechanical Hamiltonian operator  $\hat{H}(\hat{\mathbf{q}}, \hat{\mathbf{p}})$  is obtained from the corresponding classical Hamiltonian  $H(\mathbf{q}, \mathbf{p})$  by the substitutions Eqs. (171) and (172).

In the coordinate representation, the state of a quantum-mechanical system is represented by a wave function  $\Psi(\mathbf{q}, t)$ , a time-dependent function of the coordinates  $q_i$  that satisfies the equation of motion

$$i\hbar \frac{\partial \Psi(\mathbf{q}, t)}{\partial t} = \hat{H}\Psi(\mathbf{q}, t) \quad (173)$$

Clearly, our description of the quantum-mechanical system then depends on the particular Hamiltonian function  $H(\mathbf{q}, \mathbf{p})$  on which we base the construction of the Hamiltonian operator  $\hat{H}(\hat{\mathbf{q}}, \hat{\mathbf{p}})$ . For a many-electron system, for example, we may base our description either on the relativistic Hamiltonian function Eq. (168) or on the simpler nonrelativistic Hamiltonian function Eq. (169), with or without the inclusion of external electromagnetic potentials.

From Eq. (173), we find that the quantum-mechanical expectation value of an operator

may be calculated as

$$\begin{aligned} i\hbar \frac{d}{dt} \langle \Psi(\mathbf{q}, t) | \hat{\Omega}(\hat{\mathbf{q}}, \hat{\mathbf{p}}) | \Psi(\mathbf{q}, t) \rangle &= \langle \Psi(\mathbf{q}, t) | \left[ \hat{\Omega}(\hat{\mathbf{q}}, \hat{\mathbf{p}}), \hat{H} \right] | \Psi(\mathbf{q}, t) \rangle \\ &+ i\hbar \langle \Psi(\mathbf{q}, t) | \frac{\partial \hat{\Omega}(\hat{\mathbf{q}}, \hat{\mathbf{p}})}{\partial t} | \Psi(\mathbf{q}, t) \rangle \end{aligned} \quad (174)$$

The quantum-mechanical expression for the time development is thus obtained from the classical expression Eq. (34) by taking the expectation value of both sides of the equation and replacing the Poisson brackets with the corresponding commutator divided by  $i\hbar$  as indicated by Eq. (35). Equation (174) is often referred to as the Ehrenfest theorem. It shows that the expectation value of a time-independent operator  $\hat{\Omega}$  is a constant of motion if it commutes with  $\hat{H}$ .

To simplify notation, we shall in the following omit hats from the quantum-mechanical operators, the distinction between classical variables and quantum-mechanical operators always being clear from the context.

## B. The one-component one-electron Hamiltonian

From the classical Hamiltonian function Eq. (125), we obtain the following nonrelativistic Hamiltonian for an electron of mass  $m_e$  and charge  $-e$  in an electromagnetic field represented by the potentials  $\phi$  and  $\mathbf{A}$ :

$$H_{1c} = \frac{\boldsymbol{\pi}^2}{2m_e} - e\phi \quad (175)$$

where the operator for the kinetic momentum  $\boldsymbol{\pi}$  has been obtained from the corresponding classical expression Eq. (123) in accordance with the general prescription for quantization

$$\boldsymbol{\pi} = -i\hbar \boldsymbol{\nabla} + e\mathbf{A} \quad (176)$$

Applying  $\boldsymbol{\pi}^2$  to an arbitrary function  $\Psi$ , we obtain

$$\begin{aligned} \boldsymbol{\pi}^2 \Psi &= p^2 \Psi + e\mathbf{p} \cdot \mathbf{A} \Psi + e\mathbf{A} \cdot \mathbf{p} \Psi + e^2 A^2 \Psi \\ &= p^2 \Psi + e(\mathbf{p} \cdot \mathbf{A}) \Psi + 2e\mathbf{A} \cdot \mathbf{p} \Psi + e^2 A^2 \Psi \\ &= (p^2 + 2e\mathbf{A} \cdot \mathbf{p} + e^2 A^2) \Psi \end{aligned} \quad (177)$$

where, in the last step, we have assumed the Coulomb gauge  $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$ . In this gauge, therefore, the nonrelativistic Schrödinger equation Eq. (173) takes the form

$$i\hbar \frac{\partial \Psi_{1c}(\mathbf{r}, t)}{\partial t} = H_{1c} \Psi_{1c}(\mathbf{r}, t) \quad (178)$$

with the electronic Hamiltonian

$$H_{1c} = \frac{1}{2m_e} p^2 + \frac{e}{m_e} \mathbf{A} \cdot \mathbf{p} + \frac{e^2}{2m_e} A^2 - e\phi \quad (179)$$

where the first three terms together represent the kinetic-energy operator and the last term is the potential-energy operator. The term proportional to  $\mathbf{A} \cdot \mathbf{p}$ , which represents the interaction of the external field with the orbital motion of the electron, is known as the *orbital paramagnetic operator*. The corresponding *diamagnetic operator*  $e^2 A^2/2m_e$  may be viewed as a small positive correction to the interaction  $-e\phi$  of the electron with the scalar potential.

The Hamiltonian  $H_{1c}$  contains the same interaction terms as does the classical Hamiltonian function. Nevertheless, it fails when an external magnetic field is applied. The reason for this failure is simple: In the construction of the Hamiltonian, a fundamental property of the electron—namely, its spin—was not considered. Indeed, since spin has no classical counterpart for a point-charge particle such as the electron, it cannot be incorporated in the Hamiltonian by following the standard procedure for quantization of the classical Hamiltonian. In Section VI C, we shall see how spin can nevertheless be included in our description, by a small modification of the Hamiltonian.

### C. The two-component one-electron Hamiltonian

For an electron in an external magnetic field, the correct nonrelativistic Hamiltonian is obtained by a small modification to the Hamiltonian Eq. (175):

$$H_{2c} = \frac{(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2}{2m_e} - e\phi \quad (180)$$

where the vector  $\boldsymbol{\pi}$  contains the usual three kinetic-momentum operators Eq. (176), while the vector  $\boldsymbol{\sigma}$  contains three new (as yet unspecified) Hermitian operators  $\sigma_i$  that are taken to commute with  $\boldsymbol{\pi}$ . In the following, we shall first determine the form of these new operators  $\sigma_i$  from the requirement that the new Hamiltonian Eq. (180) reduces to the old one Eq. (175) in the absence of an electromagnetic field. Next, we shall examine the consequence of these operators for our description of an electron in an electromagnetic field.

Since the Hamiltonian Eq. (175) is correct in the absence of an electromagnetic field, Eq. (180) must reduce to Eq. (175) when the vector potential in Eq. (176) vanishes:

$$(\boldsymbol{\sigma} \cdot \mathbf{p})^2 = p^2 \quad (181)$$

or equivalently in terms of components

$$\sum_i \sigma_i^2 p_i^2 + \sum_{i>j} (\sigma_i \sigma_j + \sigma_j \sigma_i) p_i p_j = \sum_i p_i^2 \quad (182)$$

We therefore conclude that  $\sigma_i^2 = 1$  and  $\sigma_i \sigma_j + \sigma_j \sigma_i = 0$ , which implies that the new operators anticommute among one another:

$$[\sigma_i, \sigma_j]_+ = 2\delta_{ij} \quad (183)$$

The  $\sigma_i$  operators therefore cannot be pure numbers but must be represented by matrices of dimension greater than one. These matrices must be traceless since

$$\text{tr } \sigma_i = \text{tr } \sigma_i \sigma_k^2 = -\text{tr } \sigma_k \sigma_i \sigma_k = -\text{tr } \sigma_i \sigma_k^2 = -\text{tr } \sigma_i \quad (184)$$

where we have used Eq. (183) (with  $\sigma_k \neq \sigma_i$ ) and  $\text{tr } \mathbf{AB} = \text{tr } \mathbf{BA}$ . Since  $\sigma_i^2 = 1$ , their eigenvalues are  $\pm 1$  and we conclude that  $\sigma_i$  are Hermitian matrices of an even dimension, with the same number of positive and negative eigenvalues. The simplest such matrices are two-dimensional matrices of the general form:

$$\sigma_i = \begin{pmatrix} z_i & x_i - iy_i \\ x_i + iy_i & -z_i \end{pmatrix} \quad (185)$$

where  $(x_i, y_i, z_i)$  are real numbers. In this representation, we may write the anticommutation relation Eq. (183) in the form

$$[\sigma_i, \sigma_j]_+ = 2x_i x_j + 2y_i y_j + 2z_i z_j \quad (186)$$

The matrices Eq. (185) may therefore be thought of as representing three-dimensional vectors  $\mathbf{v}_i^T = (x_i, y_i, z_i)$  that must be orthonormal  $\mathbf{v}_i^T \mathbf{v}_j = \delta_{ij}$  for Eq. (183) to be satisfied. Clearly, there are infinitely many possible two-dimensional representations of  $\sigma_i$ , just like there are infinitely many orthonormal basis vectors in three-dimensional space. In Exercise 4, it is shown that these matrices and hence the operators  $\sigma_i$  also satisfy the commutation relations

$$[\sigma_i, \sigma_j] = 2c_i \epsilon_{ijk} \sigma_k \quad (187)$$

where  $c = \pm 1$  and where the *Levi-Civita antisymmetric symbol*  $\epsilon_{ijk}$  is  $+1$  for even permutations of the indices,  $-1$  for odd permutations, and  $0$  whenever two indices are identical. The substitution  $\boldsymbol{\sigma} = -\boldsymbol{\sigma}'$  shows that the  $\sigma'_i$  operators satisfy the same commutation relations

as  $\sigma_i$  but with opposite sign. The representations therefore occur in pairs  $\pm\boldsymbol{\sigma}$  with  $c = \pm 1$ . As seen from Eq. (180), the sign of  $\boldsymbol{\sigma}$  does not affect the Hamiltonian. Without loss of generality, we may therefore require that  $\boldsymbol{\sigma}$  is chosen such that  $c = 1$ .

Among the infinitely many valid representations of the operators  $\sigma_i$ , we shall henceforth use the standard representation provided by the three *Pauli spin matrices*:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (188)$$

which correspond to the choice  $\mathbf{v}_i = \mathbf{e}_i$  in Eq. (185) and satisfy the relations

$$[\sigma_i, \sigma_j]_+ = 2\delta_{ij} \quad (189)$$

$$[\sigma_i, \sigma_j]_- = 2i\epsilon_{ijk}\sigma_k \quad (190)$$

as may be explicitly verified by direct substitution of Eq. (188). Finally, by adding together the commutator and anticommutator in Eqs. (189) and (190), we obtain

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k \quad (191)$$

which may be taken as the fundamental property of the  $\sigma_i$  operators, noting that the commutator and anticommutator follow directly from it by substitution.

We have now determined the algebra Eq. (191) of the  $\sigma_i$  operators in the modified Hamiltonian Eq. (180) from the requirement that this Hamiltonian reduces to the original Hamiltonian Eq. (175) in the absence of an electromagnetic field. Moreover, we have found that the  $\sigma_i$  operators are represented by traceless two-by-two Hermitian matrices with eigenvalues  $\pm 1$ . The Hamiltonian  $H_{2c}$  thus becomes a *two-component Hamiltonian*, in contrast to the *one-component Hamiltonian* of  $H_{1c}$  in Eq. (175). The two-component form of the Hamiltonian Eq. (180) introduces a new degree of freedom in our description of the electron—namely, its *spin*. Spin is therefore not a relativistic property of the electron but rather a property that can be satisfactorily treated in nonrelativistic theory. In Section VII, we shall see that the two-component Hamiltonian Eq. (180) arises as the natural nonrelativistic limit of the relativistic Dirac equation.

Let us now see how spin arises from the kinetic-energy part  $(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2/2m_e$  of the Hamiltonian Eq. (180). Since the  $\sigma_i$  operators by assumption commute with the  $\pi_i$  operators, we obtain

$$(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2 = \left( \sum_i \sigma_i \pi_i \right) \left( \sum_j \sigma_j \pi_j \right) = \sum_{ij} \sigma_i \sigma_j \pi_i \pi_j = \sum_i \pi_i^2 + i \sum_{ijk} \epsilon_{ijk} \pi_i \pi_j \sigma_k \quad (192)$$

where we have used Eq. (191). This result may be written compactly as

$$(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2 = \pi^2 + i \boldsymbol{\sigma} \cdot \boldsymbol{\pi} \times \boldsymbol{\pi} \quad (193)$$

Applying  $\boldsymbol{\pi} \times \boldsymbol{\pi}$  to an arbitrary wave function and using the identity  $\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{B}$ , we find

$$\begin{aligned} (\boldsymbol{\pi} \times \boldsymbol{\pi}) \Psi &= e \mathbf{p} \times \mathbf{A} \Psi + e \mathbf{A} \times \mathbf{p} \Psi \\ &= e (\mathbf{p} \times \mathbf{A}) \Psi + e (\mathbf{p} \Psi) \times \mathbf{A} + e \mathbf{A} \times \mathbf{p} \Psi \\ &= -ie\hbar (\boldsymbol{\nabla} \times \mathbf{A}) \Psi = -ie\hbar \mathbf{B} \Psi \end{aligned} \quad (194)$$

which upon substitution in Eq. (193) and division by  $2m_e$  yields the following expression for the kinetic-energy operator

$$\frac{(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2}{2m_e} = \frac{\pi^2}{2m_e} + \frac{e\hbar}{2m_e} \mathbf{B} \cdot \boldsymbol{\sigma} \quad (195)$$

We now introduce the *spin operators*  $s_i$  by a simple scaling of the  $\sigma_i$  operators

$$\mathbf{s} = \frac{\hbar}{2} \boldsymbol{\sigma} \quad (196)$$

Substitution of  $\sigma_i = 2s_i/\hbar$  in Eq. (190) shows that the components of  $\mathbf{s}$  satisfy the usual commutation relations of angular momentum

$$[s_i, s_j] = i\hbar \epsilon_{ijk} s_k \quad (197)$$

or equivalently  $\mathbf{s} \times \mathbf{s} = i\hbar \mathbf{s}$ . Since  $\sigma_i$  has eigenvalues  $\pm 1$ , the eigenvalues of  $s_i$  are  $\pm \frac{1}{2}\hbar$ , representing the two components of a state of total angular-momentum quantum number  $1/2$ , with a total angular momentum  $\sqrt{\frac{3}{4}}\hbar$ .

In the Coulomb gauge, we may expand  $\pi^2/2m_e$  in Eq. (180) according to Eq. (177), yielding the following expression for the two-component Hamiltonian:

$$H_{2c} = \frac{1}{2m_e} p^2 + \frac{e}{m_e} \mathbf{A} \cdot \mathbf{p} + \frac{e}{m_e} \mathbf{B} \cdot \mathbf{s} + \frac{e^2}{2m_e} A^2 - e\phi \quad (198)$$

where the *orbital paramagnetic operator*  $(e/m_e)\mathbf{A} \cdot \mathbf{p}$  and the *spin paramagnetic operator*  $(e/m_e)\mathbf{B} \cdot \mathbf{s}$  represent the interactions of the external magnetic field with the orbital and spin angular momenta of the electron, respectively. The spin interaction is interpreted by associating a *magnetic moment* with the spin of the electron:

$$\mathbf{m} = -\frac{g\mu_B}{\hbar} \mathbf{s}, \quad g = 2, \quad \mu_B = \frac{e\hbar}{2m_e}, \quad (199)$$

where we have introduced the *electron g factor* and the *Bohr magneton*  $\mu_B$ . The spin paramagnetic term may now be written as a *Zeeman interaction*  $-\mathbf{B} \cdot \mathbf{m}$  of the external induction with the magnetic moment of the electron. Thus, the two-component Hamiltonian Eq. (180) differs from the one-component Hamiltonian (175) only in the presence of the Zeeman term:

$$H_{2c} = H_{1c} - \mathbf{B} \cdot \mathbf{m} \quad (200)$$

In the absence of an external magnetic field, the one- and two-component nonrelativistic Hamiltonians are identical and the two components of the wave function are not coupled.

#### D. Two-component wave functions and spin variables

The Pauli spin representation Eq. (188) has been chosen such that the basis vectors are eigenvectors of  $\mathbf{s}^2 = \mathbf{s}_x^2 + \mathbf{s}_y^2 + \mathbf{s}_z^2$  and of  $\mathbf{s}_z$  in Eq. (196):

$$\mathbf{s}^2 \boldsymbol{\alpha} = \frac{3\hbar^2}{4} \boldsymbol{\alpha}, \quad \mathbf{s}_z \boldsymbol{\alpha} = \frac{\hbar}{2} \boldsymbol{\alpha} \quad (201)$$

$$\mathbf{s}^2 \boldsymbol{\beta} = \frac{3\hbar^2}{4} \boldsymbol{\beta}, \quad \mathbf{s}_z \boldsymbol{\beta} = -\frac{\hbar}{2} \boldsymbol{\beta} \quad (202)$$

where we have introduced the  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$  unit vectors

$$\boldsymbol{\alpha} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (203)$$

This choice is consistent with considering the electron as a particle with spin 1/2 and spin quantization along the  $z$  axis. The  $\boldsymbol{\alpha}$  state has spin projection  $\hbar/2$ , whereas the  $\boldsymbol{\beta}$  state has spin projection  $-\hbar/2$ . The wave function of an electron is a function in real space and two-dimensional spin space. Using the Pauli representation to describe spin, the wave function becomes

$$\boldsymbol{\Psi}_{2c}(\mathbf{r}, t) = \begin{pmatrix} \Psi_{\alpha}(\mathbf{r}, t) \\ \Psi_{\beta}(\mathbf{r}, t) \end{pmatrix} = \Psi_{\alpha}(\mathbf{r}, t) \boldsymbol{\alpha} + \Psi_{\beta}(\mathbf{r}, t) \boldsymbol{\beta} \quad (204)$$

In this representation, the Hamiltonian Eq. (200) becomes a two-by-two Hermitian matrix, where each element contains an operator:

$$i\hbar \frac{\partial \boldsymbol{\Psi}_{2c}(\mathbf{r}, t)}{\partial t} = \mathbf{H}_{2c} \boldsymbol{\Psi}_{2c}(\mathbf{r}, t) \quad (205)$$

In the Pauli representation Eq. (188) of the two-component Hamiltonian Eq. (180), we may then write the time-dependent Schrödinger equation in the explicit form

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \Psi_\alpha(\mathbf{r}) \\ \Psi_\beta(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} \frac{\pi^2}{2m_e} - e\phi + \mu_B B_z & \mu_B (B_x - iB_y) \\ \mu_B (B_x + iB_y) & \frac{\pi^2}{2m_e} - e\phi - \mu_B B_z \end{pmatrix} \begin{pmatrix} \Psi_\alpha(\mathbf{r}) \\ \Psi_\beta(\mathbf{r}) \end{pmatrix} \quad (206)$$

where we have used the fact that  $H_{1c}$  is spin-independent and therefore appears on the diagonal of  $\mathbf{H}_{2c}$ . Equation (206) clearly shows how the two components of the wave function become coupled in the presence of an external magnetic field with a nonzero component perpendicular to the  $z$  axis.

Alternatively, we may avoid working with two components by introducing a new variable in the wave function. The two-component wave function Eq. (204) then becomes as a linear combination of alpha and beta components

$$\Psi_{2c}(\mathbf{x}, t) = \Psi_\alpha(\mathbf{r}, t)\alpha(m_s) + \Psi_\beta(\mathbf{r}, t)\beta(m_s) \quad (207)$$

where  $\mathbf{x}$  is a composite set of spatial coordinates  $\mathbf{r}$  and spin coordinate  $m_s$ , which may take on two discrete values  $m_s = -1/2$  and  $m_s = 1/2$ . The spin space is spanned by the two functions  $\alpha(m_s)$  and  $\beta(m_s)$ , whose functional form is given by the equations

$$\alpha\left(\frac{1}{2}\right) = 1, \quad \alpha\left(-\frac{1}{2}\right) = 0 \quad (208)$$

$$\beta\left(\frac{1}{2}\right) = 0, \quad \beta\left(-\frac{1}{2}\right) = 1 \quad (209)$$

Spin integration is simple sum over the two spin variables

$$\int f(m_s) dm_s = f\left(\frac{1}{2}\right) + f\left(-\frac{1}{2}\right) \quad (210)$$

and we find that the  $\alpha$  and  $\beta$  states are orthogonal

$$\int \sigma^*(m_s)\tau(m_s) dm_s = \delta_{\sigma\tau} \quad (211)$$

where we use the generic notation  $\sigma$  and  $\tau$  for the spin functions  $\alpha$  and  $\beta$ . The effect of the Pauli spin matrices Eq. (188) is given by

$$\sigma_x \alpha(m_s) = \beta(m_s), \quad \sigma_y \alpha(m_s) = -i\beta(m_s), \quad \sigma_z \alpha(m_s) = \alpha(m_s) \quad (212)$$

$$\sigma_x \beta(m_s) = \alpha(m_s), \quad \sigma_y \beta(m_s) = i\alpha(m_s), \quad \sigma_z \beta(m_s) = -\beta(m_s) \quad (213)$$



The two-component Schrödinger equation Eq. (205) may now be expressed as

$$i\hbar \frac{\partial \Psi_{2c}(\mathbf{x}, t)}{\partial t} = H_{2c} \Psi_{2c}(\mathbf{x}, t) \quad (214)$$

which, apart from the introduction of the spin variable  $m_s$ , resembles the one-component Schrödinger equation Eq. (175). The equations defining the spin functions have previously been presented in Chapter 2.

### E. Gauge transformation of the Schrödinger equation

In Section IV C, we demonstrated that the classical equations of motion are invariant to gauge transformations. In quantum mechanics, all calculated observables should likewise be invariant to such transformations. In this section, we show how a gauge transformation changes the Hamiltonian and wave function in a manner that leaves all observable properties unaffected, justifying our use of the Coulomb gauge in our construction of the Hamiltonian operator.

Let us write the time-dependent one-component Schrödinger equation of a single particle Eq. (178) in the form

$$\left( H - i\hbar \frac{\partial}{\partial t} \right) \Psi = 0 \quad (215)$$

where, in the presence of the scalar and vector potentials  $\phi$  and  $\mathbf{A}$ , the Hamiltonian and momentum operators are given by [see Eqs. (175) and (176)]:

$$H = \frac{\pi^2}{2m} + z\phi \quad (216)$$

$$\boldsymbol{\pi} = -i\hbar \boldsymbol{\nabla} - z\mathbf{A} \quad (217)$$

where  $m$  is the mass of the particle and  $z$  its charge. Clearly, a gauge transformation of the scalar and vector potentials [see Eqs. (53) and (54)]

$$\phi' = \phi - \frac{\partial f}{\partial t} \quad (218)$$

$$\mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla} f \quad (219)$$

leads to a different Hamiltonian with a different kinetic-momentum operator

$$H' = \frac{\pi'^2}{2m} + z\phi' \quad (220)$$

$$\boldsymbol{\pi}' = -i\hbar \boldsymbol{\nabla} - z\mathbf{A}' \quad (221)$$

As a result of the changes in the Hamiltonian, the wave function also changes:

$$\left(H' - i\hbar\frac{\partial}{\partial t}\right)\Psi' = 0 \quad (222)$$

The gauge-transformed potentials represent the same physical situation (i.e., the same electromagnetic field) as do the original potentials. Consequently, the transformed wave function  $\Psi'$  represents the same state as the original wave function  $\Psi$ .

To establish the relationship between  $\Psi'$  and  $\Psi$ , we first note that  $\boldsymbol{\pi}'$  and  $\boldsymbol{\pi}$  are related by a unitary transformation:

$$\begin{aligned} \boldsymbol{\pi}' &= -i\hbar\nabla - z\mathbf{A} - z\nabla f = \exp(izf/\hbar)(-i\hbar\nabla - z\mathbf{A})\exp(-izf/\hbar) \\ &= \exp(izf/\hbar)\boldsymbol{\pi}\exp(-izf/\hbar) \end{aligned} \quad (223)$$

where application of  $-i\hbar\nabla$  to the exponential introduces the term  $-z\nabla f$ . Furthermore, we have

$$\left(z\phi' - i\hbar\frac{\partial}{\partial t}\right) = \exp(izf/\hbar)\left(z\phi - i\hbar\frac{\partial}{\partial t}\right)\exp(-izf/\hbar) \quad (224)$$

from which it follows that

$$\left(H' - i\hbar\frac{\partial}{\partial t}\right) = \exp(izf/\hbar)\left(H - i\hbar\frac{\partial}{\partial t}\right)\exp(-izf/\hbar) \quad (225)$$

Comparing Eqs. (215) and (222) and using this result, we find that, if  $\Psi$  is a solution to the untransformed Schrödinger equation Eq. (215), then

$$\Psi' = \exp(izf/\hbar)\Psi \quad (226)$$

represents a solution to the gauge-transformed Schrödinger equation Eq. (222). It likewise follows that the expectation of values of the differential operators  $\boldsymbol{\pi}$  and  $i\hbar\partial/\partial t - z\phi$  as well as of any multiplicative operator  $g$  that depends on position of the electron are invariant to the simultaneous transformation of operator and wave function:

$$\langle\Psi'|\boldsymbol{\pi}'|\Psi'\rangle = \langle\Psi|\boldsymbol{\pi}|\Psi\rangle \quad (227)$$

$$\langle\Psi'|i\hbar\frac{\partial}{\partial t} - z\phi'|\Psi'\rangle = \langle\Psi|i\hbar\frac{\partial}{\partial t} - z\phi|\Psi\rangle \quad (228)$$

$$\langle\Psi'|g|\Psi'\rangle = \langle\Psi|g|\Psi\rangle \quad (229)$$

More generally, the expectation value of any operator, including the Hamiltonian operators Eqs. (175) and (180), constructed entirely from  $\boldsymbol{\pi}$  and  $i\hbar\partial/\partial t - z\phi$  and from multiplicative operators will be invariant to gauge transformations.

It should be noted that, in approximate calculations, the chosen wave-function model may or may not have the flexibility needed to represent phase changes such as that in Eq. (226). If the model is not sufficiently flexible to accommodate such changes, then our results will be affected by gauge transformations of the potentials. Therefore, unless it is possible, in some convenient manner, to fix our potentials uniquely, the results of our calculations will be ambiguous. We shall later see how such a situation arises in connection with the application of external magnetic fields and how the resulting ambiguity may be resolved.

## VII. THE ONE-ELECTRON RELATIVISTIC HAMILTONIAN

### A. The Dirac equation

As in nonrelativistic theory, our starting point is the classical Hamiltonian, which for a relativistic electron in an electromagnetic field takes the form Eq. (150)

$$H = c\sqrt{\pi^2 + m_e^2 c^2} - e\phi \quad (230)$$

with the kinetic momentum given by

$$\boldsymbol{\pi} = \mathbf{p} + e\mathbf{A} \quad (231)$$

We recall that the nonrelativistic Schrödinger equation Eq. (178) is linear with respect to time differentiation but quadratic with respect to spatial differentiations. In relativistic theory, we would like time and space coordinates to appear in a more symmetric fashion. To obtain an expression that is linear in the momenta, we must get rid of the square root in Eq. (230). To accomplish this, we write

$$\pi_x^2 + \pi_y^2 + \pi_z^2 + m_e^2 c^2 = (\alpha_x \pi_x + \alpha_y \pi_y + \alpha_z \pi_z + \alpha_0 m_e c)^2 \quad (232)$$

where the  $\alpha_i$  are four Hermitian operators that commute with the components of  $\boldsymbol{\pi}$ . Upon expansion of the right-hand side, arguments similar to those used for Eq. (183) show that the  $\alpha_i$  anticommute:

$$[\alpha_i, \alpha_j]_+ = 2\delta_{ij} \quad (233)$$

In terms of these operators, the Hamiltonian becomes linear in the momenta

$$H_{\text{Dirac}} = c\boldsymbol{\alpha} \cdot \boldsymbol{\pi} + \beta m_e c^2 - e\phi \quad (234)$$

where  $\boldsymbol{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$  and  $\beta = \alpha_0$ . Upon quantization, we now obtain the time-dependent Dirac equation:

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = H_{\text{Dirac}} \Psi(\mathbf{r}, t) \quad (235)$$

where the time and space coordinates appear in a symmetric fashion, as required by relativity.

The anticommutation relations Eq. (233) are identical to those satisfied by the Pauli spin matrices Eq. (189). However, since there are only three independent Pauli matrices and we need four operators  $\alpha_i$ , we must here use representation matrices of higher dimension. In Exercise 5, it is shown that the Dirac operators are represented by the four-dimensional matrices

$$\boldsymbol{\alpha}_i = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_i \\ \boldsymbol{\sigma}_i & \mathbf{0} \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_2 \end{pmatrix} \quad (236)$$

where  $\mathbf{I}$  is the two-by-two unit matrix and the  $\boldsymbol{\sigma}_i$  are the Pauli spin matrices Eq. (188). In this representation, therefore, the relativistic wave function has four components

$$\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix} \quad (237)$$

and the Dirac equation becomes a set of four coupled equations. For example, the time-independent Dirac equation  $H_{\text{Dirac}} \Psi = E \Psi$  may be written as

$$c(\pi_- \Psi_4 + \pi_z \Psi_3) = (E - m_e c^2 + e\phi) \Psi_1 \quad (238)$$

$$c(\pi_+ \Psi_3 - \pi_z \Psi_4) = (E - m_e c^2 + e\phi) \Psi_2 \quad (239)$$

$$c(\pi_- \Psi_2 + \pi_z \Psi_1) = (E + m_e c^2 + e\phi) \Psi_3 \quad (240)$$

$$c(\pi_+ \Psi_1 - \pi_z \Psi_2) = (E + m_e c^2 + e\phi) \Psi_4 \quad (241)$$

where we have introduced

$$\pi_{\pm} = \pi_x \pm i\pi_y \quad (242)$$

Assuming that  $m_e^2 c^4 \gg c^2 \pi^2$ , the terms on the left-hand side of Eqs. (238)–(241) may be neglected. The solutions to the Dirac equation then separate into two classes: the electronic solutions with positive eigenvalues close to  $m_e c^2$  and the positronic solutions with negative eigenvalues close to  $-m_e c^2$ .

According to Hamilton's equation Eq. (24), we obtain the velocity by taking the partial derivative of the Hamiltonian with respect to the momentum. In the nonrelativistic and relativistic cases, using the quantum-mechanical Hamiltonians Eqs. (175) and (234), we obtain

$$\mathbf{v}_{1c} = \frac{\partial H_{1c}}{\partial \mathbf{p}} = \frac{\boldsymbol{\pi}}{m_e} \quad (243)$$

$$\mathbf{v}_D = \frac{\partial H_{\text{Dirac}}}{\partial \mathbf{p}} = c\boldsymbol{\alpha} \quad (244)$$

Whereas the nonrelativistic result is readily understood, the relativistic result is puzzling. Since the eigenvalues of the Dirac operators  $\alpha_i$  are  $\pm 1$  (see Exercise 5), this result means that the electron always travels at the speed of light, seemingly contradicting experimental observations that the speed of the electron is always less than the speed of light. This paradox is resolved by associating with the motion of the electron a rapid oscillation about its mean position, known as the *Zitterbewegung*. Thus, although the instantaneous velocity of the electron is always  $c$ , its measured "average" velocity is always lower.

## B. Reduction of the Dirac equation

As chemists we are interested in the positive-energy solutions to the Dirac equation. For slow electrons, we may construct an effective nonrelativistic Hamiltonian by decoupling the positive and negative solutions to the Dirac equation, reducing the four-component Dirac Hamiltonian to an effective two-component (Pauli-type) Hamiltonian, from which the electronic states may be obtained without considering the positronic solutions.

To see how such a decoupling may be accomplished, we begin by writing the Dirac wave function in the form

$$\boldsymbol{\Psi}^D = \begin{pmatrix} \psi_e \\ \psi_p \end{pmatrix} \quad (245)$$

where the indices "e" and "p" refer to the electronic and positronic parts of the wave function. Each of the two components may be expanded in the two spin components  $\alpha$  and  $\beta$  but this expansion is not studied here. To express the Dirac Hamiltonian Eq. (234) in the form of Eq. (245), we write the  $\alpha_i$  operators of Eq. (236) as

$$\alpha_i = \sigma_i \mathbf{J} \quad (246)$$

where  $\mathbf{J}$  is the two-by-two exchange matrix

$$\mathbf{J} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (247)$$

with the two dimensions referring to the electronic and positronic components. The Dirac Hamiltonian now takes the form

$$\mathbf{H}^D = m_e c^2 \boldsymbol{\beta} + c(\boldsymbol{\sigma} \cdot \boldsymbol{\pi}) \mathbf{J} - e\phi \mathbf{I} \quad (248)$$

where  $\mathbf{I}$  is the two-by-two unit matrix and  $\boldsymbol{\beta}$  the diagonal matrix

$$\boldsymbol{\beta} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (249)$$

The Dirac equation may now be written in the form

$$\mathbf{K}^D \boldsymbol{\Psi}^D = \mathbf{0} \quad (250)$$

where we have introduced the operator

$$\mathbf{K}^D = \mathbf{H}^D - i\hbar \frac{\partial}{\partial t} \mathbf{I} = m_e c^2 \boldsymbol{\beta} + c(\boldsymbol{\sigma} \cdot \boldsymbol{\pi}) \mathbf{J} + \boldsymbol{\tau} \quad (251)$$

$$\boldsymbol{\tau} = \tau \mathbf{I} = \left( -e\phi - i\hbar \frac{\partial}{\partial t} \right) \mathbf{I} \quad (252)$$

Since  $\mathbf{I}$  and  $\boldsymbol{\beta}$  are diagonal and  $\mathbf{J}$  is purely off diagonal, we may express Eq. (250) as

$$\begin{pmatrix} m_e c^2 + \mathcal{O}(c^0) & \mathcal{O}(c) \\ \mathcal{O}(c) & -m_e c^2 + \mathcal{O}(c^0) \end{pmatrix} \begin{pmatrix} \psi_e \\ \psi_p \end{pmatrix} = \mathbf{0} \quad (253)$$

Folding the positronic component  $\psi_p$  into the electronic component  $\psi_e$ , we obtain

$$[m_e c^2 + \mathcal{O}(c^0)] \psi_e - \mathcal{O}(c) [-m_e c^2 + \mathcal{O}(c^0)]^{-1} \mathcal{O}(c) \psi_e = 0 \quad (254)$$

Since the contribution from the positronic component is of order unity, its neglect would lead to errors of the same order of magnitude as the nonrelativistic energy.

We now perform a unitary transformation of the Dirac equation Eq. (250), reducing the coupling between the electronic and positronic components such that the solution of the resulting electronic equation is correct to order  $c^{-2}$  compared with the full solution. For this purpose, we use the Foldy–Wouthuysen technique, generating a transformed Dirac equation

$$\tilde{\mathbf{K}}^D \tilde{\boldsymbol{\Psi}}^D = \mathbf{0} \quad (255)$$

with the Hamiltonian and wave function expressed in terms of an anti-Hermitian matrix  $\mathbf{S}$ :

$$\tilde{\mathbf{K}}^{\text{D}} = \exp(\mathbf{S})\mathbf{K}^{\text{D}}\exp(-\mathbf{S}) \quad (256)$$

$$\tilde{\Psi}^{\text{D}} = \exp(\mathbf{S})\Psi^{\text{D}} \quad (257)$$

Since we are only interested in rotations between the electronic and positronic components of  $\mathbf{K}^{\text{D}}$ , we use a matrix of the form

$$\mathbf{S} = \begin{pmatrix} 0 & \rho + \delta \\ -\rho + \delta & 0 \end{pmatrix} = \rho\boldsymbol{\beta}\mathbf{J} + \delta\mathbf{J} \quad (258)$$

where  $\rho$  is Hermitian and  $\delta$  anti-Hermitian. The decoupling is achieved by choosing  $\rho$  and  $\delta$  such that the first-order off-diagonal term  $c(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})\mathbf{J}$  in Eq. (251) is eliminated, introducing in the process new terms of higher orders in  $c^{-1}$ . Consider the first-order BCH expansion of Eq. (256):

$$\tilde{\mathbf{K}}^{\text{D}} = \mathbf{K}^{\text{D}} + [\mathbf{S}, \mathbf{K}^{\text{D}}] + \dots \quad (259)$$

The  $\rho$  and  $\delta$  are determined such that the highest-order commutator term  $[\mathbf{S}, m_e c^2 \boldsymbol{\beta}]$  cancels the highest-order coupling term  $c(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})$  in  $\mathbf{K}^{\text{D}}$ :

$$c(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})\mathbf{J} + [\mathbf{S}, m_e c^2 \boldsymbol{\beta}] = \mathbf{0} \quad (260)$$

Using the relations (which follow directly from the definition of  $\boldsymbol{\beta}$  and  $\mathbf{J}$ )

$$\mathbf{J}^2 = \boldsymbol{\beta}^2 = \mathbf{I} \quad (261)$$

$$\mathbf{J}\boldsymbol{\beta} = -\boldsymbol{\beta}\mathbf{J} \quad (262)$$

we then obtain

$$[c(\boldsymbol{\sigma} \cdot \boldsymbol{\pi}) - 2m_e c^2 \rho]\mathbf{J} - 2m_e c^2 \delta \boldsymbol{\beta}\mathbf{J} = \mathbf{0} \quad (263)$$

The decoupling is thus achieved by setting  $\delta = 0$  and choosing  $\rho$  such that

$$\rho = \frac{\boldsymbol{\sigma} \cdot \boldsymbol{\pi}}{2m_e c} \quad (264)$$

Note that  $\mathbf{S}$  is  $c^{-2}$  smaller than the term that it removes. In general, if the coupling element to be removed  $\mathbf{T}$  is of order  $c^n$ , then  $\mathbf{S}$  is of order  $c^{n-2}$  and the new term  $[\mathbf{S}, \mathbf{T}]$  introduced in the electronic–electronic element by the Foldy–Wouthuysen transformation is of order  $c^{n-2}c^n = c^{2n-2}$ .

We must now evaluate the terms in the expansion of  $\tilde{\mathbf{K}}^D$ , with our chosen transformation matrix  $\mathbf{S}$ . In the notation  $\tilde{\mathbf{M}} = \exp(\mathbf{S})\mathbf{M}\exp(-\mathbf{S})$ , the Dirac equation now takes the form

$$\left[ m_e c^2 \tilde{\boldsymbol{\beta}} + 2m_e c^2 \rho \tilde{\mathbf{J}} + \tilde{\boldsymbol{\tau}} \right] \tilde{\Psi}^D = \mathbf{0} \quad (265)$$

where, from Eqs. (261) and (262), we obtain the following expansions

$$\tilde{\boldsymbol{\beta}} = \left( 1 - 2\rho^2 + \frac{2}{3}\rho^4 + \dots \right) \boldsymbol{\beta} - \left( 2\rho - \frac{4}{3}\rho^3 + \dots \right) \mathbf{J} \quad (266)$$

$$\tilde{\mathbf{J}} = \left( 1 - 2\rho^2 + \frac{2}{3}\rho^4 + \dots \right) \mathbf{J} + \left( 2\rho - \frac{4}{3}\rho^3 + \dots \right) \boldsymbol{\beta} \quad (267)$$

$$\tilde{\boldsymbol{\tau}} = \left( \tau - \frac{1}{2}[\rho, [\rho, \tau]] + \dots \right) \mathbf{I} + \left( [\rho, \tau] - \frac{1}{6}[\rho, [\rho, [\rho, \tau]]] + \dots \right) \boldsymbol{\beta} \mathbf{J} \quad (268)$$

Including only the lower-order terms of  $\rho$  and using Eq. (264), we arrive at the following expression for  $\tilde{\mathbf{K}}^D$ :

$$\begin{aligned} \tilde{\mathbf{K}}^D &= m_e c^2 \left( 1 + 2\rho^2 - 2\rho^4 \right) \boldsymbol{\beta} + \left( \tau - \frac{1}{2}[\rho, [\rho, \tau]] \right) \mathbf{I} \\ &\quad + \mathcal{O}(c^{-4})\boldsymbol{\beta} + \mathcal{O}(c^{-4})\mathbf{I} + \mathcal{O}(c^{-1})\mathbf{J} + \mathcal{O}(c^{-1})\boldsymbol{\beta} \mathbf{J} \end{aligned} \quad (269)$$

The largest off-diagonal terms are now proportional to  $c^{-1}$ . This term can be removed by a further Foldy–Wouthuysen transformation of order  $c^{-3}$ , introducing new terms of order  $c^{-4}$  in the electronic–electronic element of the operator. The electronic solution to the Dirac equation to order  $c^{-2}$  may thus be obtained from the purely electronic part of  $\tilde{\mathbf{K}}^D$ :

$$\begin{aligned} \tilde{K}_{ee}^D &= m_e c^2 + \frac{1}{2m_e} (\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2 + \tau \\ &\quad - \frac{1}{8m_e^3 c^2} (\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^4 - \frac{1}{8m_e^2 c^2} [\boldsymbol{\sigma} \cdot \boldsymbol{\pi}, [\boldsymbol{\sigma} \cdot \boldsymbol{\pi}, \tau]] + \mathcal{O}(c^{-4}) \end{aligned} \quad (270)$$

The double commutator may be simplified by using the identity

$$\begin{aligned} [\boldsymbol{\pi}, \tau] &= [-i\hbar \boldsymbol{\nabla} + e\mathbf{A}, -e\phi - i\hbar \partial/\partial t] = i\hbar [\boldsymbol{\nabla}, \phi] + i\hbar [\partial/\partial t, \mathbf{A}] \\ &= i\hbar (\boldsymbol{\nabla} \phi + \partial \mathbf{A} / \partial t) = -ie\hbar \mathbf{E} \end{aligned} \quad (271)$$

where we have substituted Eq. (252) for  $\tau$  and used Eq.(48) to introduce the electric field  $\mathbf{E}$ . Introducing the resulting expression for  $\tilde{K}_{ee}^D$  in Eq. (265) and expanding in orders of  $c^{-2}$ , we arrive at the following two-component equation

$$\left( m_e c^2 + H^{(0)} + H^{(1)} + \dots \right) \psi_e = i\hbar \frac{\partial \psi_e}{\partial t} \quad (272)$$



where the zero- and first-order Hamiltonians are given by

$$H^{(0)} = \frac{1}{2m_e}(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2 - e\phi \quad (273)$$

$$H^{(1)} = -\frac{1}{8m_e^3c^2}(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^4 + \frac{ie\hbar}{8m_e^2c^2}[\boldsymbol{\sigma} \cdot \boldsymbol{\pi}, \boldsymbol{\sigma} \cdot \mathbf{E}] \quad (274)$$

The zero-order Hamiltonian is recognized as the two-component nonrelativistic Hamiltonian Eq. (180). In the next subsection, we shall study this first-order Hamiltonian in greater detail.

### C. The first-order relativistic one-electron Hamiltonian

The first-order relativistic one-electron Hamiltonian Eq.(274) consists of two terms

$$H^{(1)} = H_{\text{kin}}^{(1)} + H_{\text{pot}}^{(1)} \quad (275)$$

representing kinetic and potential corrections to the nonrelativistic Hamiltonian Eq. (273):

$$H_{\text{kin}}^{(1)} = -\frac{(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^4}{8m_e^3c^2}, \quad H_{\text{pot}}^{(1)} = \frac{ie\hbar}{8m_e^2c^2}[\boldsymbol{\sigma} \cdot \boldsymbol{\pi}, \boldsymbol{\sigma} \cdot \mathbf{E}] \quad (276)$$

The kinetic operator  $H_{\text{kin}}^{(1)}$  represents a relativistic mass-velocity correction to the nonrelativistic operator  $(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2/2m_e$ . By squaring the expressions for  $(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2$  given in Eq. (195) and using Eq. (177), we generate a number of contributions to the relativistic kinetic energy:

$$\begin{aligned} H_{\text{kin}}^{(1)} &= -\frac{\pi^4}{8m_e^3c^2} - \frac{e}{4m_e^3c^2}(\pi^2 \mathbf{B} \cdot \mathbf{s} + \mathbf{B} \cdot \mathbf{s} \pi^2) - \frac{e^2}{2m_e^3c^2}(\mathbf{B} \cdot \mathbf{s})^2 \\ &= -\frac{p^4}{8m_e^3c^2} - \frac{e}{4m_e^3c^2}[(p^2 \mathbf{A} \cdot \mathbf{p} + \mathbf{A} \cdot \mathbf{p} p^2) + (p^2 \mathbf{B} \cdot \mathbf{s} + \mathbf{B} \cdot \mathbf{s} p^2)] + \mathcal{O}(A^2) \end{aligned} \quad (277)$$

where  $\mathcal{O}(A^2)$  contains terms that are at least quadratic or bilinear in  $\mathbf{A}$  and  $\mathbf{B} = \nabla \times \mathbf{A}$ . The mass-velocity correction  $-\pi^4/8m_e^3c^2$  is always negative, reducing the kinetic energy of the electron, giving the largest relativistic correction to the total electronic energy; it may alternatively be obtained from the expansion in Eq. (151) of the classical relativistic energy in orders of  $c^{-2}$ . The terms proportional to  $\mathbf{A} \cdot \mathbf{p}$  and  $\mathbf{B} \cdot \mathbf{s}$  are mass-velocity corrections to the nonrelativistic orbital and spin paramagnetic interaction operators in Eq. (198).

To simplify the first-order potential operator  $H_{\text{pot}}^{(1)}$  in Eq. (276), we first write

$$[\boldsymbol{\sigma} \cdot \boldsymbol{\pi}, \boldsymbol{\sigma} \cdot \mathbf{E}] = \boldsymbol{\pi} \cdot \mathbf{E} - \mathbf{E} \cdot \boldsymbol{\pi} + i\boldsymbol{\sigma} \cdot (\boldsymbol{\pi} \times \mathbf{E} - \mathbf{E} \times \boldsymbol{\pi}) \quad (278)$$

where we have used

$$(\boldsymbol{\sigma} \cdot \mathbf{u})(\boldsymbol{\sigma} \cdot \mathbf{v}) = \mathbf{u} \cdot \mathbf{v} + i\boldsymbol{\sigma} \cdot \mathbf{u} \times \mathbf{v} \quad (279)$$

obtained from Eq. (191) for arbitrary  $\mathbf{u}$  and  $\mathbf{v}$ . Next, substituting  $\boldsymbol{\pi} = -i\hbar\nabla + e\mathbf{A}$  and  $\boldsymbol{\sigma} = 2\mathbf{s}/\hbar$  in Eq. (278), we arrive at the following expression for the first-order relativistic correction to the potential-energy operator:

$$H_{\text{pot}}^{(1)} = \frac{e\hbar^2}{8m_e^2c^2} (\nabla \cdot \mathbf{E}) + \frac{e}{4m_e^2c^2} \mathbf{s} \cdot (\mathbf{E} \times \boldsymbol{\pi} - \boldsymbol{\pi} \times \mathbf{E}) \quad (280)$$

The first term is the one-electron Darwin operator

$$H_{\text{D1}} = \frac{e\hbar^2}{8m_e^2c^2} (\nabla \cdot \mathbf{E}) \quad (281)$$

which is interpreted as a correction to the potential energy for the charge smearing caused by the Zitterbewegung of the electron, discussed in Section VII A. In the important special case of Coulomb attraction to a static nucleus of charge  $Ze$ , the vector potential vanishes and the electrostatic potential becomes

$$\phi^{\text{nuc}} = \frac{Ze}{4\pi\epsilon_0 r} \quad (282)$$

Using Eq. (48) and Eq. (A19), we then obtain

$$\nabla \cdot \mathbf{E}^{\text{nuc}} = -\nabla^2\phi^{\text{nuc}} = -\frac{Ze}{4\pi\epsilon_0} \nabla^2 r^{-1} = \frac{Ze}{\epsilon_0} \delta(\mathbf{r}) \quad (283)$$

so the one-electron Darwin operator becomes

$$H_{\text{D1}}^{\text{nuc}} = \frac{Ze^2\hbar^2}{8\epsilon_0 m_e^2 c^2} \delta(\mathbf{r}) \quad (284)$$

The one-electron Darwin correction is thus positive, reducing the overall attraction of the electron to the nucleus, in agreement with the notion of Zitterbewegung. The presence of  $\delta(\mathbf{r})$  may be understood from the observation that, if the electron is some distance away from the nucleus, its Zitterbewegung will not affect its mean distance from the nucleus and therefore not change the interaction. On the other hand, at the position of the nucleus, the oscillations of the electron will increase its mean separation to the nucleus, thereby reducing the interaction. The one-electron Darwin operator usually provides the second largest (positive) contribution to the electronic energy, partly canceling the (negative) mass-velocity correction.

The second contribution to the first-order relativistic Hamiltonian Eq. (280) is the one-electron spin-orbit operator:

$$H_{\text{SO1}} = \frac{e}{4m_e^2c^2} \mathbf{s} \cdot (\mathbf{E} \times \boldsymbol{\pi} - \boldsymbol{\pi} \times \mathbf{E}) \quad (285)$$

which couples the spin of the electron to its orbital motion in an external electric field  $\mathbf{E}$ . In a purely electrostatic external field with no magnetic component,  $\mathbf{E}$  is irrotational  $\boldsymbol{\nabla} \times \mathbf{E} = \mathbf{0}$  by Faraday's law Eq. (39) and we may choose  $\mathbf{A} = \mathbf{0}$ . We may then write the spin-orbit operator in the form

$$H_{\text{SO1}}^{\text{es}} = -\mathbf{m} \cdot \frac{\mathbf{E} \times \mathbf{p}}{2m_e c^2} \quad (286)$$

where  $\mathbf{m} = -(e/m_e)\mathbf{s}$ , see Eq. (199). We may now interpret the spin-orbit term as the Zeeman interaction of the electron's magnetic moment  $\mathbf{m}$  with the magnetic field  $\mathbf{E} \times \mathbf{v}/2c^2$  experienced by the electron as it moves through the electric field  $\mathbf{E}$  with velocity  $\mathbf{v} = \mathbf{p}/m_e$ . Classically, however, the field experienced by the electron is given by  $\mathbf{E} \times \mathbf{v}/c^2$  rather than by  $\mathbf{E} \times \mathbf{v}/2c^2$ . The difference occurs since the classical treatment neglects the *Thomas precession* of the electron in the external field, which reduces the interaction by a factor of two. In the special case of a Coulomb potential from a point-charge nucleus of charge  $Ze$ , the electrostatic potential is given by Eq. (282). From Eq. (48) and Eq. (A16), we find that the spin-orbit operator takes the form

$$H_{\text{SO1}}^{\text{nuc}} = \frac{Ze^2}{8\pi\epsilon_0 m_e^2 c^2} \frac{\mathbf{s} \cdot \mathbf{r} \times \boldsymbol{\pi}}{r^3} \quad (287)$$

where  $\mathbf{r} \times \boldsymbol{\pi}$  is the angular momentum of the electron about the nucleus.

#### D. The two-component Pauli Hamiltonian

The numerical constants in the two-component Hamiltonian simplify considerably in atomic units, where  $\hbar = e = m_e = a_0 = 1$ . In these units, the permittivity and permeability of free space are given by  $4\pi\epsilon_0 = 1$  and  $\mu_0 = 4\pi\alpha^2$ , respectively, where  $\alpha$  is the dimensionless fine-structure constant  $\alpha = e^2/4\pi\epsilon_0\hbar c$ . In atomic units, the speed of light in vacuum is therefore equal to the inverse of the fine-structure constant  $c = 1/\sqrt{\epsilon_0\mu_0} = \alpha^{-1} \approx 137.0$ . Therefore, the Pauli-type two-component Hamiltonian  $H^{(0)} + H^{(1)}$  of Eqs. (273) and (274)

may be written as

$$h_{\text{Pauli}} = \begin{cases} \frac{1}{2}\pi^2 - \phi & \leftarrow \text{Schrödinger} \\ -\frac{1}{8}\alpha^2\pi^4 & \leftarrow \text{mass velocity} \\ +\frac{1}{8}\alpha^2(\nabla \cdot \mathbf{E}) & \leftarrow \text{Darwin} \\ +\mathbf{B} \cdot \mathbf{s} - \frac{1}{4}\alpha^2(\pi^2 \mathbf{B} \cdot \mathbf{s} + \mathbf{B} \cdot \mathbf{s} \pi^2) - \frac{1}{2}\alpha^2(\mathbf{B} \cdot \mathbf{s})^2 & \leftarrow \text{Zeeman} \\ +\frac{1}{4}\alpha^2 \mathbf{s} \cdot (\mathbf{E} \times \boldsymbol{\pi} - \boldsymbol{\pi} \times \mathbf{E}) & \leftarrow \text{spin-orbit} \end{cases} \quad (288)$$

All relativistic corrections contain the square of the fine-structure constant  $\alpha^2 \approx 5.2 \cdot 10^{-5}$  and are therefore classified as small. Of course, this holds only for electrons moving in the weak electromagnetic fields of light atomic nuclei.

Macroscopically applied electromagnetic fields are usually weak—for example, a magnetic induction of 10 T (typical of magnetic resonance experiments) corresponds to  $4.3 \cdot 10^{-5}$  in atomic units. The Zeeman interaction term is then of the same order of magnitude as  $\alpha^2$ . Similarly, macroscopically applied electric fields (which enter the Hamiltonian through the potential  $\phi$ ) are usually weak. For example, a field strength of  $10^5 \text{ Vm}^{-1}$  (typical of the accelerator of a television set) corresponds to  $1.9 \cdot 10^{-7}$  in atomic units. Thus, most macroscopically applied electromagnetic fields may be considered weak and of the same order of magnitude as  $\alpha^2$ , although strong dynamical fields may be generated using lasers.

By contrast, microscopically generated fields may be very strong. For example, the electric field generated by a nucleus may be sufficiently strong for the validity of the Pauli-type Hamiltonian Eq. (288) to break down. Consider an electron moving in a nuclear Coulomb field, for which the classical Hamiltonian of Eq. (230) takes the form

$$H = \sqrt{m_e^2 c^4 + p^2 c^2} - \frac{Ze^2}{4\pi\epsilon_0 r} \quad (289)$$

where we have inserted Eq. (282). For a given total electronic energy, there will always be a region close to the nucleus where the motion of the electron becomes relativistic. The electron may be treated by nonrelativistic quantum mechanics only if this region is sufficiently small.

## VIII. THE TWO-ELECTRON RELATIVISTIC HAMILTONIAN

### A. The Breit equation

The Dirac equation gives a relativistic quantum-mechanical treatment of one electron. For the relativistic treatment of two electrons, we need to consider the interaction between the electrons. In Eq. (167), the Hamiltonian for the relativistic classical interaction between two electrons was given to order  $v^2/c^2$ . Just as it was impossible to set up the exact relativistic classical two-electron interaction, it is not possible to obtain the exact relativistic quantum-mechanical two-electron interaction. The relativistic quantum-mechanical Hamiltonian corresponding to the classical two-electron Hamiltonian Eq. (167) is the Breit Hamiltonian

$$H_{\text{Breit}} = H_1^{\text{D}} + H_2^{\text{D}} + H_{12}^{\text{C}} + H_{12}^{\text{G}} + H_{12}^{\text{R}} \quad (290)$$

where the one-electron Dirac Hamiltonians are given by

$$H_i^{\text{D}} = \beta_i m_e c^2 + c \boldsymbol{\alpha}_i \cdot \boldsymbol{\pi}_i - e\phi_i \quad (291)$$

and the two-electron Coulomb, Gaunt, and retardation operators by

$$H_{12}^{\text{C}} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{12}} \quad (292)$$

$$H_{12}^{\text{G}} = -\frac{e^2}{4\pi\epsilon_0} \frac{\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{r_{12}} \quad (293)$$

$$H_{12}^{\text{R}} = \frac{e^2}{8\pi\epsilon_0} \frac{r_{12}^2 (\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2) - (\boldsymbol{\alpha}_1 \cdot \mathbf{r}_{12})(\mathbf{r}_{12} \cdot \boldsymbol{\alpha}_2)}{r_{12}^3} \quad (294)$$

The two-electron part of the Breit Hamiltonian is obtained from its classical counterpart by performing the substitutions  $\boldsymbol{\pi}_i \rightarrow m_i \mathbf{v}_i \rightarrow m_e c \boldsymbol{\alpha}_i$  where, in the first step, we have ignored the higher-order contributions to  $\boldsymbol{\pi}_i$  and, in the second step, we have made the identification Eq. (244) of  $c\boldsymbol{\alpha}_i$  with the velocity operator  $\mathbf{v}_i$ . As a result, the Breit operator

$$H_{12}^{\text{B}} = H_{12}^{\text{G}} + H_{12}^{\text{R}} \quad (295)$$

does not contain the  $c^{-2}$  nor the electron masses, unlike its classical counterpart. The two-electron part of the Breit Hamiltonian Eq. (290) may be derived rigorously by the methods of quantum electrodynamics.

The Breit Hamiltonian Eq. (290) enters the Breit equation

$$i\hbar \frac{\partial \Psi(\mathbf{r}_1, \mathbf{r}_2, t)}{\partial t} = H_{\text{Breit}} \Psi(\mathbf{r}_1, \mathbf{r}_2, t) \quad (296)$$

which is a two-electron analog of the Dirac equation. Whereas the Dirac equation is a four-component equation, the Breit equation has 16 components. Also, whereas the Dirac equation is an exact relativistic equation in the sense that it is Lorentz invariant, the Breit equation is only approximately Lorentz invariant because of simplifications made in the treatment of the two-electron interactions.

## B. Matrix form of the Breit equation

The solution to the Breit equation Eq. (296) may be expressed as a four-dimensional vector

$$\Psi^{\text{B}}(\mathbf{r}_1, \mathbf{r}_2) = \begin{pmatrix} \Psi_{\text{ee}}(\mathbf{r}_1, \mathbf{r}_2) \\ \Psi_{\text{ep}}(\mathbf{r}_1, \mathbf{r}_2) \\ \Psi_{\text{pe}}(\mathbf{r}_1, \mathbf{r}_2) \\ \Psi_{\text{pp}}(\mathbf{r}_1, \mathbf{r}_2) \end{pmatrix} \quad (297)$$

where the indices “e” and “p” refer to the electronic and positronic parts of the wave function, respectively. Each of the four components may be further expanded in the four spin components of a two-electron wave function ( $\alpha\alpha, \alpha\beta, \beta\alpha, \beta\beta$ ), but this expansion will not be studied here. In agreement with the partitioning of the wave function Eq. (297), the Breit Hamiltonian Eq. (290) may be partitioned into a four-by-four matrix, where each element is an operator working on  $\mathbf{r}_1$  and  $\mathbf{r}_2$ .

Following the discussion in Section VII B, we write the Dirac Hamiltonian of each particle in the form

$$\mathbf{H}_i^{\text{D}} = m_e c^2 \boldsymbol{\beta} + c(\boldsymbol{\sigma}_i \cdot \boldsymbol{\pi}_i) \mathbf{J} - e\phi_i \mathbf{I} \quad (298)$$

in the notation of Eq. (248). In setting up a representation of the Dirac Hamiltonians in the four-dimensional space of both particles, we note that the Hamiltonian of one particle does not affect the other particle. These operators must therefore be of direct-product form:

$$\mathcal{H}_1^{\text{D}} = \mathbf{H}_1^{\text{D}} \otimes \mathbf{I}, \quad (299)$$

$$\mathcal{H}_2^{\text{D}} = \mathbf{I} \otimes \mathbf{H}_2^{\text{D}} \quad (300)$$

The two-electron Coulomb operator Eq. (292) does not couple the electronic and positronic parts of the wave function and may be written as

$$\mathcal{H}_{12}^{\text{C}} = h_{12}^{\text{C}} \mathbf{I} \otimes \mathbf{I}, \quad h_{12}^{\text{C}} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{12}} \quad (301)$$

In the Gaunt and retardation operators Eqs. (293) and (294),  $\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2$  and  $(\boldsymbol{\alpha}_1 \cdot \mathbf{r}_{12})(\mathbf{r}_{12} \cdot \boldsymbol{\alpha}_2)$  couple the different components of the wave function. Using Eq. (246), we obtain

$$\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2 = (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) \mathbf{J} \otimes \mathbf{J}, \quad (302)$$

$$(\boldsymbol{\alpha}_1 \cdot \mathbf{r}_{12})(\mathbf{r}_{12} \cdot \boldsymbol{\alpha}_2) = (\boldsymbol{\sigma}_1 \cdot \mathbf{r}_{12})(\mathbf{r}_{12} \cdot \boldsymbol{\sigma}_2) \mathbf{J} \otimes \mathbf{J} \quad (303)$$

so we have

$$\mathcal{H}_{12}^G = h_{12}^G \mathbf{J} \otimes \mathbf{J}, \quad h_{12}^G = -\frac{e^2}{4\pi\epsilon_0} \frac{\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2}{r_{12}} \quad (304)$$

$$\mathcal{H}_{12}^R = h_{12}^R \mathbf{J} \otimes \mathbf{J}, \quad h_{12}^R = \frac{e^2}{8\pi\epsilon_0} \frac{\boldsymbol{\sigma}_1 \cdot r_{12}^2 \boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1 \cdot \mathbf{r}_{12} \mathbf{r}_{12} \cdot \boldsymbol{\sigma}_2}{r_{12}^3} \quad (305)$$

and likewise

$$\mathcal{H}_{12}^B = h_{12}^B \mathbf{J} \otimes \mathbf{J}, \quad h_{12}^B = h_{12}^G + h_{12}^R \quad (306)$$

The Breit equation Eq. (296) may thus be written in matrix form as

$$i\hbar \frac{\partial \Psi_{12}(\mathbf{r}_1, \mathbf{r}_2, t)}{\partial t} = \mathcal{H}^B \Psi_{12}(\mathbf{r}_1, \mathbf{r}_2, t) \quad (307)$$

where the Breit Hamiltonian

$$\mathcal{H}^B = \mathcal{H}_1^D + \mathcal{H}_2^D + \mathcal{H}_{12}^C + \mathcal{H}_{12}^G + \mathcal{H}_{12}^R \quad (308)$$

is expressed in terms of the Dirac Hamiltonians Eqs. (299) and (300), the Coulomb operator Eq. (301), and the Gaunt and retardation operators Eqs. (304) and (305), respectively. Note that, for two noninteracting one-electron systems, the two-electron terms in Eq. (308) vanish and the solutions to Eq. (307) become the direct products of the solutions to the one-electron Dirac equations.

### C. Reduction of the Breit equation

The two-particle Breit equation Eq. (307) yields, in addition to two-electron solutions, also two-positron and mixed electron–positron solutions. In chemistry, we are only interested in two-electron solutions. We shall now show how we can transform the Breit equation Eq. (307) written in the form (by analogy with Eqs. (250))

$$\mathcal{K}^B \Psi_{12} = \mathbf{0} \quad (309)$$

$$\mathcal{K}^B = \mathcal{H}^B - i\hbar \frac{\partial}{\partial t} \mathbf{I} \otimes \mathbf{I} \quad (310)$$

to a single block that yields the electron–electron solutions. As for the transformed one-electron equation in Section VII B, we shall require that the solutions to the transformed electron–electron Hamiltonian only contains errors of order  $c^{-4}$  compared with the full Breit equation.

It is instructive to begin by examining  $\mathcal{K}^B$  in more detail. From the explicit form of the operators entering Eq. (308), we find

$$\mathcal{K}^B = \begin{pmatrix} 2m_e c^2 + \tau + h_{12}^C & c(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\pi}_1) & c(\boldsymbol{\sigma}_2 \cdot \boldsymbol{\pi}_2) & h_{12}^B \\ c(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\pi}_1) & \tau + h_{12}^C & h_{12}^B & c(\boldsymbol{\sigma}_2 \cdot \boldsymbol{\pi}_2) \\ c(\boldsymbol{\sigma}_2 \cdot \boldsymbol{\pi}_2) & h_{12}^B & \tau + h_{12}^C & c(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\pi}_1) \\ h_{12}^B & c(\boldsymbol{\sigma}_2 \cdot \boldsymbol{\pi}_2) & c(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\pi}_1) & -2m_e c^2 + \tau + h_{12}^C \end{pmatrix} \quad (311)$$

where we have introduced the Breit operator

$$h_{12}^B = h_{12}^G + h_{12}^R \quad (312)$$

and collected the time differentiation together with the scalar potential

$$\tau = -i\hbar \frac{\partial}{\partial t} - e\phi_1 - e\phi_2 \quad (313)$$

by analogy with Eq. (252). The diagonal electron–electron and positron–positron elements  $\mathcal{K}_{eeee}^B$  and  $\mathcal{K}_{pppp}^B$  contain terms of order  $c^2$  and lower, whereas the diagonal electron–positron and positron–electron elements  $\mathcal{K}_{eepp}^B$  and  $\mathcal{K}_{ppee}^B$  contain terms of order  $c^0$  and lower. Moreover, whereas the coupling of  $\mathcal{K}_{eeee}^B$  to the elements  $\mathcal{K}_{eepp}^B$ ,  $\mathcal{K}_{epee}^B$ ,  $\mathcal{K}_{eppe}^B$ , and  $\mathcal{K}_{ppee}^B$  containing  $c(\boldsymbol{\sigma}_i \cdot \boldsymbol{\pi}_i)$  is of order  $c$ , the coupling to  $\mathcal{K}_{eepp}^B$  and  $\mathcal{K}_{ppee}^B$  containing  $h_{12}^B$  is of order unity. Our task now is to reduce these couplings, employing the Foldy–Wouthuysen technique of Section VII B, suitably generalized to two particles.

Carrying out a unitary transformation of Eq. (309), we obtain

$$\tilde{\mathcal{K}}^B \tilde{\Psi}_{12} = \mathbf{0}, \quad \tilde{\mathcal{K}}^B = \exp(\mathbf{S}_{12}) \mathcal{K}^B \exp(-\mathbf{S}_{12}), \quad \tilde{\Psi}_{12} = \exp(\mathbf{S}_{12}) \Psi_{12} \quad (314)$$

where the four-dimensional anti-Hermitian operator  $\mathbf{S}_{12}$  is chosen to eliminate  $c(\boldsymbol{\sigma}_i \cdot \boldsymbol{\pi}_i)$  in Eq. (311). Since the transformation generated by  $\mathbf{S}_i = \rho_i \boldsymbol{\beta} \mathbf{J}$  with  $\rho_i = (\boldsymbol{\sigma}_i \cdot \boldsymbol{\pi}_i)/2m_e c$  accomplished this task for one electron in Section VII B, we attempt the transformation generated by the direct sum

$$\mathbf{S}_{12} = \mathbf{S}_1 \oplus \mathbf{S}_2 \quad (315)$$



noting from Eq. (B13) that the exponential of  $\mathbf{S}_{12}$  factorizes into a direct product

$$\exp(\mathbf{S}_{12}) = \exp(\mathbf{S}_1) \otimes \exp(\mathbf{S}_2) \quad (316)$$

From Eqs. (299) and (300), we then obtain for the one-particle operators

$$\tilde{\mathcal{H}}_1^D = [\exp(\mathbf{S}_1) \otimes \exp(\mathbf{S}_2)] [\mathbf{H}_1^D \otimes \mathbf{I}] [\exp(-\mathbf{S}_1) \otimes \exp(-\mathbf{S}_2)] = \tilde{\mathbf{H}}_1^D \otimes \mathbf{I} \quad (317)$$

$$\tilde{\mathcal{H}}_2^D = [\exp(\mathbf{S}_1) \otimes \exp(\mathbf{S}_2)] [\mathbf{I} \otimes \mathbf{H}_2^D] [\exp(-\mathbf{S}_1) \otimes \exp(-\mathbf{S}_2)] = \mathbf{I} \otimes \tilde{\mathbf{H}}_2^D \quad (318)$$

where we have used Eq. (B3). The one-particle part of the transformed Breit Hamiltonian is therefore the direct sum of the similarly transformed Dirac Hamiltonians, with couplings of order  $c^{-1}$  as discussed in Section VII B. Transforming the Coulomb and Breit operators in the same manner, we obtain

$$\begin{aligned} \tilde{\mathcal{H}}_{12}^C &= (h_{12}^C - \frac{1}{2} [\rho_1, [\rho_1, h_{12}^C]] - \frac{1}{2} [\rho_2, [\rho_2, h_{12}^C]]) \mathbf{I} \otimes \mathbf{I} \\ &\quad + [\rho_1, h_{12}^C] \boldsymbol{\beta} \mathbf{J} \otimes \mathbf{I} + [\rho_2, h_{12}^C] \mathbf{I} \otimes \boldsymbol{\beta} \mathbf{J} + [\rho_1, [\rho_2, h_{12}^C]] \boldsymbol{\beta} \mathbf{J} \otimes \boldsymbol{\beta} \mathbf{J} + \mathcal{O}(c^{-3}) \end{aligned} \quad (319)$$

$$\begin{aligned} \tilde{\mathcal{H}}_{12}^B &= (h_{12}^B - \frac{1}{2} [\rho_1, [\rho_1, h_{12}^B]_{++}] - \frac{1}{2} [\rho_2, [\rho_2, h_{12}^B]_{++}]) \mathbf{J} \otimes \mathbf{J} \\ &\quad + [\rho_1, h_{12}^B]_{++} \boldsymbol{\beta} \otimes \mathbf{J} + [\rho_2, h_{12}^B]_{++} \mathbf{J} \otimes \boldsymbol{\beta} + [\rho_1, [\rho_2, h_{12}^B]_{++}] \boldsymbol{\beta} \otimes \boldsymbol{\beta} + \mathcal{O}(c^{-3}) \end{aligned} \quad (320)$$

Although the Breit operator occurs only in the coupling blocks of  $\tilde{\mathcal{H}}^B$ , the  $\mathbf{S}_{12}$  transformation introduces Breit contributions  $\boldsymbol{\beta} \otimes \boldsymbol{\beta}$  into the diagonal blocks of  $\tilde{\mathcal{H}}^B$ . Collecting the electron-electron contributions of Eqs. (317)–(320), we obtain

$$\tilde{\mathcal{K}}_{\text{eeee}}^B = 2m_e c^2 + H_{\text{NR}} + H_{\text{BP}} - i\hbar \frac{\partial}{\partial t} + \mathcal{O}(c^{-4}) \quad (321)$$

where we have introduced the nonrelativistic and Breit–Pauli Hamiltonians

$$H_{\text{NR}} = \sum_{i=1}^2 \left( \frac{1}{2m_e} (\boldsymbol{\sigma}_i \cdot \boldsymbol{\pi}_i)^2 - e\phi_i \right) + g_{12} \quad (322)$$

$$\begin{aligned} H_{\text{BP}} &= - \sum_{i=1}^2 \left( \frac{1}{8m_e^3 c^2} (\boldsymbol{\sigma}_i \cdot \boldsymbol{\pi}_i)^4 - \frac{1}{2} [\rho_i, [\rho_i, e\phi_i + i\hbar \frac{\partial}{\partial t}]] \right) \\ &\quad - \frac{1}{2} [\rho_1, [\rho_1, h_{12}^C]] - \frac{1}{2} [\rho_2, [\rho_2, h_{12}^C]] + [\rho_1, [\rho_2, h_{12}^B]_{++}] \end{aligned} \quad (323)$$

which we shall take as our final effective two-electron operator. However, from the structure of the transformed Breit Hamiltonian

$$\tilde{\mathcal{K}}^B = \begin{pmatrix} 2m_e c^2 + \mathcal{O}(c^0) & \mathcal{O}(c^{-1}) & \mathcal{O}(c^{-1}) & \mathcal{O}(c^0) \\ \mathcal{O}(c^{-1}) & \mathcal{O}(c^0) & \mathcal{O}(c^0) & \mathcal{O}(c^{-1}) \\ \mathcal{O}(c^{-1}) & \mathcal{O}(c^0) & \mathcal{O}(c^0) & \mathcal{O}(c^{-1}) \\ \mathcal{O}(c^0) & \mathcal{O}(c^{-1}) & \mathcal{O}(c^{-1}) & -2m_e c^2 + \mathcal{O}(c^0) \end{pmatrix} \quad (324)$$

we suspect that further transformations may be necessary to decouple the electronic solutions from the positronic and mixed electronic–positronic solutions to order  $c^{-2}$ , noting that there are still coupling elements of order  $c^{-1}$  and  $c^0$ . We shall not attempt a further Foldy–Wouthuysen decoupling here, noting that its only effect is to introduce an additional Breit–Pauli term quadratic in the Breit operator

$$H_{\text{BP}}^{\text{c}} = \frac{h_{12}^{\text{B}}{}^2}{4m_e c^2} \quad (325)$$

which is usually neglected because the Breit operator itself has been obtained from perturbation theory.

In the Breit–Pauli Hamiltonian, the one-electron part is identical to the first-order relativistic Hamiltonian in Eq. (274), which was developed in Section VII B and discussed in Section VII C. In the following subsections, we shall consider in detail the reduced two-electron Coulomb, Gaunt, and retardation parts of the Breit–Pauli Hamiltonian Eq. (323).

#### D. Reduction of the two-electron Coulomb operator

The first-order Coulomb contribution to the Breit–Pauli Hamiltonian Eq. (323) is given by

$$H_{\text{C}} = -\frac{1}{2} \sum_{i=1}^2 [\rho_i, [\rho_i, h_{12}^{\text{C}}]] = -\frac{e^2}{32\pi\epsilon_0 m_e^2 c^2} \sum_{i=1}^2 [\boldsymbol{\sigma}_i \cdot \boldsymbol{\pi}_i, [\boldsymbol{\sigma}_i \cdot \boldsymbol{\pi}_i, r_{12}^{-1}]] \quad (326)$$

where we have inserted Eqs. (264) and (301). In the Einstein summation convention, the double commutator may be evaluated as

$$[\boldsymbol{\sigma}_1 \cdot \boldsymbol{\pi}_1, [\boldsymbol{\sigma}_1 \cdot \boldsymbol{\pi}_1, r_{12}^{-1}]] = \sigma_{1i} \sigma_{1j} [\pi_{1i}, [\pi_{1j}, r_{12}^{-1}]] + [\sigma_{1i}, \sigma_{1j}] [\pi_{1j}, r_{12}^{-1}] \pi_{1i} \quad (327)$$

With the help of the relation Eq. (191), the first and second terms become

$$\begin{aligned} \sigma_{1i} \sigma_{1j} [\pi_{1i}, [\pi_{1j}, r_{12}^{-1}]] &= -\hbar^2 (\delta_{ij} + i\epsilon_{ijk} \sigma_{1k}) (\nabla_{1i} \nabla_{1j} r_{12}^{-1}) = -\hbar^2 (\nabla_1^2 r_{12}^{-1}) \\ &= 4\pi \hbar^2 \delta(\mathbf{r}_{12}) \end{aligned} \quad (328)$$

$$\begin{aligned} [\sigma_{1i}, \sigma_{1j}] [\pi_{1j}, r_{12}^{-1}] \pi_{1i} &= 2i\epsilon_{ijk} \sigma_{1k} [\pi_{1j}, r_{12}^{-1}] \pi_{1i} = 2\hbar\epsilon_{ijk} \sigma_{1k} (\nabla_{1j} r_{12}^{-1}) \pi_{1i} \\ &= -2\hbar\epsilon_{ijk} \sigma_{1k} r_{12j} r_{12}^{-3} \pi_{1i} \end{aligned} \quad (329)$$

Inserting this result in Eq. (326), we obtain the following first-order relativistic correction to the Coulomb interaction:

$$H_{\text{C}} = -\frac{\hbar^2 e^2}{4\epsilon_0 m_e^2 c^2} \delta(\mathbf{r}_{12}) - \frac{e^2}{8\pi\epsilon_0 m_e^2 c^2} \left( \frac{\mathbf{s}_1 \cdot \mathbf{r}_{12} \times \boldsymbol{\pi}_1}{r_{12}^3} + \frac{\mathbf{s}_2 \cdot \mathbf{r}_{21} \times \boldsymbol{\pi}_2}{r_{12}^3} \right) \quad (330)$$

Here the first term is the two-electron Darwin term and the second term the two-electron spin-orbit operator. Note that these terms arise in the same manner as the Darwin and spin-orbit corrections to the one-electron Coulomb operator in the presence of point-charge nuclei, as discussed in Section VII C.

### E. Reduction of the Gaunt operator

Inserting Eqs. (304) and (264) in the Breit-Pauli Hamiltonian Eq. (323), we obtain the first-order Gaunt correction to the nonrelativistic electronic Hamiltonian:

$$H_G = [\rho_1, [\rho_2, h_{12}^G]_+]_+ = -\frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} [\boldsymbol{\sigma}_1 \cdot \boldsymbol{\pi}_1, [\boldsymbol{\sigma}_2 \cdot \boldsymbol{\pi}_2, \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 r_{12}^{-1}]_+]_+ \quad (331)$$

Upon expansion of the anticommutators, we obtain

$$H_G = -\frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} (\sigma_{1i}\sigma_{1k}\sigma_{2j}\sigma_{2k}\pi_{1i}\pi_{2j}r_{12}^{-1} + \sigma_{1i}\sigma_{1k}\sigma_{2k}\sigma_{2j}\pi_{1i}r_{12}^{-1}\pi_{2j} + \sigma_{1k}\sigma_{1i}\sigma_{2j}\sigma_{2k}\pi_{2j}r_{12}^{-1}\pi_{1i} + \sigma_{1k}\sigma_{1i}\sigma_{2k}\sigma_{2j}r_{12}^{-1}\pi_{1i}\pi_{2j}) \quad (332)$$

where each term involves two pairs of Pauli spin operators, one for each electron. Using Eq. (191), each such pair expands into a scalar (spin-free) term and a spin term. There will therefore be three distinct contributions to  $H_G$ , involving none, one, or two spin operators, respectively:

$$H_G = H_G^{\text{OO}} + H_G^{\text{SoO}} + H_G^{\text{SS}} \quad (333)$$

First, for the spin-free orbit-orbit operator, we obtain

$$\begin{aligned} H_G^{\text{OO}} &= -\frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} \delta_{ik}\delta_{jk} (\pi_{1i}\pi_{2j}r_{12}^{-1} + \pi_{1i}r_{12}^{-1}\pi_{2j} + \pi_{2j}r_{12}^{-1}\pi_{1i} + r_{12}^{-1}\pi_{1i}\pi_{2j}) \\ &= -\frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} (2\pi_{1i}r_{12}^{-1}\pi_{2i} + 2\pi_{2i}r_{12}^{-1}\pi_{1i} + [\pi_{1i}, [\pi_{2i}, r_{12}^{-1}]]) \end{aligned} \quad (334)$$

Since  $[\pi_{1i}, [\pi_{2i}, r_{12}^{-1}]] = \hbar^2(\nabla_{1i}\nabla_{2i}r_{12}^{-1}) = -\hbar^2\nabla_1^2 r_{12}^{-1}$ , this operator is equivalent to

$$H_G^{\text{OO}} = -\frac{e^2}{8\pi\epsilon_0 m_e^2 c^2} (\boldsymbol{\pi}_1 \cdot r_{12}^{-1} \boldsymbol{\pi}_2 + \boldsymbol{\pi}_2 \cdot r_{12}^{-1} \boldsymbol{\pi}_1) + \frac{e^2 \hbar^2}{4\epsilon_0 m_e^2 c^2} \delta(\mathbf{r}_{12}) \quad (335)$$

We shall later see that there is an additional contribution to the orbit-orbit operator from the retardation part of the Breit operator  $H_R$ . However, we shall first treat the other two contributions from the Gaunt operator: the spin-other-orbit and the spin-spin operators.

To obtain the spin-other-orbit operator, we retain all terms in Eq. (332) that contain one spin operator upon expansion of Eq. (191):

$$\begin{aligned}
H_G^{\text{SoO}} &= -\frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} \left[ \delta_{ik} i \epsilon_{jkl} \sigma_{2l} (\pi_{1i} \pi_{2j} r_{12}^{-1} - \pi_{1i} r_{12}^{-1} \pi_{2j} + \pi_{2j} r_{12}^{-1} \pi_{1i} - r_{12}^{-1} \pi_{1i} \pi_{2j}) \right. \\
&\quad \left. + \delta_{jk} i \epsilon_{ikl} \sigma_{1l} (\pi_{1i} \pi_{2j} r_{12}^{-1} + \pi_{1i} r_{12}^{-1} \pi_{2j} - \pi_{2j} r_{12}^{-1} \pi_{1i} - r_{12}^{-1} \pi_{1i} \pi_{2j}) \right] \\
&= -\frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} \left[ i \epsilon_{jil} \sigma_{2l} \left( [\pi_{1i}, [\pi_{2j}, r_{12}^{-1}]] + 2 [\pi_{2j}, r_{12}^{-1}] \pi_{1i} \right) \right. \\
&\quad \left. + i \epsilon_{ijl} \sigma_{1l} \left( [\pi_{2j}, [\pi_{1i}, r_{12}^{-1}]] + 2 [\pi_{1i}, r_{12}^{-1}] \pi_{2j} \right) \right] \quad (336)
\end{aligned}$$

Since  $[\pi_{1i}, [\pi_{2j}, r_{12}^{-1}]]$  is symmetric in  $i$  and  $j$ , whereas  $\epsilon_{ijl}$  is antisymmetric in  $i, j$ , and  $k$ , the first term in Eq. (336) vanishes. The third term vanishes for similar reasons, giving

$$H_G^{\text{SoO}} = -\frac{e^2 \hbar}{8\pi\epsilon_0 m_e^2 c^2} \left[ \epsilon_{jil} \sigma_{2l} (\nabla_{2j} r_{12}^{-1}) \pi_{1i} + \epsilon_{ijl} \sigma_{1l} (\nabla_{1i} r_{12}^{-1}) \pi_{2j} \right] \quad (337)$$

which yields our final expression for the spin-other-orbit operator:

$$H_G^{\text{SoO}} = -\frac{e^2}{4\pi\epsilon_0 m_e^2 c^2} \left[ \frac{\mathbf{s}_1 \cdot \mathbf{r}_{21} \times \boldsymbol{\pi}_2}{r_{12}^3} + \frac{\mathbf{s}_2 \cdot \mathbf{r}_{12} \times \boldsymbol{\pi}_1}{r_{12}^3} \right] \quad (338)$$

Unlike for the orbit-orbit operator, there is no retardation contribution to the spin-other-orbit operator. Finally, to obtain the spin-spin operator from Eq. (332), we collect all terms containing two spin operators:

$$\begin{aligned}
H_G^{\text{SS}} &= \frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} \epsilon_{ikm} \epsilon_{jkn} \sigma_{1m} \sigma_{2n} (\pi_{1i} \pi_{2j} r_{12}^{-1} - \pi_{1i} r_{12}^{-1} \pi_{2j} - \pi_{2j} r_{12}^{-1} \pi_{1i} + r_{12}^{-1} \pi_{1i} \pi_{2j}) \\
&= \frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} (\delta_{ij} \delta_{mn} - \delta_{in} \delta_{jm}) \sigma_{1m} \sigma_{2n} [\pi_{1i}, [\pi_{2j}, r_{12}^{-1}]] \\
&= \frac{e^2 \hbar^2}{16\pi\epsilon_0 m_e^2 c^2} \left[ \sigma_{1n} \sigma_{2n} (\nabla_{1i} \nabla_{1i} r_{12}^{-1}) - \sigma_{1j} \sigma_{2i} (\nabla_{1i} \nabla_{1j} r_{12}^{-1}) \right] \quad (339)
\end{aligned}$$

Invoking the identity

$$(\nabla_{1i} \nabla_{1j} r_{12}^{-1}) = -\frac{4\pi}{3} \delta_{ij} \delta(\mathbf{r}_{12}) - \frac{\delta_{ij} r_{12}^2 - 3r_{12i} r_{12j}}{r_{12}^5} \quad (340)$$

we arrive the final expression for the spin-spin operator (there is no retardation correction):

$$H_G^{\text{SS}} = \frac{e^2}{4\pi\epsilon_0 m_e^2 c^2} \left[ \frac{(\mathbf{s}_1 \cdot \mathbf{s}_2) r_{12}^2 - 3(\mathbf{s}_1 \cdot \mathbf{r}_{12})(\mathbf{r}_{12} \cdot \mathbf{s}_2)}{r_{12}^5} - \frac{8\pi}{3} \delta(\mathbf{r}_{12})(\mathbf{s}_1 \cdot \mathbf{s}_2) \right] \quad (341)$$

Note that there are two distinct contributions to the spin-spin interaction: a classical dipolar interaction and a contact interaction, which only contributes for coinciding electrons.

## F. Reduction of the retardation operator

Finally, turning our attention to the retardation part of the Breit–Pauli Hamiltonian Eq. (305), we obtain from Eqs. (264) and (323)

$$H_R = [\rho_1, [\rho_2, h_{12}^R]_+]_+ = -\frac{e^2}{32\pi\epsilon_0 m_e^2 c^2} [\boldsymbol{\sigma}_1 \cdot \boldsymbol{\pi}_1, [\boldsymbol{\sigma}_2 \cdot \boldsymbol{\pi}_2, (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\nabla}_1) (\boldsymbol{\sigma}_2 \cdot \boldsymbol{\nabla}_2) r_{12}]_+]_+ \quad (342)$$

where we have used the identity

$$\nabla_{1i} \nabla_{2j} r_{12} = -\delta_{ij} r_{12}^{-1} + r_{12i} r_{12j} r_{12}^{-3} \quad (343)$$

Upon expansion of the anticommutators, we obtain

$$H_R = -\frac{e^2}{32\pi\epsilon_0 m_e^2 c^2} \sigma_{1i} \sigma_{1k} \sigma_{2j} \sigma_{2l} [\pi_{1i} \pi_{2j} (\nabla_{1k} \nabla_{2l} r_{12}) + \pi_{1i} (\nabla_{1k} \nabla_{2j} r_{12}) \pi_{2l} + \pi_{2j} (\nabla_{1i} \nabla_{2l} r_{12}) \pi_{1k} + (\nabla_{1i} \nabla_{2j} r_{12}) \pi_{1k} \pi_{2l}] \quad (344)$$

Although not obvious, the sum of the four terms in brackets is symmetric with respect to permutations of  $i$  and  $k$  and with respect to permutations of  $j$  and  $l$ . To see this, we expand the terms as

$$\begin{aligned} \pi_{1i} \pi_{2j} (\nabla_{1k} \nabla_{2l} r_{12}) &= -\hbar^2 (\nabla_{1i} \nabla_{1k} \nabla_{2j} \nabla_{2l} r_{12}) - i\hbar (\nabla_{1k} \nabla_{2j} \nabla_{2l} r_{12}) \pi_{1i} \\ &\quad - i\hbar (\nabla_{1i} \nabla_{1k} \nabla_{2l} r_{12}) \pi_{2j} + (\nabla_{1k} \nabla_{2l} r_{12}) \pi_{1i} \pi_{2j} \end{aligned} \quad (345)$$

$$\pi_{1i} (\nabla_{1k} \nabla_{2j} r_{12}) \pi_{2l} = -i\hbar (\nabla_{1i} \nabla_{1k} \nabla_{2j} r_{12}) \pi_{2l} + (\nabla_{1k} \nabla_{2j} r_{12}) \pi_{1i} \pi_{2l} \quad (346)$$

$$\pi_{2j} (\nabla_{1i} \nabla_{2l} r_{12}) \pi_{1k} = -i\hbar (\nabla_{1i} \nabla_{2j} \nabla_{2l} r_{12}) \pi_{1k} + (\nabla_{1i} \nabla_{2l} r_{12}) \pi_{2j} \pi_{1k} \quad (347)$$

Upon substitution of these expansions into Eq. (344), it is easy to verify that the factor in brackets is symmetric in  $i$  and  $k$  and symmetric in  $j$  and  $l$ . Only the scalar part of Eq. (191) therefore contributes to  $H_R$ , which therefore does not contribute to the spin–other-orbit and spin–spin operators. After a little algebra, we obtain

$$H_R = -\frac{e^2}{32\pi\epsilon_0 m_e^2 c^2} [2\pi_{1i} (\nabla_{1i} \nabla_{2j} r_{12}) \pi_{2j} + 2\pi_{2j} (\nabla_{1i} \nabla_{2j} r_{12}) \pi_{1i} + [\pi_{1i}, [\pi_{2j}, (\nabla_{1i} \nabla_{2j} r_{12})]]] \quad (348)$$

Invoking Eq. (343) and

$$\nabla_1^2 \nabla_2^2 r_{12} = 2\nabla_1^2 r_{12}^{-1} = -8\pi\delta(\mathbf{r}_{12}) \quad (349)$$

we find that the Breit–Pauli retardation operator may be written in the form

$$\begin{aligned}
H_R = & \frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} (\boldsymbol{\pi}_1 \cdot r_{12}^{-1} \boldsymbol{\pi}_2 - \boldsymbol{\pi}_1 \cdot \mathbf{r}_{12} r_{12}^{-3} \mathbf{r}_{12} \cdot \boldsymbol{\pi}_2) \\
& + \frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} (\boldsymbol{\pi}_2 \cdot r_{12}^{-1} \boldsymbol{\pi}_1 - \boldsymbol{\pi}_2 \cdot \mathbf{r}_{12} r_{12}^{-3} \mathbf{r}_{12} \cdot \boldsymbol{\pi}_1) - \frac{e^2 \hbar^2}{4\epsilon_0 m_e^2 c^2} \delta(\mathbf{r}_{12})
\end{aligned} \tag{350}$$

Adding this operator to  $H_G^{\text{OO}}$ , we obtain the full orbit–orbit operator

$$\begin{aligned}
H^{\text{OO}} = & - \frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} (\boldsymbol{\pi}_1 \cdot r_{12}^{-1} \boldsymbol{\pi}_2 + \boldsymbol{\pi}_1 \cdot \mathbf{r}_{12} r_{12}^{-3} \mathbf{r}_{12} \cdot \boldsymbol{\pi}_2) \\
& - \frac{e^2}{16\pi\epsilon_0 m_e^2 c^2} (\boldsymbol{\pi}_2 \cdot r_{12}^{-1} \boldsymbol{\pi}_1 + \boldsymbol{\pi}_2 \cdot \mathbf{r}_{12} r_{12}^{-3} \mathbf{r}_{12} \cdot \boldsymbol{\pi}_1)
\end{aligned} \tag{351}$$

in which the contact contributions from the Gaunt and retardation operators cancel.

## IX. MANY-ELECTRON SYSTEMS AND THE BREIT-PAULI HAMILTONIAN

To construct a relativistic Hamiltonian for many electrons, we take the sum of one-electron Dirac Hamiltonians Eq. (291) and add pairwise Coulomb and Breit (Gaunt and retardation) interactions Eqs. (292)–(294), obtaining in atomic units the Dirac–Breit Hamiltonian

$$H^{\text{DB}} = \sum_i (c \boldsymbol{\alpha}_i \cdot \boldsymbol{\pi}_i - \phi_i + \beta_i c^2) + \frac{1}{2} \sum_{i \neq j} \frac{1}{r_{ij}} - \frac{1}{4} \sum_{i \neq j} \frac{r_{ij}^2 \boldsymbol{\alpha}_i \cdot \boldsymbol{\alpha}_j + \boldsymbol{\alpha}_i \cdot \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \boldsymbol{\alpha}_j}{r_{ij}^3} \tag{352}$$

Here  $\phi = \phi(\mathbf{r}_i)$  is the scalar potential at  $\mathbf{r}_i$  and we have introduced the kinetic momentum

$$\boldsymbol{\pi}_i = -i \boldsymbol{\nabla}_i + \mathbf{A}_i, \tag{353}$$

where  $\boldsymbol{\nabla}_i$  differentiates with respect to  $\mathbf{r}_i$  and  $\mathbf{A}_i = \mathbf{A}(\mathbf{r}_i)$ . In a similar manner, we adopt the notation  $\mathbf{B}_i = \mathbf{B}(\mathbf{r}_i)$  and  $\mathbf{E}_i = \mathbf{E}(\mathbf{r}_i)$ . Reducing this Hamiltonian to two-component form, we obtain the many-electron Breit–Pauli operator

$$\begin{aligned}
H^{\text{BP}} = & H^{\text{NR}} + H_{\text{K}}^{\text{MV}} + H_{\text{Z}}^{\text{MV}} + H_{\text{C}}^{\text{DW}} + H_{\text{C}}^{\text{SO}} \\
& + H_{\text{G}}^{\text{SoO}} + H_{\text{G}}^{\text{SS}} + H_{\text{G}}^{\text{OO}} + H_{\text{R}}^{\text{OO}} + \mathcal{O}(\alpha^4)
\end{aligned} \tag{354}$$

where the different contributions are listed in Fig. 2. The one-electron mass-velocity, Darwin and spin–orbit corrections of Eqs. (B1.2)–(B1.5) were treated in Section VII C and will not be discussed further here. The two-electron relativistic corrections are of three kinds: Coulomb, Gaunt, and retardation. The Coulomb corrections are analogous to the one-electron Darwin

$$H^{\text{NR}} = \frac{1}{2} \sum_i \pi_i^2 + \sum_i \mathbf{B}_i \cdot \mathbf{s}_i - \sum_i \phi_i + \frac{1}{2} \sum_{i \neq j} r_{ij}^{-1} \quad \text{nonrelativistic} \quad (\text{B1.1})$$

$$H_{\text{K}}^{\text{MV}} = -\frac{\alpha^2}{8} \sum_i \pi_i^4 \quad \text{mass-velocity} \quad (\text{B1.2})$$

$$H_{\text{Z}}^{\text{MV}} = -\frac{\alpha^2}{4} \sum_i (\pi_i^2 \mathbf{B}_i \cdot \mathbf{s}_i + \mathbf{B}_i \cdot \mathbf{s}_i \pi_i^2) - \frac{\alpha^2}{2} \sum_i (\mathbf{B}_i \cdot \mathbf{s}_i)^2 \quad \text{Zeeman mass-velocity} \quad (\text{B1.3})$$

$$H_{\text{C}}^{\text{DW}} = \frac{\alpha^2}{8} \sum_i (\nabla_i \cdot \mathbf{E}_i) - \frac{\alpha^2 \pi}{2} \sum_{i \neq j} \delta(\mathbf{r}_{ij}) \quad \text{Darwin} \quad (\text{B1.4})$$

$$H_{\text{C}}^{\text{SO}} = \frac{\alpha^2}{4} \sum_i \mathbf{s}_i \cdot (\mathbf{E}_i \times \boldsymbol{\pi}_i - \boldsymbol{\pi}_i \times \mathbf{E}_i) - \frac{\alpha^2}{2} \sum_{i \neq j} \frac{\mathbf{s}_i \cdot \mathbf{r}_{ij} \times \boldsymbol{\pi}_i}{r_{ij}^3} \quad \text{spin-orbit} \quad (\text{B1.5})$$

$$H_{\text{G}}^{\text{SoO}} = -\alpha^2 \sum_{i \neq j} \frac{\mathbf{s}_j \cdot \mathbf{r}_{ij} \times \boldsymbol{\pi}_i}{r_{ij}^3} \quad \text{spin-other-orbit} \quad (\text{B1.6})$$

$$H_{\text{G}}^{\text{SS}} = \frac{\alpha^2}{2} \sum_{i \neq j} \frac{r_{ij}^2 \mathbf{s}_i \cdot \mathbf{s}_j - 3 \mathbf{s}_i \cdot \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \mathbf{s}_j}{r_{ij}^5} - \frac{4\alpha^2 \pi}{3} \sum_{i \neq j} \delta(\mathbf{r}_{ij}) \mathbf{s}_i \cdot \mathbf{s}_j \quad \text{spin-spin} \quad (\text{B1.7})$$

$$H_{\text{G}}^{\text{OO}} = -\frac{\alpha^2}{2} \sum_{i \neq j} \frac{\boldsymbol{\pi}_i \cdot \boldsymbol{\pi}_j}{r_{ij}} + \frac{\alpha^2 \pi}{2} \sum_{i \neq j} \delta(\mathbf{r}_{ij}) \quad \text{Gaunt orbit-orbit} \quad (\text{B1.8})$$

$$H_{\text{R}}^{\text{OO}} = \frac{\alpha^2}{4} \sum_{i \neq j} \frac{\boldsymbol{\pi}_i \cdot r_{ij}^2 \boldsymbol{\pi}_j - \boldsymbol{\pi}_i \cdot \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \boldsymbol{\pi}_j}{r_{ij}^3} - \frac{\alpha^2 \pi}{2} \sum_{i \neq j} \delta(\mathbf{r}_{ij}) \quad \text{retardation orbit-orbit} \quad (\text{B1.9})$$

FIG. 2: The contributions to the Breit–Pauli Hamiltonian Eq. (354)

and spin–orbit corrections. Thus, the negative two-electron Darwin correction in Eq. (B1.4) reduces the two-electron nonrelativistic interaction and arises because of the charge smearing due to Zitterbewegung, whereas the two-electron spin–orbit operator in Eq. (B1.5) couples the spin of an electron to its orbital motion in the presence of the electric field set up by the other electrons.

The Gaunt operators represent the leading magnetic interactions between the electrons, arising from their orbital motion and spin. Thus, spin-other-orbit operator  $H_{\text{G}}^{\text{SoO}}$  of Eq. (B1.6) differs from the two-electron spin–orbit operator in  $H_{\text{C}}^{\text{SO}}$  of Eq. (B1.5) in that it does not couple the spin of one electron with its own orbital motion but rather with that of another electron. It may be interpreted as the Zeeman interaction of the magnetic moment of one electron with the magnetic induction generated by the orbital motion of another electron. Alternatively, it may be viewed as the paramagnetic interaction of the motion of one electron with the vector potential set up by the magnetic moment of another

electron. The two-electron spin–spin operator  $H_G^{SS}$  of Eq. (B1.7) may likewise be understood as the Zeeman interaction of the magnetic moment of one electron with the magnetic field generated by the spin magnetic moment of another electron. Note that there are two contributions to the spin–spin coupling in Eq. (B1.7). One contribution is the dipolar spin–spin interaction, representing the interaction between two separated dipole moments; the other contribution is the *Fermi contact interaction*, which arises when the two electrons coincide in space. The final Gaunt term, the orbit–orbit operator  $H_G^{OO}$  in Eq. (B1.8), arises because of the paramagnetic interaction of the motion of one electron with the magnetic induction generated by the motion of another electron.

Finally, the two-electron retardation operator  $H_R^{OO}$  in Eq. (B1.9) represents a correction to the instantaneous two-electron Coulomb interactions, arising since these interactions propagate with a finite rather than infinite velocity. It is usually combined with the Gaunt orbit–orbit interaction in Eq. (B1.8), yielding the total orbit–orbit operator  $H^{OO} = -\frac{\alpha^2}{4} \sum_{i \neq j} \frac{\boldsymbol{\pi}_i \cdot r_{ij}^2 \boldsymbol{\pi}_j + \boldsymbol{\pi}_i \cdot \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \boldsymbol{\pi}_j}{r_{ij}^3}$  in which there is no contact contribution. To higher orders in  $\alpha^2$ , there will further retardation corrections, including corrections to the magnetic Gaunt interactions. However, to orders higher than  $\alpha^2$ , corrections such as the Lamb shift, the leading quantum electrodynamical correction (of order  $\alpha^3$ ), become important for molecules containing light elements.

## X. THE MOLECULAR ELECTRONIC BREIT–PAULI HAMILTONIAN

In Section IX, we presented the Hamiltonian in Eq. (354) for electrons in an electromagnetic field. This Hamiltonian was obtained by reducing the relativistic Dirac–Breit Hamiltonian to nonrelativistic form and is accurate to second order in the fine-structure constant. For this Hamiltonian to be useful in molecular calculations, we must now consider the electromagnetic field experienced by electrons in molecules.

This electromagnetic field has two distinct contributions. By far the most important contribution is generated internally, from the nuclei inside the molecule. In addition, there is an external contribution, arising from sources outside the molecule, generated by neighboring molecules or applied macroscopically. Accordingly, we decompose the scalar and vector potentials entering the electronic Breit–Pauli Hamiltonian Eq. (354) into external and nuclear



contributions:

$$\phi(\mathbf{r}) = \phi_{\text{nuc}}(\mathbf{r}) + \phi_{\text{ext}}(\mathbf{r}) \quad (355)$$

$$\mathbf{A}(\mathbf{r}) = \mathbf{A}_{\text{nuc}}(\mathbf{r}) + \mathbf{A}_{\text{ext}}(\mathbf{r}) \quad (356)$$

In Section X A, we consider the nuclear potentials; next, in Section X B, we introduce the external potentials. After a brief discussion of the purely nuclear contributions to the molecular Hamiltonian in Section X C, we present in Section X D the full molecular Breit–Pauli Hamiltonian, appropriate for molecular electronic calculations in the Born–Oppenheimer approximation, obtained by substituting the potentials Eqs. (355) and (356) and their associated electromagnetic fields into the electronic Hamiltonian Eq. (354).

### A. Nuclear electromagnetic fields

In the Born–Oppenheimer approximation, we treat the atomic nuclei as a framework of stationary particles, characterized by their *charges*  $Z_K$ , their *quadrupole moments*  $\mathbf{Q}_K$ , and their *magnetic moments*  $\mathbf{M}_K$ . In addition, the nuclei have a finite size, characterized by their radial extent  $\mathcal{R}_K$ , which we here relate to the second moment of the nuclear charge distribution  $\rho_K(\mathbf{r})$ :

$$Z_K \mathcal{R}_K^2 = \int (x^2 + y^2 + z^2) \rho_K(\mathbf{r}) \, d\mathbf{r} \quad (357)$$

From the discussion in Sections III G and III H and, in particular, Eqs. (81) and (99), we arrive at the following multipole expansions of the nuclear contributions to the potentials in Eqs. (355) and (356) (in atomic units):

$$\phi_{\text{nuc}}(\mathbf{r}) = \sum_K \frac{Z_K}{r_K} - \frac{2\pi}{3} \sum_K Z_K \mathcal{R}_K^2 \delta(\mathbf{r}_K) + \frac{1}{6} \sum_K \frac{\text{tr} \mathbf{Q}_K (3\mathbf{r}_K^T \mathbf{r}_K - r_K^2 \mathbf{I}_3)}{r_K^5} \quad (358)$$

$$\mathbf{A}_{\text{nuc}}(\mathbf{r}) = \alpha^2 \sum_K \frac{\mathbf{M}_K \times \mathbf{r}_K}{r_K^3} \quad (359)$$

where we have ignored higher-order moments as well as all nonelectromagnetic interactions such as those arising from parity violation. To leading orders, the corresponding nuclear electric and magnetic fields are given by

$$\mathbf{E}_{\text{nuc}}(\mathbf{r}) = \sum_K \frac{Z_K \mathbf{r}_K}{r_K^3} \quad (360)$$

$$\mathbf{B}_{\text{nuc}}(\mathbf{r}) = \alpha^2 \sum_K \frac{3(\mathbf{M}_K \cdot \mathbf{r}_K) \mathbf{r}_K - r_K^2 \mathbf{M}_K}{r_K^5} + \frac{8\pi}{3} \alpha^2 \sum_K \delta(\mathbf{r}_K) \mathbf{M}_K \quad (361)$$

which have been obtained as  $\mathbf{E}_{\text{nuc}} = -\nabla\phi_{\text{nuc}}$  and  $\mathbf{B}_{\text{nuc}} = \nabla \times \mathbf{A}_{\text{nuc}}$  using the vector relations in Appendix A. For symmetry (parity) reasons, there are no odd-order electric moments and no even-order magnetic moments in the expansions of the nuclear potentials.

Among the different contributions to the nuclear potentials Eqs. (358) and (359), by far the largest is the electrostatic contribution from the nuclear charge  $Z_K$ —indeed, the resulting nuclear point-charge potential is responsible for generating stable molecular structures. The remaining contributions to the potentials are orders of magnitude smaller, representing small corrections to the point-charge nuclear model and being responsible for the *hyperfine interactions* in atoms and molecules.

The contact interaction in Eq. (358) depends on the size of the nucleus as measured by  $\mathcal{R}_K^2$  in Eq. (357), reducing the potential at the nucleus and increasing the total energy. Since nuclear sizes are of the order of 1 fm, the correction is small, at least for light nuclei. The finite nuclear size may alternatively be corrected for by representing the nucleus by a Gaussian charge distribution

$$\rho_K(\mathbf{r}_K) = Z_K \left(\frac{\eta_K}{\pi}\right)^{3/2} \exp(-\eta_K r_K^2) \quad (362)$$

where  $\eta_K = (3/2)\mathcal{R}_K^{-2}$ . Gaussian distributions or linear combinations of such distributions are used since they simplify the calculation of integrals over Gaussian atomic orbitals. Integrated over all space, the potential from the nucleus becomes

$$\int \frac{\rho_K(\mathbf{r}_K)}{r_K} d\mathbf{r} = \frac{Z_K \operatorname{erf}(\sqrt{\eta_K} r_K)}{r_K} \quad (363)$$

where the monotonically increasing error function  $0 \leq \operatorname{erf}(x) \leq 1$  of Eq. (9.8.22) reduces potential at the nucleus.

The magnetic moments  $\mathbf{M}_K$  arise from the nuclear spins  $\mathbf{I}_K$ , to which they are related as

$$\mathbf{M}_K = g_K \mu_N \mathbf{I}_K = \gamma_K \hbar \mathbf{I}_K \quad (364)$$

where  $\mu_N = e\hbar/2m_p$  is the *nuclear magneton*,  $g_K$  is the *nuclear g value*, and  $\gamma_K$  the *magnetogyric ratio* of the nucleus. The nuclear magneton is numerically small, about  $2.7 \cdot 10^{-4}$  in atomic units. The nuclear  $g$  value  $g_K$  is a dimensionless empirical constant, different for each nucleus. It is of order unity (its absolute value never exceeds six) and may be positive or negative. The nuclear magnetic moments are therefore roughly three orders of magnitude smaller than the magnetic moment  $\mathbf{m}$  of the electron Eq. (199). Only nuclei with spin 1/2 or greater have a nonvanishing magnetic dipole moment.

## B. External electromagnetic fields

The external contributions to the potentials Eq. (355) and (356) may have a variety of sources. We here consider the important special cases of externally applied uniform electrostatic and magnetostatic fields  $\mathbf{E}$  and  $\mathbf{B}$ . From the vector relations in Appendix A, it is easily verified that such fields are represented by the following scalar and vector potentials:

$$\phi_{\text{ext}}(\mathbf{r}) = -\mathbf{E} \cdot \mathbf{r}_O \quad (365)$$

$$\mathbf{A}_{\text{ext}}(\mathbf{r}) = \frac{1}{2}\mathbf{B} \times \mathbf{r}_O \quad (366)$$

where  $\mathbf{r}_O = \mathbf{r} - \mathbf{O}$  is the position of the electron relative to some arbitrary origin  $\mathbf{O}$ . It is instructive to see how these potentials contribute to the nonrelativistic two-component electronic Hamiltonian  $H_{2c}$  of Eq. (198). In atomic units, we obtain from Eqs. (A1) and (A3)

$$-\phi_{\text{ext}}(\mathbf{r}) = \mathbf{E} \cdot \mathbf{r}_O = -\mathbf{E} \cdot \mathbf{p}_e, \quad (367)$$

$$\mathbf{A}_{\text{ext}}(\mathbf{r}) \cdot \mathbf{p} + \mathbf{B} \cdot \mathbf{s} = \mathbf{B} \cdot \left(\frac{1}{2}\mathbf{r}_O \times \mathbf{p} + \mathbf{s}\right) = -\mathbf{B} \cdot \mathbf{m}_e \quad (368)$$

$$\frac{1}{2}A_{\text{ext}}^2 = \frac{1}{8}(\mathbf{B} \times \mathbf{r}_O)^2 = \frac{1}{8}[B^2 r_O^2 - (\mathbf{B} \cdot \mathbf{r}_O)^2] \quad (369)$$

We have here introduced the *electric and magnetic dipole-moment operators* of the electron

$$\mathbf{p}_e = -\mathbf{r}_O \quad (370)$$

$$\mathbf{m}_e = -\frac{1}{2}\mathbf{l}_O - \mathbf{s} \quad (371)$$

where  $\mathbf{l}_O = \mathbf{r}_O \times \mathbf{p}$  is orbital angular momentum of the electron relative to  $\mathbf{O}$ . Substituting these results into Eq. (198), we obtain the following Hamiltonian

$$H_{2c} = \frac{1}{2}p^2 - \mathbf{E} \cdot \mathbf{p}_e - \mathbf{B} \cdot \mathbf{m}_e + \frac{1}{8}[B^2 r_O^2 - (\mathbf{B} \cdot \mathbf{r}_O)^2] \quad (372)$$

for an electron in a uniform, static electromagnetic field. From Section VII D, we recall that the externally applied fields are usually small, typically of the order of  $\alpha^2$ .

Although the electric and magnetic dipole operators Eqs. (370) and (371) depend on  $\mathbf{O}$ , this dependence does not affect observable quantities since a change from  $\mathbf{O}$  to  $\mathbf{O}'$  merely corresponds to a gauge transformation, represented by the gauge function

$$f(\mathbf{r}, t) = -\mathbf{E} \cdot (\mathbf{O}' - \mathbf{O}) t - \frac{1}{2}\mathbf{B} \times (\mathbf{O}' - \mathbf{O}) \cdot \mathbf{r}. \quad (373)$$

In particular, although the total energy, in agreement with the classical result Eq. (129), does change as a result of the change in the scalar potential by  $\mathbf{E} \cdot (\mathbf{O}' - \mathbf{O})$ , this change only affects the energy scale—all energy differences are unaffected. By contrast, the corresponding change in the vector potential  $\frac{1}{2}\mathbf{B} \times (\mathbf{O}' - \mathbf{O})$  does not affect the energy scale, in agreement with the observation that  $\mathbf{B}$  does no work on the system since it generates a force perpendicular to the velocity of the particle Eq. (3).

### C. Purely nuclear contributions

To arrive at the final Born–Oppenheimer electronic Hamiltonian, we must also add the purely nuclear interactions. These we take to be of the form

$$\begin{aligned}
H^{\text{nuc}} = & \frac{1}{2} \sum_{K \neq L} \frac{Z_K Z_L}{R_{KL}} + \sum_K Z_K \phi_K - \sum_K \mathbf{B} \cdot \mathbf{M}_K \\
& + \frac{\alpha^2}{2} \sum_{K \neq L} \frac{R_{KL}^2 (\mathbf{M}_K \cdot \mathbf{M}_L) - 3 (\mathbf{M}_K \cdot \mathbf{R}_{KL}) (\mathbf{R}_{KL} \cdot \mathbf{M}_L)}{R_{KL}^5}
\end{aligned} \tag{374}$$

where we have included the pairwise Coulomb repulsion between nuclei of charges  $Z_K$  and  $Z_L$  and separation  $R_{KL}$ , the interaction with an external scalar potential  $\phi_K = \phi(\mathbf{R}_K)$  at the positions of the nuclei  $\mathbf{R}_K$ , the Zeeman interaction of the nuclear magnetic moments  $\mathbf{M}_K$  with a uniform external magnetic induction  $\mathbf{B}$ , and the pairwise dipolar magnetic interactions between nuclei of magnetic moments  $\mathbf{M}_K$  and  $\mathbf{M}_L$  and relative positions  $\mathbf{R}_{KL} = \mathbf{R}_K - \mathbf{R}_L$ . Unlike in the electronic spin–spin operator Eq. (B1.7), there is no nuclear–nuclear contact term since the static nuclei never occupy the same point in space.

In introducing nuclei in this chapter, we have treated them as static sources of electromagnetic fields and added, in a somewhat ad hoc manner, the nuclear interactions Eq. (374). There is, however, a way to construct the molecular Hamiltonian that introduces the nuclear terms more naturally, including those terms that involve nuclear motions. In this approach, the electronic Breit–Pauli Hamiltonian Eq. (354) is extended to nuclei by allowing the particles to have masses, charges, spins, and magnetic moments different from those of the electron. In a sense, this approach treats nuclei as anomalous electrons and may be justified provided only terms linear in the inverse nuclear masses are retained. However, this approach excludes several terms needed for the description of molecular properties—for example, the nuclear dipolar interaction in Eq. (374)—which must be retained on a phenomenological ba-

sis. At present, there is no unambiguous and rigorous procedure for constructing molecular Hamiltonians for particles other than electrons.

#### D. Survey of terms in the molecular Breit–Pauli Hamiltonian

There are many ways to collect the terms in the molecular electronic Breit–Pauli Hamiltonian, obtained by substituting the potentials Eqs. (355) and (356) and their associated electromagnetic fields into the contributions Eqs. (B1.1)–(B1.9) of the electronic Hamiltonian Eq. (354). Our classification is based on the nature of the physical mechanisms and interactions rather than on the number and kinds of particles involved in each interaction:

$$H_{\text{mol}}^{\text{BP}} = \left\{ \begin{array}{l} H_{\text{kin}} \quad \leftarrow \text{kinetic energy} \\ + H_{\text{cou}} \quad \leftarrow \text{Coulomb interactions} \\ + H_{\text{ef}} \quad \leftarrow \text{external electric interactions} \\ + H_z \quad \leftarrow \text{Zeeman interactions} \\ + H_{\text{so}} \quad \leftarrow \text{spin–orbit interactions} \\ + H_{\text{ss}} \quad \leftarrow \text{spin–spin interactions} \\ + H_{\text{oo}} \quad \leftarrow \text{orbit–orbit interactions} \\ + H_{\text{dia}} \quad \leftarrow \alpha^4 \text{ diamagnetic interactions} \end{array} \right. \quad (375)$$

Before considering the individual terms, we recall that the nuclei are treated as stationary sources of electromagnetic fields, thereby excluding all terms involving nuclear motion. Moreover, although the electrons are treated relativistically to order  $\alpha^2$ , we retain diamagnetic terms (which are of order  $\alpha^4$ , assuming that the external fields are of order  $\alpha^2$  as discussed in Section VII D) since they are spectroscopically important and needed for the calculation of magnetic properties. With these remarks, we are ready to survey the various contributions to the molecular electronic Breit–Pauli Hamiltonian.

TABLE I: Relativistic corrections to the electronic energy ( $mE_h$ ), calculated in an uncontracted cc-pCVTZ basis for H<sub>2</sub>O and in a cc-pCVDZ-quality basis for H<sub>2</sub>S

	H <sub>2</sub> O		H <sub>2</sub> S	
	RHF	CCSD(T)	RHF	CCSD(T)
mass-velocity	-251.5	-251.9	-4549.0	-4550.9
one-electron Darwin	199.9	199.9	3473.6	3473.6
two-electron Darwin	-3.4	-3.1	-34.1	-33.0
two-electron spin-spin	6.8	6.3	68.2	66.0
two-electron orbit-orbit	0.8	0.5	19.1	17.8
Breit-Pauli	-47.4	-48.3	-1022.2	-1026.5

### 1. Kinetic energy

The Breit-Pauli kinetic-energy operator is given by

$$H_{\text{kin}} = -\frac{1}{2} \sum_i \nabla_i^2 - \frac{\alpha^2}{8} \sum_i \nabla_i^4 \quad (376)$$

The first term is the usual *Newtonian kinetic-energy operator*. The second term is the *mass-velocity term*, which arises because of the relativistic dependence of mass on velocity. The mass-velocity correction is always negative and represents, together with the one-electron Darwin correction described shortly, the major contribution to the relativistic energy correction for slow electrons—see Table I, where we have tabulated the lowest-order relativistic corrections to the ground-state energies of H<sub>2</sub>O and H<sub>2</sub>S. The mass-velocity term is unbounded from below. It (and other small terms in the Breit-Pauli Hamiltonian) should therefore never be used in variational calculations.

## 2. Coulomb interactions

In the Breit–Pauli Hamiltonian Eq.(375), the Coulomb interactions are represented by the operators

$$\begin{aligned}
 H_{\text{cou}} = & - \sum_{iK} \frac{Z_K}{r_{iK}} + \frac{1}{2} \sum_{i \neq j} \frac{1}{r_{ij}} + \frac{1}{2} \sum_{K \neq L} \frac{Z_K Z_L}{R_{KL}} \\
 & + \frac{\alpha^2 \pi}{2} \sum_{iK} Z_K \delta(r_{iK}) - \frac{\alpha^2 \pi}{2} \sum_{i \neq j} \delta(r_{ij}) \\
 & + \frac{2\pi}{3} \sum_{iK} Z_K \mathcal{R}_K^2 \delta(\mathbf{r}_{iK}) - \frac{1}{6} \sum_{iK} \frac{\text{tr} \mathbf{Q}_K (3\mathbf{r}_{iK}^T \mathbf{r}_{iK} - r_{iK}^2 \mathbf{I}_3)}{r_{iK}^5}
 \end{aligned} \tag{377}$$

The first three operators are the usual nonrelativistic point-charge interactions among electrons of charge  $-1$  and nuclei of charge  $Z_K$ , followed by the Darwin corrections caused by the Zitterbewegung of the electrons. As seen from Table I, the one-electron Darwin correction provides, together with the mass-velocity correction in Eq. (376), the dominant relativistic correction to the total energies of light systems—it is almost as large as the mass-velocity correction but positive, increasing the total energy. The negative two-electron Darwin term is much less important (by about two orders of magnitude) since the repulsion between the electrons reduces their probability of being at the same point in space. There is no nuclear–nuclear Darwin correction in Eq. (377) since the nuclei are treated as stationary in the Born–Oppenheimer approximation.

As noted in Section X A, the nuclear point-charge model is not always adequate. The finite size and nonspherical charge distributions of the nuclei may then be corrected for by using the second last and last operators in Eq. (377), respectively. In these operators,  $\mathcal{R}_K$  represents the nuclear extent as defined in Eq. (357) and  $\mathbf{Q}_K$  is the nuclear quadrupole moment, interacting with the field gradient at the nucleus. Nonspherical nuclear charge distributions are observed in nuclear quadrupole resonance studies. They are also important in magnetic resonance studies since the quadrupole moment is aligned with the nuclear magnetic moment, thereby providing a coupling between nuclear spins and electronic field gradients.

### 3. External electric-field interactions

The interactions of a molecule with an externally applied scalar potential  $\phi(\mathbf{r})$  are in the Breit–Pauli Hamiltonian Eq. (375) represented by

$$H_{\text{ef}} = - \sum_i \phi_i + \sum_K Z_K \phi_K + \frac{\alpha^2}{8} \sum_i (\nabla_i \cdot \mathbf{E}_i) \quad (378)$$

where  $\phi(\mathbf{r})$  is evaluated at the positions of the particles. According to Coulomb’s law Eq. (36), the Darwin correction vanishes in vacuum. Since macroscopically applied fields are fairly uniform on a molecular scale, it is expedient to expand the Hamiltonian Eq. (378) in multipoles according to Eq. (90):

$$H_{\text{ef}} = q_{\text{tot}}\phi_0 - \mathbf{p}_{\text{tot}} \cdot \mathbf{E}_0 - \frac{1}{2} \text{tr} \mathbf{M}_{\text{tot}} \mathbf{V}_0 + \dots \quad (379)$$

where  $q_{\text{tot}}$  is the total charge of the molecule, while  $\mathbf{p}_{\text{tot}}$  and  $\mathbf{M}_{\text{tot}}$  are the dipole and second moments, respectively:

$$p_{\text{tot},\alpha} = - \sum_i r_{i\alpha} + \sum_K Z_K R_{K\alpha} \quad (380)$$

$$M_{\text{tot},\alpha\beta} = - \sum_i r_{i\alpha} r_{i\beta} + \sum_K Z_K R_{K\alpha} R_{K\beta} \quad (381)$$

In the multipole expansion Eq. (379), the electric potential  $\phi_0$ , the electric field  $\mathbf{E}_0$ , and the electric field gradient  $\mathbf{V}_0$  are evaluated at the origin of the expansion. Higher-order multipoles are rarely needed to describe macroscopic fields but are needed for highly nonuniform fields generated by neighboring molecular systems—in particular, for the calculation of weak intermolecular interactions and long-range Coulomb interactions. For a general discussion of multipole expansions, we refer to Chapter 9. The external scalar potential  $\phi$  may also contribute to the spin–orbit interaction, as discussed in Section XD 5.

### 4. Zeeman interactions

The Zeeman term in the Breit–Pauli Hamiltonian Eq. (375) describes the paramagnetic interactions of the molecule with an externally applied magnetic field  $\mathbf{B}$ :

$$H_z = -\mathbf{B} \cdot \sum_i \left( -\frac{1}{2} \mathbf{l}_{i0} - \mathbf{s}_i + \frac{1}{2} \alpha^2 \mathbf{s}_i \nabla_i^2 \right) - \mathbf{B} \cdot \sum_K \mathbf{M}_K \quad (382)$$



Here the first and second parts represent the Zeeman interactions of the electrons and the nuclei, respectively. Since the nuclear moments are of the order of  $10^{-3}$  in atomic units, the nuclear part is much smaller than the electronic part but it is important in nuclear magnetic resonance (NMR) spectroscopy, where it determines the position of the (unshielded) resonance lines in the spectra. We also note that, in the Breit–Pauli Hamiltonian Eq. (375), there are, apart from the paramagnetic interactions discussed here, a corresponding set of diamagnetic (quadratic) interactions, which are included in  $H_{\text{dia}}$  and discussed later.

There are three distinct contributions to the electronic part of  $H_z$  in Eq. (382), the first of which represents the Zeeman interaction with the magnetic moment generated by the orbital angular momentum of the electrons relative to the gauge origin  $\mathbf{l}_{i\mathbf{O}} = \mathbf{r}_{i\mathbf{O}} \times \mathbf{p}_i$ . The second and third electronic contributions to  $H_z$  constitute the Zeeman interaction with the spin of the electrons. In addition to a dominant nonrelativistic contribution from Eq. (B1.1), there is a small relativistic correction from Eq. (B1.3), which contributes to the  $g$  shift of electron paramagnetic resonance (EPR) spectroscopy. Note that, because of its dependence on the gauge origin  $\mathbf{O}$ , the Zeeman operator is not uniquely defined as discussed in Section X B.

The expectation value of an imaginary Hermitian operator  $\Omega$  vanishes for orbitally non-degenerate states, since their wave function may be chosen real:

$$\langle \text{real} | \Omega | \text{real} \rangle = \langle \text{real} | \Omega | \text{real} \rangle^* = \langle \text{real} | \Omega^\dagger | \text{real} \rangle = - \langle \text{real} | \Omega | \text{real} \rangle \quad (383)$$

For such states, the expectation value of  $\mathbf{L}_{\mathbf{O}} = \sum_i \mathbf{l}_{i\mathbf{O}}$  therefore vanishes and the orbital angular momentum is said to be quenched. Likewise, the expectation value of any triplet operator such as the spin angular momentum operator  $\mathbf{S} = \sum_i \mathbf{s}_i$  vanishes for singlet states. Closed-shell molecules therefore do not interact with external magnetic fields to first order, but the second-order diamagnetic interactions never vanish. Molecules that interact to first order are called paramagnetic, those that do not are called diamagnetic.

### 5. Spin-orbit interactions

Collecting those terms in the Breit–Pauli Hamiltonian that, to second order in the fine-structure constant, couple the motion of the electrons to particle spins (magnetic moments),

we obtain

$$\begin{aligned}
 H_{\text{so}} = & \frac{\alpha^2}{2} \sum_{iK} \frac{Z_K \mathbf{s}_i \cdot \mathbf{l}_{iK}}{r_{iK}^3} - \frac{\alpha^2}{2} \sum_{i \neq j} \frac{\mathbf{s}_i \cdot \mathbf{l}_{ij}}{r_{ij}^3} - \alpha^2 \sum_{i \neq j} \frac{\mathbf{s}_j \cdot \mathbf{l}_{ij}}{r_{ij}^3} \\
 & + \alpha^2 \sum_{iK} \frac{\mathbf{M}_K \cdot \mathbf{l}_{iK}}{r_{iK}^3} + \frac{\alpha^2}{4} \sum_i \mathbf{s}_i \cdot (\mathbf{E}_i \times \mathbf{p}_i - \mathbf{p}_i \times \mathbf{E}_i)
 \end{aligned}
 \tag{384}$$

The first two terms are the *spin-orbit operators*, which couple the spin of each electron to its own orbital motion, in the presence of the other particles (nuclei and electrons) in the system. By contrast, the third term in Eq. (384) couples the spin and orbital motion of different electrons and is known as the *spin-other-orbit operator*. Because of their triplet nature, the spin-orbit and spin-other-orbit operators do not contribute to the total energy of closed-shell systems to order  $\alpha^2$  and are not listed in Table I. On the other hand, they are responsible for the splitting of otherwise degenerate states in open-shell systems, producing, for example, the fine structure of atomic spectra—see Section 15.6.6. The spin-orbit operators are also important in providing a mechanism for intersystem crossings and phosphorescence—that is, spin-forbidden radiationless and radiative transitions between molecular electronic energy levels. Finally, in magnetic-resonance spectroscopies, these operators contribute in second order to  $g$  values and hyperfine coupling constants in EPR and to zero-field splittings in NMR.

The fourth operator in Eq. (384), which arises from the substitution of  $\mathbf{A}_{\text{nuc}}$  of Eq. (359) into the  $\sum_i \mathbf{A}_i \cdot \mathbf{p}_i$  part of Eq. (B1.1), provides a mechanism for coupling the nuclear magnetic moments  $\mathbf{M}_K$  to the orbital motion of the electrons. It is known as the *orbital hyperfine operator* or the *paramagnetic spin-orbit operator*; it is similar to the spin-other-orbit operator, which likewise couples the spin and orbital motion of different particles. Like all hyperfine operators, the orbital hyperfine interaction is weak ( $10^{-8}$  in atomic units). It contributes to nuclear shielding constants and to indirect nuclear spin-spin coupling constants in NMR, and in second order to hyperfine coupling constants in NMR. The last term in Eq. (384) depends on the external electric field and is unimportant for free molecules.

## 6. Spin–spin interactions

Collecting all terms in the Breit–Pauli operator that involve interactions between the spins (magnetic moments) of two particles, we obtain

$$\begin{aligned}
 H_{\text{ss}} = & \frac{\alpha^2}{2} \sum_{i \neq j} \left[ \frac{r_{ij}^2 \mathbf{s}_i \cdot \mathbf{s}_j - 3 \mathbf{s}_i \cdot \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \mathbf{s}_j}{r_{ij}^5} - \frac{8\pi}{3} \delta(\mathbf{r}_{ij}) \mathbf{s}_i \cdot \mathbf{s}_j \right] \\
 & - \alpha^2 \sum_{iK} \left[ \frac{r_{iK}^2 \mathbf{s}_i \cdot \mathbf{M}_K - 3 \mathbf{s}_i \cdot \mathbf{r}_{iK} \mathbf{r}_{iK} \cdot \mathbf{M}_K}{r_{iK}^5} - \frac{8\pi}{3} \delta(\mathbf{r}_{iK}) \mathbf{s}_i \cdot \mathbf{M}_K \right] \\
 & + \frac{\alpha^2}{2} \sum_{K \neq L} \frac{R_{KL}^2 (\mathbf{M}_K \cdot \mathbf{M}_L) - 3 (\mathbf{M}_K \cdot \mathbf{R}_{KL}) (\mathbf{R}_{KL} \cdot \mathbf{M}_L)}{R_{KL}^5}
 \end{aligned} \tag{385}$$

There are two distinct interactions in these operators: classical *dipolar interactions* and *Fermi-contact interactions*. Because of the smallness of the nuclear magneton, the interactions in Eq. (385) decrease by about three orders of magnitude with each nucleus introduced. The Fermi part of the two-electron spin–spin operator provides the largest two-electron relativistic correction to the energy of light systems—see Table I, which shows that the spin–spin correction is twice that of the two-electron Darwin correction but of opposite sign, in agreement with Exercise 2.1. The two-electron spin–spin operator is responsible for the zero-field splitting observable in EPR spectroscopy of triplet species (biradicals).

Next in Eq. (385) follows a hyperfine interaction operator, which couples the electronic and nuclear spins and arises from the substitution of  $\mathbf{B}_{\text{nuc}}$  of Eq. (361) into  $\sum_i \mathbf{B}_i \cdot \mathbf{s}_i$  of Eq. (B1.1), contributes to the indirect coupling of nuclear spins observed in NMR for liquids, in many cases being the dominant interaction. The Fermi-contact term is particularly important since it is isotropic, coupling the rapidly tumbling nuclear moments studied in high-resolution NMR.

The nuclear spin–spin operator in Eq. (385) is responsible for the important direct coupling of nuclear spins observable in NMR in solids but makes no contribution to the isotropic coupling in liquids. We have omitted the corresponding contact term since it is unimportant and does not arise naturally in the derivation of the Hamiltonian.

### 7. Orbit–orbit interactions

The two-electron orbit–orbit operator Eq. (B1.8) represents a relativistic correction to the two-electron Coulomb interaction, arising from the relative motion of the electrons:

$$H_{\text{oo}} = -\frac{\alpha^2}{4} \sum_{i \neq j} \frac{\mathbf{p}_i \cdot r_{ij}^2 \mathbf{p}_j + \mathbf{p}_i \cdot \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \mathbf{p}_j}{r_{ij}^3} \quad (386)$$

As seen from Table I, the orbit–orbit correction is smaller than the two-electron Darwin and spin–spin corrections but becomes more important in heavier atoms. Moreover, the orbit–orbit interaction does not split levels and does not cause transitions. It is therefore of less interest than the spin–orbit interaction.

### 8. Diamagnetic interactions

The diamagnetic interactions are terms of order  $\alpha^4$  or smaller (keeping in mind that the external magnetic induction is of order  $\alpha^2$ ), arising from expansion of the vector potential in the kinetic-momentum operators of Eq. (354). Although the Breit–Pauli Hamiltonian as such is correct only to order  $\alpha^2$ , the small diamagnetic interactions are included in our discussion since they are important for many spectroscopic parameters:

$$H_{\text{dia}} = \begin{cases} H_{\text{BB}} & \leftarrow \text{magnetizability} \\ + H_{\text{BM}} & \leftarrow \text{nuclear shielding (NMR)} \\ + H_{\text{MM}} & \leftarrow \text{indirect nuclear spin–spin coupling (NMR)} \\ + H_{\text{Bs}} & \leftarrow \text{electronic } g \text{ factor shift (EPR)} \\ + H_{\text{Ms}} & \leftarrow \text{hyperfine coupling (EPR)} \end{cases} \quad (387)$$

The first three operators in  $H_{\text{dia}}$  arise from expansion of  $\frac{1}{2}A^2$  with the vector potential given by Eq. (356):

$$H_{\text{BB}} = \frac{1}{8} \sum_i (\mathbf{B} \times \mathbf{r}_{iO})^2 \quad (388)$$

$$H_{\text{BM}} = \frac{\alpha^2}{2} \sum_{iK} \frac{(\mathbf{B} \times \mathbf{r}_{iO}) \cdot (\mathbf{M}_K \times \mathbf{r}_{iK})}{r_{iK}^3} \quad (389)$$

$$H_{\text{MM}} = \frac{\alpha^4}{2} \sum_{iKL} \frac{(\mathbf{M}_K \times \mathbf{r}_{iK}) \cdot (\mathbf{M}_L \times \mathbf{r}_{iL})}{r_{iK}^3 r_{iL}^3} \quad (390)$$

and contribute to the diamagnetic parts of the magnetizability tensor, to the nuclear shielding tensor, and to the indirect nuclear spin–spin tensor, respectively. Similarly, the last two operators in  $H_{\text{dia}}$  arise from expansion of Eq. (356) in the spin–orbit and spin-other-orbit operators Eqs. (B1.5) and (B1.6), with a simultaneous substitution of the nuclear electric field Eq. (360):

$$H_{\text{Bs}} = \frac{\alpha^2}{4} \sum_{iK} \frac{Z_K (\mathbf{B} \times \mathbf{r}_{iO}) \cdot (\mathbf{s}_i \times \mathbf{r}_{iK})}{r_{iK}^3} - \frac{\alpha^2}{4} \sum_{i \neq j} \frac{(\mathbf{B} \times \mathbf{r}_{iO}) \cdot (\mathbf{s}_i + 2\mathbf{s}_j) \times \mathbf{r}_{ij}}{r_{ij}^3} \quad (391)$$

$$H_{\text{Ms}} = \frac{\alpha^4}{2} \sum_{i,KL} \frac{Z_K (\mathbf{M}_L \times \mathbf{r}_{iL}) \cdot (\mathbf{s}_i \times \mathbf{r}_{iK})}{r_{iK}^3 r_{iL}^3} - \frac{\alpha^4}{2} \sum_{K,i \neq j} \frac{(\mathbf{M}_K \times \mathbf{r}_{iK}) \cdot (\mathbf{s}_i + 2\mathbf{s}_j) \times \mathbf{r}_{ij}}{r_{iK}^3 r_{ij}^3} \quad (392)$$

These operators couple the spin of the electron to the external magnetic field and the nuclear magnetic moments, respectively, contributing in second order to the  $g$  shifts and nuclear hyperfine coupling constants in EPR spectroscopy.

## XI. EXERCISES

1. Demonstrate that the vector potential Eq. (79) satisfies the Coulomb-gauge condition.  
Hint: In magnetostatics, the continuity relation Eq. (42) gives  $\nabla \cdot \mathbf{J} = 0$ .

2. Show that, for a particle subject to the Lorentz force Eq. (3), the generalized potential is given by Eq. (118).

(a) Use the relation  $\mathbf{B} = \nabla \times \mathbf{A}$  to show that

$$\mathbf{v} \times \mathbf{B} = \nabla (\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla) \mathbf{A} \quad (393)$$

and thus that the velocity-dependent Lorentz force  $\mathbf{F} = z (\mathbf{E} + \mathbf{v} \times \mathbf{B})$  in Eq. (3) can be written in the form

$$\mathbf{F} = z \left[ -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} + \nabla (\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla) \mathbf{A} \right] \quad (394)$$

Hint: Use the identity  $(\mathbf{a} \times \mathbf{b})_i = \sum_{jk} \epsilon_{ijk} a_j b_k$ , where the Levi-Civita antisymmetric symbol  $\epsilon_{ijk}$  is +1 for an even permutation of the indices, -1 for an odd permutation, and 0 whenever two indices are equal.

(b) Show that the Lorentz force Eq. (394) can be obtained from the generalized potential  $U = z (\phi - \mathbf{v} \cdot \mathbf{A})$  as

$$F_i = -\frac{\partial U}{\partial q_i} + \frac{d}{dt} \left( \frac{\partial U}{\partial \dot{q}_i} \right) \quad (395)$$

Hint: Use the fact that  $\mathbf{A}(\mathbf{r}, t)$  depends on  $\mathbf{r}$  and  $t$  but not on  $\mathbf{v}$ .

3. Show that, for a particle subject to the Lorentz force Eq. (3), the Hamiltonian is given by Eq. (125).

(a) Show that the first of Hamilton's equations Eq. (24) with the Hamiltonian Eq. (125) yields the generalized momentum:

$$\mathbf{p} = m\mathbf{v} + z\mathbf{A} \quad (396)$$

(b) Using Eq. (396), show that left-hand side of the second of Hamilton's equations Eq. (25) may be written in the form

$$\dot{\mathbf{p}} = m\mathbf{a} + z \frac{\partial \mathbf{A}}{\partial t} + z (\mathbf{v} \cdot \nabla) \mathbf{A} \quad (397)$$

Hint: Use the fact that, since  $\mathbf{A}(\mathbf{r}, t)$  does not depend on  $\mathbf{v}$ , its total time derivative is given by

$$\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{A} \quad (398)$$

- (c) Using Eq. (396), show that right-hand side of the second of Hamilton's equations Eq. (25) may be written in the form

$$-\nabla H = z(\mathbf{v} \cdot \nabla) \mathbf{A} + z\mathbf{v} \times (\nabla \times \mathbf{A}) - z\nabla\phi \quad (399)$$

Hint: Use the identity

$$\nabla(\mathbf{u} \cdot \mathbf{u}) = 2(\mathbf{u} \cdot \nabla)\mathbf{u} + 2\mathbf{u} \times (\nabla \times \mathbf{u}) \quad (400)$$

valid for any vector  $\mathbf{u}$  as a function of  $\mathbf{r}$ .

- (d) Substituting Eqs. (396) and (400) into the second of Hamilton's equations Eq. (25) and using Eqs. (46) and (48), show that we obtain

$$m\mathbf{a} = z\mathbf{E} + z(\mathbf{v} \times \mathbf{B}) \quad (401)$$

which are precisely Newton's equations with the Lorentz force Eq. (3).

4. In this exercise, we demonstrate that the set of two-dimensional anticommuting matrices Eq. (185) also satisfy the commutation relations Eq. (187).

- (a) Show that the commutator  $[\sigma_x, \sigma_y]$  can be expanded as

$$[\sigma_x, \sigma_y] = 2i(a\sigma_x + b\sigma_y + c\sigma_z) \quad (402)$$

with real coefficients  $a$ ,  $b$ , and  $c$ .

- (b) Anticommuting both sides of Eq. (402) with  $\sigma_x$  and then with  $\sigma_y$ , demonstrate that

$$[\sigma_x, \sigma_y] = 2ic\sigma_z \quad (403)$$

- (c) Squaring both sides of Eq. (403), demonstrate that  $c = \pm 1$ . Show that  $c$  is preserved by a cyclic permutation of  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$ .

5. In this exercise, we investigate the matrix representations of a set of Hermitian anticommuting operators:

$$[\Omega_i, \Omega_j]_+ = 2\delta_{ij}, \quad \Omega_i^\dagger = \Omega_i \quad (404)$$

- (a) Show that the eigenvalues of such operators are  $\pm 1$  and that, for each operator  $\Omega_i$ , there are equally many positive and negative eigenvalues. Hint: Demonstrate that the sum of the eigenvalues is zero.
- (b) Assume that the matrix representation of  $\Omega_0$  is two-dimensional diagonal:

$$\Omega_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (405)$$

Demonstrate that the remaining matrices have the off-diagonal structure

$$\Omega_i = \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix}, \quad i > 0 \quad (406)$$

where  $aa^* = 1$ . Show that, to within an arbitrary phase factor, there are only two different matrices that satisfy Eq. (406). There are therefore only three independent anticommuting Hermitian operators represented by two-dimensional matrices. In the standard representation, these operators are represented by the Pauli spin matrices  $\sigma_i$ .

- (c) Assume that the matrix representation of  $\Omega_0$  is four-dimensional diagonal:

$$\Omega_0 = \begin{pmatrix} \mathbf{I}_1 & \mathbf{0}_2 \\ \mathbf{0}_2 & -\mathbf{I}_2 \end{pmatrix} \quad (407)$$

Demonstrate that the remaining matrices have the off-diagonal structure

$$\Omega_i = \begin{pmatrix} \mathbf{0} & \boldsymbol{\omega}_i \\ \boldsymbol{\omega}_i & \mathbf{0} \end{pmatrix}, \quad i > 0 \quad (408)$$

where  $\boldsymbol{\omega}_i$  are a set of anticommuting Hermitian two-dimensional matrices. Taking these two-dimensional matrices to be the Pauli spin matrices, we arrive at four four-dimensional Dirac matrices  $\boldsymbol{\alpha}_i$ .



## Appendix A: Vector identities

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) \quad (\text{A1})$$

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} \quad (\text{A2})$$

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) \quad (\text{A3})$$

$$\nabla \cdot (\phi \mathbf{a}) = (\nabla \phi) \cdot \mathbf{a} + \phi (\nabla \cdot \mathbf{a}) \quad (\text{A4})$$

$$\nabla \times (\phi \mathbf{a}) = (\nabla \phi) \times \mathbf{a} + \phi (\nabla \times \mathbf{a}) \quad (\text{A5})$$

$$\nabla (\mathbf{a} \cdot \mathbf{b}) = (\mathbf{a} \cdot \nabla) \mathbf{b} + (\mathbf{b} \cdot \nabla) \mathbf{a} + \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) \quad (\text{A6})$$

$$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = (\nabla \times \mathbf{a}) \cdot \mathbf{b} - (\nabla \times \mathbf{b}) \cdot \mathbf{a} \quad (\text{A7})$$

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a} (\nabla \cdot \mathbf{b}) - \mathbf{b} (\nabla \cdot \mathbf{a}) - (\mathbf{a} \cdot \nabla) \mathbf{b} + (\mathbf{b} \cdot \nabla) \mathbf{a} \quad (\text{A8})$$

$$\nabla \cdot \nabla \times \mathbf{a} = 0 \quad (\text{A9})$$

$$\nabla \times \nabla \phi = \mathbf{0} \quad (\text{A10})$$

$$\nabla \times (\nabla \times \mathbf{a}) = \nabla (\nabla \cdot \mathbf{a}) - \nabla^2 \mathbf{a} \quad (\text{A11})$$

$$\nabla \cdot \mathbf{r} = 3 \quad (\text{A12})$$

$$\nabla \times \mathbf{r} = \mathbf{0} \quad (\text{A13})$$

$$(\mathbf{a} \cdot \nabla) \mathbf{r} = \mathbf{a} \quad (\text{A14})$$

$$\nabla_i r = \frac{r_i}{r} \quad (\text{A15})$$

$$\nabla_i r^{-1} = -\frac{r_i}{r^3} \quad (\text{A16})$$

$$\nabla_i \nabla_j r^{-1} = \frac{3r_i r_j - \delta_{ij} r^2}{r^5} - \frac{4\pi}{3} \delta_{ij} \delta(\mathbf{r}) \quad (\text{A17})$$

$$\nabla^2 r = 2r^{-1} \quad (\text{A18})$$

$$\nabla^2 r^{-1} = -4\pi \delta(\mathbf{r}) \quad (\text{A19})$$

$$(\mathbf{I}_3 \nabla^2 - \nabla \nabla^T) r = \frac{\mathbf{I}_3 r^2 + \mathbf{r} \mathbf{r}^T}{r^3} \quad (\text{A20})$$

$$(\mathbf{I}_3 \nabla^2 - \nabla \nabla^T) r^{-1} = \frac{\mathbf{I}_3 r^2 - 3\mathbf{r} \mathbf{r}^T}{r^5} - \frac{8\pi}{3} \delta(r) \quad (\text{A21})$$

## Appendix B: Matrix direct products

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} A_{11}\mathbf{B} & A_{12}\mathbf{B} & \cdots & A_{1n}\mathbf{B} \\ A_{21}\mathbf{B} & A_{22}\mathbf{B} & \cdots & A_{2n}\mathbf{B} \\ \cdots & & & \cdots \\ A_{n1}\mathbf{B} & A_{n2}\mathbf{B} & \cdots & A_{nn}\mathbf{B} \end{pmatrix} \quad (\text{B1})$$

$$\mathbf{A} \oplus \mathbf{B} = \mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{B} \quad (\text{B2})$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD} \quad (\text{B3})$$

$$(c\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (c\mathbf{B}) = c(\mathbf{A} \otimes \mathbf{B}) \quad (\text{B4})$$

$$(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C} \quad (\text{B5})$$

$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C} \quad (\text{B6})$$

$$\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} \quad (\text{B7})$$

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \quad (\text{B8})$$

$$(\mathbf{A} \otimes \mathbf{B})^\dagger = \mathbf{A}^\dagger \otimes \mathbf{B}^\dagger \quad (\text{B9})$$

$$f(\mathbf{I} \otimes \mathbf{A}) = \mathbf{I} \otimes f(\mathbf{A}) \quad (\text{B10})$$

$$f(\mathbf{A} \otimes \mathbf{I}) = f(\mathbf{A}) \otimes \mathbf{I} \quad (\text{B11})$$

$$\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr} \mathbf{A} \text{tr} \mathbf{B} \quad (\text{B12})$$

$$\exp(\mathbf{A} \oplus \mathbf{B}) = \exp(\mathbf{A}) \otimes \exp(\mathbf{B}) \quad (\text{B13})$$

TABLE II: Fundamental constants and units (2002 CODATA recommended values)

quantity	symbol	numerical value
speed of light in vacuum	$c$	$299\,792\,458\text{ m s}^{-1}$
magnetic constant	$\mu_0$	$4\pi \times 10^{-7}\text{ N A}^{-2}$
electric constant	$\epsilon_0 = \mu_0^{-1}c^{-2}$	$8.854\,187\,817\dots \times 10^{-12}\text{ F m}^{-1}$
elementary charge	$e$	$1.602\,176\,53(14) \times 10^{-19}\text{ C}$
electron mass	$m_e$	$9.109\,3826(16) \times 10^{-31}\text{ kg}$
proton mass	$m_p$	$1.672\,621\,71(29) \times 10^{-27}\text{ kg}$
Planck constant	$h$	$6.626\,0693(11) \times 10^{-34}\text{ J s}$
reduced Planck constant	$\hbar = h/2\pi$	$1.054\,571\,68(18) \times 10^{-34}\text{ J s}$
fine-structure constant	$\alpha = e^2/4\pi\epsilon_0\hbar c$	$7.297\,352\,568(24) \times 10^{-3}$
inverse fine-structure constant	$\alpha^{-1}$	$137.035\,999\,11(46)$
Rydberg constant	$R_\infty = \alpha^2 m_e c^2 / 2h$	$10\,973\,731.568\,525(73)\text{ m}^{-1}$
Bohr radius	$a_0 = \alpha / 4\pi R_\infty$	$0.529\,177\,210\,8(18) \times 10^{-10}\text{ m}$
Hartree energy	$E_h = \alpha^2 m_e c^2$	$4.359\,744\,17(75) \times 10^{-18}\text{ J}$
Bohr magneton	$\mu_B = e\hbar/2m_e$	$9.274\,009\,49(80) \times 10^{-24}\text{ J T}^{-1}$
nuclear magneton	$\mu_N = e\hbar/2m_p$	$5.050\,783\,43(43) \times 10^{-27}\text{ J T}^{-1}$
Avogadro constant	$N_A$	$6.022\,1415(10) \times 10^{23}\text{ mol}^{-1}$
molar gas constant	$R$	$8.314\,472(15)\text{ J mol}^{-1}\text{ K}^{-1}$
Boltzmann constant	$k = R/N_A$	$1.380\,6505(24) \times 10^{-23}\text{ J K}^{-1}$