

## Adiabatic motion of a neutral spinning particle in an inhomogeneous magnetic field

Robert G. Littlejohn and Stefan Weigert

*Department of Physics, University of California, Berkeley, Berkeley, California 94720*

(Received 25 January 1993)

The motion of a neutral particle with a magnetic moment in an inhomogeneous magnetic field is considered. This situation, occurring, for example, in a Stern-Gerlach experiment, is investigated from classical and semiclassical points of view. It is assumed that the magnetic field is strong or slowly varying in space, i.e., that adiabatic conditions hold. To the classical model, a systematic Lie-transform perturbation technique is applied up to second order in the adiabatic-expansion parameter. The averaged classical Hamiltonian contains not only terms representing fictitious electric and magnetic fields but also an additional velocity-dependent potential. The Hamiltonian of the quantum-mechanical system is diagonalized by means of a systematic WKB analysis for coupled wave equations up to second order in the adiabaticity parameter, which is coupled to Planck's constant. An exact term-by-term correspondence with the averaged classical Hamiltonian is established, thus confirming the relevance of the additional velocity-dependent second-order contribution.

PACS number(s): 03.65.-w, 03.20.+i

### I. INTRODUCTION

The classical experimental apparatus for observing the spin of particles is that of Stern and Gerlach, in which neutral particles, traveling through a region with an inhomogeneous magnetic field, are deflected according to their magnetic moment. Under adiabatic conditions, this situation turns out to be of considerable theoretical interest. For example, the presence of a strong or slowly varying magnetic field implies a natural separation of the dynamical variables into slow and fast ones, corresponding to the "orbital" (spatial) and spin degrees of freedom, respectively. In this case, it is possible to give an effective description of the particle motion: a fictitious magnetic field is introduced which, in an averaged sense, accounts for some of the effects of the spin degrees of freedom on the spatial motion.

The fictitious magnetic field represents a phase anholonomy, examples of which are dispersed throughout many areas of physics [1]. Such anomalies have attracted a lot of interest under the heading of "geometric phases," or "Berry's phase."

In this paper the dynamics of a neutral particle with a magnetic moment under adiabatic conditions is investigated in a systematic way. Second-order corrections (in an appropriate expansion parameter) are included throughout; thus, the present analysis goes beyond the above-mentioned phenomena which arise from considerations including first-order ones. For the neutral particle with spin, both a purely classical or a purely quantum-mechanical description can be given in a consistent way; in addition, a semiclassical approach is possible where the orbital degrees of freedom are considered as classical, whereas the spin variable remains quantized and, in particular, is not required to take on large values. Here, an adiabatic treatment of the classical and the semiclassical systems will be given, involving Lie-transform perturbation techniques, on the one hand, and the Weyl calculus

useful for semiclassical systems, on the other.

The paper is organized as follows: In Sec. II the investigation of the classical system is presented. Our study of this system was motivated by the recent work of Aharonov and Stern [2], who analyze the same system but with different techniques. In our analysis, the equations of motion are first written down in an appropriate set of coordinates, and a formal expansion parameter is introduced. Then the Hamiltonian description of the system is given. Finally, the Hamiltonian, expressed in appropriate canonical coordinates, is used to carry out a Lie-transform perturbation analysis through second order. In Sec. III we analyze the quantum-mechanical system with the orbital motion treated semiclassically. Since the Schrödinger equation of this system is a coupled-wave equation, a modification of the existing theory of systematic diagonalization of such equations up to arbitrary powers in Planck's constant (or other adiabatic parameter) is presented. Subsequently, this method is applied to calculate the symbol of the Hamiltonian operator valid through second order. Finally, the relevant physical parameters and their relations which define the adiabatic conditions are discussed in Sec. IV. In addition, the results of the two calculations are compared.

The principal results of this work are, first, the development of systematic perturbation methods for both the classical and semiclassical approaches to the motion of spinning neutral particles in inhomogeneous magnetic fields, and second, the discovery of a term in the averaged Hamiltonian which was not found in the analysis of Aharonov and Stern. In addition, our semiclassical treatment is new, since Aharonov and Stern only performed a purely classical calculation.

Regarding the perturbation methods, it must be said that in the case of neutral spinning particles in inhomogeneous magnetic fields, the terms in the equations of motion associated with Berry's phase seem to be very small in practice, so that even the first-order corrections

are probably not observable in a practical sense. Nevertheless, these corrections are of interest for several reasons. First, there is general theoretical interest in the subject of “geometrical” forces in dynamical systems, for which the spin system serves as a model. This seems to have been the motivation of Berry [3] (and in Ref. [1]), Berry and Robbins [4], and Aharonov and Stern [2]. Related to this is our own motivation in studying examples of systematic perturbation methods, developed earlier in a general way [5–7]. Second, the spin system serves as a paradigm for a more general class of Born-Oppenheimer systems, and for these the correction terms we develop, even at second order, are definitely important under some circumstances.

Regarding the extra term we find in the averaged Hamiltonian, we should say, first of all, that this term is not entirely new, for we are aware of Born-Oppenheimer calculations which seem to be equivalent to the use of this term [8]. These calculations show incidentally that the term in question is important to practice. Nevertheless, the fact that this term does not appear in the classical analysis of Aharonov and Stern [2] requires explanation. We have discussed this question in detail in Sec. IV; here we will simply remark that the relative magnitude of the new term depends on the momentum and that if the momentum is chosen large enough, the new term becomes comparable to or larger than other terms (such as the geometric scalar potential) discussed by Aharonov and Stern. The momentum can take on the requisite large values without violating the adiabatic assumption upon which the classical analysis of both Aharonov and Stern and ourselves is based. Indeed, it is not even consistent to adopt the small values of momentum assumed by Aharonov and Stern unless one assumes (as they do) that the magnetic field varies in direction but not in magnitude. Such magnetic fields exist, but are not typical. Altogether, we can say that our analysis is more general than that of Aharonov and Stern, and includes theirs as a special case.

## II. THE CLASSICAL SYSTEM

In this section the motion of a classical neutral particle in an inhomogeneous magnetic field  $\mathbf{B}(\mathbf{x})$  is investigated. The analysis of the particle motion is divided into three parts. First, Newton’s equations of motion for the system are written down explicitly in appropriate coordinates and a physical interpretation of the individual terms is given. Furthermore, a formal expansion parameter  $\epsilon$  is introduced which facilitates bookkeeping of the involved terms according to their order of magnitude. Second, a Hamiltonian formulation of the equations of motion is given, and third, a Lie-transform perturbation analysis with respect to the parameter  $\epsilon$  is applied.

### A. Equations of motion

A classical particle with mass  $m_0$  and (classical) spin  $\mathbf{s}$  moves in a magnetic field  $\mathbf{B}(\mathbf{x})$  according to the equations

$$\begin{aligned} m_0 \ddot{\mathbf{x}} &= \gamma \nabla(\mathbf{B} \cdot \mathbf{s}), \\ \dot{\mathbf{s}} &= \gamma \mathbf{s} \times \mathbf{B}. \end{aligned} \quad (1)$$

The magnetic moment of the particle is given by  $-\gamma \mathbf{s}$ , where  $\gamma$  is simply regarded as a constant of proportionality. Choosing units in which  $m_0=1$  and absorbing the number  $\gamma$  into  $\mathbf{B}$ , we obtain a slightly simpler version of Eqs. (1):

$$\begin{aligned} \ddot{\mathbf{x}} &= \nabla(\mathbf{B} \cdot \mathbf{s}), \\ \dot{\mathbf{s}} &= \frac{1}{\epsilon} \mathbf{s} \times \mathbf{B}. \end{aligned} \quad (2)$$

A dimensionless parameter  $\epsilon$  has been introduced, which indicates that the spin precession is fast compared to the spatial variables. Physically, this means that the spin precesses a large number of times about the local axis of the magnetic field before the field has changed substantially due to the spatial (or “orbital”) motion of the particle. The manner in which these equations depend on the parameter  $\epsilon$  can be derived by scaling arguments, beginning with the assumption of either a strong or a slowly varying magnetic field. This derivation will be discussed in detail in Sec. IV. It is the set of Eqs. (2) on which, eventually, a perturbation expansion is to be done, carrying things out to second order.

When we refer the components of the vector  $\mathbf{s}$  to the axes  $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ , which form an inertial frame of reference (the “space” axes), we will write  $(s_x, s_y, s_z)$ . Instead of these components, we can describe the spin equally well by the variables  $(\phi', m')$ , where  $m' = s_z$  is the  $z$  component of the spin, and  $\phi'$  is the azimuthal angle of the spin about the  $\mathbf{e}_z$  axis. The variables  $(\phi', m')$  are canonically conjugate and allow us to describe the spin in a classical Hamiltonian formalism. The length of the spin,  $s = |\mathbf{s}|$ , is a constant of motion. The prime on the variables  $(\phi', m')$  is a reminder that these variables refer to the space  $z$  axis.

The spin motion can also be described with respect to another set of axes  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  which is defined by the requirement that  $\mathbf{e}_3$  points along the local direction  $\mathbf{b}$  of the magnetic field  $\mathbf{B} = B\mathbf{b}$ . The new axes (the “body” axes) are functions of  $\mathbf{x}$ , because the magnetic field  $\mathbf{B}$  is a function of position  $\mathbf{x}$ . Geometrically, the body axes constitute a field of frames (that is, a field of orthogonal unit vectors) over the three-dimensional configuration space. The axes  $(\mathbf{e}_1, \mathbf{e}_2)$ , which span the plane perpendicular to the vector  $\mathbf{b}$ , are not uniquely determined by the requirement  $\mathbf{e}_3 = \mathbf{b}$ . In fact, any smooth choice of  $(\mathbf{e}_1, \mathbf{e}_2)$  is possible and can be regarded as a choice of gauge. In this terminology, a redefinition of the pair of basis vectors  $(\mathbf{e}_1, \mathbf{e}_2)$  over configuration space in a possibly spatially dependent manner amounts to a gauge transformation. Similar situations, involving fields of planes over configuration space, have been studied previously by Felsager and Leinaas [9] in the context of a geometric interpretation of magnetic fields and by Littlejohn [10,11] within the framework of the classical guiding-center motion.

The goal of the following calculation is to express the equations of motion for the spin in canonical coordinates referring to the body axes  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ . We denote the components of the spin with respect to these axes by  $(s_1, s_2, s_3)$ . We let  $m = s_3$ , and we let  $\phi$  be the azimuthal

angle of the spin in the  $(\mathbf{e}_1, \mathbf{e}_2)$  plane, so that

$$\begin{aligned} s_1 &= \mathbf{e}_1 \cdot \mathbf{s} = (s^2 - m^2)^{1/2} \cos \phi, \\ s_2 &= \mathbf{e}_2 \cdot \mathbf{s} = (s^2 - m^2)^{1/2} \sin \phi, \\ s_3 &= \mathbf{b} \cdot \mathbf{s} = m. \end{aligned} \quad (3)$$

Our notation is to use unprimed symbols  $(\phi, m)$  for quantities that refer to the body axes and primed symbols  $(\phi', m')$  for those referring to the space axes.

Let  $R$  be the rotation operator which maps the space axes into the body axes, interpreted in the active sense, so that  $\mathbf{e}_1 = R\mathbf{e}_x$ ,  $\mathbf{e}_2 = R\mathbf{e}_y$ ,  $\mathbf{e}_3 = \mathbf{b} = R\mathbf{e}_z$ . We write  $(\mathbf{R})_{ij} = R_{ij}$  for the matrix elements of the operator  $R$  with respect to the space axes. Thus, the three columns of  $\mathbf{R}$  are the space components of  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b})$ . The rotation  $R$  is a function of position  $\mathbf{x}$ .

To get the equations of motion for the body components  $(s_1, s_2, s_3)$ , we temporarily use indices  $a, b, c = (x, y, z)$  for space components and indices  $i, j, k = (1, 2, 3)$  for body components, so that  $(\mathbf{e}_i)_a = R_{ai}$  and  $s_i = \mathbf{e}_i \cdot \mathbf{s} = R_{ai} s_a$ . Differentiating the latter expression with respect to time, we obtain

$$\begin{aligned} \dot{s}_i &= R_{ai} \dot{s}_a + \dot{R}_{ai} s_a \\ &= \frac{1}{\epsilon} B \epsilon_{abc} R_{ai} R_{bj} R_{c3} s_j + \dot{R}_{ai} R_{aj} s_j \\ &= \frac{1}{\epsilon} B \epsilon_{ij3} s_j - (\mathbf{R}^{-1} \dot{\mathbf{R}})_{ij} s_j. \end{aligned} \quad (4)$$

The first term on the right-hand side of the last equation is the dynamical phase, corresponding to the precession of the spin about the field  $\mathbf{B}(\mathbf{x})$ , and the second one accounts for the change of frames due to the orbital motion. This term is the same as  $-(\boldsymbol{\omega} \times \mathbf{s})_i$ , where  $\boldsymbol{\omega}$  is the angular velocity. The matrix  $\mathbf{R}^{-1} \dot{\mathbf{R}}$  is antisymmetric, since

$$(\mathbf{R}^{-1} \dot{\mathbf{R}})_{ij} = \dot{\mathbf{x}} \cdot \nabla_j \mathbf{e}_i \cdot \mathbf{e}_i = -\dot{\mathbf{x}} \cdot \nabla_i \mathbf{e}_j \cdot \mathbf{e}_j, \quad (5)$$

as follows from the orthonormality of the basis vectors

$$\nabla(\mathbf{e}_i \cdot \mathbf{e}_j) = 0. \quad (6)$$

In Eq. (5) the convention of letting the  $\nabla$  act only on the operand immediately following has been adopted, so  $\nabla \mathbf{e}_i \cdot \mathbf{e}_j$  means  $(\nabla \mathbf{e}_i) \cdot \mathbf{e}_j$ , etc.

The equations of motion for the variables  $(\phi, m)$  follow from differentiating Eq. (3) with respect to  $t$ , using Eqs. (4) and (5), and solving for  $(\dot{\phi}, \dot{m})$ . After also transforming the equation of motion for  $\mathbf{x}$ , we obtain the following system of equations:

$$\begin{aligned} \dot{\mathbf{x}} &= m \nabla B + B (s^2 - m^2)^{1/2} \nabla \mathbf{b} \cdot \mathbf{c}, \\ \dot{m} &= (s^2 - m^2)^{1/2} (\dot{\mathbf{x}} \cdot \nabla \mathbf{b} \cdot \mathbf{c}), \\ \dot{\phi} &= -\frac{B}{\epsilon} - \dot{\mathbf{x}} \cdot \mathbf{A} - \frac{m}{(s^2 - m^2)^{1/2}} (\dot{\mathbf{x}} \cdot \nabla \mathbf{b} \cdot \mathbf{a}), \end{aligned} \quad (7)$$

which is exactly equivalent to Eqs. (2). Here we have introduced the unit vectors  $\mathbf{a}$  and  $\mathbf{c}$ , defined by

$$\begin{aligned} \mathbf{a} &= -\mathbf{e}_1 \sin \phi + \mathbf{e}_2 \cos \phi, \\ \mathbf{c} &= +\mathbf{e}_1 \cos \phi + \mathbf{e}_2 \sin \phi; \end{aligned} \quad (8)$$

these are functions of both variables,  $\mathbf{x}$  and  $\phi$  and they rotate with the precessing spin. In particular, the vectors  $\mathbf{c}$  and  $\mathbf{s}_1$ , the projection of  $\mathbf{s}$  onto the plane perpendicular to the magnetic field, point in the same direction, so that  $\mathbf{s}_1 = (s^2 - m^2)^{1/2} \mathbf{c}$ , and we have

$$\mathbf{s} = m \mathbf{b} + (s^2 - m^2)^{1/2} \mathbf{c}. \quad (9)$$

Our notation uses the mnemonic that  $(\mathbf{a}, \mathbf{b}, \mathbf{c})$  form a right-handed triad. These vectors satisfy

$$\mathbf{b} \times \mathbf{a} = \frac{\partial \mathbf{a}}{\partial \phi} = -\mathbf{c}, \quad \mathbf{b} \times \mathbf{c} = \frac{\partial \mathbf{c}}{\partial \phi} = +\mathbf{a}. \quad (10)$$

The gauge potential  $\mathbf{A}$  in Eq. (7) is defined by

$$\mathbf{A} = \nabla \mathbf{e}_1 \cdot \mathbf{e}_2 = -\nabla \mathbf{e}_2 \cdot \mathbf{e}_1 = \nabla \mathbf{c} \cdot \mathbf{a} = -\nabla \mathbf{a} \cdot \mathbf{c}. \quad (11)$$

The vector  $\nabla \mathbf{c} \cdot \mathbf{a} = -\nabla \mathbf{a} \cdot \mathbf{c}$  is actually independent of  $\phi$ .

Before turning to the Hamiltonian formulation, we will give a physical interpretation of the individual terms on the right-hand side of Eqs. (7). In the first equation, the force  $\nabla(\mathbf{B} \cdot \mathbf{s})$  on the particle is decomposed into two parts, the first of which,  $(\mathbf{b} \cdot \mathbf{s}) \nabla B$ , is due to the changing strength of the magnetic field, whereas the second one,  $B \nabla \mathbf{b} \cdot \mathbf{s}_1 = B \nabla \mathbf{b} \cdot \mathbf{s}$ , originates from the directional change of the field from point to point. Recall that for an orthonormal triad of vectors  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ , an infinitesimal change of any of these vectors can have components only in the plane perpendicular to it.

Next, if the particle moves in the time  $dt$  from the point  $\mathbf{x}$  to the point  $\mathbf{x}'$ , the change  $dm$  of the variable  $m$  is due solely to the change in direction of the field  $\mathbf{B}$ ,

$$dm = d\mathbf{b} \cdot \mathbf{s} = d\mathbf{b} \cdot \mathbf{s}_1 = (s^2 - m^2)^{1/2} d\mathbf{x} \cdot \nabla \mathbf{b} \cdot \mathbf{c}, \quad (12)$$

which implies the second of Eqs. (7).

The three terms contributing to the change of the angle  $\phi$  in Eqs. (7) have the following origin: First, the term proportional to  $B$  accounts for the precession of the spin about the local direction of the magnetic field,  $\mathbf{b}(\mathbf{x})$ .

The second term  $-\dot{\mathbf{x}} \cdot \mathbf{A}$ , is related to the freedom of choosing an arbitrary convention for the origin of the phase  $\phi$  in the planes perpendicular to the spatially varying vector  $\mathbf{b}$ . The following discussion of this term is an abbreviated version of the analysis presented in [11]: In order to compare the relative orientation of two pairs of basis vectors  $(\mathbf{e}_1(\mathbf{x}), \mathbf{e}_2(\mathbf{x}))$  and  $(\mathbf{e}_1(\mathbf{x}'), \mathbf{e}_2(\mathbf{x}'))$ , spanning two different planes  $\pi(\mathbf{x})$  and  $\pi(\mathbf{x}')$  at the points  $\mathbf{x}$  and  $\mathbf{x}'$ , respectively, one needs a notion of parallel transport for vectors. We begin by transporting the vectors  $(\mathbf{e}_1(\mathbf{x}), \mathbf{e}_2(\mathbf{x}))$  parallel to themselves to the point  $\mathbf{x}'$ , resulting in two vectors spanning a plane  $\bar{\pi}(\mathbf{x}')$ , which in general will not coincide with the plane  $\pi(\mathbf{x}')$ . Projecting the vectors  $(\bar{\mathbf{e}}_1(\mathbf{x}'), \bar{\mathbf{e}}_2(\mathbf{x}'))$  onto the plane  $\pi(\mathbf{x}')$ , one obtains a new set of vectors  $(\mathbf{e}'_1(\mathbf{x}'), \mathbf{e}'_2(\mathbf{x}'))$ , coplanar with the pair  $(\mathbf{e}_1(\mathbf{x}'), \mathbf{e}_2(\mathbf{x}'))$ . For infinitesimally close points  $\mathbf{x}$  and  $\mathbf{x}'$ , the orthonormality of the original vectors  $(\mathbf{e}_1(\mathbf{x}), \mathbf{e}_2(\mathbf{x}))$  is preserved in the transport process. Consequently, the transported pair of vectors is related to the vectors

$(\mathbf{e}_1(\mathbf{x}'), \mathbf{e}_2(\mathbf{x}'))$  by an infinitesimal rotation about the axis  $\mathbf{b}(\mathbf{x})$ :

$$\begin{bmatrix} \mathbf{e}'_1(\mathbf{x}') \\ \mathbf{e}'_2(\mathbf{x}') \end{bmatrix} = \begin{bmatrix} 1 & d\alpha \\ -d\alpha & 1 \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(\mathbf{x}') \\ \mathbf{e}_2(\mathbf{x}') \end{bmatrix}, \quad (13)$$

where  $d\alpha$  is the small angle of rotation. Expanding this out to first order, we find

$$d\alpha = -d\mathbf{e}_1 \cdot \mathbf{e}_2 = d\mathbf{e}_2 \cdot \mathbf{e}_1 = -d\mathbf{x} \cdot \nabla \mathbf{e}_1 \cdot \mathbf{e}_2. \quad (14)$$

Consequently, when the particle travels the distance  $d\mathbf{x}$  during the time interval  $dt$ , it acquires an additional phase

$$d\phi = -d\mathbf{x} \cdot \nabla \mathbf{e}_1 \cdot \mathbf{e}_2, \quad (15)$$

which is the origin of the second term in the equation for  $\dot{\phi}$ .

The spatial dependence of the vector  $\mathbf{b}$  is the origin of the third contribution to the last of Eqs. (7). This can be seen as follows: if the particle moves from the point  $\mathbf{x}$  to  $\mathbf{x}'$ , the new direction of the field is given by

$$\mathbf{b}(\mathbf{x}') \equiv \mathbf{b}' = \mathbf{b}(\mathbf{x}) + d\mathbf{b}(\mathbf{x}),$$

and the change  $d\mathbf{b}$  is orthogonal to  $\mathbf{b}$ . This transformation of the axis  $\mathbf{b}$  corresponds to a small rotation about some axis  $\boldsymbol{\rho}$  perpendicular to the  $(\mathbf{b}, \mathbf{b}')$  plane:  $\boldsymbol{\rho} = \mathbf{b} \times \mathbf{b}' = \mathbf{b} \times d\mathbf{b}$ , and the angle of rotation,  $\rho$ , is given by the magnitude of  $\boldsymbol{\rho}$ :  $\rho = |d\mathbf{b}|$ . In other words, keeping the vector  $\mathbf{s}$  fixed in space, a change of its coordinates and, therefore, of the angle  $\phi$  occurs as a consequence of the orbital motion. Equivalently, this change can be described by rotating the spin  $\mathbf{s}$  about the axis  $\boldsymbol{\rho}$  by a negative amount into  $\mathbf{s}'$  (a rotation given by  $-\rho$ ), while keeping the vectors of the frame fixed.

The change  $d\mathbf{s}$  of the spin vector  $\mathbf{s}$  is naturally decomposed into two components, (anti)parallel and orthogonal to the vector  $d\mathbf{b}$ , respectively. Explicitly one finds to first order in  $\rho$ ,

$$\mathbf{s}' - \mathbf{s} \equiv d\mathbf{s} = -(\mathbf{b} \times d\mathbf{b}) \times \mathbf{s} = -m\mathbf{b} + (\mathbf{s} \cdot d\mathbf{b})\mathbf{b}. \quad (16)$$

The angle  $\phi$  picks up a contribution  $d\phi$  which can be determined in the following way: Introduce the projec-

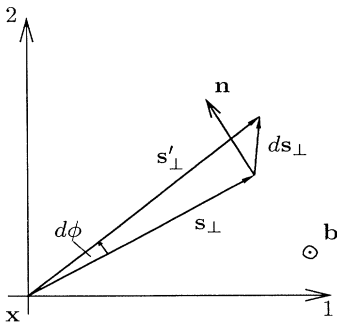


FIG. 1. Projection of the spin vector  $\mathbf{s}$  and its rotated version  $\mathbf{s}'$  onto the  $(\mathbf{e}_1, \mathbf{e}_2)$  plane perpendicular to the direction of the magnetic field  $\mathbf{b}$  at the point  $\mathbf{x}$ . The vectors  $\mathbf{n}$ ,  $\mathbf{b}$ , and  $\mathbf{s}_\perp$  are mutually orthogonal.

tion of the vector  $d\mathbf{s}$  onto the plane perpendicular to the direction of the field  $\mathbf{b}$ , and denote it by  $(d\mathbf{s})_\perp$ ; the local geometry in this plane is depicted in Fig. 1. The resulting increment of the angle  $\phi$  is equal to

$$d\phi = \frac{(d\mathbf{s})_\perp \cdot \mathbf{n}}{|\mathbf{s}_\perp|}, \quad (17)$$

where the unit vector  $\mathbf{n}$  is perpendicular to both  $\mathbf{b}$  and  $\mathbf{s}_\perp$ , that is,  $\mathbf{n} = |\mathbf{s}_\perp|^{-1} \mathbf{b} \times \mathbf{s}_\perp$ . Using Eq. (16) in Eq. (17), one obtains

$$d\phi = -\frac{m}{|\mathbf{s}_\perp|} d\mathbf{b} \cdot (\mathbf{b} \times \mathbf{c}), \quad (18)$$

which agrees with the result shown in Eqs. (7) after using Eq. (10).

## B. Hamiltonian formulation

In the following analysis it will be convenient to have a Hamiltonian formulation of the equations of motion. Such a formulation will allow us to apply Lie transformations [12] in carrying out the perturbation calculation, although the adiabatic nature of the problem will lead to a slightly nonstandard version of the Lie-transform method. In particular, we will introduce coordinates which are not quite canonical and which lead to an expression for the Poisson bracket which is ordered in the adiabatic expansion parameter  $\epsilon$ . Apart from this feature, we will follow the standard Lie-transform technique of using exponentiated Poisson brackets to generate near-identity canonical transformations.

We begin by expressing the Hamiltonian in the variables  $(\mathbf{x}, \mathbf{k}, \phi', m')$ , where  $\mathbf{k}$  is the kinetic momentum (to be distinguished from the canonical momentum  $\mathbf{p}$  to be introduced later). The Hamiltonian is just the obvious expression for the energy,

$$\begin{aligned} H(\mathbf{x}, \mathbf{k}, \phi', m') &= \frac{\mathbf{k}^2}{2} - \mathbf{B} \cdot \mathbf{s} \\ &= \frac{\mathbf{k}^2}{2} - [(s^2 - m'^2)^{1/2} (B_x \cos \phi' + B_y \sin \phi') \\ &\quad + m' B_z], \end{aligned} \quad (19)$$

which, as one can easily show, produces the equations of motion (1) precisely except for the  $\epsilon$ . The canonical variables for the slow degrees of freedom are  $(\mathbf{x}, \mathbf{k})$  and those for the fast (precession) degree of freedom are the previously mentioned set of variables  $(\phi', m')$ , which refer to the space axes  $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ . The components of the vector  $\mathbf{s}$  read in these variables:

$$\begin{aligned} s_x &= (s^2 - m'^2)^{1/2} \cos \phi', \\ s_y &= (s^2 - m'^2)^{1/2} \sin \phi', \\ s_z &= m'. \end{aligned} \quad (20)$$

To make the  $\epsilon$  come out right, one can resort to the variational principle [13]

$$\delta \int \mathbf{k} \cdot d\mathbf{x} + \epsilon m' d\phi' - H dt = 0, \quad (21)$$

instead of using Hamilton's equations directly. Because

of the  $\epsilon$  multiplying the term  $m'd\phi'$ , the set  $(\phi', m')$  represents noncanonical variables of a simple kind. The nonvanishing Poisson brackets are

$$\{x_i, k_j\} = \delta_{ij}, \quad \{\phi', m'\} = \frac{1}{\epsilon}, \quad (22)$$

and the formula for the Poisson bracket of any two functions  $f$  and  $g$  of  $(\mathbf{x}, \mathbf{k}, \phi', m')$  is

$$\{f, g\} = \frac{\partial f}{\partial \mathbf{x}} \cdot \frac{\partial g}{\partial \mathbf{k}} - \frac{\partial f}{\partial \mathbf{k}} \cdot \frac{\partial g}{\partial \mathbf{x}} + \frac{1}{\epsilon} \left[ \frac{\partial f}{\partial \phi'} \frac{\partial g}{\partial m'} - \frac{\partial f}{\partial m'} \frac{\partial g}{\partial \phi'} \right]. \quad (23)$$

In preparation for the classical perturbation analysis, it is desirable to transform the Hamiltonian to action-angle variables for the fast degree of freedom. These action-angle variables are  $(\phi, m)$ , which are related to  $(\phi', m')$  by Eqs. (3) and (20). We combine these equations into the form

$$\begin{aligned} (s^2 - m'^2)^{1/2} \cos \phi &= (s^2 - m'^2)^{1/2} (R_{11} \cos \phi' + R_{21} \sin \phi') + R_{31} m', \\ (s^2 - m'^2)^{1/2} \sin \phi &= (s^2 - m'^2)^{1/2} (R_{12} \cos \phi' + R_{22} \sin \phi') + R_{32} m', \end{aligned} \quad (24)$$

$$m = (s^2 - m'^2)^{1/2} (R_{13} \cos \phi' + R_{23} \sin \phi') + R_{33} m'.$$

These equations give us an implicit specification of the transformation  $(\phi', m') \rightarrow (\phi, m)$ ; although the transformation itself only involves two variables, we have three equations here because the equations are not independent.

In these equations the matrix  $\mathbf{R}$  depends on the slow variable  $\mathbf{x}$ . If  $\mathbf{x}$  is considered frozen or as a constant parameter, then Eqs. (24) specify a canonical transformation in the one fast degree of freedom,  $(\phi', m') \rightarrow (\phi, m)$ . A straightforward but lengthy way to show this is to solve explicitly for the new variables as functions of the old, and then to compute the Poisson brackets. A more physical way to see the same thing is simply to note that all we have done in Eqs. (24) is to refer the fast variables to the body  $z$  axis instead of to the space  $z$  axis; by spatial symmetry, if the component of the spin along one axis and the corresponding azimuthal angle are canonically conjugate, then the same must be true for any other axis, because the Poisson-bracket relations do not distinguish any preferred direction in space. Of course, in our physical problem, the body  $z$  axis is distinguished from all other axes by the fact that it points along the magnetic field, but this fact only affects the Hamiltonian, not the Poisson-bracket relations.

To say this another way, the symplectic two-form for the fast variables satisfies

$$dp \wedge dq = dm' \wedge d\phi' = s d^2\Omega,$$

where  $d^2\Omega$  is the element of solid angle in spin space. By symmetry, we must also have  $dm \wedge d\phi = s d^2\Omega$ , so

$$dm' \wedge d\phi' = dm \wedge d\phi, \quad (25)$$

or, equivalently,

$$m'd\phi' = m d\phi + df, \quad (26)$$

where  $f$  is a scalar parametrized by the matrix  $\mathbf{R}$  and otherwise depending on any independent pair of variables chosen from the set  $(\phi', m', \phi, m)$ . If  $(\phi', \phi)$  are chosen, then  $f(\phi', \phi, \mathbf{R})$  is Goldstein's  $F_1$ -type generating function [14] of the canonical transformation  $(\phi', m') \rightarrow (\phi, m)$  in the one fast degree of freedom. Later, we will regard  $f$  as a function of  $(\phi, m, \mathbf{R})$ .

Unfortunately, if the variables  $\mathbf{x}$  are unfrozen, so that Eqs. (24) can be regarded as the specification of a transformation

$$(\mathbf{x}, \mathbf{k}, \phi', m') \rightarrow (\mathbf{x}, \mathbf{k}, \phi, m),$$

then the transformation in all the degrees of freedom (both slow and fast) is no longer canonical. Therefore, we cannot simply transform the Hamiltonian to the new variables and expect Hamilton's equations to be valid. However, there is no trouble in transforming a variational principle such as in Eq. (21), since variational principles retain their validity under arbitrary coordinate transformations. Therefore, our strategy will be to transform the variational principle in Eq. (21) to the new coordinates  $(\mathbf{x}, \mathbf{k}, \phi, m)$  and then to perform subsequent transformations to restore the canonical form.

We begin with the differential form  $m'd\phi'$  in Eq. (21). We must not be too quick to substitute Eq. (26), however, because in Eq. (26), the variable  $\mathbf{x}$  was considered frozen, so that when the operator  $d$  acts on  $f$ , no terms involving  $d\mathbf{x}$  appear. If we unfreeze  $\mathbf{x}$ , then Eq. (26) must be replaced by

$$m'd\phi' = m d\phi + df - \sum_{i,j} \frac{\partial f}{\partial R_{ij}} dR_{ij}. \quad (27)$$

In this equation, the meaning of the derivatives of  $f$  with respect to  $\mathbf{R}$  depends on the independent variables upon which  $f$  is considered to depend. For example, considering  $f$  to be a function of  $(\phi, m, \mathbf{R})$ , the derivatives of  $f$  with respect to  $\mathbf{R}$  are taken at constant  $(\phi, m)$ . Equation (27) is the equation we must use to eliminate  $m'd\phi'$  in Eq. (21).

Next, we transform the Hamiltonian to the new variables. This is easy, and we find

$$H = \mathbf{k}^2/2 - mB(\mathbf{x}). \quad (28)$$

Notice that  $H$  does not depend on the new angle  $\phi$ ; however, we cannot conclude from this that  $m$  is conserved, because the variables  $(\mathbf{x}, \mathbf{k}, \phi, m)$  do not form a canonical set.

Making these transformations and neglecting the term  $df$  which will not contribute to the equations of motion, we find the variational principle in the new coordinates:

$$\delta \int \mathbf{k} \cdot d\mathbf{x} + \epsilon m d\phi - \epsilon \sum_{i,j} \frac{\partial f}{\partial R_{ij}} (\nabla R_{ij}) \cdot d\mathbf{x} - H dt = 0. \quad (29)$$

If it were not for the term involving the derivatives of  $f$ , the variables  $(\mathbf{x}, \mathbf{k}, \phi, m)$  would satisfy Poisson-bracket re-

lations which would be just like Eqs. (22) without the primes, i.e., they would be canonical apart from the factor of  $\epsilon$ . In the presence of this term, however, the variables  $(\phi, m)$  satisfy a more complex set of Poisson-bracket relations. As we will see, the term involving the derivatives of  $f$  contains the gauge structure for this problem, and much interesting physics.

The next step is to make the variational principle Eq. (29) explicit by working out the term involving the derivatives of  $f$ . It would be logical to do this by first finding the generating function  $F_1$  corresponding to Eq. (24), then by transforming it from the variables  $(\phi', \phi)$  to the variables  $(\phi, m)$ , and, finally, by differentiating it with respect to  $\mathbf{R}$  (in an appropriate sense, since the components  $R_{ij}$  are not independent). Also the obvious way to find the generating function would be to work backwards from the transformation equations (24) themselves, using the relations  $m' = \partial F_1 / \partial \phi'$ ,  $m = -\partial F_1 / \partial \phi$ . This calculation, while straightforward in principle, is very lengthy in practice.

Fortunately, there is an easier way to fill in the unknown term in the variational principle, which is to work backwards from the known equations of motion (7). We write

$$\mathbf{Q} = - \sum_{i,j} \frac{\partial f}{\partial R_{ij}} \nabla R_{ij} , \quad (30)$$

for the unknown vector appearing in Eq. (29). The vector  $\mathbf{Q}$  depends on  $(\phi, m, \mathbf{x})$  (because these are the variables upon which  $f$  depends), but not on  $\mathbf{k}$ . Then, by working out the Euler-Lagrange equations from Eq. (29), we obtain

$$\begin{aligned} \ddot{\mathbf{x}} &= m \nabla B + \dot{\mathbf{x}} \times (\nabla \times \mathbf{Q}) - \frac{\partial \mathbf{Q}}{\partial \phi} \dot{\phi} - \frac{\partial \mathbf{Q}}{\partial m} \dot{m} . \\ \dot{m} &= \dot{\mathbf{x}} \cdot \frac{\partial \mathbf{Q}}{\partial \phi} , \\ \dot{\phi} &= - \frac{B}{\epsilon} - \dot{\mathbf{x}} \cdot \frac{\partial \mathbf{Q}}{\partial m} . \end{aligned} \quad (31)$$

By comparing these equations to Eqs. (7),  $\mathbf{Q} \cdot d\mathbf{x}$  can be determined to within an exact differential. Carrying out this comparison involves some algebra, in which the vector identities of the Appendix are useful. We will omit the details of this calculation, and simply announce the result, which is

$$\mathbf{Q} = m \mathbf{A} - (s^2 - m^2)^{1/2} \nabla \mathbf{b} \cdot \mathbf{a} . \quad (32)$$

Combining Eqs. (29), (30), and (32), we can write the variational principle in the form

$$\delta \int \mathbf{p} \cdot d\mathbf{x} + \epsilon m d\phi - H dt = 0 , \quad (33)$$

where we have introduced the canonical momentum

$$\mathbf{p} = \mathbf{k} + \epsilon [m \mathbf{A} - (s^2 - m^2)^{1/2} \nabla \mathbf{b} \cdot \mathbf{a}] . \quad (34)$$

The reason for using  $\mathbf{P}$  instead of  $\mathbf{K}$  is that the variables  $(\mathbf{x}, \mathbf{p}, \phi, m)$  satisfy the simple set of Poisson-bracket relations,

$$\{x_i, p_j\} = \delta_{ij}, \quad \{\phi, m\} = \frac{1}{\epsilon} , \quad (35)$$

i.e., these variables are canonical apart from the constant factor of  $\epsilon$ . In terms of the new variables, the Hamiltonian has the form

$$\begin{aligned} H(\mathbf{x}, \mathbf{p}, \phi, m) \\ = \frac{1}{2} \{ \mathbf{p} - \epsilon [m \mathbf{A} - (s^2 - m^2)^{1/2} \nabla \mathbf{b} \cdot \mathbf{a}] \}^2 - mB . \end{aligned} \quad (36)$$

It is straightforward to check that this Hamiltonian, combined with the fundamental Poisson brackets in Eq. (35), reproduces exactly the equations of motion (7).

On the basis of experience with Berry's phase in adiabatic spin systems and in related Born-Oppenheimer systems, one would expect to see a gauge potential like  $m \mathbf{A}$  in the Hamiltonian precisely where our Hamiltonian has such a term. We see, however, that this is not the only term present; there is another term,  $(s^2 - m^2)^{1/2} \nabla \mathbf{b} \cdot \mathbf{a}$ , which is of the same order and which is gauge invariant. We notice also that Eq. (36) gives the Hamiltonian as a power series in  $\epsilon$ .

### C. Lie-transform perturbation theory

This section deals with a version of the Lie-transform perturbation theory which is adapted to the problem at hand. The basic ideas of this method are explained briefly, and, along with this discussion, the necessary formulas of the theory are collected.

In perturbation theory one typically starts with a system the behavior of which is known exactly. For classical Hamiltonian systems, this situation effectively corresponds to periodic (or integrable) motion of the unperturbed system. Then, after adding a small perturbation to the system, one investigates the modification of its properties. One approach to perturbation theory is to express the perturbation of the Hamiltonian function in action-angle variables and to remove, by a sequence of canonical transformations, the resulting angle dependence up to the desired order. To do so, the method of Lie transforms is advantageous: on the one hand, the explicit determination of the intermediate canonical transformations is simplified by the fact that the generators of the required near-identity transformations do not depend on both old and new variables; on the other hand, it is relatively simple to keep track of higher-order corrections in a systematic way. In the following analysis we use a modification of the Lie-transform method of Dragt and Finn [15], which has been reviewed by Cary [12].

The reason we must modify standard perturbation techniques is that our problem is an adiabatic problem, i.e., one in which there is a separation of time scales among the different degrees of freedom. Thus, our perturbation effects are not only represented by correction terms in the Hamiltonian, but also by the fact that the Poisson brackets among the phase-space coordinates themselves are ordered in  $\epsilon$ , as shown in Eq. (35). In our problem, the Poisson brackets are fairly simple (they are just constants); in other adiabatic problems, such as classical guiding-center motion [10], more complex perturbation methods are called for.

In the following perturbation analysis, we carry out a canonical transformation which we write in the form

$$(\mathbf{x}, \mathbf{p}, \phi, m) \rightarrow (\mathbf{X}, \mathbf{P}, \Phi, M),$$

where capitalized symbols are used for the new canonical coordinates in which the Hamiltonian is ignorable in the new rapid angle  $\Phi$ . Thus, in the new coordinates, the conjugate momentum  $M$  is a constant of motion.

Our first step is to define the Lie generator  $w = w(\mathbf{x}, \mathbf{p}, \phi, m)$ , which is a function, and the Lie operator  $L_w$  associated with it. This operator acts on phase-space functions  $A(\mathbf{x}, \mathbf{p}, \phi, m)$  according to

$$\begin{aligned} L_w A = \{w, A\} &= \frac{1}{\epsilon} \{w, A\}_f + \{w, A\}_s, \\ &= \frac{1}{\epsilon} L_{wf} A + L_{ws} A, \end{aligned} \quad (37)$$

where the subscripts  $s, f$  stand for ‘‘slow’’ and ‘‘fast’’ and where

$$\begin{aligned} L_{wf} A = \{w, A\}_f &= \frac{\partial w}{\partial \phi} \frac{\partial A}{\partial m} - \frac{\partial w}{\partial m} \frac{\partial A}{\partial \phi}, \\ L_{ws} A = \{w, A\}_s &= \frac{\partial w}{\partial \mathbf{x}} \cdot \frac{\partial A}{\partial \mathbf{p}} - \frac{\partial w}{\partial \mathbf{p}} \cdot \frac{\partial A}{\partial \mathbf{x}}. \end{aligned} \quad (38)$$

Then, for any Lie generator  $w$  and any parameter  $\lambda$ , the operator  $e^{\lambda L_w}$  is a canonical transformation, in the sense that if  $A$  is a function, then  $e^{\lambda L_w} A$  is the new function formed by transforming  $A$  to the new set of canonical variables. Here, the canonical transformation operator is to be expanded in powers of  $\epsilon$ , so one should set  $\lambda = \epsilon^2$  or some higher power of  $\epsilon$ , since the operator  $L_w$  is already  $O(1/\epsilon)$ . This splitting of the Poisson bracket into a slow and a fast part is a slightly unusual strategy in Lie transforms.

Now we introduce a sequence of Lie generators,  $w_1, w_2, \dots$ , and associated Lie operators,  $L_1 = L_{w_1}, L_2 = L_{w_2}$ , etc., and define an overall canonical transformation operator  $T$  by

$$\begin{aligned} T &= \exp(-\epsilon^2 L_1) \exp(-\epsilon^3 L_2) \exp(-\epsilon^4 L_3) \cdots \\ &= \exp(-\epsilon L_{1f} - \epsilon^2 L_{1s}) \exp(-\epsilon^2 L_{2f} - \epsilon^3 L_{2s}) \cdots \\ &= 1 - \epsilon L_{1f} + \epsilon^2 (-L_{1s} + \frac{1}{2} L_{1f}^2 - L_{2f}) + O(\epsilon^3). \end{aligned} \quad (39)$$

Likewise, it is assumed that for the old and new Hamiltonians,  $H$  and  $K$ , the expansions

$$\begin{aligned} H &= H_0 + \epsilon H_1 + \epsilon^2 H_2 + \cdots, \\ K &= K_0 + \epsilon K_1 + \epsilon^2 K_2 + \cdots \end{aligned} \quad (40)$$

hold.

All quantities in the relation  $K = T^{-1}H$  between the old and new Hamiltonians are expressed as power series in the perturbation parameter. The expansion of  $T^{-1}$  to second order in  $\epsilon$  is obtained from Eq. (39) by flipping each minus sign into a plus sign. By equating the coefficients of matching powers of  $\epsilon$ , one finds the conditions

$$\begin{aligned} K_0 &= H_0, \\ K_1 &= H_1 + L_{1f} H_0, \\ K_2 &= H_2 + L_{1f} H_1 + (L_{1s} + \frac{1}{2} L_{1f}^2 + L_{2f}) H_0. \end{aligned} \quad (41)$$

These equations contain the Lie generators  $w_1, w_2, \dots$  and the terms  $K_0, K_1, \dots$  of the new Hamiltonian as unknowns. The equations represent a hierarchy of differential equations for the Lie generators  $w_1, w_2, \dots$  and algebraic equations for the functions  $K_0, K_1, \dots$ ; the requirements that  $K$  not depend on the angle variable and that no secular terms show up (cf. [12]) imply that these equations can be uniquely split into their averaged and oscillatory parts, giving a number of equations which is equal to the number of unknowns. The method of solution is to start with the lowest-order equation of Eqs. (41) and to work one’s way to the higher-order equations step by step.

Eventually, the generators  $w_1, w_2, \dots$  can be used to specify the explicit form of the canonical transformation

$$(\mathbf{x}, \mathbf{p}, \phi, m) \rightarrow (\mathbf{X}, \mathbf{P}, \Phi, M).$$

Writing  $\mathbf{z}$  for the old variables and  $\mathbf{Z}$  for the new ones, we obtain

$$\mathbf{Z} = T\mathbf{z} = \mathbf{z} - \epsilon L_{1f} \mathbf{z} + \epsilon^2 (-L_{1s} + \frac{1}{2} L_{1f}^2 - L_{2f}) \mathbf{z} + \cdots. \quad (42)$$

#### D. The perturbation calculation

In this section we present a classical perturbation analysis of the spinning particle under the assumption that the time scales of the orbital and the precessional motion are clearly separated. In other words, the Hamiltonian Eq. (36), which is written as a power series in the small parameter  $\epsilon$ , is subjected to a systematic perturbation expansion. The calculation will include the determination of both the averaged Hamiltonian and the explicit form of the new canonical variables in which it is expressed.

It is the goal of the following calculation to determine the quantities  $K_0, K_1$ , and  $K_2$  and the new canonical coordinates  $\mathbf{Z}$  for the Hamiltonian of Eq. (36), which we write in the form

$$\begin{aligned} H_0 &= \frac{\mathbf{p}^2}{2} - mB, \\ H_1 &= -\mathbf{p} \cdot [m \mathbf{A} - (s^2 - m^2)^{1/2} \nabla \mathbf{b} \cdot \mathbf{a}], \\ H_2 &= \frac{1}{2} [m \mathbf{A} - (s^2 - m^2)^{1/2} \nabla \mathbf{b} \cdot \mathbf{a}]^2. \end{aligned} \quad (43)$$

This problem is simpler than the general case since the expansion of the Hamiltonian terminates after the second-order term.

It follows from the first of Eqs. (41) that to lowest order, the old and new Hamiltonian are equal:

$$K_0(\mathbf{X}, \mathbf{P}, M) = \frac{\mathbf{P}^2}{2} - MB(\mathbf{X}), \quad (44)$$

or, more exactly, they have the same functional dependence on their independent variables [respectively,  $(\mathbf{x}, \mathbf{p}, \phi, m)$  and  $(\mathbf{X}, \mathbf{P}, \Phi, M)$ ]. As a consequence, the second of Eqs. (41) can be written

$$\begin{aligned}
L_{1f}H_0 &= \{w_1, H_0\}_f \\
&= -B \frac{\partial w_1}{\partial \phi} = K_1 - H_1 \\
&= K_1 + \mathbf{p} \cdot [m \mathbf{A} - (s^2 - m^2)^{1/2} \nabla \mathbf{b} \cdot \mathbf{a}] .
\end{aligned} \tag{45}$$

Averaging both sides with respect to the fast variable  $\phi$  and demanding that  $K_1$  be ignorable in  $\phi$ , one finds

$$K_1(\mathbf{X}, \mathbf{P}, M) = -M \mathbf{P} \cdot \mathbf{A}(\mathbf{X}) . \tag{46}$$

Taking averaged and oscillatory parts is easy in such expressions, because any quantity involving odd powers of the vectors  $\mathbf{a}$  and  $\mathbf{c}$  is purely oscillatory, and terms not involving  $\mathbf{a}$  or  $\mathbf{c}$  at all are purely averaged. When dealing with quadratic expressions involving  $\mathbf{a}$  and  $\mathbf{c}$ , some terms turn out to be purely averaged. The general problem of projecting out averaged and oscillatory parts is a problem in Fourier series.

Equation (46) contains the first appearance of the gauge potential in the new Hamiltonian; at this order the same result could have been obtained more easily by a simple-minded method of averaging, in which one just throws away oscillatory terms. In this method, the second term in  $H_1$  in the second of Eqs. (43), being purely oscillatory, is just thrown away. However, this simple-minded method of averaging only works at lowest order; to get higher-order corrections, one must be more systematic, such as by using the Lie perturbation expansion developed here.

The oscillatory terms in Eq. (45) allow us to determine the Lie generator  $w_1$ . Projecting out the oscillatory part, we obtain

$$-B \frac{\partial w_1}{\partial \phi} = -(s^2 - m^2)^{1/2} (\mathbf{p} \cdot \nabla \mathbf{b} \cdot \mathbf{a}) , \tag{47}$$

which we integrate to obtain

$$w_1 = \frac{(s^2 - m^2)^{1/2}}{B} (\mathbf{p} \cdot \nabla \mathbf{b} \cdot \mathbf{c}) . \tag{48}$$

This generator is needed in the calculation of the second-order Hamiltonian, to which we now turn.

Plugging in the expression for  $K_1$  into the third of Eqs. (41) leads to

$$L_{2f}H_0 = \{w_2, H_0\}_f = K_2 - H_2 - \frac{1}{2} L_{1f}(H_1 + K_1) - L_{1s}H_0 . \tag{49}$$

$$\begin{aligned}
w_2 &= (s^2 - m^2)^{1/2} \left[ -\frac{m}{B} (\mathbf{A} \cdot \nabla \mathbf{b} \cdot \mathbf{c}) + \frac{1}{B^3} (\mathbf{p} \cdot \nabla B) (\mathbf{p} \cdot \nabla \mathbf{b} \cdot \mathbf{a}) - \frac{1}{B^2} (\mathbf{p} \mathbf{p} : \nabla \nabla \mathbf{b} \cdot \mathbf{a}) \right. \\
&\quad \left. - \frac{m}{B^2} (\nabla B \cdot \nabla \mathbf{b} \cdot \mathbf{a}) \right] + \frac{s^2 - m^2}{4} (\nabla \mathbf{b} \cdot \mathbf{a}) \cdot (\nabla \mathbf{b} \cdot \mathbf{c}) .
\end{aligned} \tag{55}$$

Now we can use the Lie generators  $w_1$  and  $w_2$  to find the explicit form of the canonical transformation

$$(\mathbf{x}, \mathbf{p}, \phi, m) \rightarrow (\mathbf{X}, \mathbf{P}, \Phi, M) .$$

Probably the most interesting part of this transformation is the formula which expresses  $M$  as a function of the old vari-

ables. In order to find  $K_2$ , we need the averaged part of this expression. The terms  $L_{2f}H_0$  and  $L_{1s}H_0$  are purely oscillatory, and can be dropped. Denoting the average by an overbar, we can write the averaged part of  $H_2$  as

$$\overline{H_2} \equiv \frac{1}{2\pi} \int_0^{2\pi} d\phi H_2 = \frac{m^2 \mathbf{A}^2}{2} + V_g , \tag{50}$$

where the first term completes the square for the vector potential and the second term  $V_g$  is the ‘‘geometric scalar potential’’ (denoted  $\Phi$  by Berry in [1]):

$$V_g(\mathbf{x}, m) = \frac{1}{4} (s^2 - m^2) \left[ \frac{\partial b_i}{\partial x_j} \frac{\partial b_i}{\partial x_j} \right] . \tag{51}$$

In deriving Eq. (51), we used the identity

$$\overline{\mathbf{a} \mathbf{a}} = \overline{\mathbf{c} \mathbf{c}} = \frac{1}{2} (1 - \mathbf{b} \mathbf{b}) . \tag{52}$$

However, the scalar potential in Eq. (51) is not the only second-order contribution to  $K_2$ , since the term  $\frac{1}{2} L_{1f}(H_1 + K_1)$  in Eq. (49) also has an averaged part, representing a nonlinear beating or resonance between two first-order terms. Such terms, for example, are known to be responsible for ponderomotive effects on electrons in light waves. Explicitly, we find

$$V_2(\mathbf{x}, \mathbf{p}, m) = \frac{1}{2} \overline{L_{1f}(H_1 + K_1)} = -\frac{m}{2B} |\mathbf{p} \cdot \nabla \mathbf{b}|^2 . \tag{53}$$

Altogether, the new second-order Hamiltonian is given by

$$\begin{aligned}
K(\mathbf{X}, \mathbf{P}, M) &= \frac{1}{2} (\mathbf{P} - \epsilon M \mathbf{A})^2 - MB \\
&\quad + \frac{\epsilon^2}{4} (s^2 - M^2) \left[ \frac{\partial b_i}{\partial x_j} \frac{\partial b_i}{\partial x_j} \right] \\
&\quad - \epsilon^2 \frac{M}{2B} |\mathbf{P} \cdot \nabla \mathbf{b}|^2 ,
\end{aligned} \tag{54}$$

which has the same structure as the result obtained by Berry in [1] and [2], apart from the extra term  $V_2$ , which later on will be shown to have a quantum-mechanical equivalent. We will call this extra term the ‘‘new term,’’ since it does not appear in the analysis of Aharonov and Stern [2]. This Hamiltonian is one of our main results.

By projecting out the oscillatory part of Eq. (49), we obtain an equation for the second-order generator  $w_2$ . This generator is not needed for the determination of the second-order Hamiltonian  $K_2$ , but it is needed to find the canonical transformation connecting old and new variables to second order. After some algebra, we find



ables, since  $M$  is the adiabatic invariant and is a formal constant of motion. We replace  $\mathbf{Z}$  with  $M$  and  $\mathbf{z}$  with  $m$  in Eq. (42), noting that all slow Poisson brackets involving  $m$  vanish, and we obtain

$$M = m - \epsilon \{w_1, m\}_f + \epsilon^2 \left( \frac{1}{2} \{w_1, \{w_1, m\}_f\}_f - \{w_2, m\}_f \right) + \dots \quad (56)$$

Evaluating this explicitly, we find

$$\begin{aligned} M = m - \epsilon \frac{(s^2 - m^2)^{1/2}}{B} (\mathbf{p} \cdot \nabla \mathbf{b} \cdot \mathbf{a}) \\ + \epsilon^2 \left\{ -\frac{m}{2B^2} |\mathbf{p} \cdot \nabla \mathbf{b}|^2 + (s^2 - m^2)^{1/2} \left[ \frac{m}{B} (\mathbf{A} \cdot \nabla \mathbf{b} \cdot \mathbf{a}) + \frac{\mathbf{p} \cdot \nabla B}{B^3} (\mathbf{p} \cdot \nabla \mathbf{b} \cdot \mathbf{c}) - \frac{1}{B^2} (\mathbf{p} \mathbf{p} \cdot \nabla \nabla \mathbf{b} \cdot \mathbf{c}) \right. \right. \\ \left. \left. - \frac{m}{B^2} (\nabla B \cdot \nabla \mathbf{b} \cdot \mathbf{c}) \right] - \frac{s^2 - m^2}{4B} (|\nabla \mathbf{b} \cdot \mathbf{a}|^2 - |\nabla \mathbf{b} \cdot \mathbf{c}|^2) \right\}. \quad (57) \end{aligned}$$

Of course  $M$ , the adiabatic invariant, cannot be gauge dependent, so the term involving  $\mathbf{A}$  at second order is a slight worry. But one must keep in mind that the canonical momentum  $\mathbf{p}$  itself is not gauge invariant; instead, the kinetic momentum,  $\mathbf{k} = \dot{\mathbf{x}} = \mathbf{p} - \epsilon m \mathbf{A}$  is gauge invariant. Indeed, if we express  $M$  in terms of  $\mathbf{k}$  instead of  $\mathbf{p}$ , we find that it is gauge invariant to the order we have calculated.

The fact that  $M$  is a constant of motion can be checked directly by differentiating Eq. (57) with respect to time and using Eqs. (7). Before doing this, it is convenient to transform the expression for  $M$  by eliminating  $\mathbf{p}$  in favor of  $\mathbf{k}$ . We will not repeat this calculation here, apart from noting that if the expression in Eq. (57), which is truncated at second order, is differentiated with respect to time, then the result is nonzero at order  $\epsilon^2$ . This nonzero result can be pushed to arbitrarily high order in  $\epsilon$  by going to higher order in the expansion given in Eq. (57).

Now we turn to the transformation of the variables  $\mathbf{X}$  and  $\mathbf{P}$ , noting that all fast Poisson brackets in Eq. (42) vanish. Thus, we find

$$\begin{aligned} \mathbf{X} = \mathbf{x} - \epsilon^2 \{w_1, \mathbf{x}\}_s = \mathbf{x} + \epsilon^2 \frac{\partial w_1}{\partial \mathbf{p}} \\ = \mathbf{x} + \epsilon^2 \frac{(s^2 - m^2)^{1/2}}{B} \nabla \mathbf{b} \cdot \mathbf{c}. \quad (58) \end{aligned}$$

We see that the position of the particle  $\mathbf{x}$  has small, rapid oscillations of order  $\epsilon^2$ , while the averaged position is given by  $\mathbf{X}$ . If the particle had a nonzero charge, the oscillations would be first order in  $\epsilon$ , as in classical guiding-center motion. As for the canonical momentum  $\mathbf{P}$ , we find

$$\begin{aligned} \mathbf{P} = \mathbf{p} - \epsilon^2 \nabla w_1 \\ = \mathbf{p} + \epsilon^2 (s^2 - m^2)^{1/2} \left[ \frac{\nabla B}{B^2} (\mathbf{p} \cdot \nabla \mathbf{b} \cdot \mathbf{c}) - \frac{1}{B} \mathbf{p} \cdot \nabla \nabla \mathbf{b} \cdot \mathbf{c} \right. \\ \left. - \frac{1}{B} \mathbf{A} (\mathbf{p} \cdot \nabla \mathbf{b} \cdot \mathbf{a}) \right]. \quad (59) \end{aligned}$$

Again, there is no need to worry about the gauge-dependent term at  $O(\epsilon^2)$ , since it is not  $\mathbf{P}$ , but rather the vector  $\mathbf{K} = \mathbf{P} - \epsilon M \mathbf{A}$ , which is gauge invariant. In fact,

transforming this  $K$  back to the variables  $(\mathbf{x}, \mathbf{p}, \phi, m)$ , one finds that the term in question is needed to make  $\mathbf{K}$  gauge invariant. Explicitly, we find

$$\mathbf{K} = \mathbf{k} + \epsilon^2 (s^2 - m^2)^{1/2} \left[ \frac{\nabla B}{B^2} (\mathbf{k} \cdot \nabla \mathbf{b} \cdot \mathbf{c}) - \frac{1}{B} (\mathbf{k} \cdot \nabla \mathbf{b} \cdot \mathbf{c}) \right], \quad (60)$$

which shows that the particle's velocity also has small oscillations at second order. These are removed by the Lie transforms, which leave the averaged variable  $\mathbf{K}$ .

### III. SEMICLASSICAL APPROACH

In this section the neutral particle with spin is investigated from semiclassical and adiabatic points of view. The spin degree of freedom is considered as a quantum-mechanical variable, whereas the orbital motion of the particle, corresponding to the slow degrees of freedom in the classical model, is treated semiclassically. The length of the spin is a constant (both a constant of motion and a constant in our ordering scheme, i.e., independent of our perturbation parameter  $\epsilon$ ) and is not assumed to be large with respect to  $\hbar$ . Thus, the spin degree of freedom is treated fully quantum mechanical, and all spin quantum numbers are of order unity. In addition, we assume that the magnetic field is large, so that the precession of the spin will represent a rapid degree of freedom relative to the slow orbital degrees of freedom.

We now introduce a formal ordering parameter  $\epsilon$  to represent these ordering assumptions. We begin with the Hamiltonian

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2m_0} - \gamma \mathbf{B} \cdot \hat{\mathbf{s}}, \quad (61)$$

where the hats represent quantum operators, where  $\gamma = ge/2m_s c$ , and where  $\hat{\mathbf{s}}$  is the vector of spin operators with dimensions of angular momentum. We allow the mass  $m_s$  appearing in  $\gamma$  to be different from the mass  $m_0$  of the particle.

In semiclassical treatments, one often treats  $\hbar$  as a variable, and one considers the limit  $\hbar \rightarrow 0$ , while classical

quantities such as energy and angular momentum are held fixed. Here we prefer to keep  $\hbar$  at its physical value, but to replace it in all the usual physical formulas by  $\epsilon\hbar$ , where  $\epsilon$  is dimensionless. This includes, for example, the formulas  $\hat{\mathbf{p}} = -i\hbar\nabla$ ,  $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$ , etc. It also includes the definition of the Weyl transform and the Moyal formula [6]. We also scale the spin  $\hat{\mathbf{s}}$  by  $\epsilon$ , since the quantity  $s/\hbar$  is supposed to be of order unity. Later we will let  $\epsilon \rightarrow 0$ , which will take care of the semiclassical approximation. To take care of the adiabatic approximation, we also replace the magnetic field  $\mathbf{B}$  by  $\mathbf{B}/\epsilon$ . Thus, the product  $\mathbf{B} \cdot \hat{\mathbf{s}}$  is of order  $\epsilon^0$ , since  $\hat{\mathbf{s}}$  contains  $\hbar$ . Having done this, we choose units so that  $\hbar = m_0 = 1$ , and we absorb  $\gamma$  into the definition of  $\mathbf{B}$ . The Hamiltonian becomes

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2} - \mathbf{B} \cdot \hat{\mathbf{s}}; \quad (62)$$

it does not contain  $\epsilon$  explicitly, but the  $\epsilon$  is present in auxiliary formulas such as  $\mathbf{p} = -i\epsilon\nabla$  and the commutation relations

$$[\hat{x}_i, \hat{p}_j] = i\epsilon\delta_{ij}. \quad (63)$$

On the other hand, the spin commutation relations have become

$$[\hat{s}_i, \hat{s}_j] = i\epsilon_{ijk}\hat{s}_k, \quad (64)$$

since the scaling  $\hat{\mathbf{s}} \rightarrow \epsilon\hat{\mathbf{s}}$ ,  $\hbar \rightarrow \epsilon\hbar$  has allowed us to cancel the  $\epsilon$ . Our object will be to expand the solutions of the Schrödinger equation  $\hat{H}|\psi\rangle = E|\psi\rangle$  in powers of  $\epsilon$ , in which variables such as the energy, momentum, orbital angular momentum, etc., are considered to be of order  $\epsilon^0$ , i.e., they are held fixed as  $\epsilon \rightarrow 0$ . This would be a standard semiclassical expansion, except for the largeness of  $\mathbf{B}$ ; alternatively, it is equivalent to expanding in all the  $\hbar$ 's in the original problem, except for the one hidden in the term  $\gamma\mathbf{B} \cdot \mathbf{s}$ .

A similar situation arises in problems of atomic scattering [16,17], in which one often applies semiclassical methods to the nuclear motion, in the Born-Oppenheimer approximation. In these treatments, it is customary to expand in the  $\hbar$ 's which occur explicitly in the nuclear Schrödinger equation, but not to expand in the  $\hbar$ 's which are hidden in the Born-Oppenheimer potential-energy functions (which came from solving the electronic Schrödinger equation). We will comment further on our ordering assumptions in Sec. IV.

A matrix representation of the spin operators  $\hat{\mathbf{s}}$  referring to the fixed space axes turns the Schrödinger equation into a set of coupled-wave equations,

$$\hat{H}(\hat{\mathbf{x}}, \hat{\mathbf{p}})\Psi = E\hat{1}\Psi, \quad (65)$$

where  $\hat{H}$  and  $\hat{1}$  (the identity matrix) are matrices in spin space and  $\Psi$  is a spinor whose components can be thought of as wave functions depending on  $\mathbf{x}$ . Our treatment of this equation will follow our earlier work [5-7] on the semiclassical analysis of coupled-wave equations, except that we will present some new techniques for carrying out the necessary calculations.

### A. Diagonalizing coupled-wave equations

We proceed from a general point of view, writing some given set of coupled-wave equations in the form

$$\hat{D}'(\hat{\mathbf{x}}, \hat{\mathbf{p}})\Psi = 0, \quad (66)$$

where  $\hat{D}'$  is a matrix of orbital operators, i.e., functions of  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{p}}$ . The prime indicates that these equations are in their original form, and distinguishes these equations from a transformed version to be introduced below. An example of such equations is given by Eq. (65), where  $\hat{D}' = \hat{H} - E\hat{1}$ .

In the first step we carry out a preparatory transformation which has the effect of diagonalizing  $\hat{D}'$  to lowest order. This transformation is less ambitious than the transformations introduced in [6,7], which were designed to diagonalize the matrix of wave operators to all orders. The preparatory transformation will later be followed by a sequence of further transformations which diagonalize the operator matrix to higher orders; these subsequent transformations are near-identity transformations, and can be developed in terms of exponential series of commutators, much as in the classical Lie-transform technique. The separation of the finite, preparatory transformation from the sequence of near-identity transformations has many advantages in terms of the overall elegance of the formalism and the computational effort involved in actual calculations, and it reveals the analogies with the classical calculations we have presented in Sec. II.

The preparatory transformation is specified by a unitary matrix of operators  $\hat{W}$ , which is chosen so that the transformed operator matrix,

$$\hat{W}^\dagger \hat{D}' \hat{W} = \hat{D} = \hat{D}_0 + \epsilon\hat{D}_1 + \dots,$$

is diagonal at lowest order (i.e.,  $\hat{D}_0$  is diagonal). As indicated, we drop the prime on  $\hat{D}$  after the preparatory transformation. As in [6,7],  $\hat{W}$  is determined via its symbol matrix  $W = W_0 + \epsilon W_1 + \dots$ , where the lowest-order term  $W_0$  diagonalizes the symbol matrix  $D'$ , i.e.,  $W_0^\dagger D' W_0 = D_0$ . Our notation is to omit the hats on symbol matrices, which are matrices of ordinary functions of  $(\mathbf{x}, \mathbf{p})$ . Thus, the diagonalization of  $D'$  by  $W_0$  is the purely algebraic problem of finding the eigenvectors and eigenvalues of  $D'$  as functions of  $(\mathbf{x}, \mathbf{p})$ .

As for the correction terms  $W_1, W_2$ , etc., in the symbol matrix of  $\hat{W}$ , these need only be chosen so that  $\hat{W}$  is unitary,  $\hat{W}^\dagger \hat{W} = \hat{1}$ , since at this stage we are not attempting to diagonalize  $\hat{D}'$  beyond lowest order. The determination of  $W_1$ , etc., as well as  $D_1, D_2$ , etc., is made with the help of the Moyal formula, as in [6,7]. We do not want to dwell on the details of this process, because the preparatory transformation is easy in the case of the spin Hamiltonian of interest to us.

After the preparatory transformation, we apply a sequence of further unitary transformations specified by a unitary operator  $\hat{U}$  written as a product of successive near-identity unitary transformations,

$$\hat{U} = \hat{U}_1 \hat{U}_2 \hat{U}_3 \dots, \quad (67)$$

where each  $\hat{U}_n$  is specified in terms of a Hermitian operator  $\hat{G}_n$ ,

$$\hat{U}_n = \exp(i\epsilon^n \hat{G}_n). \quad (68)$$

All the operators  $\hat{U}, \hat{U}_n, \hat{G}_n$  are actually matrices of (orbital) operators. These definitions parallel the classical Lie-transformation equations as in Eq. (39), and the purpose is similar: we shall choose the  $\hat{G}_n$ 's so that the transformed wave operator is diagonal to successively higher orders.

We denote the diagonalized wave operator by  $\hat{\Lambda}$ , so that

$$\hat{\Lambda} = \hat{U}^\dagger \hat{D} \hat{U}. \quad (69)$$

Since for any two operators  $\hat{A}, \hat{B}$  we have the identity

$$e^{-i\hat{A}} \hat{B} e^{i\hat{A}} = \hat{B} - i[\hat{A}, \hat{B}] - \frac{1}{2!}[\hat{A}, [\hat{A}, \hat{B}]] + \dots, \quad (70)$$

we can substitute Eq. (67) into Eq. (69) and expand in an infinite series of commutators of the  $\hat{G}_n$ 's with  $\hat{D}$ . This gives

$$\begin{aligned} \hat{\Lambda} &= \hat{D} - i\epsilon[\hat{G}_1, \hat{D}] \\ &+ \epsilon^2(-\frac{1}{2}[\hat{G}_1, [\hat{G}_1, \hat{D}]] - i[\hat{G}_2, \hat{D}]) + \dots \end{aligned} \quad (71)$$

Next, we transcribe this equation into symbols. The symbols of  $\hat{D}$  and  $\hat{\Lambda}$  are ordered,  $D = D_0 + \epsilon D_1 + \dots$ ,  $\Lambda = \Lambda_0 \epsilon \Lambda_1 + \dots$ , and the transcription of the commutators also produces an  $\epsilon$  ordering, in accordance with the following fact: If  $\hat{A}$  and  $\hat{B}$  are matrices of operators, then the symbol matrix of  $[\hat{A}, \hat{B}]$  is given by

$$[A, B] + \frac{i\epsilon}{2}(\{A, B\} - \{B, A\}) + O(\epsilon^2), \quad (72)$$

where the first term is the usual matrix commutator, and where the Poisson bracket of two matrices is defined by  $\{A, B\}_{\alpha\beta} = \{A_{\alpha\gamma}, B_{\gamma\beta}\}$ . With this definition, it is *not* true that  $\{A, B\} = -\{B, A\}$ . Equation (72) follows from the Moyal formula [6].

After transcribing Eq. (71) to symbols, expanding the symbol matrices  $\Lambda$  and  $D$ , and using Eq. (72), we collect the results by orders of  $\epsilon$ . We find

$$\begin{aligned} \Lambda_0 &= D_0, \\ \Lambda_1 &= D_1 - i[G_1, D_0], \\ \Lambda_2 &= D_2 - i[G_1, D_1] + \frac{1}{2}(\{G_1, D_0\} - \{D_0, G_1\}) \\ &\quad - \frac{1}{2}[G_1[G_1, D_0]] - i[G_2, D_0]. \end{aligned} \quad (73)$$

On comparing this with the classical formula (41), we see that the two have the same structure if the fast Poisson-bracket operator  $L_{nf}$  is associated with  $-i[G_n, \ ]$  and the slow one  $L_{ns}$  with  $\frac{1}{2}(\{G_n, \} - \{, G_n\})$ , where the thin space preceding (following) the comma indicates the location of the operand. In other words, the matrix operations (involving the spinor indices) are fast, and the Poisson brackets (involving the orbital variables) are slow.

In analyzing Eqs. (73), the  $D_n$ 's are taken as given, and the  $G_n$ 's are to be determined so that the  $\Lambda_n$ 's are diago-

nal. At order  $\epsilon^0$ ,  $D_0$  is already diagonal by hypothesis, so there is no work to be done. At successive orders, each equation splits into two equations when we project out the diagonal and off-diagonal parts. The diagonal part determines  $\Lambda_n$ , i.e.,  $\Lambda_n$  is the diagonal part of the right-hand side, and the off-diagonal part gives an equation for  $G_n$ , in which the left-hand side vanishes. We see that projecting out the diagonal and off-diagonal parts is the analog of projecting out the averaged and oscillatory parts, respectively, of the terms in the classical perturbation expansion. The off-diagonal equation can only be solved for the off-diagonal elements of  $G_n$ , since the diagonal elements of the commutator  $[G_n, D_0]$  vanish; the diagonal elements of  $G_n$  can be chosen to be anything we like. This freedom is analogous to the arbitrary choice of the constants of integration one finds when solving for the classical Lie generators  $w_n$ . In the two cases, the freedom corresponds to an infinitesimal gauge transformation in the sense of either Berry's phase or Hannay's angles. Finally, the formula for  $G_n$  will involve energy denominators, i.e., the differences between two eigenvalues of the original symbol matrix  $D'$ ; if these eigenvalues come close together, then the perturbation expansion developed here will break down, since the ordering scheme defined by the parameter  $\epsilon$  does not hold any more. The classical analog is the presence of a resonance, which causes the classical perturbation expansion to break down. In the quantum problem, small denominators may signal the presence of a Landau-Zener transition or a mode conversion, a well-known problem which we have studied recently [18] in the multidimensional case. In this paper we assume that the eigenvalues of  $D'$  are well separated, so that Landau-Zener transitions do not occur.

## B. The perturbation calculation

We will now apply the formalism of Sec. III A to the spin Hamiltonian, or, rather, to a slight generalization of it. The generalization consists of writing the Hamiltonian operator matrix in the form

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2} + \hat{V}(\hat{\mathbf{x}}), \quad (74)$$

where  $\hat{V}$  is a matrix depending only on  $\hat{\mathbf{x}}$ . If  $V_{\alpha\beta}(\mathbf{x})$  is identified with  $-(\mathbf{B} \cdot \mathbf{s})_{\alpha\beta}$ , where the  $\alpha\beta$  indices are spinor indices and  $\mathbf{B} = \mathbf{B}(\mathbf{x})$ , then the spin Hamiltonian in Eq. (65) is a special case of Eq. (74). The Hamiltonian in Eq. (74) is also the Hamiltonian for the nuclear motion in the Born-Oppenheimer approximation for molecules, if the electronic wave functions are expanded in a fixed basis (such as harmonic-oscillator eigenstates), instead of the more usual adiabatic basis (which depends on the nuclear coordinates  $\mathbf{x}$ ). Thus, by studying Eq. (74), we cover both the cases of the spin Hamiltonian and the Born-Oppenheimer Hamiltonian.

As above, we will set  $D' = H - E$  to proceed with the perturbation analysis. This analysis was carried out to second order in [7], but we will repeat it here in order to illustrate the simplified formalism of Sec. III A, and to

draw the parallels with the classical calculation. To begin, we carry out the preparatory transformation, for which we require the unitary operator matrix  $\hat{W}$ . Since

$$\hat{D}' = (\hat{\mathbf{p}}^2/2 - E)\hat{1} + \hat{V}(\hat{\mathbf{x}}),$$

only the potential-energy term  $\hat{V}$  is nondiagonal, and  $W(\mathbf{x})$ , the symbol matrix of  $\hat{W}$ , is simply the  $\mathbf{x}$ -dependent matrix which diagonalizes  $V(\mathbf{x})$ . There are no higher-order corrections to  $W$ , i.e.,  $W = W_0$  and  $W_1 = W_2 = \dots = 0$ . The columns of  $W$  are the  $\mathbf{x}$ -dependent eigenvectors of  $\tau^{(\mu)}$  of  $V(\mathbf{x})$ ,

$$V\tau^{(\mu)} = v^{(\mu)}\tau^{(\mu)}, \quad (75)$$

where the eigenvalues are  $v^{(\mu)}(\mathbf{x})$  and  $(\mu)$  is the "polarization index" (in the terminology of [6]), i.e., it labels the eigenvalues and eigenvectors. In the Born-Oppenheimer Hamiltonian, the eigenvalues  $v^{(\mu)}$  are just the electronic potential-energy functions, and in the spin problem,

$$v^{(\mu)}(\mathbf{x}) = -\mu B(\mathbf{x}). \quad (76)$$

where  $B(\mathbf{x}) = |\mathbf{B}(\mathbf{x})|$  and

$$\mu = -s, -s+1, \dots, +s.$$

In Born-Oppenheimer language, the preparatory transformation is just the transformation taking us from some fixed basis to the adiabatic basis. In the language of the classical perturbation theory developed in Sec. II, the preparatory transformation is the analog of the preparatory canonical transformation specified by Eqs. (24) and (34), which took us to the action-angle variables of the rapid degree of freedom.

The preparatory transformation causes  $\hat{D} = \hat{W}^\dagger \hat{D}' \hat{W}$  to acquire higher-order terms, just as the classical Hamiltonian in Eq. (36) acquired higher-order terms under the preparatory canonical transformation. These higher-order terms can either be worked out by direct substitution or by the Moyal formula. The calculation is standard, and is summarized in terms of symbol matrices by

$$\begin{aligned} D_{0\mu\nu} &= \left[ \frac{\mathbf{p}^2}{2} - E + v^{(\mu)} \right] \delta_{\mu\nu}, \\ D_{1\mu\nu} &= -\mathbf{p} \cdot \mathbf{A}_{\mu\nu}, \\ D_{2\mu\nu} &= \frac{1}{2} \sum_{\sigma} \mathbf{A}_{\mu\sigma} \cdot \mathbf{A}_{\sigma\nu}, \end{aligned} \quad (77)$$

where

$$\mathbf{A}_{\mu\nu} = i\tau_{\alpha}^{(\mu)*} \nabla \tau_{\alpha}^{(\nu)} = i\langle \mu | \nabla | \nu \rangle = \mathbf{A}_{\mu\nu}^*. \quad (78)$$

Here we have introduced Dirac bra-ket notation for operations involving the fast or spinor indices, so that  $\tau^{(\mu)}$  corresponds notationally to  $|\mu\rangle$ . The expansion of  $\hat{D}$  given in Eqs. (77) terminates at second order,  $D_0$  is diagonal, but  $D_1$  and  $D_2$  are not.

Now we apply Eqs. (73) to diagonalize  $\hat{D}$  to higher order. At order zero we have simply  $\Lambda_0 = D_0$ , so the diagonal elements  $\lambda_0^{(\mu)}$  are simply

$$\lambda_0^{(\mu)} = \frac{\mathbf{p}^2}{2} - E + v^{(\mu)}(\mathbf{x}). \quad (79)$$

At order 1 we project out the diagonal part of the first-order entry in Eqs. (73). We obtain the diagonal elements of  $\Lambda_1$  in the form

$$\lambda_1^{(\mu)} = -\mathbf{p} \cdot \mathbf{A}^{(\mu)}, \quad (80)$$

where

$$\mathbf{A}^{(\mu)}(\mathbf{x}) = i\langle \mu | \nabla | \mu \rangle = \mathbf{A}_{\mu\mu} \quad (81)$$

is the usual vector potential for the geometric phase in Born-Oppenheimer problems [19,1]. From the off-diagonal elements of the first-order entry in Eq. (73), we find the off-diagonal elements of  $G_1$ ,

$$G_{1\mu\nu} = \frac{-i\mathbf{p} \cdot \mathbf{A}_{\mu\nu}}{\lambda_0^{(\mu)} - \lambda_0^{(\nu)}} = \frac{-i\mathbf{p} \cdot \mathbf{A}_{\mu\nu}}{v^{(\mu)} - v^{(\nu)}} \quad (\mu \neq \nu). \quad (82)$$

For simplicity, we set the diagonal elements of  $G_1$  to zero.

At second order, we first substitute lower-order results into the second-order entry in Eqs. (73), obtaining

$$\begin{aligned} \Lambda_2 &= D_2 - \frac{i}{2} ([G_1, D_1] + [G_1, \Lambda_1]) \\ &\quad + \frac{1}{2} (\{G_1, \Lambda_0\} - \Lambda_0 G_1) - i [G_2, \Lambda_0], \end{aligned} \quad (83)$$

just as we did in the calculations leading up to Eq. (49). Then we project out the diagonal part of this equation, noting that  $[G_1, \Lambda_1]$ ,  $[G_2, \Lambda_0]$ ,  $\{G_1, \Lambda_0\}$ , and  $\{\Lambda_0, G_1\}$  all vanish on the diagonal. Only the terms  $D_2$  and  $[G_1, D_1]$  have diagonal elements. The result gives the diagonal elements of  $\Lambda_2$ ,

$$\lambda_2^{(\mu)} = \frac{1}{2} \sum_{\nu} |\mathbf{A}_{\mu\nu}|^2 + \sum_{\nu (\neq \mu)} \frac{|\mathbf{p} \cdot \mathbf{A}_{\mu\nu}|^2}{v^{(\mu)} - v^{(\nu)}}. \quad (84)$$

We will not compute  $G_2$ , although it would be needed in order to write out the diagonalizing transformation to second order. In this respect, we do not carry the quantum calculation as far as the classical one.

At the order to which we have worked, it was never necessary to calculate any slow Poisson brackets, although these would be needed at next order. Therefore, to this order, the diagonalization of the operator matrix  $\hat{D}$  (i.e., after the preparatory transformation) is equivalent simply to diagonalizing the symbol matrix and then converting symbols back into operators. In this sense one might say that the second term in Eq. (84) is "obvious," since it is just what we expect for second-order corrections to the eigenvalues of an almost diagonal matrix (as in ordinary quantum perturbation theory).

Now we collect our results and write the diagonal elements of  $\Lambda$  in the form

$$\begin{aligned} \lambda^{(\mu)} &= \frac{1}{2} (\mathbf{p} - \epsilon \mathbf{A}^{(\mu)})^2 + v^{(\mu)}(\mathbf{x}) - E + \frac{\epsilon^2}{2} \sum_{\nu (\neq \mu)} |\mathbf{A}_{\mu\nu}|^2 \\ &\quad + \epsilon^2 \sum_{\nu (\neq \mu)} \frac{|\mathbf{p} \cdot \mathbf{A}_{\mu\nu}|^2}{v^{(\mu)} - v^{(\nu)}} + O(\epsilon^3). \end{aligned} \quad (85)$$

If we omit the last term in this expression, we obtain simply the diagonal elements of  $D$ , carried out to second order. These are the terms discussed by Berry [1] and

Aharonov and Stern [2]; Berry makes some interesting connections with the symmetric partner of his antisymmetric phase two-form when deriving these terms, and Aharonov and Stern derive them from a classical model. Neglecting the final term, Eq. (85) represents a particle moving in a fictitious magnetic field, represented by the vector potential  $\mathbf{A}^{(\mu)}$ , and a fictitious electric field, represented by the scalar potential in the second to the last term of Eq. (85) (there is also the usual Born-Oppenheimer potential  $v^{(\mu)}$ ). The fictitious potentials have geometrical interpretations, as discussed by Berry [1]. The final term of Eq. (85) is a contribution to the averaged Hamiltonian not considered by the authors just mentioned; it is of the same order as the geometric scalar potential in the ordering scheme we have used. Its actual importance and the relation of our ordering scheme to that of other authors will be discussed in Sec. IV.

Now we specialize Eq. (85) to the case of the spin Hamiltonian. First, the Born-Oppenheimer potential energy is given already by Eq. (76). Next, the geometric vector potential  $\mathbf{A}^{(\mu)}$  is related to the classical geometric vector potential  $\mathbf{A}$  introduced by Eq. (11) by

$$\mathbf{A}^{(\mu)} = \mu \mathbf{A} = \mu \nabla \mathbf{e}_1 \cdot \mathbf{e}_2. \quad (86)$$

This is a special case of the general relation between Hannay's angles and Berry's phase [20]; the geometrical significance of  $\mathbf{A}$  was discussed by Littlejohn [10,11,21], and the specific connection between  $\mathbf{A}$  and the geometric phase occurring in the spin problem was discussed by Berry [3] and Littlejohn and Flynn [6]. Next, the geometric scalar potential has been discussed by Berry [1], who has specialized it to the case of the spin Hamiltonian. Slightly generalized, Berry's result reads

$$V_g = \frac{1}{2} \sum_{\nu (\neq \mu)} |\mathbf{A}_{\mu\nu}|^2 = \frac{s^2 - \mu^2}{4} \left[ \frac{\partial b_i}{\partial x_j} \frac{\partial b_i}{\partial x_j} \right]. \quad (87)$$

We see that with the identification of  $\mu$  with  $M$ , the geometric vector and scalar potentials in the quantum Hamiltonian (85) are the same as in the classical Hamiltonian (54).

Finally, we specialize the final term of Eq. (85) to the case of the spin system. To compute the  $\mathbf{x}$  derivatives of  $|\nu\rangle$  which occur in  $\mathbf{A}_{\mu\nu}$ , we invoke the following identity, which is derived in Appendix A of [22]:

$$\frac{\partial}{\partial \mathbf{B}} |\nu\rangle = \frac{\mathbf{B} \times \hat{\mathbf{s}}}{B^2} |\nu\rangle - |\nu\rangle \langle \nu | \frac{\partial}{\partial B} |\nu\rangle. \quad (88)$$

Therefore, if  $\nu \neq \mu$ , we have

$$\begin{aligned} \mathbf{p} \cdot \mathbf{A}_{\mu\nu} &= i \mathbf{p} \cdot \langle \mu | \frac{\partial}{\partial \mathbf{x}} | \nu \rangle = i \mathbf{p} \cdot \nabla \mathbf{B} \cdot \langle \mu | \frac{\partial}{\partial \mathbf{B}} | \nu \rangle \\ &= i [(\mathbf{p} \cdot \nabla \mathbf{B}) \mathbf{b} + B \mathbf{p} \cdot \nabla \mathbf{b}] \cdot \langle \mu | \frac{\mathbf{B} \times \hat{\mathbf{s}}}{B^2} | \nu \rangle \\ &= i \mathbf{p} \cdot \nabla \mathbf{b} \cdot \langle \mu | \mathbf{b} \times \hat{\mathbf{s}} | \nu \rangle. \end{aligned} \quad (89)$$

The magnitude  $B$  of the magnetic field has dropped out of this expression, which therefore depends only on the direction  $\mathbf{b}$  of the magnetic field. Next, we write  $\hat{\mathbf{s}}_i = \hat{\mathbf{s}} \cdot \mathbf{e}_i$ ,

where  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 = \mathbf{b})$  is the same triad of unit vectors introduced in Sec. II, so that  $(\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \hat{\mathbf{s}}_3)$  satisfy commutation relations as in Eq. (64) and

$$\hat{\mathbf{s}}_3 |\mu\rangle = \mu |\mu\rangle. \quad (90)$$

Then, noting that

$$\mathbf{b} \times \hat{\mathbf{s}} = i [\hat{\mathbf{s}}_3, \hat{\mathbf{s}}], \quad (91)$$

we have

$$\mathbf{p} \cdot \mathbf{A}_{\mu\nu} = -(\mu - \nu) \mathbf{p} \cdot \nabla \mathbf{b} \cdot \langle \mu | \hat{\mathbf{s}} | \nu \rangle. \quad (92)$$

Using Eq. (92) once to cancel the denominator and Eq. (89) once, we can write the final term of Eq. (85) in the form

$$\begin{aligned} V_2 &= -\frac{i}{B} (\mathbf{p} \cdot \nabla \mathbf{b})_i (\mathbf{p} \cdot \nabla \mathbf{b})_j \\ &\quad \times \sum_{\nu (\neq \mu)} \langle \mu | \hat{\mathbf{s}}_i | \nu \rangle \langle \nu | (\mathbf{b} \times \hat{\mathbf{s}})_j | \mu \rangle \\ &= -\frac{i}{2B} (\mathbf{p} \cdot \nabla \mathbf{b})_i (\mathbf{p} \cdot \nabla \mathbf{b})_j \langle \mu | [\hat{\mathbf{s}}_i, (\mathbf{b} \times \hat{\mathbf{s}})_j] | \mu \rangle \\ &= -\frac{\mu}{2B} |\mathbf{p} \cdot \nabla \mathbf{b}|^2. \end{aligned} \quad (93)$$

In carrying out the second equality, we add the vanishing term  $\nu = \mu$  to the sum to create a resolution of identity, and we use the reality of  $V_2$  to write the result in terms of a commutator. We see that with  $\mu$  identified with  $M$ , the term  $V_2$  is identical in form to the final term in the classical Hamiltonian of Eq. (54).

#### IV. DISCUSSION

In the following discussion we explain the physical meaning of the dimensionless expansion parameter  $\epsilon$  in both the classical and semiclassical calculations, and we explain why our classical term  $V_2$ , shown in Eqs. (53) and (54), does not appear in the classical analysis of Aharonov and Stern [2]. In addition, we show that in the semiclassical treatment of the problem, one should actually distinguish two expansion parameters. This fact implies a refinement of the ordering of the terms which appear in the expansion of the symbol of the Hamiltonian operator. Finally, we consider why the classical and semiclassical calculations have given rise to results of precisely the same form.

We begin with the classical problem. The approximation introduced in Sec. II is fundamentally an adiabatic one. Stated in physical terms, the assumption is that the environment seen by the particle, i.e., the strength and direction of the magnetic field in which the particle finds itself, does not change significantly during a single precession period. To make this quantitative, we work with equations in which all physical constants are represented, such as the equation of motion (1), and we regard the spin vector  $\mathbf{s}$  as having dimensions of angular momentum. Then we write the precession frequency as  $\omega_p = \gamma B$ , and we introduce the scale length of the magnetic field,

$$L_B = \frac{B}{|\nabla \mathbf{B}|}, \quad (94)$$

where the scalar quantity  $|\nabla \mathbf{B}|$  is any reasonable norm of the tensor  $\nabla \mathbf{B}$ ; for example, it could be the maximum absolute value of the components. We will let  $L_B$  stand for the scale length of either the direction or the magnitude of the magnetic field; later we will consider what happens when these scale lengths are different.

In terms of these quantities, we define the adiabatic ordering parameter by

$$\epsilon_a = \frac{v}{L_B \omega_p} = \frac{v |\nabla \mathbf{B}|}{\gamma B^2}, \quad (95)$$

where  $v$  is the velocity of the particle, so that  $\epsilon_a$  is the fraction of the magnetic scale length  $L_B$  covered during a precession period and so that the motion is adiabatic if  $\epsilon_a \ll 1$ . Of course, the quantity  $\epsilon_a$  is really a function of space and time, or, more properly, of the phase-space coordinates of the particle; therefore, we will regard the motion as being adiabatic only in those regions of phase space where  $\epsilon_a$  is small.

Unlike  $\epsilon_a$ , the  $\epsilon$  introduced into the classical equations of motion (2) is a constant. It is a formal parameter introduced for mathematical convenience, but it can be given a physical interpretation as follows: Suppose that in some specific experimental situation the adiabatic condition  $\epsilon_a \ll 1$  is *not* satisfied in some region of phase space of interest. Then it is possible to change the experimental apparatus to create a new situation in which the adiabatic condition does hold. There are at least two ways to do this.

One is to expand the scale length of the magnetic field by a factor  $1/\epsilon$ . This can be done, for example, by appropriately scaling the size of the magnets and the currents in them. This change causes  $\mathbf{B}(\mathbf{x})$  to be replaced by  $\mathbf{B}(\epsilon \mathbf{x})$ , and  $L_B$  to be replaced by  $L_B/\epsilon$ . Then, as we see from Eq. (95), we can always satisfy the adiabatic condition  $\epsilon_a \ll 1$  by choosing  $\epsilon$  small enough. The equations of motion in the scaled magnetic field are

$$m_0 \frac{d^2 \mathbf{x}}{dt^2} = \gamma \frac{\partial}{\partial \mathbf{x}} [\mathbf{B}(\epsilon \mathbf{x}) \cdot \mathbf{s}], \quad (96)$$

$$\frac{d\mathbf{s}}{dt} = \gamma \mathbf{s} \times \mathbf{B}(\epsilon \mathbf{x}),$$

in which we can set  $\mathbf{x}' = \epsilon \mathbf{x}$ ,  $t' = \epsilon t$  to obtain

$$m_0 \frac{d^2 \mathbf{x}'}{dt'^2} = \gamma \frac{\partial}{\partial \mathbf{x}'} [\mathbf{B}(\mathbf{x}') \cdot \mathbf{s}], \quad (97)$$

$$\frac{d\mathbf{s}}{dt'} = \frac{1}{\epsilon} \gamma \mathbf{s} \times \mathbf{B}(\mathbf{x}').$$

Note that these equations have the  $\epsilon$  in exactly the same place as in Eqs. (2).

Another way to achieve adiabaticity is to keep the scale length of the magnetic field fixed, while scaling its magnitude, say, by increasing the current in the coils. We write  $\mathbf{B}(\mathbf{x}) \rightarrow \mathbf{B}(\mathbf{x})/\delta$  for this scaling, where  $\delta$  is another quantity we allow to become small. Then Eq. (95) seems to indicate that  $\epsilon_a$  will scale as  $\delta$ , but this is in-

correct, because the velocity  $v$  also scales with  $\delta$ . Nevertheless, it turns out that  $\epsilon_a$  does go to zero as  $\delta \rightarrow 0$ . To see how the velocity scaling comes about, we first write the equations of motion in the scaled magnetic field,

$$m_0 \frac{d^2 \mathbf{x}}{dt^2} = \frac{1}{\delta} \gamma \nabla [\mathbf{B}(\mathbf{x}) \cdot \mathbf{s}], \quad (98)$$

$$\frac{d\mathbf{s}}{dt} = \frac{1}{\delta} \gamma \mathbf{s} \times \mathbf{B}(\mathbf{x}),$$

and then we set  $t = \sqrt{\delta} t'$  to get

$$m_0 \frac{d^2 \mathbf{x}}{dt'^2} = \gamma \nabla [\mathbf{B}(\mathbf{x}) \cdot \mathbf{s}], \quad (99)$$

$$\frac{d\mathbf{s}}{dt'} = \frac{1}{\sqrt{\delta}} \gamma \mathbf{s} \times \mathbf{B}(\mathbf{x}),$$

against just as in Eqs. (2), with  $\epsilon$  identified with  $\sqrt{\delta}$ .

To understand the meaning of this square root, let us return to Eqs. (98) and expand out the force term to obtain

$$m_0 \frac{d^2 \mathbf{x}}{dt^2} = \frac{1}{\delta} \gamma [(\mathbf{b} \cdot \mathbf{s}) \nabla B + B \nabla \mathbf{b} \cdot \mathbf{s}]. \quad (100)$$

The first of the terms on the right-hand side involves the rate of change of the magnitude of the magnetic field, and the second involves the rate of change of its direction. For a typical magnetic field, these two terms will be of the same order of magnitude, because of the Maxwell equation

$$\nabla \cdot \mathbf{B} = B \nabla \cdot \mathbf{b} + \mathbf{b} \cdot \nabla B = 0. \quad (101)$$

One can create situations in which the scale lengths of the magnitude and direction of the magnetic field are quite different, but for now we will assume they are comparable as in the typical case. Even so, the effects of these two terms on the particle motion are quite different, because the first term involves the component of the spin along the magnetic field, which is predominantly an averaged quantity, whereas the second term involves the component of the spin perpendicular to the magnetic field, which is predominantly an oscillatory quantity. Therefore, although both terms are large in the scaling we have chosen, as indicated by the factor  $1/\delta$ , only the first term has a significant effect on a time scale much longer than a precession period.

In particular, let us consider a time interval required for the particle to cross a distance comparable to  $L_B$ . This time is of the order of  $L_B/v$ , where  $v$  is a typical velocity in the time interval in question. Furthermore, since the first term in Eq. (100) is appreciably constant over the distance  $L_B$ , the acceleration is also appreciably constant, and the particle, even if it starts with small velocity, will achieve a velocity  $v \sim \Delta v \sim 1/v\delta$  in this interval. Therefore,  $v \sim 1/\sqrt{\delta}$ , and  $\epsilon_a \sim \sqrt{\delta} = \epsilon$ , in agreement with our earlier scaling (of the scale length of the magnetic field).

Continuing with the scaling  $\mathbf{B} \rightarrow \mathbf{B}/\delta$ , we consider the relative magnitude of the kinetic momentum  $\mathbf{k} = m_0 \mathbf{v}$  and

the geometric vector potential  $M \mathbf{A}$ . This ratio is of interest, because it indicates the relative importance of the geometrical vector potential in the dynamics of the particle. This ratio is small in our treatment, as indicated by the  $\epsilon$  in the expression  $\mathbf{p} - \epsilon M \mathbf{A}$  appearing in the Hamiltonian (54), but it is understood to be of order unity in the treatment of Aharonov and Stern [2].

We write  $\epsilon_g$  for the ratio in question, so that

$$\epsilon_g = \frac{M |\mathbf{A}|}{m_0 v} \sim \frac{s}{L_B m_0 v}. \quad (102)$$

Here the parallel component of the spin  $M$  has been taken as of the same order as the magnitude  $s$  of the spin, and  $L_B$  represents the scale length of the direction of the magnetic field. If we were to perform the scaling  $\mathbf{B}(\mathbf{x}) \rightarrow \mathbf{B}(\mathbf{x})/\delta$ , but to ignore the scaling of the velocity  $v$ , then we would conclude that  $\epsilon_g$  is not small, but because of the scaling  $v \sim 1/\sqrt{\delta}$ , we see that  $\epsilon_g$  is actually of order  $\sqrt{\delta} = \epsilon$ .

Aharonov and Stern also treat the geometrical scalar potential as of the same order as the kinetic energy (in fact, all the terms in their Hamiltonian are treated as of the same order). In our treatment, however, the geometric scalar potential is even smaller than the geometric vector potential, as indicated by the  $\epsilon^2$  multiplying the second to the last term in Eq. (54). Thus, there are definite differences in the ordering assumptions made in the two treatments. These differences are responsible, as we will show, for the appearance of the extra term in our treatment [the last term in Eq. (54)], which does not appear in the treatment of Stern and Aharonov.

The disagreement seems to arise because Aharonov and Stern imagine a magnetic field in which the scale length of the magnitude of the field is much longer than that for the direction of the field. In the extreme case, the magnetic field varies in direction, but not in magnitude. As pointed out above, such magnetic fields are not typical, but they can be constructed (but not, apparently, in a current-free region). One can also attempt to create the equivalent of a magnetic field which is constant in magnitude within a Born-Oppenheimer model such as Eq. (74). In this case the effective magnetic field is given by an expansion of the  $2 \times 2$  potential-energy matrix  $V$  in terms of Pauli matrices. Such "magnetic fields" are not subject to Maxwell equations such as Eq. (101), so that it is easier to give different scale lengths to the direction and magnitude of the field. Nevertheless, Born-Oppenheimer models in which the effective magnetic field is constant in magnitude are rather artificial.

In any case, if we assume for the sake of argument that the scale length of the magnitude of the magnetic field is effectively infinite, then the argument given above, that the velocity  $v$  scales as  $1/\sqrt{\delta}$ , must be modified. This is because the right-hand side of Eq. (100) is now purely oscillatory, and the velocity does not accumulate on time scales longer than a precession period. In the scaling  $\mathbf{B} \rightarrow \mathbf{B}/\delta$ , the precession period  $T_p$  scales as  $T_p \sim 1/\omega_p \sim \delta$ , and since the acceleration scales as  $\dot{v} \sim 1/\delta$ , we see that the velocity only accumulates to an

amount which is of order unity, i.e.,  $\delta^0$ . Thus, for this case we have  $v \sim \delta^0$ ,  $\epsilon_a \sim \delta$ , and  $\epsilon_g \sim \delta^0 = 1$ . Furthermore, we see that all the terms of the classical Hamiltonian in Eq. (54) except the last one are of order unity, while the last is of order  $\delta$  due to the  $B$  in the denominator. (The Born-Oppenheimer potential  $-MB$  is large, but we throw it away because it is constant.) Therefore, under these special assumptions regarding scale lengths, we recover the results of Aharonov and Stern, including a justification for the neglect of the new term not present in their analysis.

On the other hand, even with the scale length assumptions made by Aharonov and Stern, there is nothing to prevent us from assuming that the velocity is of order  $1/\sqrt{\delta}$  due to its initial conditions. Then because  $\Delta v$  is small, this scaling of  $v$  remains valid in the course of time. In this way we return to the scaling conclusions of our earlier analysis, where no assumptions were made about the scale length of the magnitude of the magnetic field, where all geometric terms in the Hamiltonian are small, and where the new term is of the same order as the geometric scalar potential.

Altogether, we can say that the analysis of Aharonov and Stern is a special case of ours, in which the initial conditions for the velocity are chosen to be small enough that the kinetic momentum and the geometric vector potential are of the same order of magnitude, and in which the magnetic field is assumed to have scale lengths which allow this condition to remain valid in the course of time.

Of course, in any specific physical situation, one can plug in numbers to see which terms are important, but here we wanted to rely primarily on scaling arguments, which are more general than substitution of special values. Nevertheless, in all realistic situations we can imagine involving charged particles in magnetic fields, the new term is not only important, it is actually much larger than the geometric scalar potential. Furthermore, in Born-Oppenheimer examples, realistic calculations seem to show that the new term is important in understanding physical effects such as the hyperfine structure in the spectrum of molecules [8].

The example of a magnetic field  $\mathbf{B}(\mathbf{x}) = -\mathbf{x}$  has been studied in great detail, corresponding to a particle with spin moving through a homogeneous sphere of (unphysical) monopoles [4]. When scaling the density of the magnetic monopoles by writing  $H = \mathbf{p}^2/2 + \delta^{-1} \mathbf{x} \cdot \mathbf{S}$ , the previously developed procedure seems well defined and applicable in a straightforward manner. If, however, the magnitude  $x = |\mathbf{x}|$  is considered as the expansion parameter, a different ordering of the terms in the effective Hamiltonian results. In Ref. [4] an argument is presented showing that in the adiabatic limit of large  $x$ , the additional term  $V_2$  is small compared to the geometric electric potential, the term next to the last one in Eq. (54).

Now we turn to the semiclassical treatment of the spin system. As in the classical treatment, we have the adiabatic parameter  $\epsilon_a$ , defined in Eq. (95), but now we have in addition another parameter  $\epsilon_s$  which controls the validity of the semiclassical approximation. This parameter is defined as the ratio of the deBroglie wavelength  $\lambda$  to the magnetic scale length, or

$$\epsilon_s = \frac{\lambda}{L_B} \sim \frac{\hbar}{m_0 v L_B} . \quad (103)$$

This parameter must be small in order for the semiclassical

treatment to provide reliable information about the system.

To see the relevance of the ordering parameters, let us take the Hamiltonian of Eq. (85), restore all physical constants, and write the result in the form

$$H = H_1 + H_2 + H_3 + H_4 + H_5 = p^2/2m_0 - \gamma \mu \hbar B - \frac{\hbar \mathbf{p} \cdot \mathbf{A}^{(\mu)}}{m_0} + \frac{\hbar}{m_0^2 \gamma B} \sum_{\nu (\neq \mu)} \frac{|\mathbf{p} \cdot \mathbf{A}_{\mu\nu}|^2}{\mu - \nu} + \frac{\hbar^2}{2m_0} \sum_{\nu} |\mathbf{A}_{\mu\nu}|^2 . \quad (104)$$

If we take term  $H_1 = p^2/2m_0$  as a reference to be of order unity, then we find that  $H_2$  (the Born-Oppenheimer potential energy) is of order  $\epsilon_s/\epsilon_a$ ;  $H_3$  (the first-order term involving the geometrical vector potential) is of order  $\epsilon_s$ ;  $H_4$  (the new term) is of order  $\epsilon_a \epsilon_s$ ; and  $H_5$  (the second-order term involving the geometrical vector potential, combined with the geometrical scalar potential) is of order  $\epsilon_s^2$ . In this estimate all vector potentials  $\mathbf{A}_{\mu\nu} = i \langle \mu | \nabla | \nu \rangle$  are assumed to be of the order of  $1/L_B$ .

In any given physical situation, the dimensionless parameters  $\epsilon_a$  and  $\epsilon_s$  can be estimated, and the relative magnitude of the terms can be compared. In all physically reasonable situations we can think of, the parameters  $\epsilon_s$  and  $\epsilon_a$  satisfy

$$\epsilon_s \ll \epsilon_a \ll 1 . \quad (105)$$

The parameter  $\epsilon_s$  is quite small because the scale length of reasonable magnetic fields is always much greater than the particle wavelength. As a result of this ordering, the terms in Eq. (104) are in decreasing order of magnitude, and, in particular, the new term is much larger than the geometrical scalar potential.

In our actual semiclassical treatment of the spin system in Sec. III, we introduced only one small parameter, not two. In effect, the scaling we used, namely,  $\mathbf{B} \rightarrow \mathbf{B}/\epsilon$ ,  $\hbar \rightarrow \epsilon \hbar$ , was equivalent to treating both  $\epsilon_a$  and  $\epsilon_s$  as of order  $\epsilon$ ,

$$\epsilon_a \sim \epsilon_s \sim \epsilon . \quad (106)$$

This was done for reasons of mathematical convenience, it being easier to expand in a single parameter.

Our final remarks concern the question of why the classical and semiclassical calculations gave the same answers through second order. Both calculations involved the same adiabatic assumption  $\epsilon_a \ll 1$ , but the semiclassical calculation involved the additional assumption  $\epsilon_s < 1$ . The two semiclassical assumptions were incorporated into a single scaling when we set  $B \rightarrow B/\epsilon$ ,  $\hbar \rightarrow \epsilon \hbar$  in Eq. (61). As for the classical equations of motion (1), they of course do not involve  $\hbar$ , if we regard  $\mathbf{s}$  simply as a classical angular momentum vector. But these equations could equally well be regarded as Heisenberg equations of motion or Ehrenfest relations coming from the quantum Hamiltonian (61). If we adopt this interpretation, then it is of interest to apply the semiclassical scaling to the apparently classical equations to see what will happen. The semiclassical scaling (in which the spin quantum number is held fixed and of order unity) is

equivalent to the scaling  $B \rightarrow B/\epsilon$ ,  $\mathbf{s} \rightarrow \epsilon \mathbf{s}$ , which, as is easy to see, converts Eqs. (1) into Eqs. (2). Thus, the classical and semiclassical scalings are formally equivalent. This of course does not mean that the results of the two calculations will agree at all orders, for the usual reasons in semiclassical mechanics (classical commutators vanish, Ehrenfest relations are not exact, etc.). Indeed, an examination of our perturbation theory in Sec. III will show that the semiclassical calculation will start to differ from the classical calculation at third order.

In summary, we have investigated perturbatively the adiabatic motion of a neutral spinning particle in an inhomogeneous magnetic field, carrying out the expansion through terms of second order. Our main results are the exposition of two systematic perturbation approaches to the problem, one classical and the other semiclassical. In addition, we have found an additional term in the resulting Hamiltonian, not present in the analysis of Aharonov and Stern [2].

#### ACKNOWLEDGMENTS

We are grateful for several fruitful discussions with Ady Stern and Yakir Aharonov. One of us (S.W.) gratefully acknowledges support from the Freiwillige Akademische Gesellschaft (Basel) and the Schweizer Nationalfonds. This work was supported by the U.S. Department of Energy under Contract Nos. DE-AC03-76SF00098 and W-7405-Eng-48.

#### APPENDIX: VECTOR IDENTITIES

The following vector identities are useful when working with the unit vectors  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  and the gauge vector potential  $\mathbf{A}$ . Some of the less obvious of these have been presented and proven previously in [10,11]. Dyadic notation is used in some of these identities.

$$\mathbf{a}\mathbf{a} + \mathbf{c}\mathbf{c} + \mathbf{b}\mathbf{b} = \mathbf{1} = (\text{the identity}) , \quad (A1)$$

$$\mathbf{a}\mathbf{c} - \mathbf{c}\mathbf{a} = \mathbf{b} \times , \quad (A2)$$

$$\nabla \mathbf{a} = -(\nabla \mathbf{b} \cdot \mathbf{a}) \mathbf{b} - \mathbf{A} \mathbf{c} , \quad (A3)$$

$$\nabla \mathbf{c} = -(\nabla \mathbf{b} \cdot \mathbf{c}) \mathbf{b} + \mathbf{A} \mathbf{a} , \quad (A4)$$

$$\nabla \times \mathbf{A} = (\nabla \mathbf{b} \cdot \mathbf{a}) \times (\nabla \mathbf{b} \cdot \mathbf{c}) , \quad (A5)$$

$$\nabla \times (\nabla \mathbf{b} \cdot \mathbf{a}) = + \mathbf{A} \times (\nabla \mathbf{b} \cdot \mathbf{c}) , \quad (A6)$$



$$\nabla \times (\nabla \mathbf{b} \cdot \mathbf{c}) = -\mathbf{A} \times (\nabla \mathbf{b} \cdot \mathbf{a}), \quad (\text{A7})$$

$$\frac{\partial A_i}{\partial x_j} - \frac{\partial A_j}{\partial x_i} = \epsilon_{mkl} b_m \frac{\partial b_k}{\partial x_i} \frac{\partial b_l}{\partial x_j}. \quad (\text{A8})$$

$$\nabla \times \mathbf{A} = \frac{1}{2} \mathbf{b} \left[ \left[ \frac{\partial b_i}{\partial x_j} \frac{\partial b_j}{\partial x_i} \right] - (\nabla \cdot \mathbf{b})^2 \right] + (\nabla \cdot \mathbf{b}) \nabla \mathbf{b} \cdot \mathbf{b} - \mathbf{b} \cdot \nabla \mathbf{b} \cdot \nabla \mathbf{b}, \quad (\text{A9})$$

$$(\nabla \times \mathbf{A}) \cdot \nabla \mathbf{b} = 0. \quad (\text{A10})$$

- 
- [1] *Geometric Phases in Physics*, edited by A. Shapere and F. Wilczek (World Scientific, Singapore, 1989).
- [2] Y. Aharonov and A. Stern, *Phys. Rev. Lett.* **69**, 3593 (1992).
- [3] M. V. Berry, in *Fundamental Aspects of Quantum Theory*, edited by V. Gorini and A. Frigerio (Plenum, New York, 1986).
- [4] M. V. Berry and J. M. Robbins, *Proc. R. Soc. London* (to be published).
- [5] R. G. Littlejohn and W. G. Flynn, *Phys. Rev. Lett.* **66**, 2839 (1991).
- [6] R. G. Littlejohn and W. G. Flynn, *Phys. Rev. A* **44**, 5239 (1991).
- [7] S. Weigert and R. G. Littlejohn, *Phys. Rev. A* **47**, 3506 (1993).
- [8] J. F. Babb and A. Dalgarno, *Phys. Rev. Lett.* **66**, 880 (1991).
- [9] B. Felsager and J. M. Leinaas, *Nucl. Phys. B* **166**, 162 (1980).
- [10] R. G. Littlejohn, *J. Plasma Phys.* **29**, 111 (1983).
- [11] R. G. Littlejohn, *Contemp. Math.* **28**, 151 (1984).
- [12] J. R. Cary, *Phys. Rep.* **79**, 131 (1981).
- [13] J. R. Cary and R. G. Littlejohn, *Ann. Phys. (N.Y.)* **151**, 1 (1982).
- [14] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1980).
- [15] A. J. Dragt and J. M. Finn, *J. Math. Phys.* **17**, 2215 (1976).
- [16] M. S. Child, *Semiclassical Mechanics with Molecular Applications* (Clarendon, Oxford, 1991).
- [17] E. E. Nikitin and S. Ya. Umanskii, *Theory of Slow Atomic Collisions* (Springer-Verlag, New York, 1984).
- [18] R. G. Littlejohn and W. G. Flynn, *Chaos* **2**, 149 (1992).
- [19] C. A. Mead and D. G. Truhlar, *J. Chem. Phys.* **70**, 2284 (1979).
- [20] M. V. Berry, *J. Phys. A* **18**, 15 (1985).
- [21] R. G. Littlejohn, *Phys. Rev. A* **38**, 6034 (1988).
- [22] R. G. Littlejohn and W. G. Flynn, *Phys. Rev. A* **45**, 7697 (1992).