

Chapter 5

Hamiltonian Mechanics and Single Particle Motion

Chapter 4 discusses the process of two parallel plasma loops merging together. Considering this process from the perspective of individual particle orbits, the particles must transition from being localized to one loop to orbiting symmetrically among the two loops. An analogous transition happens on the Spheromak Experiment as the individual spider legs merge into an axisymmetric plasma jet. To study the orbits and transitions, we have employed the Hamiltonian formalism and report here a general theorem of Hamiltonian mechanics that evolved out of studies of charged particle motion. The theorem states that the action integral of the fastest periodic coordinate serves as an effective Hamiltonian for the reduced or orbit-averaged system. Stated alternatively, the action integral encodes the average evolution of the system such that the average evolution of the remaining non-periodic coordinates can all be extracted from this single quantity. Being rooted in Hamiltonian mechanics, the theorem is quite general and, in essence, extends concepts from guiding center theory to a broad class of Hamiltonian systems. Moreover, the theorem provides a unified framework for obtaining the averaged evolution without having to average the individual equations of motion, which can lead to a substantial reduction in the number of computations.

This chapter is organized in the following fashion. Sec. 5.1 will state and prove the theorem formally in two stages. First, we prove the theorem in a restricted setting in which coordinates other than the periodic one are ignorable. Examples such as Kepler orbits and the relativistic $\mathbf{E} \times \mathbf{B}$ drift are provided; in the latter example, we demonstrate the relativistic coupling between orthogonal Cartesian directions and an unexpected non-uniform motion *parallel* to the magnetic field. Sec. 5.1.8 then generalizes the theorem to the adiabatic case, where the action integral is then identified as a Hamiltonian for the reduced system. The exposition of Sec. 5.1 parallels that of Ref. [91] but includes more details and examples. Sec. 5.2 explicitly shows the connection between the general theory and guiding center motion, rederiving the drift equations using the first adiabatic invariant μ as an effective Hamiltonian. These drifts include the grad-B drift, the magnetic mirror force, and the

$\mathbf{E} \times \mathbf{B}$ drift; finite Larmor radius effects are also computed for the latter. Sec. 5.2.5 then shows how the magnetic flux enclosed by a gyro-orbit can be computed from the action integral, and Sec. 5.2.6 presents a formula that ties together different calculations presented throughout the chapter. The application of the theorem to the merging of plasma loops is tackled in Chapter 6.

A comprehensive exposition of mechanics can, of course, be found in the classic texts [34, 92]. For charged particle motion through magnetic fields, Northrop provides a detailed exposition [35], but more intuitive texts are available [93, 2]. Finally, more mathematical treatments of Hamiltonian dynamics rooted in differential geometry are available [94, 95].

5.1 A Theorem on the Action Integral of Periodic Motion

The theory presented here hinges on the action integral of periodic motion, defined for a periodic coordinate Q as

$$J = \oint P_Q dQ, \quad (5.1)$$

where P_Q is the canonical momentum associated with Q , and the symbol \oint refers to integration over one full period [92, 94, 34]. Action integrals are adiabatic invariants [34, pg. 154] [94, pg. 297] and play an important role in perturbation theory [94, ch. 10] [92, ch. 12]. A brief review of action integrals in the context of action-angle variables and canonical transformations is given in Appendix C, but canonical transformations will not be needed in this chapter.

The crux of this section is that the action integral of the fastest periodic coordinate in a system acts as an effective Hamiltonian for the reduced or orbit-averaged system. Sec. 5.1.1 explains the meaning of “reduced” or “orbit-averaged” by analogy to a wall clock. In Sec. 5.1.2, a preliminary version of the theorem is stated and proved. Several examples such as Kepler motion and relativistic $\mathbf{E} \times \mathbf{B}$ drifts make the meaning of the theorem concrete and also demonstrate applications. Sec. 5.1.7 shows how the drifts of guiding center generalize to any Hamiltonian system with a periodic component. Sec. 5.1.8 then generalizes the preliminary theorem to the adiabatic case. This section concludes by applying the theorem to a mechanical system which exhibits the phenomenon of magnetic mirroring without any magnetic field.

5.1.1 A Clock Analogy

The context for this theorem will be Hamiltonian systems in which one coordinate undergoes periodic motion while the other coordinates are ignorable, but the essence of the theorem can be described by analogy to a wall clock. As the minute hand makes a full revolution and returns to its original position, say 12:00, the hour hand does not return to its original position but instead increments forward one hour or 30 degrees. One part of the system returns its starting point, but another part

does not. Let us now imagine a peculiar clock whose hour hand does not move at a constant pace; perhaps it moves faster in the beginning of the hour and slower down towards the end or perhaps even moves backwards at times. The only condition we place on the hour hand is that, when the minute hand returns to 12:00, the hour hand has moved its obligatory 30 degrees. Such an hour hand would not be very useful for estimating time in between hours even though it faithfully tells the correct time on the hour and in the long run moves at the correct *average* rate.

Analogously, many Hamiltonian systems have a periodic coordinate that acts like the minute hand and other coordinates that increment with every period. An example is shown in Fig. 5.1 where an electron undergoes planar motion in the magnetic field of a straight wire; the radial coordinate is periodic but the axial position increments with every gyration. Like the peculiar hour hand, the z position does not change at a steady pace but moves faster, slower, and even backwards depending on the phase of the motion. The crux of Sec. 5.1.2 is to compute the net displacements of the other coordinates by taking partial derivatives of the action integral of the periodic coordinate. Section 5.1.7 then devises an averaged system in which the evolution of these other coordinates is steady, uniform, and free of peculiar non-uniformities.

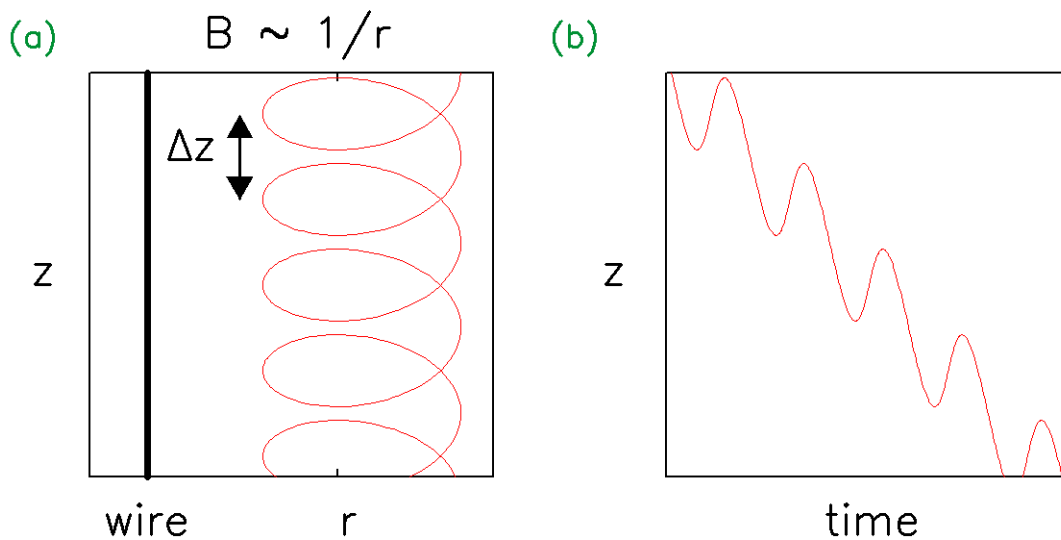


Figure 5.1: (a) An electron orbiting a current-carrying wire has periodic radial motion but displaces itself a net distance Δz with every gyration. (b) Plotting the axial position z as function of time shows the oscillatory motion superimposed over a net drift.

5.1.2 Net Displacements and Differentiating the Action Variable

We now state the theorem formally. Consider a two-dimensional time-independent Hamiltonian system where one coordinate, say η , is ignorable:

$$H = H(\xi, P_\xi, P_\eta). \quad (5.2)$$

Suppose that the ξ motion is periodic, that is, $\xi(t + \Delta t) = \xi(t)$ for some time interval Δt . The evolution of P_η is trivial:

$$\dot{P}_\eta = \frac{\partial H}{\partial P_\eta} = 0, \quad (5.3)$$

but the evolution of η is in general non-trivial. For instance, in a central force problem, the canonical angular momentum P_ϕ is constant, but $\dot{\phi} = P_\phi/mr^2$ varies with r . Returning to the general setting, over the time interval Δt , η undergoes a net displacement $\Delta\eta$: $\eta(t + \Delta t) = \eta(t) + \Delta\eta$. We claim that

$$\Delta\eta = -\frac{\partial J}{\partial P_\eta}, \quad (5.4)$$

where J is the action integral associated with ξ ,

$$J(H, P_\eta) = \oint P_\xi(H, \eta, P_\eta) d\xi. \quad (5.5)$$

In Eq. (5.5), $P_\xi(H, \eta, P_\eta)$ is obtained by solving the Hamiltonian in Eq. (5.2) for P_ξ . Eq. (5.4) states that if J is known then the net change of η during one period of ξ can be calculated without integrating the potentially complicated form of $\dot{\eta}$.

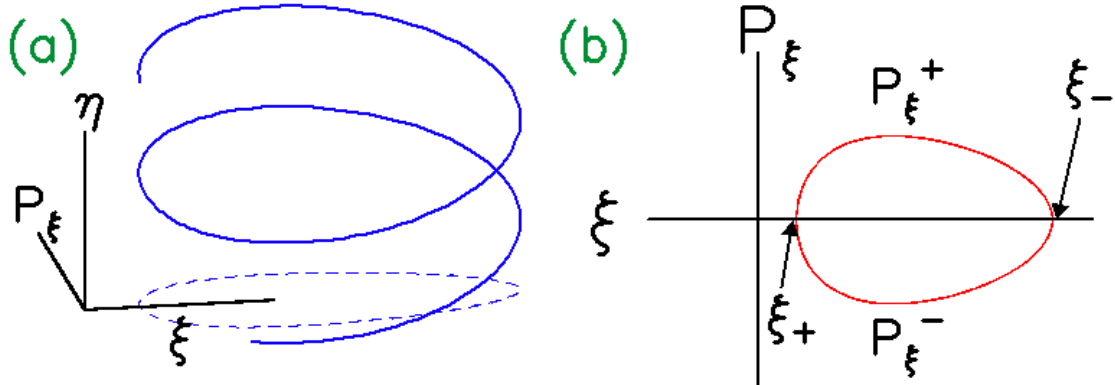


Figure 5.2: (a) The trajectory, in phase space, is helical because of the periodic motion in the ξP_ξ plane and the net displacement, or drift, in the η motion. (b) Projecting the trajectory onto the ξP_ξ plane produces a closed trajectory. The area of this curve is J . Also labeled are the turning points ξ_\pm and the two branches of P_ξ along a trajectory: the upper and lower branch denoted P_ξ^\pm .

The proof of Eq. (5.4) is surprisingly neat. First, we note that, in Eq. (5.4), there is no con-

tribution from differentiating the integral bounds because the ξ motion is periodic. To make this statement concrete, suppose that P_ξ is a double-valued function of ξ with value $P_{\xi+}$ along the upper branch and $P_{\xi-}$ along the lower branch as suggested in Fig. 5.2. Then

$$\oint P_\xi d\xi = \int_{\xi_1}^{\xi_2} P_{\xi+} d\xi + \int_{\xi_2}^{\xi_1} P_{\xi-} d\xi, \quad (5.6)$$

where ξ_1 and ξ_2 are the turning points of motion. By continuity, $P_{\xi+}(\xi = \xi_1) = P_{\xi-}(\xi = \xi_1)$ and similarly at $\xi = \xi_2$. Differentiating Eq. (5.6), one must differentiate the integral bounds in addition to the integrand:

$$\frac{\partial J}{\partial P_\eta} = \left[\int_{\xi_1}^{\xi_2} \frac{\partial P_{\xi+}}{\partial P_\eta} d\xi + \int_{\xi_2}^{\xi_1} \frac{\partial P_{\xi-}}{\partial P_\eta} d\xi \right] + \quad (5.7)$$

$$\frac{\partial \xi_1}{\partial P_\eta} [P_{\xi-}(\xi_1) - P_{\xi+}(\xi_1)] + \frac{\partial \xi_2}{\partial P_\eta} [P_{\xi+}(\xi_2) - P_{\xi-}(\xi_2)]. \quad (5.8)$$

However, the boundary terms then cancel by continuity of P_ξ at the turning points¹. Hence,

$$\frac{\partial J}{\partial P_\eta} = \oint \frac{\partial P_\xi(H, \xi, P_\xi)}{\partial P_\eta} d\xi. \quad (5.9)$$

Second, the P_ξ that appears in Eq. (5.9) is obtained by solving the Hamiltonian $H = H(\xi, P_\xi, P_\eta)$ for P_ξ along this particular trajectory. Furthermore, we can relate the $\partial P_\xi / \partial P_\eta$ to partial derivatives of H as follows. The differential of $H = H(\xi, P_\xi, P_\eta)$ is

$$dH = \frac{\partial H}{\partial \xi} d\xi + \frac{\partial H}{\partial P_\xi} dP_\xi + \frac{\partial H}{\partial P_\eta} dP_\eta. \quad (5.10)$$

In Eq. (5.9), we hold ξ and H fixed, so we set $d\xi = dH = 0$ in Eq. (5.10) and obtain

$$\frac{\partial P_\xi}{\partial P_\eta} = - \frac{\partial H / \partial P_\eta}{\partial H / \partial P_\xi}. \quad (5.11)$$

Using Eq. (5.11) and Hamilton's equations in Eq. (5.9) gives

$$\frac{\partial J}{\partial P_\eta} = - \oint \frac{\partial H / \partial P_\eta}{\partial H / \partial P_\xi} d\xi = - \oint \frac{d\eta / dt}{d\xi / dt} d\xi = -\Delta\eta. \quad (5.12)$$

This result is exact and no assumption of slowness is needed. It applies whether $\Delta\eta$ is small and the trajectory is nearly closed or whether $\Delta\eta$ is large and the trajectory is decidedly not closed. If there are extra ignorable coordinates in the system, Eq. (5.4), suitably adjusted, applies to each of them.

Eq. (5.4) generalizes the theorem [92, pg. 461] [34, pg. 156] that the period of motion is given

¹Differential geometry provides a more general framework for such proofs; see Ref. [94, pg. 197], for instance.

by a partial derivative of J with respect to H ,

$$\Delta t = \frac{\partial J}{\partial H}. \quad (5.13)$$

The resemblance between Eq. (5.4) and Eq. (5.13) is due to the fact that $(t, -H)$ can be considered a pair of canonical coordinates in extended phase space [96, pg. 15], so Eq. (5.13) can be considered a special case of the theorem presented.

5.1.3 Example: The Trivial Case

Our first example will be the trivial case in Cartesian coordinates in which the two dimensions are completely uncoupled:

$$H = \frac{P_x^2}{2m} + \frac{P_y^2}{2m} + V(x). \quad (5.14)$$

There are no forces in the y direction, so y evolves linearly, $\dot{y} = P_y/m$, and we anticipate $\Delta y = \dot{y}\Delta t = (P_y/m)\Delta t$. This is indeed obtained, for

$$J(H, P_y) = \oint \sqrt{2mH - P_y^2 - V(x)} \, dx, \quad (5.15)$$

so

$$-\frac{\partial J}{\partial P_y} = \oint \frac{P_y}{\sqrt{2mH - P_y^2 - V(x)}} \, dx. \quad (5.16)$$

However, $\Delta t = \partial J/\partial H$, and

$$\frac{\partial J}{\partial H} = \oint \frac{m}{\sqrt{2mH - P_y^2 - V(x)}} \, dx. \quad (5.17)$$

Therefore, $-\partial J/\partial P_y = (P_y/m)\Delta t$ as expected. This result is rather boring but illustrates that, for non-trivial examples to exist, there must be some coupling between the periodic and ignorable coordinate. In the following examples, we will see that this coupling can occur either through the expression for kinetic energy or through the introduction of a magnetic field.

5.1.4 Example: Kepler Motion

It is well-known that the bounded trajectories of Kepler orbits are ellipses. We therefore know, a priori, that $\Delta\phi = \pm 2\pi$, where ϕ is the polar angle in the plane of motion and the radial coordinate r is taken as the periodic coordinate. Kepler motion thus provides a non-trivial check of Eq. (5.4) and provides an example where the periodic and ignorable coordinates are coupled via the expression for kinetic energy.

The Hamiltonian for Kepler motion is

$$H = \frac{P_r^2}{2m} + \frac{P_\phi^2}{2mr^2} - \frac{mMG}{r}, \quad (5.18)$$

where m is the mass of the object in orbit, M is the central mass, and G is the gravitational constant. P_ϕ is the canonical angular momentum, $P_\phi = m\dot{\phi}r^2$, which is conserved along trajectories. The ϕ coordinate is ignorable, but the evolution of ϕ is coupled to the radial coordinate, since $\dot{\phi} = P_\phi/mr^2$. The action variable,

$$J = \oint \sqrt{2mH - P_\phi^2/r^2 + 2m\frac{mMG}{r}} dr, \quad (5.19)$$

can be efficiently evaluated by contour integration [92, pg. 468] [97, Appendix II] in the complex r -plane:

$$J = -2\pi |P_\phi| + 2\pi \frac{m^2MG}{\sqrt{2m|H|}}. \quad (5.20)$$

$\Delta\phi$ and Δt can now be found by applying Eq. (5.4) to Eq. (5.20),

$$\Delta\phi = -\frac{\partial J}{\partial P_\phi} = \pm 2\pi, \quad (5.21)$$

$$\Delta t = \frac{\partial J}{\partial H} = \pi \frac{m^2MG}{\sqrt{2m|H|}^3}. \quad (5.22)$$

The first equation tells us that ϕ always changes by 2π over a radial cycle no matter the energy or angular momentum; the \pm sign is determined by the sign of P_ϕ , i.e., the direction of rotation. This neatly proves that all Kepler orbits are closed and degenerate.² The second equation is the well-known formula for the period of a Kepler orbit.

Kepler orbits demonstrate how the periodic and ignorable coordinates can couple through the expression for kinetic energy. In Cartesian coordinates, kinetic energy takes a simple form:

$$T = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \quad (5.23)$$

for which there is no coupling. In a more general coordinate system, however, the kinetic energy must be expressed in terms of a position-dependent metric:

$$T = \frac{1}{2} mg_{ij}(q) \dot{q}^i \dot{q}^j. \quad (5.24)$$

²Degeneracy, in the setting of classical mechanics, is when the frequencies of two coordinates are not independent of each other [92, pg. 465].

For instance, in the polar coordinate system used for Kepler orbits,

$$T = \frac{1}{2}m \begin{pmatrix} \dot{r} & \dot{\phi} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \begin{pmatrix} \dot{r} \\ \dot{\phi} \end{pmatrix}. \quad (5.25)$$

If the Lagrangian is of the form $L = T - U = (1/2)g_{ij}(q)\dot{q}^i\dot{q}^j - U(q)$, then the canonical momenta are $P_i = mg_{ij}(q)\dot{q}^j$, and one can solve for the velocities \dot{q}^i as

$$\dot{q}^i = \frac{1}{m}g^{ij}(q)P_j, \quad (5.26)$$

where g^{ij} is the inverse matrix of the metric. It is the q dependence of g^{ij} that couples the periodic variable to the ignorable coordinate's velocity. In Kepler motion, $\dot{\phi} = P_\phi/mr^2$, so that the evolution of ϕ is coupled to the motion in the radial direction.

5.1.5 Example: Vector Potentials and Charged Particles

The coordinates may couple through the vector potential term of the Lagrangian for a charged particle in a magnetic field:

$$L = \frac{m}{2}v^2 + q\mathbf{v} \cdot \mathbf{A}. \quad (5.27)$$

In Cartesian coordinates, the expression for the canonical momenta is $P_i = mv_i + qA_i$, so

$$v_i = \frac{P_i - qA_i(\mathbf{r})}{m}. \quad (5.28)$$

The vector potential depends on the periodic variable and can thus couple the periodic motion to the ignorable coordinate's velocity. Particle motion in magnetic fields is not the only instance of such terms in the Lagrangian that are proportional to velocity. For instance, the transformation to a non-inertial frame introduces similar terms [34, sec. 39], and the Coriolis force $2m\mathbf{v} \times \boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is the angular velocity of the non-inertial frame, has a form suggestive of the magnetic Lorentz force $q\mathbf{v} \times \mathbf{B}$.

5.1.6 Example: Relativistic Mechanics and Relativistic $\mathbf{E} \times \mathbf{B}$ Drift

As a final example, coupling can occur in relativistic expression for kinetic energy. For a velocity-independent potential, the relativistic Lagrangian [92, Ch. 7.9] is

$$L = -mc^2\sqrt{1 - \beta^2} - V(\mathbf{x}) = -\frac{mc^2}{\gamma} - V(\mathbf{x}), \quad (5.29)$$

where $\beta = v/c$ and $\gamma = (1 - \beta^2)^{-1/2}$. The canonical momenta, in Cartesian coordinates, are

$$P_i = \frac{\partial L}{\partial v_i} = -mc^2 \frac{1}{2} \frac{-2v_i}{\sqrt{1 - v^2/c^2}} = \gamma mc^2 v_i. \quad (5.30)$$

Since γ depends on the magnitude of velocity, the relationship between v_i and P_i for any Cartesian direction also involves the velocities in the other directions. Suppose that z is ignorable, so P_z is conserved. The z velocity,

$$\dot{z} = \frac{P_z}{\gamma mc^2}, \quad (5.31)$$

is *not* necessarily constant because changes in v_x and v_y will change γ and hence v_z . Thus, the motion in different Cartesian directions can be coupled.

As an example of this coupling, we study the relativistic motion of a charged particle in crossed electric and magnetic fields. We find that the motion along the magnetic field is not constant but rather is modulated at the cyclotron frequency. Let $\mathbf{E} = E\hat{x}$ and $\mathbf{B} = B\hat{z}$ with $E/B < c$; Fig. 5.3 shows the configuration of the fields and a sample trajectory. Recall that in the non-relativistic case [2, Sec. 3.5.1] [93, Sec. 2.2.2], the particle drifts in the $-\hat{y}$ direction with a velocity $v_{E \times B}$ independent of its charge or mass,

$$\mathbf{v}_{E \times B} = \frac{\mathbf{E} \times \mathbf{B}}{B^2} = -\frac{E}{B} \hat{y}. \quad (5.32)$$

Furthermore, the velocity in the z direction is uniform and independent of the motion in the xy plane. Relativistic effects, however, couple the motion in the z direction to the motion in the xy plane so that v_z modulates at the cyclotron frequency, as shown in Fig. 5.4. This modulation contains a net drift that is computable with the Hamiltonian formalism. The full Lagrangian is

$$L = -mc^2 \sqrt{1 - v^2/c^2} - q\phi + q\mathbf{v} \cdot \mathbf{A} \quad (5.33)$$

$$= -mc^2 \sqrt{1 - v^2/c^2} + qEx + qv_y Bx, \quad (5.34)$$

where $\phi = -Ex$ and $\mathbf{A} = Bx\hat{y}$ are the appropriate potentials for the fields. The relativistic canonical momenta are the same as Eq. (5.30) with the addition of the vector potential term: $P_i = m\gamma v_i + qA_i$. The relativistic Hamiltonian can be found by performing the Legendre transform:

$$H = c\sqrt{(\mathbf{P} - q\mathbf{A})^2 + m^2c^2} + qV = c\sqrt{P_x^2 + P_z^2 + (P_y - qBx)^2 + m^2c^2} - qEx. \quad (5.35)$$

The x action,

$$J = \oint P_x dx = \oint \sqrt{\frac{(H + qEx)^2}{c^2} - (P_y - qBx)^2 - P_z^2 - m^2c^2} dx, \quad (5.36)$$

can be evaluated in closed form by expanding the quadratics and completing the square in x :

$$J = \oint \sqrt{\left(\frac{H^2}{c^2} - P_y^2 - P_z^2 - m^2 c^2\right) + 2q\left(\frac{EH}{c^2} + BP_y\right)x - q^2\left(B^2 - \frac{E^2}{c^2}\right)x^2} dx \quad (5.37)$$

$$= \oint \sqrt{c^2 - a^2(x-b)^2} dx, \quad (5.38)$$

where

$$a = q\sqrt{B^2 - E^2/c^2}, \quad (5.39)$$

$$b = \frac{1}{q} \frac{EH/c^2 + BP_y}{B^2 - E^2/c^2}, \quad (5.40)$$

$$c^2 = \frac{H^2}{c^2} - P_y^2 - P_z^2 - m^2 c^2 + \frac{(EH/c^2 + BP_y)^2}{B^2 - E^2/c^2}. \quad (5.41)$$

This integral represents the area of an ellipse of half-height c and half-width c/a . The area of the ellipse is $\pi c^2/a$, so the action is

$$J = \pi \left[\frac{H^2/c^2 - P_y^2 - P_z^2 - m^2 c^2}{q\sqrt{B^2 - E^2/c^2}} + \frac{(EH/c^2 + BP_y)^2}{q(B^2 - E^2/c^2)^{3/2}} \right]. \quad (5.42)$$

From this somewhat gruesome expression, the quantities Δt and Δz follow by partial differentiation:

$$\Delta t = \frac{\partial J}{\partial H} \quad (5.43)$$

$$= \pi \frac{(B^2 - E^2/c^2)(2H/c^2) + 2(EH/c^2 + P_y B)E/c^2}{q(B^2 - E^2/c^2)^{3/2}} \quad (5.44)$$

$$= 2\pi \frac{P_y B E/c^2 + B^2 H/c^2}{q(B^2 - E^2/c^2)^{3/2}} \quad (5.45)$$

$$\Delta z = -\frac{\partial J}{\partial P_z} = \frac{2\pi}{q\sqrt{B^2 - E^2/c^2}} P_z. \quad (5.46)$$

Together, these two quantities give the parallel drift:

$$\frac{\Delta z}{\Delta t} = \frac{B^2 - E^2/c^2}{P_y B E/c^2 + B^2 H/c^2} P_z. \quad (5.47)$$

Eq. (5.47) simplifies by writing P_y , P_z , and H in terms of initial conditions:

$$P_y = m\gamma_0 v_{y0} + qBx_0, \quad (5.48)$$

$$P_z = m\gamma_0 v_{z0}, \quad (5.49)$$

$$H = \gamma_0 m c^2 - qE x_0. \quad (5.50)$$

Substitution into Eq. (5.47) gives

$$\frac{\Delta z}{\Delta t} = \frac{B^2 - E^2/c^2}{(m\gamma_0 v_{y0} + qBx_0)BE/c^2 + B^2(\gamma_0 mc^2 - qEx_0)/c^2} P_z \quad (5.51)$$

$$= \frac{B^2 - E^2/c^2}{(m\gamma_0 v_{y0})BE/c^2 + B^2(\gamma_0 mc^2)/c^2} P_z \quad (5.52)$$

$$= \frac{B^2 - E^2/c^2}{v_{y0}BE/c^2 + B^2} v_{z0}, \quad (5.53)$$

a simplified form of Eq. (5.47). The x_0 dependence has canceled out as expected. Also, taking the non-relativistic limit by letting $c \rightarrow \infty$, we recover the non-relativistic result $\Delta z/\Delta t = v_{z0}$.

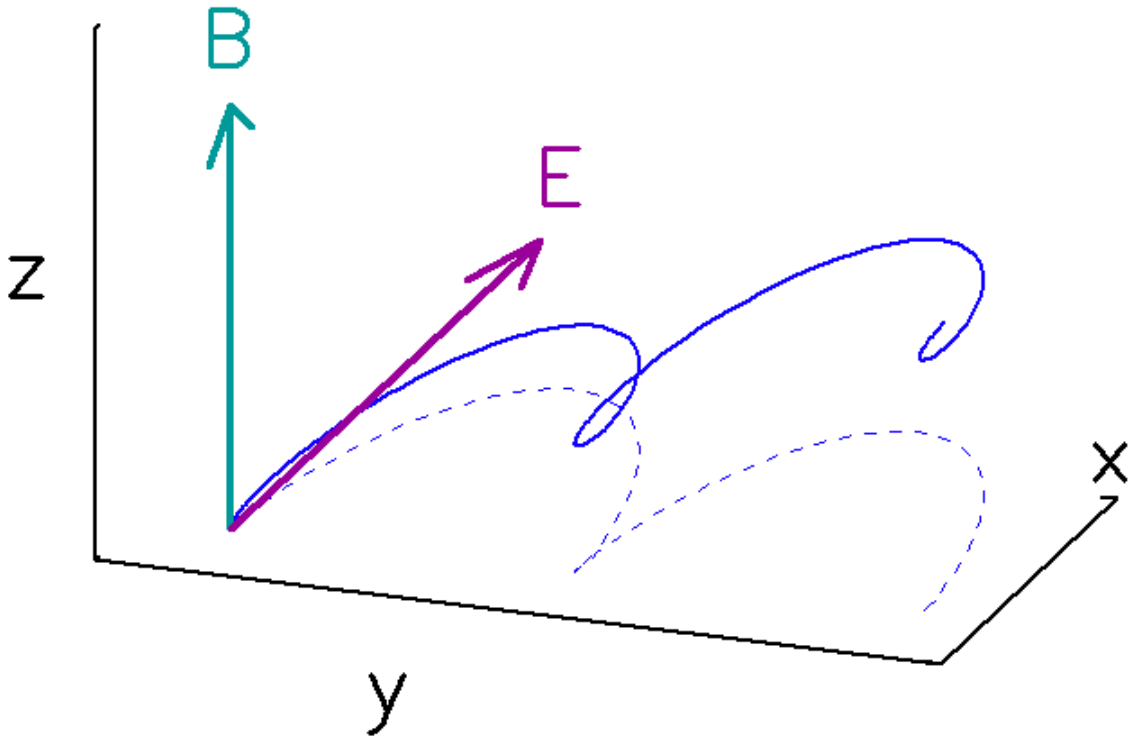


Figure 5.3: In a uniform electric and magnetic field, a charged particle will execute cycloid motion in the plane normal to \mathbf{B} and also move parallel to \mathbf{B} . Here, the solid blue line is the actual trajectory, and the dashed line is the projection of this trajectory, showing the cycloid motion. What is *not* apparent in this plot, though, is the fact that the velocity along \mathbf{B} is not uniform.

This unexpected modulation of the parallel velocity can be understood from Lorentz transformations. In the non-relativistic case, the $\mathbf{E} \times \mathbf{B}$ drift velocity is the velocity of a frame in which the electric field is transformed to zero, leaving only a magnetic field. In this frame, kinetic energy is

thus conserved, and the charged particle undergoes simple Larmor motion. The cycloid trajectory observed in the lab frame is a superposition of the uniform $\mathbf{E} \times \mathbf{B}$ drift and the circular Larmor motion. In the relativistic case, kinetic energy is likewise conserved in the $\mathbf{E} \times \mathbf{B}$ drift frame with no modulation of γ and hence no modulation of v_z . However, the transformation back to the lab frame involves a non-linear transformation of time, resulting in the modulation of the parallel velocity.

To see this effect, we first note that a Lorentz boost with velocity v_d along the y direction transforms the electric and magnetic fields as [98, Eq. (12.108)]

$$E'_x = \gamma_d(E_x + v_d B_z), \quad (5.54)$$

$$B'_z = \gamma_d\left(B_z + \frac{v_d}{c^2} E_x\right), \quad (5.55)$$

where $\gamma_d = (1 - (v_d/c)^2)^{-1/2}$. Since $E/B < c$, we can choose $v_d = -E/B$ so that the electric field vanishes in the boosted frame: $E'_x = 0$. The magnetic field becomes

$$B'_z = \gamma_d \left(B - \frac{E^2}{Bc^2} \right) = \frac{B - E^2/(Bc^2)}{\sqrt{1 - E^2/B^2c^2}} = \sqrt{B^2 - E^2/c^2}. \quad (5.56)$$

Without an electric field, the particle undergoes simple cyclotron motion, and the trajectory will look something like

$$\begin{pmatrix} ct' \\ x'(t') \\ y'(t') \\ z'(t') \end{pmatrix} = \begin{pmatrix} ct' \\ r_0 \cos \Omega t' \\ r_0 \sin \Omega t' \\ v'_{z0} t' \end{pmatrix}, \quad (5.57)$$

with $\Omega = qB'_z/m\gamma_d$ the cyclotron frequency in the drift frame [92, Eq. (7.154)]. We transform this solution back to the lab frame with the inverse Lorentz boost:

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma_d & 0 & \gamma_d \beta_d & 0 \\ 0 & 1 & 0 & 0 \\ \gamma_d \beta_d & 0 & \gamma_d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct' \\ r_0 \cos \Omega t' \\ r_0 \sin \Omega t' \\ v'_{z0} t' \end{pmatrix} \quad (5.58)$$

$$= \begin{pmatrix} \gamma_d(ct' + \beta_d r_0 \sin \Omega t') \\ r_0 \cos \Omega t' \\ \gamma_d(\beta_d t' + r_0 \sin \Omega t') \\ v'_{z0} t' \end{pmatrix} \quad (5.59)$$

The z velocity is not uniform because we have to express z as a function of t rather than t' , which

is achieved by inverting the zeroth component of the above transformation:

$$ct = \gamma_d (ct' + \beta_d r_0 \sin \Omega t'). \quad (5.60)$$

It is interesting to note that z evolves linearly with t' but not with t . We shall return to this curious fact in Sec. 5.1.7.

5.1.7 Drifts and Averaged Coordinates

The quantity $\Delta\eta$ is the net change in η over the course of a ξ cycle. In general, η may have oscillated greatly over this cycle, but, whatever the particulars of its motion, it ends up changing by $\Delta\eta$. This allows us to define an averaged η velocity, or drift, as

$$\frac{\Delta\eta}{\Delta t} = -\frac{\partial J/\partial P_\eta}{\partial J/\partial H}. \quad (5.61)$$

This velocity is constant and equal to $\lim_{t \rightarrow \infty} \eta(t)/t$. It captures the net rate of change of η without the details of the η evolution. Equation (5.61) generalizes the particle drifts associated with guiding center theory.

The velocity in Eq. (5.61) can be derived from Hamilton's equations by writing H as a function of J . That is, the functional form $J = J(H, P_\eta)$ can be inverted for H as a function of P_η and J : $H = H(J, P_\eta)$. Then, the differential of $J = J(H, P_\eta)$,

$$dJ = \frac{\partial J}{\partial H} dH + \frac{\partial J}{\partial P_\eta} dP_\eta, \quad (5.62)$$

implies that, by setting $dJ = 0$,

$$\frac{\partial H(J, P_\eta)}{\partial P_\eta} = -\frac{\partial J/\partial P_\eta}{\partial J/\partial H} = \frac{\Delta\eta}{\Delta t}. \quad (5.63)$$

Equation (5.63) has a form suggestive of Hamilton's equation: an η velocity was obtained by differentiation of the Hamiltonian with respect to P_η . However, the velocity obtained is the *average* one, where as Hamilton's equations applied to the original system $H = H(\xi, P_\xi, P_\eta)$ gives the *instantaneous* η velocity:

$$\frac{\partial H(\xi, P_\xi, P_\eta)}{\partial P_\eta} = \dot{\eta} \neq \frac{\Delta\eta}{\Delta t} = \frac{\partial H(J, P_\eta)}{\partial P_\eta}. \quad (5.64)$$

Thus, when H is written as a function of J in place of ξ and P_ξ , the velocity furnished by Hamilton's equations is the average or drift velocity. This result is discussed in Appendix C where it is seen that a canonical transformation from (ξ, P_ξ) to action-angle coordinates simultaneously transforms η to its average version.

Let us recall the peculiar hour hand of Sec. 5.1.1 that has a non-uniform speed even though it moves the requisite 30 degrees every hour. This hour hand is analogous to the ignorable coordinate η , whose evolution is likewise not necessarily uniform. We have seen that writing $H = H(J, P_\eta)$ instead of $H = H(\xi, P_\xi, P_\eta)$ transforms the ignorable coordinate into its averaged form; continuing the analogy, this is like transforming the peculiar hour hand into a well-behaved version that moves at a uniform rate and is much more reliable for reading time, as suggested by Fig. 5.5. Similarly, the relativistic example in Sec. 5.1.6 shows that boosting to the drift frame causes the z coordinate to evolve linearly in drift frame time whereas its evolution in lab frame time is non-uniform. In general, replacing the periodic coordinates ξ and P_ξ with the constant of motion J eliminates the oscillatory components of η .

5.1.8 Adiabatic Evolution and Reduced Systems

Eq. (5.4) is proved under the assumption that η is ignorable. Here, we relax the requirement that η is ignorable and instead allow the oscillations to evolve adiabatically. By adiabatic, we mean that the orbits with η dependence resemble the orbits of the ignorable case but that the orbit parameters slowly change over the course of many periods. This relaxation of assumptions not only allows a broader class of systems to be studied but also yields a new result: the action integral J serves as an effective Hamiltonian for a reduced system in which all the remaining non-periodic coordinates are replaced by their averaged versions.

We proceed by adding η dependence to the Hamiltonian, $H = H(\xi, P_\xi, \eta, P_\eta)$, but regard (η, P_η) as slowly varying parameters of the ξ oscillation. This assumption allows the action integral J to be defined³, in which case it will be an adiabatic invariant [34, pg. 154] [94, pg. 297]. To make these statements precise, we separate H into two parts:

$$H(\xi, P_\xi, \eta, P_\eta) = H_{\text{loc}}(\xi, P_\xi, \eta, P_\eta) + H_{\text{ext}}(\eta, P_\eta). \quad (5.65)$$

H_{loc} is a *local* Hamiltonian that describes the ξ oscillations; because H_{loc} contains all the ξ and P_ξ dependence, the ξ and P_ξ evolution depend only on H_{loc}

$$\frac{d\xi}{dt} = \frac{\partial H_{\text{loc}}}{\partial P_\xi} \quad \frac{dP_\xi}{dt} = -\frac{\partial H_{\text{loc}}}{\partial \xi}. \quad (5.66)$$

We presume that η and P_η play the role of slowly varying parameters in H_{loc} ; that is, over the course of a single ξ period, η and P_η can be held fixed to good approximation. The evolution of

³Arbitrary η dependence could result in large variations over a single ξ period to the point where ξ is no longer periodic. In such cases, the action integral, or any first integral of motion, will not exist, and such systems are termed non-integrable. The transition from an integrable system to a non-integrable system through introduction of a perturbation, such as the η dependence described here, is the topic of KAM theory [94, Appendix 8] [96, Sec. 1.4].

these “parametric coordinates” is determined by both the local and the *external* Hamiltonian H_{ext} :

$$\frac{d\eta}{dt} = \frac{\partial H}{\partial P_\eta} \quad \frac{dP_\eta}{dt} = -\frac{\partial H}{\partial \eta} \quad (5.67)$$

In essence, then, H_{ext} describes the way in which the ξ oscillations interact with the outside world by influencing the evolution of the oscillation parameters. This scheme, by construction, ensures the existence and adiabatic invariance of J , defined as in Eq. (6.20) but now with η dependence, as follows. Traditionally, the adiabatic invariance of the action integral is demonstrated by adding a time-dependent parameter, say $\lambda(t)$, to a Hamiltonian system, $H = H(\xi, P_\xi; \lambda(t))$ [34, pg. 154]. Here, we treat η and P_η as the slowly varying parameters of the ξ oscillation and thereby effectively reduce the present system $H = H(\xi, P_\xi, \eta, P_\eta)$ to the traditional case $H = H(\xi, P_\xi; \lambda(t))$. As in Ref. [34, pg. 154], we assume it is a good approximation to hold the parametric coordinates η and P_η fixed while evaluating the ξ action integral. However, an important distinction of our treatment is that Eq. (5.65) is time-independent and thus conservative, whereas the Hamiltonian $H = H(\xi, P_\xi; \lambda(t))$ is time-dependent and hence not conservative. Of course, the local and external systems exchange energy but the total energy, $E = E_{\text{loc}}(t) + E_{\text{ext}}(t)$, is conserved.

The added η dependence in J allows the derivation of ΔP_η , the net change in the previously conserved canonical momentum, in addition to $\Delta\eta$. Indeed,

$$\frac{\partial J}{\partial \eta} = \oint \frac{\partial P_\xi}{\partial \eta} d\xi = - \oint \frac{\partial H / \partial \eta}{\partial H / \partial J} d\xi = - \oint -\frac{\dot{P}_\eta}{\dot{\xi}} d\xi = \Delta P_\eta \quad (5.68)$$

Equation (5.68) together with Eq. (5.4) yields the following system of equations:

$$\frac{\partial J}{\partial \eta} = \Delta P_\eta, \quad \frac{\partial J}{\partial P_\eta} = -\Delta\eta. \quad (5.69)$$

Equations (5.69) have the makings of a Hamiltonian system with $-J$ serving as the Hamiltonian. They are *precisely* Hamiltonian as follows. We define discretized derivatives $d\eta/dt = \Delta\eta/\Delta t$ and $dP_\eta/dt = \Delta P_\eta/\Delta t$ that capture the net rates of change of η and P_η . Upon invocation of a rescaled time τ normalized by the ξ period:

$$d\tau = \frac{dt}{\Delta t}, \quad (5.70)$$

Eqs. (5.69) become

$$\frac{d\eta}{d\tau} = \frac{\partial}{\partial P_\eta} (-J), \quad (5.71)$$

$$\frac{dP_\eta}{d\tau} = -\frac{\partial}{\partial \eta} (-J). \quad (5.72)$$

This system of equations is Hamiltonian with $-J$ playing the role of the Hamiltonian and τ playing

the role of time. We note that τ is the angle variable conjugate to J ; it evolves linearly in time and increments by one for every cycle.

We show that $\Delta\eta$ and ΔP_η can both be split into two terms, local and external. This observation is not needed for the proof that J is a Hamiltonian for the reduced system but does offer some insight. We evaluate $J = \oint P_\xi d\xi$ by first solving for P_ξ in the Eq. (5.65); this means inverting the equation $H_{\text{loc}}(\xi, P_\xi, \eta, P_\eta) = H - H_{\text{ext}}(\eta, P_\eta)$. In this equation, H and H_{ext} appear in the combination $H - H_{\text{ext}}$, so P_ξ and J depend only on $H - H_{\text{ext}}$, which is just H_{loc} . In a more intuitive sense, J depends only on H_{loc} because H_{loc} is sufficient to prescribe the ξ dynamics. Up to now, we have written J as $J = J(H, \eta, P_\eta)$, but the above discussion suggests that J is rather a function of H_{loc} :

$$J = J(H, \eta, P_\eta) = \tilde{J}(H_{\text{loc}}(\eta, P_\eta), \eta, P_\eta) = \tilde{J}(H - H_{\text{ext}}(\eta, P_\eta), \eta, P_\eta). \quad (5.73)$$

The symbol $\tilde{J}(H_{\text{loc}}, \eta, P_\eta)$ is used simply to denote the difference in functional dependence from $J(H, \eta, P_\eta)$. This distinction may appear to be a matter of semantics, but we note that $\tilde{J}(H_{\text{loc}}, \eta, P_\eta)$ depends on η and P_η both implicitly through H_{ext} and also explicitly. Accordingly, $\Delta\eta$ and ΔP_η have two terms: one term comes from the explicit dependence and one from the implicit dependence. The former is the drift of the system, e.g., the displacements suffered by η and P_η as a result of the ξ motion completing a full cycle, while the latter is the slow change of η and P_η due to H_{ext} . In the example provided in Sec. 5.1.9, these two terms will be computed explicitly.

We obtain a Hamiltonian for the averaged system in regular time in which partial differentiation of H , written as a function of J , gives the discretized derivatives defined above. To see this, we use the differential of $J = J(H, \eta, P_\eta)$,

$$dJ = \frac{\partial J}{\partial H} dH + \frac{\partial J}{\partial \eta} d\eta + \frac{\partial J}{\partial P_\eta} dP_\eta, \quad (5.74)$$

to evaluate partial derivatives of $H(J, \eta, P_\eta)$:

$$\frac{\partial H}{\partial P_\eta} = -\frac{\partial J / \partial P_\eta}{\partial J / \partial H} = \frac{\Delta\eta}{\Delta t} = \frac{d\eta}{dt} \quad (5.75)$$

$$-\frac{\partial H}{\partial \eta} = -\frac{\partial J / \partial \eta}{\partial J / \partial H} = \frac{\Delta P_\eta}{\Delta t} = \frac{dP_\eta}{dt}, \quad (5.76)$$

giving the discretized derivatives. As in Sec. 5.1.7, writing H as a function of J rather than (ξ, P_ξ) turns the velocities given by Hamilton's equations into the drift velocities. We can further refine $H = H(J, \eta, P_\eta)$ into local and external parts by solving Eq. (5.73) for H_{loc} , i.e., $H_{\text{loc}} = H_{\text{loc}}(J, \eta, P_\eta)$. Since $H_{\text{loc}} = H - H_{\text{ext}}$,

$$H = H_{\text{loc}}(J, \eta, P_\eta) + H_{\text{ext}}(\eta, P_\eta). \quad (5.77)$$

The term $H_{\text{loc}}(J, \eta, P_\eta)$ is an adiabatic potential [99] and is, in essence, the residue of averaging the

periodic ξ motion. In fact, for systems approximating a harmonic oscillator, $J = 2\pi H_{\text{loc}}/\omega(\eta, P_\eta)$, so the adiabatic potential is $H_{\text{loc}} = J\omega/(2\pi)$, showing that J acts like an electrostatic charge and $\omega(\eta, P_\eta)$ acts like an electrostatic potential. The magnitude of this effective “ J charge” depends on the amplitude of the ξ oscillation. The use of $-J$ as a Hamiltonian with normalized time τ and the use of H , written as a function of J , with regular time are entirely equivalent⁴. Practically, though, there are techniques to evaluate J directly [97, Appendix II], so using $-J$ as the Hamiltonian spares one from inverting J for H , which might not be analytically feasible, as will be in the case in Sec. 6.1.

5.1.9 Example: Particle in a Groove

The following system was proposed by Prof. Bellan as a mechanical analog of magnetic mirroring [2, Sec. 3.5.6] [93, pg. 29], a phenomenon where charged particles are reflected from regions of strong magnetic field strength. It is well-known that μ -invariance is responsible for this mirroring phenomenon: as the field strength increases, the energy of the particle’s Larmor motion also increases as per Eq. (1.10) to the point when the Larmor energy equals the particle’s total kinetic energy. The particle then has no kinetic energy left to move along the field line and is thus reflected. We shall show that this principle is not restricted to charged particles in magnetic fields.

Consider a marble in a long groove where the steepness of the groove varies with position, as shown in Fig. 5.6. The height h of the groove can be expressed as a function of x and y , where y is the distance along the groove and x is the distance across the groove,

$$h(x, y) = \frac{1}{2}\kappa x^2 (1 + \alpha y^2) + \frac{1}{2}\lambda y^2. \quad (5.78)$$

The concavity and steepness of the groove is determined by the curvature λ in the y direction and κ in the x direction. α couples the steepness of the x direction to the y coordinate. The exact Lagrangian and Hamiltonian for a marble in such a groove is not trivial⁵, but, in the same approximation in which a pendulum can be considered a harmonic oscillator, the Hamiltonian for this system is

$$H = \frac{P_x^2}{2m} + \frac{P_y^2}{2m} + \frac{1}{2}mg\kappa x^2 (1 + \alpha y^2) + \frac{1}{2}mg\lambda y^2. \quad (5.80)$$

Presuming that the particle’s y position changes slowly relative to the oscillations across the groove

⁴In fact, in Hamiltonian mechanics, one can use any momentum as a Hamiltonian for the system so long as the associated coordinate is used as the “time,” or orbit parameter [96, pg. 15][100].

⁵The exact Lagrangian equals the Lagrangian for a particle in gravitational field with the constraint that the particle remain on the groove. Using the Lagrange multiplier Λ to enforce the constraint:

$$L = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz + \Lambda (z - h(x, y)), \quad (5.79)$$

for which the z equation of motion is $m\ddot{z} = -mg + \Lambda$. The approximation used here, like approximating a pendulum as a simple harmonic oscillator, ignores the acceleration in the z direction, so $\Lambda = mg$. Using this value of Λ in the exact Lagrangian gives the approximate Hamiltonian in Eq. (5.80).

(i.e., $|\alpha|$ and $|\lambda|$ are small compared to κ), the y -dependent frequency of x oscillation is

$$\omega(y) = \sqrt{g\kappa}\sqrt{1 + \alpha y^2}. \quad (5.81)$$

We identify $H_{\text{loc}} = P_x^2/2m + m\omega^2 x^2/2$, the energy of x oscillations. The x action is then that of a harmonic oscillator, $J = 2\pi H_{\text{loc}}/\omega(y)$, and Eq. (5.77) becomes

$$H(J, y, P_y) = \frac{P_y^2}{2m} + \frac{1}{2}mg\lambda y^2 + 2\pi\omega(y)J. \quad (5.82)$$

Equation (5.76) gives

$$\dot{P}_y = -\frac{\partial H(J, y, P_y)}{\partial y} = -mg\lambda y + 2\pi\frac{d\omega(y)}{dy}J. \quad (5.83)$$

The first term in Eq. (5.83) is the average force of the external Hamiltonian, in this case the pull of gravity along the groove. The second term, though, comes from differentiating the local Hamiltonian and is due to the adiabatic invariance of J . This average force in the y direction is due to the change in x -oscillation energy as the particle finds itself in a narrower or wider groove.

The adiabatic term in Eq. (5.83) can oppose the pull of gravity, leading to surprising behavior. Suppose that λ is negative so that the groove is concave down along in the y direction. Intuition dictates that a marble placed in such a groove will rattle its way down the groove with increasing velocity, as shown in Fig. 5.7. However, with the right choice of initial conditions, the adiabatic force can have a negative value, meaning it acts as a restoring force in the y direction and, if sufficiently strong, can overwhelm the contribution from λ and give oscillatory y motion! This is the mechanical analog of a magnetic mirror and has been verified by direct numerical integration as shown in Fig. 5.8.

For oscillatory y motion, Eq. (5.82) admits an action integral in the y direction. Denote this second action by K , which can be considered a new Hamiltonian for the x averaged system. This is an example of a two-tier hierarchy of action variables, or a wheel within a wheel [91]. For the reduced system, J plays the role of a conserved quantity, so we develop an analog of Eq. (5.4):

$$\frac{\partial K}{\partial J} = \oint \frac{\partial P_y(H, J, y)}{\partial J} dy = \oint \frac{1}{\partial J/\partial P_y} dy \quad (5.84)$$

$$= \oint \frac{-1}{dy/d\tau} dy = -\Delta\tau, \quad (5.85)$$

where Eq. (5.74) has been used to evaluate $\partial P_y/\partial J$ and Eq. (5.72) has been used to evaluate $\partial J/\partial P_y$. Since τ counts x cycles, $-\partial K/\partial J$ gives the number of x cycles per y cycle. If this quantity is rational, then the trajectory is closed, which is of interest when quantizing the system, as there is sometimes a one-to-one correspondence between periodic classical trajectories and quantum energy levels [101].

5.2 Charged Particle Motion in a Magnetic Field

Having discussed at length the general theory, we now focus on charged particle motion in magnetic fields. Sec. 5.2.1 shows that, in the guiding center approximation, the action integral J reduces to the first adiabatic invariant μ . Since J acts as a Hamiltonian for the reduced system, it follows that μ is a Hamiltonian for the guiding center drifts, and we shall explicitly recover several of the guiding center drifts from the general theory. Sec. 5.2.5 show an unexpected application of the theory: computing the magnetic flux enclosed by a particle's orbit by differentiating the action variable with respect to the particle's charge. We finally conclude the chapter with a theorem that relates these various partial derivatives of J .

5.2.1 The Connection Between J and μ

The action integral J is clearly important but often difficult to evaluate in closed form. In the same limit as the guiding center approximation, however, the integral for J can be approximated and is proportional to μ , the first adiabatic invariant [33, p. 16]. This approximation also shows that the guiding center limit is really a harmonic-oscillator approximation to the full problem.

Consider a magnetic field in the z direction and with x dependence: $\mathbf{B} = B_z(x)\hat{z}$. A charged particle will drift in the y direction due to the gradient in the magnetic field strength [2, Sec. 3.5.2] [93, Sec. 2.3] and will move freely in the z direction. The action J can be written in integral form using the vector potential

$$\mathbf{A} = A_y(x)\hat{y} = \left(\int_{x_0}^x B_z(x')dx' \right) \hat{y}, \quad (5.86)$$

and Hamiltonian

$$H = \frac{P_x^2}{2m} + \frac{(P_y - qA_y(x))^2}{2m} + \frac{P_z^2}{2m}, \quad (5.87)$$

from which

$$J(H, P_y, P_z) = \oint P_x dx = \oint \sqrt{2mH - P_z^2 - (P_y - qA_y(x))^2} dx. \quad (5.88)$$

To help with notation, let us define H_\perp as the kinetic energy in the Larmor motion (e.g. *not* in the z direction):

$$H_\perp = H - \frac{P_z^2}{2m}. \quad (5.89)$$

We evaluate J by first defining the guiding center position x_{gc} and then performing a Taylor expansion about this point. Let x_{gc} be the x location where the y velocity vanishes:

$$0 = m\dot{y}(x_{gc}) = P_y - qA_y(x_{gc}). \quad (5.90)$$

If the orbit were perfectly circular, as is assumed in guiding center theory, then x_{gc} would coincide with the center of the circle. x_{gc} is a function of P_y , and, for future reference, a partial differentiation of Eq. (5.90) with respect to P_y gives

$$\frac{\partial x_{\text{gc}}}{\partial P_y} = \frac{1}{qB_z(x_{\text{gc}})}. \quad (5.91)$$

We now perform a Taylor expansion of A_y around x_{gc} :

$$A_y(x) \approx A_y(x_{\text{gc}}) + \left. \frac{dA_y}{dx} \right|_{x_{\text{gc}}} (x - x_{\text{gc}}) \quad (5.92)$$

$$= A_y(x_{\text{gc}}) + B_z(x_{\text{gc}})(x - x_{\text{gc}}). \quad (5.93)$$

This expansion allows us to approximate J . Letting $x' = x - x_{\text{gc}}$,

$$J \approx \oint \sqrt{2mH_{\perp} - (qB_z(x_{\text{gc}})x')^2} dx' \quad (5.94)$$

$$= |qB_z(x_{\text{gc}})| \oint \sqrt{\frac{2mH_{\perp}}{q^2 B_z(x_{\text{gc}})^2} - (x')^2} dx' \quad (5.95)$$

The integral now represents the area enclosed by a circle of radius $\sqrt{2mH_{\perp}}/|qB|$, so

$$J \approx \pi \frac{2mH_{\perp}}{|qB_z(x_{\text{gc}})|} = 2\pi \frac{m}{|q|} \left(\frac{mv_L^2}{2|B_z(x_{\text{gc}})|} \right) = 2\pi \frac{m}{|q|} \mu, \quad (5.96)$$

where μ is defined as in Eq. (1.10). Thus, J is proportional to, but not equal to, μ in the guiding center limit. The importance of this derivation is the functional dependence of B on x_{gc} and hence on P_y ; this dependence is key in deriving the grad-B drift below. Note that the absolute value of q guarantees that the action J is positive⁶ for both positively and negatively charged particles.

In reducing J to μ , we have used the same approximations that reduce any potential well to a harmonic oscillator. To see this, note that the Hamiltonian in Eq. (5.87) can be viewed as a one-dimensional system with an effective potential energy $U_{\text{eff}} = (P_y - qA_y(x))^2/2m = (m/2)v_y(x)^2$ [2, pg. 119]. That is, the effective potential is equal to the kinetic energy in the y direction, and x_{gc} is clearly the location of the minimum of this potential well. As any potential well can be approximated as a harmonic oscillator by performing a second-order Taylor expansion of the potential around the minimum, Eq. (5.94) is equivalent to the action integral for a harmonic oscillator. The action for a harmonic oscillator is $J = 2\pi\omega H$, and, indeed, the Eq. (1.10) is proportional⁷ to $2\pi\omega H$ with ω equal to the gyrofrequency. The guiding center limit is thus a harmonic-oscillator approximation.

⁶ J must be positive for such systems because $J = \oint P_x dx$ and $P_x = m\dot{x}$. Therefore, the contour of integration for J is clockwise in the rP_r plane, corresponding to a positive value of J , regardless of the particle charge.

⁷ The constant of proportionality between between J and μ is $m/|q|$, which is suggestive of the gyromagnetic ratio which related the angular momentum of a charged rotating rigid body to its magnetic moment [98, pg. 252].

Section 5.1.8 shows that the action J can be used as a Hamiltonian for the reduced system. Having shown that J reduces to μ in the guiding center limit, we expect that μ will serve as a Hamiltonian for the guiding center motion. Section 5.2.2 shows that Eq. (5.94) for μ does indeed contain the grad-B drift of guiding center theory thanks to the definition of x_{gc} in Eq. (5.90). Sec. 5.2.3 extends the analysis to include parallel dynamics and the magnetic mirror force, and Sec. 5.2.4 adds an electrostatic field to study the $\mathbf{E} \times \mathbf{B}$ drift.

5.2.2 The Grad-B Drift

For the magnetic fields used in the previous section, a charged particle will drift in the y direction due to the gradient in field strength. We show that this grad-B drift, which is well-established in guiding center theory, can be computed using the Hamiltonian formalism. Δt is given by Eq. (5.13) and (5.96),

$$\frac{\partial J}{\partial H} = 2\pi \frac{m}{|q| B_z(x_{gc})} = \Delta t, \quad (5.97)$$

and is the usual gyroperiod. Δy is given by a partial derivative with respect to P_y :

$$\Delta y = -\frac{\partial J}{\partial P_y} = 2\pi m \frac{H_{\perp}}{|q| B_z(x_{gc})^2} \frac{\partial B_z}{\partial x} \frac{dx_{gc}}{dP_y} \quad (5.98)$$

$$= 2\pi m \frac{H_{\perp}}{|q| B_z(x_{gc})^2} \frac{\partial B_z}{\partial x} \left(\frac{1}{q B_z(x_{gc})} \right), \quad (5.99)$$

where Eq. (5.91) was used to evaluate dx_{gc}/dP_y . The drift velocity in the y direction, computed through the Hamiltonian formalism, is

$$\frac{\Delta y}{\Delta t} = \frac{H_{\perp}}{q B_z(x_{gc})^2} \frac{\partial B_z}{\partial x}. \quad (5.100)$$

Let us compare Eq. (5.100) to the grad-B drift [2, eq. 3.89] from the guiding center approximation:

$$\mathbf{v}_{\nabla B} = -\frac{H_{\perp}}{q B^3} \nabla B \times \mathbf{B} \quad (5.101)$$

$$= -\frac{H_{\perp}}{q B^3} \left(\frac{\partial B_z}{\partial x} \hat{x} \right) \times (B_z \hat{z}) \quad (5.102)$$

$$= \frac{H_{\perp}}{q B_z^2} \frac{\partial B_z}{\partial x} \hat{y}, \quad (5.103)$$

which is in perfect agreement with Eq. (5.100).

5.2.3 Parallel Dynamics: The Magnetic Mirror Force

We now consider a magnetic field whose strength slowly changes in the z direction as well as in the x direction. The parallel dynamics become non-trivial: the gyrofrequency changes with the field

strength, the parallel motion can slow down, and magnetic mirroring can occur. We show that μ serves as a Hamiltonian for both the parallel dynamics in addition to the grad-B drift discussed above.

We first make some assumptions on the magnetic field. Because $\partial_z B_z \neq 0$, there must be a second component of B to ensure $\nabla \cdot \mathbf{B} = 0$. We therefore allow B_x to be non-zero. Since the z dependence of B_z is mild, B_x must be small. All of this can be encapsulated by the vector potential

$$\mathbf{A} = A_y(x, z)\hat{y} = \tilde{A}(x, \epsilon z)\hat{y}; \quad (5.104)$$

that is, the vector potential is some function of x and ϵz . The ϵ in the z dependence allows us to keep track of the relative magnitudes of terms. For instance, the magnetic field components are

$$B_z = \frac{\partial A_y}{\partial x}, \quad B_x = -\frac{\partial A_y}{\partial z}, \quad (5.105)$$

so that B_z is zeroth order but B_x is small to first order. Likewise, $\partial_z B_z$ is first order, and $\partial_z B_x$ is second order. We shall only work to first order. In this limit, the magnitude of the magnetic field is

$$B = \sqrt{B_z^2 + B_x^2} \approx B_z, \quad (5.106)$$

and the unit vector of \mathbf{B} is

$$\hat{B} = \frac{B_z \hat{z} + B_x \hat{x}}{B} \approx \hat{z} + \frac{B_x}{B_z} \hat{x}. \quad (5.107)$$

To proceed with the Hamiltonian formalism, we modify our definition of x_{gc} to be a function of both P_y and z defined by

$$0 = \dot{y}(x_{\text{gc}}(P_y, z)) = P_y - qA_y(x_{\text{gc}}(P_y, z), z). \quad (5.108)$$

Implicit differentiation gives a modified version of Eq. (5.91):

$$\frac{\partial x_{\text{gc}}}{\partial P_y} = \frac{1}{qB_z}, \quad (5.109)$$

$$\frac{\partial x_{\text{gc}}}{\partial z} = -\frac{\partial A_y / \partial z}{\partial A_y / \partial x} = \frac{B_x}{B_z}. \quad (5.110)$$

The form of J remains the same,

$$J = 2\pi \frac{m}{|q|} \frac{H_{\perp}}{B_z(x_{\text{gc}}(P_y, z), z)}, \quad (5.111)$$

but there is now z dependence both in the magnetic field and also in the guiding center position. Neither $\partial J / \partial H$ nor $\partial J / \partial P_y$ change in form except for the z dependence added to B_z , so the grad-

B drift is the same as Eq. (5.100). The calculation of ΔP_z , however, becomes interesting, as the evaluation of $\partial J/\partial z$ must include differentiation of the z dependence of both B_z and also x_{gc} . Using Eqs. (5.109) and (5.110), we get

$$\frac{\partial J}{\partial z} = -2\pi \frac{m}{|q|} \frac{H_{\perp}}{B_z^2} \frac{dB_z(x_{\text{gc}}, z)}{dz} \quad (5.112)$$

$$= -2\pi \frac{m}{|q|} \frac{H_{\perp}}{B_z^2} \left[\frac{\partial B_z}{\partial x} \frac{\partial x_{\text{gc}}}{\partial z} + \frac{\partial B_z}{\partial z} \right] \quad (5.113)$$

$$= -2\pi \frac{m}{|q|} \frac{H_{\perp}}{B_z^2} \left[\frac{\partial B_z}{\partial x} \frac{B_x}{B_z} + \frac{\partial B_z}{\partial z} \right] \quad (5.114)$$

$$= -2\pi \frac{m}{|q|} \frac{H_{\perp}}{B_z^2} \left[\frac{B_x}{B_z} \frac{\partial}{\partial x} + \frac{\partial}{\partial z} \right] B_z \quad (5.115)$$

$$= -2\pi \frac{m}{|q|} \frac{\mu}{B_z} \hat{B} \cdot \nabla B_z. \quad (5.116)$$

Then

$$\frac{\Delta P_z}{\Delta t} = \frac{\partial J/\partial z}{\partial J/\partial H} = -\mu \hat{B} \cdot \nabla B_z. \quad (5.117)$$

We can work out the dynamics as predicted by the guiding center approximation and compare them with the conclusions of the previous paragraph. We start with the ∇B drift and show that it is unaltered by the added z dependence. While the gradient of the magnetic field strength is modified to first order ϵ :

$$\nabla B \approx \nabla B_z = \frac{\partial B_z}{\partial x} \hat{x} + \frac{\partial B_z}{\partial z} \hat{z}, \quad (5.118)$$

the grad-B drift remains the same as Eq. (5.103) to first order:

$$\mathbf{v}_{\nabla B} = -\frac{H_{\perp}}{qB^3} \nabla B \times \mathbf{B} \quad (5.119)$$

$$= -\frac{H_{\perp}}{qB^3} \left(\frac{\partial B_z}{\partial x} \hat{x} + \frac{\partial B_z}{\partial z} \hat{z} \right) \times (B_x \hat{x} + B_z \hat{z}) \quad (5.120)$$

$$= \frac{H_{\perp}}{qB^2} \frac{\partial B_z}{\partial x} \hat{y}. \quad (5.121)$$

The term $B_x \cdot \partial_z B_z$ is second order ϵ and thus is dropped. Guiding center theory predicts a second drift, the curvature drift [2, Sec. 3.5.2] [93, Sec. 2.3.2], but this is zero to first order:

$$\mathbf{v}_c = -\frac{mv_{\parallel}^2}{qB^2} \hat{B} \cdot \nabla \hat{B} \times \mathbf{B} \quad (5.122)$$

$$= -\frac{mv_{\parallel}^2}{qB_z^2} \left[\left(\frac{\partial}{\partial z} + \frac{B_x}{B_z} \frac{\partial}{\partial x} \right) \left(\hat{z} + \frac{B_x}{B_z} \hat{x} \right) \right] \times (B_z \hat{z} + B_x \hat{x}), \quad (5.123)$$

because the term in square brackets is second order ϵ . The parallel dynamics comes from μ conser-

vation:

$$\mu B = H - P_z^2/2m. \quad (5.124)$$

Taking a total time derivative of both sides,

$$\mu \mathbf{v}_d \cdot \nabla B = -P_z \dot{P}_z/m, \quad (5.125)$$

where \mathbf{v}_d is the drift velocity. Since B does not vary in the y -direction, $\mathbf{v}_d \cdot \nabla B = \dot{z} \partial_z B$, and dividing by Eq. (5.125) by $\dot{z} = P_z/m$ gives

$$\mu \hat{B} \cdot \nabla B = -\dot{P}_z, \quad (5.126)$$

which agrees with Eq. (5.117).

5.2.4 $\mathbf{E} \times \mathbf{B}$ Drifts

This section incorporates an electrostatic field $\mathbf{E}(x, y) = -\nabla V(x, y)$ in addition to a magnetic field of the form $\mathbf{B} = B_z(x) \hat{z}$. The Hamiltonian becomes

$$H = \frac{P_x^2}{2m} + \frac{(P_y - qA_y(x))^2}{2m} + qV(x, y), \quad (5.127)$$

and the action integral is

$$J(H, y, P_y) = \oint \sqrt{2m(H - qV(x, y)) - (P_y - qA_y(x))^2} dx. \quad (5.128)$$

We would like to approximate this integral and express the result in a form similar to the action J_0 for the purely magnetic case:

$$J_0(H, P_y) = \oint \sqrt{2mH - (P_y - qA_y(x))^2} dx \approx \pi \frac{2mH}{|q| B_z(x_{gc})}. \quad (5.129)$$

The coarsest approximation to Eq. (5.128) is sufficient to recover the $\mathbf{E} \times \mathbf{B}$ drift [2, ch. 3.5.1] [93, Sec. 2.2.2] in addition to the grad-B drift. More accurate approximations of Eq. (5.128) produce higher-order corrections to the $\mathbf{E} \times \mathbf{B}$ velocity. We relate one such higher-order correction to a finite Larmor radius effect. In Appendix D, we relate a second such higher-order correction with a pondermotive-like force induced by the cyclotron motion.

The coarsest approximation to Eq. (5.128) is to assume that, over the course of a gyro-orbit, $V(x, y)$ can be approximated by its value at the guiding center, $V(x, y) \approx V(x_{gc}, y)$. With this approximation,

$$J(H, y, P_y) \approx J_0(H - qV(x_{gc}, y), P_y), \quad (5.130)$$

The period Δt is the same as with no potential, but Δy has a new term that comes from the P_y dependence of x_{gc} in V :

$$\Delta y = -\frac{\partial J}{\partial P_y} = -\frac{\partial J_0}{\partial H} \frac{\partial}{\partial P_y} (H - qV) - \frac{\partial J_0}{\partial P_y} = \frac{\partial J_0}{\partial H} \frac{\partial_x V}{B_z} - \frac{\partial J_0}{\partial P_y}, \quad (5.131)$$

where Eq. (5.91) has been used. The drift in the y direction is thus

$$\frac{\Delta y}{\Delta t} = \frac{\partial_x V}{B_z} - \frac{\partial J_0 / \partial P_y}{\partial J_0 / \partial H}. \quad (5.132)$$

The second term is the grad-B drift discussed in Sec. 5.2.2, and we denote this quantity as $v_{\nabla B}$. The first term, $\partial_x V / B$, is readily identified as the y component of the $\mathbf{E} \times \mathbf{B}$ drift:

$$\left(\frac{\mathbf{E} \times \mathbf{B}}{B^2} \right)_y = \frac{(E_x \hat{x}) \times (B_z \hat{z})}{B_z^2} = \frac{\partial_x V}{B_z}. \quad (5.133)$$

The x component of the $\mathbf{E} \times \mathbf{B}$ drift is also contained in the Hamiltonian formalism and can be computed by taking Δx_{gc} and using the definition of x_{gc} from Eq. (5.90) to obtain $\Delta P_y = qB_z(x_{gc})\Delta x_{gc}$. Then, since

$$\Delta P_y = \frac{\partial J}{\partial y} = \frac{\partial J_0}{\partial H} \left(-q \frac{\partial V}{\partial y} \right), \quad (5.134)$$

the drift in the x direction is

$$\frac{\Delta x_{gc}}{\Delta t} = -\frac{\partial_y V}{B_z}. \quad (5.135)$$

We can retain the two next higher-order terms of the expansion of V to yield more accurate forms of the action integral. We proceed by using the expansion $V(x, y) \approx V(x_{gc}, y) + \partial_x V \delta x$ in the integrand of Eq. (5.128) and then completing the square in δx :

$$J \approx \oint \sqrt{2m(H - qV - \partial_x V \delta x) - (qB_z \delta x)^2} dx \quad (5.136)$$

$$= \oint \sqrt{2m \left(H - qV + \frac{m}{2} \left(\frac{\partial_x V}{B_z} \right)^2 \right) - q^2 B_z^2 \left(\delta x + \frac{m \partial_x V}{q B_z^2} \right)^2} dx. \quad (5.137)$$

This integral for J resembles Eq. (5.128) but with two modifications. First, the oscillation center is shifted by the amount $m \partial_x V / q B_z^2$. As discussed in Appendix D, this shift occurs because our definition of guiding center no longer minimizes the effective potential once a potential has been added. However, in evaluating Eq. (5.137), this shift will have no bearing on the value of the action integral⁸. The second modification is the addition of a new potential term $(m/2)(\partial_x V / B_z)^2$, which is the kinetic energy of the $\mathbf{E} \times \mathbf{B}$ drift. In Appendix D, this new potential is proposed to be a

⁸Put succinctly, J is the phase space area enclosed by the trajectory in the xP_x plane and thus is clearly not changed by an offset.

pondermotive-like potential induced by the cyclotron motion. Eq. (5.137) allows us to express J in terms of J_0 :

$$J(H, P_y) = J_0 \left(H - qV(x_{\text{gc}}, y, t) + \frac{m}{2} \left(\frac{\partial_x V}{B_z} \right)^2, P_y \right), \quad (5.138)$$

which serves as a Hamiltonian for the guiding center motion. The effect of the new potential term on the drift will be discussed shortly.

If the potential $V(x, y)$ is expanded one term further, $V(x, y) \approx V(x_{\text{gc}}, y) + \partial_x V \delta x + (1/2) \partial_{xx} V (\delta x)^2$, then the second-order term effectively modifies the magnetic field strength:

$$J \approx \oint \sqrt{2m \left(H - qV - \partial_x V \delta x - \frac{1}{2} q \partial_{xx} V \delta x^2 \right) - (qB_z \delta x)^2} dx \quad (5.139)$$

$$= \sqrt{2m (H - qV - \partial_x V \delta x) - (q^2 B_z^2 + m q \partial_{xx} V) (\delta x)^2} dx \quad (5.140)$$

$$= \sqrt{2m (H - qV - \partial_x V \delta x) - (qB_z^* \delta x)^2} dx, \quad (5.141)$$

where B_z^* is

$$B_z^* = \sqrt{B_z^2 + \frac{m}{q} \partial_{xx} V}. \quad (5.142)$$

One can then proceed as before and obtain Eq. (5.138) with B_z replaced with B_z^* .

$$J(H, P_y) = J_0 \left(H - qV(x_{\text{gc}}, y, t) + \frac{m}{2} \left(\frac{\partial_x V}{B_z^*} \right)^2, P_y \right), \quad (5.143)$$

This is the most accurate approximation of the action integral and has many instances of implicit P_y dependence through x_{gc} in V and B_z^* .

The added potential term in Eq. (5.138) contributes a new term to Δy , leading to a higher-order drift. The new term in Δy is obtained by differentiating the $(m/2)(\partial_x V/B_z)^2$ in Eq. (5.138) with respect to P_y . To organize the growing number of terms, we write $\Delta y = (\Delta y)_0 + (\Delta y)_1$, where $(\Delta y)_0$ consists of the terms from Eq. (5.131) and $(\Delta y)_1$ is the new higher-order term. Then

$$(\Delta y)_1 = -\frac{\partial J_0}{\partial H} \frac{\partial}{\partial P_y} \left(\frac{m}{2} \left(\frac{\partial_x V}{B_z} \right)^2 \right) \quad (5.144)$$

$$= -\frac{\partial J_0}{\partial H} \left(\frac{\partial_x V}{B_z} \right) \frac{\partial}{\partial P_y} \left(m \frac{\partial_x V}{B_z} \right) \quad (5.145)$$

$$= -\frac{\partial J_0}{\partial H} \left(\frac{\partial_x V}{B_z} \right) \frac{dx_{\text{gc}}}{dP_y} \frac{\partial}{\partial x} \left(m \frac{\partial_x V}{B_z} \right). \quad (5.146)$$

We then rewrite $\partial_x V/B_z$ as $v_{\text{E} \times \text{B}}$ and also use Eq. (5.91) to write $dx_{\text{gc}}/dP_y = 1/qB_z$. Then

$$(\Delta y)_1 = -\frac{\partial J_0}{\partial H} v_{\text{E} \times \text{B}} \frac{m}{qB_z} \frac{\partial}{\partial x} (v_{\text{E} \times \text{B}}). \quad (5.147)$$

This new term in Δy leads to new terms in the drift velocity $\Delta y/\Delta t$. Combining Eq. (5.131) with Eq. (5.147) gives

$$v_d = v_{\nabla B} + v_{\mathbf{E} \times \mathbf{B}} - v_{\mathbf{E} \times \mathbf{B}} \frac{m}{qB_z} \frac{\partial}{\partial x} (v_{\mathbf{E} \times \mathbf{B}}). \quad (5.148)$$

Since the higher-order drift involves a gradient of $v_{\mathbf{E} \times \mathbf{B}}$, one might regard it as a “grad-E” drift. However, in Appendix D.1 we show that the existence of this higher-order drift really depends on how one defines the guiding center.

Finally, we show that Eq. (5.143) can be used to derive a finite Larmor radius effect discussed by Chen in [93, sec. 2.4]. A finite Larmor radius effect is a higher-order correction to guiding center theory due to the fact that the particle’s orbit has non-zero size. Following Chen, we assume that the magnetic field is uniform and work in the guiding center limit so that J_0 is given by Eq. (5.96) with the modified field strength B^* . Then

$$J(H, P_y) = 2\pi m \frac{H - qV(x_{\text{gc}}) + (m/2)(\partial_x V/B_z)^2}{|q| B_z^*}. \quad (5.149)$$

We have already computed the contribution to Δy from differentiation of both qV and $(m/2)(\partial_x V/B_z)^2$. Here, we will focus on the contribution from differentiating B_z^* :

$$(\Delta y)_2 = -\frac{\partial J}{\partial P_y} = 2\pi m \frac{H - qV(x_{\text{gc}})}{|q| B_z^{*2}} \frac{\partial B_z^*}{\partial P_y}. \quad (5.150)$$

Using the definition of B_z^* from Eq. (5.142) and bearing in mind that B_z is constant, we have

$$\frac{\partial B_z^*}{\partial P_y} = \frac{1}{2B_z^{*2}} \frac{m}{q} \frac{\partial(\partial_{xx} V)}{\partial P_y} \quad (5.151)$$

$$= \frac{m}{2qB_z^*} \frac{\partial^3 V}{\partial x^3} \frac{\partial x_{\text{gc}}}{\partial P_y} \quad (5.152)$$

$$= \frac{m}{2q^2 B_z^* B_z} \frac{\partial^3 V}{\partial x^3}, \quad (5.153)$$

where Eq. (5.91) has been used to evaluate $\partial x_{\text{gc}}/\partial P_y$. At this point, we take $B_z^* \approx B_z$ and evaluate the resulting drift:

$$\frac{(\Delta y)_2}{\Delta t} = \frac{H - qV(x_{\text{gc}})}{|q| B^{*2}} \frac{m}{2qB_z} \frac{\partial^3 V}{\partial x^3}. \quad (5.154)$$

Since B_z is constant and $H - qV = (m/2)v_{\perp}^2$, we can write

$$\frac{\Delta y}{\Delta t} = \frac{m^2 v_{\perp}^2}{|q| B^{*2}} \frac{1}{4qB^*} \frac{\partial^2}{\partial x^2} \left(\frac{E_x}{B_z} \right) \quad (5.155)$$

$$= \frac{1}{4} r_L^2 \frac{\partial^2}{\partial x^2} \left(\frac{E_x}{B_z} \right), \quad (5.156)$$

which is equal to the finite Larmor radius correction to the $\mathbf{E} \times \mathbf{B}$ drift given in Ref. [93, sec. 2.4].

5.2.5 Magnetic Flux Enclosed by a Gyro-Orbit

Here, we show that the action integral J can be used to compute the magnetic flux enclosed by a gyro-orbit. As magnetic flux is only defined through a closed surface and charged particle trajectories are typically not closed, we evaluate the flux using the orbit in the drift frame. We give an alternate derivation of the same formula in Appendix C.

For any orbit in an electromagnetic field for which one coordinate, say ξ , is periodic, we have

$$\frac{\partial J}{\partial q} = \oint \frac{\partial P_\xi}{\partial q} d\xi = \oint \frac{\partial H(\mathbf{r}, \mathbf{P}; q)/\partial q}{\partial H(\mathbf{r}, \mathbf{P}; q)/\partial P_\xi} d\xi \quad (5.157)$$

$$= - \oint \left[-\frac{\mathbf{p} - q\mathbf{A}}{m} \cdot \mathbf{A} + V \right] \frac{d\xi}{\xi} \quad (5.158)$$

$$= - \int_{t_0}^{t_0+\Delta t} [-\mathbf{v} \cdot \mathbf{A} + V] dt \quad (5.159)$$

$$= \int \mathbf{A} \cdot d\mathbf{l} - \int_{t_0}^{t_0+\Delta t} V dt. \quad (5.160)$$

The line integral of \mathbf{A} is over one period of motion; since the motion is in general not periodic, we cannot identify this integral as the magnetic flux enclosed by the orbit. However, in the drift frame, the orbit is closed by definition, and the integral of \mathbf{A} over the orbit in the drift would indeed yield a well-defined flux. We extract this integral from Eq. (5.159) by writing the lab frame velocity as $\mathbf{v} = \mathbf{v}' + \mathbf{v}_d$, where \mathbf{v}_d is the drift velocity and \mathbf{v}' is the particle velocity as seen in the drift frame. We then have

$$\frac{\partial J}{\partial q} = - \int_{t_0}^{t_0+\Delta t} [-(\mathbf{v}' + \mathbf{v}_d) \cdot \mathbf{A} + V] dt \quad (5.161)$$

$$= \oint \mathbf{A} \cdot d\mathbf{l}' + \int_{t_0}^{t_0+\Delta t} \mathbf{v}_d \cdot \mathbf{A} dt - \int_{t_0}^{t_0+\Delta t} V dt. \quad (5.162)$$

The first term is the flux, which we denote by Φ . The second term⁹ can be rewritten using $q\mathbf{A} = \mathbf{P} - m\mathbf{v}$. As the drift is in a direction of symmetry, the component of \mathbf{P} along the drift is conserved. We then obtain

$$\frac{\partial J}{\partial q} = \Phi + \frac{1}{q} (\mathbf{P} - m\mathbf{v}_d) \cdot \Delta\mathbf{r} - \int_{t_0}^{t_0+\Delta t} V dt, \quad (5.163)$$

showing the relationship between $\partial J/\partial q$ and Φ . If there is a non-zero potential V , then its average over one period must be computed.

We mention in passing that $\oint \mathbf{A} \cdot d\mathbf{l}$ is related to the phase shift due to the Aharonov-Bohm effect [102], so the quantity $\partial J/\partial q$, when no potential is present, could be of particular value in

⁹We note that, if one had formally boosted to the drift frame in the non-relativistic limit, then the potential V would be transformed precisely by the amount $-\mathbf{v}_d \cdot \mathbf{A}$ while the vector potential would not be transformed. This can be seen by boosting the four-vector $(-V, c\mathbf{A})$ in the non-relativistic limit.

quantum systems.

5.2.6 Drift, Action, and Euler's Formula

We conclude this chapter with a formula that relates the various partial derivatives of J . The derivation here utilizes Euler's formula for homogenous functions. An alternate derivation that integrates the action over special phase space contours is included in Appendix D.5; the derivation there uses a technique developed by Montgomery in the case of rigid body dynamics [103].

Suppose that, for a charged particle in a magnetic field, the x coordinate is periodic and the y and z coordinates are ignorable. We write H , which is kinetic energy, as $(m/2)v^2$. Then J is a homogenous function of degree one in the variables v , P_y , P_z , and q :

$$J = \oint \sqrt{m^2 v^2 - (P_y - qA_y)^2 - (P_z - qA_z)^2} dx. \quad (5.164)$$

Applying Euler's theorem of homogenous functions to J , we obtain

$$J(v, P_y, P_z; q) = v \frac{\partial J}{\partial v} + q \frac{\partial J}{\partial q} + P_y \frac{\partial J}{\partial P_y} + P_z \frac{\partial J}{\partial P_z}, \quad (5.165)$$

or

$$J(H, P_y, q) = 2H\Delta t + q \frac{\partial J}{\partial q} + P_y \frac{\partial J}{\partial P_y} + P_z \frac{\partial J}{\partial P_z}. \quad (5.166)$$

The flux enclosed by a gyro-orbit is given by Eq. (D.26), which can be rewritten as

$$q\Phi = q \frac{\partial J}{\partial q} + P_y \frac{\partial J}{\partial P_y} + P_z \frac{\partial J}{\partial P_z} + mv_d^2 \frac{\partial J}{\partial H}. \quad (5.167)$$

Using Eq. (5.167) in Eq. (5.165),

$$J = q\Phi + 2 \left(H - \frac{1}{2}mv_d^2 \right) \Delta t. \quad (5.168)$$

This formula ties together the action integral, the flux, and the period, and the drift velocity. As discussed in Sec. 5.2.1, J becomes proportional to μ in the guiding center limit. In the same limit, the drift-velocity energy is negligible compared to the total energy H , so

$$2\pi \frac{m}{|q|} \mu = q\Phi + 2H\Delta t. \quad (5.169)$$

5.3 Conclusions

We have shown that, in any Hamiltonian system where one coordinate undergoes periodic evolution, the action integral for that coordinate encodes the evolution of the system once the periodic motion

has been averaged over. In fact, this action integral acts as a Hamiltonian for the reduced system and provides the averaged equations of motion provided that time is measured in periods of motion. This formalism can be used to rederive the drifts of the guiding center approximation using μ as a Hamiltonian and can even be used to derive the flux enclosed by a gyro-orbit, but the scope of the results goes far beyond charged particle motion in magnetic field.

In regards to future work, the curvature drift from guiding center theory has not yet been derived from the Hamiltonian formalism, and this could be a very interesting problem. As the curvature drift involves motion along curved field lines, one approach would be to rotate the coordinate system to align with the magnetic field after each period of motion. It is conceivable that the Hamiltonian formalism could be used to compute the required rotation matrix; this framework is reminiscent of an $SO(3)$ fiber bundle [104, ch. 1] over the ξP_ξ plane. Another application would be to situations where the guiding center approximation breaks down. Guiding center theory is essentially a harmonic-oscillator approximation, but the Hamiltonian theory depends on periodic motion that is not necessarily harmonic. Therefore, the latter is more robust and may supplant guiding center theory in cases of large orbit sizes or vanishing magnetic fields. Indeed, the beginnings of such work are seen in the next chapter.

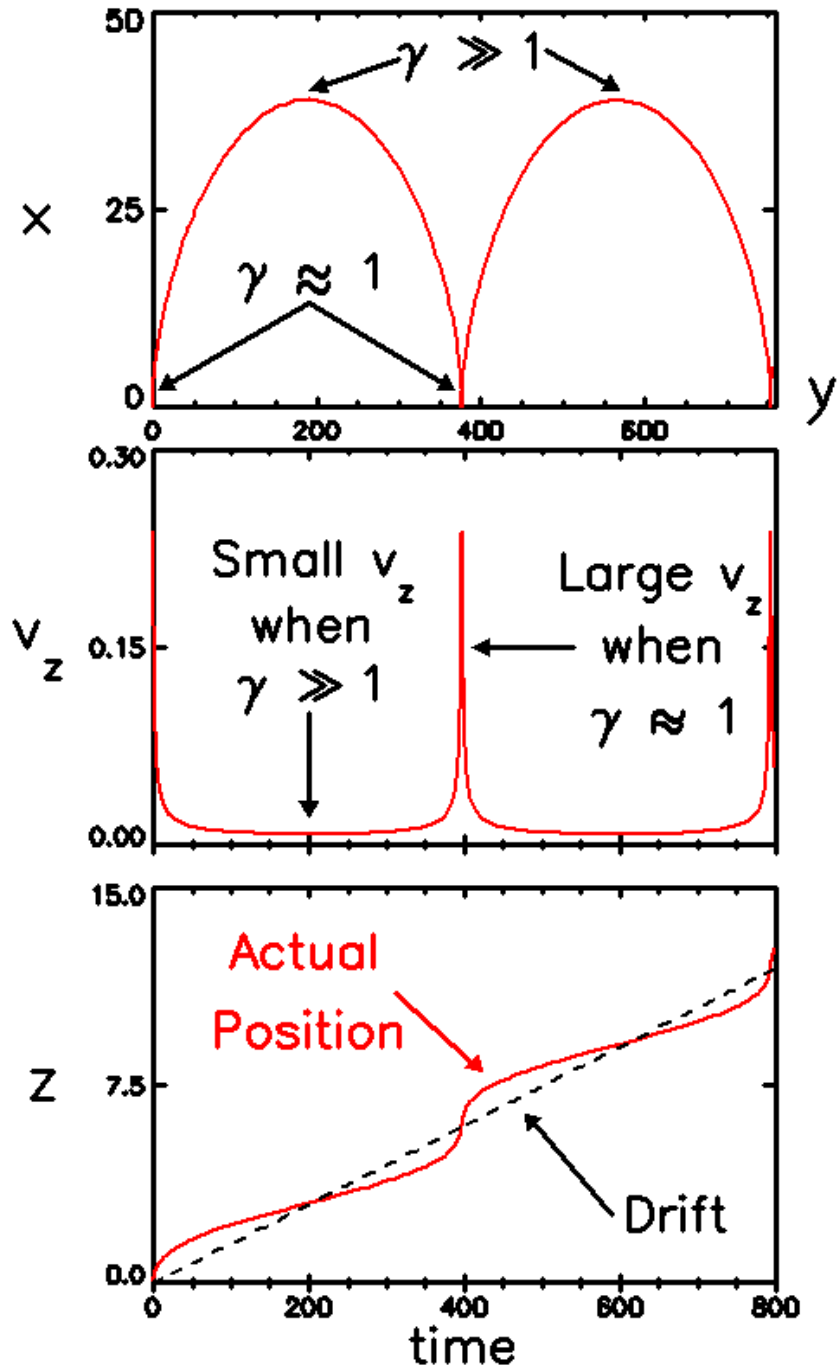


Figure 5.4: The top plot shows the cycloid motion in the xy plane. γ is large at the top of the cycloid, where the electric field has accelerated the particle to relativistic velocities. γ is approximately one at the bottom of the cycloid where the particle is nearly at rest. The middle plot shows the z velocity as a function of time. When γ is large, v_z is small in accordance with Eq. (5.28), but when $\gamma \approx 1$ v_z is large. The bottom plot shows the effects of this modulated velocity on z position, drawn with a solid red line. The dashed line is the theoretical drift value predicted by Eq. (5.47); the agreement is perfect.

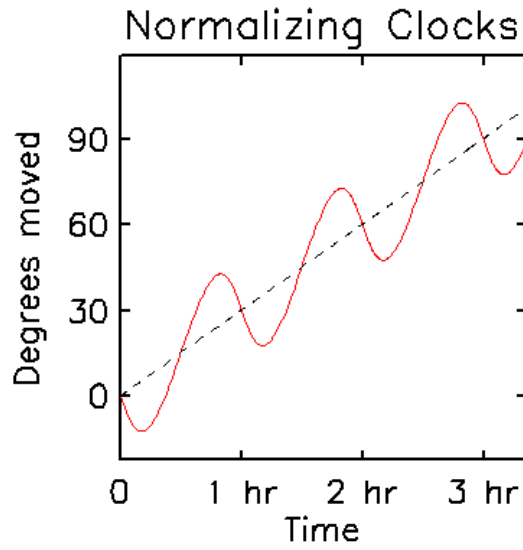


Figure 5.5: A peculiar hour hand with a non-uniform rate is plotted in red; it is analogous to the η motion. Although the hour hand moves at the correct average rate, the dashed black line, representative of the drift $\Delta\eta/\Delta t$, is much more convenient for reading time.

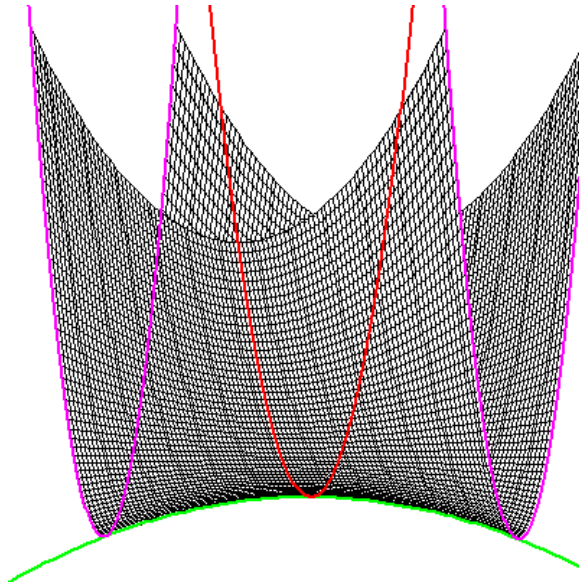


Figure 5.6: A side view of the groove. The green line runs along the spine of the groove which is in the y direction; you can clearly see that the groove is concave down in this direction. The red parabola is a cross-section of the groove at its apex where the groove is widest. The pink parabolas are cross-sections further down the groove and are steeper and more narrow than the red one.

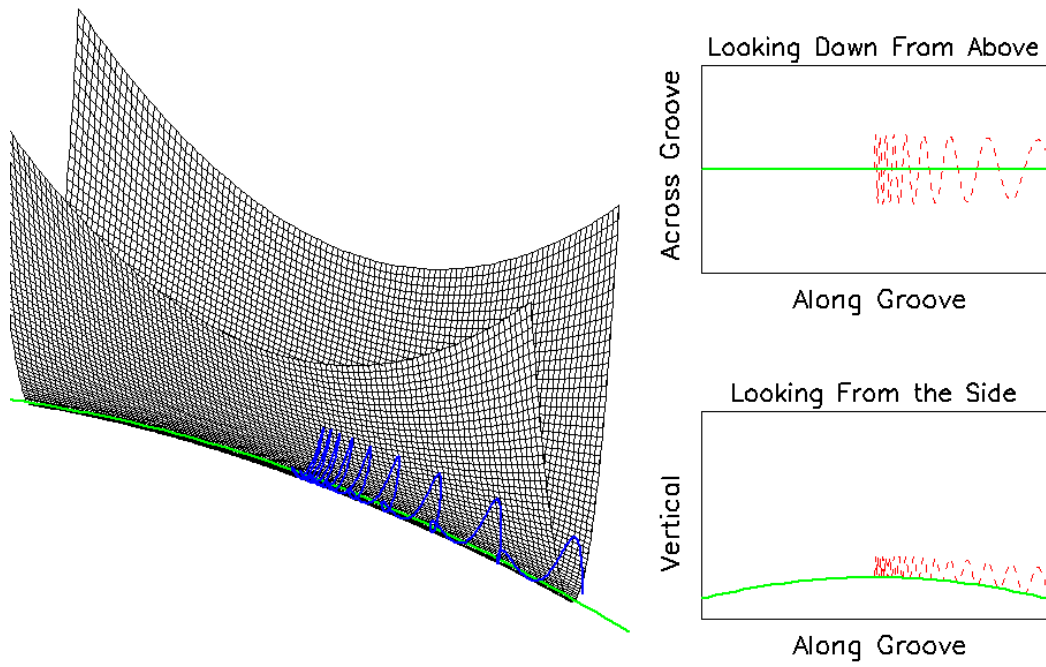


Figure 5.7: Intuition dictates that a particle placed in a saddle-like groove will eventually roll down the groove with increasing velocity. Indeed, this is observed in this numerical simulation. y the direction along the groove, and z the vertical direction. The Hamiltonian is given by Eq. (5.80) with $m = 1$, $\kappa = 1$, $\alpha = 1$ and $\lambda = -0.01$, and particle starts at $x = y = 0$ with $v_x(0) = 0.1$ and $v_y(0) = -0.004$.

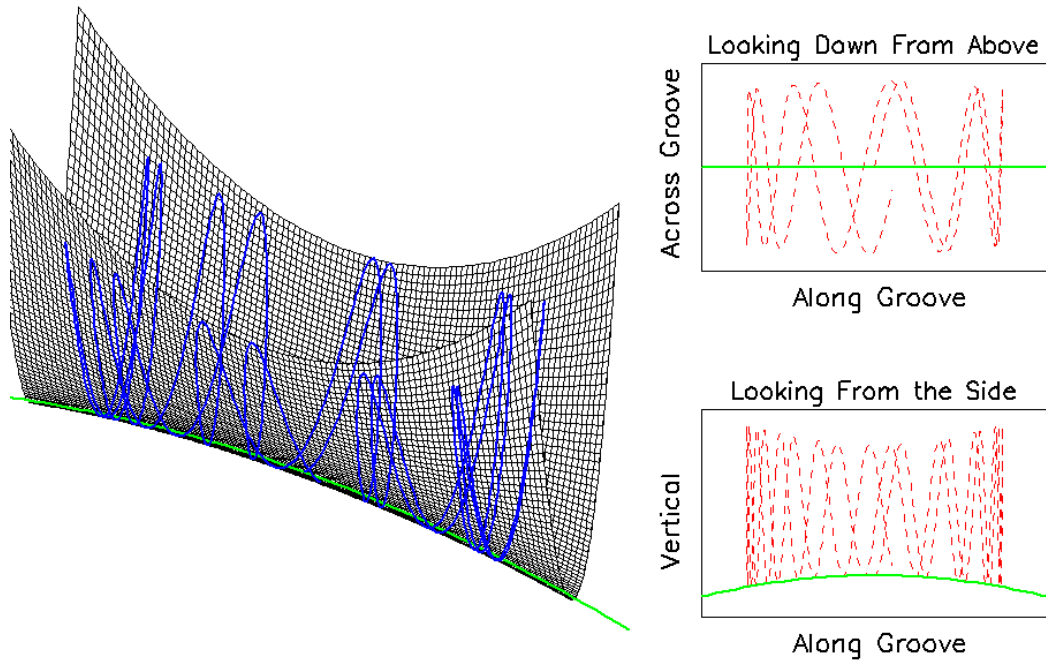


Figure 5.8: With the right initial conditions, a particle in a thin saddle-like groove can undergo oscillatory motion due to the narrowing of the groove. x is the direction across the groove, y the direction along the groove, and z the vertical direction. The Hamiltonian is given by Eq. (5.80) with $m = 1$, $\kappa = 1$, $\alpha = 1$ and $\lambda = -0.01$, and particle starts at $x = y = 0$ with $v_x(0) = 0.25$ and $v_y(0) = -0.1$.