

Energetic-particle orbits near rational flux surfaces in stellarators: I. Passing particles

Thomas E. Foster^{1,2}[†], Felix I. Parra^{1,2}, Roscoe B. White²,
José Luis Velasco³, Iván Calvo³, and Elizabeth J. Paul⁴

¹Department of Astrophysical Sciences, Princeton University,
Peyton Hall, Princeton, NJ 08544, USA

²Princeton Plasma Physics Laboratory, PO Box 451, Princeton, NJ 08543, USA

³Laboratorio Nacional de Fusión, CIEMAT, E-28040 Madrid, Spain

⁴Department of Applied Physics and Applied Mathematics,
Columbia University, New York, NY 10027, USA

(compiled on 4 December 2025)

Recent simulations (White *et al.*, 2022; White, 2022, 2025) have shown that, even when the magnetic field of a stellarator possesses nested toroidal flux surfaces, the orbits of passing energetic particles can exhibit islands. These ‘drift islands’ arise near rational flux surfaces, where they are likely to enhance alpha-particle transport—flattening the alpha density profile locally—unless they can be avoided by suitable design of the stellarator magnetic field. To investigate how this might be achieved, we derive an equation for the drift-island shape in a general stellarator. This result follows from the solution to a more fundamental problem: that of calculating the orbits of passing particles near a rational flux surface. We show that these orbits are determined by conservation of an adiabatic invariant associated with the closed rational-surface field lines. We use this ‘transit adiabatic invariant’ to prove that there are no drift islands, for all passing particles, if and only if the magnetic field satisfies a weaker version of the Cary & Shasharina (1997a) condition for omnigenicity; we call such magnetic fields ‘cyclometric’. The drift-island width scales as $\sim (\rho_\star \delta / s)^{1/2} a$ ($\rho_\star$ is the normalized gyroradius, δ is the deviation from cyclometry, s is the magnetic shear, and a is the minor radius), so large drift islands could arise in low-shear stellarators that are insufficiently cyclometric. To ensure accurate results for very energetic particles, we compute higher-order corrections to the transit invariant. Our calculations agree extremely well with ASCOT5 guiding-centre and full-orbit simulations of alpha particles in reactor-scale equilibria, even at 3.5 MeV. Finally, we show how our results can also be derived using Hamiltonian perturbation theory, which provides a systematic framework for calculating passing-particle orbits on both rational and irrational surfaces.

1. Introduction

In stellarator reactors, deuterium–tritium fusion will produce energetic, 3.5 MeV alpha particles. To minimise losses of energy from the plasma and avoid unsustainable damage to the reactor walls, these fusion-born alphas must be well confined by the stellarator magnetic field, at least until their energy is transferred back to the thermal plasma by

[†] Email address for correspondence: thomas.foster@princeton.edu

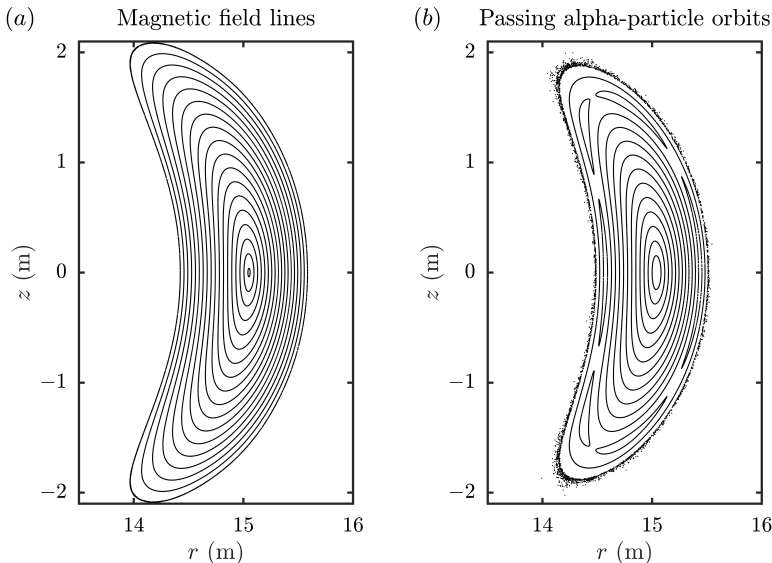


FIGURE 1. Poincaré plots, constructed in the $\phi = 0$ plane using cylindrical coordinates (r, ϕ, z) , for (a) magnetic field lines and (b) guiding-centre orbits of passing alpha particles, in a W7-X equilibrium scaled up to reactor size. The alpha particles in (b) are co-passing relative to the magnetic field and have energy 3.5 MeV and pitch-angle $\mu/\mathcal{E} = 0.143 \text{ T}^{-1}$, where μ is their magnetic moment and $\mathcal{E}$ is their kinetic energy per unit mass. The chosen value of $\mu/\mathcal{E}$ coincides with the trapped-passing boundary at the last closed flux surface, so the orbits in (b) are barely passing. The field lines and guiding-centre orbits were computed using the ASCOT5 code; more details about the equilibrium and simulation method are provided in §3.4.

collisions. This collisional slowing-down process takes $\sim 0.1 \text{ s}$ for typical reactor parameters (densities of $\sim 10^{20} \text{ m}^{-3}$ and temperatures of $\sim 15 \text{ keV}$), which is long enough for the alpha particles, travelling at $\sim 10^7 \text{ ms}^{-1}$, to complete many orbits through the device. Together with the fact that their $\sim 5 \text{ cm}$ gyroradii are a small fraction of the minor radius of a reactor-scale stellarator, this means fusion-born alphas are commonly modelled using collisionless, guiding-centre equations (e.g. Lotz *et al.*, 1992; Bader *et al.*, 2021; Velasco *et al.*, 2021; LeViness *et al.*, 2022; Paul *et al.*, 2022; Albert *et al.*, 2023; Bindel *et al.*, 2023; Figueiredo *et al.*, 2024).

Recently, White *et al.* (2022) simulated the collisionless, guiding-centre motion of alpha particles in stellarators and observed a striking difference between the behaviour of *passing* particles in stellarators and tokamaks (see also White, 2022, 2025). It is well known that, in tokamaks, axisymmetry constrains the magnetic field lines to lie on nested toroidal flux surfaces and the particle orbits to lie on nested toroidal drift surfaces. If axisymmetry is broken by a magnetic perturbation, rational flux surfaces break into magnetic islands and nearby drift surfaces break into ‘drift islands’ (e.g. Brambilla & Lichtenberg, 1973; Mynick, 1993*a,b*). In stellarators—which are inherently nonaxisymmetric—White *et al.* (2022) found that the orbits of passing particles generally contain drift islands, even when the magnetic field is unperturbed and still possesses a full set of nested toroidal flux surfaces (as shown in figure 1; see also Grad, 1967; Marcal, 1982; Wobig & Pfirsch, 2001; White, 2022, 2025; Chambliss *et al.*, 2025).

These ‘equilibrium-induced’ drift islands grow with energy, raising concerns that they could be a source of enhanced alpha-particle transport specific to stellarators (Paul *et al.*, 2022; White *et al.*, 2022; White, 2022, 2025; Chambliss *et al.*, 2025). Wide islands,

in the low-collisionality regimes relevant for fusion-born alphas, allow particles to make large radial excursions through the device. Furthermore, collisions or orbit-class transitions will enable particles to ‘hop’ in and out of these wide orbits, expediting their progress towards the reactor walls. As a result, drift islands are likely to be sites of enhanced neoclassical transport, unless they can be avoided by suitable design of the stellarator magnetic field.

To investigate how this might be achieved, we derive simple equations for the drift-island shape in a general stellarator. In this paper, we focus on passing-particle orbits, while in Paper II (Foster *et al.*, 2025), we generalize our theory to include trapped particles that complete multiple toroidal transits before bouncing, whose orbits can also form wide drift islands. The question of how drift islands modify *collisional* transport of alpha particles around rational surfaces is left for future work.

The drift-island shape follows from the solution to an even more fundamental theoretical problem: that of calculating the orbits of passing particles near a rational flux surface. We estimate the width of these orbits in §2. Then, in §3, we show that these orbits are determined by conservation of an adiabatic invariant associated with the closed rational-surface field lines. We derive this ‘transit adiabatic invariant’ by expanding the guiding-centre equations for particles close to the rational surface, then applying an orbit-averaging procedure. Similar adiabatic invariants have been used to describe locally passing particles in stellarators with a small rotational transform per field period or per ripple well (e.g. Dobrott & Frieman, 1971; Cary *et al.*, 1988; Todoroki, 1993; Rome, 1995; Smirnova, 1997). These adiabatic invariants described in previous work are not the same as the transit invariant; one difference is that the transit invariant depends explicitly on which low-order rational surface the particles are close to. An adiabatic invariant more similar to the transit invariant was derived by Hastie *et al.* (1967) (see also Bhattacharjee *et al.*, 1992) for passing particles in a magnetic field with either closed field lines or a small rotational transform. Our work generalizes the derivation of Hastie *et al.* (1967) to the case of particles near low-order rational surfaces in a stellarator with a rotational transform of order unity. Furthermore, we compute higher-order corrections to the transit invariant (see below and §6) and compare with simulations, conducted using the ASCOT5 code (Varje *et al.*, 2019), of particle orbits in various modern stellarator designs.

In §4, we discuss how the drift-island width depends on properties of the stellarator equilibrium field, such as its number of field periods, aspect ratio, and magnetic shear. In §5, we derive a necessary and sufficient condition for the drift-island width to vanish for all passing particles; we refer to magnetic fields that satisfy this condition as ‘cyclometric’. We also consider nearly cyclometric fields and determine how the drift-island width scales with the deviation from cyclometry.

To model fusion-born alphas in a reactor, it is crucial that our calculations remain valid even at high energies. However, for sufficiently energetic particles, the accuracy of orbit-averaging, which we use to obtain the transit invariant, has been questioned (e.g. Paul *et al.*, 2022). To ensure accurate results even at high energies, in §6 we calculate higher-order corrections to the transit invariant. We derive these corrections by expanding in $\sqrt{\rho_\star} \ll 1$, where $\rho_\star = \rho/L$ is the energetic-particle gyroradius divided by a typical length scale of the magnetic field. The appropriate expansion parameter is $\sqrt{\rho_\star}$ because the orbit width for passing particles near a rational surface scales with $\sqrt{\rho_\star}$, in contrast with the familiar scaling with $\rho_\star$ for particles on irrational surfaces (these estimates are obtained in §2). When the higher-order corrections are included, our analytical results are in excellent agreement with simulations, indicating that orbit-averaged guiding-centre

equations and the transit adiabatic invariant can accurately model passing alpha particles, even at 3.5 MeV.

Finally, in §7, we show how Hamiltonian perturbation theory can be used to calculate the orbits of passing particles on both rational and irrational flux surfaces in a general stellarator. This framework provides a clear criterion for which flux surfaces are ‘rational enough’ that particle orbits in their vicinity need to be described using the transit invariant.

2. Drift islands: physical picture and estimates

Drift islands arise in the orbits of passing particles near rational flux surfaces. The drift-island width is, therefore, equal to the orbit width of these particles, by which we mean the typical size of the radial excursions they make from their initial flux surface. In §2.1, we estimate this orbit width using a physical picture of the particle motion. It is well known that the orbit width of passing particles typically scales with $\rho_\star$ (e.g. Bishop & Smith, 1966; Galeev et al., 1969). However, we will find that, near low-order rational surfaces, the orbit width is wider, scaling instead with $\sqrt{\rho_\star}$ (Grad, 1967). In §2.2, we show that the same estimates can also be obtained using an analogy with magnetic islands.

2.1. Orbit width of passing particles near rational surfaces

First, we review a standard argument that the orbit width of passing particles typically scales with $\rho_\star$ (e.g. Helander, 2014) and point out why it does not apply to particles near rational flux surfaces.

A generic flux surface has an irrational rotational transform. Therefore, each magnetic field line covers the surface densely and a particle streaming along one of these field lines will explore the entire surface (see figure 2(a)). This means that the net radial drift accumulated by the particle over long times is proportional to the flux-surface average of its radial drift velocity divided by its parallel streaming rate, which weights each region of the flux surface according to how long the particle spends there (e.g. Helander, 2014). It can be shown that this flux-surface average vanishes, so passing particles do not experience a secular radial drift. Thus, the width w of a generic passing orbit can be estimated using the size of the radial drift velocity, $|\mathbf{v}_d| \sim \rho_\star v$ (v is the particle speed), and the typical time required for this drift to average out:

$$w \sim |\mathbf{v}_d| \frac{L}{v} \sim \rho_\star L. \quad (2.1)$$

This estimate is accurate for most particles but is incorrect for particles close to a low-order rational surface. Near a rational surface, the magnetic field lines almost close on themselves, as shown in figure 2(c), so a particle streaming along one of these field lines will take a very long time to explore the entire surface. In this time, the particle can drift radial distances larger than $\sim \rho_\star L$.

To estimate the orbit width of particles near a rational surface, it is helpful to have in mind the following physical picture. A particle streaming along a closed field line on the rational surface will not sample the entire surface, which means it can experience a net radial drift. This drift will carry the particle onto neighbouring flux surfaces, where the field lines no longer perfectly close on themselves (assuming the magnetic shear is non-zero); compare figures 2(b) and 2(c). By streaming along these nearly closed field lines, the particle will migrate tangentially across the rational surface—this is clearest

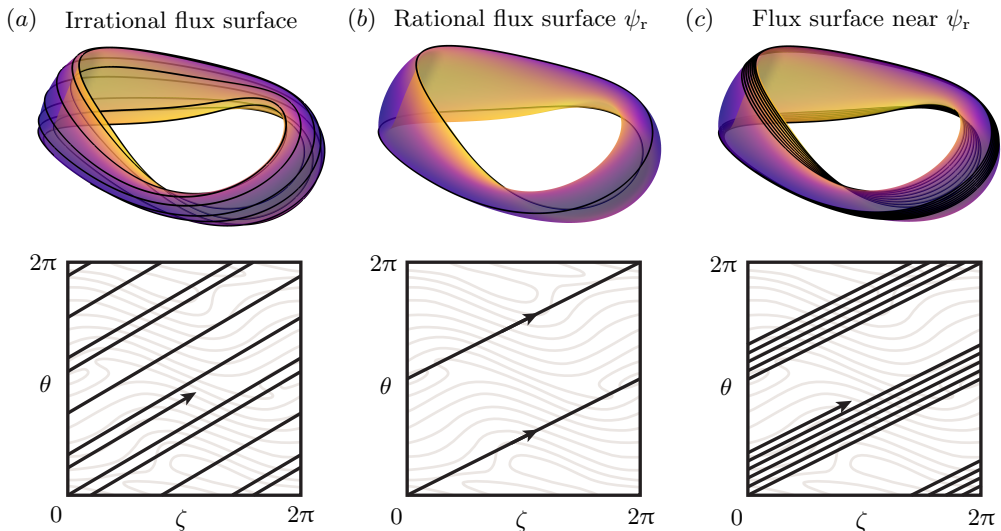


FIGURE 2. (top row) Illustration of how the rotational transform of a flux surface affects the magnetic field lines in physical space, with colour showing the magnetic field strength. For the flux surface shape, we used a precise-QA equilibrium from Landreman *et al.* (2021). (bottom row) Cartoons showing how the rotational transform affects the field lines in straight-field-line coordinates for a single-field-period stellarator, with grey contours representing the magnetic field strength. (a) On an irrational surface, field lines cover the surface densely and never close. (b) On a rational surface ψ_r , field lines close on themselves. (c) Near the rational surface ψ_r , field lines require many toroidal turns to explore the entire surface.

from a picture like figure 2(c). Eventually, its radial drift, which averages to zero over the entire rational surface, must change sign. The particle will then drift back towards the rational surface but will overshoot. By streaming along nearly closed field lines again, the particle will migrate tangentially across the rational surface, in the opposite direction to before, back to its starting point. In a Poincaré plot, this cyclic motion traces out a drift island.

To estimate the width of such an orbit, it is convenient to use straight-field-line coordinates (ψ, θ, ζ) (e.g. Helander, 2014). The radial coordinate ψ is the toroidal magnetic flux contained by the flux surface divided by 2π , and θ and ζ are the poloidal and toroidal angles, respectively. Let the rotational transform be $\iota(\psi)$. Suppose the particle is near a low-order rational surface, $\psi = \psi_r$, with $\iota(\psi_r) = N/M$, for coprime integers N and $M > 0$. The orbit width is the radial distance the particle can drift in the time required for it to migrate across a significant fraction of the rational surface by streaming along nearly closed field lines, such as those depicted in figure 2(c). We denote this orbit width, measured by the variation in the particle's ψ coordinate, by $\Delta\psi$.

Consider the motion of the particle from $\zeta = 0$ to $\zeta = 2\pi M$, where the orbit nearly ‘bites its tail’; throughout the paper, this motion is referred to as a ‘transit’. In a single transit, θ changes by $\Delta\theta \simeq 2\pi M[\iota(\psi) - N/M]$, as shown in figure 3. Assuming $M \sim N \sim 1$, this means $\Delta\theta \sim \iota'_r \Delta\psi$, where $\iota'_r := (d\iota/d\psi)(\psi = \psi_r)$ is related to the global magnetic shear by $s := |(\psi/\iota)(d\iota/d\psi)|$. The number of transits required for the particle to move over a significant portion of the flux surface, by streaming along nearly closed field lines, is $\sim 1/\Delta\theta$. Thus, the time required for the particle to migrate over the flux surface is

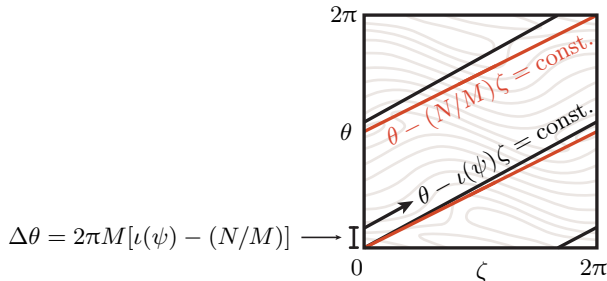


FIGURE 3. Cartoon showing the difference between a closed curve of constant $\theta - (N/M)\zeta$ (shown in red) and a field line (black) in straight-field-line coordinates on a flux surface near the rational surface. In this case, the stellarator has a single field period, $N = 1$, and $M = 2$. Grey contours represent the magnetic field strength. The change in poloidal angle, $\Delta\theta$, when the field line is followed for one full transit ($\Delta\zeta = 2\pi M$) is indicated.

$\sim (1/\iota'_t \Delta\psi)(L/v)$.¹ In this time, the change in ψ due to drifts is

$$\Delta\psi \sim \left(\frac{1}{\iota'_t \Delta\psi} \right) (\rho_* \delta) \Psi_t, \quad (2.2)$$

where Ψ_t is a typical value of the toroidal flux through a flux surface, and δ is a factor that reduces the radial drift rate in nearly omnigenous stellarators and certain other configurations (viz. cyclometric configurations; we introduce cyclometry and define δ more precisely in §5). Note that $\delta \sim 1$ in an unoptimized device. Solving (2.2) for $\Delta\psi$ gives

$$\frac{\Delta\psi}{\Psi_t} \sim \left(\frac{\rho_* \delta}{s} \right)^{1/2}. \quad (2.3)$$

As claimed, the orbit width for particles near rational surfaces, which is equivalent to the drift-island width, scales with $\sqrt{\rho_*}$ instead of ρ_* . For context, birth alpha particles in a reactor-scale stellarator might have $\rho_* \sim 0.03$ (using parameters from Ku et al. (2008) or Drevlak et al. (2014); ρ_* will be lower in a high-field stellarator). Stellarator designs usually have s between 0.01 and 1 (Landreman & Paul, 2022; Komori et al., 2006) and δ between 10^{-4} and 10^{-1} (Velasco et al., 2023). Thus, for birth alphas, estimate (2.3) suggests that the drift-island width can reach $\sim 50\%$ of the minor radius.

Estimate (2.3) does not capture the dependence of the island width on N and M , since we ordered $N \sim M \sim 1$ and disregarded order-unity factors. This dependence will be revealed by the full calculation in §3. Physically, if the flux surface is rational, but with high order (meaning M is large), the particle will sample most of the rational surface before the field lines close, so the radial drift will very nearly average out and the particle will not drift far. We will find that the island width typically decays exponentially with $|N|$ and $|M|$.

¹Here, we neglect the tangential drift, which could also be responsible for moving the particle across the flux surface. This can be justified *a posteriori* given the island-width estimate (2.3). The change in θ in one transit due to the tangential drift is $\sim \rho_*$, while (2.3) implies that the change due to streaming along sheared field lines is $\sim (\rho_* \delta s)^{1/2}$, which usually dominates. If the drift term dominates, the drift islands simply shift to a new radial location, as explained in §5.3.

2.2. Drift-island width by analogy with magnetic islands

The drift-island width can also be estimated using well-known facts about magnetic islands. This is possible because it can be shown that passing orbits follow the field lines of a modified magnetic field. To demonstrate this, we write the guiding-centre velocity, whose direction determines the trajectory of a particle's guiding centre through space, as²

$$\frac{d\mathbf{X}}{dt} = \frac{v_{\parallel}}{B} \left(\mathbf{B} + \nabla \times \left(\frac{v_{\parallel} \mathbf{B}}{\Omega} \right) \right). \quad (2.4)$$

Here, $\mathbf{X}$ is the guiding-centre position, t is time, $\mathbf{B}$ is the magnetic field, $B := |\mathbf{B}|$ is the magnetic field strength, and $\Omega := ZeB/mc$ is the gyrofrequency, where Ze and m are the particle charge and mass, respectively, and c is the speed of light (we use CGS units). All functions are evaluated at the guiding-centre position. The parallel velocity in (2.4) is

$$v_{\parallel} := \sigma \sqrt{2(\mathcal{E} - \mu B)}, \quad (2.5)$$

where $\mathcal{E}$ is the kinetic energy per unit mass of the particle, μ is its magnetic moment, and $\sigma := v_{\parallel}/|v_{\parallel}|$, and the curl in (2.4) is taken at fixed $\mathcal{E}$ and μ . We have ignored the electrostatic potential in (2.5) because it is negligible for sufficiently energetic particles.

According to (2.4), the particle does not perfectly follow magnetic field lines. Instead, it moves along the field lines of the divergence-free field (Morozov & Solov'ev, 1959, 1966)

$$\mathbf{B}^* := \mathbf{B} + \nabla \times \left(\frac{v_{\parallel} \mathbf{B}}{\Omega} \right), \quad (2.6)$$

where $\mathcal{E}$ and μ are treated as fixed parameters. For passing particles, this field $\mathbf{B}^*$ exists throughout a toroidal region and is a small perturbation to the real magnetic field $\mathbf{B}$, which we assume has nested toroidal flux surfaces. It is well known that when a magnetic field with nested toroidal flux surfaces is perturbed, rational flux surfaces break up into magnetic islands. Therefore, the field lines of $\mathbf{B}^*$, which actually represent guiding-centre trajectories through space, will exhibit islands: these are the drift islands.

If a magnetic field with nested toroidal flux surfaces is modified by a small perturbation, the width of the resulting magnetic islands scales with the square root of the ratio of the perturbation amplitude to the magnetic shear (e.g. Boozer, 2005). Applying this result to the field in (2.6), where the size of the perturbation is $\sim vB/\Omega L \sim \rho_{\star} B$, we find that the drift-island width scales with $(\rho_{\star}/s)^{1/2}$.

The scaling with $\sqrt{\rho_{\star}}$ is recovered,³ but what about the factor of $\sqrt{\delta}$, which reduces the drift-island width in optimized configurations? Later in the article (in §5.2, where we give a precise definition of δ), we will explain how the argument above may be extended to obtain this factor. To summarize, we will show that, in a perfectly cyclometric stellarator, $\mathbf{B}^*$ is integrable to higher order in $\rho_{\star}$. Therefore, in a nearly cyclometric stellarator, the perturbation to $\mathbf{B}^*$ that breaks integrability comes from the magnetic drift due to the

²Equation (2.4) gives the correct velocity along the field lines to order v and across the field lines to order $\rho_{\star} v$ (Hazeltine & Waelbroeck, 2018). In White *et al.* (1982), it was pointed out that the right side of this equation could be divided by the near-identity factor $1 + (v_{\parallel}/\Omega)(\hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}})$ to ensure that Liouville's theorem is satisfied exactly. Since we are only interested in the direction of the guiding-centre velocity, this modification is unnecessary.

³This is a special case of a general result concerning nearly integrable Hamiltonian systems: perturbations of size $\sim \varepsilon$ (relative to the unperturbed Hamiltonian) that remove integrability cause invariant tori in phase space to break up into islands of width $\sim \sqrt{\varepsilon}$ around resonances (e.g. Chirikov, 1979; Lichtenberg & Lieberman, 2013).

non-cyclometric piece of the field. The size of this perturbation is $\sim \rho_* \delta B$. Using this perturbation size in the magnetic-island-width formula leads to the scaling (2.3) that we found previously for the drift-island width, including the factor of $\sqrt{\delta}$.

3. Drift islands and the transit adiabatic invariant

In this section, we derive an equation for the drift-island shape by expanding the guiding-centre equations around a rational surface (§§3.1–3.2), which will reveal the existence of an adiabatic invariant (§3.3). The shape of the drift islands, as they would appear in a Poincaré section, can be visualized by plotting the level curves of this invariant (§3.4).

The calculations in this section rely on three main assumptions. First, we use the guiding-centre equations; in §A we compare guiding-centre and full-orbit simulations to verify that the guiding-centre model is accurate in reactor-scale stellarators even for the most energetic alpha particles. Second, we assume that the stellarator magnetic field does not change in time and that it possesses nested toroidal flux surfaces. In fact, this assumption is not necessary; in §B we explain how the analysis can be generalized to stellarators whose magnetic field contains magnetic islands. Finally, we assume that drift-island overlap does not occur. It is well known that, when the drift islands around two nearby rational surfaces become large enough to overlap, the islands are replaced by a region of stochastic orbits (Chirikov, 1979), a phenomenon not described by our theory. The assumption of no drift-island overlap is discussed further in §4.3.

3.1. Coordinate system

Near a rational surface, passing particles rapidly execute almost-closed orbits by streaming along almost-closed field lines. Meanwhile, they slowly drift perpendicular to the field lines and they slowly migrate tangentially across the rational surface as described in §2.1. It will be useful to choose coordinates that separate the fast and slow timescales.

In §2, we introduced the straight-field-line coordinates (ψ, θ, ζ) . We will still use ζ , which varies on the fast timescale of a transit, and ψ , which changes on a slower timescale due to drifts. Instead of the poloidal angle θ , we will now employ a combination of angles that changes more slowly,⁴ namely

$$\eta := \theta - \frac{N}{M} \zeta. \quad (3.1)$$

As above, N/M is the rotational transform of the rational surface ψ_r , where N and $M > 0$ are coprime integers. On the rational surface, η coincides with the usual field-line label $\alpha := \theta - \iota(\psi) \zeta$. On nearby flux surfaces, however, curves of constant η remain closed, with the helicity of the rational-surface field lines, so these curves are not perfectly tangent to the magnetic field. We can think of η as an extension of the rational-surface field-line label to nearby flux surfaces, or as a way of telling us which rational-surface field line a particle is close to (see figure 4).

We take $\zeta \in [0, 2\pi M)$ to describe position along each constant- η curve. Then, each of these curves can be uniquely labelled using $\eta \in [0, 2\pi/M)$.

⁴An analogous change of coordinates is routinely used in the theory of motion close to a resonance in quasi-integrable Hamiltonian systems. In this context, the resulting coordinate, which is a linear combination of angle coordinates that does not change for particles exactly on-resonance, is called the ‘slow angle’ (e.g. Lichtenberg & Lieberman, 2013; Hamilton et al., 2023).

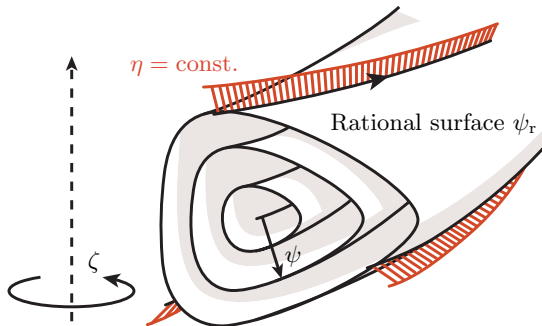


FIGURE 4. Diagram of the coordinate system that we use. The outermost flux surface displayed here is the rational surface ψ_r , on which a magnetic field line is drawn. The constant- η surface that coincides, on ψ_r , with this field line is shown in red.

3.2. Orbit calculation

The coordinates $(\psi(\mathbf{X}), \eta(\mathbf{X}), \zeta(\mathbf{X}))$ of a particle's guiding-centre position $\mathbf{X}$ evolve according to the guiding-centre equations

$$\frac{d\psi}{dt} = \mathbf{v}_d \cdot \nabla \psi, \quad (3.2a)$$

$$\frac{d\eta}{dt} = v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \eta + \mathbf{v}_d \cdot \nabla \eta, \quad (3.2b)$$

$$\frac{d\zeta}{dt} = v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta. \quad (3.2c)$$

Here, $\hat{\mathbf{b}} := \mathbf{B}/B$ is the unit vector parallel to the magnetic field, $v_{\parallel}$ is the parallel velocity, defined in (2.5), and

$$\mathbf{v}_d := \frac{v_{\parallel}}{B} \nabla \times \left(\frac{v_{\parallel} \mathbf{B}}{\Omega} \right) \quad (3.2d)$$

is the drift velocity, where the curl in (3.2d) is taken at fixed $\mathcal{E}$ and μ . All functions of position in (3.2) are evaluated at the guiding-centre position $\mathbf{X}$. As a reminder, we did not include the electrostatic potential in the definition (2.5) of the parallel velocity, which means the $\mathbf{E} \times \mathbf{B}$ drift does not appear in (3.2d), because potential variations have a negligible effect on the motion of sufficiently energetic particles. Potential variations can easily be incorporated into the following calculations by including the potential in $v_{\parallel}$.

The physical interpretation of the expressions for $d\psi/dt$ and $d\zeta/dt$ is straightforward: ψ changes due to the radial magnetic drift, while ζ changes due to parallel streaming along field lines. Meanwhile, the expression for $d\eta/dt$ in (3.2b) consists of two terms. The $\mathbf{v}_d \cdot \nabla \eta$ ‘drift term’ represents the magnetic drift pushing the particle tangentially across the flux surface. The $v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \eta$ ‘shear term’ accounts for changes in η due to the slight misalignment between the magnetic field lines, along which the particle rapidly streams, and the curves of constant η on the particle's current flux surface (see figure 3). This misalignment increases with distance from the rational surface, since the shear term can be expressed as

$$v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \eta = v_{\parallel} \left(\iota(\psi) - \frac{N}{M} \right) \hat{\mathbf{b}} \cdot \nabla \zeta. \quad (3.3)$$

We now introduce the assumption that the particle is close to the rational surface, in the

sense that

$$\left| \iota(\psi) - \frac{N}{M} \right| \ll 1. \quad (3.4)$$

In a low-shear device, this condition may be satisfied everywhere in the plasma. In a higher-shear stellarator, condition (3.4) is equivalent to $|\iota'_r(\psi - \psi_r)| \ll 1$; this defines narrow regions around low-order rational surfaces where our calculations are applicable.

When (3.4) is satisfied, the rate of change of η is small, as expected. We refer to ψ and η as the ‘slow coordinates’. To predict the evolution of the slow coordinates over many nearly periodic transits, we average, at fixed ψ and η , over the transit motion,

$$\left\langle \frac{d\psi}{dt} \right\rangle_t = \langle \mathbf{v}_d \cdot \nabla \psi \rangle_t, \quad (3.5a)$$

$$\left\langle \frac{d\eta}{dt} \right\rangle_t = \langle v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \eta \rangle_t + \langle \mathbf{v}_d \cdot \nabla \eta \rangle_t. \quad (3.5b)$$

Here, we have defined the transit average

$$\langle \dots \rangle_t := \frac{1}{\tau_t} \oint \frac{(\dots)}{|v_{\parallel}| \hat{\mathbf{b}} \cdot \nabla \zeta} d\zeta, \quad (3.6a)$$

where the transit time is

$$\tau_t := \oint \frac{1}{|v_{\parallel}| \hat{\mathbf{b}} \cdot \nabla \zeta} d\zeta. \quad (3.6b)$$

Throughout the paper, the closed-contour integral symbol denotes a line integral along a curve of constant ψ and η , in the direction of the magnetic field (without loss of generality, we assume $\hat{\mathbf{b}} \cdot \nabla \zeta > 0$ so that the transit time is positive). Therefore, in (3.6), where the integration variable is the toroidal angle, the integrals are taken from $\zeta = 0$ to $\zeta = 2\pi M$.

We need to evaluate the transit averages appearing in (3.5). Using (3.2d), the drifts appearing in these equations may be written as

$$\mathbf{v}_d \cdot \nabla \psi = \frac{v_{\parallel}}{B} \nabla \cdot \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \times \nabla \psi \right), \quad (3.7a)$$

$$\mathbf{v}_d \cdot \nabla \eta = \frac{v_{\parallel}}{B} \nabla \cdot \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \times \nabla \eta \right). \quad (3.7b)$$

The Jacobian for the (ψ, η, ζ) coordinate system is $((\nabla \psi \times \nabla \eta) \cdot \nabla \zeta)^{-1} = (\mathbf{B} \cdot \nabla \zeta)^{-1}$, so evaluating the divergences in (3.7) leads to

$$\mathbf{v}_d \cdot \nabla \psi = v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta \left[\partial_{\eta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\eta} \mathbf{x} \right) \right], \quad (3.8a)$$

$$\mathbf{v}_d \cdot \nabla \eta = v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta \left[\partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) - \partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) \right]. \quad (3.8b)$$

Here, $\mathbf{x}(\psi, \theta, \zeta)$ denotes the point in physical space with coordinates (ψ, θ, ζ) . Throughout the paper, unless otherwise specified, the partial-derivative symbols ∂_{ψ} , ∂_{η} and ∂_{ζ} always represent partial derivatives with respect to one of the coordinates (ψ, η, ζ) , with the other two held fixed (and with $\mathcal{E}$ and μ fixed). Thus, the vectors $\partial_{\psi} \mathbf{x}$, $\partial_{\eta} \mathbf{x}$ and $\partial_{\zeta} \mathbf{x}$ are the contravariant basis vectors for our straight-field-line coordinate system. Using (3.8), the transit-averaged drifts are

$$\langle \mathbf{v}_d \cdot \nabla \psi \rangle_t = \frac{mc}{Ze\tau_t} \partial_{\eta} I, \quad (3.9a)$$

$$\langle \mathbf{v}_d \cdot \nabla \eta \rangle_t = -\frac{mc}{Ze\tau_t} \partial_{\psi} I, \quad (3.9b)$$

where

$$I(\psi, \eta, \mathcal{E}, \mu) := \oint |v_{\parallel}| \hat{\mathbf{b}} \cdot \partial_{\zeta} \mathbf{x} \, d\zeta. \quad (3.10)$$

Formulas (3.9) are reminiscent of the usual identities for the drift rate of trapped particles (see, e.g. Helander, 2014, equations (60) and (61)). Interestingly, formulas (3.9) are exact in the following sense: their derivation, starting with (3.7), did not require anywhere that the particle is close to the rational surface ψ_r . We will make use of this fact in §6.3. Of course, transit-averaging is only meaningful if the integration contour approximates the actual particle trajectory, which is only the case when $|\iota - N/M| \ll 1$.

We also need the transit average of the shear term, which follows from (3.3),

$$\langle v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \eta \rangle_t = \frac{2\pi M \sigma}{\tau_t} \left(\iota(\psi) - \frac{N}{M} \right). \quad (3.11)$$

This equation is physically transparent: in one transit, which takes time τ_t , the change in η due to streaming is $2\pi M \sigma [\iota(\psi) - N/M]$, as shown in figure 3.

Finally, using results (3.9) and (3.11) in (3.5), we find that the transit-averaged rate of change of the slow coordinates is

$$\left\langle \frac{d\psi}{dt} \right\rangle_t = \frac{\sigma mc}{Ze\tau_t} \partial_{\eta} \mathcal{I}, \quad (3.12a)$$

$$\left\langle \frac{d\eta}{dt} \right\rangle_t = -\frac{\sigma mc}{Ze\tau_t} \partial_{\psi} \mathcal{I}. \quad (3.12b)$$

Here, we have introduced the ‘transit adiabatic invariant’

$$\mathcal{I}(\psi, \eta, \mathcal{E}, \mu, \sigma) := \sigma I(\psi, \eta, \mathcal{E}, \mu) - \frac{2\pi M Ze}{mc} \int_{\psi_r}^{\psi} \left(\iota(\psi') - \frac{N}{M} \right) d\psi', \quad (3.13)$$

which, according to (3.12), is conserved on average over multiple transits:

$$\left\langle \frac{d\mathcal{I}}{dt} \right\rangle_t = \partial_{\psi} \mathcal{I} \left\langle \frac{d\psi}{dt} \right\rangle_t + \partial_{\eta} \mathcal{I} \left\langle \frac{d\eta}{dt} \right\rangle_t = 0. \quad (3.14)$$

The existence of the adiabatic invariant (3.13) for passing particles near a rational flux surface is one of the key results of this paper.

3.3. Transit adiabatic invariant

We now discuss the physical meaning of the transit adiabatic invariant. Its definition in (3.13) is a sum of two terms. The first is a ‘kinetic’ term $I(\psi, \eta, \mathcal{E}, \mu)$, which can be simplified using the fact that $\hat{\mathbf{b}}$ is parallel to $\partial_{\zeta} \mathbf{x}$ to leading order in $|\iota - N/M| \ll 1$,

$$\hat{\mathbf{b}} = (\hat{\mathbf{b}} \cdot \nabla \zeta) \left[\partial_{\zeta} \mathbf{x} + \left(\iota(\psi) - \frac{N}{M} \right) \partial_{\eta} \mathbf{x} \right] \simeq (\hat{\mathbf{b}} \cdot \nabla \zeta) \partial_{\zeta} \mathbf{x}. \quad (3.15)$$

Then,

$$I \simeq \oint |v_{\parallel}| \, dl. \quad (3.16)$$

Here, l denotes arc length along the integration contour, which, as before, is the closed curve obtained by holding the values of ψ and η fixed. The kinetic term, in the form (3.16), is similar to the usual second adiabatic invariant for trapped particles (e.g. Helander, 2014).

The second term in (3.13) is a ‘geometric’ term that depends only on the distance $\psi - \psi_r$

from the rational surface; it is independent of $\mathcal{E}$ and μ . This term has a simple physical interpretation: it is proportional to the ‘helical flux’ $\chi_r(\psi)$ (Boozer, 2016) passing between flux surface ψ and the rational surface ψ_r ,

$$\chi_r(\psi) := \int_{\psi_r}^{\psi} \left(\iota(\psi') - \frac{N}{M} \right) d\psi'. \quad (3.17)$$

The reason $\chi_r(\psi)$ is called the helical flux is as follows. It is straightforward to check that $\mathbf{B} = \nabla \times \mathbf{A}$ for a vector potential $\mathbf{A} := (\psi - \psi_r) \nabla \eta - \chi_r(\psi) \nabla \zeta$. Then, we have

$$\oint \mathbf{A} \cdot d\mathbf{x} = -2\pi M \chi_r(\psi). \quad (3.18)$$

The name ‘helical flux’ is appropriate because we can apply Stokes’ Theorem to the line integral on the left side of (3.18) to conclude that $-2\pi M \chi_r(\psi)$ equals the magnetic flux through the two-dimensional surface bounded by the two curves labelled by (ψ, η) and (ψ_r, η) , which will look similar to the red helical strip in figure 4.

Using formulas (3.10) and (3.18) for the kinetic and geometric terms, respectively, the transit adiabatic invariant can be expressed as

$$\mathcal{I} = \oint \left(v_{\parallel} \hat{\mathbf{b}} + \frac{Ze}{mc} \mathbf{A} \right) \cdot d\mathbf{x}. \quad (3.19)$$

Equation (3.19) is similar to the standard formula

$$\mathcal{A} = \oint \mathbf{p} \cdot d\mathbf{q} \quad (3.20)$$

for adiabatic invariants in Hamiltonian systems (Hazeltine & Waelbroeck, 2018). Here, $\mathbf{q}$ is the vector of generalized position coordinates and $\mathbf{p}$ is the associated canonical momentum. In Cartesian coordinates, the canonical momentum for a charged particle in a magnetic field is $\mathbf{p} = m\mathbf{v} + (Ze/c)\mathbf{A}$. For trapped particles, (3.20) leads to the usual second adiabatic invariant because the vector potential term in the canonical momentum cancels out upon integrating along both halves of the bounce motion (along the field line and back). However, for passing particles, this vector potential term *does not* vanish, because the integral is taken along a closed loop circulating in one direction around the stellarator.

3.4. Drift-island shape

We have shown that passing-particle orbits near a rational surface ψ_r conserve the transit adiabatic invariant $\mathcal{I}$. Plotting the level sets of $\mathcal{I}$ allows us to visualize these orbits, and any drift islands they trace out, as they would appear in a Poincaré plot. For example, plotting the level sets of $\mathcal{I}$ in the (η, ψ) plane shows the drift islands as they would appear in a Poincaré plot of ψ versus θ at a fixed toroidal angle.

Before making such a plot, we can further simplify formula (3.13) for $\mathcal{I}$. So far, the only assumption we have made regarding proximity to the rational surface is $|\iota - N/M| \ll 1$, which, if the shear is low, may still be satisfied at relatively large values of $|\psi - \psi_r|$. We now make the stronger assumption that $|\psi - \psi_r| \ll \Psi_t$, since the drift islands usually fall inside a small region around ψ_r . We will see, in §5.3, that this is equivalent to assuming that the magnetic shear satisfies $s \gg \rho_*$. The condition $|\psi - \psi_r| \ll \Psi_t$ allows us, in (3.13), to replace $I(\psi, \eta, \mathcal{E}, \mu)$ with its value at the rational surface, for which we write

$$I_r(\eta, \mathcal{E}, \mu) := I(\psi_r, \eta, \mathcal{E}, \mu) = \oint_r \frac{|v_{\parallel}|}{\hat{\mathbf{b}} \cdot \nabla \zeta} d\zeta. \quad (3.21)$$

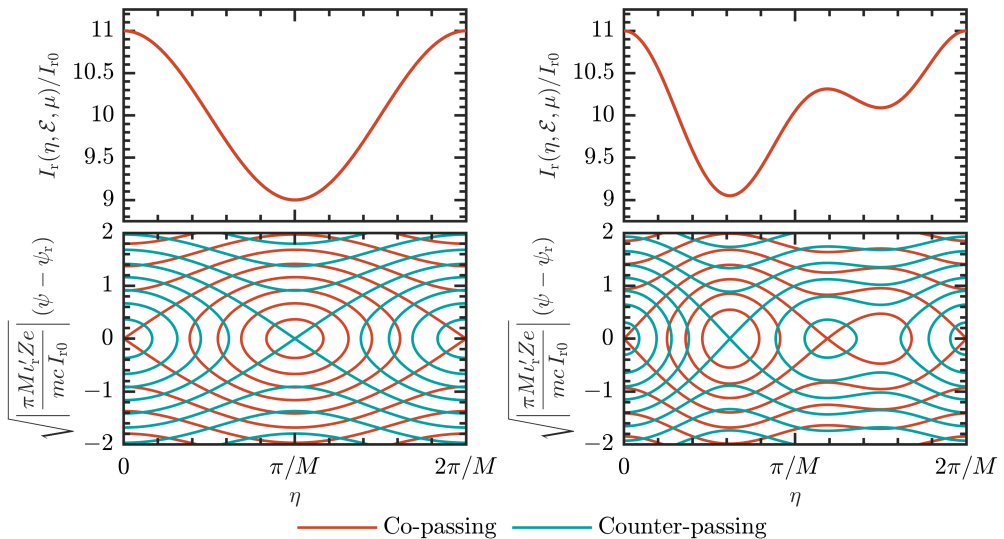


FIGURE 5. (top row) Examples of possible $I_r(\eta, \mathcal{E}, \mu)$, normalized by a typical value $I_{r0} \sim vL$, plotted against η . (bottom row) Particle orbits near the rational surface, visualized by plotting level sets of the invariant $\mathcal{I}_r(\psi, \eta, \mathcal{E}, \mu, \sigma)$ (defined in (3.22)) corresponding to the I_r functions above. These level sets are plotted against $\psi - \psi_r$ (normalized to be dimensionless) on the vertical axis and η on the horizontal axis. Co-passing ($\sigma = 1$) and counter-passing ($\sigma = -1$) orbits are plotted in red and blue, respectively. For these plots, ι'_r was assumed to be negative.

Throughout the paper, the subscript ‘r’ denotes an integral taken along the rational-surface field line labelled by (ψ_r, η) . With this replacement, together with the approximation $\iota(\psi) - N/M \simeq \iota'_r(\psi - \psi_r)$, the transit invariant becomes⁵

$$\mathcal{I}(\psi, \eta, \mathcal{E}, \mu, \sigma) \simeq \mathcal{I}_r(\psi, \eta, \mathcal{E}, \mu, \sigma) := \sigma I_r(\eta, \mathcal{E}, \mu) - \frac{\pi M Z e}{mc} \iota'_r(\psi - \psi_r)^2. \quad (3.22)$$

The level sets of this function are drawn in the (η, ψ) plane in figure 5 for two possible forms of $I_r(\eta, \mathcal{E}, \mu)$. In the top row of figure 5, we show the two example functional forms of $I_r(\eta, \mathcal{E}, \mu)$. In the bottom row, we plot the levels sets of $\mathcal{I}_r(\psi, \eta, \mathcal{E}, \mu, \sigma)$ corresponding to the I_r functions above, for both possible signs of σ : co-passing ($\sigma = 1$) and counter-passing ($\sigma = -1$) orbits are plotted in red and blue, respectively. The two functional forms of $I_r(\eta, \mathcal{E}, \mu)$ in the top row were chosen because they give rise to qualitatively different drift-island shapes in the bottom row; the case on the left is more representative of what is observed in real stellarators.

In each case, the level sets of $\mathcal{I}_r(\psi, \eta, \mathcal{E}, \mu, \sigma)$ display a clear island structure. We also see that co- and counter-passing particles trace out different drift islands. The reason is that, in (3.22), the sign of I_r is different for co- and counter-passing particles, while the sign of the geometric term is not. This is a consequence of the fact that, in the canonical momentum $\mathbf{p} = m\mathbf{v} + (Ze/c)\mathbf{A}$, the first term depends on the direction of circulation around the torus, while the second term does not.

The discussion in this subsection shows that the transit-invariant equation, (3.13) or (3.22),

⁵The approximation (3.22) for the transit invariant is equivalent to the standard ‘pendulum approximation’ for the Hamiltonian close to a resonance in a quasi-integrable Hamiltonian system (e.g. Lichtenberg & Lieberman, 2013; Hamilton *et al.*, 2023).

100 keV alphas in reactor-scale W7-X

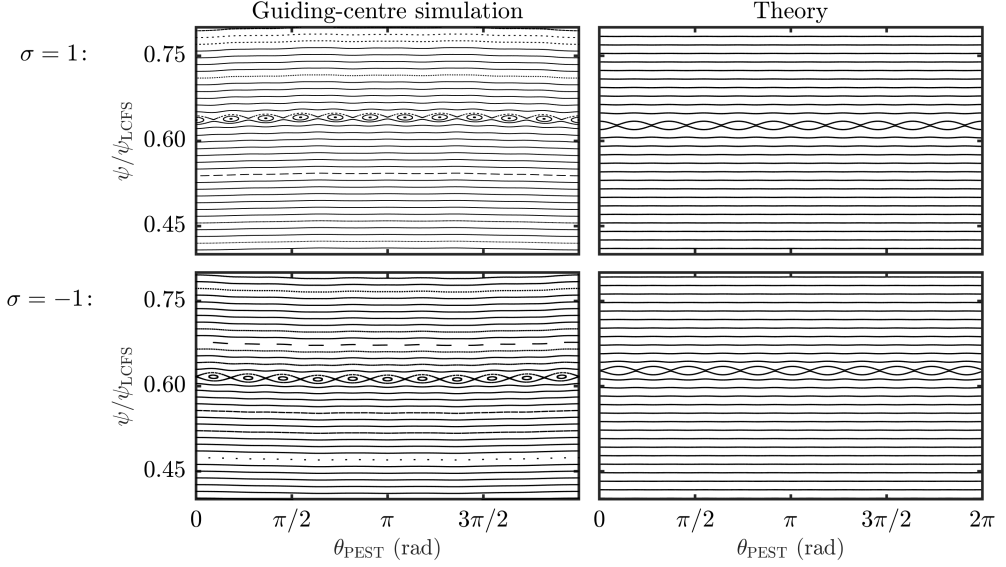


FIGURE 6. (left panels) Poincaré plot of the guiding-centre orbits of passing alpha particles in a reactor-scale W7-X equilibrium, simulated using ASCOT5. These Poincaré plots are taken at cylindrical angle $\phi = 0$. The radial coordinate is ψ normalized by its value at the last closed flux surface, ψ_{LCFS} . The coordinate on the horizontal axis is the poloidal straight-field-line angle θ_{PEST} that corresponds to the cylindrical toroidal angle. The alphas are freely passing, with energy 100 keV and $\mu = 0$. They are co-passing ($\sigma = 1$) in the upper panel and counter-passing ($\sigma = -1$) in the lower panel. (right panels) Level sets of $\mathcal{I}_r(\psi, \eta, \mathcal{E}, \mu, \sigma)$, defined in (3.22), for particles near the $\iota = 10/11$ rational surface. In each panel on the right, the sign of σ is the same as in the corresponding panel on the left.

3.5 MeV alphas in reactor-scale W7-X

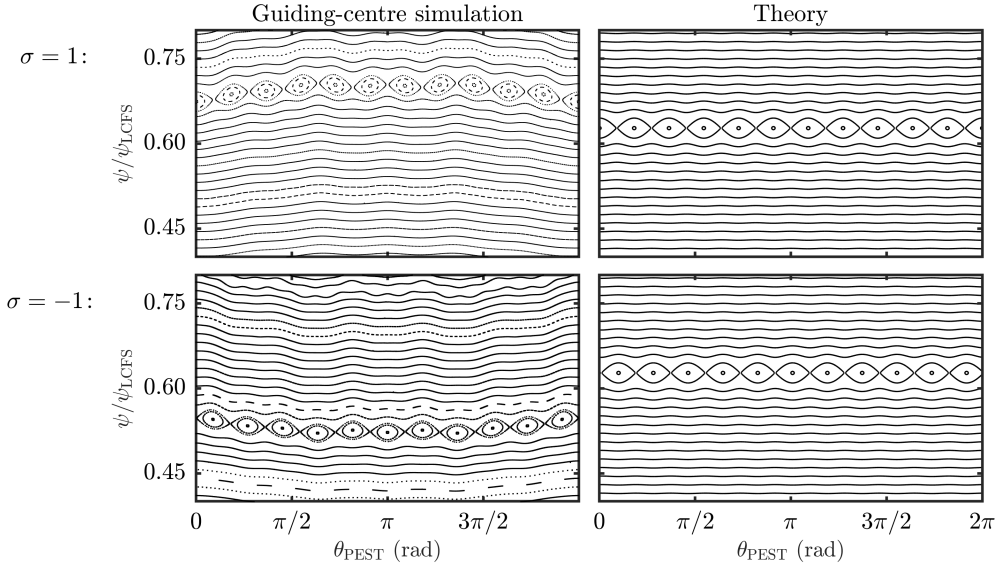


FIGURE 7. Poincaré plot and theory comparison similar to figure 6 but for alphas with energy 3.5 MeV. We see that the theory predicts the drift-island width accurately but fails to capture the θ dependence of the shape of the island chain or its radial shift away from the rational surface.

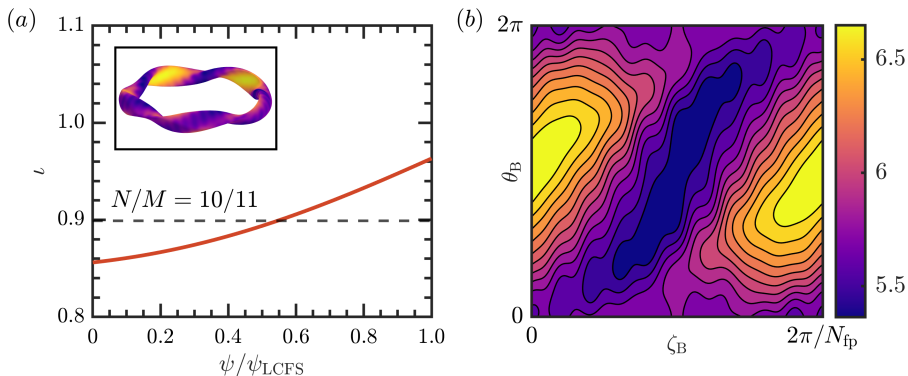


FIGURE 8. (a) Rotational transform of the W7-X equilibrium we use to trace alpha orbits in [figures 6](#) and [7](#). (b) Plot of the magnetic field strength in Boozer coordinates on the rational surface with $\iota = N/M = 10/11$.

can be thought of as an equation for the drift-island shape. In [figures 6](#) and [7](#), we compare this theory with Poincaré plots obtained from collisionless, guiding-centre simulations conducted using the ASCOT5 orbit-following code ([Varje et al., 2019](#)). The guiding-centre equations simulated by ASCOT5 are not exactly the same as the guiding-centre equations we used for our calculations because ASCOT5 includes relativistic effects; this makes very little difference to the alpha-particle trajectories at the relevant energies..

The equilibrium used for these simulations was a $\beta = 2\%$ W7-X ([Beidler et al., 1990](#)) standard-configuration equilibrium scaled up to reactor size (throughout this article, all equilibria were scaled to the same volume, 444 m^3 , and volume-averaged magnetic field strength, 5.86 T , as the ARIES-CS reactor design, as in [Bader et al. \(2021\)](#)). We provide more information about this equilibrium in [figure 8](#). The rotational-transform profile shows that there is a rational surface with $\iota = N/M = 10/11$ in the plasma; drift islands form around this flux surface. This W7-X equilibrium is from the examples directory for the DESC stellarator optimization package ([Dudt & Kolemen, 2020](#)). We used a scaled-up DESC equilibrium to evaluate our theoretical expressions. In order to load the magnetic-field data into ASCOT5, we created a VMEC file ([Hirshman et al., 1986](#)) for the equilibrium and evaluated the magnetic field on a cylindrical grid using the EXTENDER code ([Drevlak et al., 2005](#)).

In [figures 6](#) and [7](#), we see that the level sets of $\mathcal{I}_r$ accurately match the Poincaré plots from the simulations at 100 keV but differ at 3.5 MeV . Typically, the theory begins to deviate from the simulations at higher energies. Nevertheless, while [\(3.13\)](#) may not satisfactorily capture all of the features of the drift islands at high energies, it does usually give an accurate prediction of the drift-island *width*. For this reason, in the next two sections, we use [\(3.13\)](#) to investigate how the drift-island width depends on certain properties of the stellarator equilibrium. Later, in [§6](#), we will carry out a more accurate calculation of the drift-island shape that precisely matches the simulations even at very high energies.

4. Drift-island width in different stellarators

Although the lowest-order transit adiabatic invariant fails to capture all features of the drift islands at high energies, it does provide an accurate prediction of the drift-island width. We can, therefore, use (3.13) or (3.22) to find out how the stellarator magnetic field should be designed to ensure that any drift islands are small. In this section, we investigate how the drift-island width depends on properties of the equilibrium magnetic field, such as the number of field periods (§4.1), aspect ratio (§4.2), and shear (§4.3).

4.1. Number of field periods

First, we discuss how the drift-island width depends on the number of field periods, N_{fp} . In simulations, White et al. (2022) and White (2022) observed drift islands only around rational surfaces whose values of N share common factors with N_{fp} ; they describe this as a ‘resonance with the modulation of the background field’. Below, we explain this observation by showing that the drift-island width predicted by (3.22) is significantly larger if N and N_{fp} share common factors.

For particles that cross the rational surface at $\eta = \eta_r$, (3.22) is equivalent to

$$\psi - \psi_r = \pm \sqrt{\frac{mc\sigma [I_r(\eta, \mathcal{E}, \mu) - I_r(\eta_r, \mathcal{E}, \mu)]}{\pi M Ze \iota'_r}}. \quad (4.1)$$

The drift-island half-width is the maximum value of this square root, so we must determine the maximum value of $|I_r(\eta, \mathcal{E}, \mu) - I_r(\eta_r, \mathcal{E}, \mu)|$ as η varies. Recall that I_r is defined by

$$I_r = \frac{Ze\sigma}{mc} \oint_r V d\zeta, \quad (4.2)$$

where we introduce $V := (mc/Ze)(v_{\parallel}/\hat{\mathbf{b}} \cdot \nabla \zeta)$ for later convenience. We may expand V in (4.2) as a Fourier series in the angular coordinates,

$$V = \sum_{(p,q) \in \mathbb{Z}^2} V_{pq} e^{i(p\theta - q\zeta)}. \quad (4.3)$$

We assume that the straight-field-line coordinates are chosen to have the same periodicity as the magnetic field strength; that is, the straight-field-line coordinates are chosen so that $\hat{\mathbf{b}} \cdot \nabla \zeta$ is periodic under ζ -increments of $2\pi/N_{\text{fp}}$. An arbitrary straight-field-line coordinate system need not have this property—Boozer coordinates are one example of a coordinate system that does. With this choice, many of the coefficients V_{pq} are zero; the fact that there are N_{fp} field periods means the only non-zero harmonics have $q = N_{\text{fp}}n$ for some $n \in \mathbb{Z}$. Therefore, we have

$$I_r = \frac{Ze\sigma}{mc} \int_0^{2\pi M} \sum_{(p,n) \in \mathbb{Z}^2} V_{p, N_{\text{fp}}n} e^{ip\eta} e^{i((N/M)p - N_{\text{fp}}n)\zeta} d\zeta \quad (4.4)$$

since the integral in (4.2) is taken over a curve with $\eta = \theta - (N/M)\zeta$ fixed. In (4.4), the only terms in the sum which do not integrate to zero are those with

$$Np - MN_{\text{fp}}n = 0. \quad (4.5)$$

The solutions to this linear Diophantine equation take the form $(p, n) = (k\widetilde{M}, k\widetilde{N})$ for some $k \in \mathbb{Z}$, where $\widetilde{M} := MN_{\text{fp}}/\text{gcd}(N, N_{\text{fp}})$, $\widetilde{N} := NN_{\text{fp}}/\text{gcd}(N, N_{\text{fp}})$, and we have used the fact that N and M are coprime. Here, $\text{gcd}(a, b)$ stands for the greatest common divisor of a and b ; thus, $|\widetilde{N}|$ is the smallest positive integer divisible by both N and N_{fp} ,

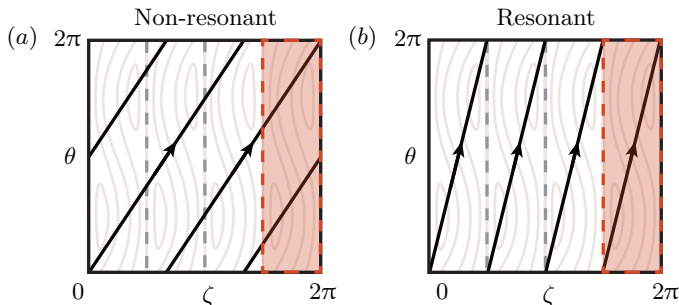


FIGURE 9. Illustration of the difference between (a) a generic rational surface (in this case, $M = 2$, $N = 3$, and $N_{fp} = 4$) and (b) a rational surface for which N shares common factors with N_{fp} (in this case, $M = 1$, $N = 4$, and $N_{fp} = 4$). When N and N_{fp} share common factors, field lines visit equivalent points in different field periods. In each case, one field period is highlighted in red.

while $\widetilde{M}$ is the integer that makes $\widetilde{N}/\widetilde{M} = N/M$. Consequently, (4.4) is equivalent to

$$I_r = \frac{2\pi M Z e \sigma}{mc} \sum_{k \in \mathbb{Z}} V_{k\widetilde{M}, k\widetilde{N}} e^{ik\widetilde{M}\eta}. \quad (4.6)$$

Typically, the dominant harmonics are those with the smallest values of $|p|$ and $|q|$, so we can approximate I_r by keeping the $k = 0$ and $k = \pm 1$ terms only. Then,

$$I_r \approx \frac{2\pi M Z e \sigma}{mc} \left(V_{00} + V_{\widetilde{M}\widetilde{N}} e^{i\widetilde{M}\eta} + V_{\widetilde{M}\widetilde{N}}^* e^{-i\widetilde{M}\eta} \right), \quad (4.7)$$

where the asterisk denotes complex conjugation. Substituting this approximation for I_r into (4.1), we find that the drift-island half-width is

$$\Delta\psi_{hw} \approx \sqrt{\left| \frac{8V_{\widetilde{M}\widetilde{N}}}{\iota_r'} \right|}. \quad (4.8)$$

We learn that the drift-island width in a stellarator is determined by resonant harmonics of V . If the magnetic field is analytic, then the amplitudes of these harmonics decay exponentially with $|N|$ and $|M|$, which means the drift-island width also decays exponentially with $|N|$ and $|M|$.

Using (4.8), we can now explain why the drift islands are widest when N and N_{fp} share common factors: increasing $\gcd(N, N_{fp})$ reduces $|\widetilde{N}|$ and $|\widetilde{M}|$, which typically makes $|V_{\widetilde{M}\widetilde{N}}|$ larger. There is a simple physical explanation for this effect (see figure 9). We remind the reader that, on irrational surfaces, the radial drift experienced by passing particles rapidly averages out because they sample the entire flux surface. Meanwhile, near rational surfaces, passing particles experience the same radial drift again and again, so they accumulate much larger radial displacements. If there are multiple field periods, then particles can experience the same radial drift before returning to their starting point; their orbit simply has to pass through a point *equivalent* to their initial position but in a different field period. For such particles, the extent to which their radial drifts reinforce, instead of averaging out, is amplified, leading to wider drift islands. For a rational-surface field line to pass through a point equivalent to the initial point but in a different field period, N_{fp} and N must share common factors; for example, see figure 9(b). The more factors N and N_{fp} share, the more times particles will experience the same drifts before returning to their starting points, and the larger the drift islands will be. As an example

of this effect, [figures 6 and 7](#) show that drift islands arise in W7-X at the rational surface with $\iota = 10/11$, for which the numerator $N = 10$ is divisible by $N_{\text{fp}} = 5$.

In summary, to avoid wide drift islands, the most important low-order rationals for the rotational-transform profile to avoid are those whose numerators N share common factors with N_{fp} .

4.2. Aspect ratio

Usually, the major radius R of a stellarator is much larger than the minor radius a ; most stellarators constructed to date have values of $\epsilon := a/R$ around 0.1–0.3 ([Beidler et al., 2011](#)). However, so far, in our estimates of the drift-island width, we have ordered gradients of the field using a single length scale L . In this section, we investigate how the drift-island width scales in a *large-aspect-ratio* stellarator with $\epsilon \ll 1$. For this discussion, we define $\rho_\star := \rho/a$; this is the only subsection in which we do not take $a \sim R \sim L$.

It will be convenient to use Boozer coordinates $(\psi, \theta_B, \zeta_B)$ ([Boozer, 1981](#)), since $\hat{\mathbf{b}} \cdot \nabla \zeta_B = B/(G_{\text{pol}} + \iota G_{\text{tor}})$ depends on the angular coordinates via the magnetic field strength only. Here, $cG_{\text{tor}}(\psi)/2$ is the toroidal current inside flux surface ψ and $cG_{\text{pol}}(\psi)/2$ is the poloidal current linked outside the flux surface. If we define $G(\psi) := G_{\text{pol}}(\psi) + (N/M)G_{\text{tor}}(\psi)$, then in Boozer coordinates [\(4.2\)](#) reads

$$I_r = G_r \oint_r \frac{\sqrt{2(\mathcal{E} - \mu B)}}{B} d\zeta_B, \quad (4.9)$$

where $G_r := G(\psi_r)$. In a large-aspect-ratio stellarator with $\beta := 8\pi P/B^2 \lesssim \epsilon$ (P is pressure), the magnetic field strength satisfies

$$B(\psi, \theta_B, \zeta_B) = B_0(\zeta_B) + B_1(\psi, \theta_B, \zeta_B), \quad (4.10)$$

where $B_1 \sim \epsilon B_0$ ([d’Herbement et al., 2022](#)). The currents in such a stellarator satisfy $G_{\text{tor}} \sim RB$ and $G_{\text{pol}} \sim aB$. We expand I_r in $\epsilon \ll 1$ by substituting [\(4.10\)](#) into [\(4.9\)](#) and ordering $v_{\parallel} \sim v$ for passing particles. The result is

$$I_r = G_r \oint_r \frac{\sqrt{2(\mathcal{E} - \mu B_0)}}{B_0} \left(1 - \frac{2\mathcal{E} - \mu B_0}{2(\mathcal{E} - \mu B_0)} \frac{B_1}{B_0} + O(\epsilon^2) \right) d\zeta_B. \quad (4.11)$$

The first term in [\(4.11\)](#) is independent of η , since B_0 depends on ζ_B only, so $I_r(\eta, \mathcal{E}, \mu) - I_r(\eta_r, \mathcal{E}, \mu) \sim \epsilon v R \sim v a$. Using this result in [\(4.1\)](#), with $\Psi_t \sim a^2 B$ and $\iota'_r \sim s/a^2 B$, we find that the drift-island width in a large-aspect-ratio stellarator is

$$\frac{\Delta\psi}{\Psi_t} \sim \left(\frac{\rho_\star}{s} \right)^{1/2}. \quad (4.12)$$

Thus, increasing the aspect ratio by making the major radius larger, at fixed $\rho_\star = \rho/a$, does not reduce the drift-island width as a fraction of the minor radius.

Interestingly, there is one case where increasing the aspect ratio in this way does lead to smaller drift islands. This is the case of stellarators with ‘mirror ratios close to unity’, which means the variation of the magnetic field strength along each field line is small in ϵ , as in a large-aspect-ratio tokamak. The magnetic field strength in such a device takes the form ([d’Herbement et al., 2022](#))

$$B(\psi, \theta_B, \zeta_B) = B_0 + B_1(\psi, \theta_B, \zeta_B), \quad (4.13)$$

where $B_1 \sim \epsilon B_0$ and B_0 is now a constant rather than a function of ζ_B . With [\(4.13\)](#), and

assuming that $\mathcal{E} - \mu B_0 \gg \mu B_1$ for passing particles sufficiently far from the trapped–passing boundary, the expansion (4.11) for I_r becomes

$$I_r = \frac{G_r \sqrt{2(\mathcal{E} - \mu B_0)}}{B_0} \left(2\pi M - \frac{2\mathcal{E} - \mu B_0}{2(\mathcal{E} - \mu B_0)} \frac{1}{B_0} \oint_r B_1 d\zeta_B + O(\epsilon^2) \right). \quad (4.14)$$

Important information about the integral appearing in (4.14) can be derived from MHD force balance. In MHD equilibrium, it is well known that, for the current density on ψ_r to be divergence-free, the ‘Hamada condition’ (Hamada, 1962) must be satisfied:

$$P'_r \partial_\eta \left(\oint_r \frac{dl}{B} \right) = 0, \quad (4.15)$$

where $P'_r := (dP/d\psi)(\psi = \psi_r)$ is the pressure gradient at the rational surface. For a large-aspect-ratio magnetic field with mirror ratio close to unity and a non-zero pressure gradient, expanding the Hamada condition (4.15) to order ϵ gives

$$\partial_\eta \left(\oint_r B_1 d\zeta_B \right) = O(\epsilon^2 B). \quad (4.16)$$

This tells us that the integral in (4.14) does not vary with η to leading order in $\epsilon \ll 1$, which implies $I_r(\eta, \mathcal{E}, \mu) - I_r(\eta_r, \mathcal{E}, \mu) \sim \epsilon^2 v R$.⁶ Therefore, the drift-island width becomes

$$\frac{\Delta\psi}{\Psi_t} \sim \left(\frac{\epsilon \rho_\star}{s} \right)^{1/2}, \quad (4.17)$$

which is smaller than the result (4.12) for a generic large-aspect-ratio stellarator. Thus, if the mirror ratio is close to unity and the Hamada condition is satisfied with a non-zero pressure gradient, then, for passing particles sufficiently far from the trapped–passing boundary, a large aspect ratio helps to reduce the drift-island width as a fraction of the minor radius.

4.3. Magnetic shear

We conclude this section with a few observations about the relationship between the drift-island width and the equilibrium magnetic shear. Scaling (2.3) shows that the drift-island width increases as the shear decreases. It might, therefore, seem that drift-island overlap is more likely to occur when the shear is low. However, the distance between two given rational surfaces scales with $1/s$, while the drift-island width scales with $1/\sqrt{s}$. Thus, as s decreases, the rational surfaces move apart faster than the drift islands get wider (similar reasoning is valid for islands due to other types of resonance, e.g. Paul *et al.*, 2023). Therefore, drift-island overlap is, in fact, less likely to occur at low shear, for a given $\rho_\star$.

In principle, drift-island overlap will eventually occur if the particle energy is increased (i.e. making $\rho_\star$ larger) at a fixed s . In our simulations, we did not observe drift-island

⁶This result can also be obtained using the equations for the radial drift velocity and MHD equilibrium in a large-aspect-ratio stellarator derived by d’Herbemont *et al.* (2022). Their equation (3.19) can be combined with their (3.14) and (B4), provided $P' \neq 0$, to conclude that $\mathbf{v}_d \cdot \nabla \psi$ is the partial derivative with respect to l of a complicated function, up to corrections of order $\epsilon^2 \rho_\star v \Psi_t / L$. This partial derivative vanishes under a transit average for passing particles that are sufficiently far from the trapped–passing boundary. So, the transit-averaged radial drift rate of these particles is order $\epsilon^2 \rho_\star v \Psi_t / L$, a factor of ϵ smaller than it would be in a generic large-aspect-ratio stellarator.

overlap for passing alpha particles away from the trapped–passing boundary⁷ in any of the stellarator equilibria that we studied, even at energies of 3.5 MeV or more. As a result, we expect stochastic transport due to drift-island overlap not to be a relevant transport mechanism for passing alpha particles in the equilibrium magnetic field of a reactor-scale stellarator.

Low magnetic shear has one other important consequence for the drift islands: it allows them to shift, radially, away from their associated rational surface. This phenomenon will be discussed in §5.3 and §6.

5. Cyclometric stellarators

We have shown that scaling certain parameters of the stellarator magnetic field—for example, increasing its aspect ratio or magnetic shear—can reduce the drift-island width. In this section, we investigate whether there are stellarator magnetic fields for which the drift-island width vanishes completely. In §5.1, we show that there are no drift islands in perfectly omnigeneous stellarators, which is equivalent to a result in appendix C of [Burby et al. \(2023\)](#). This is fortunate because modern stellarator designs are invariably optimized for omnigeneity (an omnigeneous stellarator is one in which the bounce-averaged radial drift velocity of every trapped particle vanishes; see, e.g. [Hall & McNamara, 1975](#); [Cary & Shasharina, 1997a,b](#); [Parra et al., 2015](#)). Surprisingly, we also find that there are no drift islands in some highly non-omnigeneous configurations: we show that the drift-island width is zero for all passing particles if and only if the magnetic field has a property we call *cyclometry*. We discuss nearly cyclometric stellarators in §5.2 and determine how the drift-island width scales with the deviation from cyclometry. The scaling changes when the magnetic shear is extremely low (viz. $s \lesssim \rho_*$); we discuss nearly cyclometric stellarators with such low values of the magnetic shear in §5.3.

5.1. Stellarators without drift islands

According to (4.1), the width of any drift islands around rational surface ψ_r will be zero if and only if

$$\partial_\eta I_r = \partial_\eta \left(\oint_r \sqrt{2(\mathcal{E} - \mu B)} dl \right) = 0 \quad (5.1)$$

for all passing particles ([Grad, 1967](#); [Burby et al., 2023](#)). Physically, (5.1) ensures that the net radial drift vanishes for particles that stream all the way around a closed field line on the rational surface.

In this subsection, we assume that (5.1) is satisfied for all passing particles and investigate what this implies for the magnetic field. In fact, our arguments only require that there exists some open neighbourhood $\mathfrak{N}$ of zero such that (5.1) is satisfied for all passing particles with $\mu/\mathcal{E} \in \mathfrak{N}$. The arguments in this subsection only apply to energetic particles, for which the electrostatic potential is negligible, but could be generalized to thermal particles if the electrostatic potential is close enough to a flux function that the dominant radial drift is the magnetic drift. For low-collisionality main ions in nearly omnigeneous stellarators ([Calvo et al., 2017](#)) and large-aspect-ratio stellarators ([d’Herbemont et al., 2022](#)), the electrostatic potential satisfies this condition in some, but not all, collisionality regimes,

⁷From a Poincaré plot, it can be difficult to say whether drift-island overlap is occurring close to the trapped–passing boundary and causing orbit stochasticity there. This is because, usually, the trapped–passing boundary is already surrounded by a region of stochastic orbits due to particles undergoing transitions between different trapping states.

as discussed in [Calvo *et al.* \(2018\)](#). Initially, we assume the magnetic field is a smooth, well-behaved function; we consider more general stellarators with discontinuous magnetic field strengths in §5.1.4.

5.1.1. Constraint on the minimum and maximum field strength

Suppose (5.1) is satisfied for passing particles with $\mu/\mathcal{E}$ in some open neighbourhood $\mathfrak{N}$ of zero. Then, we can Taylor expand the integrand in (5.1) to find

$$\sqrt{2\mathcal{E}} \sum_{n=1}^{\infty} c_n \left(\frac{\mu}{\mathcal{E}}\right)^n \partial_{\eta} \left(\oint_{\mathbf{r}} B^n d\mathbf{l} \right) = 0, \quad c_n = \begin{cases} \frac{1}{2^{2n}} \frac{(2n)!}{(1-2n)(n!)^2} & \text{if } n > 0, \\ 1 & \text{if } n = 0, \end{cases} \quad (5.2)$$

for values of $\mu/\mathcal{E} \in \mathfrak{N}$ that are within the radius of convergence. For these values of $\mu/\mathcal{E}$, the series in (5.2) is identically zero, so its coefficients must all vanish. This means

$$\partial_{\eta} \left(\oint_{\mathbf{r}} B^n d\mathbf{l} \right) = 0 \quad \text{for } n = 0, 1, 2, \dots \quad (5.3)$$

For example, when $n = 0$, (5.3) says that every field line on the rational surface is a closed curve with equal length. More generally, if $Y(B, \mathcal{E}, \mu, \sigma)$ is any analytic function of B , then (5.3) implies

$$\partial_{\eta} \left(\oint_{\mathbf{r}} Y(B(\psi_{\mathbf{r}}, \eta, l), \mathcal{E}, \mu, \sigma) d\mathbf{l} \right) = 0. \quad (5.4)$$

We can use (5.4), with a specific choice of Y , to derive a constraint on the minimum and maximum values of B along any field line on the rational surface. Let these values be $B_m(\eta)$ and $B_M(\eta)$, respectively, for a field line η . We will use (5.4) to prove, by contradiction, that $B_m(\eta)$ and $B_M(\eta)$ must be independent of η .

Starting with B_m , suppose $B_m(\eta_1) = B_1$ and $B_m(\eta_2) = B_2 \neq B_1$ for two field lines η_1 and η_2 . Without loss of generality, assume $B_1 < B_2$. Now, choose Y in (5.4) to be highly concentrated around B_1 and extremely small for $B \geq B_2$ (we need Y to be ‘extremely small’, rather than zero, because an analytic Y that vanishes for $B \geq B_2$ would have to be zero everywhere). This choice would make the integral in (5.4) much larger on field line η_1 than it is on field line η_2 , but (5.4) says these integrals must be equal. This contradiction proves that, in fact, $B_m(\eta_1) = B_m(\eta_2)$. Since η_1 and η_2 were arbitrary, we conclude that B_m must be independent of η . Identical reasoning can be used to show that B_M must also be independent of η .

It is well known that a similar constraint must be satisfied by the value of B at any local minimum or maximum of the field strength on an omnigenous flux surface ([Cary & Shasharina, 1997a,b](#); [Parra *et al.*, 2015](#)). The traditional proof of this result is quite different from the argument we gave above; it is worth explaining the reason for this difference. The precise statement of the constraint for omnigenous flux surfaces is as follows. If we use α as a field-line label and write B_{lm} (B_{lM}) for the value of B at a local minimum (maximum), then, on an omnigenous flux surface, we must have $\partial_{\alpha} B_{lm} = 0$ ($\partial_{\alpha} B_{lM} = 0$), for all α . The usual proof of $\partial_{\alpha} B_{lm} = 0$ considers the orbits of deeply trapped particles that only sample field-strength values close to B_{lm} in a single well ([Cary & Shasharina, 1997a,b](#); [Parra *et al.*, 2015](#)). We cannot use this approach to prove that $B_m(\eta)$ is independent of η because passing particles stream over multiple wells and never sample just the values of the field strength near B_m .

5.1.2. Constraint on the well widths

Next, we use (5.1) to obtain a constraint on the magnetic field involving a quantity we call the well width, $\Delta l_W(\eta, B)$. To define $\Delta l_W(\eta, B)$, consider a field line η and let B satisfy $B_m \leq B \leq B_M$. In general, as l varies, there will be multiple intervals within which $B(\psi_r, \eta, l) \leq B$. We refer to these intervals as different ‘wells’; they are the regions accessible to trapped particles with $\mu/\mathcal{E} = 1/B$. Then, $\Delta l_W(\eta, B)$ is the width of well W , treated as an interval in l (examples are shown in figure 10(b)).

The sum of the well widths of every well going around a rational-surface field line η will be denoted by $L(\eta, B)$:

$$L(\eta, B) := \sum_{W \in \mathcal{W}(\eta, B)} \Delta l_W(\eta, B), \quad (5.5)$$

where $\mathcal{W}(\eta, B)$ is the set of wells, for a given B , along this field line. We will show that, if (5.1) is satisfied for all passing particles with $\mu/\mathcal{E} \in \mathfrak{N}$, then $L(\eta, B)$ must be independent of η . To prove this statement, we begin by changing integration variables in

$$I_r = \oint_r \sqrt{2(\mathcal{E} - \mu B(\psi_r, \eta, l))} dl \quad (5.6)$$

from l to $B = B(\psi_r, \eta, l)$. Taking into account the multi-valued nature of this substitution, we find

$$I_r = \int_{B_m}^{B_M} \sqrt{2(\mathcal{E} - \mu B)} \sum_i \left| \frac{\partial l_i}{\partial B} \right| dB, \quad (5.7)$$

where $l_i(\eta, B)$ is the l coordinate of the i th point at which the field strength equals B along the field line (ψ_r, η) . The sign of $\partial l_i / \partial B$ alternates on opposite sides of each well, so $\sum_i |\partial l_i / \partial B| = \partial L / \partial B$. Thus, integrating by parts gives

$$I_r = \sqrt{2(\mathcal{E} - \mu B_M)} L_M + \int_{B_m}^{B_M} \frac{\mu L(\eta, B)}{\sqrt{2(\mathcal{E} - \mu B)}} dB. \quad (5.8)$$

The quantity $L_M := L(\eta, B_M)$ is the total length of the closed field line η . In §5.1.1, we showed that L_M , B_m , and B_M must all be independent of η . Therefore, the only η dependence in (5.8) is contained in $L(\eta, B)$, inside the integral. Since the integration limits are constants, we cannot apply the Abel-inversion technique of Cary & Shasharina (1997b) to prove that $\partial_\eta I_r = 0$ implies $\partial_\eta L(\eta, B) = 0$. Instead, as before, we use the fact that $\mu/\mathcal{E} < 1/B_M \leq 1/B$ to Taylor expand the integrand, finding

$$I_r = \sqrt{2(\mathcal{E} - \mu B_M)} L_M + \frac{\mu}{\sqrt{2\mathcal{E}}} \sum_{n=0}^{\infty} \frac{1}{2^{2n}} \frac{(2n)!}{(n!)^2} \left(\frac{\mu}{\mathcal{E}} \right)^n \int_{B_m}^{B_M} B^n L(\eta, B) dB. \quad (5.9)$$

For I_r to be independent of η for passing particles with $\mu/\mathcal{E} \in \mathfrak{N}$ and within the radius of convergence of the sum in (5.9), every coefficient of $(\mu/\mathcal{E})^n$ in the sum must be independent of η ; that is,

$$\partial_\eta \left(\int_{B_m}^{B_M} B^n L(\eta, B) dB \right) = 0 \quad \text{for } n = 0, 1, 2, \dots \quad (5.10)$$

A standard result in real analysis states that two continuous real functions on a closed interval with equal moments must be equal (e.g. Rudin, 1976, p.169). Applying this result to $L(\eta_1, B)$ and $L(\eta_2, B)$, for any two field lines η_1 and η_2 , we conclude that $L(\eta, B)$ must be independent of η .⁸ We refer to magnetic fields with this property as ‘cyclometric’.

⁸Note that this argument uses the fact that B_m and B_M are independent of η . Also note that

Cyclometry is a strong constraint, from which we can deduce two further statements about the magnetic field. First of all, (5.4) must hold, regardless of whether the integrand is analytic; that is,

$$\partial_\eta \left(\oint_{\mathbf{r}} Y(B(\psi_{\mathbf{r}}, \eta, l), \mathcal{E}, \mu, \sigma) d\mathbf{l} \right) = 0, \quad (5.11)$$

for any $Y(B, \mathcal{E}, \mu, \sigma)$. To show this, we change variables in (5.11) from l to $B(\psi_{\mathbf{r}}, \eta, l)$, integrate by parts, and use $\partial_\eta L(\eta, B) = 0$. Second, on the rational surface, there cannot be any tangential intersections between field lines and contours of constant B . Therefore, these contours must close toroidally, helically, or poloidally; the magnetic field must be ‘pseudosymmetric’ (Mikhailov *et al.*, 2002; Skovoroda, 2005). We provide a proof of this second statement in §C. As in §5.1.1, the proof is different from the proof of the analogous result for omnigenous stellarators. In omnigenous stellarators, the possibility of tangential intersections can be ruled out by considering particles that are trapped in wells near a hypothetical tangential intersection. This kind of argument cannot be used in the cyclometric case because passing particles stream over multiple wells and never sample just the values of the field strength near a tangential intersection. Note that the proof we give in §C uses the fact that the magnetic field is smooth. Cyclometric fields that are not pseudosymmetric are possible if the field strength is allowed to have discontinuities; we discuss this possibility further in §5.1.4.

5.1.3. Boozer-coordinate formulation

Using Boozer coordinates, it is possible to define cyclometry without making reference to arc lengths along field lines. This alternative formulation will be necessary when we discuss nearly cyclometric magnetic fields in §§5.2–5.3. Furthermore, it provides a convenient way of extending the definition of cyclometry to flux surfaces other than the rational surface $\psi_{\mathbf{r}}$, which will be important in §5.3.

The condition that cyclometric fields must satisfy can be written in Boozer coordinates as

$$\partial_\eta I = \partial_\eta \left(G \oint \frac{\sqrt{2(\mathcal{E} - \mu B)}}{B} d\zeta_{\mathbf{B}} \right) = 0 \quad \text{for passing particles with } \mu/\mathcal{E} \in \mathfrak{N}. \quad (5.12)$$

Here, we have used $I(\psi, \eta, \mathcal{E}, \mu)$ instead of $I_{\mathbf{r}}(\eta, \mathcal{E}, \mu)$; this will allow us to define cyclometry on a general flux surface. In writing (5.12), we also used the fact that $\hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x}$ becomes $[G_{\text{pol}} + (N/M)G_{\text{tor}}]/B = G/B$ in Boozer coordinates. Note that, while $I(\psi, \eta, \mathcal{E}, \mu)$ is defined on any flux surface, its definition involves an integration curve with rotational transform equal to a specified rational number, N/M . Cyclometry on a general flux surface ψ is only useful if N/M is approximately equal to the magnetic rotational transform, $\iota(\psi)$, on that surface.

Let $\Delta\zeta_{\mathbf{B},W}(\psi, \eta, B)$ be the cumulative change in the Boozer toroidal angle when we move, along the curve of constant $\eta = \theta_{\mathbf{B}} - (N/M)\zeta_{\mathbf{B}}$, between the endpoints of well W . We define wells on a general flux surface in a similar manner to wells on the rational surface, except as intervals along curves of constant η instead of along field lines. Let

$$Z_{\mathbf{B}}(\psi, \eta, B) := \sum_{W \in \mathcal{W}(\eta, B)} \Delta\zeta_{\mathbf{B},W}(\psi, \eta, B), \quad (5.13)$$

which is the Boozer-coordinate analogue of (5.5). Following exactly the same steps as in

we are careful to apply this argument to $L(\eta, B)$, which is a continuous function of B , rather than to $\partial_\eta L(\eta, B)$, which need not be a continuous function of B , as can be seen from (C6).

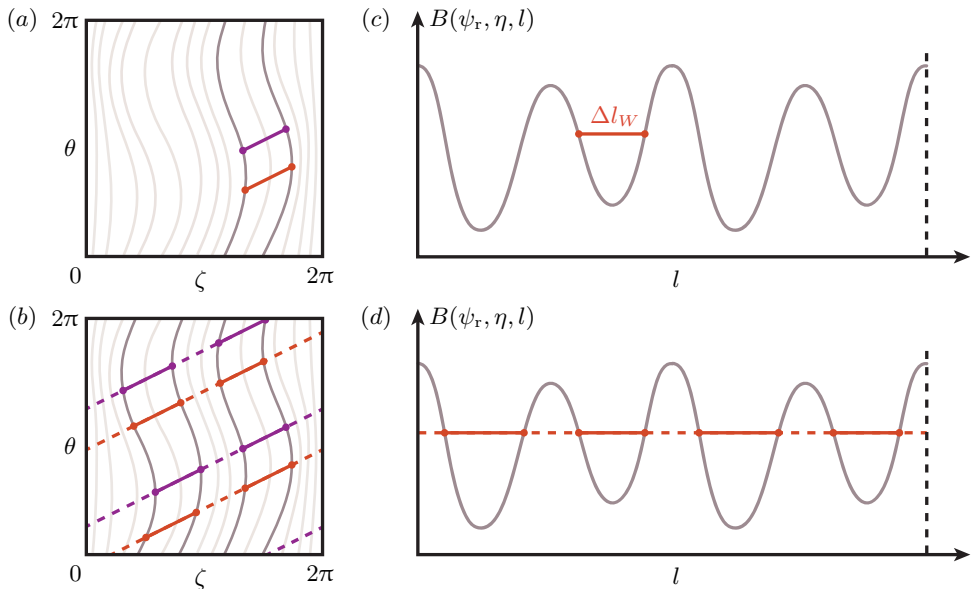


FIGURE 10. Difference between omnigenous (isometric) and cyclometric magnetic fields on a rational surface (with $N_{\text{fp}} = 1$). (a) An omnigenous magnetic field. The red and purple lines show two wells with the same bounce-point magnetic field strength, which must have equal widths. (b) Sketch of the field strength against distance along a field line in this omnigenous field, with a single well indicated. (c) A cyclometric magnetic field. The red and purple dashed lines represent two closed field lines. The solid line segments show the different wells along these field lines, for a given bounce-point field strength. The sum of the widths of these wells must be the same for each field line. (d) Sketch of the field strength as a function of distance along one of these field lines, with the well widths that must be summed to obtain $L(\eta, B)$ indicated.

§5.1, we can show that (5.12) is satisfied if and only if $Z_B(\psi, \eta, B)$ is independent of η , for all B . Thus, cyclometry is equivalent to the condition that, in Boozer coordinates, the sum of the toroidal-angle well widths is the same along each curve of constant η , for all B .

5.1.4. Discussion

We have shown that no drift islands exist at a given rational surface if and only if the magnetic field is cyclometric on that surface. Cyclometry is a weaker condition than omnigenicity; in an omnigenous field, the width of each well individually must be independent of the field-line label.⁹ Thus, perfectly omnigenous stellarators are automatically cyclometric on every rational surface and do not contain drift islands (Burby et al., 2023). The difference between omnigenous and cyclometric magnetic fields is illustrated in figure 10.

The fact that omnigenous fields are cyclometric on rational surfaces raises the question of whether cyclometric fields might possess other properties of omnigenous fields. For example, it is well known that omnigenous fields have a contour of maximum B that is a straight line in Boozer coordinates, and also that analytic, non-quasisymmetric omnigenous fields do not exist (Cary & Shasharina, 1997b). In fact, neither of these properties carry over to cyclometric fields because the proofs in the omnigenous case require the rotational transform to be irrational. To demonstrate this, figure 11 shows

⁹This property is sometimes called ‘isometry’ (Cary & Shasharina, 1997b; Skovoroda, 2005), which is the reason we chose the name ‘cyclometry’.

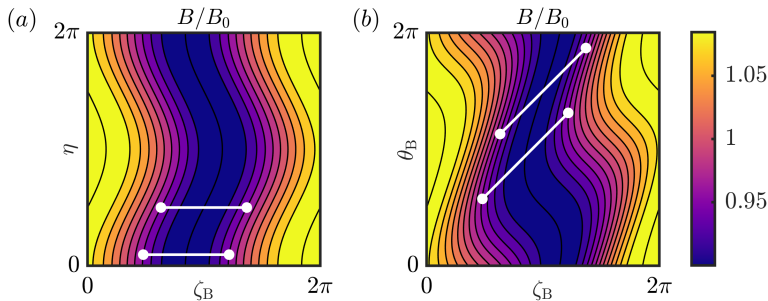


FIGURE 11. Example of how an analytic, omnigenous magnetic field strength may be constructed on a rational surface. (a) First, it is simplest to write down such a field as a function of ζ_B and $\eta = \theta_B - (N/M)\zeta_B$. In a contour plot of B in the (ζ_B, η) plane, magnetic field lines are horizontal lines. Thus, for the field strength to be omnigenous, the horizontal distance between opposing sides of each well must be independent of η . This can be achieved, for example, if the field strength is simply shifted horizontally by different amounts at each η . In the case shown, the field strength is $B(\eta, \zeta_B) = B_0[1 + 0.1 \cos(\zeta_B + 0.5 \cos \eta)]$, where B_0 is a normalization constant, and $M = 1$. (b) Contour plot of the same field strength in the conventional (ζ_B, θ_B) plane, obtained by changing the vertical-axis coordinate from η to $\theta_B = \eta + (N/M)\zeta_B$, which amounts to shearing the plot vertically. In this case, we took the rational surface to have $N = M = 1$.

how an analytic, omnigenous magnetic field strength, with a contour of maximum B that is not straight in Boozer coordinates, may be constructed when the rotational transform is rational.

Omnigenous magnetic fields are automatically cyclometric on rational surfaces but the converse is not true. That is, it is possible for there to be no drift islands around a highly non-omnigenous rational surface. The reason is that the field may be cyclometric despite the fact that individual wells widen or shrink as η varies; an example is shown in [figure 12](#).

Interestingly, there is a class of magnetic fields for which cyclometry and omnigenicity *are* equivalent. This is the case if, for every B , there is only one (non-identical) well along each rational-surface field line; then, $L(\eta, B)$ is a multiple of the width $\Delta l(\eta, B)$ of this single well. This can occur, for example, if the field-strength contours wrap around the torus toroidally and the rotational transform is $\iota = 1/M$, for some $M \in \mathbb{Z}^{>0}$ (see [figure 13\(a\)](#)). It can also occur if the field-strength contours wrap around the torus poloidally, there is a single well in every field period, and ι is a multiple of N_{fp} (see [figure 13\(b\)](#)). This second scenario is probably of academic interest only because the required rotational transform is very high. For magnetic fields such as these, cyclometry forces $\Delta l(\eta, B)$ to be independent of η for all B , and hence cyclometry implies omnigenicity on ψ_r .

To complete our discussion of cyclometric magnetic fields, it is worth pointing out that some of the constraints derived in this subsection may be relaxed if we permit the magnetic field strength to be a discontinuous function. Recently, by allowing the field strength to be discontinuous, more general omnigenous magnetic field strengths were discovered ([Velasco et al., 2024](#); [Calvo et al., 2025](#); [Velasco et al., 2025](#)). These *piecewise omnigenous* field strengths are automatically cyclometric on rational surfaces because their values of Z_B (see (5.13)) are independent of η on ψ_r . Similarly, it is possible to construct more general cyclometric field strengths that are discontinuous; we refer to these fields as *piecewise cyclometric*. The idea behind this construction is easiest to understand with a picture such as [figure 14](#). This figure shows a piecewise-cyclometric magnetic field strength

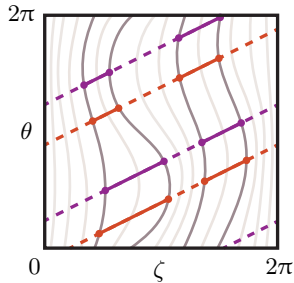


FIGURE 12. Cartoon depiction of a magnetic field strength (with $N_{fp} = 1$) that is cyclometric but not isometric. The sum of the well widths along the red and purple field lines is the same, even though individual wells change size.

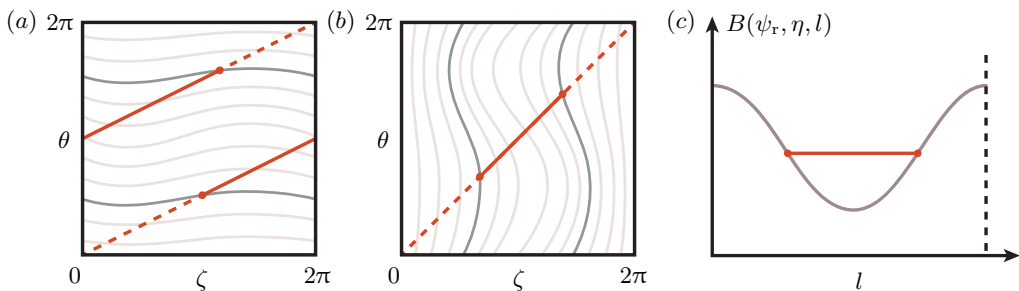


FIGURE 13. Cartoon showing two magnetic field strengths (with $N_{fp} = 1$) in straight-field-line coordinates for which cyclometry implies omnigenity. (a) A case with $\iota = 1/2$. (b) A case with $\iota = 1$. (c) Sketch of how the field strength varies, in both cases, as a function of distance along a closed field line.

that assumes only two values, B_m and B_M . In [figure 14\(a\)](#), the field strength is plotted against ζ_B and η . In these coordinates, the constant- η curves are straight horizontal lines; measured along these lines, the widths $\Delta\zeta_{B,W}$ of the purple regions change as η varies. Thus, on ψ_r , a particle trapped in a single well (within which the field strength is B_m) would experience a changing well width as η varies, so this field strength is not omnigenous. Crucially, when the width of one purple region increases, the corresponding width of the other purple region along the same horizontal line decreases, so that sum of these widths, Z_B , is constant. Therefore, this field strength is piecewise cyclometric. In [figure 14\(b\)](#), the same field strength is replotted against ζ_B and θ_B , which amounts to shearing the plot in [figure 14\(a\)](#) vertically. Since there are closed constant- B contours that do not encircle the torus toroidally, helically, or poloidally, this figure demonstrates that piecewise-cyclometric fields do not need to be pseudosymmetric.

The results of this subsection may be useful for stellarator optimization. In a stellarator containing a low-order rational surface, the size of any drift islands could be reduced by optimizing the magnetic field for cyclometry on that surface. Since cyclometry is a property of the magnetic field only, this may be more convenient than optimizing for $\partial_\eta I_r = 0$ at a number of different pitch angles. Of course, the drift islands could also be removed by optimizing for better omnigenity around the rational surface; however, the fact that cyclometry can be achieved even when there are relatively large deviations from omnigenity means cyclometry is a less stringent optimization target that might be more compatible with other objective functions.

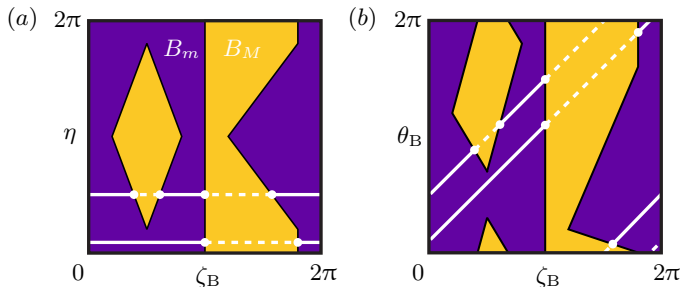


FIGURE 14. Plots of the magnetic field strength on a flux surface, in Boozer coordinates, for a piecewise-cyclometric field. The field strength is B_m in the purple region and B_M in the orange region. (a) Plot in (η, ζ_B) coordinates, where $\eta = \theta_B - (N/M)\zeta_B$. The white lines are two curves of constant η on the flux surface. (b) Plot of the same field strength in conventional (θ_B, ζ_B) coordinates, assuming $N = 1$ and $M = 2$. The same constant- η curves are shown in white.

5.2. Nearly cyclometric stellarators

In real stellarators, small deviations from perfect cyclometry are inevitable. In this subsection, we investigate how the size of these deviations affects the drift-island width. For this discussion, we assume an order-unity aspect ratio (i.e. $\epsilon \sim 1$).

For any flux surface ψ close to ψ_r , we will refer to a magnetic field as ‘nearly cyclometric on ψ ’ if it is possible to split

$$B(\psi, \theta_B, \zeta_B) = B_0(\psi, \theta_B, \zeta_B) + B_1(\psi, \theta_B, \zeta_B), \quad (5.14)$$

where B_0 is a cyclometric field strength and $B_1 \sim \delta B_0$, with $\delta \ll 1$, is a small perturbation. That is, B_0 is a function of the Boozer angles for which Z_B in (5.13) is independent of η , for all B .

In this subsection, we only consider magnetic fields that are nearly cyclometric on the rational surface ψ_r . Expanding (4.9) for such a field, using (5.14), we find

$$I_r = G_r \oint_r \frac{\sqrt{2(\mathcal{E} - \mu B_0)}}{B_0} \left(1 - \frac{2\mathcal{E} - \mu B_0}{2(\mathcal{E} - \mu B_0)} \frac{B_1}{B_0} + O(\delta^2) \right) d\zeta_B. \quad (5.15)$$

Since B_0 is cyclometric, (5.12) is satisfied for this magnetic field strength. This tells us that the leading term in (5.15) is independent of η , so $I_r(\eta, \mathcal{E}, \mu) - I_r(\eta_r, \mathcal{E}, \mu) \sim \delta v L$. Therefore, the drift-island width in a nearly cyclometric stellarator scales as¹⁰

$$\frac{\Delta\psi}{\Psi_t} \sim \left(\frac{\rho_\star \delta}{s} \right)^{1/2}. \quad (5.16)$$

This result shows that optimizing a stellarator for omnigenity should reduce the width of any drift islands, unless this optimization also forces the magnetic shear to be very

¹⁰One might think that (5.16) requires the radial drift due to B_1 to be smaller, pointwise, than the radial drift due to B_0 , by a factor of δ . In fact, the argument leading to (5.16) is valid even when the radial drift due to B_1 is of the same order as the drift due to B_0 ; this occurs when B_1 varies on short lengthscales $\sim \delta L$, so that its gradient is $\nabla B_1 \sim \nabla B_0$. The fact that (5.16) is valid even when the perturbation varies on lengthscales $\sim \delta L$ means it is possible to use (5.16) to derive the drift-island width (4.12) for large-aspect-ratio stellarators. To do this, we set $L \sim R$ and consider B_1 in (4.10) as a deviation from omnigenity of order ϵ varying on lengthscales $a \sim \epsilon L$ (we must also remember that $\rho_\star$ in the large-aspect-ratio case is defined by $\rho_\star = \rho/a$).

low. In the next subsection, we show that, when the shear is low, scaling (5.16) must be modified.

Scaling (5.16) agrees with the estimate obtained in §2.1. In §2.2, we used an analogy with magnetic islands to argue that the drift-island width scales with $(\rho_*/s)^{1/2}$, but we did not obtain the factor of $\sqrt{\delta}$ which reduces the drift-island width in nearly cyclometric stellarators. Now we are in a position to explain how the argument of §2.2 could be extended to obtain this factor. The key point is that, in a perfectly cyclometric field, the field $\mathbf{B}^*$ that the guiding centres follow is integrable (up to order ρ_*^2 terms, which are beyond the level of precision of our calculations so far). Therefore, the size of the perturbation that should be used in the formula for magnetic-island width is not the $\sim \rho_*$ drift term in $\mathbf{B}^*$ (see (2.6)), but rather the $\sim \rho_*\delta$ drift due to the non-cyclometric piece of the magnetic field.¹¹ With this modification, the magnetic-island width formula also predicts scaling (5.16).

5.3. Nearly cyclometric stellarators with low shear

Stellarators that have been optimized to be highly quasisymmetric or quasi-isodynamic—and, therefore, highly cyclometric—often have very low magnetic shear (e.g. Landreman & Paul, 2022; Goodman et al., 2023; Nies et al., 2024). This may be concerning in light of scaling (5.16), which suggests that, even if δ is small, a low enough shear will produce large drift islands. Fortunately, it turns out that (5.16) does not apply when the shear is sufficiently low. In this subsection, we show that, as the shear is reduced, the drift-island width saturates at $\Delta\psi \sim \sqrt{\delta}\Psi_t$ once $s \lesssim \rho_*$. We demonstrate this by deriving an equation for the drift-island shape in a nearly cyclometric stellarator with $s \sim \rho_*$.

Importantly, when $s \sim \rho_*$, the drift islands shift radially away from their associated rational surface. This radial shift has been observed in simulations of alpha particles in stellarators (e.g. White et al., 2022; White, 2022; Chambliss et al., 2025) and, in some cases, can cause the drift islands to move all the way into or out of the plasma volume (White et al., 2022; White, 2022). The same radial shift is known to affect runaway-electron and energetic ion orbits in perturbed tokamaks (e.g. de Rover et al., 1996; Motojima & Shishkin, 1999; Heyn et al., 2012; Matsuyama et al., 2014; Shinohara et al., 2020; He et al., 2020; Liu et al., 2021), as well as thermal-ion orbits in the presence of small-scale neoclassical tearing modes (Imada et al., 2018; Dudkovskaia et al., 2023). Physically, the radial shift is caused by the tangential magnetic drift, for the following reason. Drift islands form around closed orbits, which occur when the rotational transform of the passing orbits—the *kinetic rotational transform*—is rational. Tangential drifts cause the kinetic rotational transform $\bar{\iota}$ to deviate from the magnetic rotational transform by an amount $\bar{\iota} - \iota \sim \rho_*$ (the precise definition that we use for $\bar{\iota}$ is not needed for the present discussion, but it can be found in (7.29)). As a result, the flux surface ψ_s on which the kinetic rotational transform assumes a given rational value is shifted, with respect to the corresponding rational surface ψ_r , by $\psi_s - \psi_r \sim (\rho_*/s)\Psi_t$, as shown in figure 15. This estimate for the radial shift will be confirmed by the calculations below. Note that the shift becomes order unity when $s \sim \rho_*$.

¹¹ For this order- $\rho_*\delta$ term to be more important than order- ρ_*^2 corrections that we neglect, we need $\delta \gg \rho_*$. In fact, all our estimates and calculations up to this point require $\delta \gg \rho_*$; this condition ensures the particle drifts a radial distance of $\Delta\psi \sim (\rho_*\delta)\Psi_t$ in a single transit, as we assumed for our estimates in §2.1. Once $\delta \lesssim \rho_*$, the stellarator is close enough to perfect cyclometry that the radial distance drifted in a single transit is affected by higher-order guiding-centre drifts, which we do not consider in this article. Thus, for particles with normalized gyroradius ρ_* , optimization for cyclometry beyond $\delta \sim \rho_*$ does not reduce the orbit width any further.

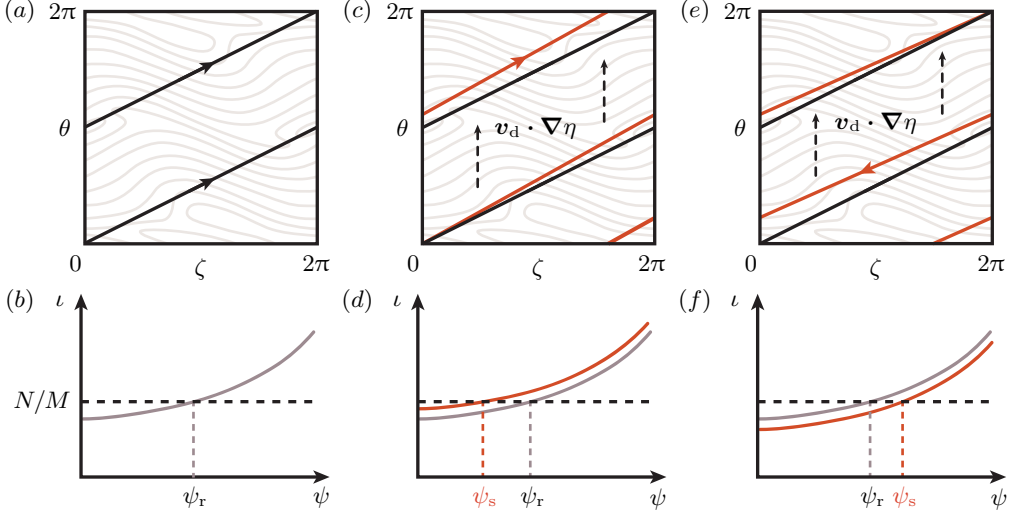


FIGURE 15. Cartoon illustrating how the tangential magnetic drift causes the drift islands to shift radially. (a) A magnetic field line on a rational surface is plotted in straight-field-line coordinates. Grey contours represent the magnetic field strength. (b) The rotational-transform profile for this magnetic field, with the rational surface ψ_r indicated. (c) Now consider a passing particle that streams along this field line, while experiencing a tangential drift in the poloidal direction, as shown. The resulting orbit is depicted in red. The slope of this line is the kinetic rotational transform, which, in this case, is larger than ι . (d) The red curve shows the profile of the kinetic rotational transform on each flux surface for this particle. The drift islands will appear on the surface ψ_s , where the kinetic rotational transform is rational. Thus, the islands shift away from ψ_r by an amount that scales with the size of the tangential drift and scales inversely with the shear. (e) A passing particle streaming in the opposite direction to the particle in (c) and (d) still experiences the same tangential drift as that particle, so the kinetic rotational transform is now lower than ι . (f) This will cause the drift islands to shift radially in the opposite direction.

To study particle orbits in a device with $s \sim \rho_*$, we can still use the transit-invariant formula (3.13). The derivation of (3.13) in §3 remains valid because η and ψ still vary slowly compared with the fast angle ζ . However, the simplified island-shape equation (3.22) is no longer correct since it assumes that the drift islands are close to ψ_r .

We first determine the shifted location of the drift islands, ψ_s . The closed particle orbits at the centre (O-point) of the drift islands must have $\langle d\eta/dt \rangle_t = 0$. In nearly cyclometric stellarators, equation (3.12b) for $\langle d\eta/dt \rangle_t$, which includes tangential drifts, reads

$$\left\langle \frac{d\eta}{dt} \right\rangle_t = \frac{2\pi M \sigma}{\tau_t} \left(\iota(\psi) - \frac{N}{M} \right) - \frac{mc}{Ze\tau_t} \partial_\psi I_0 + O\left(\frac{\rho_* \delta v}{L} \right). \quad (5.17)$$

Here, $I_0(\psi, \mathcal{E}, \mu)$ is the leading-order term when $I(\psi, \eta, \mathcal{E}, \mu)$, defined in (3.10), is expanded for a nearly cyclometric magnetic field: $I = I_0 + I_1 + O(\delta^2 v L)$, where

$$I_0 := G \oint \frac{\sqrt{2(\mathcal{E} - \mu B_0)}}{B_0} d\zeta_B, \quad (5.18a)$$

$$I_1 := -G \oint \frac{2\mathcal{E} - \mu B_0}{B_0 \sqrt{2(\mathcal{E} - \mu B_0)}} \frac{B_1}{B_0} d\zeta_B. \quad (5.18b)$$

The stellarator needs to be nearly cyclometric to ensure that, despite the low shear, the drift-island width is smaller than the minor radius. In §5.2, we considered magnetic

fields that are nearly cyclometric on the rational surface ψ_r . Now that the drift islands can shift an order-unity distance away from ψ_r , the magnetic field strength on ψ_r is no longer relevant; instead, we need to assume that the magnetic field is nearly cyclometric in a region containing ψ_s . The reader might wonder whether, in a device with $s \sim \rho_*$, cyclometry is actually a stronger condition than necessary to avoid drift islands at ψ_s . After all, we derived cyclometry in §5.1 by requiring that the transit-averaged radial drift vanishes on ψ_r for *all* passing particles; but, when $s \sim \rho_*$, the transit-averaged radial drift only needs to vanish on ψ_s for those passing particles *whose drift islands arise on* ψ_s . Nevertheless, it turns out that cyclometry is still necessary; in fact, to remove drift islands from a stellarator with $s \sim \rho_*$, every flux surface that the islands can shift to needs to be cyclometric. We prove this claim in §D.

Resuming our calculation of the shifted island position, we find that ψ_s satisfies

$$2\pi M\sigma \left(\iota(\psi_s) - \frac{N}{M} \right) - \frac{mc}{Ze} (\partial_\psi I_0)_s = 0, \quad (5.19)$$

since (5.17) must vanish at ψ_s . Here, the subscript ‘s’ indicates quantities evaluated at ψ_s . Note that (5.19) implies $\psi_s - \psi_r \sim (\rho_*/s)\Psi_t$, confirming our estimate above. Also, note that ψ_s depends on $\mathcal{E}$ and μ , which means the drift islands appear at different radial positions for particles with different speeds and pitch angles. Furthermore, the drift islands shift radially in opposite directions for co- and counter-passing particles; this can be understood from figure 15. Therefore, in a stellarator whose rotational transform profile has been designed to pass through a low-order rational at the plasma edge to allow an island divertor, drift islands will shift into the plasma for either co- or counter-passing alpha particles. These drift islands may be wide because the low-order rational surface at the edge typically has $N = N_{fp}$, which, as explained in §4.1, promotes large drift islands.

To obtain an equation for the drift-island shape, we expand $\mathcal{I}$ for particles that satisfy $\psi - \psi_s \sim \sqrt{\delta}\Psi_t$. We find

$$\begin{aligned} \mathcal{I}(\psi, \eta) = & \sigma \left(I_{0s} + I_{1s}(\eta) + \frac{1}{2} (\partial_\psi^2 I_0)_s (\psi - \psi_s)^2 \right) \\ & - \frac{2\pi MZe}{mc} \int_{\psi_r}^{\psi_s} \left(\iota(\psi') - \frac{N}{M} \right) d\psi' - \frac{\pi MZe}{mc} \iota'_s (\psi - \psi_s)^2 + O(\delta^{3/2}vL), \end{aligned} \quad (5.20)$$

where we have suppressed any velocity-space arguments. The level sets of $\mathcal{I}(\psi, \eta)$ describe the drift-island shape in a low-shear device; instead of (4.1), which was valid when $s \gg \rho_*$, we now have

$$\psi - \psi_s = \pm \sqrt{\frac{mc\sigma [I_{1s}(\eta, \mathcal{E}, \mu) - I_{1s}(\eta_s, \mathcal{E}, \mu)]}{\pi MZe [\iota'_s - (mc\sigma/2\pi MZe)(\partial_\psi^2 I_0)_s]}}, \quad (5.21)$$

for particles that cross ψ_s at $\eta = \eta_s$. This equation confirms that our initial estimate $\psi - \psi_s \sim \sqrt{\delta}\Psi_t$ was correct.

From (5.21), we learn that there are two important modifications to the drift-island shape when the shear is small. First, the radial shift means that all functions of position are evaluated at ψ_s instead of ψ_r . Second, the magnetic shear ι'_r has been replaced with $\iota'_s - (mc\sigma/2\pi MZe)(\partial_\psi^2 I_0)_s$, which is the shear in the kinetic rotational transform of the particle orbits. The new term $(mc/2\pi MZe)(\partial_\psi^2 I_0)_s$ is the radial derivative of the (transit-averaged) tangential drift rate, which is an order- ρ_* correction. Even when the magnetic shear becomes vanishingly small, this term will allow particles to move tangentially across ψ_s and will set a finite width for the drift islands. Indeed, taking the subsidiary limit

$s \ll \rho_*$ in (5.21), we see that the island width $\Delta\psi$ does not increase beyond $\sim \sqrt{\delta}\Psi_t$, as claimed at the start of this subsection. Once $s \ll \rho_*$, the orbits essentially follow level sets of $I(\psi, \eta, \mathcal{E}, \mu)$, which is the adiabatic invariant discussed by [Hastie et al. \(1967\)](#) for passing particles in magnetic fields with closed field lines (see also [Elbarmi et al., 2020](#)).¹² Interestingly, in this limit, there is no longer any difference between the drift islands for co- and counter-passing particles. In a stellarator with $s \ll \rho_*$, it is possible to describe all passing orbits using an adiabatic invariant because, as explained in §4.3, stochastic orbits due to drift-island overlap do not occur at such low values of the magnetic shear.

6. Higher-order transit adiabatic invariant

In §3, we used the transit adiabatic invariant to determine the drift-island shape. Our results agreed well with simulations at lower energies but began to deviate at higher energies. This is unsurprising because we expanded in $\rho_* \ll 1$, which means we neglected higher-order corrections that become more important at larger energies. In this section, we derive a more accurate equation for the drift-island shape by calculating the first-order correction to the transit invariant. We use a method similar to the one described in appendix B of [Cary et al. \(1986\)](#). The resulting, improved drift-island shape agrees very well with simulations, even for alpha particles at 3.5 MeV.

6.1. Higher-order adiabatic invariant

The calculation in §3 used two small parameters: $\rho_* \ll 1$ and $|\iota - N/M| \ll 1$. The higher-order calculation is more straightforward if there is only one expansion parameter. Since we are interested in the particles that move around an island of width $\sim (\rho_* \delta/s)^{1/2} \Psi_t$, in this section we order $s \sim \delta \sim 1$ and $\psi - \psi_r \sim \sqrt{\rho_*} \Psi_t$.¹³

When this ordering is satisfied, the results of §3 can be summarized as follows. Both η and ψ evolve slowly, at rates $d\eta/dt \sim \sqrt{\rho_*} v/L$ and $d\psi/dt \sim \rho_* \Psi_t v/L$. The transit invariant $\mathcal{I}$ is also slowly varying (viz. $d\mathcal{I}/dt \sim \sqrt{\rho_*} v^2$) because it depends only on the slow coordinates η and ψ . The special importance of $\mathcal{I}$ lies in the fact that changes in $\mathcal{I}$ do not accumulate over long times: the transit average $\langle d\mathcal{I}/dt \rangle_t \sim \rho_* v^2$ is an order (in $\sqrt{\rho_*}$) smaller than the instantaneous value of $d\mathcal{I}/dt$. According to the general theory of adiabatic invariants ([Lenard, 1959](#); [Kruskal, 1962](#); [Henrard, 1993](#)), it should be possible to construct an asymptotic series

$$\bar{\mathcal{I}}(\psi, \eta, \zeta, \mathcal{E}, \mu, \sigma) = \mathcal{I}(\psi, \eta, \mathcal{E}, \mu, \sigma) + \mathcal{I}^{(1/2)}(\psi, \eta, \zeta, \mathcal{E}, \mu, \sigma) + \dots, \quad (6.1)$$

with $\mathcal{I}^{(n/2)} \sim \rho_*^{n/2} vL$, that is conserved to all orders in $\sqrt{\rho_*}$. Any finite truncation of this series, up to and including the term $\mathcal{I}^{(n/2)}$, gives a quantity $\bar{\mathcal{I}}$ that has a small time derivative, $d\bar{\mathcal{I}}/dt \sim \rho_*^{(n+1)/2} v^2$, but an even smaller transit-averaged time derivative, $\langle d\bar{\mathcal{I}}/dt \rangle_t \sim \rho_*^{(n+2)/2} v^2$. Therefore, knowledge of the higher-order terms in (6.1) allows us to more precisely constrain the position of the particle over long times.

In this section, we calculate $\mathcal{I}^{(1/2)}$. The resulting higher-order transit invariant changes by less than $\sim \rho_* vL$ over one transit and will therefore capture order- ρ_* effects, such as

¹²A similar result is known to hold for one-and-a-half dimensional Hamiltonian systems whose unperturbed Hamiltonian $\mathcal{H}_0(J)$ (J is the action coordinate) has $\Omega(J) := \partial \mathcal{H}_0 / \partial J = 0$, which is the appropriate generalization of the low-magnetic-shear condition (see [Arnold et al., 2006](#), the remark after Theorem 17, p.187).

¹³Ordering $\delta \sim 1$ allows us to study both non-omnigeneous and omnigeneous stellarators, since we may take the subsidiary limit $\delta \ll 1$ to describe an optimized device.

tangential drifts (which cause the drift islands to shift radially, as discussed in §5.3) and finite radial orbit widths.

6.2. Guiding-centre equations

For this derivation, we must ensure that our guiding-centre equations of motion are correct to a high enough order in ρ_* . When $\psi - \psi_r \sim \sqrt{\rho_*} \Psi_t$, the guiding-centre equations (3.2), including error estimates, read

$$\frac{d\psi}{dt} = \mathbf{v}_d \cdot \nabla \psi + O\left(\rho_*^2 \frac{v \Psi_t}{L}\right), \quad (6.2a)$$

$$\frac{d\eta}{dt} = v_{\parallel} \left(\iota(\psi) - \frac{N}{M} \right) \hat{\mathbf{b}} \cdot \nabla \zeta + \mathbf{v}_d \cdot \nabla \eta + O\left(\rho_*^{3/2} \frac{v}{L}\right), \quad (6.2b)$$

$$\frac{d\zeta}{dt} = v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta + O\left(\rho_* \frac{v}{L}\right). \quad (6.2c)$$

The order- ρ_*^2 error in (6.2a) is due to higher-order guiding-centre drifts. The order- $\rho_*^{3/2}$ error in (6.2b) is due to parallel drifts projected in the $\nabla \eta$ direction, while the order- ρ_* error in (6.2c) is due to the magnetic drift in the $\nabla \zeta$ direction.

To find $\mathcal{I}^{(1/2)}$, we will need the equations of motion for ψ , η , and ζ to one order (in $\sqrt{\rho_*}$) beyond leading order. Thus, the guiding-centre equations (6.2) are sufficiently accurate for our purposes. In §E, we explain why we do not need to consider variations in the magnetic moment, despite the fact that its rate-of-change, $d\mu/dt \sim \rho_* v^3 / BL$, is as large as, for example, the higher-order corrections to $d\eta/dt$ that we must retain in (6.2).

6.3. Calculation of higher-order invariant

We proceed to impose $d\bar{\mathcal{I}}/dt = 0$ order-by-order on (6.1). At lowest order, we find

$$\partial_{\psi} \mathcal{I} \frac{d\psi}{dt} + \partial_{\eta} \mathcal{I} \frac{d\eta}{dt} + \partial_{\zeta} \mathcal{I}^{(1/2)} \frac{d\zeta}{dt} = 0 \quad (6.3)$$

because the time derivatives of the coordinates satisfy $d\psi/dt \sim \rho_* \Psi_t v / L$, $d\eta/dt \sim \sqrt{\rho_*} v / L$ and $d\zeta/dt \sim v / L$, and $\mathcal{I}$ has a large radial gradient: $\partial_{\psi} \mathcal{I} \sim (1/\sqrt{\rho_*}) v L / \Psi_t$. Let

$$\Gamma_{\psi} := \frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta}, \quad \Gamma_{\eta} := \iota(\psi) - \frac{N}{M} + \frac{\mathbf{v}_d \cdot \nabla \eta}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta}. \quad (6.4)$$

Integrating (6.3) to find $\mathcal{I}^{(1/2)}$ gives

$$\mathcal{I}^{(1/2)}(\psi, \eta, \zeta) = - \left(\partial_{\psi} \mathcal{I} \int_0^{\zeta} \Gamma_{\psi} d\zeta' + \partial_{\eta} \mathcal{I} \int_0^{\zeta} \Gamma_{\eta} d\zeta' \right) + \mathcal{I}^{(1/2)}(\psi, \eta, 0). \quad (6.5)$$

We no longer write out the velocity-space arguments $(\mathcal{E}, \mu, \sigma)$. In (6.5), we have found the dependence of $\mathcal{I}^{(1/2)}$ on ζ . However, we still need to determine the constant of integration $\mathcal{I}^{(1/2)}(\psi, \eta, 0)$, which depends on the slow coordinates only. Its value is fixed by a solvability condition derived from the next-order analogue of (6.3):

$$\left(\partial_{\psi} \mathcal{I} + \partial_{\psi} \mathcal{I}^{(1/2)} \right) \frac{d\psi}{dt} + \left(\partial_{\eta} \mathcal{I} + \partial_{\eta} \mathcal{I}^{(1/2)} \right) \frac{d\eta}{dt} + \left(\partial_{\zeta} \mathcal{I}^{(1/2)} + \partial_{\zeta} \mathcal{I}^{(1)} \right) \frac{d\zeta}{dt} = 0. \quad (6.6)$$

We can eliminate the unknown quantity $\mathcal{I}^{(1)}$ from this equation by dividing by $d\zeta/dt$ and integrating over ζ :

$$\partial_{\psi} \mathcal{I} \oint \Gamma_{\psi} d\zeta + \partial_{\eta} \mathcal{I} \oint \Gamma_{\eta} d\zeta + \oint \left(\Gamma_{\psi} \partial_{\psi} \mathcal{I}^{(1/2)} + \Gamma_{\eta} \partial_{\eta} \mathcal{I}^{(1/2)} \right) d\zeta = 0. \quad (6.7)$$

We found, back in (3.12), that

$$\oint \Gamma_\psi d\zeta = \frac{mc}{Ze} \partial_\eta \mathcal{I}, \quad \oint \Gamma_\eta d\zeta = -\frac{mc}{Ze} \partial_\psi \mathcal{I}. \quad (6.8)$$

Therefore, the first two terms in (6.7) cancel. This cancellation had to occur at order $\sqrt{\rho_\star} v^2$, as in (6.3), since $\mathcal{I}$ is an adiabatic invariant to lowest order. The cancellation also occurs at order $\rho_\star v^2$ because the transit-averaged identities (3.9a)–(3.9b) are satisfied exactly, as we pointed out in §3.2. Therefore, (6.7) becomes

$$0 = \oint \left(\Gamma_\psi \partial_\psi \mathcal{I}^{(1/2)} + \Gamma_\eta \partial_\eta \mathcal{I}^{(1/2)} \right) d\zeta. \quad (6.9)$$

In §F, we show that, after substituting (6.5) into (6.9), the resulting equation can be manipulated into the Poisson-bracket form

$$\partial_\psi F \partial_\eta \mathcal{I} - \partial_\eta F \partial_\psi \mathcal{I} = 0, \quad (6.10)$$

where

$$F(\psi, \eta) := \mathcal{I}^{(1/2)}(\psi, \eta, 0) - \frac{Ze}{mc} \iota'_r(\psi - \psi_r) \oint (\zeta' - \pi M) \Gamma_\psi d\zeta' + O(\rho_\star v L). \quad (6.11)$$

Equation (6.10) implies that $F(\psi, \eta)$ depends on ψ and η only through $\mathcal{I}$. Thus, we can write $F(\psi, \eta) = f(\mathcal{I}(\psi, \eta))$, for some function f . Substituting this result into (6.5) gives

$$\mathcal{I}^{(1/2)}(\psi, \eta, \zeta) = \frac{Ze}{mc} \iota'_r(\psi - \psi_r) \left(\int_0^\zeta 2\pi M \Gamma_\psi d\zeta' + \oint (\zeta' - \zeta - \pi M) \Gamma_\psi d\zeta' \right) + f(\mathcal{I}). \quad (6.12)$$

In (6.12), we have determined $\mathcal{I}^{(1/2)}$ up to a function of the lowest-order adiabatic invariant. We have complete freedom to pick $f(\mathcal{I})$ —we explain why in §G—so we choose $f = 0$ for simplicity. Furthermore, we make a negligible error by evaluating the line integrals in (6.12) at the rational surface ψ_r , which, as above, we represent using a subscript ‘r’ next to the integral sign. Overall, the higher-order transit adiabatic invariant is

$$\begin{aligned} \bar{\mathcal{I}} = \sigma I - \frac{2\pi M Ze}{mc} \int_{\psi_r}^\psi \left(\iota(\psi') - \frac{N}{M} \right) d\psi' + \frac{Ze}{mc} \iota'_r(\psi - \psi_r) \oint_r (\zeta' - \zeta - \pi M) \Gamma_\psi d\zeta' \\ + \frac{Ze}{mc} \iota'_r(\psi - \psi_r) \int_{0,r}^\zeta 2\pi M \Gamma_\psi d\zeta', \end{aligned} \quad (6.13)$$

where

$$\Gamma_\psi = \frac{\mathbf{v}_d \cdot \nabla \psi}{v_\parallel \hat{\mathbf{b}} \cdot \nabla \zeta} = \frac{mc}{Ze} [\partial_\eta (v_\parallel \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x}) - \partial_\zeta (v_\parallel \hat{\mathbf{b}} \cdot \partial_\eta \mathbf{x})]. \quad (6.14)$$

Equation (6.13) is the main result of this section. Before considering the implications of (6.13) for the drift-island shape, we first discuss certain consistency checks that (6.13) must pass in order for the higher-order transit invariant to be well-defined.

The first requirement is that $\bar{\mathcal{I}}$ must be a continuous function of position on each flux surface. As written, (6.13) does not manifestly possess this property. The reason is that the higher-order correction terms in (6.13) contain the toroidal angle explicitly, but the definitions we adopted for our coordinates restricted $\eta \in [0, 2\pi/M)$ and $\zeta \in [0, 2\pi M)$, so there are adjacent points on any given flux surface which have very different values of ζ . For example, suppose we pick a point and increase its poloidal angle until the point crosses the curve $\eta = 2\pi/M$. Then, the η coordinate of this point will jump discontinuously from

$\eta = 2\pi/M$ to $\eta = 0$, while its ζ coordinate will change discontinuously by a multiple of 2π . It is not necessarily clear that the value of $\bar{\mathcal{I}}$ is unchanged by these discontinuous jumps.

In fact, (6.13) can be rewritten in a form which explicitly shows that $\bar{\mathcal{I}}$ is a continuous function of position. To achieve this, we change the dummy integration variables in the higher-order correction terms in (6.13) to

$$\xi := \begin{cases} \zeta' - \zeta & \text{if } \zeta \leq \zeta' \leq 2\pi M \\ 2\pi M + \zeta' - \zeta & \text{if } 0 \leq \zeta' \leq \zeta, \end{cases} \quad (6.15)$$

which removes any reference to an arbitrary origin for the toroidal angle coordinate. After this change of variables, (6.13) becomes

$$\bar{\mathcal{I}} = \sigma I - \frac{2\pi M Ze}{mc} \int_{\psi_r}^{\psi} \left(\iota(\psi') - \frac{N}{M} \right) d\psi' + \frac{Ze}{mc} \iota'_r(\psi - \psi_r) \oint_r (\xi - \pi M) \Gamma_\psi d\xi. \quad (6.16)$$

The line integral in the last term of (6.16) starts at $\xi = 0$ (the particle position) instead of at the arbitrary origin $\zeta = 0$, and extends along the curve of constant η for a full transit. This geometric prescription for the integration curve, together with the fact that the integrand depends on ξ and η in a continuous manner, ensures that this line integral does not change discontinuously anywhere on the flux surface. Therefore, the higher-order transit invariant is a continuous function of position.

The second requirement which (6.13) needs to satisfy is that the value of $\bar{\mathcal{I}}$ must be the same, up to terms of order $\rho_\star vL$, in any set of straight-field-line coordinates. This is a form of gauge invariance. It is not obvious that $\bar{\mathcal{I}}$ is gauge invariant in this way because of the explicit toroidal-angle dependence in the higher-order correction terms in (6.13) or (6.16). Furthermore, the curves in physical space corresponding to fixed $\eta = \theta - (N/M)\zeta$ (which we used to define the kinetic term I) are different in different straight-field-line coordinate systems. In §H, we prove explicitly that formula (6.13) is, nevertheless, gauge invariant.

6.4. Higher-order island shape

We now use the higher-order transit invariant, defined by (6.13) or (6.16), to determine the drift-island shape more accurately. Since the higher-order transit invariant depends not just on ψ and η but also on ζ , the shape of the drift islands varies slightly as we move toroidally around the device. To visualize the islands as they would appear in a Poincaré plot at a fixed toroidal angle, we must plot the level sets of $\bar{\mathcal{I}}(\psi, \eta, \zeta, \mathcal{E}, \mu, \sigma)$ with ζ fixed; for this discussion, we fix $\zeta = 0$.

The modifications to the drift-island shape can be most easily understood by Taylor expanding (6.16) around the rational surface, remembering that $\psi - \psi_r \sim \sqrt{\rho_\star} \Psi_t$ and discarding terms of order $\rho_\star vL$ or smaller:

$$\begin{aligned} \bar{\mathcal{I}}(\psi, \eta, 0, \mathcal{E}, \mu, \sigma) &\simeq \sigma I_r(\eta, \mathcal{E}, \mu) + \sigma(\psi - \psi_r) \partial_\psi I(\psi_r, \eta, \mathcal{E}, \mu) \\ &\quad - \frac{\pi M Ze}{mc} \iota'_r(\psi - \psi_r)^2 - \frac{\pi M Ze}{3mc} \iota''_r(\psi - \psi_r)^3 \\ &\quad + \frac{Ze}{mc} \iota'_r(\psi - \psi_r) \oint_r (\xi - \pi M) \Gamma_\psi d\xi, \end{aligned} \quad (6.17)$$

where $\iota''_r := (d^2\iota/d\psi^2)(\psi = \psi_r)$. This invariant coincides with (3.22) to lowest order but includes new terms that modify the drift-island shape in the following ways.

The first line of (6.17) contains the familiar term $\sigma I_r(\eta, \mathcal{E}, \mu)$ plus a linear (in $\psi - \psi_r$)

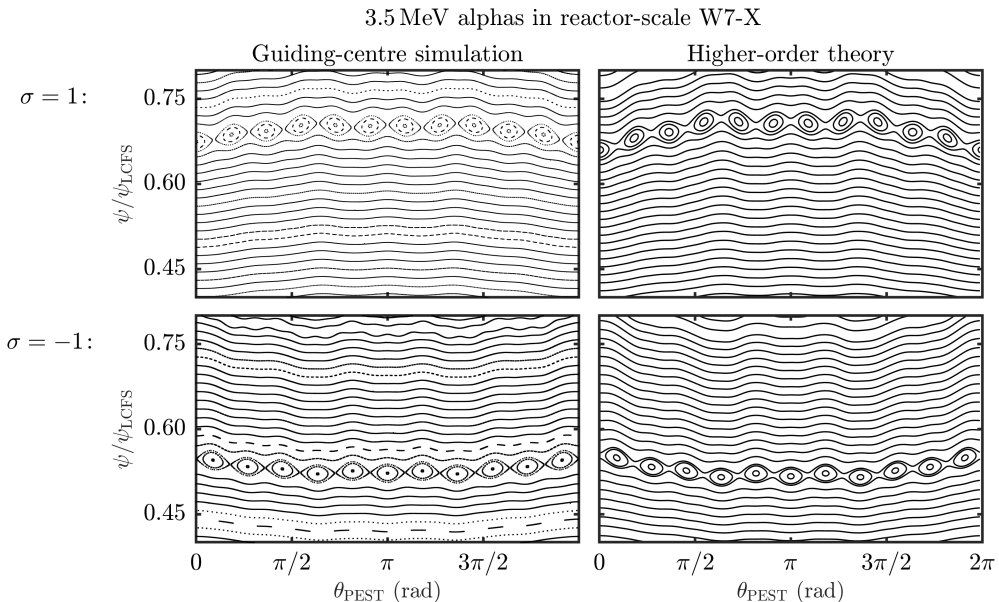


FIGURE 16. (left panels) Here, we repeat the Poincaré plots of [figure 7](#) for passing alpha particles with energy 3.5 MeV and $\mu = 0$ in a reactor-scale W7-X equilibrium. (right panels) This time, we plot the level sets of the higher-order transit invariant $\bar{\mathcal{I}}(\psi, \eta, 0, \mathcal{E}, \mu, \sigma)$, given by [\(6.16\)](#).

correction that shifts the centre of the islands radially. This correction is proportional to the transit average of the particle’s tangential magnetic drift, according to [\(3.9b\)](#). Therefore, this term allows the tangential drift to shift the position of the drift islands radially, as discussed in [§5.3](#).

In the second line of [\(6.17\)](#), the geometric term now has a cubic correction involving ι_r'' . This shows that, at higher order, the island shape is sensitive to the curvature of the ι -profile around the rational surface.

Finally, in the third line of [\(6.17\)](#), we find another linear term that shifts the islands radially. This is the only term that takes different values as we follow a curve of constant ψ and η around the stellarator: it is different at $\zeta = 0, 2\pi, \dots, 2\pi(M - 1)$. Therefore, this is the term that allows the higher-order transit invariant to describe deformed island chains in which each island in the chain has a slightly different shape and radial location.

In [figure 16](#), we compare the higher-order drift-island shape of [\(6.17\)](#) with the simulations presented earlier, in [figure 7](#). It is clear that, with the higher-order corrections, the theory does a much better job of capturing the drift-island shape and position.

7. Hamiltonian perturbation theory for passing particles

We have shown that the orbits of passing particles near a rational flux surface can be calculated using the transit adiabatic invariant. We derived this invariant using a local expansion in which $\iota - N/M$ was ordered as small. There are two disadvantages to such an expansion. First, given the initial coordinates of a particle, there is no clear criterion for which rational surface should play the role of ψ_r ; after all, there are infinitely many rational surfaces to choose from. Second, according to the general theory of nearly integrable Hamiltonian systems, when the drift islands from two nearby rational surfaces

become large enough to overlap, the islands are replaced by a region of stochastic orbits (Chirikov, 1979). It is difficult to see that such a transition should occur when using the local expansion.

In this section, we address these issues by describing another method—based on Hamiltonian perturbation theory—for calculating the orbits of passing particles. This method can be applied to particles on rational or irrational surfaces. Within this framework, a finite set of ‘low-order’ rational surfaces are naturally identified, around which the particle orbits must be described using the transit adiabatic invariant.

The version of Hamiltonian perturbation theory that we use is essentially the Hori (1966) and Deprit (1969) Lie-transform approach to Poincaré–von Zeipel perturbation theory (for reviews, see Cary, 1981; Morbidelli, 2002; Arnold et al., 2006). However, we avoid Lie transforms and instead use Littlejohn’s phase-space Lagrangian formulation (Littlejohn, 1983; Cary & Brizard, 2009; Parra & Calvo, 2011). The phase-space Lagrangian makes it straightforward to use non-canonical coordinates, which we find more convenient for describing guiding-centre motion. The calculation is similar to the calculation of the distorted magnetic surfaces in a perturbed magnetic field, to which the phase-space Lagrangian method has also been applied (e.g. Cary & Littlejohn, 1983; Viana et al., 2023; Escande & Momo, 2024; Chu et al., 2025).

7.1. Phase-space Lagrangian formulation

Littlejohn (1983) showed that the guiding-centre equations of motion can be obtained from a phase-space Lagrangian of the form

$$\mathcal{L}_{\text{gc}} := \left(\mathbf{A} + \frac{v_{\parallel}}{\Omega} \mathbf{B} + \dots \right) \cdot \frac{d\mathbf{X}}{dt} - \left(\frac{mc}{Ze} \right)^2 \mu \frac{d\varphi}{dt} - \frac{mc}{Ze} \mathcal{E}. \quad (7.1)$$

The phase-space coordinates employed here are the guiding-centre position $\mathbf{X}$, the energy per unit mass $\mathcal{E}$, the magnetic moment μ , and the gyrophase φ ; these coordinates must be defined appropriately in order for the Lagrangian to take this form (e.g. Littlejohn, 1983; Parra & Calvo, 2011). All other symbols have the same meaning as before, and functions of position, such as the parallel velocity $v_{\parallel} = \sigma \sqrt{2(\mathcal{E} - \mu B)}$, are evaluated at the guiding-centre position.

In (7.1), the ellipses indicate that higher-order corrections (in powers of $\rho_{\star} \ll 1$) can be included so that the associated guiding-centre equations are accurate to higher order, thus describing the particle motion accurately over longer times. In this section, we ignore these higher-order corrections because we have argued (in §6) that, to calculate the first correction $\mathcal{I}^{(1/2)}$ to the transit adiabatic invariant, the usual guiding-centre equations—without higher-order drifts—are sufficiently accurate. It is worth bearing in mind that the calculation we describe in this section could be continued to arbitrarily high order in $\rho_{\star}$, but this would require higher-order accuracy in the starting phase-space Lagrangian and in the definitions of the guiding-centre coordinates.

We are only interested in the shape of the guiding-centre trajectory through space, not the rate at which the particle advances along this trajectory. As a result, the calculation is simpler if we parametrize position along the trajectory using the toroidal angle ζ instead of time t . The relevant phase-space Lagrangian is then

$$\mathcal{L}_{\text{gc}}^{\zeta} := \left(\mathbf{A} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \right) \cdot \frac{d\mathbf{X}}{d\zeta} - \left(\frac{mc}{Ze} \right)^2 \mu \frac{d\varphi}{d\zeta} - \frac{mc}{Ze} \mathcal{E} \frac{dt}{d\zeta}, \quad (7.2)$$

where t is now treated as a phase-space coordinate. Since $\mathcal{E}$ and μ are both constant,

and since we are not interested in calculating how φ or t evolve with ζ , we can lower the dimensionality of phase space by using, instead of (7.2), the reduced Lagrangian

$$\mathcal{L} := \left(\mathbf{A} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \right) \cdot \frac{d\mathbf{X}}{d\zeta}. \quad (7.3)$$

In (7.3), $\mathcal{E}$ and μ are treated as fixed parameters. This fact will be important whenever we add a total ζ derivative to the Lagrangian (7.3). There will be no need to include $\mathcal{E}$ or μ derivatives when using the chain rule (i.e. $df/d\zeta = \partial_{\psi}f(d\psi/d\zeta) + \partial_{\theta}f(d\theta/d\zeta) + \partial_{\zeta}f$). The reduced Lagrangian (7.3) can also be obtained by enforcing $\mathcal{E} = \text{const.}$ and $\mu = \text{const.}$ as holonomic constraints in (7.2) (e.g. Goldstein *et al.*, 2002). We can think of (7.3) as the field-line Lagrangian (Cary & Littlejohn, 1983) for the fictitious magnetic field $\mathbf{B}^*$ which, as explained in §2.2, the guiding centre follows. We will use (7.3) to calculate the guiding-centre orbits of passing particles in a stellarator magnetic field.

7.2. Drift surfaces

First, we express the Lagrangian in the straight-field-line coordinates (ψ, θ, ζ) appropriate for magnetic fields with nested toroidal flux surfaces. For the vector potential, we can take (e.g. Helander, 2014)

$$\mathbf{A} := \psi \nabla \theta - \chi(\psi) \nabla \zeta, \quad (7.4)$$

where $\chi(\psi) := \int_0^{\psi} \iota(\psi') d\psi'$ is the poloidal flux contained inside flux surface ψ . Then, (7.3) becomes

$$\mathcal{L} = \left(\psi + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \frac{d\theta}{d\zeta} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \frac{d\psi}{d\zeta} - \chi + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x}. \quad (7.5)$$

Throughout this section, the partial derivatives ∂_{ψ} , ∂_{θ} and ∂_{ζ} are taken with respect to one of the coordinates (ψ, θ, ζ) , with the other two held fixed. Note that this changes the definition of ∂_{ζ} , which in previous sections denoted a derivative at fixed ψ and η .

It is useful to think of (7.5) as a small perturbation to the Lagrangian for the field lines of the magnetic field, which is

$$\mathcal{L}_B := \psi \frac{d\theta}{d\zeta} - \chi(\psi). \quad (7.6)$$

The Euler-Lagrange equations of motion corresponding to (7.6) are simply

$$\frac{d\psi}{d\zeta} = 0, \quad \frac{d\theta}{d\zeta} = \iota(\psi). \quad (7.7)$$

The first equation states that $\mathbf{B}$ has flux surfaces, while the second states that θ and ζ are straight-field-line coordinates. The reason these equations are so simple is that $\mathcal{L}_B$ is independent of the angles θ and ζ .¹⁴ By contrast, the equations of motion for the guiding-centre Lagrangian (7.5) are

$$\frac{d\psi}{d\zeta} = \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] + O(\rho_{\star}^2 \Psi_t), \quad (7.8a)$$

$$\frac{d\theta}{d\zeta} = \iota(\psi) - \left[\partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) \right] + O(\rho_{\star}^2). \quad (7.8b)$$

In comparison with (7.7), each equation here has an additional, order- $\rho_{\star}$ term that depends on both angles θ and ζ ; physically, these terms represent spatially varying drifts. The

¹⁴The coordinates ψ and θ are action-angle coordinates for the Hamiltonian system described by phase-space Lagrangian $\mathcal{L}_B$.

key idea of Hamiltonian perturbation theory is to look for a coordinate transformation that removes the angle dependence from the Lagrangian (7.5), so that the guiding-centre equations of motion become as simple as (7.7).

Since the angle-dependent terms in (7.5) are small in $\rho_\star$, we will look for a *near-identity* coordinate transformation of the form

$$\psi = \bar{\psi} + \psi^{(1)}(\bar{\psi}, \bar{\theta}, \zeta), \quad (7.9a)$$

$$\theta = \bar{\theta} + \theta^{(1)}(\bar{\psi}, \bar{\theta}, \zeta), \quad (7.9b)$$

where $\psi^{(1)} \sim \rho_\star \Psi_t$ and $\theta^{(1)} \sim \rho_\star$ are small corrections. At this stage, (7.9) is a completely general near-identity coordinate transformation. Shortly, we will make a particular choice for $\psi^{(1)}$ and $\theta^{(1)}$ that eliminates any dependence on $\bar{\theta}$ and ζ from the Lagrangian (or, at least, relegates this dependence to higher order). Equation (7.9) defines the new coordinates $\bar{\psi}$ and $\bar{\theta}$ implicitly, analogous to how the guiding-centre position is often defined implicitly by

$$\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}(\mathbf{X}, \mathcal{E}, \mu, \varphi), \quad (7.10)$$

where $\mathbf{x}$ is the particle position, $\mathbf{X}$ is the guiding-centre position, and $\boldsymbol{\rho}$ is the gyroradius vector.

Using (7.9) to change coordinates in the Lagrangian gives

$$\begin{aligned} \mathcal{L} = & \left(\bar{\psi} + \psi^{(1)} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \left(\frac{d\bar{\theta}}{d\zeta} + \frac{d\theta^{(1)}}{d\zeta} \right) + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \left(\frac{d\bar{\psi}}{d\zeta} + \frac{d\psi^{(1)}}{d\zeta} \right) \\ & - \chi - \iota \psi^{(1)} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x}, \end{aligned} \quad (7.11)$$

where we drop terms of order $\rho_\star^2$ from the Lagrangian throughout the rest of this section.¹⁵ From here on, functions of position without arguments written explicitly are assumed to be evaluated at the point with magnetic coordinates $(\bar{\psi}(\psi, \theta, \zeta), \bar{\theta}(\psi, \theta, \zeta), \zeta)$, rather than at the actual guiding-centre position (ψ, θ, ζ) .

Recall that a total time derivative can be added to a Lagrangian without modifying the associated equations of motion. To remove the derivatives of the as-yet-undetermined functions $\psi^{(1)}$ and $\theta^{(1)}$ from (7.11), we subtract the total ζ derivative

$$\frac{d}{d\zeta} \left[\left(\bar{\psi} + \psi^{(1)} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \theta^{(1)} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \psi^{(1)} \right]. \quad (7.12)$$

Furthermore, we anticipate that we may need to add other ζ derivatives to the Lagrangian to put it in a form that has no angle dependence. To allow for this possibility, we subtract $dS^{(1)}/d\zeta$ for some ‘generating function’ $S^{(1)} \sim \rho_\star \Psi_t$, which we will choose later. The result is

$$\begin{aligned} \mathcal{L} \equiv & \left(\bar{\psi} + \psi^{(1)} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} - \partial_{\theta} S^{(1)} \right) \frac{d\bar{\theta}}{d\zeta} + \left(-\theta^{(1)} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} - \partial_{\psi} S^{(1)} \right) \frac{d\bar{\psi}}{d\zeta} \\ & - \chi - \iota \psi^{(1)} + \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} - \partial_{\zeta} S^{(1)}, \end{aligned} \quad (7.13)$$

where we use the symbol $\equiv$ to denote equality up to a total ζ derivative. We can remove

¹⁵More precisely: if we represent our Lagrangian in the form $\mathcal{L} = \Lambda_{\psi}(d\bar{\psi}/d\zeta) + \Lambda_{\theta}(d\bar{\theta}/d\zeta) - \mathcal{H}$, then we drop terms of order $\rho_\star^2$ from the symplectic parts Λ_{ψ} and Λ_{θ} , and from the Hamiltonian $\mathcal{H}$. Such terms are as small as the higher-order terms that we neglected, *ab initio*, in (7.1).

some of the angle dependence in (7.13) by picking

$$\psi^{(1)} := -\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} + \partial_{\theta} S^{(1)}, \quad (7.14a)$$

$$\theta^{(1)} := \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} - \partial_{\psi} S^{(1)}, \quad (7.14b)$$

for whatever choice of $S^{(1)}$ we eventually make. This leaves

$$\mathcal{L} \equiv \bar{\psi} \frac{d\bar{\theta}}{d\zeta} - \chi + V - (\partial_{\zeta} S^{(1)} + \iota \partial_{\theta} S^{(1)}), \quad (7.15)$$

where, as in §4.1,

$$V = \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot (\partial_{\zeta} \mathbf{x} + \iota \partial_{\theta} \mathbf{x}) = \frac{mc}{Ze} \frac{v_{\parallel}}{\hat{\mathbf{b}} \cdot \nabla \zeta}. \quad (7.16)$$

The remaining angle dependence in (7.15) would be eliminated if we could choose $S^{(1)}$ to satisfy the magnetic differential equation

$$\partial_{\zeta} S^{(1)} + \iota \partial_{\theta} S^{(1)} = V - \langle V \rangle. \quad (7.17)$$

Here, we have introduced the average over the poloidal and toroidal angles

$$\langle \dots \rangle := \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} (\dots) d\theta d\zeta. \quad (7.18)$$

We included the term $\langle V \rangle$ in (7.17) because the left-hand side clearly vanishes under this average, so the right-hand side must as well. Unfortunately, we will find that (7.17) is not always solvable for $S^{(1)}$. Furthermore, even when (7.17) is solvable, the solutions have certain deficiencies that make them unsuitable for completing the construction of our near-identity coordinate transformation. In the next subsection, we discuss these issues with (7.17) in more detail.

7.3. Solvability of the magnetic differential equation

For a general inhomogeneous term V , magnetic differential equations of the form (7.17) have been extensively studied (e.g. [Kruskal & Kulsrud, 1958](#); [Newcomb, 1959](#); [Hamada, 1962](#); [Rüssmann, 1975](#); [Imbert-Gérard *et al.*, 2024](#)). In this subsection, we investigate when (7.17) can be solved in the specific case where $V = (mc/Ze)(v_{\parallel}/\hat{\mathbf{b}} \cdot \nabla \zeta)$.

Equation (7.17) specifies the rate of change of $S^{(1)}$ along magnetic field lines, since it can be recast as

$$\hat{\mathbf{b}} \cdot \nabla S^{(1)} = \frac{mc}{Ze} \left(v_{\parallel} - \hat{\mathbf{b}} \cdot \nabla \zeta \left\langle \frac{v_{\parallel}}{\hat{\mathbf{b}} \cdot \nabla \zeta} \right\rangle \right) := U. \quad (7.19)$$

Symbolically, the solution to this equation is

$$S^{(1)} \stackrel{?}{=} \int^l U dl', \quad (7.20)$$

where the integral is taken along field lines. However, it is unclear whether (7.20) defines a continuous function of position on each flux surface. On an irrational surface, two arbitrarily close points could be connected by an arbitrarily large distance along a single field line, meaning their values of $S^{(1)}$ —calculated by integrating U along this line—could be very different. The problem with (7.20) is that it does not ‘know about’ the toroidal periodicity of the domain.

This observation suggests we look for a periodic solution to (7.17) using a Fourier expansion.

Let

$$S^{(1)} = \sum_{(p,q) \in \mathbb{Z}^2} S_{pq}^{(1)} e^{i(p\theta - q\zeta)}, \quad V = \sum_{(p,q) \in \mathbb{Z}^2} V_{pq} e^{i(p\theta - q\zeta)}. \quad (7.21)$$

In terms of these Fourier coefficients, (7.17) becomes

$$-i(q - \iota p) S_{pq}^{(1)} = \begin{cases} V_{pq} & \text{if } (p, q) \neq (0, 0), \\ 0 & \text{if } (p, q) = (0, 0). \end{cases} \quad (7.22)$$

Formally, this means

$$S^{(1)} \stackrel{?}{=} \sum_{(p,q) \in \mathbb{Z}^2 \setminus \{(0,0)\}} \frac{iV_{pq}}{q - \iota p} e^{i(p\theta - q\zeta)}, \quad (7.23)$$

where we have chosen the solution with zero average over the poloidal and toroidal angles ($S^{(1)}$ is determined up to an arbitrary additive constant).

Now, the question of whether the magnetic differential equation (7.17) is solvable becomes the question of whether the Fourier series in (7.23) converges. This series will not, typically, converge if the rotational transform ι is rational: in this case, there are combinations of p and q that make the denominator vanish, which is called a ‘resonant’ denominator. It turns out that the series also diverges at certain irrational values of ι (Rüssmann, 1975; Yoccoz, 1992).

The series (7.23) converges in two physically relevant scenarios, which are summarized below.

- (i) The series converges for certain ‘highly irrational’ values of ι . For example, a sufficient condition for the series to converge for all analytic V is that ι satisfies the following Diophantine condition: there are constants $C > 0$ and $\varsigma > 0$ such that

$$\left| \iota - \frac{N}{M} \right| > \frac{C}{|M|^{2+\varsigma}}, \quad (7.24)$$

for all integers N and M (see Arnold, 1963b, §3). This condition prevents ι from being too close to any rational number. When (7.24) is satisfied, the terms in the series (7.23) decay fast enough for convergence because the numerators V_{pq} decay exponentially for analytic V while the denominators decay at most algebraically.

- (ii) The series converges if the numerators vanish whenever the denominators do. This is clearly a necessary condition for convergence; interestingly, it is also sufficient, according to the following result of Newcomb (1959). Let (ψ_1, ψ_2) be an interval in which the shear $\iota'(\psi)$ never vanishes. The series (7.23) converges for all $\psi \in (\psi_1, \psi_2)$ if (1) V is sufficiently smooth and (2) for every rational surface $\psi_{MN} \in (\psi_1, \psi_2)$ and for all $k \in \mathbb{Z}$, we have $V_{kM, kN}(\psi_{MN}) = 0$, where N and $M > 0$ are coprime integers defined by $\iota(\psi_{MN}) = N/M$. When condition (2) is satisfied, certain terms in the series in (7.23) have an indeterminate ‘0/0’ form at rational surfaces. This indeterminacy arises because (7.22) does not fix the quantities $S_{kM, kN}^{(1)}(\psi_{MN})$; as a result, we can freely choose their values. To ensure continuity with the irrational surfaces, we must choose

$$S_{kM, kN}^{(1)}(\psi_{MN}) = -\frac{i\partial_\psi V_{kM, kN}(\psi_{MN})}{kM\iota'(\psi_{MN})}, \quad (7.25)$$

and we assume that this choice has been made when interpreting (7.23) on rational surfaces.

The condition $V_{kM,kN}(\psi_{MN}, \mathcal{E}, \mu, \sigma) = 0$ (where we have temporarily restored the velocity-space arguments) arises naturally in the context of passing-particle orbits in stellarators. We showed in §4.1 that this condition is equivalent to the absence of drift islands around rational surface ψ_{MN} , for particles of energy $\mathcal{E}$ and magnetic moment μ . If $V_{kM,kN}(\psi_{MN}, \mathcal{E}, \mu, \sigma) = 0$ for all passing $\mathcal{E}$ and μ , then the magnetic field must be cyclometric on ψ_{MN} , as discussed in §5.2. Since omnigenous stellarators are always cyclometric, this condition is automatically satisfied in omnigenous devices. The fact that $S^{(1)}$ —and, therefore, $\bar{\psi}$ —is always well-defined in perfectly omnigenous stellarators is unsurprising, as it is well-known that (ψ, α, B) coordinates may be used to derive a formula equivalent to (7.30) for $\bar{\psi}$ which does not involve a sum and which, therefore, never suffers from convergence issues (e.g. Helander & Nührenberg, 2009; Landreman & Catto, 2012).

To summarize, the Fourier series (7.23) sometimes converges, for example when ι is ‘highly irrational’ or when the stellarator is perfectly cyclometric in a neighbourhood of the flux surface of interest. Unfortunately, even in these cases, the resulting $S^{(1)}$ cannot be used to complete the construction of our near-identity coordinate transformation, for the following reasons.

- (i) In the case of highly irrational flux surfaces, the Fourier series for $S^{(1)}$ converges on these surfaces. However, it still (generically) diverges on the dense set of surfaces that do not satisfy (7.24); this set includes, for example, all rational surfaces. Therefore, $S^{(1)}$ must be discontinuous in the radial direction. Our perturbative change of coordinates required $S^{(1)}$ to have a well-defined radial derivative—in particular, the definition of $\theta^{(1)}$ involved $\partial_\psi S^{(1)}$ —so this is a problem.
- (ii) In the case of cyclometric stellarators, the Fourier series for $S^{(1)}$ diverges if any harmonic $V_{kM,kN}(\psi_{MN})$ is non-zero, regardless of how small the harmonic may be. A formalism that breaks down in the presence of an infinitesimal deviation from cyclometry is inadequate for describing real devices.

In the next subsection, we explain how these problems associated with the resonant denominators may be circumvented.

7.4. Dealing with resonances

Our approach is based on two key ideas (other methods exist; see e.g. Morbidelli, 2002). The first idea is that we do not need to eliminate the angle dependence in the Lagrangian completely; we only need to eliminate it to order ρ_* . To use this idea, we follow Arnold (1963a) and introduce an ‘ultraviolet cutoff’ by splitting the perturbation into $V = V^{<K} + V^{\geq K}$, where $V^{<K}$ and $V^{\geq K}$ are defined as follows. The first term, $V^{<K}$, contains all harmonics of V that have $|p| + |q| < K$ for some cutoff $K > 0$. The second term, $V^{\geq K}$, contains all the remaining harmonics. Crucially, we define the cutoff K to be large enough that $V^{\geq K}$ is an order (in ρ_*) smaller than $V^{<K}$; thus $V^{\geq K} \sim \rho_*^2 \Psi_t$. For a sufficiently smooth V , we can always find a K large enough to satisfy this requirement. The ultraviolet cutoff is useful because $V^{\geq K}$ is small enough that it can be neglected in the Lagrangian, so we only need to eliminate the angle dependence contained in $V^{<K}$.

Resuming our calculation from (7.15), we can write

$$\mathcal{L} \equiv \bar{\psi} \frac{d\bar{\theta}}{d\zeta} - \chi + V^{<K} - (\partial_\zeta S^{(1)} + \iota \partial_\theta S^{(1)}), \quad (7.26)$$

since we are dropping terms of order $\rho_\star^2$ or smaller from the Lagrangian. The calculation now proceeds as before, but with V replaced by the smoother function $V^{<K}$. Thus, to eliminate the angle dependence in (7.26), we take

$$S^{(1)} := \sum_{(p,q) \in \mathbb{I}^{<K}} \frac{i(V^{<K})_{pq}}{q - \iota p} e^{i(p\theta - q\zeta)}, \quad (7.27)$$

where the index set is $\mathbb{I}^{<K} := \{(p, q) \in \mathbb{Z}^2 : |p| + |q| < K \text{ and } (p, q) \neq (0, 0)\}$. Equation (7.27) replaces the problematic definition (7.23). The series in (7.27) converges on all flux surfaces except a finite set of rational surfaces: those that resonate with one of the finitely many harmonics of $V^{<K}$. We will refer to these surfaces as ‘low-order’ rational surfaces.

To avoid these divergences on low-order rational surfaces, we introduce the second key idea: our perturbative construction does not need to be valid for all passing particles simultaneously. Instead, we can restrict attention to particles in a region that does not contain any low-order rational surfaces; in such a region, the rotational transform is, in some sense, ‘sufficiently irrational’. From now on, we assume $\psi \in \mathfrak{R}_0$, where $\mathfrak{R}_0$ is a spatial region that does not contain any of the low-order rational surfaces. Note that we will slightly refine the definition of $\mathfrak{R}_0$ below, in [definition 7.1](#).

Within $\mathfrak{R}_0$, the series (7.27) converges and the near-identity coordinate transformation specified by (7.9), (7.14) and (7.27) is well-defined. The Lagrangian in these new coordinates is

$$\mathcal{L} \equiv \bar{\psi} \frac{d\bar{\theta}}{d\zeta} - \bar{\chi}(\bar{\psi}), \quad (7.28)$$

where $\bar{\chi} := \chi - \langle V \rangle$. We have achieved our aim of removing angle dependence from the Lagrangian; any angle dependence is now hidden in higher-order terms that have been suppressed in (7.28).

According to (7.28), the new coordinates $\bar{\psi}$ and $\bar{\theta}$ evolve with ζ in a trivial way. One equation of motion simply states that $\bar{\psi}$ is constant (to leading order; its ζ derivative can be non-zero at order $\rho_\star^2 \Psi_t$ due to the suppressed higher-order terms). Physically, this means passing particles on irrational surfaces do not experience a secular radial drift over long times of order $(1/\rho_\star)(L/v)$, as stated in [§2](#). The other equation of motion is

$$\frac{d\bar{\theta}}{d\zeta} = \iota(\bar{\psi}) + \frac{mc}{Ze} \partial_\psi \left\langle \frac{v_\parallel}{\hat{\mathbf{b}} \cdot \nabla \zeta} \right\rangle := \bar{\iota}(\bar{\psi}), \quad (7.29)$$

which means $\bar{\theta}$ changes with ζ at a constant rate (again, there may be deviations from this behaviour at order $\rho_\star^2$). This rate $\bar{\iota}$ is the kinetic rotational transform of the orbit; it is determined by a combination of the rotational transform of the field and the flux-surface-averaged tangential drift—which is a small correction—as discussed in [§5.3](#).

Upon inverting the coordinate transformation (7.9) to first order in $\rho_\star$, we find that

$$\bar{\psi}(\psi, \theta, \zeta) = \psi + \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\theta \mathbf{x} \right)(\psi, \theta, \zeta) - (\partial_\theta S^{(1)})(\psi, \theta, \zeta) + O(\rho_\star^2 \Psi_t) \quad (7.30)$$

is a conserved quantity. This is an equation for the passing-particle drift surfaces in regions of sufficiently irrational rotational transform in any stellarator. It is consistent with a less

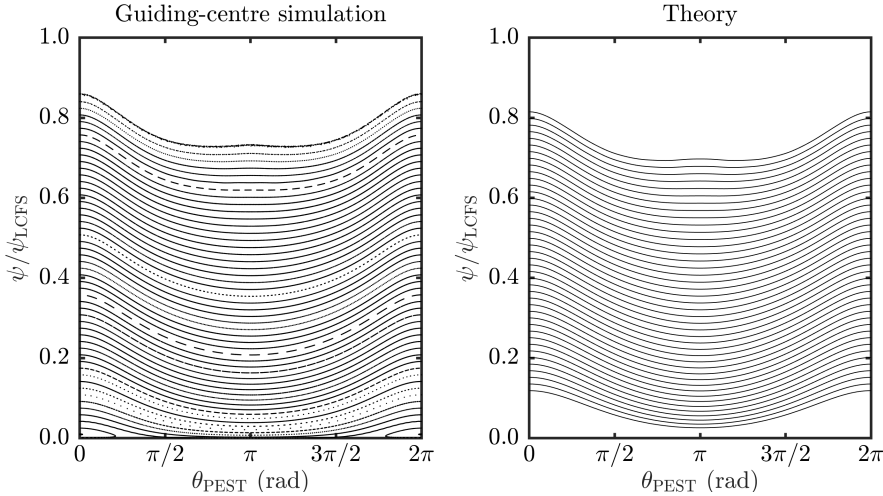


FIGURE 17. (a) Poincaré plot of the guiding-centre orbits of passing alpha particles in a reactor-scale LHD equilibrium, simulated using ASCOT5. The alphas have energy 3.5 MeV, $\mu/\mathcal{E} = 0.111 \text{ T}^{-1}$, and $\sigma = 1$. Drift surfaces beyond $\psi/\psi_{\text{LCFS}} \approx 0.8$ are missing because we have not plotted orbits that crossed the last closed flux surface. (b) Level sets of $\bar{\psi}$, defined by (7.30). The plotted contour levels were chosen as follows: a contour level was computed for each particle by averaging $\bar{\psi}$ over the points plotted in (a) for that particle.

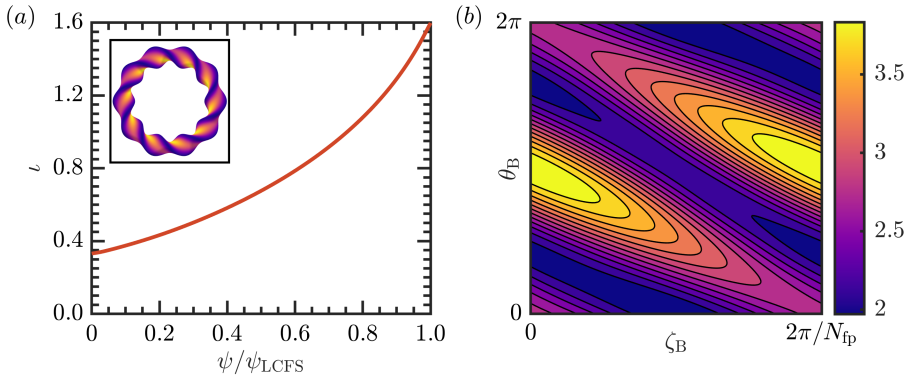


FIGURE 18. (a) Rotational transform of the LHD equilibrium we use to trace alpha orbits in figure 17. (b) Plot of the magnetic field strength in Boozer coordinates on the $\iota = 1$ flux surface.

general result derived in Tykhyi & Yakovenko (2006, equation (12)) and Kolesnichenko *et al.* (2006, appendix A) for a magnetic field whose field strength is equal to a constant plus a small angle-dependent term, provided an ultraviolet cutoff is used to make the Fourier series in these references well-behaved.

In figure 17, we compare this calculation of particle orbits in a region of sufficiently irrational rotational transform with an ASCOT5 simulation for a reactor-scale LHD equilibrium. More detail about the equilibrium is provided in figure 18. Despite the fact that the rotational transform passes through many rational values, there are no visible drift islands; the same is true at all other pitch angles. This must be a consequence of the high shear and smallness of any resonant harmonics of V .

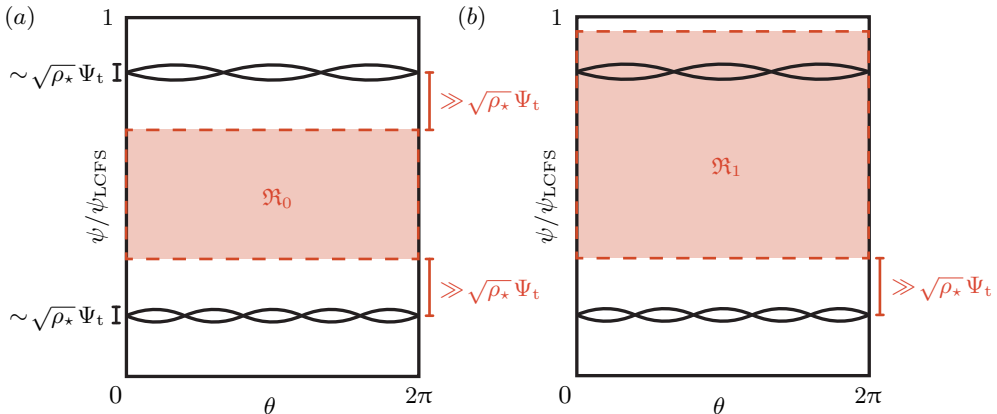


FIGURE 19. Diagrams showing the regions of validity for the orbit calculations presented in §§7.4–7.5. In each case, a Poincaré plot is drawn with drift-island chains around two low-order rational surfaces. (a) Region $\mathfrak{R}_0$ must remain much further than $\sim\sqrt{\rho_*}\Psi_t$ from either low-order rational surface. (b) Region $\mathfrak{R}_1$ contains one of the low-order rational surfaces and must remain much further than $\sim\sqrt{\rho_*}\Psi_t$ from the other.

Before moving on, it is instructive to see how the above calculation fails close to a low-order rational surface. Let the nearest low-order rational surface to $\mathfrak{R}_0$ be ψ_r . The calculation can fail because, when $|\psi - \psi_r|$ is too small for some $\psi \in \mathfrak{R}_0$, $S^{(1)}$ becomes large enough to violate our orderings. For the relations $\psi^{(1)} \ll \Psi_t$ and $\theta^{(1)} \ll 1$ to hold, we need $S^{(1)} \ll \Psi_t$ and $\partial_\psi S^{(1)} \ll 1$. From (7.27), $S^{(1)}$ diverges as $S^{(1)} \sim \rho_* \Psi_t / \iota'_r(\psi - \psi_r)$ (we show this explicitly in §I—see (I 11)), which means $\partial_\psi S^{(1)} \sim \rho_* \Psi_t / \iota'_r(\psi - \psi_r)^2$. Thus, for our construction to be valid, we need $|\psi - \psi_r| \gg (\rho_*/s)^{1/2} \Psi_t$. The lower bound in this inequality is the width of the drift islands around ψ_r . Therefore, for our coordinate transformation to be valid throughout $\mathfrak{R}_0$, we need $\mathfrak{R}_0$ to stay much further than a drift-island width away from any low-order rational surfaces, as shown in figure 19(a). For reference, we now state the complete definition of $\mathfrak{R}_0$.

Definition 7.1. $\mathfrak{R}_0$ is an interval (ψ_1, ψ_2) whose distance to any low-order rational surface, ψ_r , satisfies $\text{dist}(\psi_r, \mathfrak{R}_0) \gg (\rho_*/s)^{1/2} \Psi_t$.¹⁶

Of course, we would also like to be able to describe orbits near the low-order rational surfaces, since this is where drift islands can appear. We consider this problem in the next subsection.

7.5. Drift surfaces in the presence of a single resonance

By modifying the method above, it is possible to calculate orbits in a region $\mathfrak{R}_1$ that contains a single low-order rational surface. Let this surface be ψ_r , with $\iota(\psi_r) = N/M$ as usual. The key idea here is that we can still eliminate most of the angle dependence from the Lagrangian by following the same steps as before. However, to avoid divergences, we must exclude all harmonics that are resonant at ψ_r from the Fourier expansion of $S^{(1)}$.

¹⁶Note that, as $\rho_* \rightarrow 0$, the condition $V^{\geq K} \sim \rho_* V^{< K}$ means we must have $K \rightarrow \infty$. As K increases, the number of low-order rational surfaces grows without bound. Thus, any region $\mathfrak{R}_0$ that satisfies definition 7.1 must decrease in size as $\rho_* \rightarrow 0$. In order to construct a theory that remains valid within a finite region as $\rho_* \rightarrow 0$, we could fix the value of K and use $K \gg 1$ as an additional expansion parameter, independent of ρ_* .

First, we split the perturbation as $V = V_r + V_{nr}^{<K} + V_{nr}^{\geq K}$, where

$$V_r := \sum_{k \in \mathbb{Z}} V_{kM, kN} e^{ik(M\theta - N\zeta)} \quad (7.31)$$

contains all resonant harmonics of V , $V_{nr}^{<K}$ contains all non-resonant harmonics with $|p| + |q| < K$, and $V_{nr}^{\geq K}$ contains the non-resonant harmonics with $|p| + |q| \geq K$. As before, we choose K large enough that $V_{nr}^{\geq K} \sim \rho_\star^2 \Psi_t$. Then, the Lagrangian in (7.15) becomes

$$\mathcal{L} \equiv \bar{\psi}_{nr} \frac{d\bar{\theta}_{nr}}{d\zeta} - \chi + V_r + V_{nr}^{<K} - (\partial_\zeta S_{nr}^{(1)} + \iota \partial_\theta S_{nr}^{(1)}). \quad (7.32)$$

In moving from (7.15) to (7.32), we have introduced new notation because we are now dealing with particles in $\mathfrak{R}_1$ instead of $\mathfrak{R}_0$. To avoid confusion, when we are working in $\mathfrak{R}_1$ we denote the generating function by $S_{nr}^{(1)}$ and the corresponding barred coordinates by $\bar{\psi}_{nr}$ and $\bar{\theta}_{nr}$. To be explicit, $\bar{\psi}_{nr}$ and $\bar{\theta}_{nr}$ are defined by

$$\psi = \bar{\psi}_{nr} + \psi^{(1)}(\bar{\psi}_{nr}, \bar{\theta}_{nr}, \zeta), \quad (7.33a)$$

$$\theta = \bar{\theta}_{nr} + \theta^{(1)}(\bar{\psi}_{nr}, \bar{\theta}_{nr}, \zeta), \quad (7.33b)$$

where

$$\psi_{nr}^{(1)} := -\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_\theta \mathbf{x} + \partial_\theta S_{nr}^{(1)}, \quad (7.34a)$$

$$\theta_{nr}^{(1)} := \frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_\psi \mathbf{x} - \partial_\psi S_{nr}^{(1)}, \quad (7.34b)$$

analogous to (7.9) and (7.14). This time, however, we choose

$$S_{nr}^{(1)} := \sum_{(p,q) \in \mathbb{I}_{nr}^{\leq K}} \frac{i(V_{nr}^{<K})_{pq}}{q - \iota p} e^{i(p\theta - q\zeta)}, \quad (7.35)$$

where $\mathbb{I}_{nr}^{\leq K} := \{(p, q) \in \mathbb{Z}^2 : |p| + |q| < K \text{ and } (p, q) \neq (kM, kN) \text{ for any } k \in \mathbb{Z}\}$ contains only non-resonant harmonics. This choice of generating function satisfies

$$\partial_\zeta S_{nr}^{(1)} + \iota \partial_\theta S_{nr}^{(1)} = V_{nr}^{<K}, \quad (7.36)$$

so the Lagrangian (7.32) becomes

$$\mathcal{L} \equiv \bar{\psi}_{nr} \frac{d\bar{\theta}_{nr}}{d\zeta} - \chi + \sum_{k \in \mathbb{Z}} V_{kM, kN} e^{ik(M\bar{\theta}_{nr} - N\zeta)}. \quad (7.37)$$

We can write (7.37) in a simple form by introducing the ζ -average

$$\langle \dots \rangle_\zeta := \frac{1}{2\pi M} \oint (\dots) d\zeta, \quad (7.38)$$

where we integrate over a curve of constant ψ and η as usual. This average is useful because it extracts Fourier modes with the helicity of the rational surface:

$$\langle V \rangle_\zeta = \sum_{k \in \mathbb{Z}} V_{kM, kN} e^{ikM\eta} = V_r. \quad (7.39)$$

Then, using $\bar{\eta}_{nr} := \bar{\theta}_{nr} - (N/M)\zeta$ as a coordinate instead of $\bar{\theta}_{nr}$, the Lagrangian (7.37) becomes

$$\mathcal{L} \equiv \bar{\psi}_{nr} \frac{d\bar{\eta}_{nr}}{d\zeta} - \mathcal{H}_r(\bar{\psi}_{nr}, \bar{\eta}_{nr}), \quad (7.40)$$

where $\mathcal{H}_r := \chi_r - \langle V \rangle_\zeta$ and χ_r is the helical flux introduced back in (3.17). This Lagrangian is the generalization of (7.28) to the case where there is a single low-order rational surface in the domain of interest.

In (7.40), the dependence on the fast angle ζ has been eliminated. Thus, in the presence of a single resonance, although we could not eliminate both angular coordinates as in §7.4, it is still possible to significantly simplify the Lagrangian. The ζ independence of this Lagrangian implies conservation of

$$\mathcal{H}_r(\bar{\psi}_{\text{nr}}, \bar{\eta}_{\text{nr}}) = \int_{\psi_r}^{\bar{\psi}_{\text{nr}}} \left(\iota(\psi') - \frac{N}{M} \right) d\psi' - \frac{1}{2\pi M} \frac{mc}{Ze} \oint v_{\parallel} \hat{\mathbf{b}} \cdot (\partial_\zeta \mathbf{x} + \iota \partial_\theta \mathbf{x}) d\zeta'. \quad (7.41)$$

Here, as in the Lagrangian, all functions are evaluated at a ψ coordinate of $\bar{\psi}_{\text{nr}}(\psi, \eta, \zeta)$ and an η coordinate of $\bar{\eta}_{\text{nr}}(\psi, \eta, \zeta)$. To express the conserved quantity (7.41) in the original guiding-centre coordinates, let $\bar{\chi}_r(\psi, \eta, \zeta) := \mathcal{H}_r(\bar{\psi}_{\text{nr}}(\psi, \eta, \zeta), \bar{\eta}_{\text{nr}}(\psi, \eta, \zeta))$. Substituting the definitions of $\bar{\psi}_{\text{nr}}$ and $\bar{\eta}_{\text{nr}}$ into (7.41) and Taylor expanding, we find

$$\begin{aligned} \bar{\chi}_r(\psi, \eta, \zeta) = & \int_{\psi_r}^{\psi} \left(\iota(\psi') - \frac{N}{M} \right) d\psi' - \frac{1}{2\pi M} \frac{mc}{Ze} \oint v_{\parallel} \hat{\mathbf{b}} \cdot (\partial_\zeta \mathbf{x} + \iota \partial_\theta \mathbf{x}) d\zeta' \\ & + \left(\iota(\psi) - \frac{N}{M} \right) \left(\frac{mc}{Ze} v_{\parallel} \hat{\mathbf{b}} \cdot \partial_\theta \mathbf{x} - \partial_\theta S_{\text{nr}}^{(1)} \right) + O(\rho_\star^2 \Psi_t), \end{aligned} \quad (7.42)$$

where all functions on the right are now evaluated at (ψ, η, ζ) . This conserved quantity $\bar{\chi}_r$ determines the orbits of all passing particles in $\mathfrak{R}_1$, just as $\bar{\psi}$ did for particles in $\mathfrak{R}_0$. Region $\mathfrak{R}_1$ contains both the low-order rational surface ψ_r , around which the orbits trace out drift islands, and irrational surfaces away from ψ_r , where the drift surfaces are intact (to this order). Both types of orbit are described by a single invariant $\bar{\chi}_r$, whose definition (7.42) is uniformly valid throughout $\mathfrak{R}_1$.

In §I, we prove that, in two opposite limits, (7.42) is consistent with results from earlier in the paper. For particles close to ψ_r , conservation of $\bar{\chi}_r$ is equivalent to conservation of the transit adiabatic invariant, including the higher-order corrections derived in §6. For particles on irrational surfaces far from ψ_r , conservation of $\bar{\chi}_r$ is equivalent to conservation of $\bar{\psi}$, defined by (7.30) for particles in a region, $\mathfrak{R}_0$, of sufficiently irrational rotational transform.

In §7.4, we discussed a restriction on region $\mathfrak{R}_0$: it needed to remain much more than a drift-island width away from any low-order rational surfaces. Similar reasoning shows that $\mathfrak{R}_1$ must remain much more than a drift-island width away from any low-order rational surfaces other than ψ_r , as shown in figure 19(b).

These restrictions on $\mathfrak{R}_0$ and $\mathfrak{R}_1$ can only be satisfied if the drift-island width is much smaller than the distance between low-order rational surfaces. This condition may not be satisfied if $\rho_\star$ is insufficiently small; then, our perturbative construction would fail. Such a failure could be anticipated on physical grounds because, once the drift-island width is comparable to the distance between rational surfaces, drift-island overlap can occur. Overlapping islands are known to give rise to stochastic orbits (Chirikov, 1979), which cannot be described by conservation of an invariant such as $\bar{\psi}$ or $\bar{\chi}_r$.

In figures 20–23, we compare the analytical results of this subsection with ASCOT5 simulations. Figures 20(a) and 20(b) show a Poincaré plot of simulated guiding-centre orbits of freely passing, 3.5 MeV alpha particles with $\mu = 0$ and $\sigma = -1$ in a reactor-scale NCSX equilibrium. The equilibrium was taken from the DESC examples directory;

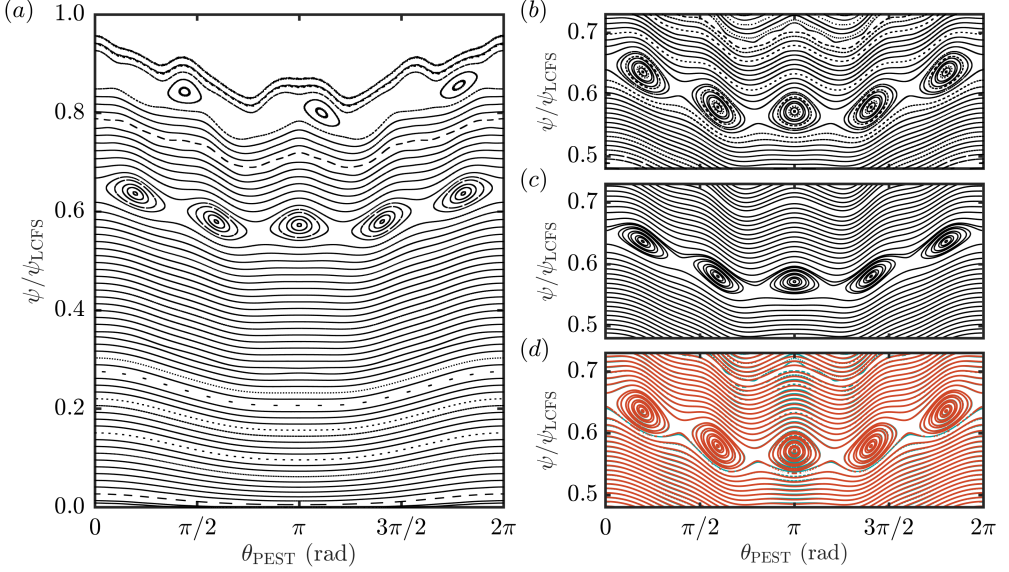


FIGURE 20. (a) Poincaré plot, at cylindrical angle $\phi = 0$, of passing alpha orbits in a reactor-scale NCSX equilibrium, simulated using ASCOT5. The alphas have energy 3.5 MeV, $\mu = 0$, and $\sigma = -1$. Multiple chains of drift islands are visible. (b) Zoomed-in view of the $(3, 5)$ drift-island chain. (c) Level sets of $\bar{\chi}_r(\psi, \eta, \zeta = 0)$, computed using (7.42) and plotted for particles in a region containing the $\iota = 3/5$ rational surface. (d) Similar to (c), except we plot the level sets of the higher-order invariant $\mathcal{H}_{2,r}(\bar{\psi}_{nr}(\psi, \eta, \zeta = 0), \bar{\eta}_{nr}(\psi, \eta, \zeta = 0))$, defined by (7.48)–(7.52), in red. We also underlay the simulation data from (b) in blue; the close agreement between the theory and simulation results is evident. In (c) and (d), the contour levels are equal to the average value of the plotted invariant for each particle in the ASCOT5 simulation shown in (b), as in figure 17.

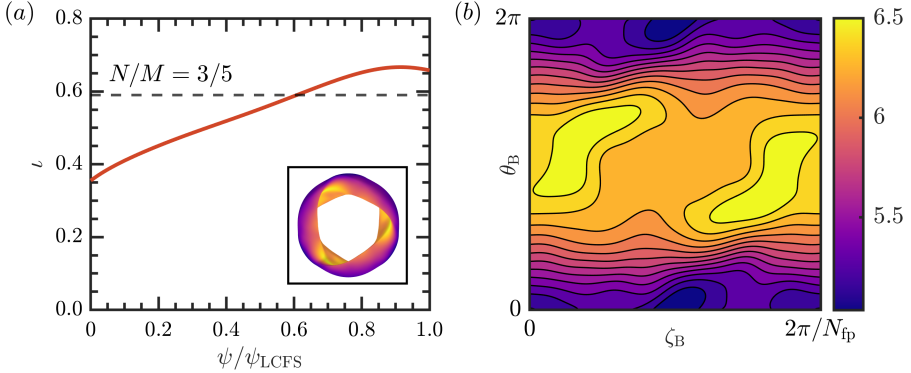


FIGURE 21. (a) Rotational transform of the NCSX equilibrium we use to trace alpha orbits in figure 20. (b) Plot of the magnetic field strength in Boozer coordinates on the rational surface with $\iota = N/M = 3/5$.

more information about it is provided in figure 21. There is a rational surface with $\iota = N/M = 3/5$ in the plasma—note that this rational surface has $N = N_{fp}$, which, according to the discussion in §4.1, promotes large drift islands. In figure 20(c), we have plotted the level sets of $\bar{\chi}_r(\psi, \eta, \zeta = 0)$, taking the low-order resonance in $\mathfrak{R}_1$ to be

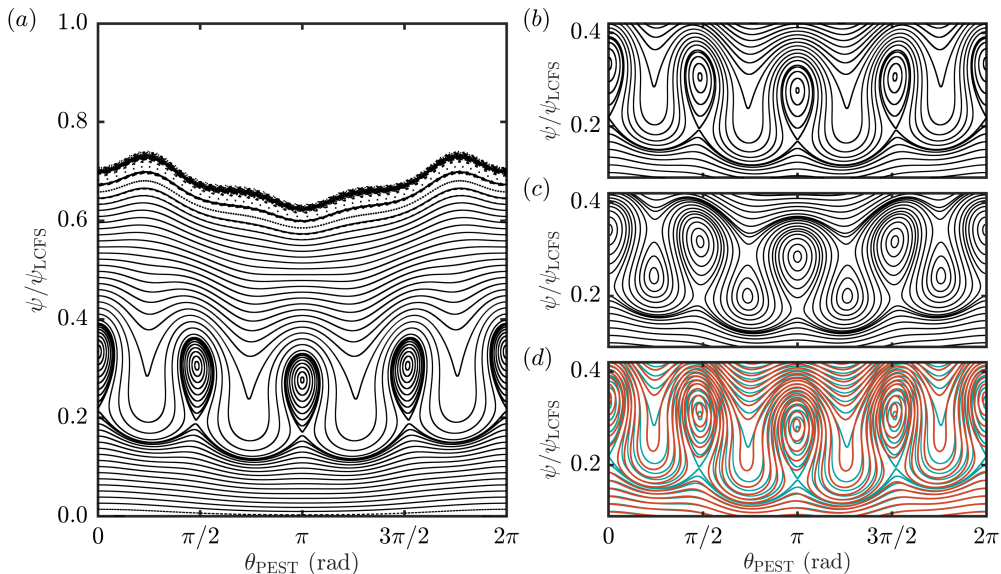


FIGURE 22. Similar to [figure 20](#) but for a reactor-scale WISTELL-A equilibrium. (a) Poincaré plot, at cylindrical angle $\phi = 0$, for passing alpha particles with energy 3.5 MeV, $\mu/\mathcal{E} = 0.1425 \text{ T}^{-1}$, and $\sigma = 1$. (b) Zoomed-in view of the prominent drift-island chain. (c) Level sets of $\bar{\chi}_r(\psi, \eta, \zeta = 0)$, plotted for particles in a region containing the $N = 1, M = 1$ resonance. (d) Similar to (c), except we plot the level sets of $\mathcal{H}_{2,r}(\bar{\psi}_{\text{nr}}(\psi, \eta, \zeta = 0), \bar{\eta}_{\text{nr}}(\psi, \eta, \zeta = 0))$, in red. We also replot the simulation data from (b) in blue to allow easy comparison. In (c) and (d), the contour levels are equal to the average value of the plotted invariant for each particle in the ASCOT5 simulation shown in (b), as in [figure 17](#).

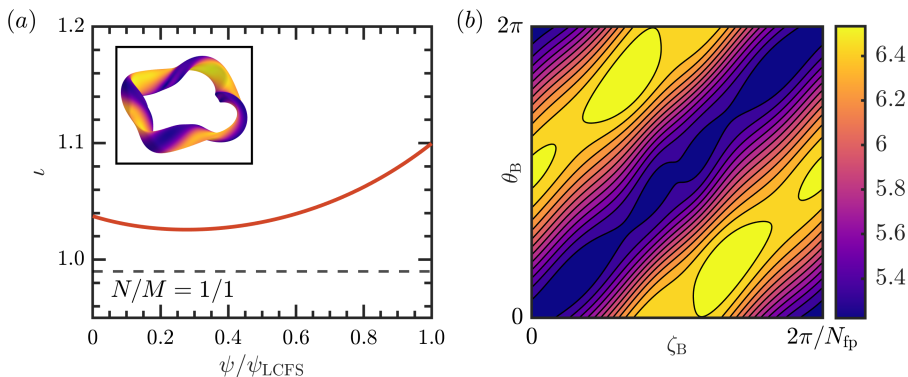


FIGURE 23. (a) Rotational transform of the WISTELL-A equilibrium we use to trace alpha orbits in [figure 22](#). (b) Plot of the magnetic field strength in Boozer coordinates on the flux surface $\psi/\psi_{\text{LCFS}} = 0.3$.

$(N, M) = (3, 5)$. [Figure 20\(d\)](#) shows the level sets of a higher-order version of the $\bar{\chi}_r$ invariant, which will be introduced in [§7.6](#). We see that conservation of $\bar{\chi}_r$ provides a very accurate prediction of the simulated alpha-particle orbits.

Similarly, [figures 22\(a\) and 22\(b\)](#) show a Poincaré plot, constructed from an ASCOT5

simulation, for barely passing, 3.5 MeV alpha particles with $\mu/\mathcal{E} = 0.1425 \text{ T}^{-1}$ and $\sigma = 1$ in a reactor-scale WISTELL-A quasi-helically symmetric equilibrium (Bader *et al.*, 2020). This equilibrium was taken from the DESC examples directory. Despite the fact that the rotational-transform profile shown in figure 23 avoids the low-order rational value $\iota = 1$, wide drift islands corresponding to an $N = M$ resonance are visible in the Poincaré plot. This occurs because the kinetic-rotational-transform profile of the co-passing orbits is lowered—relative to the magnetic rotational-transform profile shown in figure 23—by the tangential magnetic drift, as discussed in §5.3. This allows the kinetic rotational transform $\bar{\iota}$ to pass through $\bar{\iota} = 1$; equivalently, it means the $N = M$ resonance shifts into the plasma. In figure 22(c), we plot contours of $\bar{\chi}_r(\psi, \eta, \zeta = 0)$, taking the low-order resonance in $\mathfrak{R}_1$ to be $(N, M) = (1, 1)$. Figure 22(d) shows contours of the higher-order version of $\bar{\chi}_r$ that will be discussed in §7.6. Once again, the agreement between theory and simulations is excellent.

Before we move on to discuss higher-order corrections, we explain some interesting features of the drift islands in figure 22. First, these drift islands are particularly large. This may be surprising, since figure 23 shows that WISTELL-A has a relatively small quasimagnetism error. One reason for their size is that these drift islands are situated near a local minimum of the rotational-transform profile (around $\psi/\psi_{\text{LCFS}} = 0.3$; see figure 23). This means they are located in a region where the magnetic shear is small, and low shear tends to produce wider drift islands. A second reason why these islands are large is that they correspond to an $N = M = 1$ resonance. This resonance is very low order and has $\tilde{N} = N_{\text{fp}}$ (in the notation of §4.1), which promotes wide drift islands for the reasons set out in §4.1.

The drift islands in figure 22 also have an unusual shape. They do not resemble the phase portrait of a pendulum, as might be expected from the simplified lowest-order transit invariant formula (5.20) that we derived for stellarators with $s \sim \rho_*$ (this is the appropriate ordering to use here because the drift islands in figure 22 are not located close to their associated rational surface; in fact, there is no rational surface with $\iota = N/M = 1$ in the plasma). Their unusual shape is caused by the magnetic shear vanishing near the resonance, so the linear approximation $\iota(\psi) - N/M \simeq \iota'_s(\psi - \psi_s)$ —which we used to derive (5.20) for drift islands near a shifted resonant surface ψ_s —is not appropriate here. Nevertheless, provided this approximation is not used, the transit invariant (3.13), or the more complicated invariants plotted in figures 22(c) and 22(d), still correctly describe the shape of these drift islands.¹⁷ The fact that $\iota'_s \approx 0$ at the location of the drift islands means the usual island-width estimate $\Delta\psi \sim (\rho_* \delta/s)^{1/2} \Psi_t$ does not apply in this case. However, we can still estimate the island width using the method of §2.1 if we employ $\Delta\theta \sim \iota''_s(\psi - \psi_s)^2$, where $\iota''_s := (d^2\iota/d\psi^2)(\psi = \psi_s)$, for the amount by which a near-resonant orbit fails to close after one transit. With this modification, we obtain a corrected island-width scaling of $\Delta\psi \sim (\rho_* \delta \Psi_t / \iota''_s)^{1/3}$. The fact that this width scales with $\rho_*^{1/3}$, instead of the typical scaling with $\sqrt{\rho_*}$, is another way of understanding that the drift islands are wider if they arise near a local minimum of $\iota(\psi)$.

A final noteworthy feature of the drift islands in figure 22 is that they only exist for a very narrow range of pitch angles. This is because the pitch angle must be close to the trapped–passing boundary, where the tangential distance through which a particle drifts in a single transit is largest (this tangential distance diverges logarithmically as

¹⁷More generally, when an integrable Hamiltonian map is perturbed, unusual island shapes like those of figure 22 are known to arise if the unperturbed map is a ‘nontwist’ map (del-Castillo-Negrete *et al.*, 1996).

μ approaches the trapped–passing boundary). A large tangential drift is needed so that the kinetic-rotational-transform profile is lowered enough—relative to the magnetic rotational-transform profile shown in [figure 23](#)—to reach the low-order rational value $N/M = 1$. In the case shown in [figure 22](#), reducing μ by a few percent removes the drift islands, since μ is then too far from the trapped–passing boundary for the $N/M = 1$ resonance to be present in the plasma. The fact that wide drift islands can arise at extremely specific pitch angles suggests that a scan across many different pitch angles could be necessary to check that a given stellarator design does not contain any large drift islands like these.

7.6. Higher-order Hamiltonian perturbation theory

To conclude this section, we explain how these Hamiltonian calculations may be systematically extended to higher order. We begin with the higher-order calculation of orbits in $\mathfrak{R}_0$. The Lagrangian [\(7.28\)](#) that we derived in [§7.4](#) for particles in $\mathfrak{R}_0$ takes the form

$$\mathcal{L} \equiv (\bar{\psi} + \Lambda_{\theta 2}(\bar{\psi}, \bar{\theta}, \zeta)) \frac{d\bar{\theta}}{d\zeta} + \Lambda_{\psi 2}(\bar{\psi}, \bar{\theta}, \zeta) \frac{d\bar{\psi}}{d\zeta} - \bar{\chi}(\bar{\psi}) + \mathcal{H}_2(\bar{\psi}, \bar{\theta}, \zeta), \quad (7.43)$$

where $\Lambda_{\theta 2}$, $\Lambda_{\psi 2}$, and $\mathcal{H}_2$ are order $\rho_\star^2$ and contain all higher-order terms that we previously suppressed in this Lagrangian. These terms contain angle dependence which complicates the higher-order equations of motion.

We can now repeat the procedure that lead to [\(7.43\)](#). That is, we look for a near-identity change of coordinates from $(\bar{\psi}, \bar{\theta})$ to a new set $(\bar{\bar{\psi}}, \bar{\bar{\theta}})$, defined implicitly by

$$\bar{\psi} = \bar{\bar{\psi}} + \psi^{(2)}(\bar{\bar{\psi}}, \bar{\bar{\theta}}, \zeta), \quad (7.44a)$$

$$\bar{\theta} = \bar{\bar{\theta}} + \theta^{(2)}(\bar{\bar{\psi}}, \bar{\bar{\theta}}, \zeta), \quad (7.44b)$$

where $\psi^{(2)} \sim \rho_\star^2 \Psi_t$ and $\theta^{(2)} \sim \rho_\star^2$. We will choose $\psi^{(2)}$ and $\theta^{(2)}$ so that any dependence on $\bar{\bar{\theta}}$ and ζ in the Lagrangian is removed—or rather, pushed to even higher order. We will need to add the total ζ derivative of a new generating function $S^{(2)} \sim \rho_\star^2 \Psi_t$ to achieve this. Then, with a suitable choice of $\psi^{(2)}$ and $\theta^{(2)}$, the Lagrangian can be put in the form

$$\mathcal{L} \equiv \bar{\bar{\psi}} \frac{d\bar{\bar{\theta}}}{d\zeta} - \bar{\bar{\chi}} + V_2 - (\partial_\zeta S^{(2)} + \iota \partial_\theta S^{(2)}), \quad (7.45)$$

for some $V_2(\bar{\bar{\psi}}, \bar{\bar{\theta}}, \zeta) \sim \rho_\star^2 \Psi_t$ whose precise form is not important for understanding the general method. In [\(7.45\)](#), we are suppressing angle-dependent terms of order $\rho_\star^3 \Psi_t$.

As before, in order to avoid resonant denominators in $\mathfrak{R}_0$, we can split $V_2 = V_2^{<K_2} + V_2^{>K_2}$, for some new cutoff K_2 that makes $V_2^{>K_2} \sim \rho_\star^3 \Psi_t$ (typically, K_2 will be larger than the cutoff K that we introduced at the previous order). Then, we choose $S^{(2)}$ to solve the magnetic differential equation

$$\partial_\zeta S^{(2)} + \iota \partial_\theta S^{(2)} = V_2^{<K_2} - \langle V_2^{<K_2} \rangle. \quad (7.46)$$

We can solve for $S^{(2)}$ as a Fourier series as long as $\mathfrak{R}_0$ does not contain any resonances with harmonics of $V_2^{<K_2}$. The final result is a Lagrangian of the form

$$\mathcal{L} \equiv \bar{\bar{\psi}} \frac{d\bar{\bar{\theta}}}{d\zeta} - \bar{\bar{\chi}}(\bar{\bar{\psi}}), \quad (7.47)$$

where $\bar{\bar{\chi}} := \bar{\chi} - \langle V_2 \rangle$. The equations of motion tell us that $\bar{\bar{\psi}}(\psi, \theta, \zeta)$ is approximately

conserved; its ζ derivative is $d\bar{\psi}/d\zeta \sim \rho_\star^3 \Psi_t$, so $\bar{\psi}$ is conserved more accurately than the quantity $\bar{\psi}$ that we constructed at first order in (7.30). Therefore, $\bar{\psi}(\psi, \theta, \zeta) = \text{const.}$ is a more accurate equation for the drift surfaces of passing particles in a region of sufficiently irrational rotational transform.

This procedure can be iterated to obtain more and more accurate equations for the drift surfaces. However, it cannot be continued indefinitely. Eventually, for any given $\mathfrak{R}_0$, we will reach an order at which the angle-dependent perturbation $V_n \sim \rho_\star^n \Psi_t$ in the Lagrangian contains a harmonic that resonates with a rational surface in $\mathfrak{R}_0$ and has an amplitude of order $\rho_\star^n \Psi_t$. Such a harmonic is too large to be ‘hidden’ in a higher-order remainder term. From this point onwards, we need to use the approach of §7.5 to describe motion in a region containing a single resonant surface. If we wish to keep going to higher order beyond this point, then we need to know how to continue the calculation of particle orbits in $\mathfrak{R}_1$ to higher order.¹⁸

The higher-order calculation of orbits in $\mathfrak{R}_1$ proceeds in a similar manner: we iteratively look for near-identity coordinate transformations that simplify the Lagrangian. The difference is that we look for transformations that eliminate all angle dependence except for dependence on the resonant combination $\eta = \theta - (M/N)\zeta$. At every order, we will obtain an integrable Lagrangian of the form (7.40), which gives a higher-order conserved quantity generalizing (7.42).

The calculation of orbits in $\mathfrak{R}_1$ is carried out to next order in §J. In this calculation, we neglect the order- $\rho_\star^2$ correction to Littlejohn’s guiding-centre Lagrangian (7.1). Neglecting this term means we are calculating higher-order corrections to the orbit that would be observed in a guiding-centre code, such as ASCOT5, that uses equations derived from the usual guiding-centre Lagrangian (without higher-order corrections). The results of this calculation are summarized below.

The higher-order conserved quantity can be expressed in the form

$$\mathcal{H}_{2,r}(\bar{\psi}_{\text{nr}}, \bar{\eta}_{\text{nr}}) := \chi_r(\bar{\psi}_{\text{nr}}) - \langle V \rangle_\zeta(\bar{\psi}_{\text{nr}}, \bar{\eta}_{\text{nr}}) - \langle V_2 \rangle_\zeta(\bar{\psi}_{\text{nr}}, \bar{\eta}_{\text{nr}}) = \text{const.}, \quad (7.48)$$

where $\bar{\eta}_{\text{nr}} = \bar{\theta}_{\text{nr}} - (N/M)\zeta$ and V_2 is defined by

$$V_2 := V_{\text{nr}}^{\geq K} + (\psi_{\text{nr}}^{(1)} \partial_\psi V_r + \theta_{\text{nr}}^{(1)} \partial_\theta V_r) + \frac{1}{2} \iota' (\psi_{\text{nr}}^{(1)})^2 - \psi_{\text{nr}}^{(1)} (\partial_\zeta \theta_{\text{nr}}^{(1)} + \iota \partial_\theta \theta_{\text{nr}}^{(1)}). \quad (7.49)$$

The new coordinates $\bar{\psi}_{\text{nr}}$ and $\bar{\eta}_{\text{nr}}$ appearing in (7.48) are defined by

$$\bar{\psi}_{\text{nr}} = \bar{\psi}_{\text{nr}} + \psi_{\text{nr}}^{(2)}(\bar{\psi}_{\text{nr}}, \bar{\theta}_{\text{nr}}, \zeta), \quad (7.50a)$$

$$\bar{\theta}_{\text{nr}} = \bar{\theta}_{\text{nr}} + \theta_{\text{nr}}^{(2)}(\bar{\psi}_{\text{nr}}, \bar{\theta}_{\text{nr}}, \zeta), \quad (7.50b)$$

where

$$\psi_{\text{nr}}^{(2)} := \psi_{\text{nr}}^{(1)} \left[\partial_\theta \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\psi \mathbf{x} \right) - \partial_\psi \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\theta \mathbf{x} \right) \right] + \theta_{\text{nr}}^{(1)} \partial_\theta \psi_{\text{nr}}^{(1)} + \partial_\theta S_{\text{nr}}^{(2)}, \quad (7.51a)$$

$$\theta_{\text{nr}}^{(2)} := \theta_{\text{nr}}^{(1)} \left[\partial_\theta \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\psi \mathbf{x} \right) - \partial_\psi \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\theta \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_\psi \psi_{\text{nr}}^{(1)} - \partial_\psi S_{\text{nr}}^{(2)}, \quad (7.51b)$$

¹⁸For completeness, we point out another issue that may be encountered at higher orders. At each order, we calculate the corrections $\psi^{(n)}$, $\theta^{(n)}$, and $S^{(n)}$, which are assumed to be smaller, by a factor of $\rho_\star \ll 1$, than $\psi^{(n-1)}$, $\theta^{(n-1)}$, and $S^{(n-1)}$, respectively. These corrections typically contain coefficients that grow rapidly with n . Therefore, n will eventually become large enough that the n th-order corrections are no smaller than the corrections from the previous order, violating our orderings.

and the generating function is

$$S_{\text{nr}}^{(2)} := \sum_{(p,q) \in \mathbb{I}_{\text{nr}}^{<K_2}} \frac{i(V_{2,\text{nr}}^{<K_2})_{pq}}{q - \iota p} e^{i(p\theta - q\zeta)}, \quad (7.52)$$

for $V_{2,\text{nr}}^{<K_2}$ defined by $V_2 = \langle V_2 \rangle_\zeta + V_{2,\text{nr}}^{<K_2} + V_{2,\text{nr}}^{\geq K_2}$.

We compare this higher-order theory with ASCOT5 simulations in [figures 20\(d\) and 22\(d\)](#). The agreement between the theory and simulations is striking, and consistent across all orbits in all equilibria that we have examined. Thus, perturbative calculations based on an asymptotic expansion in $\rho_\star$ are able to describe the orbits of passing alpha particles in reactor-scale stellarators with impressive accuracy, even at 3.5 MeV.

7.7. Discussion

The calculation in this section was motivated, in part, by the need to find a clear criterion for which flux surfaces count as ‘rational’ for the purposes of calculating passing-particle orbits; that is, around which rational surfaces must particle orbits be described using the transit adiabatic invariant? There are infinitely many rational surfaces, so it cannot be the case that we can choose any one of them to be ψ_r .

The answer obtained in this section is that the transit adiabatic invariant describes particle orbits around those rational surfaces that resonate with low-order harmonics of $V = (mc/Ze)(v_\parallel/\hat{\mathbf{b}} \cdot \nabla \zeta)$. These harmonics depend on the velocity-space coordinates $\mathcal{E}$ and μ , so it may be possible to treat a flux surface as ‘irrational’ for one particle despite the fact that it must be treated as ‘rational’ for another. As we continue the calculation to higher order in $\rho_\star \ll 1$, the set of rational surfaces that give rise to a resonant denominator gets larger and larger; higher-order calculations are able to detect smaller and smaller drift islands. Therefore, the answer to which flux surfaces count as ‘rational’ also depends on the order of accuracy of our calculation. This makes sense physically: at higher order, we compute the particle orbits more accurately and we can, therefore, follow a given particle for a longer time. As a result, higher-order calculations can detect orbits that close on themselves after longer times, which occur on higher-order rational surfaces.

8. Conclusions

In this paper, we have calculated the guiding-centre orbits of passing particles in a general stellarator. These orbits contain drift islands around rational flux surfaces, as observed in recent simulations ([White et al., 2022](#); [White, 2022, 2025](#)). We derived simple equations for the drift-island shape and used them to investigate how the stellarator magnetic field might be designed to ensure that these islands are small. Our key findings are as follows.

- (i) We characterized stellarator magnetic fields in which the drift-island width vanishes for all passing particles, finding that they must be *cyclometric* (see [§5.1](#)). Cyclometry is similar to the well-known [Cary & Shasharina \(1997a\)](#) condition for omnigenity, but is less restrictive. Therefore, to remove a set of drift islands around a rational surface, optimizing for cyclometry on that surface may be more compatible with other objectives than simply optimizing for better omnigenity.
- (ii) The drift-island width usually scales as $w \sim (\rho_\star \delta/s)^{1/2} a$, where $\rho_\star$ is the energetic-particle gyroradius divided by a typical length scale of the field, δ is the deviation from cyclometry, s is the magnetic shear, and a is the minor radius. Therefore, wide

drift islands may arise in stellarators that are insufficiently cyclometric and have low magnetic shear. The drift islands can shift away from their associated rational surface by a radial distance $\sim (\rho_*/s)a$. Therefore, when there is a low-order rational surface at the plasma edge, as is the case in stellarators with an island divertor, wide drift islands may shift into the plasma for either co- or counter-passing alphas.

- (iii) Passing particles near a low-order rational surface conserve the *transit adiabatic invariant*, defined in (3.13). We calculated higher-order corrections to the transit invariant, which become more important at higher energies. The resulting theory agrees extremely well with guiding-centre and full-orbit simulations conducted using ASCOT5. This suggests that the transit invariant and orbit-averaged guiding-centre equations can accurately describe the motion of passing alpha particles in reactor-scale stellarators, even at 3.5 MeV.
- (iv) We showed that the orbits of passing particles in a general stellarator, on both rational and irrational surfaces, can be systematically calculated using Hamiltonian perturbation theory. These calculations match simulations almost perfectly (e.g. figure 20), indicating that passing alpha orbits in stellarator reactors can be modelled using an asymptotic expansion in $\rho_* \ll 1$, even at the birth energy.

Acknowledgements

We are grateful to Richard Nies, Xu Chu and Matt Kunz for many helpful conversations.

Funding

This manuscript has been authored by Princeton University/Princeton Plasma Physics Laboratory under contract number DE-AC02-09CH11466 with the U.S. Department of Energy. The United States Government retains a non-exclusive, paid-up, irrevocable, world-wide licence to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purpose(s). This research was supported in part by Grants No. PID2021-123175NB-I00 and PID2024-155558OB-I00, funded by Ministerio de Ciencia, Innovación y Universidades/Agencia Estatal de Investigación/10.13039/501100011033 and by ERDF/EU.

Appendix A. Comparison of guiding-centre and full-orbit simulations

To verify that the guiding-centre model accurately describes the orbits of alpha particles in a stellarator reactor, we compared some of our guiding-centre simulations with full-orbit ASCOT5 simulations. This is particularly important in light of recent concerns that the guiding-centre model may not be valid for energetic alphas (Burby *et al.*, 2025). figures 24 and 25 show side-by-side Poincaré plots created from the guiding-centre and full-orbit simulations. To construct the Poincaré plot in the full-orbit cases, we used ASCOT5 to compute the coordinates of the guiding centre from the particle coordinates at each timestep and recorded a point in the plot whenever the guiding centre crossed the Poincaré plane. These figures show excellent agreement between the guiding-centre and full-orbit simulations. The agreement was equally good in other equilibria and for other particle energies and pitch angles (for passing particles and trapped particles away from the trapped-passing boundary) that we checked.

These results indicate that observed discrepancies between the alpha-particle losses in

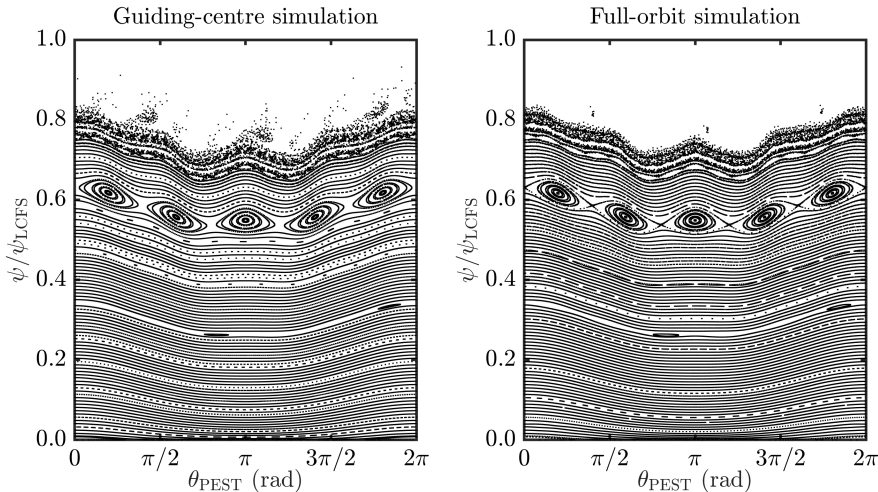


FIGURE 24. Comparison of guiding-centre and full-orbit simulations for passing alpha particles with energy 3.5 MeV, $\mu/\mathcal{E} = 0.096 \text{ T}^{-1}$, and $\sigma = 1$, in a reactor-scale NCSX equilibrium. We provide more detail about this equilibrium in [figure 21](#).

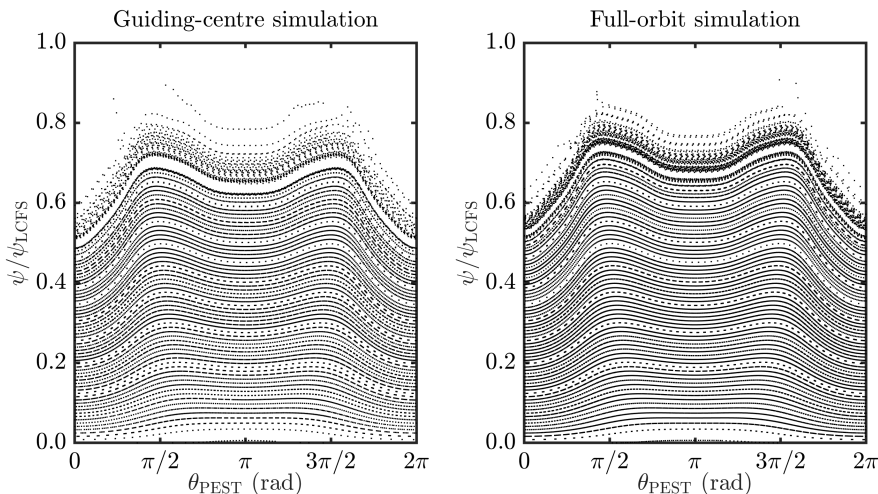


FIGURE 25. Comparison of guiding-centre and full-orbit simulations for trapped alpha particles with energy 3.5 MeV and $\mu/\mathcal{E} = 0.177 \text{ T}^{-1}$, in a CIEMAT-QI equilibrium with $N_{\text{fp}} = 4$ and $\beta \sim 4\%$. CIEMAT-QI is a modern quasi-isodynamic stellarator design ([Sánchez et al., 2023](#); [Velasco et al., 2023](#); [García-Regaña et al., 2024](#)); we used it for the trapped-particle comparison because trapped alphas are rapidly lost from the other, less well-optimized reactor-scale equilibria considered in this article. To construct these Poincaré plots for trapped particles, we recorded a point whenever the guiding centre crossed the $\phi = \pi/4$ plane (ϕ is cylindrical angle; this plane is roughly midway between the particles' bounce points) with $v_{\parallel} > 0$.

guiding-centre and full-orbit simulations (e.g. [de Assunção, 2023](#); [Rodrigues et al., 2024](#); [Carbajal et al., 2025](#)) are not caused by breakdown of the perturbative guiding-centre expansion. Instead, our results support the explanation proposed by [Rodrigues et al. \(2024\)](#) for the primary cause of these discrepancies: full-orbit simulations should lose additional particles through the last closed flux surface (LCFS) because the guiding centre of a particle can remain inside the LCFS even though the particle itself crosses the LCFS

as it completes its gyro-orbit. The additional lost particles are those whose orbits carry their guiding centre within roughly a gyroradius of the LCFS; note that such particles can be born on flux surfaces far from the edge if their orbit width is large. Although the number of alpha particles that satisfy this condition is small when $\rho_\star \ll 1$ (relative to the total number of alphas in the device), the alpha losses could still be enhanced by a considerable factor in cases where the losses in a guiding-centre simulation are very small to begin with, as might be the case for a well-optimized stellarator.

Appendix B. Relaxing the assumption of flux surfaces

Throughout the paper, we assume that the magnetic field possesses nested toroidal flux surfaces. The reader may be concerned that this assumption is inappropriate for particles near rational surfaces, since rational surfaces can easily be broken into magnetic islands by small magnetic perturbations. In this appendix, we argue that the assumption of nested flux surfaces is not a problem for sufficiently energetic particles.

Our orbit calculations remain accurate in the presence of magnetic islands around the rational surface ψ_r provided the magnetic-island width is much smaller than the drift-island width. The physical reason is that particles circulating around a drift island spend most of their orbit in regions where the flux surfaces are intact. To make this argument more rigorous, we proceed to explain how our calculations can be modified to include the effect of magnetic islands whose width is comparable to the drift-island width.

Suppose the magnetic field is $\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1$, where $\mathbf{B}_0$ possesses nested toroidal flux surfaces and $\mathbf{B}_1$ is a small island-generating perturbation. To make the magnetic-island width comparable to the drift-island width, we order $\mathbf{B}_1 \sim \rho_\star \mathbf{B}_0$. Then, for particles a distance $\psi - \psi_r \sim (\rho_\star/s)^{1/2} \Psi_t$ from the rational surface, the equations of motion for the guiding centre, (3.2), become

$$\frac{d\psi}{dt} = v_\parallel \frac{\mathbf{B}_1}{B_0} \cdot \nabla \psi + \mathbf{v}_d \cdot \nabla \psi, \quad (\text{B } 1a)$$

$$\frac{d\eta}{dt} = v_\parallel \hat{\mathbf{b}} \cdot \nabla \eta, \quad (\text{B } 1b)$$

$$\frac{d\zeta}{dt} = v_\parallel \hat{\mathbf{b}} \cdot \nabla \zeta, \quad (\text{B } 1c)$$

to leading order. Here, the magnetic coordinates (ψ, θ, ζ) are constructed from the unperturbed field $\mathbf{B}_0$, as are $B_0 = |\mathbf{B}_0|$, $\hat{\mathbf{b}} = \mathbf{B}_0/B_0$, and

$$v_\parallel(\mathbf{x}, \mathcal{E}, \mu, \sigma) = \sigma \sqrt{2(\mathcal{E} - \mu B_0)}, \quad (\text{B } 1d)$$

$$\mathbf{v}_d(\mathbf{x}, \mathcal{E}, \mu) = \frac{v_\parallel}{B_0} \nabla \times \left(\frac{v_\parallel \mathbf{B}_0}{\Omega_0} \right), \quad (\text{B } 1e)$$

where $\Omega_0 = ZeB_0/mc$. To this order, the same equations would arise if the guiding centre moved along field lines of

$$\mathbf{B}^\star = \mathbf{B}_0 + \underbrace{\mathbf{B}_1}_{\sim \rho_\star \mathbf{B}_0} + \underbrace{\nabla \times \left(\frac{v_\parallel \mathbf{B}_0}{\Omega_0} \right)}_{\sim \rho_\star \mathbf{B}_0}. \quad (\text{B } 2)$$

We see that the drift-island-generating perturbation in $\mathbf{B}^\star$ now consists of two parts: a part which is simply the magnetic perturbation $\mathbf{B}_1$, and a part due to the magnetic drift caused by $\mathbf{B}_0$.

We can calculate the drift-island shape from (B 1) by following the same steps as in §3. The resulting island-shape is the same as (3.22) but with $I_r(\eta, \mathcal{E}, \mu)$ replaced by

$$I_r^{\text{dm}}(\eta, \mathcal{E}, \mu, \sigma) := \frac{Ze\sigma}{mc} \oint \mathbf{A}_1 \cdot \partial_\zeta \mathbf{x} d\zeta + \oint |v_\parallel| \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x} d\zeta, \quad (\text{B } 3)$$

where $\mathbf{B}_1 = \nabla \times \mathbf{A}_1$. Therefore, the particle orbit traces out islands whose shape is determined by a combination of drifts and the magnetic perturbation; we call these ‘drift-magnetic’ islands. Higher-order corrections to the shape of the drift-magnetic islands may be derived using the methods of §6 and §7.

Using (B 3), we can investigate when it is legitimate to assume that the field has nested toroidal flux surfaces for the purposes of calculating particle orbits around a rational surface ψ_r . This assumption is legitimate when the second term in (B 3) dominates the dependence of I_r^{dm} on η . This occurs when $|\mathbf{B}_1| \ll \rho_* |\mathbf{B}_0|$ or, more precisely, when the harmonics of $\hat{\mathbf{b}} \cdot \mathbf{A}_1 / \hat{\mathbf{b}} \cdot \nabla \zeta$ that are resonant at ψ_r are much smaller than the corresponding harmonics of $(mc/Ze)(v_\parallel / \hat{\mathbf{b}} \cdot \nabla \zeta)$.

Appendix C. Proof that field-strength contours cannot intersect field lines tangentially in cyclometric magnetic fields

In §5.1.2, we state that, if the magnetic field is cyclometric on a rational surface ψ_r , then there cannot be tangential intersections between field-strength contours and magnetic field lines on ψ_r . In this appendix, we provide a proof of this claim, assuming the field strength is a smooth, well-behaved function.

We will use the following coordinates. All quantities will be evaluated at the rational surface ψ_r ; therefore, we suppress any ψ arguments. We describe position on ψ_r using coordinates η and l , where l is arc length along the field lines. We use ∂_l and ∂_η to denote partial derivatives with respect to one of l or η , respectively, with the other held fixed.

Suppose a field-strength contour and a field line intersect tangentially, as shown in figure 26. At the intersection point, we have (1) $\partial_l B = 0$ and, provided the intersection point does not happen to be a stationary point of B on the flux surface, (2) $\partial_\eta B \neq 0$. We will prove that, in a cyclometric field, $\partial_\eta B = 0$ at any point where $\partial_l B = 0$. This means there cannot be tangential intersections between field-strength contours and field lines in a cyclometric field, since it is not possible to have tangential intersections only at the stationary points of B .

To prove that $\partial_\eta B = 0$ wherever $\partial_l B = 0$, our strategy will be as follows. We will consider two rational-surface field lines labeled by $\eta = \eta_*$ and $\eta = \eta_* + \delta\eta$, where $|\delta\eta| \ll 1$, as well as two values attained by the magnetic field strength on these field lines, namely $B = B_*$ and $B = B_* + \delta B$, where δB is small. We will derive an expression for $\delta L := L(\eta_* + \delta\eta, B_* + \delta B) - L(\eta_*, B_* + \delta B)$. In a cyclometric field, δL must be independent of $\delta\eta$, for all δB . We will find that this implies $\partial_\eta B = 0$ at each point where $\partial_l B = 0$ along η_* .

First, we introduce some helpful notation. Along any rational-surface field line η , there will be a number of points at which the magnetic field strength takes a given value B . As in §5.1.2, we write $l_i(\eta, B)$ for the l -coordinates of these points, which we label using a discrete index i . Also, we set $l_{i*} := l_i(\eta_*, B_*)$.

Suppose $\partial_l B(\eta_*, l_{i*})$ is non-zero for some i . Then, for sufficiently small $\delta\eta$ and δB , every

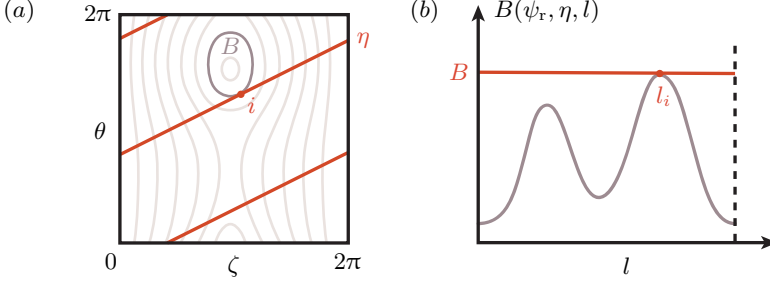


FIGURE 26. (a) Plot, in straight-field-line coordinates, of a field line (shown in red) that has a tangential intersection with a field-strength contour. At the intersection point i , we must have $\partial_l B = 0$. (b) Sketch of the field strength as a function of distance along this field line.

point $l_i(\eta_*, B_* + \delta B)$ on field line η_* where the field strength is $B_* + \delta B$ is accompanied by a nearby point $l_i(\eta_* + \delta\eta, B_* + \delta B)$ on field line $\eta_* + \delta\eta$ where the field strength is also $B_* + \delta B$; we assume these nearby points are indexed by the same i . Then, their l -coordinates are related by

$$l_i(\eta_* + \delta\eta, B_* + \delta B) - l_i(\eta_*, B_* + \delta B) \simeq -\frac{\partial_\eta B(\eta_*, l_{i*})}{\partial_l B(\eta_*, l_{i*})} \delta\eta. \quad (\text{C } 1)$$

When $\partial_l B(\eta_*, l_{i*})$ is non-zero for all l_{i*} , we can compute δL by summing (C 1) over i , and accounting for whether the change in l_i increases the size of a well (contributing positively to δL) or decreases the size of a well (contributing negatively to δL). We find

$$\delta L = - \sum_{i \in \mathcal{P}(\eta_*, B_*)} \frac{\partial_\eta B(\eta_*, l_{i*})}{|\partial_l B(\eta_*, l_{i*})|} \delta\eta, \quad (\text{C } 2)$$

where $\mathcal{P}(\eta_*, B_*)$ denotes the set of points along field line η_* where the field strength is B_* .

If $\partial_l B(\eta_*, l_{i*}) = 0$ at any of the points l_{i*} , then these points will need to be treated separately, since (C 1) is incorrect for them. These points are the locations of local minima or maxima when B is plotted against l along field line η_* . For simplicity, we do not consider magnetic fields with points of inflection when B is plotted against l , although it is not difficult to generalize (C 4) below to include such points.

Consider a point l_{i*} with $\partial_l B(\eta_*, l_{i*}) = 0$ and suppose that it is a local maximum, meaning $\partial_l^2 B(\eta_*, l_{i*}) < 0$. Then, there are two wells on either side of l_{i*} that touch at l_{i*} . If the field strength is changed from B_* to $B_* + \delta B$, there are two possibilities, as shown in figure 27. If $\delta B > 0$, then the two wells on either side of l_{i*} merge into one larger well. In this case, the point l_{i*} does not smoothly connect to another point in $\mathcal{P}(\eta_*, B_* + \delta B)$. On the other hand, if $\delta B < 0$, then the two wells on either side of l_{i*} move apart. In this case, the point l_{i*} splits into two points in $\mathcal{P}(\eta_*, B_* + \delta B)$, separated by a region where $B(\eta_*, l) > B_* + \delta B$; we will use i_- and i_+ to index these two points. The distance between them is

$$l_{i_+}(\eta_*, B_* + \delta B) - l_{i_-}(\eta_*, B_* + \delta B) \simeq 2\sqrt{\frac{2\delta B}{\partial_l^2 B(\eta_*, l_{i*})}}. \quad (\text{C } 3)$$

Now suppose l_{i*} is a local minimum of B along η_* . We can think of such a point as being

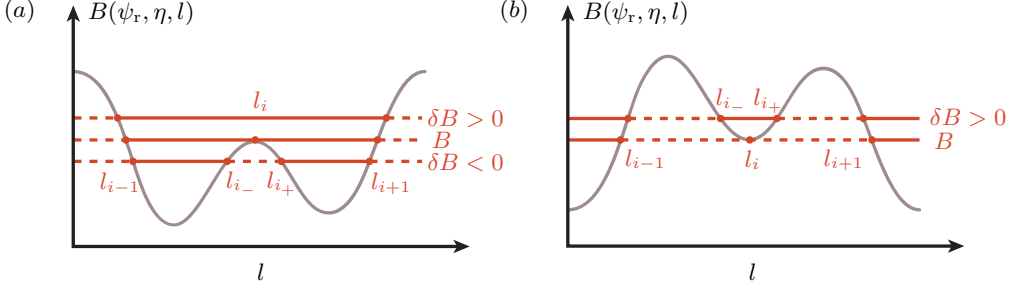


FIGURE 27. Sketch of how the wells (indicated by solid red lines) corresponding to a field strength B change when B is slightly increased ($\delta B > 0$) or decreased ($\delta B < 0$). In each case, there is a point l_i at which the field strength is B and $\partial_l B = 0$. (a) In this case, $\partial_l^2 B < 0$ at l_i . If $\delta B > 0$, the wells on either side of l_i merge. If $\delta B < 0$, these wells move apart. (b) In this case, $\partial_l^2 B > 0$ at l_i . If $\delta B > 0$, a new well opens up. If $\delta B < 0$, there will be no well here.

the location of an infinitesimal well. Then, when $\delta B < 0$, the well disappears. When $\delta B > 0$, the well widens and l_{i*} splits into two points l_{i-} and l_{i+} , which also satisfy (C3).

The local minima or maxima behave in a similar way when we simultaneously change the magnetic field by δB and the field-line label by $\delta \eta$. Therefore, the generalization of (C2) to the case where some l_i are local maxima or minima is

$$\begin{aligned} \delta L \simeq & - \sum_{i \in \mathcal{Q}(\eta_*, B_*)} \frac{\partial_\eta B(\eta_*, l_{i*})}{|\partial_l B(\eta_*, l_{i*})|} \delta \eta \\ & - \sum_{i \in \mathcal{P}_M(\eta_*, B_*)} \left[2\sqrt{\Xi\left(\frac{2[\delta B - \delta \eta \partial_\eta B(\eta_*, l_{i*})]}{\partial_l^2 B(\eta_*, l_{i*})}\right)} - 2\sqrt{\Xi\left(\frac{2\delta B}{\partial_l^2 B(\eta_*, l_{i*})}\right)} \right] \\ & + \sum_{i \in \mathcal{P}_m(\eta_*, B_*)} \left[2\sqrt{\Xi\left(\frac{2[\delta B - \delta \eta \partial_\eta B(\eta_*, l_{i*})]}{\partial_l^2 B(\eta_*, l_{i*})}\right)} - 2\sqrt{\Xi\left(\frac{2\delta B}{\partial_l^2 B(\eta_*, l_{i*})}\right)} \right]. \quad (\text{C4}) \end{aligned}$$

In (C4), $\Xi(x) := x\Theta(x)$, where $\Theta(x)$ is the Heaviside step function, and we write $\mathcal{P}_M(\eta, B)$ for the subset of $\mathcal{P}(\eta, B)$ that are local maxima, $\mathcal{P}_m(\eta, B)$ be the subset of $\mathcal{P}(\eta, B)$ that are local minima, and $\mathcal{Q}(\eta, B)$ for all other points in $\mathcal{P}(\eta, B)$.

In (C4), for sufficiently small $\delta \eta$ and δB , the terms linear in $\delta \eta$ are negligible compared to the terms containing square roots. Assuming $B_* < B_M$, we can choose δB to be a large enough positive number that $\delta B - \delta \eta \partial_\eta B(\eta_*, l_{i*}) > 0$ for all i in $\mathcal{P}_M(\eta_*, B_*) \cup \mathcal{P}_m(\eta_*, B_*)$, over a range of $\delta \eta$. Then,

$$\delta L \simeq \sum_{i \in \mathcal{P}_m(\eta_*, B_*)} \left[2\sqrt{\frac{2[\delta B - \delta \eta \partial_\eta B(\eta_*, l_{i*})]}{\partial_l^2 B(\eta_*, l_{i*})}} - 2\sqrt{\frac{2\delta B}{\partial_l^2 B(\eta_*, l_{i*})}} \right]. \quad (\text{C5})$$

For δL to be independent of $\delta \eta$ within this range, we need

$$\frac{\partial \delta L}{\partial \delta \eta} \bigg|_{\delta B} = \sum_{i \in \mathcal{P}_m(\eta_*, B_*)} \frac{-\sqrt{2} \partial_\eta B(\eta_*, l_{i*})}{\sqrt{[\partial_l^2 B(\eta_*, l_{i*})][\delta B - \delta \eta \partial_\eta B(\eta_*, l_{i*})]}} = 0. \quad (\text{C6})$$

Since $1/\sqrt{\delta B - \delta \eta \partial_\eta B(\eta_*, l_{i*})}$ are linearly independent functions of $\delta \eta$ and δB when all the values of $\partial_\eta B(\eta_*, l_{i*})$ are different from each other, (C6) requires $\partial_\eta B(\eta_*, l_{i*}) = 0$ at every local minimum.

Similar reasoning, with δB chosen to be sufficiently negative, allows us to conclude that $\partial_\eta B(\eta_*, l_{i*}) = 0$ at every local maximum. Therefore, $\partial_\eta B = 0$ at all the points where $\partial_l B = 0$, which is what we set out to prove.

Appendix D. Proof that low-shear ($s \sim \rho_*$) stellarators without drift islands must be cyclometric

In §5.1, we prove that cyclometry is the necessary and sufficient condition for there to be no drift islands, for all passing particles, at a given rational surface ψ_r . Our proof uses the fact that a particle orbit contains no drift islands if and only if its transit-averaged radial drift vanishes on ψ_r . This follows from (3.22), which is valid as long as any resonant (closed) passing orbits, around which drift islands could form, are located near ψ_r .

In stellarators with a very low magnetic shear, however, the drift islands can shift significant radial distances away from ψ_r , as discussed in §5.3. When the shear is as low as $s \sim \rho_*$, the shifted position of the drift islands depends on $\mathcal{E}$ and μ . In order to remove the drift islands at a given flux surface ψ , it is only necessary to ensure that the transit-averaged radial drift vanishes *for those passing particles that have closed orbits on this flux surface*. This is less stringent than requiring that the transit-averaged radial drift vanishes on this flux surface for all passing particles. Therefore, one might wonder whether cyclometry is a stronger condition than necessary for the elimination of drift islands from a stellarator with $s \sim \rho_*$.

In this appendix, we extend the arguments of §5.1 to stellarators with $s \sim \rho_*$ and show that, to avoid drift islands in these devices, cyclometry is still required. In fact, stellarators with $s \sim \rho_*$ need to be cyclometric throughout a volume—on all flux surfaces that the drift islands could shift to—rather than on a single rational surface.

As usual, we assume the rotational transform is close to N/M . Then, as discussed in §3.4, passing orbits trace out level sets of the transit invariant $\mathcal{I}(\psi, \eta, \mathcal{E}, \mu, \sigma)$. Drift islands are present when $\mathcal{I}$ has isolated local extrema in the (η, ψ) plane, which represent drift-island O-points, as shown in figure 28(a). Therefore, we must ensure that either

- (i) $\mathcal{I}$ has no local extrema in the plasma, or
- (ii) $\mathcal{I}$ only attains its extremal values on continuous curves in the (η, ψ) plane, as shown in figure 28(b), and never at isolated points.

Local extrema of $\mathcal{I}$ represent resonant passing orbits that close on themselves after precisely M toroidal and N poloidal turns (this can be seen from (3.12)). Condition (i) ensures that there are no such closed orbits in the plasma, which is a trivial way to avoid any associated drift islands. We are interested in stellarators that do contain such closed orbits. In such devices, condition (ii) must be satisfied, which ensures the closed orbits are surrounded by topologically toroidal drift surfaces instead of drift islands. We proceed to investigate what this implies for the magnetic field.

For this discussion, it will be convenient to change our velocity-space coordinates: instead of $\mathcal{E}$ and μ , we will employ the particle speed v , defined by $\mathcal{E} = v^2/2$, and the pitch-angle coordinate $\lambda := \mu/\mathcal{E}$. The key idea is that, whereas in §5.1 the cyclometry condition follows from the fact that the drift-island width vanishes for a range of values of λ , here (for stellarators with $s \sim \rho_*$) we must use the fact that the drift-island width vanishes even as we vary λ and v simultaneously.

Let $\psi = \psi_s(\eta, v, \lambda, \sigma)$ be the curve in the (η, ψ) plane on which $\mathcal{I}$ attains its extremal

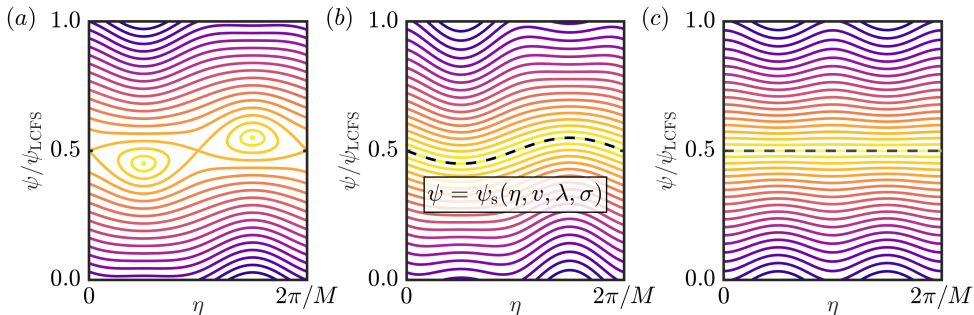


FIGURE 28. Three possible topologies of the level sets of $\mathcal{I}$ in a contour plot in the (η, ψ) plane. (a) Typically, $\mathcal{I}$ will have isolated extrema, which represent the O-points of the drift islands. (b) A case where $\mathcal{I}$ reaches its extremal value along a continuous curve. (c) Here, $\mathcal{I}$ attains its extremal value on the flux surface $\psi/\psi_{\text{LCFS}} = 0.5$.

value (see figure 28(b)). We begin by proving that $\psi_s(\eta, v, \lambda, \sigma)$ must be independent of η —that is, $\mathcal{I}$ attains its extremal value on a flux surface, as shown in figure 28(c)—for all passing particles. By definition of an extremum, every point along the curve satisfies

$$\partial_\eta \mathcal{I} = \sigma \partial_\eta I = 0 \quad (\text{D } 1a)$$

and

$$\partial_\psi \mathcal{I} = \sigma \partial_\psi I - \frac{2\pi MZe}{mc} \left(\iota - \frac{N}{M} \right) = 0. \quad (\text{D } 1b)$$

It will be useful to work with these conditions in Boozer coordinates because the definition of cyclometry on a general flux surface, given in §5.1.3, was formulated in Boozer coordinates. Conditions (D 1a) and (D 1b) are equivalent to

$$\partial_\eta \left(G \oint \frac{\sqrt{1 - \lambda B}}{B} d\zeta_B \right) = 0 \quad (\text{D } 2a)$$

and

$$\sigma v \partial_\psi \left(G \oint \frac{\sqrt{1 - \lambda B}}{B} d\zeta_B \right) - \frac{2\pi MZe}{mc} \left(\iota - \frac{N}{M} \right) = 0, \quad (\text{D } 2b)$$

respectively. We should think of (D 2b) as defining the curve $\psi = \psi_s(\eta, v, \lambda, \sigma)$, and (D 2a) as an additional condition that must be enforced at each point along this curve. The reason (D 2a) should not be used to define the curve is that this equation cannot always be inverted for ψ ; it may, for all passing λ , be satisfied throughout a finite area in (η, ψ) space. For example, (D 2a) holds everywhere in a stellarator that is cyclometric on every flux surface. Equation (D 2b), on the other hand, does not have this problem. While it is—in principle—possible that (D 2b) might hold throughout a finite area for a particular value of v , this cannot be the case for all v . Thus, for generic v , we can invert (D 2b) to determine $\psi = \psi_s(\eta, v, \lambda, \sigma)$.

We now use the fact that (D 2a) and (D 2b) hold for particles with different speeds v to prove that $\psi_s(\eta, v, \lambda, \sigma)$ must be independent of η —that is, $\mathcal{I}$ attains its extremal value on a flux surface, as shown in figure 28(c)—for all passing particles. Differentiating (D 2a) with respect to v and (D 2b) with respect to η , we find that

$$(\partial_v \psi_s) \partial_\psi \partial_\eta \left(G \oint \frac{\sqrt{1 - \lambda B}}{B} d\zeta_B \right) = 0 \quad (\text{D } 3a)$$

and

$$\sigma v \partial_\eta \partial_\psi \left(G \oint \frac{\sqrt{1-\lambda B}}{B} d\zeta_B \right) + \sigma v (\partial_\eta \psi_s) \partial_\psi^2 \left(G \oint \frac{\sqrt{1-\lambda B}}{B} d\zeta_B \right) = 0, \quad (\text{D } 3b)$$

are satisfied at each point along the curve $\psi = \psi_s(\eta, v, \lambda, \sigma)$. Since $\partial_v \psi_s$ is generically nonzero, (D 3a) implies

$$\partial_\psi \partial_\eta \left(G \oint \frac{\sqrt{1-\lambda B}}{B} d\zeta_B \right) = 0. \quad (\text{D } 4)$$

Then, using the commutativity of partial derivatives, (D 3b) becomes

$$(\partial_\eta \psi_s) \partial_\psi^2 \left(G \oint \frac{\sqrt{1-\lambda B}}{B} d\zeta_B \right) = 0. \quad (\text{D } 5)$$

Generically, the integral in (D 5) has a nonzero second derivative with respect to ψ , so $\partial_\eta \psi_s = 0$. Hence, $\psi_s(\eta, v, \lambda, \sigma)$ must be independent of η .

From here on, we assume that, whenever $\mathcal{I}$ has an extremum in the plasma for some v , λ , and σ , the extremal value is attained everywhere on a flux surface $\psi_s(v, \lambda, \sigma)$, as shown in figure 28(c), on which (D 2a) and (D 2b) are satisfied. Then, we can fix our attention on a single flux surface, ψ , and, as we vary λ , we can imagine simultaneously adjusting v in such a way that $\psi_s(v, \lambda, \sigma) = \psi$ remains constant.¹⁹ In other words, we are considering the passing particles whose orbits are resonant on flux surface ψ ; their speed $v_s(\psi, \lambda, \sigma)$ is given by

$$v_s = \frac{2\pi M Z e}{mc} \left(\iota - \frac{N}{M} \right) / \partial_\psi \left(\sigma G \oint \frac{\sqrt{1-\lambda B}}{B} d\zeta_B \right). \quad (\text{D } 6)$$

Only one sign of σ gives a positive v_s , which means resonant orbits only occur at ψ for particles circulating in one direction.

For all possible passing λ , we need to impose condition (D 2a) on flux surface ψ . In fact, as in §5.1, we will only need (D 2a) to hold for all λ in some open neighbourhood of zero, provided the denominator in (D 6) does not vanish at $\lambda = 0$; this assumption is satisfied on generic flux surfaces. From this point onwards, we can follow essentially the same steps that we took in §5.1 to obtain the cyclometry constraint for stellarators with higher shear. Wherever we Taylor expanded in $\mu/\mathcal{E}$ in §5.1, we now Taylor expand in λ . This leads to the result that $Z_B(\psi, \eta, B)$, defined in (5.13), must be independent of η , which means flux surface ψ must be cyclometric. Identical reasoning could be used for any flux surface ψ on which v_s , defined in (D 6), at $\lambda = 0$ is below the maximum alpha-particle speed. Physically, this means that cyclometry is necessary on all flux surfaces on which resonant passing orbits are possible for alpha particles with energies below 3.5 MeV.

Appendix E. Why μ -variation can be neglected when calculating $\mathcal{I}^{(1/2)}$

In our calculation of $\mathcal{I}^{(1/2)}$ in §6, we neglect changes in μ . This is inconsistent because $d\mu/dt$ (gyroaveraged) is non-zero at order $\rho_* v^3/B L$, so a term of the form $\partial_\mu \mathcal{I} (d\mu/d\zeta)$ should appear in (6.6). In this appendix, we explain why neglecting changes in μ in this way still leads to the right formula for $\mathcal{I}^{(1/2)}$ at order $\sqrt{\rho_*} v L$.

¹⁹This reasoning is valid unless the stellarator has $|\iota - N/M| \ll \rho_*$ everywhere. In such a device, the second term in (D 2b) can be neglected and the resulting equation cannot be solved for v_s . We do not consider this possibility further because, for $|\iota - N/M| \ll \rho_*$ to hold everywhere, the magnetic shear must satisfy $s \ll \rho_*$, which is much smaller than the already low $s \sim \rho_*$ ordering adopted here.

The main idea is that μ is the lowest-order term in a series, analogous to (6.1), for an improved adiabatic invariant $\bar{\mu}$ that is conserved to all orders in $\rho_\star$:

$$\bar{\mu}(\mathbf{X}, \mathcal{E}, \mu, \varphi, \sigma) = \mu + \mu^{(1)}(\mathbf{X}, \mathcal{E}, \mu, \varphi, \sigma) + \mu^{(2)}(\mathbf{X}, \mathcal{E}, \mu, \varphi, \sigma) + \dots, \quad (\text{E } 1)$$

where $\mu^{(n)} \sim \rho_\star^n v^2 / B$ and φ is the gyrophase. For example, the first correction is

$$\mu^{(1)} = \frac{v_\parallel^2}{B} (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} \cdot \boldsymbol{\rho}) + \frac{Ze}{2mc} v_\parallel (\boldsymbol{\rho} \cdot \nabla \hat{\mathbf{b}} \cdot (\hat{\mathbf{b}} \times \boldsymbol{\rho})) - \frac{mc}{2Ze} \frac{v_\parallel \mu}{B} \hat{\mathbf{b}} \cdot (\nabla \times \hat{\mathbf{b}}) + O\left(\frac{\rho_\star^2 v^2}{B}\right), \quad (\text{E } 2)$$

where $\boldsymbol{\rho} = (mc/ZeB)(\hat{\mathbf{b}} \times \mathbf{v})$ is the gyroradius vector and all functions of position are evaluated at the guiding-centre position $\mathbf{X}$ (e.g. Parra & Calvo, 2011). Imagine using $\bar{\mu}$ as a coordinate instead of μ in §6. Then, it would be completely rigorous to neglect $d\bar{\mu}/dt$ to the relevant order. Thus, the calculation would proceed exactly as in §6, but with $\bar{\mu}$ replacing μ everywhere. We would obtain the same formula (6.13) for the higher-order transit invariant, but with $\bar{\mu}$ replacing μ . This replacement only modifies (6.13) at order $\rho_\star vL$, which is higher-order than $\mathcal{I}^{(1/2)}$. Therefore, our formula for $\mathcal{I}^{(1/2)}$ is correct to the order of interest.

There is a subtlety in this argument: the higher-order magnetic moment $\bar{\mu}$ in (E 1) contains a term proportional to $\hat{\mathbf{b}} \cdot (\nabla \times \hat{\mathbf{b}}) \propto j_\parallel / B$, where $j_\parallel$ is the parallel current density. The equations of ideal MHD with nested flux surfaces can make this current density singular around rational surfaces (Helander, 2014), which would cause $\mu^{(1)}$ to diverge. This divergence indicates that the guiding-centre ordering is broken, which is expected because a sufficiently narrow current sheet creates a magnetic field with sharp gradients. In a real stellarator, singular currents will be regulated by the formation of magnetic islands (e.g. Reiman & Boozer, 1984; Cary & Kotschenreuther, 1985; Hegna & Bhattacharjee, 1989) or by transport processes (e.g. Boozer, 1981; Hazeltine & Catto, 2015), which add non-ideal terms to the MHD equations. Nevertheless, layers of finite but enhanced current density may still arise around rational surfaces. It is therefore an assumption in this work that any such layers are insignificant enough that the guiding-centre expansion remains valid as particles move across rational surfaces.

Appendix F. Manipulations required for the higher-order correction to the transit invariant

In this appendix, we show how to obtain (6.10). We will repeatedly use the fact that radial derivatives can make certain functions an order larger in $\sqrt{\rho_\star}$; in particular, $\partial_\psi \mathcal{I} \simeq -(2\pi MZe/mc) \iota'_r(\psi - \psi_r) \sim vL/\sqrt{\rho_\star} \Psi_t$ and $\partial_\psi^2 \mathcal{I} \simeq -(2\pi MZe/mc) \iota''_r \sim vL/\rho_\star \Psi_t^2$. Note, however, that $\partial_\eta \mathcal{I}$ does not have a large radial gradient: $\partial_\psi \partial_\eta \mathcal{I} \sim vL/\Psi_t \ll \partial_\psi \mathcal{I}$.

First, substituting (6.5) into (6.9), we find

$$\begin{aligned} 0 = \frac{mc}{Ze} & \left[\partial_\psi \mathcal{I}^{(1/2)}(\psi, \eta, 0) \partial_\eta \mathcal{I} - \partial_\eta \mathcal{I}^{(1/2)}(\psi, \eta, 0) \partial_\psi \mathcal{I} \right] \\ & - \oint \Gamma_\psi \partial_\psi \left(\partial_\psi \mathcal{I} \int_0^\zeta \Gamma'_\psi d\zeta' + \partial_\eta \mathcal{I} \int_0^\zeta \Gamma'_\eta d\zeta' \right) d\zeta \\ & - \oint \Gamma_\eta \partial_\eta \left(\partial_\psi \mathcal{I} \int_0^\zeta \Gamma'_\psi d\zeta' + \partial_\eta \mathcal{I} \int_0^\zeta \Gamma'_\eta d\zeta' \right) d\zeta. \quad (\text{F } 1) \end{aligned}$$

Here, we write Γ'_ψ and Γ'_η for the values of Γ_ψ and Γ_η at toroidal angle ζ' (this notation is only necessary to avoid ambiguity in certain double integrals in this appendix, so we

do not use this notation in the main text). The last two lines of (F 1) contain four double integrals, which we simplify as follows. For the first integral, we have

$$\oint \Gamma_\psi \partial_\psi \left(\partial_\psi \mathcal{I} \int_0^\zeta \Gamma'_\psi d\zeta' \right) d\zeta \simeq \partial_\psi^2 \mathcal{I} \oint \Gamma_\psi \left(\int_0^\zeta \Gamma'_\psi d\zeta' \right) d\zeta, \quad (\text{F } 2)$$

due to the large radial gradient of $\mathcal{I}$. Realizing that

$$2\Gamma_\psi \int_0^\zeta \Gamma'_\psi d\zeta' = \partial_\zeta \left[\left(\int_0^\zeta \Gamma'_\psi d\zeta' \right)^2 \right], \quad (\text{F } 3)$$

we find

$$\oint \Gamma_\psi \left(\int_0^\zeta \Gamma'_\psi d\zeta' \right) d\zeta = \frac{1}{2} \frac{mc}{Ze} \partial_\eta \mathcal{I} \oint \Gamma_\psi d\zeta, \quad (\text{F } 4)$$

where we have employed (6.8) to write this expression in a form that will be convenient later. Then, using $\partial_\psi^2 \mathcal{I} \simeq -(2\pi M Ze/mc) \iota'_r$, the first integral becomes

$$\oint \Gamma_\psi \partial_\psi \left(\partial_\psi \mathcal{I} \int_0^\zeta \Gamma'_\psi d\zeta' \right) d\zeta \simeq -(\partial_\eta \mathcal{I}) \pi M \iota'_r \oint \Gamma_\psi d\zeta. \quad (\text{F } 5)$$

The second integral on the middle line of (F 1) is

$$\oint \Gamma_\psi \partial_\psi \left(\partial_\eta \mathcal{I} \int_0^\zeta \Gamma'_\eta d\zeta' \right) d\zeta \simeq (\partial_\eta \mathcal{I}) \iota'_r \oint \zeta \Gamma_\psi d\zeta, \quad (\text{F } 6)$$

where we have used $\Gamma_\eta \simeq \iota'_r(\psi - \psi_r)$ (by (6.4)) and the fact that $\partial_\eta \mathcal{I}$ does not have a large radial gradient.

For the first integral on the last line of (F 1), we have

$$\oint \Gamma_\eta \partial_\eta \left(\partial_\psi \mathcal{I} \int_0^\zeta \Gamma'_\psi d\zeta' \right) d\zeta \simeq (\partial_\psi \mathcal{I}) \iota'_r(\psi - \psi_r) \partial_\eta \left[\oint \left(\int_0^\zeta \Gamma'_\psi d\zeta' \right) d\zeta \right], \quad (\text{F } 7)$$

because $\partial_\psi \partial_\eta \mathcal{I}$ is small. Integrating by parts gives

$$\oint \left(\int_0^\zeta \Gamma'_\psi d\zeta' \right) d\zeta = 2\pi M \oint \Gamma_\psi d\zeta - \oint \zeta \Gamma_\psi d\zeta, \quad (\text{F } 8)$$

from which we obtain

$$\oint \Gamma_\eta \partial_\eta \left(\partial_\psi \mathcal{I} \int_0^\zeta \Gamma'_\psi d\zeta' \right) d\zeta \simeq -(\partial_\psi \mathcal{I}) \iota'_r(\psi - \psi_r) \partial_\eta \left(\oint (\zeta - 2\pi M) \Gamma_\psi d\zeta \right). \quad (\text{F } 9)$$

Finally, the last integral in (F 1) can be simplified using

$$\begin{aligned} \oint \Gamma_\eta \partial_\eta \left(\partial_\eta \mathcal{I} \int_0^\zeta \Gamma'_\eta d\zeta' \right) d\zeta &\simeq \iota_r'^2 (\psi - \psi_r)^2 (\partial_\eta^2 \mathcal{I}) \oint \zeta d\zeta \\ &\simeq -(\partial_\psi \mathcal{I}) \pi M \iota'_r (\psi - \psi_r) \oint \Gamma_\psi d\zeta, \end{aligned} \quad (\text{F } 10)$$

where the last line follows from (6.8).

Replacing the integrals in (F 1) with the four expressions (F 5), (F 6), (F 9) and (F 10) gives (6.10).

Appendix G. Origin of the arbitrary constant in $\mathcal{I}^{(1/2)}$

In this appendix, we explain why an undetermined function $f(\mathcal{I})$ arises, in (6.12), in our calculation of the higher-order transit adiabatic invariant, and why we are free to set this function to zero. The fact that arbitrary choices like this one need to be made in the construction of higher-order adiabatic invariants is well known (e.g. [Lenard, 1959](#)).

In §6, we find $\mathcal{I}^{(1/2)}$ by enforcing the condition

$$\left\langle \frac{d(\mathcal{I} + \mathcal{I}^{(1/2)})}{dt} \right\rangle_t \sim \rho_*^{3/2} v^2, \quad (\text{G } 1)$$

which makes $\mathcal{I} + \mathcal{I}^{(1/2)}$ an approximately conserved quantity. For any small, but otherwise arbitrary, function $f(\mathcal{I}) \sim \rho_*^{1/2} \mathcal{I}$, we also have

$$\left\langle \frac{d(\mathcal{I} + \mathcal{I}^{(1/2)} + f(\mathcal{I}))}{dt} \right\rangle_t \sim \rho_*^{3/2} v^2, \quad (\text{G } 2)$$

which means $\mathcal{I}^{(1/2)} + f(\mathcal{I})$ is an equally valid choice for the higher-order correction to the transit adiabatic invariant. Since f is arbitrary, we can choose $f = 0$.

This freedom in the definition of $\mathcal{I}^{(1/2)}$ is related to the fact that a function of an adiabatic invariant is also an adiabatic invariant. Equation (G 1) states that $\mathcal{I} + \mathcal{I}^{(1/2)}$ is a approximately conserved quantity. For any function g , $g(\mathcal{I} + \mathcal{I}^{(1/2)})$ must also be approximately conserved. Suppose $g(\mathcal{I}) = \mathcal{I} + f(\mathcal{I})$, for some small $f(\mathcal{I}) \sim \rho_*^{1/2} \mathcal{I}$. Expanding $g(\mathcal{I} + \mathcal{I}^{(1/2)}) \simeq \mathcal{I} + \mathcal{I}^{(1/2)} + f(\mathcal{I})$, we learn that $\mathcal{I} + \mathcal{I}^{(1/2)} + f(\mathcal{I})$ is approximately conserved, so $\mathcal{I}^{(1/2)} + f(\mathcal{I})$ is an equally valid choice for the higher-order correction to the transit adiabatic invariant.

Appendix H. Gauge invariance of the transit adiabatic invariant

In this appendix, we verify explicitly that formula (6.13) gives the same value for $\bar{\mathcal{I}}$ in any straight-field-line coordinate system (up to terms of order $\rho_* v L$). We will use a new set of straight-field-line coordinates $(\psi, \tilde{\theta}, \tilde{\zeta})$, defined by

$$\tilde{\zeta} = \zeta + \nu(\psi, \theta, \zeta), \quad (\text{H } 1a)$$

$$\tilde{\theta} = \theta + \iota(\psi) \nu(\psi, \theta, \zeta), \quad (\text{H } 1b)$$

where $\nu(\psi, \theta, \zeta)$ is 2π -periodic in both its θ and ζ arguments. This is not the most general transformation between two sets of straight-field-line coordinates: we could also add two flux functions whose ratio is not the rotational transform $\iota(\psi)$ to the right sides of (H 1a) and (H 1b). We ignore this possibility because adding a flux function to either θ or ζ is equivalent to changing the location where that angle is zero, and (6.16) shows that this would not change the value of $\bar{\mathcal{I}}$ (as discussed in §6.3).

To begin, we introduce some helpful notation. We write $(\psi, \tilde{\eta}(\psi, \eta, \zeta), \tilde{\zeta}(\psi, \eta, \zeta))$ for the new coordinates describing the point (ψ, η, ζ) . Similarly, we write $(\psi, \eta(\psi, \tilde{\eta}, \tilde{\zeta}), \zeta(\psi, \tilde{\eta}, \tilde{\zeta}))$ for the old coordinates corresponding to $(\psi, \tilde{\eta}, \tilde{\zeta})$. For any other function $f(\psi, \eta, \zeta)$ of the old coordinates, we define

$$\tilde{f}(\psi, \tilde{\eta}, \tilde{\zeta}) := f(\psi, \eta(\psi, \tilde{\eta}, \tilde{\zeta}), \zeta(\psi, \tilde{\eta}, \tilde{\zeta})). \quad (\text{H } 2)$$

Then, the definition of the new coordinates can be expressed in two equivalent ways:

$$\tilde{\zeta} = \zeta + \nu(\psi, \eta, \zeta) \quad \text{or} \quad \zeta = \tilde{\zeta} - \tilde{\nu}(\psi, \tilde{\eta}, \tilde{\zeta}), \quad (\text{H } 3a)$$

$$\tilde{\eta} = \eta + \left(\iota(\psi) - \frac{N}{M} \right) \nu(\psi, \eta, \zeta) \quad \text{or} \quad \eta = \tilde{\eta} - \left(\iota(\psi) - \frac{N}{M} \right) \tilde{\nu}(\psi, \tilde{\eta}, \tilde{\zeta}). \quad (\text{H } 3b)$$

For the rest of this appendix, we suppress the $(\mathcal{E}, \mu, \sigma)$ arguments of any functions that depend on the particle velocity.

We write $\tilde{\tilde{\mathcal{K}}}(\psi, \tilde{\eta}, \tilde{\zeta})$ for the value of the higher-order transit invariant that someone working exclusively in the new coordinates would obtain using (6.13). To be explicit, this means $\tilde{\tilde{\mathcal{K}}}(\psi, \tilde{\eta}, \tilde{\zeta})$ is defined by

$$\begin{aligned} \tilde{\tilde{\mathcal{K}}}(\psi, \tilde{\eta}, \tilde{\zeta}) := & \oint (v_{\parallel} \widehat{\mathbf{b}} \cdot \partial_{\tilde{\zeta}} \mathbf{x})(\psi, \tilde{\eta}, \tilde{\zeta}') d\tilde{\zeta}' - \frac{2\pi MZe}{mc} \int_{\psi_r}^{\psi} \left(\iota(\psi') - \frac{N}{M} \right) d\psi' \\ & + \frac{Ze}{mc} \iota'_r(\psi - \psi_r) \left[\int_0^{\tilde{\zeta}} 2\pi M \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \widehat{\mathbf{b}} \cdot \nabla \tilde{\zeta}} \right) (\psi_r, \tilde{\eta}, \tilde{\zeta}') d\tilde{\zeta}' \right. \\ & \left. + \oint (\tilde{\zeta}' - \tilde{\zeta} - \pi M) \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \widehat{\mathbf{b}} \cdot \nabla \tilde{\zeta}} \right) (\psi_r, \tilde{\eta}, \tilde{\zeta}') d\tilde{\zeta}' \right]. \end{aligned} \quad (\text{H } 4)$$

To prove the gauge invariance of (6.13), our aim will be to show that, in fact, $\tilde{\tilde{\mathcal{K}}}(\psi, \tilde{\eta}, \tilde{\zeta}) = \tilde{\mathcal{I}}(\psi, \eta, \zeta)$ (up to terms of order $\rho_{\star} vL$). Using (H 2), we find

$$\begin{aligned} \tilde{\tilde{\mathcal{K}}}(\psi, \tilde{\eta}, \tilde{\zeta}) = & \underbrace{\oint (v_{\parallel} \widehat{\mathbf{b}} \cdot \partial_{\tilde{\zeta}} \mathbf{x})(\psi, \eta(\psi, \tilde{\eta}, \tilde{\zeta}'), \zeta(\psi, \tilde{\eta}, \tilde{\zeta}')) d\tilde{\zeta}'}_{\textcircled{1}} - \underbrace{\frac{2\pi MZe}{mc} \int_{\psi_r}^{\psi} \left(\iota(\psi') - \frac{N}{M} \right) d\psi'}_{\textcircled{2}} \\ & + \frac{Ze}{mc} \iota'_r(\psi - \psi_r) \left[\int_0^{\tilde{\zeta}} 2\pi M \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \widehat{\mathbf{b}} \cdot \nabla \tilde{\zeta}} \right) (\psi_r, \eta(\psi_r, \tilde{\eta}, \tilde{\zeta}'), \zeta(\psi_r, \tilde{\eta}, \tilde{\zeta}')) d\tilde{\zeta}' \right. \\ & \left. + \oint (\tilde{\zeta}' - \tilde{\zeta} - \pi M) \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \widehat{\mathbf{b}} \cdot \nabla \tilde{\zeta}} \right) (\psi_r, \eta(\psi_r, \tilde{\eta}, \tilde{\zeta}'), \zeta(\psi_r, \tilde{\eta}, \tilde{\zeta}')) d\tilde{\zeta}' \right]_{\textcircled{3}}. \end{aligned} \quad (\text{H } 5)$$

In the arguments of the integrands here, $\eta(\psi, \tilde{\eta}, \tilde{\zeta}')$ is *not* the value of η corresponding to the guiding-centre position: for that we would write $\eta(\psi, \tilde{\eta}, \tilde{\zeta})$, or simply η . The difference between $\eta(\psi, \tilde{\eta}, \tilde{\zeta}')$ and η is order $\sqrt{\rho_{\star}}$ and, therefore, cannot be neglected in $\textcircled{1}$. This difference is given by

$$\eta(\psi, \tilde{\eta}, \tilde{\zeta}') = \eta + \left(\iota(\psi) - \frac{N}{M} \right) [\tilde{\nu}(\psi, \tilde{\eta}, \tilde{\zeta}) - \tilde{\nu}(\psi, \tilde{\eta}, \tilde{\zeta}')]. \quad (\text{H } 6)$$

Thus, we should Taylor expand the integrand of $\textcircled{1}$ in its η argument using (H 6). The result is

$$\begin{aligned} \textcircled{1} \simeq & \oint (v_{\parallel} \widehat{\mathbf{b}} \cdot \partial_{\zeta} \mathbf{x})(\psi, \eta, \zeta') d\zeta' \\ & + \iota'_r(\psi - \psi_r) \oint [\nu(\psi_r, \eta, \zeta) - \nu(\psi_r, \eta, \zeta')] \partial_{\eta} (v_{\parallel} \widehat{\mathbf{b}} \cdot \partial_{\zeta} \mathbf{x})(\psi_r, \eta, \zeta') d\zeta' \\ & - \iota'_r(\psi - \psi_r) \oint (\partial_{\zeta} \nu v_{\parallel} \widehat{\mathbf{b}} \cdot \partial_{\eta} \mathbf{x})(\psi_r, \eta, \zeta') d\zeta' + O(\rho_{\star} vL), \end{aligned} \quad (\text{H } 7)$$

where we have also employed

$$\partial_{\tilde{\zeta}} \mathbf{x} = \partial_{\tilde{\zeta}} \zeta [\partial_{\zeta} \mathbf{x} - \iota'_r(\psi - \psi_r) \partial_{\zeta} \nu \partial_{\eta} \mathbf{x}] + O(\rho_{\star} L) \quad (\text{H } 8)$$

and changed the integration variable from $\tilde{\zeta}'$ to $\zeta' = \zeta(\psi, \tilde{\eta}, \tilde{\zeta}')$ (which means $d\zeta' = \partial_{\tilde{\zeta}} \zeta d\tilde{\zeta}'$).

We also need to manipulate the terms labelled ③ (the higher-order correction terms). In these terms, we can use $\eta(\psi, \tilde{\eta}, \tilde{\zeta}) \simeq \eta$ and $\hat{\mathbf{b}} \cdot \nabla \zeta \simeq \partial_{\tilde{\zeta}} \zeta \hat{\mathbf{b}} \cdot \nabla \tilde{\zeta}$, which are both correct up to relative errors of order $\sqrt{\rho_{\star}}$. Changing integration variables to $\zeta' = \zeta(\psi, \tilde{\eta}, \tilde{\zeta}')$, as above, we obtain

$$\begin{aligned} \textcircled{3} \simeq \frac{Ze}{mc} \iota'_r(\psi - \psi_r) & \left[\int_{\zeta_0}^{\zeta} 2\pi M \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta} \right) (\psi_r, \eta, \zeta') d\zeta' \right. \\ & + \int_{\zeta_0}^{\zeta_0 + 2\pi M} (\zeta' - \zeta + \nu(\psi_r, \eta, \zeta')) \\ & \quad \left. - \nu(\psi_r, \eta, \zeta) - \pi M \right) \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta} \right) (\psi_r, \eta, \zeta') d\zeta' \right] \quad (\text{H } 9) \end{aligned}$$

(up to an error of order $\rho_{\star} vL$), where $\zeta_0 = \zeta(\psi_r, \tilde{\eta}, \tilde{\zeta}=0)$. In fact, we can replace ζ_0 with 0 in (H 9). This can be seen by performing, in (H 9), the following manipulations:

$$\begin{aligned} & \int_{\zeta_0}^{\zeta_0 + 2\pi M} (\zeta' - \zeta + \nu(\psi_r, \eta, \zeta') - \nu(\psi_r, \eta, \zeta) - \pi M) \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta} \right) (\psi_r, \eta, \zeta') d\zeta' \\ & = \left(\int_0^{2\pi M} - \int_0^{\zeta_0} + \int_{2\pi M}^{2\pi M + \zeta_0} \right) \left[(\zeta' - \zeta + \nu(\psi_r, \eta, \zeta')) \right. \\ & \quad \left. - \nu(\psi_r, \eta, \zeta) - \pi M \right) \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta} \right) (\psi_r, \eta, \zeta') \right] d\zeta' \\ & = \int_0^{2\pi M} (\zeta' - \zeta + \nu(\psi_r, \eta, \zeta') - \nu(\psi_r, \eta, \zeta) - \pi M) \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta} \right) (\psi_r, \eta, \zeta') d\zeta' \\ & \quad + \int_0^{\zeta_0} 2\pi M \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta} \right) (\psi_r, \eta, \zeta') d\zeta'. \quad (\text{H } 10) \end{aligned}$$

Finally, using (6.14) and integrating by parts, we obtain

$$\begin{aligned} \textcircled{3} \simeq \frac{Ze}{mc} \iota'_r(\psi - \psi_r) & \left[\int_0^{\zeta} 2\pi M \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta} \right) (\psi_r, \eta, \zeta') d\zeta' \right. \\ & \quad \left. + \int_0^{2\pi M} (\zeta' - \zeta - \pi M) \left(\frac{\mathbf{v}_d \cdot \nabla \psi}{v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \zeta} \right) (\psi_r, \eta, \zeta') d\zeta' \right] \\ & + \iota'_r(\psi - \psi_r) \int_0^{2\pi M} [\nu(\psi_r, \eta, \zeta') - \nu(\psi_r, \eta, \zeta)] \partial_{\eta} (v_{\parallel} \hat{\mathbf{b}} \cdot \partial_{\zeta} \mathbf{x}) (\psi_r, \eta, \zeta') d\zeta' \\ & + \iota'_r(\psi - \psi_r) \int_0^{2\pi M} (\partial_{\zeta} \nu v_{\parallel} \hat{\mathbf{b}} \cdot \partial_{\eta} \mathbf{x}) (\psi_r, \eta, \zeta') d\zeta'. \quad (\text{H } 11) \end{aligned}$$

Summing ①, ②, and ③ gives $\tilde{\mathcal{K}}(\psi, \tilde{\eta}, \tilde{\zeta}) = \bar{\mathcal{I}}(\psi, \eta, \zeta)$, up to terms of order $\rho_{\star} vL$. Therefore, formula (6.13) for the higher-order transit invariant is gauge invariant, as claimed.

Appendix I. Limiting forms of $\bar{\chi}_r$

In §7.5, we show that passing orbits in the region $\mathfrak{R}_1$ (which contains a single low-order rational surface) conserve $\bar{\chi}_r$, defined in (7.42). In this appendix, we show that this conservation law reproduces the correct behaviour for orbits close to (§I.1) or far from (§I.2) the rational surface.

I.1. Close to the rational surface

First, we discuss particles whose distance to the rational surface is on the order of a drift-island width. In §6, we described such particles by ordering $s \sim 1$ and $\psi - \psi_r \sim \sqrt{\rho_\star} \Psi_t$. Here, we will show that, when this ordering is satisfied, $\bar{\chi}_r$ becomes (up to a constant factor) the transit adiabatic invariant, including the higher-order corrections derived in §6.

With $s \sim 1$ and $\psi - \psi_r \sim \sqrt{\rho_\star} \Psi_t$, formula (7.42) for $\bar{\chi}_r$ becomes

$$\begin{aligned} \bar{\chi}_r(\psi, \eta, \zeta) = & \int_{\psi_r}^{\psi} \left(\iota(\psi') - \frac{N}{M} \right) d\psi' - \frac{1}{2\pi M} \frac{mc}{Ze} \oint v_{\parallel} \hat{\mathbf{b}} \cdot \left(\partial_{\zeta} \mathbf{x} + \frac{N}{M} \partial_{\theta} \mathbf{x} \right) d\zeta' \\ & + \frac{mc}{Ze} \iota'_r(\psi - \psi_r) \left(v_{\parallel} \hat{\mathbf{b}} \cdot \partial_{\theta} \mathbf{x} - \frac{1}{2\pi M} \oint v_{\parallel} \hat{\mathbf{b}} \cdot \partial_{\theta} \mathbf{x} d\zeta' \right) \\ & - \iota'_r(\psi - \psi_r) \partial_{\theta} S_{\text{nr}}^{(1)} + O(\rho_\star^2 \Psi_t). \end{aligned} \quad (\text{I1})$$

In the second and third lines of (I1), all functions of position except $\psi - \psi_r$ are evaluated on the rational surface, at (ψ_r, θ, ζ) . To make further progress, we need to know $\partial_{\theta} S_{\text{nr}}^{(1)}(\psi_r, \theta, \zeta)$. We make a negligible error by restoring, in the definition (7.35) of $S_{\text{nr}}^{(1)}$, all non-resonant harmonics with $|p| + |q| \geq K$. This gives

$$S_{\text{nr}}^{(1)}(\psi_r, \theta, \zeta) \simeq \sum_{(p,q) \in \mathbb{I}_{\text{nr}}^{<K}} \frac{i(V_{\text{nr}}^{<K})_{pq}(\psi_r)}{q - Np/M} e^{i(p\theta - q\zeta)} + \sum_{(p,q) \in \mathbb{I}_{\text{nr}}^{\geq K}} \frac{i(V_{\text{nr}}^{\geq K})_{pq}(\psi_r)}{q - Np/M} e^{i(p\theta - q\zeta)}, \quad (\text{I2})$$

where $\mathbb{I}_{\text{nr}}^{\geq K} := \{(p, q) \in \mathbb{Z}^2 : |p| + |q| \geq K \text{ and } (p, q) \neq (kM, kN) \text{ for any } k \in \mathbb{Z}\}$. Restoring these harmonics is allowed because the fact that we are working precisely on the rational surface means their size can be bounded as follows. Since $|Mq - Np| \geq 1$ for non-resonant harmonics, we have

$$\left| \sum_{(p,q) \in \mathbb{I}_{\text{nr}}^{\geq K}} \frac{i(V_{\text{nr}}^{\geq K})_{pq}(\psi_r)}{q - Np/M} e^{i(p\theta - q\zeta)} \right| \leq M \sum_{(p,q) \in \mathbb{I}_{\text{nr}}^{\geq K}} |(V_{\text{nr}}^{\geq K})_{pq}(\psi_r)|. \quad (\text{I3})$$

If the Fourier series of $V(\psi_r, \theta, \zeta)$ is absolutely convergent, then the sum on the right side of (I3) converges and will be small when K is large. Hence, the approximate form of $S_{\text{nr}}^{(1)}(\psi_r, \theta, \zeta)$ in (I2) satisfies

$$\left(\partial_{\zeta} + \frac{N}{M} \partial_{\theta} \right) S_{\text{nr}}^{(1)}(\psi_r, \theta, \zeta) \simeq V(\psi_r, \theta, \zeta) - \langle V \rangle_{\zeta}(\psi_r, \theta, \zeta). \quad (\text{I4})$$

This magnetic differential equation is useful because we can solve it by integrating along rational-surface field lines to derive a simple integral formula for $S_{\text{nr}}^{(1)}$. The result is

$$S_{\text{nr}}^{(1)}(\psi_r, \theta, \zeta) \simeq \int_{0,r}^{\zeta} (V - \langle V \rangle_{\zeta}) d\zeta' + h(\eta), \quad (\text{I5})$$

for some function $h(\eta)$. This function can be determined using the fact that $S_{\text{nr}}^{(1)}$ contains

no $e^{ik(M\theta - N\zeta)}$ harmonics, for any $k \in \mathbb{Z}$, according to definition (7.35). Therefore, we may impose $\langle S_{\text{nr}}^{(1)} \rangle_\zeta = 0$. This gives

$$\begin{aligned} \langle S_{\text{nr}}^{(1)} \rangle_\zeta &\simeq \frac{1}{2\pi M} \oint_{\text{r}} \left(\int_{0,\text{r}}^\zeta V \, d\zeta' \right) d\zeta - \frac{1}{2\pi M} \oint_{\text{r}} \zeta \langle V \rangle_\zeta d\zeta + h(\eta) \\ &= \pi M \langle V \rangle_\zeta - \frac{1}{2\pi M} \oint_{\text{r}} \zeta V \, d\zeta + h(\eta), \end{aligned} \quad (\text{I6})$$

where we have integrated by parts. Solving for h , we find

$$S_{\text{nr}}^{(1)}(\psi_{\text{r}}, \theta, \zeta) \simeq \int_{0,\text{r}}^\zeta V \, d\zeta' + \frac{1}{2\pi M} \oint_{\text{r}} (\zeta' - \zeta - \pi M) V \, d\zeta'. \quad (\text{I7})$$

Now we substitute this result into (I1). Using $\partial_\theta = \partial_\eta$ and $\partial_\zeta \mathbf{x} + (N/M) \partial_\theta \mathbf{x} = \partial_\zeta \mathbf{x}|_\eta$, where we have temporarily introduced the notation $\partial_\zeta \mathbf{x}|_\eta$ to indicate a ζ derivative holding η fixed, as well as $V = (mc/Ze) v_\parallel \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x}|_\eta$ on the rational surface, we find

$$\begin{aligned} \bar{\chi}_{\text{r}}(\psi, \eta, \zeta) &\simeq \int_{\psi_{\text{r}}}^\psi \left(\iota(\psi') - \frac{N}{M} \right) d\psi' - \frac{1}{2\pi M} \frac{mc}{Ze} \oint v_\parallel \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x}|_\eta d\zeta' \\ &\quad + \frac{mc}{Ze} \iota'_{\text{r}}(\psi - \psi_{\text{r}}) \left(v_\parallel \hat{\mathbf{b}} \cdot \partial_\eta \mathbf{x} - \int_{0,\text{r}}^\zeta \partial_\eta (v_\parallel \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x}|_\eta) d\zeta' \right. \\ &\quad \left. + \frac{1}{2\pi M} \oint_{\text{r}} [(\zeta' - \zeta - \pi M) \partial_\eta (v_\parallel \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x}|_\eta) - v_\parallel \hat{\mathbf{b}} \cdot \partial_\eta \mathbf{x}] d\zeta' \right). \end{aligned} \quad (\text{I8})$$

From here on, we use ∂_ζ to denote ζ derivatives at fixed η . Integrating by parts allows us to express $\bar{\chi}_{\text{r}}$ in the form

$$\begin{aligned} \bar{\chi}_{\text{r}}(\psi, \eta, \zeta) &\simeq \int_{\psi_{\text{r}}}^\psi \left(\iota(\psi') - \frac{N}{M} \right) d\psi' - \frac{1}{2\pi M} \frac{mc}{Ze} \oint v_\parallel \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x} d\zeta' \\ &\quad - \frac{mc}{Ze} \iota'_{\text{r}}(\psi - \psi_{\text{r}}) \left(\int_{0,\text{r}}^\zeta [\partial_\eta (v_\parallel \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x}) - \partial_\zeta (v_\parallel \hat{\mathbf{b}} \cdot \partial_\eta \mathbf{x})] d\zeta' \right. \\ &\quad \left. + \frac{1}{2\pi M} \oint_{\text{r}} (\zeta' - \zeta - \pi M) [\partial_\eta (v_\parallel \hat{\mathbf{b}} \cdot \partial_\zeta \mathbf{x}) - \partial_\zeta (v_\parallel \hat{\mathbf{b}} \cdot \partial_\eta \mathbf{x})] d\zeta' \right), \end{aligned} \quad (\text{I9})$$

which is proportional to the higher-order transit adiabatic invariant (6.13), as claimed.

I.2. Far from the rational surface

Next, we discuss particles that are far from the rational surface, by which we mean $\iota - N/M \sim 1$. The particle orbits that satisfy this ordering are contained in a subregion of $\mathfrak{R}_1$ that lies much further than a drift-island width from the rational surface ψ_{r} . This subregion is a permissible choice for the region $\mathfrak{R}_0$ introduced in §7.4, according to definition 7.1. Therefore, for such particles, we expect conservation of $\bar{\chi}_{\text{r}}$ to be equivalent to conservation of $\bar{\psi}$, defined by (7.30). We proceed to show that this is the case.

First, since the formula for $\bar{\psi}$ involves $S^{(1)}$ while the formula for $\bar{\chi}_{\text{r}}$ involves $S_{\text{nr}}^{(1)}$, it will be useful to relate these two functions. The difference between $S^{(1)}$ and $S_{\text{nr}}^{(1)}$ consists entirely of resonant harmonics:

$$S^{(1)} = S_{\text{nr}}^{(1)} + \sum_{\substack{k \in \mathbb{Z} \setminus \{0\} \\ |kM| + |kN| < K}} \frac{iV_{kM, kN}}{k(N - \iota M)} e^{ik(M\theta - N\zeta)}. \quad (\text{I10})$$

This difference clearly blows up as $\psi \rightarrow \psi_r$; however, we are now working in a subregion of $\mathfrak{R}_1$ that excludes this divergence. Thus, since $\iota - N/M \sim 1$, we make a small error by restoring, in the sum, the harmonics with $|p| + |q| \geq K$:

$$S^{(1)} \simeq S_{\text{nr}}^{(1)} + \sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{iV_{kM, kN}}{k(N - \iota M)} e^{ik(M\theta - N\zeta)}. \quad (\text{I11})$$

Only the θ derivatives of $S^{(1)}$ and $S_{\text{nr}}^{(1)}$ appear in $\bar{\psi}$ and $\bar{\chi}_r$, so we differentiate (I11) with respect to θ and find

$$\partial_\theta S^{(1)} \simeq \partial_\theta S_{\text{nr}}^{(1)} + \frac{\langle V \rangle_\zeta - \langle V \rangle}{\iota - N/M}. \quad (\text{I12})$$

With this approximation, equation (7.42) for $\bar{\chi}_r$ becomes

$$\bar{\chi}_r(\psi, \eta, \zeta) \simeq \int_{\psi_r}^\psi \left(\iota(\psi') - \frac{N}{M} \right) d\psi' + \left(\iota(\psi) - \frac{N}{M} \right) \left(\frac{mc}{Ze} v_\parallel \hat{\mathbf{b}} \cdot \partial_\theta \mathbf{x} - \partial_\theta S^{(1)} \right) - \langle V \rangle. \quad (\text{I13})$$

To lowest order, our conserved quantity is $\bar{\chi}_r(\psi, \eta, \zeta) \simeq \chi_r(\psi)$, which means ψ is constant. Therefore, it must be possible to re-express conservation of $\bar{\chi}_r$ in the form

$$\psi + \psi^{[1]}(\psi, \theta, \zeta) + \psi^{[2]}(\psi, \theta, \zeta) + \dots = \text{const.}, \quad (\text{I14})$$

where $\psi_t \gg \psi^{[1]} \gg \psi^{[2]} \gg \dots$. Note that we are using square brackets in the superscripts to distinguish these corrections from $\psi^{(1)}$ and $\psi^{(2)}$ defined in (7.9) and (7.44), respectively. Equation (I14) is consistent with $\bar{\chi}_r(\psi, \theta, \zeta) = \text{const.}$ if we take

$$\psi^{[1]}(\psi, \theta, \zeta) := \frac{mc}{Ze} v_\parallel \hat{\mathbf{b}} \cdot \partial_\theta \mathbf{x} - \partial_\theta S^{(1)}. \quad (\text{I15})$$

Note that we are free to add any function of ψ to $\psi^{[1]}$, by reasoning similar to that of §G. With the choice (I15), we recover (7.30) as claimed.

Appendix J. Higher-order drift surfaces in the presence of a single resonance

In this appendix, we show how the higher-order conservation law (7.48) for particles in region $\mathfrak{R}_1$ is obtained. To begin, we must retrace the steps leading up to Lagrangian (7.37) and restore all terms of order $\rho_\star^2$ that we previously suppressed. We find

$$\begin{aligned} \mathcal{L} \equiv & \left(\bar{\psi}_{\text{nr}} + \psi_{\text{nr}}^{(1)} \left[\partial_\psi \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\theta \mathbf{x} \right) - \partial_\theta \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\psi \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_\theta \psi_{\text{nr}}^{(1)} \right) \frac{d\bar{\theta}_{\text{nr}}}{d\zeta} \\ & + \left(\theta_{\text{nr}}^{(1)} \left[\partial_\psi \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\psi \mathbf{x} \right) - \partial_\psi \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\theta \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_\psi \psi_{\text{nr}}^{(1)} \right) \frac{d\bar{\psi}_{\text{nr}}}{d\zeta} \\ & - \chi - \frac{1}{2} \iota' (\psi_{\text{nr}}^{(1)})^2 + \psi_{\text{nr}}^{(1)} \left[\partial_\psi \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\zeta \mathbf{x} \right) - \partial_\zeta \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\psi \mathbf{x} \right) \right] \\ & + \theta_{\text{nr}}^{(1)} \left[\partial_\theta \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\zeta \mathbf{x} \right) - \partial_\zeta \left(\frac{v_\parallel}{\Omega} \mathbf{B} \cdot \partial_\theta \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_\zeta \psi_{\text{nr}}^{(1)} + V_r + V_{\text{nr}}^{\geq K}, \quad (\text{J1}) \end{aligned}$$

where all functions are evaluated at $(\bar{\psi}_{\text{nr}}, \bar{\theta}_{\text{nr}}, \zeta)$. Here, and for the remainder of this appendix, we drop terms of order $\rho_\star^3$ or smaller. Next, we make the near-identity change

of coordinates defined by

$$\bar{\psi}_{\text{nr}} = \bar{\bar{\psi}}_{\text{nr}} + \psi_{\text{nr}}^{(2)}(\bar{\bar{\psi}}_{\text{nr}}, \bar{\bar{\theta}}_{\text{nr}}, \zeta), \quad (\text{J } 2a)$$

$$\bar{\theta}_{\text{nr}} = \bar{\bar{\theta}}_{\text{nr}} + \theta_{\text{nr}}^{(2)}(\bar{\bar{\psi}}_{\text{nr}}, \bar{\bar{\theta}}_{\text{nr}}, \zeta). \quad (\text{J } 2b)$$

Substituting (J 2) into (J 1) gives

$$\begin{aligned} \mathcal{L} \equiv & \left(\bar{\bar{\psi}}_{\text{nr}} + \psi_{\text{nr}}^{(2)} + \psi_{\text{nr}}^{(1)} \left[\partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right. \right. \\ & \left. \left. - \partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\theta} \psi_{\text{nr}}^{(1)} \right) \left(\frac{d\bar{\bar{\theta}}_{\text{nr}}}{d\zeta} + \frac{d\psi_{\text{nr}}^{(2)}}{d\zeta} \right) \\ & + \left(\theta_{\text{nr}}^{(1)} \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) - \partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\psi} \psi_{\text{nr}}^{(1)} \right) \left(\frac{d\bar{\bar{\psi}}_{\text{nr}}}{d\zeta} + \frac{d\psi_{\text{nr}}^{(2)}}{d\zeta} \right) \\ & - \chi - \iota \psi_{\text{nr}}^{(2)} - \frac{1}{2} \iota' (\psi_{\text{nr}}^{(1)})^2 + \psi_{\text{nr}}^{(1)} \left[\partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) \right] \\ & + \theta_{\text{nr}}^{(1)} \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\zeta} \psi_{\text{nr}}^{(1)} + V_{\text{r}} + V_{\text{nr}}^{\geq K}, \end{aligned} \quad (\text{J } 3)$$

where all functions are now evaluated at $(\bar{\bar{\psi}}_{\text{nr}}, \bar{\bar{\theta}}_{\text{nr}}, \zeta)$. We now subtract the following total ζ derivative from the Lagrangian:

$$\begin{aligned} \frac{d}{d\zeta} \left\{ \left(\bar{\bar{\psi}}_{\text{nr}} + \psi_{\text{nr}}^{(2)} + \psi_{\text{nr}}^{(1)} \left[\partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) - \partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\theta} \psi_{\text{nr}}^{(1)} \right) \theta_{\text{nr}}^{(2)} \right. \\ \left. + \left(\theta_{\text{nr}}^{(1)} \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) - \partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\psi} \psi_{\text{nr}}^{(1)} \right) \psi_{\text{nr}}^{(2)} + S_{\text{nr}}^{(2)} \right\}, \end{aligned} \quad (\text{J } 4)$$

where $S_{\text{nr}}^{(2)} \sim \rho_{\star}^2 \Psi_{\text{t}}$ is an as-yet-undetermined function. The result is

$$\begin{aligned} \mathcal{L} \equiv & \left(\bar{\bar{\psi}}_{\text{nr}} + \psi_{\text{nr}}^{(2)} + \psi_{\text{nr}}^{(1)} \left[\partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right. \right. \\ & \left. \left. - \partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\theta} \psi_{\text{nr}}^{(1)} - \partial_{\theta} S_{\text{nr}}^{(2)} \right) \frac{d\bar{\bar{\theta}}_{\text{nr}}}{d\zeta} \\ & + \left(-\theta_{\text{nr}}^{(2)} + \theta_{\text{nr}}^{(1)} \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) - \partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\psi} \psi_{\text{nr}}^{(1)} - \partial_{\psi} S_{\text{nr}}^{(2)} \right) \frac{d\bar{\bar{\psi}}_{\text{nr}}}{d\zeta} \\ & - \chi - \iota \psi_{\text{nr}}^{(2)} - \frac{1}{2} \iota' (\psi_{\text{nr}}^{(1)})^2 + \psi_{\text{nr}}^{(1)} \left[\partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) \right] \\ & + \theta_{\text{nr}}^{(1)} \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\zeta} \psi_{\text{nr}}^{(1)} + V_{\text{r}} + V_{\text{nr}}^{\geq K} - \partial_{\zeta} S_{\text{nr}}^{(2)}. \end{aligned} \quad (\text{J } 5)$$

In order to eliminate the ζ -dependent terms in the symplectic parts of this Lagrangian, we pick

$$\psi_{\text{nr}}^{(2)} := \psi_{\text{nr}}^{(1)} \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) - \partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] + \theta_{\text{nr}}^{(1)} \partial_{\theta} \psi_{\text{nr}}^{(1)} + \partial_{\theta} S_{\text{nr}}^{(2)}, \quad (\text{J } 6a)$$

$$\theta_{\text{nr}}^{(2)} := \theta_{\text{nr}}^{(1)} \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) - \partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] - \theta_{\text{nr}}^{(1)} \partial_{\psi} \psi_{\text{nr}}^{(1)} - \partial_{\psi} S_{\text{nr}}^{(2)}, \quad (\text{J } 6b)$$

for whatever choice of $S_{\text{nr}}^{(2)}$ we eventually make. This leaves

$$\begin{aligned} \mathcal{L} \equiv \bar{\bar{\psi}}_{\text{nr}} \frac{d\bar{\bar{\theta}}_{\text{nr}}}{d\zeta} - \chi - \frac{1}{2} \iota' (\psi_{\text{nr}}^{(1)})^2 + V_{\text{r}} + V_{\text{nr}}^{\geq K} \\ + \iota \psi_{\text{nr}}^{(1)} \left[\partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) - \partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) \right] \\ + \psi_{\text{nr}}^{(1)} \left[\partial_{\psi} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\psi} \mathbf{x} \right) \right] \\ + \theta_{\text{nr}}^{(1)} \left[\partial_{\theta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\zeta} \mathbf{x} \right) - \partial_{\zeta} \left(\frac{v_{\parallel}}{\Omega} \mathbf{B} \cdot \partial_{\theta} \mathbf{x} \right) \right] \\ - \theta_{\text{nr}}^{(1)} (\partial_{\zeta} \psi_{\text{nr}}^{(1)} + \iota \partial_{\theta} \psi_{\text{nr}}^{(1)}) - (\partial_{\zeta} S_{\text{nr}}^{(2)} + \iota \partial_{\theta} S_{\text{nr}}^{(2)}), \end{aligned} \quad (\text{J } 7)$$

After some algebra, using the relations (7.34), (7.36), and $(v_{\parallel}/\Omega) \mathbf{B} \cdot (\partial_{\zeta} \mathbf{x} + \iota \partial_{\theta} \mathbf{x}) \simeq V_{\text{r}} + V_{\text{nr}}^{<K}$ to rewrite several terms, this Lagrangian simplifies to

$$\begin{aligned} \mathcal{L} \equiv \bar{\bar{\psi}}_{\text{nr}} \frac{d\bar{\bar{\theta}}_{\text{nr}}}{d\zeta} - \chi + V_{\text{r}} + V_{\text{nr}}^{\geq K} + (\psi_{\text{nr}}^{(1)} \partial_{\psi} V_{\text{r}} + \theta_{\text{nr}}^{(1)} \partial_{\theta} V_{\text{r}}) + \frac{1}{2} \iota' (\psi_{\text{nr}}^{(1)})^2 \\ - \psi_{\text{nr}}^{(1)} (\partial_{\zeta} \theta_{\text{nr}}^{(1)} + \iota \partial_{\theta} \theta_{\text{nr}}^{(1)}) - (\partial_{\zeta} S_{\text{nr}}^{(2)} + \iota \partial_{\theta} S_{\text{nr}}^{(2)}). \end{aligned} \quad (\text{J } 8)$$

Let V_2 be defined by

$$V_2 := V_{\text{nr}}^{\geq K} + (\psi_{\text{nr}}^{(1)} \partial_{\psi} V_{\text{r}} + \theta_{\text{nr}}^{(1)} \partial_{\theta} V_{\text{r}}) + \frac{1}{2} \iota' (\psi_{\text{nr}}^{(1)})^2 - \psi_{\text{nr}}^{(1)} (\partial_{\zeta} \theta_{\text{nr}}^{(1)} + \iota \partial_{\theta} \theta_{\text{nr}}^{(1)}), \quad (\text{J } 9)$$

as announced in (7.49). Then, the Lagrangian takes the form

$$\mathcal{L} \equiv \bar{\bar{\psi}}_{\text{nr}} \frac{d\bar{\bar{\theta}}_{\text{nr}}}{d\zeta} - \chi + V_{\text{r}} + V_2 - (\partial_{\zeta} S_{\text{nr}}^{(2)} + \iota \partial_{\theta} S_{\text{nr}}^{(2)}). \quad (\text{J } 10)$$

Let $V_2 = \langle V_2 \rangle_{\zeta} + V_{2,\text{nr}}^{<K_2} + V_{2,\text{nr}}^{\geq K_2}$. We choose $S_{\text{nr}}^{(2)}$ to remove the angle dependence contained in the low-order, non-resonant harmonics $V_{2,\text{nr}}^{<K_2}$; that is, we take

$$S_{\text{nr}}^{(2)} := \sum_{(p,q) \in \mathbb{I}_{\text{nr}}^{\leq K_2}} \frac{i(V_{2,\text{nr}}^{<K_2})_{pq}}{q - \iota p} e^{i(p\theta - q\zeta)}, \quad (\text{J } 11)$$

as in (7.52), so that $\partial_{\zeta} S_{\text{nr}}^{(2)} + \iota \partial_{\theta} S_{\text{nr}}^{(2)} = V_{2,\text{nr}}^{<K_2}$. If we define $\bar{\eta}_{\text{nr}} = \bar{\bar{\theta}}_{\text{nr}} - (N/M)\zeta$, then our final result may be written

$$\mathcal{L} \equiv \bar{\bar{\psi}}_{\text{nr}} \frac{d\bar{\eta}_{\text{nr}}}{d\zeta} - \chi_{\text{r}} + \langle V \rangle_{\zeta} + \langle V_2 \rangle_{\zeta}. \quad (\text{J } 12)$$

This Lagrangian does not depend on ζ , so the Hamiltonian

$$\mathcal{H}_{2,\text{r}}(\bar{\bar{\psi}}_{\text{nr}}, \bar{\eta}_{\text{nr}}) = \chi_{\text{r}}(\bar{\bar{\psi}}_{\text{nr}}) - \langle V \rangle_{\zeta}(\bar{\bar{\psi}}_{\text{nr}}, \bar{\eta}_{\text{nr}}) - \langle V_2 \rangle_{\zeta}(\bar{\bar{\psi}}_{\text{nr}}, \bar{\eta}_{\text{nr}}) \quad (\text{J } 13)$$

is conserved, as claimed.

REFERENCES

- ALBERT, C. G., BUCHHOLZ, R., KASILOV, S. V., KERNBICHLER, W. & RATH, K. 2023 Alpha particle confinement metrics based on orbit classification in stellarators. *J. Plasma Phys.* **89** (3), 955890301.
- ARNOLD, V. I. 1963*a* Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Mathematical Surveys*

- 18** (5), 9–36, translated into English in *Collected Works: Representations of Functions, Celestial Mechanics and KAM Theory, 1957–1965*, Springer, 2009, pp. 267–294.
- ARNOLD, V. I. 1963*b* Small denominators and problems of stability of motion in classical and celestial mechanics. *Russian Mathematical Surveys* **18** (6), 85.
- ARNOLD, V. I., KOZLOV, V. V., NEISHTADT, A. I. & IACOB, I. 2006 *Mathematical Aspects of Classical and Celestial Mechanics*. Springer.
- DE ASSUNÇÃO, M. 2023 Integration algorithms for charged-particle dynamics in magnetised plasmas. Master’s thesis, Técnico Lisboa.
- BADER, A., ANDERSON, D. T., DREVLAK, M., FABER, B. J., HEGNA, C. C., HENNEBERG, S., LANDREMAN, M., SCHMITT, J. C., SUZUKI, Y. & WARE, A. 2021 Modeling of energetic particle transport in optimized stellarators. *Nucl. Fusion* **61** (11), 116060.
- BADER, A., FABER, B. J., SCHMITT, J. C., ANDERSON, D. T., DREVLAK, M., DUFF, J. M., FRERICH, H., HEGNA, C. C., KRUGER, T. G., LANDREMAN, M. & OTHERS 2020 Advancing the physics basis for quasi-helically symmetric stellarators. *J. Plasma Phys.* **86** (5), 905860506.
- BEIDLER, C., GRIEGER, G., HERRNEGGER, F., HARMEYER, E., KISSLINGER, J., LOTZ, W., MAASSBERG, H., MERKEL, P., NÜHRENBERG, J., RAU, F. & OTHERS 1990 Physics and engineering design for Wendelstein VII-X. *Fusion Technology* **17** (1), 148–168.
- BEIDLER, C. D., ALLMAIER, K., ISAEV, M. Y., KASILOV, S. V., KERNBICHLER, W., LEITOLD, G. O., MAASSBERG, H., MIKKELSEN, D. R., MURAKAMI, S., SCHMIDT, M. & OTHERS 2011 Benchmarking of the mono-energetic transport coefficients—results from the international collaboration on neoclassical transport in stellarators (ICNTS). *Nucl. Fusion* **51** (7), 076001.
- BHATTACHARJEE, A., SCHREIBER, G. M. & TAYLOR, J. B. 1992 Geometric phase, rotational transforms, and adiabatic invariants in toroidal magnetic fields. *Phys. Fluids B* **4** (9), 2737–2739.
- BINDEL, D., LANDREMAN, M. & PADIDAR, M. 2023 Direct optimization of fast-ion confinement in stellarators. *Plasma Phys. Controlled Fusion* **65** (6), 065012.
- BISHOP, A. S. & SMITH, C. G. 1966 Microscopic treatment of classical containment in the stellarator. *Phys. Fluids* **9** (7), 1380–1385.
- BOOZER, A. H. 1981 Plasma equilibrium with rational magnetic surfaces. *Phys. Fluids* **24**, 1999.
- BOOZER, A. H. 2005 Physics of magnetically confined plasmas. *Reviews of Modern Physics* **76** (4), 1071.
- BOOZER, A. H. 2016 Runaway electrons and magnetic island confinement. *Phys. Plasmas* **23** (8), 082514.
- BRAMBILLA, M. & LICHTENBERG, A. J. 1973 Drift-surface-island formation and particle diffusion in toroidal plasma. *Nucl. Fusion* **13** (4), 517.
- BURBY, J. W., MACKAY, R. S. & NAIK, S. 2023 Isodrastic magnetic fields for suppressing transitions in guiding-centre motion. *Nonlinearity* **36** (11), 5884.
- BURBY, J. W., MALDONADO, I. A., RUTH, M., MESSENGER, D. A. & CARBAJAL, L. 2025 Nonperturbative guiding center model for magnetized plasmas. *Phys. Rev. Lett.* **134** (17), 175101.
- CALVO, I., PARRA, F. I., VELASCO, J. L. & ALONSO, J. A. 2017 The effect of tangential drifts on neoclassical transport in stellarators close to omnigenity. *Plasma Phys. Controlled Fusion* **59** (5), 055014.
- CALVO, I., VELASCO, J. L., HELANDER, P. & PARRA, F. I. 2025 Piecewise omnigenous stellarators with zero bootstrap current. *Physical Review E* **112** (2), L023201.
- CALVO, I., VELASCO, J. L., PARRA, F. I., ALONSO, J. A. & GARCÍA-REGAÑA, J. M. 2018 Electrostatic potential variations on stellarator magnetic surfaces in low collisionality regimes. *J. Plasma Phys.* **84** (4), 905840407.
- CARBAJAL, L., VARELA, J., BADER, A., GUTTENFELDER, W., CERFON, A., SCHMITT, J. C., MORRISSEY, J., HEGNA, C. C., CANIK, J. M., MANDELL, N. R. & OTHERS 2025 Alpha-particle confinement in Infinity Two Fusion Pilot Plant baseline plasma design. *J. Plasma Phys.* pp. 1–30.
- CARY, J. R. 1981 Lie transform perturbation theory for Hamiltonian systems. *Physics Reports* **79** (2), 129–159.

- CARY, J. R. & BRIZARD, A. J. 2009 Hamiltonian theory of guiding-center motion. *Reviews of Modern Physics* **81** (2), 693.
- CARY, J. R., ESCANDE, D. F. & TENNYSON, J. L. 1986 Adiabatic-invariant change due to separatrix crossing. *Physical Review A* **34** (5), 4256.
- CARY, J. R., HEDRICK, C. L. & TOLLIVER, J. S. 1988 Orbits in asymmetric toroidal magnetic fields. *Phys. Fluids* **31** (6), 1586–1600.
- CARY, J. R. & KOTSCHENREUTHER, M. 1985 Pressure induced islands in three-dimensional toroidal plasma. *Phys. Fluids* **28** (5), 1392–1401.
- CARY, J. R. & LITTLEJOHN, R. G. 1983 Noncanonical Hamiltonian mechanics and its application to magnetic field line flow. *Annals of Physics* **151** (1), 1–34.
- CARY, J. R. & SHASHARINA, S. G. 1997a Helical plasma confinement devices with good confinement properties. *Phys. Rev. Lett.* **78** (4), 674.
- CARY, J. R. & SHASHARINA, S. G. 1997b Omnigenity and quasihelicity in helical plasma confinement systems. *Phys. Plasmas* **4** (9), 3323–3333.
- CHAMBLISS, A., PAUL, E. & HUDSON, S. 2025 Fast particle trajectories and integrability in quasisymmetric and quasihelical stellarators. *J. Plasma Phys.* **91** (3), E74.
- CHIRIKOV, B. V. 1979 A universal instability of many-dimensional oscillator systems. *Physics Reports* **52** (5), 263–379.
- CHU, X., QIAN, T., BERLINGER, B., HAMMOND, K., HAQUE, M., SALIA, S., WILLIAMS, D. J., ZHU, E. & ZARNSTORFF, M. C. 2025 Error field measurements and correction on muse permanent magnet stellarator. *Plasma Phys. Controlled Fusion*, in press.
- DEL-CASTILLO-NEGRETE, D., GREENE, J. M. & MORRISON, P. J. 1996 Area preserving nontwist maps: periodic orbits and transition to chaos. *Physica D: Nonlinear Phenomena* **91** (1–2), 1–23.
- DEPRIT, A. 1969 Canonical transformations depending on a small parameter. *Celestial Mechanics* **1** (1), 12–30.
- D’HERBEMONT, V., PARRA, F. I., CALVO, I. & VELASCO, J. L. 2022 Finite orbit width effects in large aspect ratio stellarators. *J. Plasma Phys.* **88** (5), 905880507.
- DOBROTT, D. & FRIEMAN, E. A. 1971 Magnetic and drift surfaces using a new stellarator expansion. *Phys. Fluids* **14** (2), 349–360.
- DREVLAK, M., GEIGER, J., HELANDER, P. & TURKIN, Y. 2014 Fast particle confinement with optimized coil currents in the W7-X stellarator. *Nucl. Fusion* **54** (7), 073002.
- DREVLAK, M., MONTICELLO, D. & REIMAN, A. 2005 PIES free boundary stellarator equilibria with improved initial conditions. *Nucl. Fusion* **45** (7), 731.
- DUDKOVSKAIA, A. V., CONNOR, J. W., DICKINSON, D., HILL, P., IMADA, K., LEIGH, S. & WILSON, H. R. 2023 Drift kinetic theory of neoclassical tearing modes in tokamak plasmas: polarisation current and its effect on magnetic island threshold physics. *Nucl. Fusion* **63** (12), 126040.
- DUDT, D. W. & KOLEMEN, E. 2020 DESC: a stellarator equilibrium solver. *Phys. Plasmas* **27** (10).
- ELBARMÍ, E., SENGUPTA, W. & WEITZNER, H. 2020 Charged particle dynamics near an x-point of a non-symmetric magnetic field with closed field lines. *J. Plasma Phys.* **86** (2), 905860209.
- ESCANDE, D. F. & MOMO, B. 2024 Description of magnetic field lines without arcana. *Reviews of Modern Plasma Physics* **8** (1), 16.
- FIGUEIREDO, P. A., JORGE, R., FERREIRA, J. & RODRIGUES, P. 2024 Energetic particle tracing in optimized quasi-symmetric stellarator equilibria. *J. Plasma Phys.* **90** (2), 905900207.
- FOSTER, T. E., PARRA, F. I. & WHITE, R. B. 2025 Energetic-particle orbits near rational flux surfaces in stellarators: II. Trapped particles. In prep.
- GALEEV, A. A., SAGDEEV, R. Z., FURTH, H. P. & ROSENBLUTH, M. N. 1969 Plasma diffusion in a toroidal stellarator. *Phys. Rev. Lett.* **22** (11), 511.
- GARCÍA-REGAÑA, J. M., CALVO, I., SANCHEZ, E., THIENPOND, H., VELASCO, J. L. & CAPITÁN, J. A. 2024 Reduced electrostatic turbulence in the quasi-isodynamic stellarator configuration CIEMAT-QI4. *Nucl. Fusion* **65** (1), 016036.
- GOLDSTEIN, H., POOLE, C. & SAFKO, J. 2002 *Classical Mechanics*, 3rd edn. Boston: Addison Wesley.
- GOODMAN, A. G., MATA, K. C., HENNEBERG, S. A., JORGE, R., LANDREMAN, M., PLUNK,

- G. G., SMITH, H. M., MACKENBACH, R. J. J., BEIDLER, C. D. & HELANDER, P. 2023 Constructing precisely quasi-isodynamic magnetic fields. *J. Plasma Phys.* **89** (5), 905890504.
- GRAD, H. 1967 Toroidal containment of a plasma. *Phys. Fluids* **10** (1), 137–154.
- HALL, L. S. & MCNAMARA, B. 1975 Three-dimensional equilibrium of the anisotropic, finite-pressure guiding-center plasma: Theory of the magnetic plasma. *Phys. Fluids* **18** (5), 552–565.
- HAMADA, S. 1962 Hydromagnetic equilibria and their proper coordinates. *Nucl. Fusion* **2** (1-2), 23.
- HAMILTON, C., TOLMAN, E. A., ARZAMASSKIY, L. & DUARTE, V. N. 2023 Galactic bar resonances with diffusion: an analytic model with implications for bar–dark matter halo dynamical friction. *The Astrophysical Journal* **954** (1), 12.
- HASTIE, R. J., TAYLOR, J. B. & HAAS, F. A. 1967 Adiabatic invariants and the equilibrium of magnetically trapped particles. *Annals of Physics* **41** (2), 302–338.
- HAZELTINE, R. D. & CATTO, P. J. 2015 Plasma current resonance in asymmetric toroidal systems. *Phys. Plasmas* **22** (9).
- HAZELTINE, R. D. & WAELEBROECK, F. L. 2018 *The Framework of Plasma Physics*. CRC Press.
- HE, K., SUN, Y., WAN, B. N., GU, S. & JIA, M. 2020 Roles of primary and sideband resonances in the confinement of energetic passing ions under resonant magnetic perturbations. *Nucl. Fusion* **60** (12), 126027.
- HEGNA, C. C. & BHATTACHARJEE, A. 1989 Magnetic island formation in three-dimensional plasma equilibria. *Phys. Fluids B* **1** (2), 392.
- HELANDER, P. 2014 Theory of plasma confinement in non-axisymmetric magnetic fields. *Rep. Prog. Phys.* **77** (8), 087001.
- HELANDER, P. & NÜHRENBERG, J. 2009 Bootstrap current and neoclassical transport in quasi-isodynamic stellarators. *Plasma Phys. Controlled Fusion* **51** (5), 055004.
- HENRARD, J. 1993 The adiabatic invariant in classical mechanics. In *Dynamics Reported: Expositions in Dynamical Systems*, pp. 117–235. Springer.
- HEYN, M. F., IVANOV, I. B., KASILOV, S. V., KERNBICHLER, W., LOARTE, A., NEMOV, V. V. & RUNOV, A. M. 2012 On the confinement of passing alpha particles in a tokamak-reactor with resonant magnetic field perturbations shielded by plasma currents. *Nucl. Fusion* **52** (5), 054010.
- HIRSHMAN, S. P., MERKEL, P. & OTHERS 1986 Three-dimensional free boundary calculations using a spectral Green’s function method. *Computer Physics Communications* **43** (1), 143–155.
- HORI, G. 1966 Theory of general perturbation with unspecified canonical variable. *Publications of the Astronomical Society of Japan* **18**, 287.
- IMADA, K., WILSON, H. R., CONNOR, J. W., DUDKOVSKAIA, A. V. & HILL, P. 2018 Nonlinear kinetic ion response to small scale magnetic islands in tokamak plasmas: neoclassical tearing mode threshold physics. *Phys. Rev. Lett.* **121** (17), 175001.
- IMBERT-GÉRARD, L.-M., PAUL, E. J. & WRIGHT, A. M. 2024 *An Introduction to Stellarators: From Magnetic Fields to Symmetries and Optimization*. SIAM.
- KOLESNICHENKO, Y. I., LUTSENKO, V. V., WELLER, A., WERNER, A., WOBIG, H., YAKOVENKO, Y. V., GEIGER, J. & ZEGENHAGEN, S. 2006 Effects of fast-ion-orbit width on Alfvén instabilities in stellarators: a general theory and its application to a W7-AS experiment. *Nucl. Fusion* **46** (8), 753.
- KOMORI, A., MORISAKI, T., MUTOH, T., SAKAKIBARA, S., TAKEIRI, Y., KUMAZAWA, R., KUBO, S., IDA, K., MORITA, S., NARIHARA, K. & OTHERS 2006 Overview of progress in LHD experiments. *Fusion Science and Technology* **50** (2), 136–145.
- KRUSKAL, M. 1962 Asymptotic theory of Hamiltonian and other systems with all solutions nearly periodic. *Journal of Mathematical Physics* **3** (4), 806–828.
- KRUSKAL, M. D. & KULSRUD, R. M. 1958 Equilibrium of a magnetically confined plasma in a toroid. *Phys. Fluids* **1** (4), 265–274.
- KU, L.-P., GARABEDIAN, P. R., LYON, J., TURNBULL, A., GROSSMAN, A., MAU, T. K., ZARNSTORFF, M. & TEAM, ARIES 2008 Physics design for ARIES-CS. *Fusion Science and Technology* **54** (3), 673–693.

- LANDREMAN, M. & CATTO, P. J. 2012 Omnigenity as generalized quasisymmetry. *Phys. Plasmas* **19** (5).
- LANDREMAN, M., MEDASANI, B. & ZHU, C. 2021 Stellarator optimization for good magnetic surfaces at the same time as quasisymmetry. *Phys. Plasmas* **28** (9).
- LANDREMAN, M. & PAUL, E. 2022 Magnetic fields with precise quasisymmetry for plasma confinement. *Phys. Rev. Lett.* **128** (3), 035001.
- LENARD, A. 1959 Adiabatic invariance to all orders. *Annals of Physics* **6** (3), 261–276.
- LEVINESS, A., SCHMITT, J. C., LAZERSON, S. A., BADER, A., FABER, B. J., HAMMOND, K. C. & GATES, D. A. 2022 Energetic particle optimization of quasi-axisymmetric stellarator equilibria. *Nucl. Fusion* **63** (1), 016018.
- LICHTENBERG, A. J. & LIEBERMAN, M. A. 2013 *Regular and Stochastic Motion*. Springer Science & Business Media.
- LITTLEJOHN, R. G. 1983 Variational principles of guiding centre motion. *J. Plasma Phys.* **29** (1), 111–125.
- LIU, Y., LI, L., LOARTE, A., PINCHES, S. D. & POLEVOI, A. 2021 Drift orbit islands of energetic particles due to 3D fields in ITER. *Nucl. Fusion* **61** (10), 106029.
- LOTZ, W., MERKEL, P., NÜHRENBURG, J. & STRUMBERGER, E. 1992 Collisionless alpha-particle confinement in stellarators. *Plasma Phys. Controlled Fusion* **34** (6), 1037.
- MARCAL, M. A. 1982 *Magnetic and Drift Surfaces in Toroidal Plasma Equilibria*. New York University.
- MATSUYAMA, A., YAGI, M., KAGEI, Y. & NAKAJIMA, N. 2014 Drift resonance effect on stochastic runaway electron orbit in the presence of low-order magnetic perturbations. *Nucl. Fusion* **54** (12), 123007.
- MIKHAILOV, M. I., SHAFRANOV, V. D., SUBBOTIN, A. A., ISAEV, M. Y., NÜHRENBURG, J., ZILLE, R. & COOPER, W. A. 2002 Improved α -particle confinement in stellarators with poloidally closed contours of the magnetic field strength. *Nucl. Fusion* **42** (11), L23.
- MORBIDELLI, A. 2002 *Modern Celestial Mechanics: Aspects of Solar System Dynamics*. London: Taylor & Francis.
- MOROZOV, A. I. & SOLOV'EV, L. S. 1959 Integrals of the drift equations. In *Soviet Physics Doklady*, , vol. 4, p. 1031.
- MOROZOV, A. I. & SOLOV'EV, L. S. 1966 Motion of charged particles in electromagnetic fields. *Reviews of Plasma Physics*, vol. 2, pp. 201–297. New York: Consultants Bureau.
- MOTOJIMA, O. & SHISHKIN, A. A. 1999 Drift island motion in helical plasma and its use for ash removal and high-energy ion injection. *Plasma Phys. Controlled Fusion* **41** (2), 227.
- MYNICK, H. E. 1993a Stochastic transport of MeV ions by low- n magnetic perturbations. *Phys. Fluids B* **5** (7), 2460–2467.
- MYNICK, H. E. 1993b Transport of energetic ions by low- n magnetic perturbations. *Phys. Fluids B* **5** (5), 1471–1481.
- NEWCOMB, W. A. 1959 Magnetic differential equations. *Phys. Fluids* **2** (4), 362–365.
- NIES, R., PAUL, E. J., PANICI, D., HUDSON, S. R. & BHATTACHARJEE, A. 2024 Exploration of the parameter space of quasisymmetric stellarator vacuum fields through adjoint optimisation. *J. Plasma Phys.* **90** (6), 905900620.
- PARRA, F. I. & CALVO, I. 2011 Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry. *Plasma Phys. Controlled Fusion* **53** (4), 045001.
- PARRA, F. I., CALVO, I., HELANDER, P. & LANDREMAN, M. 2015 Less constrained omnigenous stellarators. *Nucl. Fusion* **55** (3), 033005.
- PAUL, E. J., BHATTACHARJEE, A., LANDREMAN, M., ALEX, D., VELASCO, J. L. & NIES, R. 2022 Energetic particle loss mechanisms in reactor-scale equilibria close to quasisymmetry. *Nucl. Fusion* **62** (12), 126054.
- PAUL, E. J., MYNICK, H. E. & BHATTACHARJEE, A. 2023 Fast-ion transport in quasisymmetric equilibria in the presence of a resonant Alfvénic perturbation. *J. Plasma Phys.* **89** (5), 905890515.
- REIMAN, A. & BOOZER, A. H. 1984 Island formation and destruction of flux surfaces in three-dimensional MHD equilibria. *Phys. Fluids* **27** (10), 2446–2454.
- RODRIGUES, P., ASSUNÇÃO, M., GOMES, E., FERREIRA, J. & JORGE, R. 2024 Guiding centre, full gyromotion, and fast-ion losses in stellarators. In *Proceedings of the 50th Conference on Controlled Fusion and Plasma Physics*. P2.093.

- ROME, J. A. 1995 Orbit topology in conventional stellarators in the presence of electric fields. *Nucl. Fusion* **35** (2), 195.
- DE ROVER, M., LOPES CARDOZO, N. J. & MONTVAI, A. 1996 Motion of relativistic particles in axially symmetric and perturbed magnetic fields in a tokamak. *Phys. Plasmas* **3** (12), 4478–4488.
- RUDIN, W. 1976 *Principles of Mathematical Analysis*. New York: McGraw-hill.
- RÜSSMANN, H. 1975 On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. *Dynamical Systems, Theory and Applications* **38**, 598–624.
- SÁNCHEZ, E., VELASCO, J. L., CALVO, I. & MULAS, S. 2023 A quasi-isodynamic configuration with good confinement of fast ions at low plasma β . *Nucl. Fusion* **63** (6), 066037.
- SHINOHARA, K., BIERWAGE, A., MATSUYAMA, A., SUZUKI, Y., MATSUNAGA, G., HONDA, M., SUMIDA, S. & KIM, J. 2020 Efficient estimation of drift orbit island width for passing ions in a shaped tokamak plasma with a static magnetic perturbation. *Nucl. Fusion* **60** (9), 096032.
- SKOVORODA, A. A. 2005 3D toroidal geometry of currentless magnetic configurations with improved confinement. *Plasma Phys. Controlled Fusion* **47** (11), 1911.
- SMIRNOVA, M. S. 1997 Improvement of the highly energetic passing particle confinement in torsatrons and heliotrons. *Phys. Plasmas* **4** (7), 2584–2596.
- TODOROKI, J. 1993 Theory of longitudinal adiabatic invariant in the helical torus. *Journal of the Physical Society of Japan* **62** (5), 1562–1582.
- TYKHYI, A. V. & YAKOVENKO, Y. V. 2006 Invariants of fast ion motion in stellarators. *Ukrainian Journal of Physics* **51** (11-12), 1077–1082.
- VARJE, J., SÄRKIMÄKI, K., KONTULA, J., OLLUS, P., KURKI-SUONIO, T., SNICKER, A., HIRVIJOKI, E. & ÄKÄSLÖMPOLO, S. 2019 High-performance orbit-following code ASCOT5 for Monte Carlo simulations in fusion plasmas. *E-print arXiv:1908.02482*.
- VELASCO, J. L., CALVO, I., ESCOTO, F. J., SÁNCHEZ, E., THIENPOND, H. & PARRA, F. I. 2024 Piecewise omnigenous stellarators. *Phys. Rev. Lett.* **133** (18), 185101.
- VELASCO, J. L., CALVO, I., MULAS, S., SÁNCHEZ, E., PARRA, F. I., CAPPA, A. & OTHERS 2021 A model for the fast evaluation of prompt losses of energetic ions in stellarators. *Nucl. Fusion* **61** (11), 116059.
- VELASCO, J. L., CALVO, I., SÁNCHEZ, E. & PARRA, F. I. 2023 Robust stellarator optimization via flat mirror magnetic fields. *Nucl. Fusion* **63** (12), 126038.
- VELASCO, J. L., SÁNCHEZ, E. & CALVO, I. 2025 Exploration of the parameter space of piecewise omnigenous stellarator magnetic fields. *Nucl. Fusion* **65** (5), 056012.
- VIANA, R. L., MUGNAINE, M. & CALDAS, I. L. 2023 Hamiltonian description for magnetic field lines in fusion plasmas: A tutorial. *Phys. Plasmas* **30** (9).
- WHITE, R. 2022 Resonant alpha particle loss in stellarators. *Phys. Plasmas* **29** (9), 092504.
- WHITE, R., BIERWAGE, A. & ETHIER, S. 2022 Poor confinement in stellarators at high energy. *Phys. Plasmas* **29** (5), 052511.
- WHITE, R. B. 2025 Poor confinement in stellarators at high energy II. *Phys. Plasmas* **32** (4).
- WHITE, R. B., BOOZER, A. H. & HAY, R. 1982 Drift Hamiltonian in magnetic coordinates. *Phys. Fluids* **25** (3), 575–576.
- WOBIG, H. & PFIRSCH, D. 2001 On guiding centre orbits of particles in toroidal systems. *Plasma Phys. Controlled Fusion* **43** (5), 695.
- YOCCOZ, J.-C. 1992 An introduction to small divisors problems. In *From Number Theory to Physics*, pp. 659–679. Springer.