

Finite orbit width effects in large aspect ratio stellarators

Vincent d'Herbement ^{1,2}, Felix I. Parra ^{3,†}, Iván Calvo ⁴ and José Luis Velasco ⁴

¹Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford OX1 3PU, UK

²Mines ParisTech, 75 272 Paris, France

³Princeton Plasma Physics Laboratory, Princeton, NJ 08540, USA

⁴Laboratorio Nacional de Fusión, CIEMAT, 28040 Madrid, Spain

(Received 2 February 2022; revised 2 September 2022; accepted 5 September 2022)

New orbit-averaged equations for low collisionality neoclassical fluxes in large aspect ratio stellarators with mirror ratios close to unity are derived. The equations retain finite orbit width effects by employing the second adiabatic invariant J as a velocity-space coordinate and they have been implemented in the orbit-averaged neoclassical code KNOSOS (Velasco *et al.*, *J. Comput. Phys.*, vol. 418, 2020, 109512; Velasco *et al.*, *Nucl. Fusion*, vol. 61, 2021, 116013). The equations are used to study the $1/\nu$ regime and the lower collisionality regimes. For generic large aspect ratio stellarators with mirror ratios close to unity, as the collision frequency decreases, the $1/\nu$ regime transitions directly into the ν regime, without passing through a $\sqrt{\nu}$ regime. An explicit formula for the neoclassical fluxes in the ν regime is obtained. The formula includes the effect of particles that transition between different types of wells. While these transitions produce stochastic scattering independent of the value of the collision frequency in velocity space, the diffusion in real space remains proportional to the collision frequency. The $\sqrt{\nu}$ regime is only recovered in large aspect ratio stellarators close to omnigenity: large aspect ratio stellarators with large mirror ratios and optimized large aspect ratio stellarators with mirror ratios close to unity. Neoclassical transport in large aspect ratio stellarators with large mirror ratios can be calculated with the orbit-averaged equations derived by Calvo *et al.* (*Plasma Phys. Control. Fusion*, vol. 59, 2017, 055014). In these stellarators, the $\sqrt{\nu}$ regime exists in the collisionality interval $(a/R) \ln(R/a) \ll \nu_{ii} Ra / \rho_i v_{ii} \ll R/a$. In optimized large aspect ratio stellarators with mirror ratios close to unity, the $\sqrt{\nu}$ regime occurs in an interval of collisionality that depends on the deviation from omnigenity δ : $\delta^2 |\ln \delta| \ll \nu_{ii} Ra / \rho_i v_{ii} \ll 1$. Here, ν_{ii} is the ion–ion collision frequency, ρ_i and v_{ii} are the ion gyroradius and thermal speed, and a and R are the minor and major radii.

Key words: fusion plasma, plasma confinement

1. Introduction

In stellarators, trapped particles can move a significant distance away from their initial flux surface even in the absence of collisions or turbulent fluctuations. Due to these large

† Email address for correspondence: fparradi@pppl.gov

orbits, stellarator collisional transport at low collision frequencies (Kovrizhnykh 1984) is of the order of or larger than the turbulent transport, dominating energy transport in the core (Dinklage *et al.* 2013, 2018).

The width of the trapped-particle orbits is of the order of the size of the stellarator unless the stellarator is (i) optimized (Calvo *et al.* 2017), i.e. close to omnigeneous (Cary & Shasharina 1997a,b; Parra *et al.* 2015), or (ii) the stellarator has a small inverse aspect ratio $\epsilon := a/R \ll 1$ (Ho & Kulsrud 1987), where R and a are the characteristic values of the major and minor radii of the stellarator, respectively.

In the case of large aspect ratio stellarators, the ion orbit width is determined by the balance between the component of the $\mathbf{E} \times \mathbf{B}$ drift that is tangential to the flux surface, and the component of the ion curvature and ∇B drifts that is perpendicular to the flux surface – the large $\mathbf{E} \times \mathbf{B}$ drift is mostly parallel to flux surfaces and its small radial component is comparable to or smaller than the average radial component of the curvature and ∇B drifts (Calvo *et al.* 2017).

To estimate the width of the ion orbits, we assume that the electric field is the gradient of an electric potential, and that the potential has a variation of order T_i/e across the minor radius a , where T_i is the ion temperature and e is the proton charge. The lowest-order value of the radial electric field is set by the need to maintain ambipolarity. The $\mathbf{E} \times \mathbf{B}$ drift is of order

$$\mathbf{v}_E \sim \rho_{i*} v_{ti}, \quad (1.1)$$

whereas the curvature and ∇B drifts are smaller by a factor of ϵ because they are proportional to the gradient of the magnetic field \mathbf{B} , $|\nabla \mathbf{B}| \sim B/R$,

$$\mathbf{v}_{Mi} \sim \epsilon \rho_{i*} v_{ti}. \quad (1.2)$$

Here

$$\rho_{i*} := \frac{\rho_i}{a} \ll 1 \quad (1.3)$$

is the normalized ion gyroradius, $\rho_i := v_{ti}/\Omega_i$ is the ion gyroradius, $v_{ti} := \sqrt{2T_i/m_i}$ is the ion thermal speed, $\Omega_i := Z_i e B / m_i c$ is the ion gyrofrequency, $B := |\mathbf{B}|$ is the magnitude of the magnetic field, $Z_i e$ and m_i are the ion charge and mass, respectively, and c is the speed of light. The radial motion due to the drifts is not secular because it averages out once the $\mathbf{E} \times \mathbf{B}$ drift has moved the particle several times around the stellarator. The typical length of the orbit parallel to the flux surfaces is of order a , giving a characteristic orbital time of the order of

$$\frac{a}{|\mathbf{v}_E|} \sim \frac{1}{\rho_{i*}} \frac{a}{v_{ti}}. \quad (1.4)$$

During this time interval, the radial component of the curvature and ∇B drifts (and, in some cases, the $\mathbf{E} \times \mathbf{B}$ drift) leads to an orbit width of order

$$w \sim \frac{a}{|\mathbf{v}_E|} |\mathbf{v}_{Mi}| \sim \epsilon a, \quad (1.5)$$

that is, the width of these orbits is smaller than the characteristic size of the stellarator, although it is much larger than the typical width of orbits in tokamaks, of order ρ_i .

For small collision frequencies (see (1.6) for a precise ordering for the collision frequency), the large width of the trapped-particle orbits has called into question the validity of local models for neoclassical transport. Here, ‘local’ refers to models that calculate neoclassical fluxes through a surface of interest using only the electric

field, the magnetic field and certain radial gradients of the magnetic field at that flux surface, whereas ‘global’ codes need the electric and magnetic field of the flux surfaces neighbouring the flux surface of interest. The most naive way to obtain a local model is to zero out the radial component of the drifts in certain terms of the drift kinetic equation (Sugama *et al.* 2016; Paul *et al.* 2017), but it has been noted that there are different ways in which this could be done, none of them necessarily consistent (Paul *et al.* 2017). Moreover, global neoclassical codes (Satake *et al.* 2006) have shown that neoclassical fluxes depend on parameters that do not appear in simplified drift kinetic models without radial drifts, such as the magnetic shear (Matsuoka *et al.* 2015; Huang *et al.* 2017). Calvo *et al.* (2017) used closeness to omnigenity to derive local orbit-averaged equations without having to artificially zero out the radial magnetic drifts. In this article, we use another expansion parameter, the inverse aspect ratio, to derive a different set of self-consistent local orbit-averaged equations that is valid for a wide class of stellarators: large aspect ratio stellarators with a mirror ratio close to unity. In our derivation, we do not start by assuming that the distribution function is Maxwellian or that the problem can be solved using a local equation, but we derive these properties from the expansion. The equations presented in this article coincide with the low collisionality limit of the equations in the DKES code (Hirshman *et al.* 1986) to lowest order in the small inverse aspect ratio expansion, but differ to higher order. The radial energy flux derived from the new equations in this paper, calculated using a modified version of the code KNOSOS (Velasco *et al.* 2020, 2021), has been shown to be close to the energy flux calculated by DKES in several experimentally relevant configurations (Velasco *et al.* 2021).

There is a subtlety in the derivation of the orbit-averaged equations for large aspect ratio stellarators with mirror ratios close to unity. Given the smallness of the orbit width in ϵ , it is tempting to neglect the radial drifts when calculating the lowest-order particle motion. However, in large aspect ratio stellarators with mirror ratios close to unity, the radial displacement of the particles is sufficiently large to affect the trapped-particle motion to lowest order. Indeed, trapped particles in this type of large aspect ratio stellarators have very small parallel velocities of order $\sqrt{\epsilon}v_{ii}$, and small changes in energy of order ϵT_i affect their trajectories, causing trapped particles to become passing and *vice versa*. Radial displacements of order ϵa are small compared with the size of the stellarator, but they lead to changes in energy of order ϵT_i due to the work done by the radial electric field. Our new equations keep the necessary finite orbit width effects by using the second adiabatic invariant as a velocity-space coordinate.

Our derivation of a local model is valid for collisionalities as small as

$$v_{i*} \sim \rho_{i*} \ll 1, \quad (1.6)$$

where

$$v_{i*} := \frac{Rv_{ii}}{v_{ii}} \quad (1.7)$$

is the collisionality,

$$v_{ii} := \frac{4\sqrt{\pi} Z_i^4 e^4 n_i \ln \Lambda}{3 m_i^{1/2} T_i^{3/2}} \quad (1.8)$$

is the ion–ion collision frequency (Braginskii 1958), n_i is the ion density and $\ln \Lambda$ is the Coulomb logarithm. We analyse the behaviour of our new equations for large aspect ratio stellarators with mirror ratios close to unity in the limit $v_{i*} \gg \rho_{i*}$, in which we recover the $1/\nu$ regime (Ho & Kulsrud 1987), and in the limit $v_{i*} \ll \rho_{i*}$. Surprisingly, for $v_{i*} \ll \rho_{i*}$, a rigorous expansion of our equations does not lead to the $\sqrt{\nu}$ regime (Galeev *et al.*

1969) for generic large aspect ratio stellarators with mirror ratio close to unity. Instead, the limit $v_{i*} \ll \rho_{i*}$ gives the ν regime (Mynick 1983). In this regime, particles follow their collisionless orbits for long times, moving away from their initial flux surface a distance of order ϵa , as explained above. Particles can only move to a flux surface further away than ϵa by having several collisions interrupt their orbits, thus leading to a radial flux that is proportional to the collision frequency. Importantly, trapped particles remain a distance of order ϵa away from their initial flux surface even when they undergo transitions between different types of wells and these transitions make their motion stochastic (Beidler, Hitchon & Shohet 1987). To treat these transitions between different types of wells, we do not need to introduce in the equations the transition probabilities calculated by Cary, Escande & Tennyson (1986).

There is a class of stellarators for which the $\sqrt{\nu}$ regime exists for $v_{i*} \ll \rho_{i*}$: stellarators close to omnigenity (Calvo *et al.* 2017). In stellarators far from omnigenity, the transitions between different types of wells of certain trapped particles smear out the $\sqrt{\nu}$ velocity-space boundary layer (Mynick 1983). We show that large aspect ratio stellarators with large mirror ratios are close to omnigenity and hence neoclassical transport in them can be calculated using the equations derived by Calvo *et al.* (2017). We also consider large aspect ratio stellarators with mirror ratios close to unity that are close to omnigenity, finding results that are consistent with our previous work on this area (Calvo *et al.* 2018).

Throughout the paper we focus on ion transport. In § 2 we remind the reader of the kinetic equations for a general stellarator in the limit $v_{i*} \sim \rho_{i*} \ll 1$. In § 3 we discuss the MagnetoHydroDynamic (MHD) equilibrium equations for $\epsilon \ll 1$, making a distinction between large aspect ratio stellarators with mirror ratios close to unity and large aspect ratio stellarators with large mirror ratios. Most of the rest of the paper is dedicated to large aspect ratio stellarators with mirror ratios close to unity, with the only exception being § 8.1. In § 4 we propose a new set of velocity-space coordinates that are necessary to simplify the expansion in $\epsilon \ll 1$ for large aspect ratio stellarators with mirror ratios close to unity, and in § 5 we finally perform the expansion in ϵ for the ion distribution function and the electric potential in this type of large aspect ratio stellarators. In §§ 6 and 7 we study the cases $v_{i*} \gg \rho_{i*}$ and $v_{i*} \ll \rho_{i*}$ for large aspect ratio stellarators with mirror ratios close to unity. In § 8 we consider large aspect ratio stellarators close to omnigenity. We divide our discussion of large aspect ratio stellarators close to omnigenity into two parts: in one we show that large aspect ratio stellarators with large mirror ratios are close to omnigenity, and in the other, we study optimized large aspect ratio stellarators with mirror ratios close to unity. We conclude in § 9.

2. Drift kinetic equation for ions in a generic stellarator

We assume $a \sim R$ in this section, but we keep the distinction between a and R in our estimates in preparation for the expansion in small inverse aspect ratio in §§ 3, 4, 5 and 8.

We assume that the magnetic field \mathbf{B} is constant in time and hence the electric field \mathbf{E} satisfies $\mathbf{E}(\mathbf{x}, t) = -\nabla\phi(\mathbf{x}, t)$. The electric potential ϕ is of order T_i/e , has a characteristic length scale of the order of the minor radius a , and is determined by the quasineutrality equation

$$Z_i \int f_i(\mathbf{x}, \mathbf{v}, t) d^3v = \int f_e(\mathbf{x}, \mathbf{v}, t) d^3v, \quad (2.1)$$

where $f_i(\mathbf{x}, \mathbf{v}, t)$ and $f_e(\mathbf{x}, \mathbf{v}, t)$ are the distribution functions of ions and electrons, \mathbf{x} and \mathbf{v} are the particle's Cartesian position and velocity, and t is time. Throughout the paper, we assume that the electrons can be modelled with the modified Maxwell–Boltzmann

response

$$\int f_e(\mathbf{x}, \mathbf{v}, t) d^3v = \hat{n}_e(r(\mathbf{x}), t) \exp\left(\frac{e\phi(\mathbf{x}, t)}{T_e(r(\mathbf{x}), t)}\right), \tag{2.2}$$

where $T_e(r, t)$ is the temperature of the electrons and $\hat{n}_e(r, t)$ is the density of electrons in the absence of ϕ . Note that \hat{n}_e and T_e are flux functions that only depend on the flux surface label $r(\mathbf{x})$. We define r more carefully below.

We use the drift kinetic equation (Hazeltine 1973) to obtain the ion distribution function. To describe velocity space, we choose the coordinates $\{\mathcal{E}, \mu, \sigma, \varphi\}$, where $\mathcal{E} := v^2/2 + Z_i e\phi/m_i$ is the total energy per unit mass, $\mu := v_\perp^2/2B$ is the magnetic moment, σ is the sign of the parallel velocity and φ is the gyrophase, defined such that

$$\frac{\mathbf{v}_\perp}{v_\perp} = \hat{\mathbf{e}}_\perp(\mathbf{x}, \varphi) := \cos \varphi \hat{\mathbf{e}}_1(\mathbf{x}) + \sin \varphi \hat{\mathbf{e}}_2(\mathbf{x}). \tag{2.3}$$

Here, \mathbf{v}_\perp is the component of \mathbf{v} perpendicular to the magnetic field, $v_\perp := |\mathbf{v}_\perp|$ and $\hat{\mathbf{e}}_1(\mathbf{x})$ and $\hat{\mathbf{e}}_2(\mathbf{x})$ are two unit vectors that form an orthonormal basis with the unit vector parallel to the magnetic field $\hat{\mathbf{b}}(\mathbf{x}) = \mathbf{B}/B$ and satisfy $\hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2 = \hat{\mathbf{b}}$. In the coordinates $\{\mathcal{E}, \mu, \sigma, \varphi\}$, the velocity-space volume element is

$$d^3v = \frac{d\mathcal{E} d\mu d\varphi}{|(\nabla_v \mathcal{E} \times \nabla_v \mu) \cdot \nabla_v \varphi|} = \frac{B}{|v_\parallel|} d\mathcal{E} d\mu d\varphi, \tag{2.4}$$

where

$$v_\parallel(\mathbf{x}, \mathcal{E}, \mu, \sigma) := \sigma \sqrt{2\left(\mathcal{E} - \mu B(\mathbf{x}) - \frac{Z_i e\phi(\mathbf{x})}{m_i}\right)} \tag{2.5}$$

is the parallel velocity.

The distribution function can be split into its gyroaverage, $\bar{f}_i := (2\pi)^{-1} \int_0^{2\pi} f_i d\varphi$, and the gyrophase-dependent piece $\tilde{f}_i := f_i - \bar{f}_i$. The gyrophase-dependent piece is of order $\rho_{i*} \tilde{f}_i$ and can be neglected in the limit $\rho_{i*} \sim v_{i*}$ because the fluxes of particles and energy depend to lowest order on the much larger gyroaveraged piece \bar{f}_i , unlike the fluxes in the neoclassical regimes of moderate collisionality ($v_{i*} \sim 1$) in which the piece of the gyroaveraged distribution function that matters for transport is small in ρ_{i*} (see the discussion at the end of this section for more details). Since the ion–electron collisions can be neglected within an expansion in the electron-to-ion mass ratio, the equation for \bar{f}_i is

$$\partial_t \bar{f}_i + \dot{\mathbf{x}} \cdot \nabla \bar{f}_i + \dot{\mathcal{E}} \partial_{\mathcal{E}} \bar{f}_i + \dot{\mu} \partial_\mu \bar{f}_i = C_{ii}[\bar{f}_i, \bar{f}_i](1 + O(\rho_{i*}^2)) + \bar{S}_i(1 + O(\rho_{i*})). \tag{2.6}$$

Here, S_i is a source term representing fuelling and heating, and \bar{S}_i is the gyroaverage of S_i . The particle motion can be split into parallel motion and perpendicular drifts,

$$\dot{\mathbf{x}} := \left(v_\parallel + \frac{m_i c \mu}{Z_i e} \hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}}\right) \hat{\mathbf{b}} + \mathbf{v}_{Mi} + \mathbf{v}_E + O(\rho_{i*}^2 v_{ii}), \tag{2.7}$$

where

$$\mathbf{v}_{Mi} := \frac{1}{\Omega_i} \hat{\mathbf{b}} \times (v_\parallel^2 \boldsymbol{\kappa} + \mu \nabla B) \tag{2.8}$$

are the curvature and ∇B drifts, collectively known as magnetic drift, $\kappa := \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}$ is the curvature of the magnetic field lines and

$$\mathbf{v}_E := \frac{c}{B} \hat{\mathbf{b}} \times \nabla \phi \tag{2.9}$$

is the $\mathbf{E} \times \mathbf{B}$ drift. The time derivative of the total energy is

$$\dot{\mathcal{E}} := \frac{Z_i e}{m_i} \partial_t \phi + O\left(\rho_{i*}^2 \frac{v_{ii}^3}{a}\right). \tag{2.10}$$

The time derivative of the magnetic moment is

$$\dot{\mu} := v_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left(\frac{v_{\parallel} \mu}{\Omega_i} \hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}} \right) + O\left(\rho_{i*}^2 \frac{v_{ii}^3}{aB}\right). \tag{2.11}$$

The ion–ion collision operator C_{ii} is a Fokker–Planck collision operator,

$$C_{ii}[f_a, f_b] := \gamma_{ii} \nabla_v \cdot (\nabla_v \nabla_v H[f_b] \cdot \nabla_v f_a - f_a \nabla_v L[f_b]), \tag{2.12}$$

where $\gamma_{ii} := 2\pi Z_i^4 e^4 \ln \Lambda / m_i^2$, and the Rosenbluth potentials H and L (Rosenbluth, MacDonald & Judd 1957) are the functionals

$$H[f](\mathbf{v}) := \int f(\mathbf{v}') |\mathbf{v} - \mathbf{v}'| d^3 v' \tag{2.13}$$

and

$$L[f](\mathbf{v}) := 2 \int \frac{f(\mathbf{v}')}{|\mathbf{v} - \mathbf{v}'|} d^3 v'. \tag{2.14}$$

Note that, in our notation, the first argument of C_{ii} refers to the distribution function that is evaluated at the velocity \mathbf{v} of interest, whereas the second argument refers to the distribution function that is integrated to obtain the Rosenbluth potentials. In the coordinates $\{\mathcal{E}, \mu, \sigma, \varphi\}$, the Fokker–Planck collision operator is given by

$$C_{ii}[f_a, f_b] = \gamma_{ii} |v_{\parallel}| \partial_{\mathcal{E}} \left[\frac{1}{|v_{\parallel}|} (H_{\mathcal{E}\mathcal{E}}[f_b] \partial_{\mathcal{E}} f_a + H_{\mathcal{E}\mu}[f_b] \partial_{\mu} f_a - L_{\mathcal{E}}[f_b] f_a) \right] + \gamma_{ii} |v_{\parallel}| \partial_{\mu} \left[\frac{1}{|v_{\parallel}|} (H_{\mu\mathcal{E}}[f_b] \partial_{\mathcal{E}} f_a + H_{\mu\mu}[f_b] \partial_{\mu} f_a - L_{\mu}[f_b] f_a) \right], \tag{2.15}$$

where we have used the fact that f_a does not depend on the gyrophase φ , and we have defined $H_{pq}[f] := \nabla_v p \cdot \nabla_v \nabla_v H[f] \cdot \nabla_v q$ and $L_p[f] := \nabla_v p \cdot \nabla_v L[f]$, with $p = \mathcal{E}, \mu$ and $q = \mathcal{E}, \mu$. Note that, in (2.7), (2.10), (2.11) and on the right-hand side of (2.6), we have indicated the size of terms associated with the gyrophase-dependent piece of the distribution function $\tilde{f}_i \sim \rho_{i*} \bar{f}_i$ that we have neglected – the errors in the collision operator C_{ii} are smaller than expected due to its gyrotopropy.

Instead of the Cartesian coordinates \mathbf{x} , it is convenient to use spatial coordinates that conform to the shape of the magnetic field. From here on, we use $\{r, \alpha, l\}$, where r is a flux surface label with units of length and of the order of the minor radius a , α is a poloidal angle that labels magnetic field lines on a given flux surface and l is the arc length of

the magnetic field line and is used to determine the position along the field line. In these coordinates, the magnetic field can be written as

$$\mathbf{B} = \Psi'_t(r) \nabla r \times \nabla \alpha, \tag{2.16}$$

where $\Psi_t(r)$ is the toroidal magnetic flux within the flux surface r divided by 2π , and $\Psi'_t := d\Psi_t/dr$. In these coordinates, the unit vector $\hat{\mathbf{b}}$ is given by

$$\hat{\mathbf{b}} = \partial_r \mathbf{x}, \tag{2.17}$$

and the element of volume is

$$d^3x = \frac{dr d\alpha dl}{|(\nabla r \times \nabla \alpha) \cdot \nabla l|} = \frac{\Psi'_t}{B} dr d\alpha dl. \tag{2.18}$$

The equations for general stellarators with $v_{i*} \sim \rho_{i*}$ were derived in § 3.1 of Calvo *et al.* (2017). Here, we generalize the work done in Calvo *et al.* (2017), and change the presentation in places to make the derivation of the large aspect ratio stellarator equations easier. We remind the reader that in this section we assume $\epsilon \sim 1$, but that we will perform a subsidiary expansion in $\epsilon \ll 1$ in the rest of the paper. We expand the drift kinetic system of equations in $\rho_{i*} \ll 1$ assuming $v_{i*} \sim \rho_{i*}$ and

$$\partial_t \sim \frac{S_i}{f_i} \sim v_{ii} \sim \rho_{i*} \frac{v_{ii}}{R}. \tag{2.19}$$

Our assumption for the size of the time derivative and the source S_i might be surprising, but it is justified by the fact that the final equation is consistent. Physically, these estimates are the result of the particle orbits being comparable to the size of the device. Thus, the time derivative and the source must compensate for both direct particle losses ($\partial_t \sim S_i/f_i \sim \rho_{i*} v_{ii}/R$) and, for particles in confined orbits, losses due to collisions ($\partial_t \sim S_i/f_i \sim v_{ii}$). With these assumptions, we can write \bar{f}_i as

$$\bar{f}_i = \bar{f}_i^{(0)} + \bar{f}_i^{(1)} + \dots, \tag{2.20}$$

with $\bar{f}_i^{(n)} \sim \rho_{i*}^n \bar{f}_i^{(0)}$. To lowest order in ρ_{i*} , (2.6) gives

$$v_{\parallel} \partial_{\parallel} \bar{f}_i^{(0)} = 0. \tag{2.21}$$

To solve this equation, we need to distinguish between passing and trapped particles. The function

$$U(r, \alpha, l, \mu, t) := \mu B(r, \alpha, l) + \frac{Z_i e \phi(r, \alpha, l, t)}{m_i} \tag{2.22}$$

is an effective potential for the motion parallel to the magnetic field line. If \mathcal{E} is larger than the maximum of U on a flux surface, $U_M(r, \mu, t)$, the parallel velocity in (2.5) never vanishes and the particle is a passing particle. If \mathcal{E} is smaller than $U_M(r, \mu, t)$, the parallel velocity vanishes at least at two bounce points, $l_{bL,W}(r, \alpha, \mathcal{E}, \mu, t)$ and $l_{bR,W}(r, \alpha, \mathcal{E}, \mu, t)$, defined by $\mathcal{E} - U(r, \alpha, l_{bL,W}, \mu, t) = 0 = \mathcal{E} - U(r, \alpha, l_{bR,W}, \mu, t)$ (the subscripts L and R refer to ‘left’ and ‘right’, respectively; see figure 1). Note that, for given values of \mathcal{E} and μ , a trapped particle can be located inside several different U wells. We will use the discrete index W to distinguish between these wells, where W takes Roman numeral values (see figure 1).

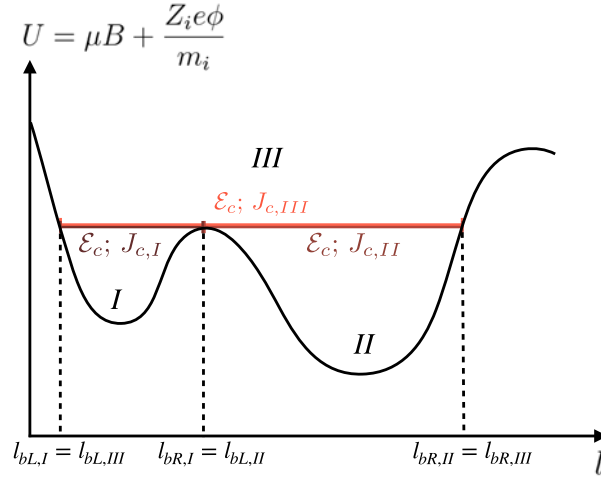


FIGURE 1. Sketch of the effective potential $U := \mu B + Z_i e \phi / m_i$ as a function of l .

On an ergodic flux surface where a single field line connects any two points, (2.21) implies that $\bar{f}_i^{(0)}$ must be independent of α for passing particles. For trapped particles, due to continuity at the bounce points $l_{bL,W}$ and $l_{bR,W}$ and (2.21), $\bar{f}_i^{(0)}$ cannot depend on σ . Using these conditions, we write $\bar{f}_i^{(0)}$ as

$$\bar{f}_i^{(0)} = \begin{cases} g_{i,W}(r, \alpha, \mathcal{E}, \mu, t), & \text{for } \mathcal{E} \leq U_M(r, \mu, t), \\ h_i(r, \mathcal{E}, \mu, \sigma, t), & \text{for } \mathcal{E} > U_M(r, \mu, t), \end{cases} \quad (2.23)$$

where $g_{i,W}$ is defined only in the trapped-particle region, $\mathcal{E} \leq U_M(r, \mu, t)$, and h_i is defined only in the passing-particle region, $\mathcal{E} > U_M(r, \mu, t)$. In most of this article, we will consider stellarators in which the effective potential U only reaches the maximum value $U_M(r, \mu, t)$ at a finite number of points on the flux surface r (the exception is § 8, where we discuss a case with contours $U = U_M$ that are lines that wrap around the flux surface: the omnigenous stellarator). In an ergodic flux surface where the value U_M is only reached at a finite number of points, there is one barely trapped particle orbit with $\mathcal{E} = U_M$ that covers the entire flux surface between its two bounce points (except for possibly a subset of points that has no area, i.e. a segment of magnetic field line of finite length might connect two of the maxima of U , but not all of them). If a surface-covering barely-trapped-particle orbit did not exist, we would be able to join all the points with $U = U_M$ with a single magnetic field line that closes on itself, contradicting the initial assumption that the surface is ergodic. We denote the well index of this surface-covering barely trapped particle as W_{bt} . For $W = W_{bt}$ and $\mathcal{E} = U_M$, $g_{i,W}$ does not depend on α , and we can impose the boundary conditions

$$g_{i,W_{bt}}(r, \alpha, U_M(r, \mu, t), \mu, t) = h_i(r, U_M(r, \mu, t), \mu, \sigma, t) \quad (2.24)$$

and

$$\partial_{\mathcal{E}} g_{i,W_{bt}}(r, \alpha, U_M(r, \mu, t), \mu, t) = \partial_{\mathcal{E}} h_i(r, U_M(r, \mu, t), \mu, \sigma, t). \quad (2.25)$$

These conditions imply that h_i cannot depend on σ at $\mathcal{E} = U_M$.

To next order in ρ_{i*} , (2.6) gives

$$v_{\parallel} \partial_l \left(\bar{f}_i^{(1)} + \frac{v_{\parallel} \mu}{\Omega_i} \hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}} \partial_{\mu} \bar{f}_i^{(0)} \right) + \partial_l \bar{f}_i^{(0)} + \frac{Z_i e}{m_i} \partial_t \phi \partial_{\mathcal{E}} \bar{f}_i^{(0)} + (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot (\nabla \alpha \partial_{\alpha} \bar{f}_i^{(0)} + \nabla r \partial_r \bar{f}_i^{(0)}) = C_{ii}[\bar{f}_i^{(0)}, \bar{f}_i^{(0)}] + \bar{S}_i. \tag{2.26}$$

We proceed to eliminate $\bar{f}_i^{(1)}$ from the equation. For trapped particles, we divide equation (2.26) by $|v_{\parallel}|$, sum over the two possible values of σ and integrate over l between bounce points to obtain

$$\partial_t g_{i,W} + \frac{Z_i e}{m_i} \langle \partial_t \phi \rangle_{\tau,W} \partial_{\mathcal{E}} g_{i,W} + \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla \alpha \rangle_{\tau,W} \partial_{\alpha} g_{i,W} + \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} \partial_r g_{i,W} = \langle C_{ii}[g_{i,W}, \bar{f}_i^{(0)}] \rangle_{\tau,W} + \langle \bar{S}_i \rangle_{\tau,W}, \tag{2.27}$$

where we have used the transit average

$$\langle \dots \rangle_{\tau,W} := \frac{1}{\tau_W} \sum_{\sigma} \int_{l_{bL,W}}^{l_{bR,W}} \frac{(\dots)}{|v_{\parallel}|} dl, \tag{2.28}$$

and

$$\tau_W := 2 \int_{l_{bL,W}}^{l_{bR,W}} \frac{dl}{|v_{\parallel}|}, \tag{2.29}$$

is the period of a trapped-particle orbit. For passing particles, we divide equation (2.26) by $|v_{\parallel}|$ and we integrate over l and α to find

$$\left\langle \frac{B}{|v_{\parallel}|} \right\rangle_{fs} \partial_t h_i + \frac{Z_i e}{m_i} \left\langle \frac{B}{|v_{\parallel}|} \partial_t \phi \right\rangle_{fs} \partial_{\mathcal{E}} h_i + \left\langle \frac{B}{|v_{\parallel}|} C_{ii}[h_i, \bar{f}_i^{(0)}] \right\rangle_{fs} = \left\langle \frac{B}{|v_{\parallel}|} \bar{S}_i \right\rangle_{fs}, \tag{2.30}$$

where we have defined the flux surface average

$$\langle \dots \rangle_{fs} := \frac{1}{V'} \int_0^{2\pi} d\alpha \int_0^{L(r,\alpha)} dl \frac{\Psi'_t}{B} (\dots). \tag{2.31}$$

Here, $L(r, \alpha)$ is the length along the magnetic field line between the two points where the magnetic field line crosses the curve defined by $l = 0$, and

$$V'(r) := \int_0^{2\pi} d\alpha \int_0^{L(r,\alpha)} dl \frac{\Psi'_t}{B} \tag{2.32}$$

is the derivative with respect to r of the volume $V(r)$ contained within the flux surface r . To obtain (2.30), we have written the radial component of the drifts as $(\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r = (v_{\parallel}/\Omega_i) \nabla \cdot (v_{\parallel} \hat{\mathbf{b}} \times \nabla r)$ to find

$$\left\langle \frac{B}{|v_{\parallel}|} (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \right\rangle_{fs} = 0. \tag{2.33}$$

Equation (2.27) cannot be used at values of \mathcal{E} that are junctures of three or more types of wells. We show an example of such a value of \mathcal{E} in figure 1. At these junctures, particles can and, in most cases, will transition from one type of well to another. In general, the

value of \mathcal{E} at which several types of well coincide, $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$, depends on r, α, μ and t . Around these values of \mathcal{E} , there are boundary layers, thin in \mathcal{E} , where the dependence of \tilde{f}_i on l cannot be neglected (see, for example, Nemov *et al.* 1999; Calvo *et al.* 2014). These boundary layers impose continuity in $g_{i,W}$ across these junctures. The derivatives of $g_{i,W}$ with respect \mathcal{E} and μ are not necessarily continuous. The derivatives of $g_{i,W}$ with respect to \mathcal{E} and μ on the different wells of a juncture are related to each other by two conditions: the combination $\partial_\mu \mathcal{E}_c \partial_\mathcal{E} g_{i,W} + \partial_\mu g_{i,W}$ is continuous across a juncture, and the collisional flux in velocity space across a juncture must be conserved. The combination $\partial_\mu \mathcal{E}_c \partial_\mathcal{E} g_{i,W} + \partial_\mu g_{i,W}$ is continuous at $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$ because $g_{i,W}$ is continuous at $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$. In the example of figure 1, continuity of $g_{i,W}$ at $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$ imposes

$$\begin{aligned} g_{i,I}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t) &= g_{i,II}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t) \\ &= g_{i,III}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t). \end{aligned} \tag{2.34}$$

for all μ . Differentiating this expression with respect to μ , we find

$$\begin{aligned} \partial_\mu \mathcal{E}_c(r, \alpha, \mu, t) \partial_\mathcal{E} g_{i,I}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t) &+ \partial_\mu g_{i,I}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t) \\ &= \partial_\mu \mathcal{E}_c(r, \alpha, \mu, t) \partial_\mathcal{E} g_{i,II}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t) + \partial_\mu g_{i,II}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t) \\ &= \partial_\mu \mathcal{E}_c(r, \alpha, \mu, t) \partial_\mathcal{E} g_{i,III}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t) + \partial_\mu g_{i,III}(r, \alpha, \mathcal{E}_c(r, \alpha, \mu, t), \mu, t), \end{aligned} \tag{2.35}$$

that is, the combination $\partial_\mu \mathcal{E}_c \partial_\mathcal{E} g_{i,W} + \partial_\mu g_{i,W}$ is continuous. The other condition for the derivatives of $g_{i,W}$ at $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$ is conservation of particle number in phase space. For example, for the case represented in figure 1, one needs to calculate the particles that are leaving wells I and II due to collisions, and then enforce that they enter well III . This velocity-space flux continuity condition is manipulated in Appendix A to give the following relation between the derivatives of $g_{i,W}$ with respect of \mathcal{E} on different sides of the juncture:

$$\begin{aligned} \tau_I [\langle H_{\mathcal{E}\mathcal{E}} \tilde{f}_i^{(0)} \rangle_{\tau,I} - 2 \langle H_{\mathcal{E}\mu} \tilde{f}_i^{(0)} \rangle_{\tau,I} \partial_\mu \mathcal{E}_c + \langle H_{\mu\mu} \tilde{f}_i^{(0)} \rangle_{\tau,I} (\partial_\mu \mathcal{E}_c)^2] \partial_\mathcal{E} g_{i,I} \\ + \tau_{II} [\langle H_{\mathcal{E}\mathcal{E}} \tilde{f}_i^{(0)} \rangle_{\tau,II} - 2 \langle H_{\mathcal{E}\mu} \tilde{f}_i^{(0)} \rangle_{\tau,II} \partial_\mu \mathcal{E}_c + \langle H_{\mu\mu} \tilde{f}_i^{(0)} \rangle_{\tau,II} (\partial_\mu \mathcal{E}_c)^2] \partial_\mathcal{E} g_{i,II} \\ = \tau_{III} [\langle H_{\mathcal{E}\mathcal{E}} \tilde{f}_i^{(0)} \rangle_{\tau,III} - 2 \langle H_{\mathcal{E}\mu} \tilde{f}_i^{(0)} \rangle_{\tau,III} \partial_\mu \mathcal{E}_c + \langle H_{\mu\mu} \tilde{f}_i^{(0)} \rangle_{\tau,III} (\partial_\mu \mathcal{E}_c)^2] \partial_\mathcal{E} g_{i,III}. \end{aligned} \tag{2.36}$$

The relation between the derivatives $\partial_\mu g_{i,W}$ on each side of the juncture can be obtained from (2.36) by using the fact that the combination $\partial_\mu \mathcal{E}_c \partial_\mathcal{E} g_{i,W} + \partial_\mu g_{i,W}$ is continuous.

Equations (2.1), (2.2), (2.27), (2.30) and (2.36) are the same as (31), (33) and (37) of Calvo *et al.* (2017) but for the inclusion of sources and time derivatives, and a different treatment of the split of $\tilde{f}_i^{(0)}$ between trapped and passing particles. These equations are radially non-local and lead to very large transport and to a non-Maxwellian distribution function. In Calvo *et al.* (2017), closeness to omnigenity was employed to derive radially local equations for a near-Maxwellian distribution function, but here we will use an expansion in the small inverse aspect ratio ϵ .

Equations (2.1), (2.2), (2.27), (2.30) and (2.36) are noticeably different from the usual neoclassical equations (Hinton & Hazeltine 1976), derived assuming $v_{i*} \sim 1$. In the limit $v_{i*} \sim 1$, to lowest order in ρ_{i*} , the gyroaveraged ion distribution is a stationary Maxwellian with density n_i and temperature T_i that only depend on the flux label $r, \tilde{f}_i^{(0)} = f_{Mi}$, and the

electrostatic potential is a flux function to lowest order in ρ_* , $\phi(r, \alpha, l, t) = \phi^{(0)}(r, t) + \phi^{(1)}(r, \alpha, l, t) + \dots$. The next-order corrections in $\rho_{i*}, \bar{f}_i^{(1)}$ and $\phi^{(1)}$, are determined by

$$v_{\parallel} \partial \bar{f}_i^{(1)} - C_{ii}^{\ell}[\bar{f}_i^{(1)}] = -\mathbf{v}_{Mi} \cdot \nabla r \partial_r f_{Mi} \tag{2.37}$$

and the quasineutrality equation, respectively. Here, C_{ii}^{ℓ} is the linearized collision operator, discussed further in § 5.2. The density and temperature in the Maxwellian f_{Mi} are calculated using particle and energy conservation equations. The neoclassical fluxes in these conservation equations are integrals of $\bar{f}_i^{(1)} \mathbf{v}_{Mi} \cdot \nabla r$, and hence scale as ρ_{i*}^2 . These neoclassical fluxes give a typical time scale for changes in density and temperature of $\partial_t \sim \rho_{i*}^2 v_{ii}$. For comparison, in a generic stellarator with $v_{i*} \sim \rho_{i*}$, (2.1), (2.2), (2.27), (2.30) and (2.36) show that the distribution function need not be close to a Maxwellian, and the typical time scale for transport is $\partial_t \sim v_{ii} \sim \rho_{i*} v_{ii}/R$. The difference between the orderings $v_{i*} \sim 1$ and $v_{i*} \sim \rho_{i*}$ is due to the typical radial separation between the position of a particle and the flux surface r that the particle started at. For $v_{i*} \sim 1$, particles collide often, and hence the radial drift does not have time to act and move the particle away from its initial radial position more than a distance of the order of the ion gyroradius ρ_i between collisions. For smaller collision frequencies ($\rho_{i*} \ll v_{i*} \ll 1$), particles in a stellarator drift out distances of order ρ_i/v_{i*} (Ho & Kulsrud 1987), giving higher and higher transport as the collision frequency decreases until eventually, for $v_{i*} \sim \rho_{i*}$, the separation between the initial radial position of the particle and its typical position becomes of the order of the minor radius of the device, a . In this regime, orbits are as large as the device, and transport occurs by either direct losses, giving the typical time scale $\partial_t \sim \rho_{i*} v_{ii}/R$, or, for particles in confined orbits, by collisions, giving $\partial_t \sim v_{ii}$. By expanding in closeness to omnigenity (Calvo *et al.* 2017) or in the small inverse aspect ratio (this article), one can recover that the distribution function is close to a Maxwellian, and that the radial flux of particles and energy is determined by a higher-order correction to that Maxwellian. The equations for these higher corrections in general look similar to (2.37), but the parallel streaming term is replaced by the drift in the α -direction. The correction to the Maxwellian in this case does not scale with ρ_{i*} as the expansion parameter is not ρ_{i*} but closeness to omnigenity or the inverse aspect ratio. One can devise equations for the correction to the Maxwellian that recover both orderings $v_{i*} \sim 1$ and $v_{i*} \sim \rho_{i*}$ by including both parallel streaming and drifts in the α -direction – we show in Appendix G that the equations in DKES are an example of this, recovering both the $v_{i*} \sim 1$ and $v_{i*} \sim \rho_{i*}$ limits for large aspect ratio stellarators.

3. MHD equilibria in large aspect ratio stellarators

In the coordinates $\{r, \alpha, l\}$, a large aspect ratio stellarator shape is

$$\mathbf{x}(r, \alpha, l) = \mathbf{x}_0(l) + \mathbf{x}_1(r, \alpha, l) + \mathbf{x}_2(r, \alpha, l) + \dots, \tag{3.1}$$

where $\mathbf{x}_0(l) \sim R$ is the magnetic axis, and $\mathbf{x}_n(r, \alpha, l) \sim \epsilon^n R$. We assume that

$$\partial_r \sim \frac{1}{a}, \quad \partial_{\alpha} \sim 1, \quad \partial_l \sim \frac{1}{R}. \tag{3.2}$$

Note that the expansion in (3.1) is not the Garren & Boozer (1991) polynomial expansion because we are not assuming that $\mathbf{x}_n(r, \alpha, l)$ is proportional to r^n . The Garren & Boozer (1991) expansion is a particular case of the expansion used here.

The values that $\mathbf{x}(r, \alpha, l)$ can take are constrained by the definition of arc length,

$$|\partial_l \mathbf{x}|^2 = 1, \tag{3.3}$$

and by the MHD force balance equation,

$$\nabla_{\perp} \left(P + \frac{B^2}{8\pi} \right) = \frac{B^2}{4\pi} \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}, \tag{3.4}$$

where $\nabla_{\perp} := \nabla - \hat{\mathbf{b}} \hat{\mathbf{b}} \cdot \nabla$ is the projection of the gradient in the plane perpendicular to the magnetic field, and $P(r)$ is the total plasma pressure, which is a flux function. We project equation (3.4) on $\partial_r \mathbf{x}$ and $\partial_{\alpha} \mathbf{x}$ to obtain

$$P' + \partial_r \left(\frac{B^2}{8\pi} \right) - \partial_l \left(\frac{B^2}{8\pi} \right) \partial_l \mathbf{x} \cdot \partial_r \mathbf{x} = \frac{B^2}{4\pi} \partial_l^2 \mathbf{x} \cdot \partial_r \mathbf{x} \tag{3.5}$$

and

$$\partial_{\alpha} \left(\frac{B^2}{8\pi} \right) - \partial_l \left(\frac{B^2}{8\pi} \right) \partial_l \mathbf{x} \cdot \partial_{\alpha} \mathbf{x} = \frac{B^2}{4\pi} \partial_l^2 \mathbf{x} \cdot \partial_{\alpha} \mathbf{x}, \tag{3.6}$$

where $P' := dP/dr$. To solve these equations, we need to obtain the magnitude of the magnetic field B from $\mathbf{x}(r, \alpha, l)$. Using (2.18), we find that the magnitude of the magnetic field is given by

$$B = \frac{\Psi'_l}{(\partial_r \mathbf{x} \times \partial_{\alpha} \mathbf{x}) \cdot \partial_l \mathbf{x}}. \tag{3.7}$$

We expand the MHD equilibrium equations in $\epsilon \ll 1$ by assuming that the plasma pressure is sufficiently small to satisfy

$$\beta := \frac{8\pi P}{B^2} \lesssim \epsilon \ll 1. \tag{3.8}$$

We are particularly interested in the magnitude of the magnetic field, given by

$$B(r, \alpha, l) = B_0(r, \alpha, l) + B_1(r, \alpha, l) + \dots, \tag{3.9}$$

where $B_n \sim \epsilon^n B_0 \ll 1$. To lowest order in ϵ , (3.5) and (3.6) become $\partial_r (B_0^2/8\pi) = 0$ and $\partial_{\alpha} (B_0^2/8\pi) = 0$. The solution to these equations is that $B_0(l)$ can only be a function of l . As a result, the lowest-order version of (3.7),

$$(\partial_r \mathbf{x}_1 \times \partial_{\alpha} \mathbf{x}_1) \cdot \hat{\mathbf{b}}_0 = \frac{\Psi'_l(r)}{B_0(l)} \tag{3.10}$$

cannot depend on α . Here, $\hat{\mathbf{b}}_0(l) := d\mathbf{x}_0/dl$ is the unit vector parallel to the magnetic axis. Note that condition (3.10) implies that

$$\Psi'_l \sim B_0 a. \tag{3.11}$$

Condition (3.10) limits the choice of $\mathbf{x}_1(r, \alpha, l)$. The function $\mathbf{x}_1(r, \alpha, l)$ must satisfy two constraints in addition to satisfying (3.10): the first-order correction to (3.3) and the conservation of electric current. These two extra constraints will not be needed for rest of the article, but we give them in Appendix B for completeness. For the rest of this paper, we only need to know that the three scalar constraints discussed above can be satisfied by

choosing the three components of $\mathbf{x}_1(r, \alpha, l)$ wisely, i.e. the large aspect ratio expansion is self-consistent.

The first-order correction B_1 can be calculated using MHD force balance. Keeping only the first order terms in ϵ in (3.5) and (3.6), we find

$$P' + \frac{B_0}{4\pi} \partial_r B_1 - \frac{B_0}{4\pi} \frac{dB_0}{dl} \hat{\mathbf{b}}_0 \cdot \partial_r \mathbf{x}_1 = \frac{B_0^2}{4\pi} \frac{d^2 \mathbf{x}_0}{dl^2} \cdot \partial_r \mathbf{x}_1 \tag{3.12}$$

and

$$\frac{B_0}{4\pi} \partial_\alpha B_1 - \frac{B_0}{4\pi} \frac{dB_0}{dl} \hat{\mathbf{b}}_0 \cdot \partial_\alpha \mathbf{x}_1 = \frac{B_0^2}{4\pi} \frac{d^2 \mathbf{x}_0}{dl^2} \cdot \partial_\alpha \mathbf{x}_1. \tag{3.13}$$

Equations (3.12) and (3.13) can be integrated to find

$$B_1(r, \alpha, l) = B_0(l) \kappa_0(l) \cdot \mathbf{x}_1(r, \alpha, l) + \frac{dB_0(l)}{dl} \hat{\mathbf{b}}_0(l) \cdot \mathbf{x}_1(r, \alpha, l) - \frac{4\pi P(r)}{B_0(l)}, \tag{3.14}$$

where $\kappa_0(l) := d^2 \mathbf{x}_0 / dl^2$ is the curvature of the magnetic axis.

For a given \mathbf{x}_1 , we can calculate the components of the drifts that we need to solve (2.27). In a general stellarator, the radial and α components of the magnetic drift are

$$\mathbf{v}_{Mi} \cdot \nabla r = -\frac{m_i c}{Z_i e \Psi'_t} (v_{\parallel}^2 \partial_\alpha \mathbf{x} \cdot \partial_{ll}^2 \mathbf{x} + \mu \partial_\alpha B - \mu \partial_l B \partial_\alpha \mathbf{x} \cdot \partial_l \mathbf{x}) \tag{3.15}$$

and

$$\mathbf{v}_{Mi} \cdot \nabla \alpha = \frac{m_i c}{Z_i e \Psi'_t} (v_{\parallel}^2 \partial_r \mathbf{x} \cdot \partial_{ll}^2 \mathbf{x} + \mu \partial_r B - \mu \partial_l B \partial_r \mathbf{x} \cdot \partial_l \mathbf{x}), \tag{3.16}$$

and the same components of the $\mathbf{E} \times \mathbf{B}$ drift are

$$\mathbf{v}_E \cdot \nabla r = -\frac{c}{\Psi'_t} (\partial_\alpha \phi - \partial_l \phi \partial_\alpha \mathbf{x} \cdot \partial_l \mathbf{x}) \tag{3.17}$$

and

$$\mathbf{v}_E \cdot \nabla \alpha = \frac{c}{\Psi'_t} (\partial_r \phi - \partial_l \phi \partial_r \mathbf{x} \cdot \partial_l \mathbf{x}). \tag{3.18}$$

To lowest order in $\epsilon \ll 1$, the expressions for the magnetic drift become

$$\mathbf{v}_{Mi} \cdot \nabla r = -\frac{m_i c (v_{\parallel}^2 + \mu B_0)}{Z_i e B_0 \Psi'_t} \partial_\alpha \left(B_1 - \frac{dB_0}{dl} \hat{\mathbf{b}}_0 \cdot \mathbf{x}_1 \right) + O(\epsilon^2 \rho_{i*} v_{ii}) \sim \epsilon \rho_{i*} v_{ii} \tag{3.19}$$

and

$$\begin{aligned} \mathbf{v}_{Mi} \cdot \nabla \alpha &= \frac{m_i c (v_{\parallel}^2 + \mu B_0)}{Z_i e B_0 \Psi'_t} \partial_r \left(B_1 - \frac{dB_0}{dl} \hat{\mathbf{b}}_0 \cdot \mathbf{x}_1 \right) + \frac{4\pi m_i c P' v_{\parallel}^2}{Z_i e B_0^2 \Psi'_t} \\ &+ O\left(\epsilon^2 \rho_{i*} \frac{v_{ii}}{a}\right) \sim \epsilon \rho_{i*} \frac{v_{ii}}{a}, \end{aligned} \tag{3.20}$$

where we have used (3.14) to write the magnetic drift components as derivatives of B_1 . Similarly, the radial and α components of the $\mathbf{E} \times \mathbf{B}$ drift are

$$\mathbf{v}_E \cdot \nabla r = -\frac{c}{\Psi'_t} \partial_\alpha \phi + O(a |\partial_l \ln \phi| \rho_{i*} v_{ii}) \sim \rho_{i*} v_{ii} \tag{3.21}$$

and

$$\mathbf{v}_E \cdot \nabla \alpha = \frac{c}{\Psi'_l} \partial_r \phi + O(|\partial_l \ln \phi| \rho_{i*} v_{ii}) \sim \rho_{i*} \frac{v_{ii}}{a}, \tag{3.22}$$

to lowest order in $\epsilon \ll 1$. For the $\mathbf{E} \times \mathbf{B}$ drift, we have emphasized that the size of the first-order corrections in ϵ is proportional to the derivative of ϕ with respect to l . The fact that the next-order corrections only depend on $\partial_l \phi$ is important because we show in § 4 that the potential is a flux function to lowest order, making these corrections even smaller than first order in ϵ .

For most of this article (§§ 4–7 and § 8.2), we will focus on large aspect ratio stellarators with constant B_0 , that is, $dB_0/dl = 0$. According to (3.10), stellarators with constant B_0 must have flux surfaces such that the area of a cut of a flux surface r through a plane perpendicular to the magnetic axis cannot depend on the position along the magnetic axis. Indeed, this area is given by

$$A(r, l) := \int_0^r dr' \int_0^{2\pi} d\alpha [\partial_r \mathbf{x}_1(r', \alpha, l) \times \partial_\alpha \mathbf{x}_1(r', \alpha, l)] \cdot \hat{\mathbf{b}}_0(l) = \frac{2\pi \Psi_l(r)}{B_0}. \tag{3.23}$$

From here on, we refer to these large aspect ratio stellarators as stellarators with mirror ratios close to unity because the ratio between the maximum and the minimum of B on a flux surface (mirror ratio) is $1 + O(\epsilon) \simeq 1$. We focus on large aspect ratio stellarators with mirror ratios close to unity for two reasons: (i) their description requires careful analysis and an unintuitive choice of velocity-space coordinates, and (ii) these stellarators with mirror ratio close to unity are extremely common – see, for example, the maps of magnetic field magnitude B in Beidler *et al.* (2011) that show mirror ratios in the interval 1.05–1.2.

To have a mirror ratio significantly different from unity in a large aspect ratio stellarator, B_0 must depend on l . In this case, the mirror ratio is $B_{0,M}/B_{0,m}$, where $B_{0,M}$ and $B_{0,m}$ are the maximum and minimum of $B_0(l)$, respectively. We are not aware of any large aspect ratio stellarators with mirror ratios significantly different from unity that have been built. Despite this fact, we will study large aspect ratio stellarators with mirror ratios significantly different from unity (from here on, ‘with large mirror ratios’ for short) in § 8.1, where we will consider $B_0(l)$ to be a general function of l . This type of stellarator is always close to omnigenous and hence one can use the formalism developed by Calvo *et al.* (2017) to calculate neoclassical transport in them.

4. New velocity-space coordinates for large aspect ratio stellarators with mirror ratios close to unity

We first consider the possibility of the potential $\phi(\mathbf{x}, t)$ being very different from a flux function, that is, $\partial_\alpha \phi \neq 0$ and $\partial_l \phi \neq 0$. We show that this is not possible in a large aspect ratio stellarator with mirror ratios close to unity, that is, large aspect ratio stellarators with constant B_0 . If ϕ is not a flux function, the variation of v_{\parallel} within a flux surface is dominated by the variation of ϕ ,

$$v_{\parallel} \simeq \sigma \sqrt{2 \left(\mathcal{E} - \mu B_0 - \frac{Z_i e \phi(r, \alpha, l, t)}{m_i} \right)}. \tag{4.1}$$

Thus, for a general ϕ , trapped particles satisfy $\mathcal{E} \leq U_M \simeq \mu B_0 + Z_i e \phi_M(r, t)/m_i$, where $\phi_M(r, t)$ is the maximum of ϕ on the flux surface. The parallel velocity of trapped particles

is of order v_{ii} , and the fraction of trapped particles is of order unity. In this case, the quasineutrality equation (2.1) becomes

$$\begin{aligned}
 &4\pi Z_i \int_0^\infty d\mu \int_{\mu B_0 + Z_i e\phi_M/m_i}^{\mu B_0 + Z_i e\phi_m/m_i} d\mathcal{E} \frac{B_0}{\sqrt{2(\mathcal{E} - \mu B_0 - Z_i e\phi/m_i)}} \sum_{W \in \mathcal{W}} g_{i,W}(r, \alpha, \mathcal{E}, \mu, t) \\
 &+ 2\pi Z_i \sum_\sigma \int_0^\infty d\mu \int_{\mu B_0 + Z_i e\phi_M/m_i}^{\mu B_0 + Z_i e\phi_m/m_i} d\mathcal{E} \frac{B_0 h_i(r, \mathcal{E}, \mu, \sigma, t)}{\sqrt{2(\mathcal{E} - \mu B_0 - Z_i e\phi/m_i)}} \\
 &= \hat{n}_e(r, t) \exp\left(\frac{e\phi}{T_e(r, t)}\right), \tag{4.2}
 \end{aligned}$$

to lowest order in ϵ . The function $g_{i,W}$ is defined for $\mathcal{E} \leq U_M$, but it is zero for the values of \mathcal{E} outside of well W (in the example of figure 1, $g_{i,I}$ and $g_{i,II}$ are zero for $\mathcal{E} > \mathcal{E}_c$, and $g_{i,III}$ is zero for $\mathcal{E} < \mathcal{E}_c$). Then, the sum of $g_{i,W}$ over the well index W gives a continuous function of \mathcal{E} . Note that the sum over W is performed over a subset \mathcal{W} of all possible wells. Set \mathcal{W} depends on the location where the quasineutrality is being evaluated because any well W has a limited range of values of l . In the example in figure 1, l in well I is between $l_{bL,I}$ and $l_{bR,I}$, and l in well II is between $l_{bL,II}$ and $l_{bR,II}$. When the ion density is evaluated for $l \in [l_{bL,I}, l_{bR,I}]$, set \mathcal{W} should include wells I and III , but exclude well II . Note that set \mathcal{W} is independent of l in a finite region of l around most points in the stellarator (e.g. in the case of figure 1, set \mathcal{W} includes wells I and III for $l \in [l_{bL,I}, l_{bR,I}]$). Thus, the quasineutrality equation (4.2) only depends on ϕ , r and α around most spatial points, giving a solution ϕ that can only depend on r and α , that is, $\partial_l \phi = 0$. As we are considering only ergodic flux surfaces, $\partial_l \phi = 0$ implies that $\partial_\alpha \phi = 0$, and hence ϕ is a flux function, $\phi = \phi(r)$. Note that this is an arbitrary flux function because we can choose it at will using the free function $\hat{n}_e(r, t)$.

Since the electric potential is a flux function to lowest order in ϵ , we write it as

$$\phi(r, \alpha, l, t) = \phi_0(r, t) + \phi_{3/2}(r, \alpha, l, t) + \dots, \tag{4.3}$$

where $\phi_0(r, t) \sim T_i/e$ is a flux function, and $\phi_{3/2}(r, \alpha, l, t) \sim \epsilon^{3/2} T_i/e$. We show that the correction to the lowest-order flux function is small in $\epsilon^{3/2}$ in § 5.2. Due to expansion (4.3), (4.1) is a bad approximation for trapped particles. Instead, we need to use

$$v_{\parallel} = \sigma \sqrt{2 \left(\mathcal{E}_1 - \mu B_1(r, \alpha, l) - \frac{Z_i e \phi_{3/2}(r, \alpha, l, t)}{m_i} \right) [1 + O(\epsilon)]}, \tag{4.4}$$

where the quantity $\mathcal{E}_1 := \mathcal{E} - \mu B_0 - Z_i e \phi_0(r, t)/m_i$ must be of order ϵv_{ii}^2 for trapped particles – otherwise, v_{\parallel} would not vanish. Thus, the characteristic size of the parallel velocity of trapped particles is $v_{\parallel} \sim \sqrt{\epsilon} v_{ii}$, and the fraction of trapped particles is of order $\sqrt{\epsilon}$.

Before expanding the ion distribution function in ϵ , we need new velocity-space coordinates for trapped and passing particles. We discuss the velocity-space coordinates for trapped and passing particles in §§ 4.1 and 4.2, respectively.

4.1. Velocity-space coordinates for trapped particles

Due to the smallness of v_{\parallel} and to the expansion in (4.3), the magnetic and $\mathbf{E} \times \mathbf{B}$ drifts in (3.19), (3.20), (3.21) and (3.22) simplify to

$$(\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r = -\frac{m_i c}{Z_i e \Psi'_t} \left(\mu \partial_\alpha B_1 + \frac{Z_i e}{m_i} \partial_\alpha \phi_{3/2} \right) + O(\epsilon^2 \rho_{i*} v_{ii}) \sim \epsilon \rho_{i*} v_{ii} \tag{4.5}$$

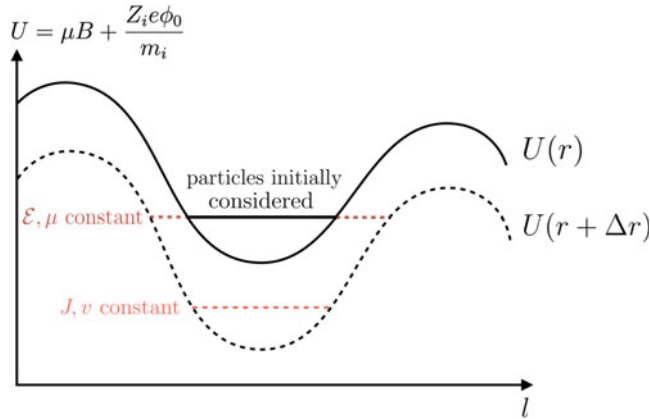


FIGURE 2. Trajectories of trapped particles in a given U well in the (\mathcal{E}, l) plane when one moves from a given flux surface to a neighbouring one keeping either J or \mathcal{E} constant. The shape of the U well does not change much whereas the whole well moves up and down due to the change in potential ϕ .

and

$$(\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla \alpha = \frac{c\phi'_0}{\Psi'_i} [1 + O(\epsilon)] \sim \rho_{i*} \frac{v_{ii}}{a}, \tag{4.6}$$

for trapped particles. Here $\phi'_0 := \partial_r \phi_0$.

Since the radial component of the drifts is small in ϵ , it is tempting to neglect the term proportional to the radial drift in (2.27). Unfortunately, this term cannot be neglected because the radial derivative of $g_{i,W}$ is very large,

$$\partial_r \ln g_{i,W}(r, \alpha, \mathcal{E}, \mu, t) \sim \frac{1}{\epsilon a}. \tag{4.7}$$

In order to understand estimate (4.7), one has to keep in mind that the potential changes significantly with radius. Indeed, the change in potential due to radial displacement $\Delta r \sim \epsilon a$ is $\Delta r \phi'_0 \sim \epsilon T_i/e$, and this change in potential energy means that v_{\parallel} has to change by $\sqrt{Z_i e \Delta r \phi'_0 / m_i} \sim \sqrt{\epsilon} v_{ii}$ if we keep \mathcal{E} constant when varying r . Hence, surprisingly, trapped particles with the same value of \mathcal{E} and separated only by $\Delta r \sim \epsilon a$ occupy very different heights with respect to the minimum of the U well. This situation is represented in figure 2. The difference in the height of the particle with respect to the minimum of the U well leads to the large radial derivative in (4.7). Note that the characteristic length of $g_{i,W}$, $\Delta r \sim \epsilon a$, is of the order of the width w of the particle orbits in (1.5), indicating that we need to keep the radial drifts in (2.27).

Importantly, the radial derivative of $g_{i,W}$ is very large because we are holding \mathcal{E} fixed and, as a result, the radial derivative of $g_{i,W}$ is related to the radial derivative of ϕ_0 . In contrast with the electric potential, the magnitude of the magnetic field does not change much if r is changed by $\Delta r \sim \epsilon a$ because its characteristic length of variation is R . Indeed, the change in B is $\Delta r \partial_r B \simeq \Delta r \partial_r B_1 \sim \epsilon^2 B_0$. The lack of rapid variation in B suggests using velocity-space coordinates that, when held constant, do not change the height of the trapped particle with respect to the minimum of the B well. The coordinates v and

$$\lambda := \frac{v_{\perp}^2}{v^2 B} \tag{4.8}$$

are often used because the variation of v along a orbit is small in ϵ and λ is equal to $1/B$ at the bounce points. To see that v does not change much during an orbit, note that in a large aspect ratio stellarator, the parallel velocity of a trapped particle is very small, giving $v \simeq v_{\perp} \simeq \sqrt{2\mu B_0}$, where B_0 is constant. Conversely, λ is not constant along trajectories. Writing λ as

$$\lambda = \frac{\mu}{\mathcal{E} - Z_i e \phi / m_i} = \frac{\mu}{\mathcal{E} - Z_i e \phi_0 / m_i} [1 + O(\epsilon^{3/2})] \tag{4.9}$$

shows that λ changes by an amount of order ϵ/B_0 along the orbit because of the variation of ϕ_0 . Changes of order ϵ/B_0 in λ are important for trapped particles because the interval of λ values where we find trapped particles, $\lambda \in [B_M^{-1}, B_m^{-1}]$, has a length of order ϵ/B_0 . Here $B_m(r)$ and $B_M(r)$ are the minimum and maximum of B on flux surface r , respectively.

Instead of the usual coordinates v and λ , we propose two other coordinates. In § 5, we will discover that v is not constant to a sufficiently high order in ϵ for one of the calculations that we perform: $v = \sqrt{2(\mathcal{E} - Z_i e \phi / m_i)}$ causes problems because it introduces dependence on l through $\phi_{3/2}$. Due to this limitation, we choose

$$\bar{v}(r, \mathcal{E}, t) := \sqrt{2 \left(\mathcal{E} - \frac{Z_i e \phi_0(r, t)}{m_i} \right)} \tag{4.10}$$

to be one of our velocity-space coordinates. Regarding λ , using it as a velocity-space coordinate is inconvenient because it is not constant in time. Fortunately, there is another quantity that gives the same information as λ and is constant in time: the second adiabatic invariant

$$\begin{aligned} J_W(r, \alpha, \mathcal{E}, \mu, t) &:= 2 \int_{l_{bL,W}}^{l_{bR,W}} |v_{\parallel}| dl \\ &= 2 \int_{l_{bL,W}}^{l_{bR,W}} \sqrt{2 \left(\mathcal{E} - \mu B(r, \alpha, l) - \frac{Z_i e \phi(r, \alpha, l, t)}{m_i} \right)} dl. \end{aligned} \tag{4.11}$$

We use the second adiabatic invariant as a velocity-space coordinate for trapped particles. Employing $\partial_{\mathcal{E}} J_W = \tau_W \sim \epsilon^{-1/2} R / v_{ti}$ and $\partial_{\mu} J_W = -\tau_W \langle B \rangle_{\tau,W} \sim \epsilon^{-1/2} B_0 R / v_{ti}$ to find

$$\nabla_v J_W = \partial_{\mathcal{E}} J_W \nabla_v \mathcal{E} + \partial_{\mu} J_W \nabla_v \mu = \tau_W v_{\parallel} \hat{\mathbf{b}} + \left(1 - \frac{\langle B \rangle_{\tau,W}}{B} \right) \tau_W \mathbf{v}_{\perp}, \tag{4.12}$$

we can calculate the velocity-space volume element,

$$d^3 v = \frac{d\bar{v} dJ d\varphi}{|(\nabla_v \bar{v} \times \nabla_v J_W) \cdot \nabla_v \varphi|} = \frac{\bar{v} B}{\tau_W |v_{\parallel}| \langle B \rangle_{\tau,W}} d\bar{v} dJ d\varphi = \frac{\bar{v}}{\tau_W |v_{\parallel}|} d\bar{v} dJ d\varphi [1 + O(\epsilon)]. \tag{4.13}$$

Note that we use the symbol J_W when the second adiabatic invariant is considered a function of $r, \alpha, \mathcal{E}, \mu, t$ and the type of well W , whereas we employ the symbol J without the subscript W when the second adiabatic invariant is a coordinate.

Previous work (Hazeltine & Catto 1981; Calvo *et al.* 2017) has used J as a coordinate, but as a replacement for the radial coordinate r instead of as a replacement for the pitch-angle variable λ . Note that the three quantities \mathcal{E}, μ and J are all desirable coordinates because they are constant along particle trajectories. One of the three variables works as a proxy

for the radial position of the particle, whereas the other two represent velocity space. In three-dimensional magnetic fields close to omnigenity (Hazeltine & Catto 1981; Calvo *et al.* 2017), J is a flux function to lowest order and hence it can be used to replace r . For trapped particles in large aspect ratio stellarators with mirror ratios close to unity, the kinetic energy associated with the parallel velocity is negligible compared with the perpendicular kinetic energy $\mu B \simeq \mu B_0$, giving $\mathcal{E} \simeq \mu B_0 + Z_i e \phi_0(r, t)/m_i$. As a result, the total energy determines the radial position r for a given μ through the potential energy $Z_i e \phi_0(r, t)/m_i$, and μ and J are the velocity-space coordinates – note that $\bar{v} \simeq \sqrt{2\mu B_0}$. We make the relation between the total energy and the radial position obvious in § 5.4, where we discuss the motion of the deeply trapped particles to demonstrate the advantages of our formulation.

We remind the reader that we are changing from the coordinates \mathcal{E} and μ to the coordinates \bar{v} and J to reduce the size of the derivative $\partial_r \ln g_{i,W}$ from $(\epsilon a)^{-1}$ to a^{-1} . We will not be able to show that the derivative of $g_{i,W}$ with respect to r holding \bar{v} and J constant is small in ϵ until § 5.2, but we can now show that if

$$\partial_r \ln g_{i,W}(r, \alpha, \bar{v}, J, t) \sim \frac{1}{a}, \tag{4.14}$$

the derivative of $g_{i,W}$ with respect to r holding \mathcal{E} and μ fixed is as large as estimated in (4.7). Indeed, by using the chain rule, we find

$$\begin{aligned} \partial_r \ln g_{i,W}(r, \alpha, \mathcal{E}, \mu, t) &= \partial_r \ln g_{i,W}(r, \alpha, \bar{v}, J, t) + \partial_r \bar{v}(r, \mathcal{E}, t) \partial_{\bar{v}} \ln g_{i,W}(r, \alpha, \bar{v}, J, t) \\ &\quad + \partial_r J_W(r, \alpha, \mathcal{E}, \mu, t) \partial_J \ln g_{i,W}(r, \alpha, \bar{v}, J, t). \end{aligned} \tag{4.15}$$

Noting that

$$\partial_r \bar{v} = -\frac{Z_i e \phi'_0}{m_i \bar{v}} \sim \frac{v_{ii}}{a}, \tag{4.16}$$

$$\partial_r J_W = -2 \int_{l_{bL,W}}^{l_{bR,W}} \frac{\mu \partial_r B + Z_i e \partial_r \phi / m_i}{|v_{\parallel}|} dl \simeq -\frac{Z_i e \phi'_0 \tau_W}{m_i} \sim \frac{v_{ii}}{\epsilon^{3/2}}, \tag{4.17}$$

$$\partial_{\bar{v}} \ln g_{i,W}(r, \alpha, \bar{v}, J, t) \sim \frac{1}{v_{ii}} \tag{4.18}$$

and

$$\partial_J \ln g_{i,W}(r, \alpha, \bar{v}, J, t) \sim \frac{1}{\sqrt{\epsilon} v_{ii} R}, \tag{4.19}$$

we obtain estimate (4.7).

4.2. Velocity-space coordinates for passing particles

For passing particles, the variable J cannot be defined. For most passing particles, $v \simeq \bar{v} := \sqrt{2(\mathcal{E} - Z_i e \phi_0/m_i)}$ and $v_{\parallel} \simeq \sigma \sqrt{2(\mathcal{E} - \mu B_0 - Z_i e \phi_0/m_i)}$ are approximate constants of the motion. For this reason, we use the coordinates \bar{v} and $\xi := v_{\parallel}/v$ for passing particles. The velocity-space volume element in these variables is

$$d^3 v = \frac{d\bar{v} d\xi d\varphi}{|(\nabla_{\bar{v}} \bar{v} \times \nabla_{\bar{v}} \xi) \cdot \nabla_{\bar{v}} \varphi|} = \bar{v} v d\bar{v} d\xi d\varphi = \bar{v}^2 d\bar{v} d\xi d\varphi [1 + O(\epsilon^{3/2})]. \tag{4.20}$$

For the few passing particles with $|\xi| \sim \sqrt{\epsilon} \ll 1$, ξ is not approximately constant, but we show in Appendix D that this region of phase space can be treated as a boundary

condition for the passing-particle distribution function h_i at $\xi = 0$. [Appendix D](#) should be read after having gone through §§ [5.1](#) and [5.2](#) and [Appendix C](#).

5. Ion distribution function and potential in low collisionality large aspect ratio stellarators with mirror ratios close to unity

We proceed to expand equations (2.27), (2.30) and (2.36) in the inverse aspect ratio $\epsilon \ll 1$ assuming that B_0 is independent of l and that $\rho_{i*} \sim \nu_{i*}$. Based on estimates (4.5) and (4.6) for the components of the $E \times B$ and magnetic drifts, we will show in § 5.3 that we need to order the time derivative and the source as

$$\partial_t \sim \frac{S_i}{f_i} \sim \epsilon^{3/2} \nu_{ii}. \tag{5.1}$$

This estimate is the result of a subsidiary expansion in $\epsilon \ll 1$ of the radially global equations for generic stellarators in § 2. Note that it is consistent with assumption (2.19) in § 2 when $\epsilon \sim 1$.

We expand $g_{i,W}$ and h_i in $\epsilon \ll 1$ assuming that $\rho_{i*} \sim \nu_{i*}$ and using the estimates in (5.1). For $g_{i,W}$, we find

$$g_{i,W} = g_{i,0,W} + g_{i,1,W} + g_{i,3/2,W} + g_{i,2,W} \dots, \tag{5.2}$$

where $g_{i,n,W} \sim \epsilon^n g_{i,0,W}$. For h_i , the expansion gives

$$h_i = h_{i,0} + h_{i,3/2} + \dots, \tag{5.3}$$

where $h_{i,n} \sim \epsilon^n h_{i,0}$. The corrections $g_{i,3/2,W}$, $g_{i,2,W}$ and $h_{i,3/2}$ are important because their size determines the boundary conditions for $g_{i,1,W}$, but in the end we do not need to calculate them.

We need to consider half-integer powers of ϵ in our expansion for two reasons: (i) the size of the trapped-particle region in velocity space is small by a factor of order $\sqrt{\epsilon}$, and (ii) boundary condition (2.25) introduces these half-integer powers naturally, as we will see shortly.

We organize the calculation as follows. In § 5.1 we argue that the lowest-order solution is a Maxwellian. In § 5.2 we obtain the equation and boundary conditions for $g_{i,1,W}$. Finally, in § 5.3 we obtain the transport equations for the ion density and temperature. In § 5.4, we summarize the equations, we compared them with those implemented in DKES (Hirshman *et al.* 1986) and we illustrate how they work by discussing the behaviour of deeply trapped particles.

5.1. *Lowest-order ion distribution function*

The proof that the distribution is a stationary Maxwellian to lowest order in ϵ does not follow the derivation used in standard neoclassical theory – for examples of the usual procedure, see § V.4 of Hinton & Hazeltine (1976) or § 3.1 of Calvo *et al.* (2017). In the typical derivation, the radial drifts are small and can be neglected in the lowest-order kinetic equation. By calculating a certain moment of this simplified lowest-order kinetic equation, one obtains an entropy equation for an infinitesimal volume between two flux surfaces close to each other. In this equation, the entropy production due to collisions within the infinitesimal volume cannot be compensated by any outward flow of entropy because the radial drifts are negligible. Consequently, in steady state, the average entropy production must vanish to lowest order, leading to a distribution function close to Maxwellian. In large aspect ratio stellarators with mirror ratios close to unity, we cannot follow this procedure because, for $\nu_{i*} \sim \rho_{i*}$, the typical frequency associated with the

radial drift, $\langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau, w} / a \sim \rho_{i*} v_{ii} / R$, is comparable to the collision frequency v_{ii} . As a result, trapped particles move a significant radial distance in the time that the distribution function evolves towards a Maxwellian. This motion can, in principle, disrupt the natural evolution of the distribution function towards a Maxwellian.

We proceed to argue that the distribution function is close to a Maxwellian in large aspect ratio stellarator with mirror ratios close to unity even though the trapped-particle radial drifts are large. Collisions can drive the distribution function close to a Maxwellian because only trapped particles drift off flux surfaces. The number of trapped particles is small by $\sqrt{\epsilon}$ and, for this reason, despite the considerable radial displacements of trapped particles, the radial flux of entropy due to trapped particles is not sufficient to compensate for the large entropy production associated with a distribution function far from Maxwellian. Moreover, trapped particles themselves are close to Maxwellian despite their large displacements because they exchange momentum and energy with passing particles much faster than the ion–ion collision frequency v_{ii} would suggest. Indeed, due to the small parallel velocity of trapped particles, $v_{\parallel} \sim \sqrt{\epsilon} v_{ii}$, grazing collisions can detract them, leading to an effective collision frequency $v_{ii} / \epsilon \gg \rho_{i*} v_{ii} / R$.

Appendix C contains the calculations that show that the lowest-order distribution function is Maxwellian. In this appendix, we first argue that collisions force the lowest-order trapped-particle distribution function $g_{i,0,w}(r, \alpha, \bar{v}, J, t)$ to be independent of α and J . Continuity at the trapped–passing-particle boundary then imposes that the trapped-particle distribution function be the passing-particle distribution function $h_{i,0}(r, \bar{v}, \xi, t)$ at small pitch angles $\xi \sim \sqrt{\epsilon} \ll 1$,

$$g_{i,0,w}(r, \alpha, \bar{v}, J, t) = h_{i,0}(r, \bar{v}, 0, t). \tag{5.4}$$

When we examine the barely-passing-particle region of velocity space, we find that, to the order of interest, the collisional flux from the trapped-particle region to the passing-particle region is negligible. Thus, passing particles collide with each other without exchanging significant momentum and energy with trapped particles and as a result their distribution function $h_{i,0}$ becomes Maxwellian. The trapped particles, however, have a role to play. The lack of collisional flux imposes that the derivative of $h_{i,0}(r, \bar{v}, \xi, t)$ with respect to ξ at small values of $\xi \sim \sqrt{\epsilon} \ll 1$ must be close to zero. Thus, the Maxwellian must have zero average flow, that is, the friction between trapped and passing particles damps the average flow to subsonic levels. Finally, noting that the passing-particle distribution function $h_{i,0}$ cannot depend on α or l , we find that $h_{i,0}$ is a Maxwellian with density n_i and temperature T_i that are flux functions,

$$h_{i,0}(r, \bar{v}, \xi, t) = f_{Mi}(r, \bar{v}, t) := n_i(r, t) \left(\frac{m_i}{2\pi T_i(r, t)} \right)^{3/2} \exp \left(-\frac{m_i \bar{v}^2}{2T_i(r, t)} \right). \tag{5.5}$$

The lowest-order trapped-particle distribution function $g_{i,0,w}$ is also a Maxwellian because of (5.4),

$$g_{i,0,w}(r, \alpha, \bar{v}, J, t) = f_{Mi}(r, \bar{v}, t). \tag{5.6}$$

5.2. Corrections to the lowest-order solutions for the ion distribution function and the electric potential

Since the lowest-order distribution function is a Maxwellian, from here on we need to use the linearized collision operator $C_{ii}^{\ell}[f] := C_{ii}[f, f_{Mi}] + C_{ii}[f_{Mi}, f]$. The linearized Fokker–Planck collision operator is composed of two terms: a differential part $C_{ii,D}^{\ell}$ and

an integral part $C_{ii,I}^\ell$,

$$C_{ii}^\ell[f] = C_{ii,D}^\ell[f] + C_{ii,I}^\ell[f]. \tag{5.7}$$

The differential part of the linearized collision operator is

$$C_{ii,D}^\ell[f] := \nabla_v \cdot \left[\frac{v_{ii,\perp} f_{Mi}}{4} (v^2 I - \mathbf{v}\mathbf{v}) \cdot \nabla_v \left(\frac{f}{f_{Mi}} \right) + \frac{v_{ii,\parallel} f_{Mi}}{2} \mathbf{v}\mathbf{v} \cdot \nabla_v \left(\frac{f}{f_{Mi}} \right) \right], \tag{5.8}$$

where I is the unit matrix,

$$v_{ii,\perp}(r, v, t) := \frac{4\gamma_{ii}}{v^3} \partial_v H[f_{Mi}] = \frac{3\sqrt{2\pi} v_{ii}}{2} \frac{v_{ii}^3}{v^3} [\text{erf}(v/v_{ii}) - \chi(v/v_{ii})], \tag{5.9}$$

is the pitch-angle scattering frequency and

$$v_{ii,\parallel}(r, v, t) := \frac{2\gamma_{ii}}{v^2} \partial_{vv}^2 H[f_{Mi}] = \frac{3\sqrt{2\pi} v_{ii}}{2} \frac{v_{ii}^3}{v^3} \chi(v/v_{ii}) \tag{5.10}$$

is the energy diffusion frequency. Here, the function $H[f_{Mi}](r, v, t)$ only depends on the velocity through its magnitude v , $\text{erf}(x) := (2/\sqrt{\pi}) \int_0^x \exp(-s^2) ds$ is the error function and

$$\chi(x) := \frac{1}{2x^2} \left[\text{erf}(x) - \frac{2x}{\sqrt{\pi}} \exp(-x^2) \right]. \tag{5.11}$$

The integral part of the collision operator is

$$C_{ii,I}^\ell[f] := \gamma_{ii} \nabla_v \cdot (\nabla_v \nabla_v H[f] \cdot \nabla_v f_{Mi} - f_{Mi} \nabla_v L[f]). \tag{5.12}$$

Using the linearized collision operator, we proceed to obtain an equation and boundary conditions for $g_{i,1,W}$, to discuss the equation for the passing-particle distribution function, and to show that the correction to the lowest-order potential $\phi_0(r)$ is of order $\epsilon^{3/2} T_i/e$.

5.2.1. Equation for $g_{i,1,W}$

We first rewrite (2.27) using the coordinates \bar{v} and J , and the fact that $g_{i,0,W} = f_{Mi}$ and $h_{i,0} = f_{Mi}$. We neglect the time derivatives and the source S_i because of the estimates in (5.1). We also use the fact that the derivatives with respect to J , $\partial_J \sim \epsilon^{-1/2}/v_{ii}R$, are larger by $\epsilon^{-1/2}$ than the derivatives with respect to \bar{v} , $\partial_{\bar{v}} \sim v_{ii}^{-1}$. This difference in size is particularly important for the linearized collision operator that can be approximated by

$$C_{ii,D}^\ell[g_{i,1,W}] \simeq \nabla_v J_W \cdot \partial_J \left[\left(\frac{v_{ii,\perp}}{4} (v^2 I - \mathbf{v}\mathbf{v}) + \frac{v_{ii,\parallel}}{2} \mathbf{v}\mathbf{v} \right) \cdot \nabla_v J_W \partial_J g_{i,1,W} \right]. \tag{5.13}$$

Finally, we take into account that, according to (4.5) and (4.6), the component of the drifts in the radial direction is much smaller than the component in the α direction. With all these considerations, to lowest order in ϵ , (2.27) becomes

$$\begin{aligned} & \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla \alpha \rangle_{\tau,W} (\partial_\alpha g_{i,1,W} + \partial_\alpha J_W \partial_J g_{i,1,W}) \\ & + \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} (\partial_r f_{Mi} + \partial_r \bar{v} \partial_{\bar{v}} f_{Mi} + \partial_r J_W \partial_J g_{i,1,W}) \\ & = \left\langle \nabla_v J_W \cdot \partial_J \left[\left(\frac{v_{ii,\perp}}{4} (v^2 I - \mathbf{v}\mathbf{v}) + \frac{v_{ii,\parallel}}{2} \mathbf{v}\mathbf{v} \right) \cdot \nabla_v J_W \partial_J g_{i,1,W} \right] \right\rangle_{\tau,W}, \end{aligned} \tag{5.14}$$

where the derivatives of $g_{i,W}$ with respect to r and α are performed holding \bar{v} and J fixed, whereas the derivatives of \bar{v} and J_W with respect to the same variables are performed

holding \mathcal{E} and μ constant. To simplify (5.14), we need to use the fact that J is an adiabatic invariant. Employing the exact expressions (3.15) and (3.17), we find

$$\langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} = \frac{m_i c}{Z_i e \Psi'_t \tau_W} \partial_\alpha J_W; \tag{5.15}$$

similarly, using (3.16) and (3.18), we obtain

$$\langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla \alpha \rangle_{\tau,W} = -\frac{m_i c}{Z_i e \Psi'_t \tau_W} \partial_r J_W. \tag{5.16}$$

Thus, the second adiabatic invariant satisfies

$$\langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla \alpha \rangle_{\tau,W} \partial_\alpha J_W + \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} \partial_r J_W = 0. \tag{5.17}$$

Employing the lowest-order expression (4.5), we find

$$\langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} \partial_r \bar{v} = \frac{c \phi'_0}{\bar{v} \Psi'_t} \left\langle \frac{\bar{v}^2}{2B_0} \partial_\alpha B_1 + \frac{Z_i e}{m_i} \partial_\alpha \phi_{3/2} \right\rangle_{\tau,W} [1 + O(\epsilon)] \sim \epsilon \rho_{i*} \frac{v_{Ti}^2}{a}. \tag{5.18}$$

With this result, (5.17) and employing the fact that $\nabla_v J_W \simeq \tau_W v_{\parallel} \hat{\mathbf{b}}$ for trapped particles (see (4.12)), (5.14) can be rewritten as

$$\frac{c \phi'_0}{\Psi'_t} \partial_\alpha g_{i,1,W} - \frac{\bar{v}^2 v_{i,\perp}(r, \bar{v}, t)}{4} \partial_J (\tau_W J \partial_J g_{i,1,W}) = \frac{m_i c \phi'_0 \bar{v}^2}{2 Z_i e B_0 \Psi'_t} \langle \partial_\alpha B_1 \rangle_{\tau,W} \mathcal{Y}_{f_{Mi}}, \tag{5.19}$$

where

$$\mathcal{Y}_i(r, \bar{v}, t) := \frac{n'_i}{n_i} + \frac{Ze \phi'_0}{T_i} + \left(\frac{m_i \bar{v}^2}{2T_i} - \frac{3}{2} \right) \frac{T'_i}{T_i}, \tag{5.20}$$

and the quantities n'_i and T'_i are $\partial_r n_i$ and $\partial_r T_i$, respectively. Importantly, note that r only appears as a parameter in (5.14), and hence, the derivative of $g_{i,1,W}$ with respect to r is determined by the variation of the coefficients in (5.14), giving $\partial_r \ln g_{i,1,W} \sim a^{-1}$, as announced in §4.1.

We can rewrite (5.19) in a more convenient form. Using the variable

$$\bar{\lambda}(r, \mathcal{E}, \mu, t) := \frac{\mu}{\mathcal{E} - Z_i e \phi_0(r, t)/m_i} = \lambda + O\left(\frac{\epsilon^{3/2}}{B_0}\right), \tag{5.21}$$

the coordinate J can be expressed as

$$J_W(r, \alpha, \bar{v}, \bar{\lambda}) \simeq 2\bar{v} \int_{l_{bL,0,W}}^{l_{bR,0,W}} \sqrt{1 - \bar{\lambda} B_0 - \frac{B_1(r, \alpha, l)}{B_0}} dl \tag{5.22}$$

to lowest order in ϵ . Here, $l_{bL,0,W}(r, \alpha, \bar{\lambda})$ and $l_{bR,0,W}(r, \alpha, \bar{\lambda})$ are the approximate bounce points, determined by the equations $B_1(r, \alpha, l_{bL,0,W})/B_0 = 1 - \bar{\lambda} B_0 = B_1(r, \alpha, l_{bR,0,W})/B_0$. We can invert (5.22) to obtain $\bar{\lambda}_W(r, \alpha, \bar{v}, J, t)$ as a function of r, α, \bar{v}, J and the well

index W . We can also calculate $\partial_\alpha \bar{\lambda}_W$ by differentiating equation (5.22) with respect to α holding r, \bar{v} and J constant,

$$0 \simeq - \int_{l_{bL,0,W}}^{l_{bR,0,W}} \frac{B_0^{-1} \partial_\alpha B_1}{\sqrt{1 - \bar{\lambda}_W B_0 - B_1/B_0}} dl - B_0 \partial_\alpha \bar{\lambda}_W \int_{l_{bL,0,W}}^{l_{bR,0,W}} \frac{1}{\sqrt{1 - \bar{\lambda}_W B_0 - B_1/B_0}} dl. \tag{5.23}$$

This expression gives

$$\langle \partial_\alpha B_1 \rangle_{\tau,W} \simeq -B_0^2 \partial_\alpha \bar{\lambda}_W. \tag{5.24}$$

We can use this result to rewrite (5.19) as

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,1,W} - \frac{\bar{v}^2 \nu_{ii,\perp}(r, \bar{v}, t)}{4} \partial_J (\tau_W J \partial_J g_{i,1,W}) = -\frac{c\phi'_0}{\Psi'_t} \partial_\alpha r_{i,1,W} \mathcal{R}ifMi, \tag{5.25}$$

where

$$r_{i,1,W}(r, \alpha, \bar{v}, J, t) := \frac{m_i B_0 \bar{v}^2}{2Z_i e \phi'_0(r, t)} [\bar{\lambda}_W(r, \alpha, \bar{v}, J, t) - \lim_{J \rightarrow \infty} \bar{\lambda}_{Wbt}(r, \alpha, \bar{v}, J, t)] \sim \epsilon \alpha \tag{5.26}$$

is the lowest-order radial displacement of the particle. We have defined $r_{i,1,W}$ such that it goes to zero at $J \rightarrow \infty$. This limit corresponds to the surface-filling barely trapped particle with $\mathcal{E} = U_M(r, \mu, t)$ and $W = W_{bt}$. Note that $\bar{\lambda}_W(r, \alpha, \bar{v}, J, t)$ does not depend on α for $J \rightarrow \infty$.

In addition to (5.25), we need the conditions to be imposed in junctures of several types of wells. With the new variables \bar{v} and J , these junctures of different types of wells happen at particular values of $J, J = J_{c,W}(r, \alpha, \bar{v}, t)$, which depend on r, α, \bar{v}, t and the type of well W . Note, for example, the juncture in figure 1: the value of J for the juncture is different for each type of well. These values of J are related by

$$J_{c,I}(r, \alpha, \bar{v}, t) + J_{c,II}(r, \alpha, \bar{v}, t) = J_{c,III}(r, \alpha, \bar{v}, t). \tag{5.27}$$

As we noted in § 2, the function $g_{i,W}$ is continuous across the juncture, but the derivatives $\partial_{\mathcal{E}} g_{i,W}$ and $\partial_\mu g_{i,W}$ are discontinuous. There are two conditions that we use to relate the discontinuous derivatives on different sides of the juncture. On the one hand, the combination $\partial_\mu \mathcal{E}_c \partial_{\mathcal{E}} g_{i,W} + \partial_\mu g_{i,W}$, which is the derivative along the boundary $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$, is continuous. On the other hand, the discontinuous derivatives $\partial_{\mathcal{E}} g_{i,W}$ around the juncture are related to each other by (2.36). In the new variables \bar{v} and J , the derivative along the boundary $J = J_{c,W}(r, \alpha, \bar{v}, t)$ is $\partial_{\bar{v}} g_{i,1,W} + \partial_J J_{c,W} \partial_J g_{i,1,W}$, and hence this combination of $\partial_{\bar{v}} g_{i,1,W}$ and $\partial_J g_{i,1,W}$ is continuous across the juncture. To finish our discussion of the junctures, we need to rewrite (2.36) in the new coordinates \bar{v} and J . We first note that expression (2.36) vanishes to lowest order in ϵ because $g_{i,0,W}$ and $h_{i,0}$ are Maxwellians. Thus, (2.36) becomes

$$\begin{aligned} \tau_I & \left[\langle H_{\mathcal{E}\mathcal{E}}[fMi] \rangle_{\tau,I} - 2 \langle H_{\mathcal{E}\mu}[fMi] \rangle_{\tau,I} \partial_\mu \mathcal{E}_c + \langle H_{\mu\mu}[fMi] \rangle_{\tau,I} (\partial_\mu \mathcal{E}_c)^2 \right] \partial_{\mathcal{E}} g_{i,1,I} \\ & + \tau_{II} \left[\langle H_{\mathcal{E}\mathcal{E}}[fMi] \rangle_{\tau,II} - 2 \langle H_{\mathcal{E}\mu}[fMi] \rangle_{\tau,II} \partial_\mu \mathcal{E}_c + \langle H_{\mu\mu}[fMi] \rangle_{\tau,II} (\partial_\mu \mathcal{E}_c)^2 \right] \partial_{\mathcal{E}} g_{i,1,II} \\ & = \tau_{III} \left[\langle H_{\mathcal{E}\mathcal{E}}[fMi] \rangle_{\tau,III} - 2 \langle H_{\mathcal{E}\mu}[fMi] \rangle_{\tau,III} \partial_\mu \mathcal{E}_c + \langle H_{\mu\mu}[fMi] \rangle_{\tau,III} (\partial_\mu \mathcal{E}_c)^2 \right] \partial_{\mathcal{E}} g_{i,1,III}. \end{aligned} \tag{5.28}$$

Note that the perturbations due to $H_{pq}[h_{i,3/2}]$ and $H_{pq}[g_{i,1,W}]$ can be neglected because they are small in ϵ – in the case of $H_{pq}[g_{i,1,W}]$ because the integral that gives $H_{pq}[g_{i,1,W}]$

is over a region of velocity space small by $\sqrt{\epsilon}$. Using the definitions (5.9) and (5.10), we find $H_{\mathcal{E}\mathcal{E}} = \nabla_v \mathcal{E} \cdot \nabla_v \nabla_v H[f_{Mi}] \cdot \nabla_v \mathcal{E} = v^2 \partial_{vv}^2 H[f_{Mi}] = v^4 v_{ii,\parallel} / 2\gamma_{ii}$, $H_{\mathcal{E}\mu} = \nabla_v \mathcal{E} \cdot \nabla_v \nabla_v H[f_{Mi}] \cdot \nabla_v \mu = (v_{\perp}^2 / B) \partial_{vv}^2 H[f_{Mi}] = v^2 v_{\perp}^2 v_{ii,\parallel} / 2\gamma_{ii} B$ and $H_{\mu\mu} = \nabla_v \mu \cdot \nabla_v \nabla_v H[f_{Mi}] \cdot \nabla_v \mu = (v_{\parallel}^2 v_{\perp}^2 / v^3 B^2) \partial_{vv}^2 H[f_{Mi}] + (v_{\perp}^4 / v^2 B^2) \partial_{vv}^2 H[f_{Mi}] = v_{\parallel}^2 v_{\perp}^2 v_{ii,\perp} / 4\gamma_{ii} B^2 + v_{\perp}^4 v_{ii,\parallel} / 2\gamma_{ii} B^2$. Moreover, $\mathcal{E}_c(r, \alpha, \mu, t)$ for the example in figure 1 is given by

$$\mathcal{E}_c(r, \alpha, \mu, t) = \mu B_{IM}(r, \alpha) + \frac{Z_i e \phi_0(r, t)}{m_i} + O(\epsilon^{3/2} v_{ii}^2) \tag{5.29}$$

to lowest order in ϵ . Here, $B_{IM}(r, \alpha)$ is the local maximum of B on magnetic field line (r, α) on which the juncture occurs. Then, $\partial_{\mu} \mathcal{E}_c = B_{IM} + O(\epsilon^{3/2} B_0)$. As a result of all these considerations, we find

$$\begin{aligned} H_{\mathcal{E}\mathcal{E}}[f_{Mi}] - 2H_{\mathcal{E}\mu}[f_{Mi}] \partial_{\mu} \mathcal{E}_c + H_{\mu\mu}[f_{Mi}] (\partial_{\mu} \mathcal{E}_c)^2 &= \frac{v_{\parallel}^2 v_{\perp}^2 B_{IM}^2 v_{ii,\perp}}{4\gamma_{ii} B^2} \\ &+ \frac{v_{ii,\parallel}}{2\gamma_{ii}} \left(v_{\parallel}^2 + v_{\perp}^2 \left(1 - \frac{B_{IM}}{B} \right) \right)^2. \end{aligned} \tag{5.30}$$

Note that $v_{\parallel} \sim \sqrt{\epsilon} v_{ii}$ and that $1 - B_{IM}/B \sim \epsilon$. Then, expression (5.30) simplifies to

$$H_{\mathcal{E}\mathcal{E}}[f_{Mi}] - 2H_{\mathcal{E}\mu}[f_{Mi}] \partial_{\mu} \mathcal{E}_c + H_{\mu\mu}[f_{Mi}] (\partial_{\mu} \mathcal{E}_c)^2 = \frac{v_{\parallel}^2 \bar{v}^2 v_{ii,\perp}(r, \bar{v}, t)}{4\gamma_{ii}} [1 + O(\epsilon)]. \tag{5.31}$$

With this result and the fact that $\tau_W \langle v_{\parallel}^2 \rangle_{\tau, W} = J$, we can simplify (5.28) to

$$\frac{\bar{v}^2 v_{ii,\perp}}{4\gamma_{ii}} \left(J_{c,I} \lim_{J \rightarrow J_{c,I}} \partial_{\mathcal{E}} g_{i,1,I} + J_{c,II} \lim_{J \rightarrow J_{c,II}} \partial_{\mathcal{E}} g_{i,1,II} \right) = \frac{\bar{v}^2 v_{ii,\perp}}{4\gamma_{ii}} J_{c,III} \lim_{J \rightarrow J_{c,III}} \partial_{\mathcal{E}} g_{i,1,III}. \tag{5.32}$$

To obtain a final expression, we need to rewrite $\partial_{\mathcal{E}} g_{i,1,W}(r, \alpha, \mathcal{E}, \mu, t)$ as a linear combination of the derivatives of $g_{i,1,W}(r, \alpha, \bar{v}, J, t)$ with respect to \bar{v} and J . Using $\partial_{\mathcal{E}} g_{i,1,W} = \partial_{\mathcal{E}} \bar{v} \partial_{\bar{v}} g_{i,1,W} + \partial_{\mathcal{E}} J \partial_J g_{i,1,W}$, $\partial_{\mathcal{E}} \bar{v} = \bar{v}^{-1} \sim v_{ii}^{-1}$, $\partial_{\mathcal{E}} J_W = \tau_W \sim \epsilon^{-1/2} R / v_{ii}$, $\partial_{\bar{v}} g_{i,1,W} \sim g_{i,1,W} / v_{ii}$ and $\partial_J g_{i,1,W} \sim \epsilon^{-1/2} g_{i,1,W} / v_{ii} R$, we can approximate $\partial_{\mathcal{E}} g_{i,1,W} \simeq \tau_W \partial_J g_{i,1,W}$ to lowest order in ϵ . This approximation might seem to be in contradiction with the fact that the two terms in the combination $\partial_{\bar{v}} g_{i,1,W} + \partial_{\bar{v}} J_{c,W} \partial_J g_{i,1,W}$ are of the same order, a property that we have used earlier in this paragraph. Note, however, that the function $J_{c,W}(r, \alpha, \bar{v}, t)$ does not have a large derivative with respect to \bar{v} – indeed, when $\phi_{3/2}$ is neglected, $J_{c,W}$ is simply proportional to \bar{v} , giving $\partial_{\bar{v}} J_{c,W} \sim \epsilon^{1/2} R$, which should be compared with the scaling with ϵ of $\partial_{\mathcal{E}} J_W = \tau_W \sim \epsilon^{-1/2} R / v_{ii}$. Using $\partial_{\mathcal{E}} g_{i,1,W} \simeq \tau_W \partial_J g_{i,1,W}$, (5.32) finally becomes

$$J_{c,I} \lim_{J \rightarrow J_{c,I}} \tau_I \partial_J g_{i,1,I} + J_{c,II} \lim_{J \rightarrow J_{c,II}} \tau_{II} \partial_J g_{i,1,II} = J_{c,III} \lim_{J \rightarrow J_{c,III}} \tau_{III} \partial_J g_{i,1,III}. \tag{5.33}$$

This equation gives the relationship among the derivatives $\partial_J g_{i,1,W}$ on different sides of the juncture. The orbit periods τ_W diverge at junctures because particles spend a logarithmically large time at local maxima of U , where the velocity v_{\parallel} vanishes. For this reason, we have to consider the discontinuities of the combination $\tau_W \partial_J g_{i,1,W}$ instead of the discontinuities of $\partial_J g_{i,1,W}$.

Condition (2.24) determines the boundary condition for $g_{i,1,W}$ at $J \rightarrow \infty$. According to expansion (5.3), there is no correction to $h_{i,0}$ of order ϵf_{Mi} , and as a result, the boundary condition for $g_{i,1,W}$ is

$$\lim_{J \rightarrow \infty} g_{i,1,W_{bt}}(r, \alpha, \bar{v}, J, t) = 0. \tag{5.34}$$

We discuss this boundary condition in more detail in Appendix D.

5.2.2. Correction to the passing-particle distribution function

To obtain boundary condition (5.34), we had to use the fact that the first correction to $h_{i,0}$ is $h_{i,3/2}$. Indeed, had there been a correction to the passing-particle distribution function of order ϵf_{Mi} , we could not have imposed condition (5.34).

We proceed to show that expansion (5.3) is consistent and that, indeed, the largest correction to $h_{i,0}$ is of order $\epsilon^{3/2} h_{i,0}$. The correction to $h_{i,0}$ can be driven by the integral contribution of the trapped-particle distribution function $g_{i,1,W}$ to the linearized collision operator, by the time derivatives of n_i and T_i , by the source S_i and by boundary condition (2.25) that requires that the derivatives of $g_{i,W}$ and h_i with respect to \mathbf{v} are continuous at the trapped–passing boundary. The contribution of $g_{i,1,W}$ to the integral piece of the linearized collision operator is $C_{ii,l}^\ell [g_{i,1,W}] \sim \epsilon^{3/2} v_{ii} f_{Mi}$ because $g_{i,1,W}$ is defined in a region of velocity space small by $\sqrt{\epsilon}$. As a result, along with the time derivatives of n_i and T_i and the source S_i , $C_{ii,l}^\ell [g_{i,1,W}]$ drives a piece of h_i that is of order $\epsilon^{3/2} f_{Mi}$, as demonstrated by the expansion of (2.30) in $\epsilon \ll 1$,

$$C_{ii}^\ell [h_{i,3/2}] + \langle C_{ii,l}^\ell [g_{i,1,W}] \rangle_{fs} + \left[\frac{\partial_t n_i}{n_i} + \frac{\partial_t T_i}{T_i} \left(\frac{m_i \bar{v}^2}{T_i} - \frac{3}{2} \right) \right] f_{Mi} = \langle \bar{S}_i \rangle_{fs}, \tag{5.35}$$

where we have used $B \simeq B_0$ and the fact that v_{\parallel} is independent of α and l for most passing particles. Recall that the time derivatives in (2.30) are performed holding \mathcal{E} and μ fixed, and that $\tilde{f}_i^{(0)}$ in the term $\langle (B/|v_{\parallel}|) C_{ii} [h_i, \tilde{f}_i^{(0)}] \rangle_{fs}$ in (2.30) includes both the passing- and trapped-particle distribution functions h_i and $g_{i,W}$. Thus, the integral collisional contribution $\langle C_{ii,l}^\ell [g_{i,1,W}] \rangle_{fs}$ is a result of the term $\langle (B/|v_{\parallel}|) C_{ii} [h_i, \tilde{f}_i^{(0)}] \rangle_{fs}$ in (2.30).

To finish our discussion of the correction to the passing-particle distribution function, we need to consider the boundary condition (2.25). This condition establishes that the derivatives with respect to \mathbf{v} of the trapped- and passing-particle distribution functions must be continuous across the trapped–passing boundary. It can also be viewed as a flux continuity condition: the collisional flux driven by the trapped-particle distribution function across the trapped–passing boundary must be the flux into the passing-particle region, driving a correction to the passing-particle distribution. The collisional flux across the trapped–passing boundary $J \rightarrow \infty$ driven by $g_{i,1,W}$ is, according to the collision operator in (5.25), $-(\bar{v}^2 v_{ii,\perp}/2) \tau_W J \partial_J g_{i,1,W}$ for $J \rightarrow \infty$. If this flux were different from zero, it would drive a correction to the passing-particle distribution function of order $\epsilon^{1/2} f_{Mi}$. We can obtain this result by imposing continuity of derivatives across the trapped–passing boundary. Due to $v_{\parallel} \sim \sqrt{\epsilon} v_{ii}$, $\nabla_v g_{i,1,W} \sim \sqrt{\epsilon} f_{Mi}/v_{ii}$. This gradient must be equal to the gradient of the correction to the passing-particle distribution function, $\nabla_v (h_i - f_{Mi}) \sim (h_i - f_{Mi})/v_{ii}$, giving the incorrect estimate $h_i - f_{Mi} \sim \sqrt{\epsilon} f_{Mi}$, as announced. This estimate is invalid because there is no net collisional flux across the trapped–passing boundary due to $g_{i,1,W}$, a property that we prove below. Moreover, we will see that there is no flux due to $g_{i,3/2,W}$ and hence the collisional flux across the trapped–passing boundary is due to the gradients in velocity space of $g_{i,2,W}$. Since $\nabla_v g_{i,2,W} \sim \epsilon^{3/2} f_{Mi}/v_{ii}$, the correction to the passing-particle distribution function is indeed $h_{i,3/2}$ (see Appendix D for more detail on how the flux continuity condition across the

barely-passing-particle region leads to this result). These arguments show that boundary condition (5.34) for $g_{i,1,W}$ is justified because the largest correction to the passing-particle distribution function is indeed of order $\epsilon^{3/2}f_{Mi}$. We proceed to show that $\tau_W J \partial_J g_{i,1,W}$ and $\tau_W J \partial_J g_{i,3/2,W}$ vanish at the trapped–passing boundary $J \rightarrow \infty$, that is, that neither $g_{i,1,W}$ nor $g_{i,3/2,W}$ drive a collisional flux from the trapped-particle region into the passing-particle region.

We start by showing

$$\lim_{J \rightarrow \infty} \tau_{W_{bt}} J \partial_J g_{i,1,W_{bt}}(r, \alpha, \bar{v}, J, t) = 0. \tag{5.36}$$

We first integrate equation (5.25) over the region in J where orbits in well W with given values of α and \bar{v} exist, $J \in [J_{m,W}(r, \alpha, \bar{v}, t), J_{M,W}(r, \alpha, \bar{v}, t)]$,

$$\begin{aligned} & \frac{c\phi'_0}{\Psi'_t} \partial_\alpha \left[\int_{J_{m,W}}^{J_{M,W}} (g_{i,1,W} + r_{i,1,W} \Upsilon f_{Mi}) \, dJ \right] \\ & - \frac{c\phi'_0}{\Psi'_t} \partial_\alpha J_{M,W} [g_{i,1,W}(r, \alpha, \bar{v}, J_{M,W}, t) + r_{i,1,W}(r, \alpha, \bar{v}, J_{M,W}, t) \Upsilon f_{Mi}] \\ & + \frac{c\phi'_0}{\Psi'_t} \partial_\alpha J_{m,W} [g_{i,1,W}(r, \alpha, \bar{v}, J_{m,W}, t) + r_{i,1,W}(r, \alpha, \bar{v}, J_{m,W}, t) \Upsilon f_{Mi}] \\ & - \frac{\bar{v}^2 v_{ii,\perp}}{4} \left(J_{M,W} \lim_{J \rightarrow J_{M,W}} \tau_W \partial_J g_{i,1,W} - J_{m,W} \lim_{J \rightarrow J_{m,W}} \tau_W \partial_J g_{i,1,W} \right) = 0. \end{aligned} \tag{5.37}$$

The second and third terms in this equation are included to cancel the derivatives with respect to α of the limits of the integral in the first term. The values of $J_{m,W}$ and $J_{M,W}$ are either 0, ∞ or values at which there is a juncture between different wells. We proceed to integrate equation (5.37) for all the values of α allowed in well W for a given \bar{v} . The interval $[\alpha_{L,W}(r, \bar{v}, t), \alpha_{R,W}(r, \bar{v}, t)]$ is the region in α where orbits in well W with a given value of \bar{v} exist. The limits $\alpha_{L,W}$ and $\alpha_{R,W}$ exist for two different reasons:

- (i) either well W closes at $\alpha_{L,W}$ and $\alpha_{R,W}$ because $J_{m,W} = J_{M,W}$; or
- (ii) well W extends to all values of α and hence $\alpha_{L,W} = 0$ and $\alpha_{R,W} = 2\pi$.

In figure 3, we give an example in which one of the wells, well IV , disappears ($J_{m,IV} = J_{M,IV}$ at $\alpha = \alpha_{R,IV}$) and a new well, well VI , appears ($J_{m,VI} \neq J_{M,VI}$ for $\alpha > \alpha_{R,IV}$). Not all cases in which a well disappears are as straightforward as the example in figure 3. We discuss a pathological example in Appendix E, where we also show that choices can be made such that either $J_{m,W} = J_{M,W}$ or the distribution function is periodic at $\alpha_{L,W}$ and $\alpha_{R,W}$. In either case, integrating equation (5.37) over α , we find

$$\begin{aligned} & - \frac{c\phi'_0}{\Psi'_t} \int_{\alpha_{L,W}}^{\alpha_{R,W}} \left[\partial_\alpha J_{M,W} (g_{i,1,W}(r, \alpha, \bar{v}, J_{M,W}, t) + r_{i,1,W}(r, \alpha, \bar{v}, J_{M,W}, t) \Upsilon f_{Mi}) \right. \\ & \quad \left. - \partial_\alpha J_{m,W} (g_{i,1,W}(r, \alpha, \bar{v}, J_{m,W}, t) + r_{i,1,W}(r, \alpha, \bar{v}, J_{m,W}, t) \Upsilon f_{Mi}) \right] \, d\alpha \\ & - \frac{\bar{v}^2 v_{ii,\perp}}{4} \int_{\alpha_{L,W}}^{\alpha_{R,W}} \left(J_{M,W} \lim_{J \rightarrow J_{M,W}} \tau_W \partial_J g_{i,1,W} - J_{m,W} \lim_{J \rightarrow J_{m,W}} \tau_W \partial_J g_{i,1,W} \right) \, d\alpha \\ & = 0. \end{aligned} \tag{5.38}$$

Summing over all possible well indices W , and using the fact that, at the juncture of several wells, $g_{i,1,W}$ is continuous and $\tau_W \partial_J g_{i,1,W}$ satisfies (5.33), several terms cancel and we

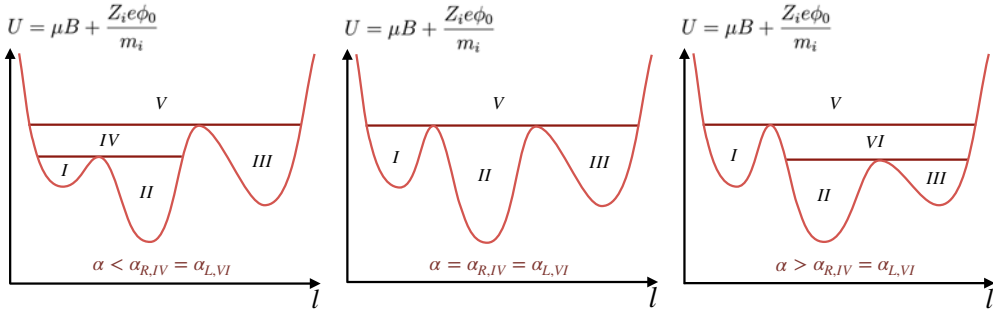


FIGURE 3. Example of limits $\alpha_{L,W}$ and $\alpha_{R,W}$. Well IV disappears at $\alpha = \alpha_{R,IV}$, and at that same value of α , a new well VI appears, giving $\alpha_{L,VI} = \alpha_{R,IV}$.

find the result in (5.36). To illustrate the different cancellations that lead to (5.36), we consider the juncture of wells I, II and IV in figure 3, characterized by $J_{c,I}$, $J_{c,II}$ and $J_{c,IV} = J_{c,I} + J_{c,II}$. For well I, we find $J_{m,I} = 0$ and $J_{M,I} = J_{c,I}$. Similarly, for well II, we have $J_{m,II} = 0$ and $J_{M,II} = J_{c,II}$. For well IV we only know the minimum value $J_{m,IV} = J_{c,IV}$. In the sum of (5.38) over all wells, the terms proportional to $J_{m,I}$, $J_{m,II}$ and their derivatives with respect to α vanish. The contributions from the juncture of wells I, II and IV in figure 3 are then

$$\begin{aligned}
 & -\frac{c\phi'_0}{\Psi'_t} \left[\int_{\alpha_{L,I}}^{\alpha_{R,I}} \partial_\alpha J_{c,I} (g_{i,1,I}(r, \alpha, \bar{v}, J_{c,I}, t) + r_{i,1,I}(r, \alpha, \bar{v}, J_{c,I}, t) \Upsilon f_{Mi}) \, d\alpha \right. \\
 & \quad + \int_{\alpha_{L,II}}^{\alpha_{R,II}} \partial_\alpha J_{c,II} (g_{i,1,II}(r, \alpha, \bar{v}, J_{c,II}, t) + r_{i,1,II}(r, \alpha, \bar{v}, J_{c,II}, t) \Upsilon f_{Mi}) \, d\alpha \\
 & \quad \left. - \int_{\alpha_{L,IV}}^{\alpha_{R,IV}} \partial_\alpha J_{c,IV} (g_{i,1,IV}(r, \alpha, \bar{v}, J_{c,IV}, t) + r_{i,1,IV}(r, \alpha, \bar{v}, J_{c,IV}, t) \Upsilon f_{Mi}) \, d\alpha \right] \\
 & - \frac{\bar{v}^2 \nu_{ii,\perp}}{4} \left[\int_{\alpha_{L,I}}^{\alpha_{R,I}} J_{c,I} \lim_{J \rightarrow J_{c,I}} \tau_I \partial_J g_{i,1,I} \, d\alpha + \int_{\alpha_{L,II}}^{\alpha_{R,II}} J_{c,II} \lim_{J \rightarrow J_{c,II}} \tau_{II} \partial_J g_{i,1,II} \, d\alpha \right. \\
 & \quad \left. - \int_{\alpha_{L,IV}}^{\alpha_{R,IV}} J_{c,IV} \lim_{J \rightarrow J_{c,IV}} \tau_{IV} \partial_J g_{i,1,IV} \, d\alpha \right] + \dots \\
 & = 0.
 \end{aligned} \tag{5.39}$$

The ellipsis points \dots here indicates that there are more terms corresponding to other junctures that we have not included in the equation. The first three lines of (5.39) cancel each other because of continuity of $g_{i,1,W}$ and $r_{i,1,W}$ across the juncture and the fact that $\partial_\alpha J_{c,I} + \partial_\alpha J_{c,II} = \partial_\alpha J_{c,IV}$. The displacement $r_{i,1,W}$ is continuous across junctures because each juncture has a single value of $\bar{\lambda}$, $\bar{\lambda} = \bar{\lambda}_c(r, \alpha, \bar{v})$. Lines four and five of (5.39) vanish because of condition (5.33).

The correction $g_{i,3/2,W}$ to the trapped-particle distribution function is shown to be independent of α and J in Appendix F. Using \bar{v} instead of v is crucial to obtain this result (see Appendix F for more details on the role of \bar{v} in the derivation). By continuity across the trapped–passing boundary, we also find that

$$g_{i,3/2,W}(r, \alpha, \bar{v}, J, t) = h_{i,3/2}(r, \bar{v}, 0, t). \tag{5.40}$$

Thus, $\tau_W J \partial_J g_{i,3/2,W}$ is zero at the trapped–passing boundary.

5.2.3. *Electric potential $\phi_{3/2}$*

To determine the piece of the electric potential that is not a flux function, we use quasineutrality (2.1). To calculate the integral of $\bar{f}_i^{(0)}$ over velocity space, we employ $\bar{v} \simeq v + Z_i e \phi_{3/2} / m_i v$ to perform a Taylor expansion around v , finding

$$g_{i,W}(r, \alpha, \bar{v}, J, t) = f_{Mi}(r, v, t) + g_{i,1,W}(r, \alpha, v, J, t) + O(\epsilon^{3/2} f_{Mi}), \tag{5.41}$$

and

$$h_i(r, \bar{v}, \xi, t) = f_{Mi}(r, v, t) + h_{i,3/2}(r, v, \xi, t) - \frac{Z_i e \phi_{3/2}(r, \alpha, l, t)}{T_i(r)} f_{Mi}(r, v, t) + O(\epsilon^2 f_{Mi}). \tag{5.42}$$

Then, the lowest-order quasineutrality equation (2.1) gives

$$\left(\frac{en_e}{T_e} + \frac{Z_i^2 n_i}{T_i} \right) \phi_{3/2} = Z_i \int \sum_{W \in \mathcal{W}} g_{i,1,W} d^3v, \tag{5.43}$$

showing that the next-order correction to the flux function ϕ_0 is indeed small by $\epsilon^{3/2}$, as predicted in § 4. Here, $n_e \simeq \hat{n}_e \exp(e\phi_0/T_e)$, and $\int g_{i,1,W} d^3v \sim \epsilon^{3/2} n_i$ because, again, the region of velocity space where $g_{i,1,W}$ is defined is small by $\sqrt{\epsilon}$. Note that we need not include $\int h_{i,3/2} d^3v$ in the quasineutrality equation because this integral does not depend on α or l to lowest order in ϵ , and hence can be absorbed into the definition of $n_i(r)$.

5.3. *Ion transport equations*

We finish by integrating equations (2.27) and (2.30) to find equations for n_i and T_i . Before we integrate, we rewrite the equations in a convenient form. Using (5.15) and (5.16) and employing $Z_i e \langle \partial_t \phi \rangle_{\tau,W} / m_i = -\tau_W^{-1} \partial_t (\tau_W \langle v_{\parallel}^2 \rangle_{\tau,W})$ and $\partial_{\mathcal{E}} (\tau_W \langle v_{\parallel}^2 \rangle_{\tau,W}) = \tau_W$, (2.27) can be written as

$$\begin{aligned} & \frac{1}{\tau_W} \partial_t (\tau_W g_{i,W}) + \frac{1}{\tau_W} \partial_{\mathcal{E}} \left(\frac{Z_i e \tau_W}{m_i} \langle \partial_t \phi \rangle_{\tau,W} g_{i,W} \right) \\ & + \frac{1}{\tau_W} \partial_{\alpha} (\tau_W \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla \alpha \rangle_{\tau,W} g_{i,W}) \\ & + \frac{1}{\tau_W} \partial_r (\tau_W \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} g_{i,W}) = \langle C_{ii} [g_{i,W}, \bar{f}_i^{(0)}] \rangle_{\tau,W} + \langle \bar{S}_i \rangle_{\tau,W}. \end{aligned} \tag{5.44}$$

Using $Z_i e \langle (B/|v_{\parallel}|) \partial_t \phi \rangle_{fs} / m_i = -\partial_t \langle B|v_{\parallel}| \rangle_{fs}$ and $\partial_{\mathcal{E}} \langle B|v_{\parallel}| \rangle_{fs} = \langle B/|v_{\parallel}| \rangle_{fs}$, (2.30) can be written as

$$\partial_t \left\langle \left\langle \frac{B}{|v_{\parallel}|} \right\rangle_{fs} h_i \right\rangle + \partial_{\mathcal{E}} \left(\frac{Z_i e}{m_i} \left\langle \frac{B}{|v_{\parallel}|} \partial_t \phi \right\rangle_{fs} h_i \right) + \left\langle \frac{B}{|v_{\parallel}|} C_{ii} [h_i, \bar{f}_i^{(0)}] \right\rangle_{fs} = \left\langle \frac{B}{|v_{\parallel}|} \bar{S}_i \right\rangle_{fs}. \tag{5.45}$$

Multiplying (5.44) by τ_W and (5.45) by 1, integrating over α , \mathcal{E} , μ and φ , and summing over both signs of σ and over the well index W , we find ion particle conservation equation

$$\partial_t n_i + \frac{1}{V_i} \partial_r (V_i \Gamma_i) = \left\langle \int S_i d^3v \right\rangle_{fs}, \tag{5.46}$$

where

$$\Gamma_i := 4\pi \left\langle \int_0^\infty d\mu \int_U^{U_M} d\mathcal{E} \frac{B}{|v_\parallel|} \sum_{W \in \mathcal{W}} g_{i,W} \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} \right\rangle_{\text{fs}} \quad (5.47)$$

is the particle flux. Note that the passing-particle piece h_i does not contribute due to property (2.33). Using the expansion in (5.2), the integration variables $v \simeq \bar{v}$ and J and the lowest-order expression for the radial drifts in (4.5), and employing equation (5.24) to relate the drifts to $r_{i,1,W}$, defined in (5.26), the particle flux becomes

$$\Gamma_i \simeq \frac{2\pi c\phi'_0}{B_0 V'} \int_0^\infty dv \sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \int_{J_{m,W}}^{J_{M,W}} dJ v g_{i,1,W} \partial_\alpha r_{i,1,W} \sim \epsilon^{5/2} \rho_{i*} n_i v_{ii}. \quad (5.48)$$

To obtain the order of magnitude estimate, we have used $\phi'_0 \sim T_i/ea$, $V' \sim Ra$, $J \sim \sqrt{\epsilon} v_{ii} R$ and $r_{i,1,W} \sim \epsilon a$. The order of magnitude estimate in (5.48) justifies estimates (5.1).

The estimate for the size of the particle flux in (5.48) can also be obtained from a random walk argument. In the introduction, in (1.5), we argued that trapped particle orbits had a radial width $w \sim \epsilon a$ due to the smallness of the magnetic drift compared with the $\mathbf{E} \times \mathbf{B}$ drift, $|v_M|/|v_E| \sim \epsilon$. These orbits are interrupted by collisions that have an effective collision frequency ν_{ii}/ϵ , causing a radial random walk with steps of length w every time ϵ/ν_{ii} . The effective collision frequency ν_{ii}/ϵ is larger than ν_{ii} because small angle collisions can easily detrap particles with $v_\parallel \sim \sqrt{\epsilon} v_{ii}$ by providing a small amount of parallel momentum. Since we assume $\nu_{i*} \sim \rho_{i*}$, $\nu_{ii}/\epsilon \sim \rho_{i*} v_{ii}/a$. Thus, the random walk diffusion coefficient associated with trapped-particle orbits is $D \sim \sqrt{\epsilon} w^2 \rho_{i*} v_{ii}/a$. The factor of $\sqrt{\epsilon}$ in the estimate of the diffusion coefficient is due to the fact that only trapped particles, which are a fraction of order $\sqrt{\epsilon} \ll 1$ of the total number of particles, participate in the diffusion. Using $D \sim \epsilon^{5/2} \rho_{i*} a v_{ii}$ and $\Gamma_i \sim D |\nabla n_i| \sim D n_i/a$, we obtain the estimate in (5.48).

Multiplying (5.44) and (5.45) by $m_i \mathcal{E}$ and integrating over velocity space, we find the ion energy conservation equation

$$\begin{aligned} \partial_t \left(\frac{3}{2} n_i T_i + Z_i e n_i \phi_0 \right) + \frac{1}{V'} \partial_r [V' (Q_i + Z_i e \phi_0 \Gamma_i)] = Z_i e n_i \partial_t \phi_0 \\ + \left\langle \int \left(\frac{1}{2} m_i v^2 + Z_i e \phi_0 \right) S_i d^3 v \right\rangle_{\text{fs}}, \end{aligned} \quad (5.49)$$

where

$$Q_i \simeq \frac{\pi m_i c \phi'_0}{B_0 V'} \int_0^\infty dv \sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \int_{J_{m,W}}^{J_{M,W}} dJ v^3 g_{i,1,W} \partial_\alpha r_{i,1,W} \sim \epsilon^{5/2} \rho_{i*} n_i T_i v_{ii} \quad (5.50)$$

is the ion energy flux. Note that, using (5.46), (5.49) can also be written as

$$\partial_t \left(\frac{3}{2} n_i T_i \right) + \frac{1}{V'} \partial_r (V' Q_i) = -Z_i e \phi'_0 \Gamma_i + \left\langle \int \frac{1}{2} m_i v^2 S_i d^3 v \right\rangle_{\text{fs}}. \quad (5.51)$$

The expressions for the particle and energy fluxes given in (5.48) and (5.50) would seem to suggest that both fluxes are proportional to the radial electric field ϕ'_0 . This is, however, not the case as $r_{i,1,W}$, defined in (5.26), is inversely proportional to ϕ'_0 . Thus, ϕ'_0 only enters in the fluxes through its influence on the correction to the distribution function $g_{i,1,W}$.

5.4. Summary

To summarize, for $\rho_{i*} \sim v_{i*}$, the ion distribution function is Maxwellian to lowest order. The density and temperature of this Maxwellian can be calculated using particle and energy conservation equations once the particle and energy fluxes in (5.48) and (5.50) are obtained. To calculate these fluxes, we must obtain the correction $g_{i,1,W}$ that is only defined in the trapped-particle region. The equation for $g_{i,1,W}$ is (5.25). This equation must be solved along with boundary condition (5.34) and relation (5.33) for the junctures of different types of wells.

Importantly, due to the presence of the small correction $\phi_{3/2}$ to the potential, we had to use the velocity-space coordinates \bar{v} and J , defined in (4.10) and (4.11), and the approximate pitch-angle variable $\bar{\lambda}$, defined in (5.21). These variables were crucial to show that boundary condition (5.34) applies, but once this is done, we can use the approximations $\bar{v} \simeq v$ and $\bar{\lambda} \simeq \lambda$, where λ is defined in (4.8). From here on, we replace \bar{v} with v and $\bar{\lambda}$ with λ .

The same set of equations that we have obtained can be derived from the kinetic equation implemented in DKES (Hirshman *et al.* 1986) in the limit given by $\rho_{i*} \sim v_{i*} \ll 1$ and $\epsilon \ll 1$. The procedure to derive (5.25), (5.33) and (5.34) from the DKES kinetic equation is similar to the method described in § 2 and in this section. We sketch the derivation in Appendix G. There are three differences with our derivation that are worth mentioning.

- (i) The DKES equations assume from the start that the lowest-order distribution is Maxwellian and that the potential is an exact flux function.
- (ii) We use the second adiabatic invariant as a variable because it remains constant as the particle moves. However, the DKES kinetic equation does not ensure that the second adiabatic invariant remains constant. Instead, the DKES kinetic equation maintains the quantity

$$\hat{J}_W := 2 \int_{l_{bL,W}}^{l_{br,W}} B |v_{\parallel}| dl \quad (5.52)$$

constant. In the expansion in $\epsilon \ll 1$, this quantity is approximately proportional to J , $\hat{J} \simeq B_0 J$.

- (iii) The perturbation to the passing-particle distribution function h_i calculated by DKES is of order $\epsilon^{3/2} f_{Mi}$, but unlike the physical correction $h_{i,3/2}$, determined by (5.35), the DKES correction to the passing-particle distribution function is not driven by the time derivatives of density, potential and temperature, the heating and fuelling sources or the integral terms of the collision operator. In DKES, the perturbation to the passing-particle distribution function is only driven by the collisional flux from and to the trapped-particle region because DKES only evolves the perturbation to the Maxwellian and uses a simple pitch-angle scattering operator without integral contributions.

The three previous differences mean that, even though the DKES kinetic equation leads to the same equations as the full kinetic equation to lowest order in $\rho_{i*} \sim v_{i*} \ll 1$ and $\epsilon \ll 1$, the higher-order equations given by DKES will be very different from the physical ones. The higher-order equations merit further study because they might be important even for small values of ϵ : for example, $g_{i,1,W}$ is not zero at the trapped-passing boundary, but of order $\sqrt{\epsilon} g_{i,1,W}$, and this boundary value is determined by the perturbation to the passing-particle distribution function, $h_i - f_{Mi}$, that we have neglected. DKES cannot determine $h_i - f_{Mi}$ in the limit given by $\rho_{i*} \sim v_{i*} \ll 1$ and $\epsilon \ll 1$.

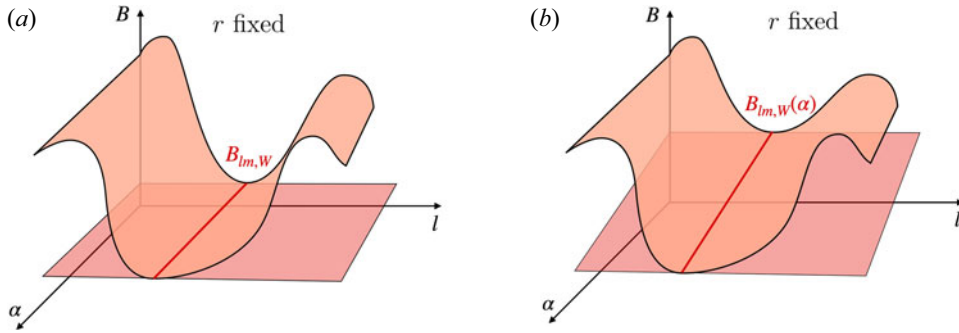


FIGURE 4. (a) A particle deeply trapped in well W can move across field lines along the red line in the particular case in which the local minima of B , $B_{lm,w}(r, \alpha)$, are independent of α . The local minima of B in omnigenous magnetic fields behave in this manner. (b) In the more general case in which $B_{lm,w}$ varies with α , deeply trapped particles cannot move along the red line to precess around the flux surface unless they are allowed to move radially as well.

Finally, to demonstrate the advantages of (5.25), (5.33) and (5.34), we consider the deeply trapped particles. These particles were singled out as problematic for, for example, the first version of the local orbit-averaged code KNOSOS (Velasco *et al.* 2020). The first version of KNOSOS was valid only for stellarators close to omnigenity (Calvo *et al.* 2017), and as a result, the radial displacement of particles is neglected to lowest order in the expansion in closeness to omnigenity. This approximation is valid for most particles, but fails for deeply and barely trapped particles because these particles cannot move across magnetic field lines within the same flux surface if they do not move simultaneously along ∇r (across flux surfaces).

We proceed to explain this problem in detail. First, we consider a particle deeply trapped in a given magnetic well W on a given flux surface r . The deeply trapped particles have a small J and hence must have a parallel velocity close to zero to keep J constant. Their total energy per unit mass is, to lowest order in ϵ ,

$$\mathcal{E} = \mu B_{lm,w}(r, \alpha) + \frac{Z_i e}{m_i} \phi_0(r, t), \tag{5.53}$$

where $B_{lm,w}(r, \alpha)$ denotes the local minimum of B in well W along the magnetic field line determined by r and α . Since the total energy \mathcal{E} is a constant of the motion, a deeply trapped particle can move across magnetic field lines of the same flux surface r only if the value of $B_{lm,w}$ remains the same while it does so, as in the case represented in figure 4(a). In the case of omnigenity, treated by Cary & Shasharina (1997a,b) and Parra *et al.* (2015), $B_{lm,w}$ remains the same for different values of α on a given flux surface. Thus, deeply trapped particles precess around the flux surface without moving radially.

In a generic large aspect ratio stellarator with mirror ratio close to unity, the deeply trapped particles cannot move across magnetic field lines on a given flux surface, as shown on figure 4(b). However, since the electric potential ϕ_0 varies with r , it is possible for v_{\parallel} to be approximately equal to zero after a displacement $\Delta\alpha$ if the particle moves across flux surfaces by a distance Δr . Using (5.53), we find that $\Delta\alpha$ and Δr must satisfy

$$0 = \mu \partial_{\alpha} B_{lm,w}(\alpha, r) \Delta\alpha + \frac{Z_i e \phi_0'(r, t)}{m_i} \Delta r, \tag{5.54}$$

where we have neglected the radial derivative of $B_{lm,w}(r, \alpha)$ because it is small by a factor of ϵ compared with the radial derivative of ϕ_0 . This is consistent with the velocities $d\alpha/dt$ and dr/dt obtained from the transit average drifts in (4.5) and (4.6),

$$\frac{d\alpha}{dt} \simeq \frac{c\phi'_0}{\Psi'_i} \tag{5.55}$$

and

$$\frac{dr}{dt} \simeq -\frac{m_i c v^2}{2Z_i e B_0 \Psi'_i} \partial_\alpha B_{1,lm,w}, \tag{5.56}$$

where $B_{1,lm,w}(r, \alpha)$ is the local minimum of B_1 in well W along the magnetic field line determined by r and α . Moreover, (5.54) is also consistent with the definition of the radial displacement $r_{i,1,w}$ given in (5.26). Indeed, to lowest order in ϵ , $\lambda = B_{lm,w}^{-1}$ for deeply trapped particles, and $\lambda \rightarrow B_M^{-1}$ for $J \rightarrow \infty$, giving

$$r_{i,1,w} \simeq \frac{m_i v^2}{2Z_i e \phi'_0} \frac{B_{1,M}(r) - B_{1,lm,w}(r, \alpha)}{B_0}. \tag{5.57}$$

This equation is a solution to (5.55) and (5.56), and it shows how the particle moves radially to increase and decrease its electric energy to keep its total energy constant.

We can draw the path followed by a deeply trapped particle with $J = 0$ on an (l, α) plane, sketched in figure 5 for $\phi'_0 > 0$ and $\Psi'_i > 0$. The particle moves towards increasing α following the local minima $B_{lm,w}$. Simultaneously, the particle moves back and forth across flux surfaces, increasing r when $\partial_\alpha B < 0$ or decreasing r when $\partial_\alpha B > 0$. It is interesting to note that

$$\frac{\Delta r}{\Delta \alpha} \sim \epsilon a, \tag{5.58}$$

showing that the oscillating displacement across flux surfaces is much smaller than the precession around it. Equation (5.25) reproduces this motion for deeply trapped particles.

6. The $1/\nu$ regime in large aspect ratio stellarators with mirror ratios close to unity

In this section, we briefly consider the limit $\nu_{i*} \gg \rho_{i*}$. In this limit, we can neglect the term $(c\phi'_0/\Psi'_i) \partial_\alpha g_{i,1,w}$ in (5.25), giving to lowest order

$$-\frac{v^2 v_{ii,\perp}(v)}{4} \partial_J (\tau_w J \partial_J g_{i,1,w}) \simeq -\frac{c\phi'_0}{\Psi'_i} \partial_\alpha r_{i,1,w} \Upsilon_i f_{Mi}. \tag{6.1}$$

This equation can be integrated using boundary condition (5.34) and relation (5.33) at the junctures of several types of wells. The final result is a trapped-particle distribution function of order $g_{i,1,w} \sim \epsilon \rho_{i*} f_{Mi} / \nu_{i*}$. Using this result in (5.48) and (5.50) for the particle and energy fluxes, we obtain

$$\Gamma_i \sim \frac{\rho_{i*}}{\nu_{i*}} \epsilon^{5/2} \rho_{i*} n_i v_{ti}, \quad Q_i \sim \frac{\rho_{i*}}{\nu_{i*}} \epsilon^{5/2} \rho_{i*} n_i T_i v_{ti}. \tag{6.2a,b}$$

These are the typical $1/\nu$ regime order-of-magnitude estimates (Ho & Kulsrud 1987).

Estimates (6.2a,b) can be obtained using a random walk argument. For $\nu_{i*} \gg \rho_{i*}$, a typical trapped particle moves radially due to the radial magnetic drift $\mathbf{v}_{Mi} \cdot \nabla r \sim \epsilon \rho_{i*} v_{ti}$ until a collision detraps it. As a result, particles move a distance $\mathbf{v}_{Mi} \cdot \nabla r / \nu_{\text{eff}}$ between

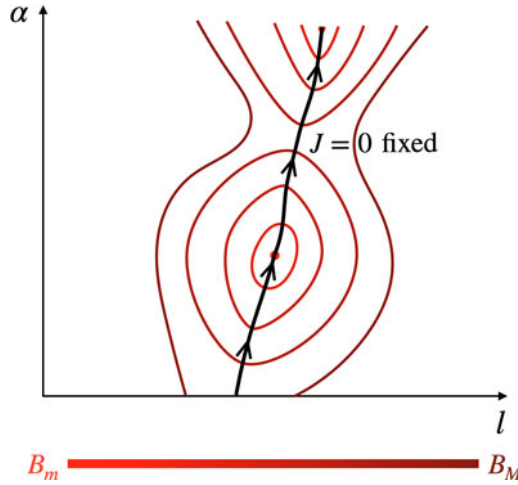


FIGURE 5. Map of the magnetic field magnitude B in an (l, α) plane. In black with arrows, we sketch the path followed by particles with $J = 0$ in a stellarator with $\phi'_0 > 0$ and $\Psi'_t > 0$. The coordinate r is not constant, and its variation around its average value has the opposite sign to the value of $B_1 = B - B_0$, as shown in (5.57).

collisions, where $v_{\text{eff}} \sim v_{ii}/\epsilon$ is the effective collision frequency for trapped particles. After each detrapping collision, a trapped particle becomes passing and moves along the magnetic field line until another low angle collision reduces its parallel velocity sufficiently to trap it in a different well. The magnitude and sign of $\mathbf{v}_{Mi} \cdot \nabla r$ in the new well will in general be different from the sign and magnitude of the radial drift in the original well because the particle has moved a significant distance along the magnetic field line while being a passing particle. For this reason, the radial displacement of a typical trapped particle is similar to a random walk with a characteristic step size $\mathbf{v}_{Mi} \cdot \nabla r / v_{\text{eff}} \sim (\rho_{i*} / v_{i*}) \epsilon a \ll \epsilon a$ and a characteristic time $1/v_{\text{eff}} \sim \epsilon / v_{ii}$. The corresponding diffusion coefficient is $D \sim \sqrt{\epsilon} (\mathbf{v}_{Mi} \cdot \nabla r / v_{\text{eff}})^2 v_{\text{eff}}$, where $\sqrt{\epsilon}$ is the estimate for the trapped-particle fraction. With $D \sim \epsilon^{5/2} \rho_{i*}^2 a v_{ii} / v_{i*}$ and using $\Gamma_i \sim D |\nabla n_i| \sim D n_i / a$ and $Q_i \sim n_i D |\nabla T_i| \sim D n_i T_i / a$, we recover estimates (6.2a,b).

7. The ν regime in large aspect ratio stellarators with mirror ratios close to unity

In this section, we study the limit $v_{i*} \ll \rho_{i*}$. We expand $g_{i,1,W}$ in $v_{i*} / \rho_{i*} \ll 1$,

$$g_{i,1,W} = g_{i,1,W}^{(0)} + g_{i,1,W}^{(1)} + \dots, \tag{7.1}$$

with $g_{i,1,W}^{(n)} \sim (v_{i*} / \rho_{i*})^n \epsilon f_{Mi}$. We proceed to expand (5.25) in $v_{i*} / \rho_{i*} \ll 1$. In § 7.1, we solve the lowest-order version of (5.25) and we show that the lowest-order distribution function does not give rise to radial fluxes of particles or energy. In § 7.2, we go to next order in $v_{i*} / \rho_{i*} \ll 1$. Finally, in § 7.3, we calculate the particle and energy fluxes.

7.1. Lowest-order distribution function

Equation (5.25) becomes, to lowest order in $v_{i*} / \rho_{i*} \ll 1$,

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,1,W}^{(0)} = -\frac{c\phi'_0}{\Psi'_t} \partial_\alpha r_{i,1,W} \Upsilon_i f_{Mi}. \tag{7.2}$$

Using the fact that $\Upsilon_{f_{Mi}}$ does not vary with α , we obtain the expression

$$g_{i,1,W}^{(0)}(r, \alpha, v, J, t) = -r_{i,1,W} \Upsilon_{f_{Mi}} + K_{i,W}(r, v, J, t), \quad (7.3)$$

where the function $K_{i,W}$ does not depend on α .

Due to the existence of junctures between different types of wells such as the one sketched in figure 1, the function $K_{i,W}$ can be independent of J in large regions of velocity space. Before we explain why, we need to consider what happens with well junctures in the limit $v_{i*}/\rho_{i*} \ll 1$. In this limit, it is in general impossible to impose continuity of $g_{i,1,W}$ across the juncture, or condition (5.33) that relates the derivatives $\partial_J g_{i,1,W} \simeq \partial_J g_{i,1,W}^{(0)}$ on different sides of a juncture to each other. This is hardly surprising since continuity of $g_{i,1,W}$ and condition (5.33) are a result of collisions, and we have neglected collisions to lowest order. In reality, continuity and condition (5.33) are not satisfied in appearance only, because boundary layers where collisions become important can form around well junctures.

The rest of this subsection is split into three parts. In the first one, we show that collisional boundary layers only form around certain junctures, and we find a condition that $K_{i,W}$ must satisfy when these boundary layers form. In the second part, we study an example that illustrates the shape of $K_{i,W}$. We finish with a general discussion.

7.1.1. *Junctures of different types of wells for very small collision frequencies*

Not all well junctures are problematic and need a boundary layer. Whether a juncture has a boundary layer or not depends on the direction of the velocity in α , $c\phi'_0/\Psi'_t$, and the derivatives of the juncture coordinates $J_{c,W}(r, \alpha, v)$ with respect to α . To explain this further, we consider the juncture in figure 1. Some particles in well *I* leave this well if the time derivative of $J_{c,I}(r, \alpha, v)$ is negative. Indeed, if a particle with $J = J_{c,I}(r(t), \alpha(t), v(t))$ at time t were to stay in well *I*, it would have to do so keeping its second adiabatic invariant constant. However, such a particle would find itself with $J > J_{c,I}(r(t + \Delta t), \alpha(t + \Delta t), v(t + \Delta t))$ at $t + \Delta t$. Since there are no possible orbits with $J > J_{c,I}$ in well *I*, the particle must have moved into well *II* or well *III*, and in doing so, the value of J of the particle has changed abruptly. This is consistent with J being an adiabatic invariant: it is only constant when the motion is a slowly changing periodic orbit, a description that does not apply to a transition from one well to another. These transitions between different types of wells can be understood to be a result of the conservation of phase-space volume. The second adiabatic invariant is the phase-space volume inside a trapped orbit. When $J_{c,I}$ becomes lower than the J of a particle, the phase-space volume of the particle orbit cannot be contained inside well *I* and the particle ‘spills over’ into wells *II* and *III* (Dobrott & Greene 1971; Cary *et al.* 1986).

Equivalently to particles moving out of well *I* when the time derivative of $J_{c,I}$ is negative, some particles move into well *I* if the time derivative of $J_{c,I}$ is positive. As we will see shortly, the fact that the time derivative of $J_{c,I}$ is positive implies that particles are leaving well *II*, well *III* or both, and since these particles must move into another well, it is intuitive that conservation of phase-space volume will force these particles to move to a well where there is ‘space’. Well *I*, with a positive time derivative of $J_{c,I}$, is such a well because at time t all particles in that well have $J \leq J_{c,I}(r(t), \alpha(t), v(t))$, and hence, if no particle were to move into well *I*, at time $t + \Delta t$ there would be no particles that have second adiabatic invariant between $J_{c,I}(r(t), \alpha(t), v(t))$ and $J_{c,I}(r(t + \Delta t), \alpha(t + \Delta t), v(t + \Delta t))$.

To summarize, some particles move out of well *I* if the time derivative of $J_{c,I}$ is negative and some particles move into well *I* if the time derivative of $J_{c,I}$ is positive. Similarly, some particles transition out of well *II* if the time derivative of $J_{c,II}$ is negative, and some transition into it if the time derivative of $J_{c,II}$ is positive. The signs reverse for well *III*

because the orbits in this well have J larger than $J_{c,III}$. Thus, particles move out of well *III* if the time derivative of $J_{c,III}$ is positive, and transition into it if the time derivative of $J_{c,III}$ is negative.

Since both r and v have time derivatives that are small in ϵ , the time derivative of $J_{c,W}(r, \alpha, v)$ is simply

$$\frac{dJ_{c,W}}{dt} \simeq v_E \cdot \nabla \alpha \partial_\alpha J_{c,W} \simeq \frac{c\phi'_0}{\Psi'_t} \partial_\alpha J_{c,W}. \tag{7.4}$$

Importantly, condition (5.27) imposes that $\partial_\alpha J_{c,I} + \partial_\alpha J_{c,II} = \partial_\alpha J_{c,III}$. With this equality, it is easy to check that, if some particles are moving into well *I*, i.e. $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,I} > 0$, some particles must be leaving well *II*, $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,II} < 0$, well *III*, $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,III} > 0$, or both. It cannot be the case that particles are moving into all three wells, or leaving all three wells.

Armed with the results above, we can start discussing the collisional boundary layers. There are two distinct cases to consider:

- (i) the signs of $c\phi'_0/\Psi'_t$ and $\partial_\alpha J_{c,W}$ are such that some particles leave one of the wells and move into the other two wells; in this case, there is no boundary layer around the juncture; or
- (ii) the signs of $c\phi'_0/\Psi'_t$ and $\partial_\alpha J_{c,W}$ are such that some particles leave two of the wells and move into the third well; in this case, a thin boundary layer forms around the juncture.

To illustrate what happens in the case that some particles leave one of the wells and enter the other two, we consider the following example. With the sign choices $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,I} < 0$, $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,II} > 0$ and $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,III} < 0$, some particles leave well *I* and enter well *II* and *III*. For small collisionality, kinetic equation (7.2) imposes that the piece of the distribution function $K_{i,W}$ is continuous along particle trajectories. Since particles are leaving well *I* and moving into well *II* and well *III*, the distribution function in wells *II* and *III* must be the same as in well *I*, that is, $K_{i,II} = K_{i,I}$ and $K_{i,III} = K_{i,I}$. Thus, there is no need for a boundary layer because the distribution is continuous. Note that this process is irreversible: if we reverse time, particles leave wells *II* and *III* and move into well *I*, and we will see shortly that this leads, in principle, to discontinuities in the distribution function. In our equations, collisions are the only element that introduces irreversibility, so it is surprising that they do not seem to play a role in the discussion. In reality, collisions have led to this solution indirectly. If one assumes that, due to some unknown mechanism, more particles move into well *II* than into well *III*, giving $K_{i,II} > K_{i,III}$, the distribution function is discontinuous across the juncture. Fokker–Planck collisions impose continuity in the distribution function, and as a result we need a boundary layer around such a hypothetical juncture. Appendix H shows that it is impossible to construct such a boundary layer, a result that implies that the distribution function $K_{i,W}$ must be continuous across junctures where particles leave one well to move into the other two wells.

We proceed to consider a juncture in which particles leave two wells and enter into the third well. For example, consider the situation that arises if the signs of ϕ'_0 , Ψ'_t and the derivatives of $J_{c,I}$, $J_{c,II}$ and $J_{c,III}$ with respect to α imply that particles leave wells *I* and *II* to enter well *III*. In this case, in general, $K_{i,I}$ and $K_{i,II}$ are different and we need to obtain a value for $K_{i,III}$. One can think of this problem as particles flowing into the juncture from wells *I* and *II*, mixing inside the juncture, and leaving through well *III* with a new distribution function $K_{i,III}$. A collisional boundary layer, described in Appendix H, forms around any juncture in which particles move out of two wells and

into the third. The result of this boundary layer is simply a balance among the phase-space fluxes of particles across the juncture. The infinitesimal element of phase-space volume integrated over the gyrophase and l can be constructed from (2.18) and (4.13), and it is $(2\pi\Psi'_t v/B_0) dr d\alpha dv dJ$ to lowest order in ϵ . With this phase-space volume element, we can calculate the number of particles crossing $J = J_{c,w}$. One first replaces the variable J with $\Delta J_W := J - J_{c,w}$. With this new set of coordinates, the infinitesimal phase-space volume becomes $(2\pi\Psi'_t v/B_0) dr d\alpha dv d\Delta J_W$. Then, the rate at which phase-space volume crosses the boundary $J = J_{c,w}$, equivalent to $\Delta J_W = 0$, is

$$\frac{2\pi\Psi'_t v}{B_0} \frac{d\Delta J_W}{dt} dr d\alpha dv = -\frac{2\pi c\phi'_0 v}{B_0} \partial_\alpha J_{c,w} dr d\alpha dv, \tag{7.5}$$

where we have used (7.4) and the fact that the time derivative of J is zero to obtain $d\Delta J_W/dt \simeq -(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,w}$. Using (7.5), we find that the number of particles crossing the boundary is $-(2\pi c\phi'_0 v g_{i,1,w}^{(0)}/B_0) \partial_\alpha J_{c,w} dr d\alpha dv$. Balance between the three fluxes leaving and entering the juncture in figure 1 gives

$$\begin{aligned} & -\frac{2\pi c\phi'_0 v}{B_0} g_{i,1,I}^{(0)} \partial_\alpha J_{c,I} dr d\alpha dv - \frac{2\pi c\phi'_0 v}{B_0} g_{i,1,II}^{(0)} \partial_\alpha J_{c,II} dr d\alpha dv \\ & = -\frac{2\pi c\phi'_0 v}{B_0} g_{i,1,III}^{(0)} \partial_\alpha J_{c,III} dr d\alpha dv. \end{aligned} \tag{7.6}$$

We can simplify this expression by dividing by $-(2\pi c\phi'_0 v/B_0) dr d\alpha dv$. Furthermore, using (7.3), we find the relationship

$$\begin{aligned} & (K_{i,I} - r_{i,1,I} \Upsilon f_{Mi}) \partial_\alpha J_{c,I} + (K_{i,II} - r_{i,1,II} \Upsilon f_{Mi}) \partial_\alpha J_{c,II} \\ & = (K_{i,III} - r_{i,1,III} \Upsilon f_{Mi}) \partial_\alpha J_{c,III}. \end{aligned} \tag{7.7}$$

The definition $r_{i,1,w}$ in (5.26) implies that $r_{i,I} = r_{i,II} = r_{i,III}$ at the juncture because all the particles on the juncture have the same value of $\bar{\lambda} \simeq \lambda$. With this result and the fact that condition (5.27) gives $\partial_\alpha J_{c,I} + \partial_\alpha J_{c,II} = \partial_\alpha J_{c,III}$, all the terms that contain $r_{i,1,w}$ cancel each other, leaving

$$K_{i,I} \partial_\alpha J_{c,I} + K_{i,II} \partial_\alpha J_{c,II} = K_{i,III} \partial_\alpha J_{c,III}. \tag{7.8}$$

Equation (7.8) can be used to calculate $K_{i,III}$ given $K_{i,I}$ and $K_{i,II}$. Note that, due to condition (5.27), $K_{i,III}$ is equal to both $K_{i,I}$ and $K_{i,II}$ if $K_{i,I} = K_{i,II}$.

The discussion in this section is related to the probabilities of transition calculated by Cary *et al.* (1986). If instead of the particle distribution function, one considers individual particle motion across a juncture, the problem with junctures is reversed. When a particle leaves one well to move into two other wells, the particle cannot be split into two. The exact position of the trapped particle along its quasi-periodic motion (that is, its phase) when it reaches the juncture determines the well that it moves into. Orbit-averaged motion ignores this phase, but one can assume that particles are equally distributed along this phase and, using this assumption, obtain the probability of transitioning into one well or the other. Following Cary *et al.* (1986), if a particle leaves well I , $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,I} < 0$, to move into either well II , $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,II} > 0$, or well III , $(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,III} < 0$, the probability that it moves into well II is $-\partial_\alpha J_{c,II}/\partial_\alpha J_{c,I}$, whereas the probability that it transitions to well III is $\partial_\alpha J_{c,III}/\partial_\alpha J_{c,I}$. Thus, for single particle motion, the derivatives of $J_{c,w}$ with respect to α enter in the problem, but they apply to the junctures in which the particle leaves one well to move into one of the other two wells, instead of applying to the case in which particles leave two wells to transition into the third well. In this latter type of juncture,

a single particle leaving one of the wells does not have any other choice but to move into the well that is accepting particles. To summarize, our formulation does not require transition probabilities because it is based on the distribution function instead of single particle motion and because, by keeping collisions, it develops the collisional boundary layers described in [Appendix H](#).

7.1.2. Example

After explaining how the distribution function behaves around junctures of different types of wells, we proceed to argue that the function $K_{i,W}$ is independent of J in large regions of velocity space. To illustrate the problem, we consider the simple situation sketched in [figure 6](#). In this figure, we show the dependence of the effective potential $U \simeq \mu B + Z_i e \phi_0 / m_i$ on l for several values of α . We consider particle motion in this effective potential for a radial electric field that forces particles to move in the direction of increasing α , that is, $c\phi'_0 / \Psi'_t > 0$. Due to the changes in the U profile with α , we can divide the motion into four steps:

- (i) Step (a) \rightarrow (b). The minimum of well II decreases with α , while well I does not change. Thus, $\partial_\alpha J_{c,II} > 0$ and $\partial_\alpha J_{c,III} > 0$, but $\partial_\alpha J_{c,I} = 0$. Since $c\phi'_0 / \Psi'_t > 0$, the discussion of the previous section implies that particles in well III with low values of J transition into well II. Particles cannot transition into well I because $\partial_\alpha J_{c,I} = 0$.
- (ii) Step (b) \rightarrow (c). The minimum of well I decreases, while well II does not change. Hence, $\partial_\alpha J_{c,I} > 0$, $\partial_\alpha J_{c,II} = 0$ and $\partial_\alpha J_{c,III} > 0$, and particles in well III with low values of J transition into well I. Particles cannot transition into well II because $\partial_\alpha J_{c,II} = 0$.
- (iii) Step (c) \rightarrow (d). The minimum of well II increases until it reaches the value it had in (a), while well I does not change. As a result, $\partial_\alpha J_{c,I} = 0$, $\partial_\alpha J_{c,II} < 0$ and $\partial_\alpha J_{c,III} < 0$, and particles that had transitioned into well II during step (i) go back into well III, but their value of J is greater than their initial one. Particles cannot transition into well I because $\partial_\alpha J_{c,I} = 0$.
- (iv) Step (d) \rightarrow (e). The minimum of well I increases until it reaches the value it had in (a), while well II does not change. Then, $\partial_\alpha J_{c,I} < 0$, $\partial_\alpha J_{c,II} = 0$ and $\partial_\alpha J_{c,III} < 0$, and particles that had transitioned into well I during step (ii) go back into well III, but their value of J is lower than their initial one. Particles cannot transition into well II because $\partial_\alpha J_{c,II} = 0$.

In the configuration in [figure 6](#), there is only one juncture between wells I, II and III. This juncture is characterized by the functions $J_{c,I}(r, \alpha, v)$, $J_{c,II}(r, \alpha, v)$ and $J_{c,III}(r, \alpha, v)$, sketched in [figure 7\(b\)](#). Each of these functions has a maximum and a minimum in α that we denote $J_{c,W,M}(r, v)$ and $J_{c,W,m}(r, v)$, respectively. Particles in well I with values of J smaller than $J_{c,I,m}(r, v)$ never transition to wells II or III. Similarly, particles in well II with J smaller than $J_{c,II,m}(r, v)$ and particles in well III with J larger than $J_{c,III,M}(r, v)$ do not transition into the other two wells. In these regions of velocity space, the dependence of $K_{i,W}$ on J will be determined in § 7.2.

The situation is very different for particles in well III that have $J_{c,III,m}(r, v) < J < J_{c,III,M}(r, v)$. These particles transition into wells I and II and eventually transition back into well III. There is a value of J , $J_{I,II}(r, v) := J_{c,I,m}(r, v) + J_{c,II,M}(r, v)$, such that particles in well III with $J_{c,III,m} < J < J_{I,II}$ transition into well II, and particles with $J_{I,II} < J < J_{c,III,M}$ transition into well I.

To understand what happens for particles with $J_{c,III,m} < J < J_{c,III,M}$, we define the function $J_{III,out}(J_{III,in})$. This function determines the value of the second adiabatic invariant of a particle in well III after it has transitioned into well I or well II and has then

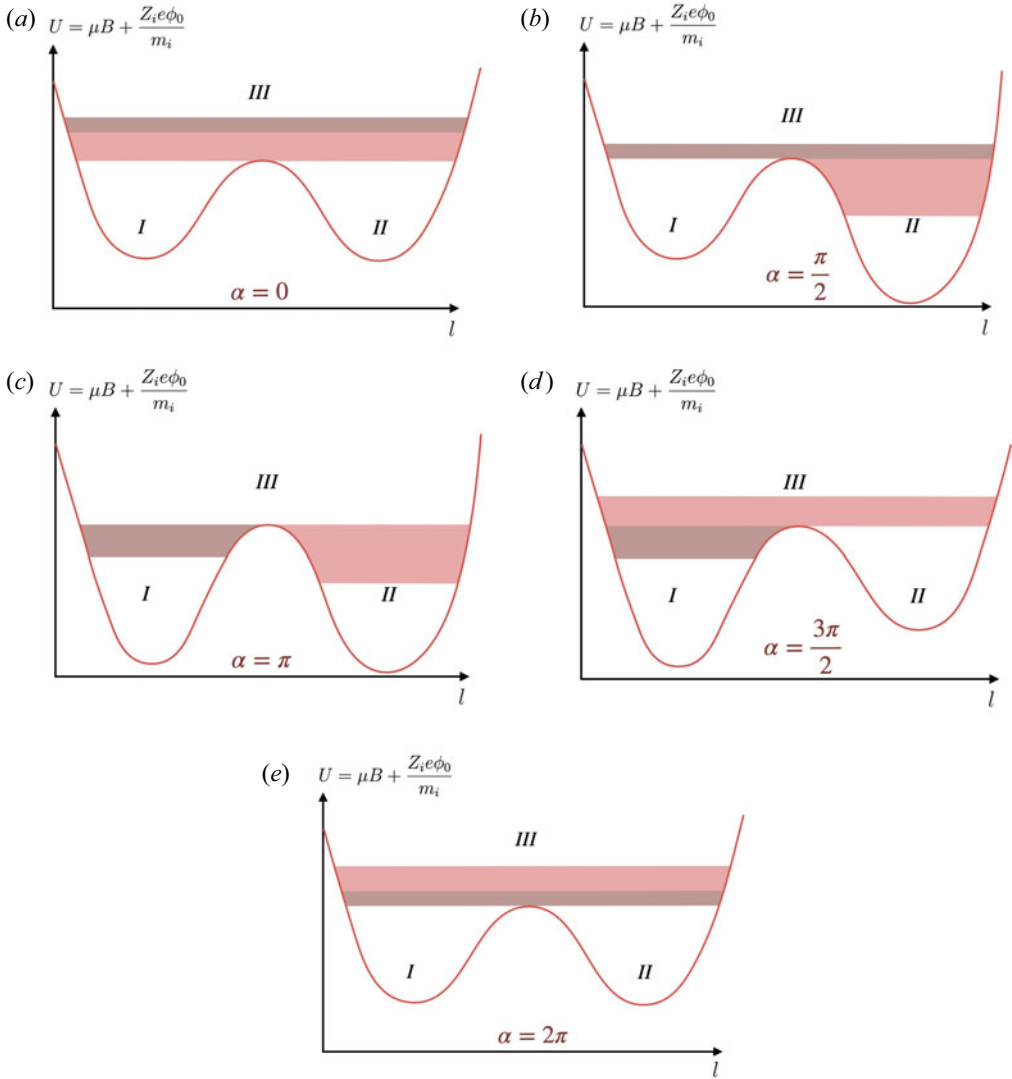


FIGURE 6. Example of effective potential $U(\alpha, l)$. Each panel corresponds to a particular value of α .

transitioned back to well III. The value $J_{III,out}$ of the adiabatic invariant after the particle has transition in and out of well I or well II depends only on the value that the adiabatic invariant had before the particle transitioned, $J_{III,in}$. For the magnetic field configuration in figure 6, the function $J_{III,out}(J_{III,in})$ is sketched in figure 7(a) as a light pink straight line for particles that transition into well II ($J_{c,III,m} < J < J_{I,II}$), and as a dark pink straight line for particles that transition into well I ($J_{I,II} < J < J_{c,III,M}$). To sketch the function $J_{III,out}(J_{III,in})$, we have used the fact that

$$J_{III,out}(J_{III,in}) = \begin{cases} J_{III,in} + J_{c,I,M} - J_{c,I,m}, & \text{for } J_{c,III,m} < J < J_{I,II}, \\ J_{III,in} - J_{c,II,M} + J_{c,II,m}, & \text{for } J_{I,II} < J < J_{c,III,M}. \end{cases} \quad (7.9)$$

Using the function $J_{III,out}(J_{III,in})$, we can schematically describe particle motion for particles with $J_{c,III,m} < J < J_{c,III,M}$. This motion is shown in figure 7(a) as a black

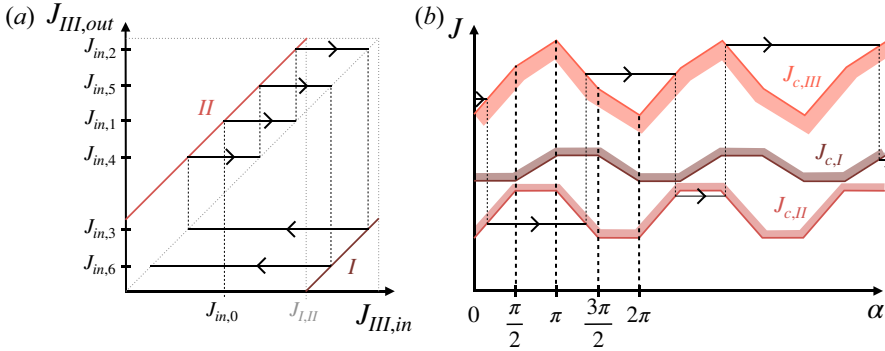


FIGURE 7. Evolution of the value of J of a particle in the magnetic field sketched in figure 6. We start with a particle in well III with initial value of the second adiabatic invariant $J = J_{in,0}$. Panel (a) shows the second adiabatic invariant of particles with $J = J_{III,in}$ after these particles has transition into well I (dark pink straight line) or well II (light pink straight line), and then back to well III. The trajectory of the particle with initial second adiabatic invariant $J_{in,0}$ is represented by the black staircase-like line. Panel (b) shows the initial part of the same trajectory in the plane (α, J) . The pink lines are $J_{c,I}(\alpha)$, $J_{c,II}(\alpha)$ and $J_{c,III}(\alpha)$.

staircase-like line with arrows. We start with a particle with initial second adiabatic invariant $J_{in,0} < J_{I,II}$. After its transitions in and out of well II, this particle has acquired the second adiabatic invariant $J_{III,out}(J_{in,0})$, which is necessarily in the interval $[J_{c,III,m}, J_{c,III,M}]$. As a result, the particle will transition again into well I or well II. This fact is shown in figure 7(a) by noting that $J_{III,out}(J_{in,0})$ becomes $J_{in,1}$, and this particle in turn will transition in and out of well II to acquire the second adiabatic invariant $J_{III,out}(J_{in,1}) = J_{III,out}(J_{III,out}(J_{in,0}))$. In figure 7(a), the fact that $J_{in,1} = J_{III,out}(J_{in,0})$ is represented as a horizontal solid line that joins $J_{III,out}(J_{in,0})$ with the diagonal line $J_{III,out} = J_{III,in}$, shown as a dotted line, and the fact that $J_{in,1}$ becomes $J_{III,out}(J_{in,1})$ is represented as a vertical dashed line from the diagonal to $J_{III,out}(J_{in,1})$. Thus, in the $J_{III,out}$ vs $J_{III,in}$ graph, particle trajectories are the staircase-like lines sketched in figure 7(a). The initial part of the same trajectory is sketched in figure 7(b) in the (α, J) plane.

Figure 7(a) shows that any particle that starts with $J_{c,III,m} < J < J_{c,III,M}$ ends up sampling all possible values of the second adiabatic invariant between $J_{c,III,m}$ and $J_{c,III,M}$ unless the function $J_{III,out}(J_{III,in})$ is very specific – a possible special case of this kind is represented in figure 8, where particle trajectories close on themselves in a loop in the (J, α) plane. We expect magnetic field magnitude profiles to satisfy the necessary constraints that lead to configurations similar to the one in figure 8 only in a countable number of flux surfaces. Thus, in general, particles in well III with J values between $J_{c,III,m}$ and $J_{c,III,M}$ sample all possible values of J between $J_{c,III,m}$ and $J_{c,III,M}$. Since $K_{i,W}$ is constant along particle trajectories (it is independent of α and it is continuous at the junctures of several wells because both $g_{i,1,W}$ and $r_{i,1,W}$ are continuous there), we find that $K_{i,III}$ does not depend on J for $J_{c,III,m} < J < J_{c,III,M}$. For the same reasons, $K_{i,I}$ in the interval $[J_{c,I,m}, J_{c,I,M}]$ and $K_{i,II}$ in the interval $[J_{c,II,m}, J_{c,II,M}]$ are both independent of J and equal to the value of $K_{i,III}$ in the interval $[J_{c,III,m}, J_{c,III,M}]$. In figure 9, we sketch $K_{i,W}$ for the configuration in figure 6.

7.1.3. General considerations

We have used the example in figure 6 to illustrate the dependence of $K_{i,W}$ on J because of its simplicity. In this particular configuration, particles do not have the option to choose

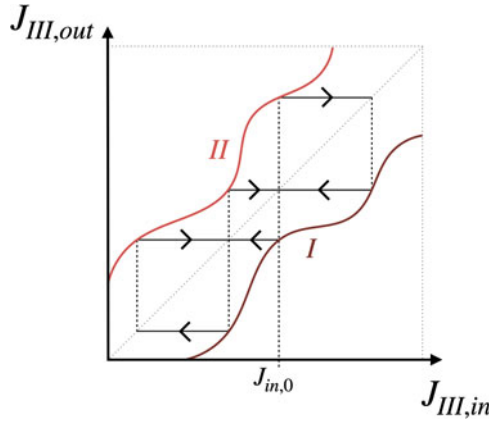


FIGURE 8. Function $J_{III,out}(J_{III,in})$ for which particle trajectories describe a loop and do not sample a finite interval of J .

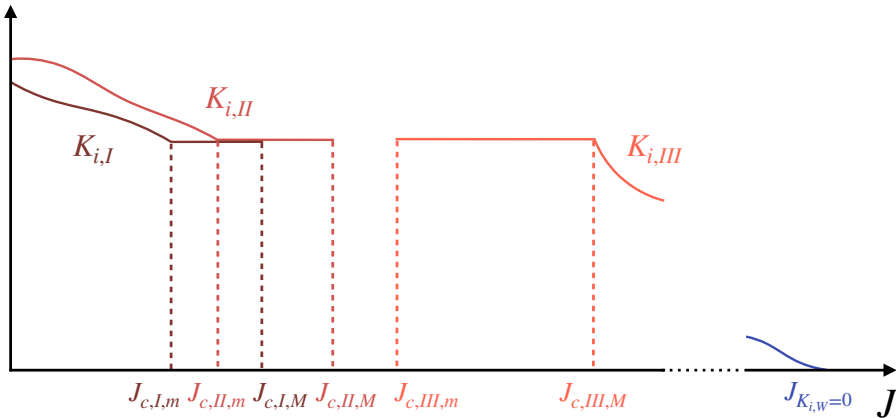


FIGURE 9. Sketch of the function $K_{i,w}$ for the case in figure 6.

between two wells, and particles from two different wells are not forced to move into the same well at the same time (this last case requires using (7.8)). Despite the simplicity of the example, we believe that only very specific configurations allow for a $K_{i,w}$ that is not constant in J in regions with junctures of different types of wells. We note that (7.8) accepts a constant $K_{i,w}$ as a solution, and our belief is that in most cases this is the only solution. Interestingly, this $K_{i,w}$ solution does not have discontinuities in $K_{i,w}$ across junctures, and hence does not have the collisional boundary layers described in Appendix H. These layers disappear only to lowest order, as we will see that they are necessary for the next-order correction in the $v_{i*}/\rho_{i*} \ll 1$ expansion.

For large values of J , there are always junctures of different types of wells (see figure 10 and the appendix of Boozer & Gardner 1990), so we expect $K_{i,w}$ to be independent of J above certain value of J . Moreover, since $r_{i,1,w} \rightarrow 0$ for $J \rightarrow \infty$ by definition (5.26), boundary condition (5.34) imposes that $K_{i,w} = 0$ in this region. From here on, the value of J at which $K_{i,w}$ vanishes is denoted by $J_{K_{i,w}=0}$ – see figure 9 for an example of $J_{K_{i,w}=0}$.

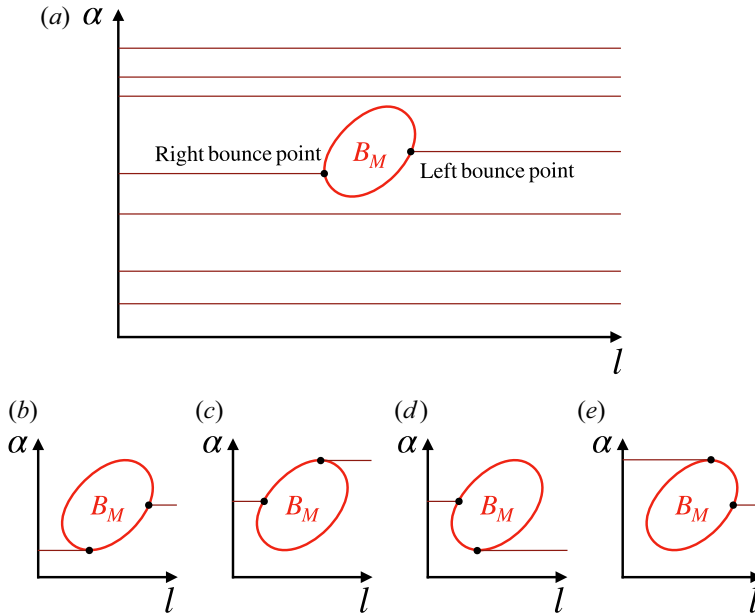


FIGURE 10. Sketch of trajectories with large J in the (l, α) plane. For large J , λ must be close to $1/B_M$, and hence it must have bounce points at a value of B close to B_M . In the figure, we sketch the contour $B = 1/\lambda \approx B_M$ as a thick red line (we have assumed that there is only one maximum of B). The total trajectory of the particle is sketched in panel (a) as a thin red line. The best way to identify a trajectory with a given λ is to determine the location of the bounce points (note that the left bounce point $l_{bL,W}$ is on the right of the figure and the right bounce point $l_{bR,W}$ is on the left). Particles with the kind of trajectory shown in panel (a) are exposed to four junctures with other wells, as demonstrated by panels (b–e). In panel (b), if the particle moves towards negative α , it will transition to another type of well, similarly to panel (d). In panels (c,e), the particle transitions when it moves towards positive α .

We finish by proving that the particle and energy fluxes due to $g_{i,1,W}^{(0)}$ vanish. From (5.48), we obtain that the flux of particles is

$$\Gamma_i \simeq \frac{2\pi c\phi'_0}{B_0 V'} \int_0^\infty dv \sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \int_{J_{m,W}}^{J_{M,W}} dJ v \left[\partial_\alpha (r_{i,1,W} K_{i,W}) - \partial_\alpha \left(\frac{1}{2} r_{i,1,W}^2 \Upsilon i f_{Mi} \right) \right]. \tag{7.10}$$

The particle flux vanishes due to the integral over α and the fact that $r_{i,1,W}$ and $K_{i,W}$ are continuous across junctures between several wells. A similar proof shows that the ion heat flux vanishes to lowest order in $v_{i*}/\rho_{i*} \ll 1$.

7.2. Next-order distribution function

To next order in $v_{i*}/\rho_{i*} \ll 1$, we would expect (5.25) to give

$$\frac{c\phi'_0}{\Psi'_i} \partial_\alpha g_{i,1,W}^{(1)} = \frac{v^2 v_{i,\perp}}{4} \partial_J (\tau_W J \partial_J g_{i,1,W}^{(0)}). \tag{7.11}$$

We explain at the end of this section that, in the intervals of J where there are junctures between different wells and hence $\partial_J K_{i,W} = 0$, (7.11) is not valid. Before discussing this

case, we consider the intervals of J where particles never transition into other wells. Then, (7.11) gives

$$\partial_\alpha g_{i,1,W}^{(1)} = \frac{\Psi'_t v^2 v_{ii,\perp}}{4c\phi'_0} \partial_J [\tau_W J (\partial_J K_{i,W} - \partial_J r_{i,1,W} \mathcal{Y}f_{Mi})]. \tag{7.12}$$

For these values of J , $g_{i,1,W}^{(1)}$ is defined for all α . For this reason, if we average equation (7.12) over α , the first term vanishes, leaving the equation for $K_{i,W}$

$$\partial_J [J (\langle \tau_W \rangle_\alpha \partial_J K_{i,W} - \langle \tau_W \partial_J r_{i,1,W} \rangle_\alpha \mathcal{Y}f_{Mi})] = 0, \tag{7.13}$$

where

$$\langle \dots \rangle_\alpha := \frac{1}{2\pi} \int_0^{2\pi} (\dots) d\alpha. \tag{7.14}$$

Solving (7.13) for $K_{i,W}$ allows one to integrate equation (7.12) with periodic boundary conditions in α .

Equation (7.13) determines the dependence of $K_{i,W}$ on J within the intervals of J where there are no junctures of different types of wells. If the interval of J considered includes $J = 0$, where $\tau_W = 0$, the integral of equation (7.13) is

$$\partial_J K_{i,W} = \frac{\langle \tau_W \partial_J r_{i,1,W} \rangle_\alpha}{\langle \tau_W \rangle_\alpha} \mathcal{Y}f_{Mi}. \tag{7.15}$$

In the example in figures 6 and 7, this solution is valid in well I for $J < J_{c,I,m}$, and in well II for $J < J_{c,II,m}$. Note that solution (7.15) leads to a discontinuity in $\partial_J K_{i,W}$ at $J_{c,I,m}$ and $J_{c,II,m}$ because intervals of J with $\partial_J K_{i,W} = 0$ start at these values of J . These discontinuities in the derivatives of $K_{i,W}$ mean that there are boundary layers where the lowest-order equation (7.2) is not valid and one needs to keep the collision operator to lowest order (see Appendix I.1 for a discussion of these boundary layers). More importantly, the finite values of $\partial_J K_{i,W}$ at $J = J_{c,I,m}$ and $J_{c,II,m}$ also mean that there is collisional flux into the region $J > J_{c,I,m}$ of well I and the region $J > J_{c,II,m}$ of well II . These phase-space fluxes are rapidly transported into well III by the $\mathbf{E} \times \mathbf{B}$ drift in the direction α and through jumps in J due transitions to other wells. This phase-space flux must then leave at $J_{c,III,M}$, where the region with $\partial_J K_{i,W} = 0$ ends. Conservation of collisional flux determines $\partial_J K_{i,III}$ at $J_{c,III,M}$,

$$\begin{aligned} & J_{c,I,m} \lim_{J \rightarrow J_{c,I,m}^-} (\langle \tau_I \rangle_\alpha \partial_J K_{i,I} - \langle \tau_I \partial_J r_{i,1,I} \rangle_\alpha \mathcal{Y}f_{Mi}) \\ & + J_{c,II,m} \lim_{J \rightarrow J_{c,II,m}^-} (\langle \tau_{II} \rangle_\alpha \partial_J K_{i,II} - \langle \tau_{II} \partial_J r_{i,1,II} \rangle_\alpha \mathcal{Y}f_{Mi}) \\ & = J_{c,III,M} \lim_{J \rightarrow J_{c,III,M}^+} (\langle \tau_{III} \rangle_\alpha \partial_J K_{i,III} - \langle \tau_{III} \partial_J r_{i,1,III} \rangle_\alpha \mathcal{Y}f_{Mi}). \end{aligned} \tag{7.16}$$

(See Appendix I.1 for a derivation of this condition using the equation for the higher-order correction $g_{i,1,W}^{(1)}$.) Since solution (7.15) is valid in wells I and II , condition (7.16) simply

becomes

$$\lim_{J \rightarrow J_{c,III,M}^+} \langle \tau_{III} \rangle_\alpha \partial_J K_{i,III} = \lim_{J \rightarrow J_{c,III,M}^+} \langle \tau_{III} \partial_J r_{i,1,III} \rangle_\alpha \Upsilon f_{Mi}. \tag{7.17}$$

With this condition, we can integrate equation (7.13) to obtain

$$\partial_J K_{i,III} = \frac{\langle \tau_{III} \partial_J r_{i,1,III} \rangle_\alpha}{\langle \tau_{III} \rangle_\alpha} \Upsilon f_{Mi}, \tag{7.18}$$

that is, solution (7.15) is valid in well III even though this well does not have particles with $J = 0$. Following this procedure for higher and higher values J , one can see that solution (7.15) is in fact valid in every well. With $\partial_J K_{i,W}$ calculated, we can integrate it to obtain $K_{i,W}$ imposing $K_{i,W} = 0$ at $J = J_{K_{i,W}=0}$.

The fact that $K_{i,W}$ vanishes for $J > J_{K_{i,W}=0}$ implies that the solution for $g_{i,1,W}^{(0)}$ for large values of J is simply

$$\begin{aligned} g_{i,1,W}^{(0)} &= -r_{i,1,W} \Upsilon f_{Mi} = -\frac{m_i v^2 B_0}{2Z_i e \phi'_0} \left(\lambda_W - \frac{1}{B_M} \right) \Upsilon f_{Mi} \\ &\simeq -\frac{m_i v^2 B_0}{2Z_i e \phi'_0} \left(\lambda_W - \frac{1}{B_0} + \frac{B_{1,M}}{B_0^2} \right) \Upsilon f_{Mi}, \end{aligned} \tag{7.19}$$

where we have substituted the definition of $r_{i,1,W}$ (see (5.26)) and we have used the approximation $\bar{\lambda}_W \simeq \lambda_W$. Surprisingly, this solution does not satisfy property (5.36) (note that $\partial_J \lambda_W \simeq -2/v^2 \tau_W B_0$) even though property (5.36) is a consequence of the equations for $g_{i,1,W}$. This apparent contradiction is resolved by a boundary layer where the lowest-order equation (7.2) is not valid because one needs to keep the collision operator to lowest order (see Appendix I.2 for a brief discussion of this boundary layer).

We finish our discussion of the next-order correction $g_{i,1,W}^{(1)}$ by considering its value in the regions where $\partial_J K_{i,W} = 0$. Naively, for this region of velocity space, (7.11) gives

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,1,W}^{(1)} = -\frac{v^2 v_{ii,\perp} \Upsilon f_{Mi}}{4} \partial_J (\tau_W J \partial_J r_{i,1,W}). \tag{7.20}$$

This equation has to be solved in regions where there are junctures of different types of well, and as a result, there can be discontinuities in $g_{i,1,W}^{(1)}$ across the junctures. Following the same procedure that we developed in Appendix H, we find that there are boundary layers in the junctures where particles leave two of the wells to go into the third, and that the result of these boundary layers is that $g_{i,1,W}^{(1)}$ must satisfy an equation similar to (7.8),

$$g_{i,1,I}^{(1)} \partial_\alpha J_{c,I} + g_{i,1,II}^{(1)} \partial_\alpha J_{c,II} = g_{i,1,III}^{(1)} \partial_\alpha J_{c,III}. \tag{7.21}$$

Despite their apparent validity, (7.20) and (7.21) cannot be solved. Using the definition of $r_{i,1,W}$ in (5.26) and $\partial_J \lambda_W \simeq -2/v^2 \tau_W B_0$, we find that (7.20) is

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,1,W}^{(1)} = \frac{m_i v^2 v_{ii,\perp}}{4Z_i e \phi'_0} \Upsilon f_{Mi}. \tag{7.22}$$

Then, $g_{i,1,W}^{(1)}$ increases with α , posing a problem of continuity: due to the ergodic nature of the trajectories in the regions of velocity space where $\partial_J K_{i,W} = 0$, if we calculate $g_{i,1,W}^{(1)}$ using (7.22), the values of $g_{i,1,W}^{(1)}$ at two contiguous values of J would in general be very

different because the lengths of the paths in α needed to reach these similar values of J starting from the same phase-space point are very different. Such large differences for contiguous values of J mean that we cannot neglect the collision operator in the regions where $\partial_J K_{i,W} = 0$, and instead of (7.22), we need to integrate

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,1,W}^{(1)} - \frac{v^2 v_{ii,\perp}}{4} \partial_J (\tau_W J \partial_J g_{i,1,W}^{(1)}) = \frac{m_i v^2 v_{ii,\perp}}{4 Z_i e \phi'_0} \Upsilon f_{Mi}. \tag{7.23}$$

For this equation to be valid, the characteristic size of ∂_J must be at least $\partial_J \sim \sqrt{\rho_{i*}/v_{i*}}/\sqrt{\epsilon} v_{ii} R \gg 1/\sqrt{\epsilon} v_{ii} R$, and hence we expect solution $g_{i,1,W}^{(1)}$ to have oscillatory character in J . Due to the inclusion of collisions, the solutions to (7.23) contain the boundary layers that appear around junctures of different types of wells, described in Appendix H. At the junctures of different wells, we need to apply continuity of $g_{i,1,W}^{(1)}$ and condition (5.33), which at this order is

$$J_{c,I} \lim_{J \rightarrow J_{c,I}} \tau_I \partial_J g_{i,1,I}^{(1)} + J_{c,II} \lim_{J \rightarrow J_{c,II}} \tau_{II} \partial_J g_{i,1,II}^{(1)} = J_{c,III} \lim_{J \rightarrow J_{c,III}} \tau_{III} \partial_J g_{i,1,III}^{(1)}. \tag{7.24}$$

We discuss the boundary conditions for $g_{i,1,W}^{(1)}$ in regions with $\partial_J K_{i,W} = 0$ in Appendix I.1.

7.3. Radial fluxes

In § 7.1 we showed that substituting $g_{i,1,W}^{(0)}$ into (5.48) for the particle flux gives zero. Thus, the particle flux is determined by $g_{i,1,W}^{(1)}$ and the boundary layers that we have described in Appendices H and I. To account for these higher-order effects, we manipulate (5.48): we integrate by parts in α and we use (5.25) to rewrite $\partial_\alpha g_{i,1,W}$ in terms of the collision operator and $r_{i,1,W}$. We find

$$\Gamma_i = -\frac{2\pi c\phi'_0}{B_0 V'} \int_0^\infty dv \sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \int_{J_{m,W}}^{J_{M,W}} dJ v \left[\frac{v^2 v_{ii,\perp} \Psi'_t}{4c\phi'_0} r_{i,1,W} \partial_J (\tau_W J \partial_J g_{i,1,W}) - \partial_\alpha \left(\frac{1}{2} r_{i,1,W}^2 \Upsilon f_{Mi} \right) \right]. \tag{7.25}$$

This expression is not useful because, inside the boundary layers described in Appendices H and I and in the regions of phase space where particles undergo transitions between different types of wells, the collision operator applied on $g_{i,1,W}$ gives large contributions to the integrand that vanish upon integration. To avoid this delicate cancellation, we integrate by parts in J to find

$$\Gamma_i = \frac{\pi \Psi'_t}{2B_0 V'} \int_0^\infty dv \sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \int_{J_{m,W}}^{J_{M,W}} dJ v^3 v_{ii,\perp} \tau_W J \partial_J r_{i,1,W} \partial_J g_{i,1,W}. \tag{7.26}$$

In this expression, we can neglect the higher-order corrections to $g_{i,1,W}^{(0)}$, and hence

$$\Gamma_i = -\frac{\pi \Psi'_t}{2B_0 V'} \int_0^\infty dv v^3 v_{ii,\perp} \Upsilon f_{Mi} \sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \left[\int_{\partial_J K_{i,W}=0} dJ \tau_W J (\partial_J r_{i,1,W})^2 + \int_{\partial_J K_{i,W} \neq 0} dJ \tau_W J \left((\partial_J r_{i,1,W})^2 - \frac{\partial_J r_{i,1,W}}{\langle \tau_W \rangle_\alpha} \langle \tau_W \partial_J r_{i,1,W} \rangle_\alpha \right) \right], \tag{7.27}$$

where we have used (7.3), and in the integral in J , we have distinguished the region where $\partial_J K_{i,W}$ vanishes from the region where it is equal to the value given in (7.15). Using

$\partial_J \lambda_W \simeq -2/v^2 B_0 \tau_W$ to write $\partial_J r_{i,1,W} \simeq -m_i/Z_i e \phi'_0 \tau_W$, the particle flux simplifies to

$$\Gamma_i = -\frac{\pi m_i^2 \Psi'_t}{2Z_i^2 e^2 \phi_0'^2 B_0 V'} \int_0^\infty dv v^3 v_{ii,\perp} \Upsilon_i f_{Mi} \sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \left[\int_{\partial_J K_{i,W}=0} dJ \frac{J}{\tau_W} + \int_{\partial_J K_{i,W} \neq 0} dJ J \left(\frac{1}{\tau_W} - \frac{1}{\langle \tau_W \rangle_\alpha} \right) \right]. \tag{7.28}$$

Note that the contribution from the particles in phase-space regions where $\partial_J K_{i,W} \neq 0$ is reduced by a factor of $(1 - \tau_W / \langle \tau_W \rangle_\alpha)$, that is, this expression shows that particles that suffer transitions between different types of wells cause higher fluxes.

A similar procedure to the one that we have shown above gives the energy flux

$$Q_i = -\frac{\pi m_i^3 \Psi'_t}{4Z_i^2 e^2 \phi_0'^2 B_0 V'} \int_0^\infty dv v^5 v_{ii,\perp} \Upsilon_i f_{Mi} \sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \left[\int_{\partial_J K_{i,W}=0} dJ \frac{J}{\tau_W} + \int_{\partial_J K_{i,W} \neq 0} dJ J \left(\frac{1}{\tau_W} - \frac{1}{\langle \tau_W \rangle_\alpha} \right) \right]. \tag{7.29}$$

The fluxes in (7.28) and (7.29) are inversely proportional to $\phi_0'^2$ and hence diverge when ϕ'_0 vanishes. The reason for this divergence is that the $\mathbf{E} \times \mathbf{B}$ drift was assumed to be much larger than the radial component of the drifts in (4.5). For $ea\phi'_0/T_i \sim \epsilon$, this assumption is not satisfied, and the first term in (7.3) becomes of the same order as the Maxwellian f_{Mi} , giving $g_{i,1,W}^{(0)} \sim f_{Mi}$, that is, the distribution function is not close to a Maxwellian. This is a manifestation of the fact that orbit widths become comparable to the minor radius of the stellarator. To summarize, formulas (7.28) and (7.29) are not valid for $ea\phi'_0/T_e \lesssim \epsilon$. At these values of the electric field, the radially global equations (2.1), (2.2), (2.27), (2.30) and (2.36) must be used. Importantly, even for $ea\phi'_0/T_e \sim 1$, energetic particles with energies larger than $eR\phi'_0 \sim m_i v_{ii}^2/\epsilon$ have ∇B and curvature drifts that are comparable to or larger than the $\mathbf{E} \times \mathbf{B}$ drift. This means that, unless the stellarator is close to omnigenity (Calvo *et al.* 2017; Catto 2019), energetic particles produced by fusion reactions, radiofrequency heating or neutral beams have to be modelled using the radially global equations (2.1), (2.2), (2.27), (2.30) and (2.36). These radially global orbit-averaged equations will be valid as long as the gyroradius and the banana orbit width of energetic particles are sufficiently small. In particular, the finite banana orbit width mechanism for energetic particle transport proposed by Goldston, White & Boozer (1981) is negligible in the limit in which the orbit-averaged equations (2.1), (2.2), (2.27), (2.30) and (2.36) are valid.

We finish by estimating the size of the particle and the energy flux. Using $\Psi'_t \sim aB_0$, $\phi'_0 \sim T_i/ea$, $V' \sim Ra$, $J \sim \sqrt{\epsilon} v_{ii} R$, $\tau_W \sim R/\sqrt{\epsilon} v_{ii}$ and $\Upsilon_i \sim 1/a$, we find

$$\Gamma_i \sim \epsilon^{5/2} v_{i*} n_i v_{ii}, \quad Q_i \sim \epsilon^{5/2} v_{i*} n_i T_i v_{ii}. \tag{7.30a,b}$$

These estimates can be obtained using a random walk estimate. As we explained in the introduction, in (1.5), the width of a typical trapped-particle orbit is $w \sim \epsilon a$. Since the effective collision frequency for trapped particles is $\nu_{\text{eff}} \sim v_{ii}/\epsilon$, trapped particles undergo a random walk with steps of length w every time ϵ/v_{ii} . The resulting diffusion coefficient is $D \sim \sqrt{\epsilon} w^2 \nu_{\text{eff}}$, where $\sqrt{\epsilon}$ is an estimate for the trapped-particle fraction. The estimates for the fluxes are then obtained from $D \sim \epsilon^{3/2} a^2 v_{ii}$, $\Gamma_i \sim D |\nabla n_i| \sim D n_i/a$ and $Q_i \sim n_i D |\nabla T_i| \sim D n_i T_i/a$.

8. Large aspect ratio stellarators close to omnigenicity

In this section we study large aspect ratio stellarators close to omnigenicity. A magnetic field is omnigenous when it satisfies $\partial_\alpha J_W = 0$ for all trapped particles, that is,

$$\partial_\alpha J_W = -v \int_{l_{bL,W}}^{l_{bR,W}} \frac{\lambda \partial_\alpha B}{\sqrt{1 - \lambda B}} dl = 0, \quad (8.1)$$

for all λ . As explained by Cary & Shasharina (1997a,b) and Parra *et al.* (2015), this condition imposes stringent constraints on how the magnitude B depends on α and l . Of these constraints, two are particularly important for our discussion.

- (i) The maxima of B on a flux surface are not on isolated points, but on lines that close on themselves. Moreover, those lines where the maxima lie are separated from each other by such a distance that any particle that travels between these maxima has only one possible value of J , which we denote by $J_{\text{om},M}$ because it is also the maximum value that J can take.
- (ii) In junctures of several types of wells like the one depicted in figure 1 (see Parra *et al.* (2015) for a discussion of the existence of omnigenous magnetic fields with junctures of different types of wells), the values of the second adiabatic invariant at the juncture are independent of α , that is, $\partial_\alpha J_{c,W} = 0$. This means that there are no transitions from one type of well to another.

Collisional transport in omnigenous stellarators is very small, and for this reason designing stellarators close to omnigenicity is of interest. We consider large aspect ratio stellarators that are close to omnigenicity. We find two distinct types. In § 8.1, we show that large aspect ratio stellarators with large mirror ratios are close to omnigenicity, and hence one can use the equations derived by Calvo *et al.* (2017) to calculate neoclassical fluxes in these stellarators. In § 8.2, we study optimized large aspect ratio stellarators with mirror ratios close to unity.

8.1. Large aspect ratio stellarators with large mirror ratios

In previous sections, we have discussed in detail large aspect ratio stellarators with mirror ratios close to unity, an approximation that describes well many modern stellarators. However, it is possible that increasing the mirror ratio will be of interest in the future. For this reason, we consider large aspect ratio stellarators with $B_0(l)$ a general function of l . Note that such stellarators satisfy (8.1) to lowest order because $B_0(l)$ is independent of α . This means that $B_1(r, \alpha, l)$ can be considered to be the deviation from omnigenicity. We can then use the formulation by Calvo *et al.* (2017) by replacing the expansion parameter δ , which measures the size of the deviation from omnigenicity, with ϵ .

Using the same techniques as Calvo *et al.* (2017), it is possible to prove that the lowest-order solution for the ion distribution function is a stationary Maxwellian f_{Mi} with density and temperature that are flux functions, and that the potential is a flux function to lowest order, $\phi(r, \alpha, l, t) \simeq \phi_0(r, t) \sim T_i/e$. The correction to the Maxwellian is only significantly different from zero for trapped particles, for which it satisfies $g_{i,1,W} \sim \epsilon f_{Mi}$. As for large aspect ratio stellarators with mirror ratios close to unity, the typical radial width of a trapped-particle orbit is ϵa because the radial magnetic drift is small by ϵ compared with the component of the $\mathbf{E} \times \mathbf{B}$ drift parallel to the flux surface. Apart from these similarities, the formulation by Calvo *et al.* (2017) differs from the equations presented in previous sections of this paper in significant ways. For $B_{0,M}/B_{0,m} - 1 \sim 1$, the typical parallel velocity of trapped particles is not small compared with the thermal speed, and the fraction

of trapped particles is of order unity. The energy gained or lost by work done by the radial electric field during a trapped particle’s radial displacement is insufficient to affect trapped particles with $v_{\parallel} \sim v_{ii}$, unlike in large aspect ratio stellarators with mirror ratios close to unity, where trapped particles have $v_{\parallel} \sim \sqrt{\epsilon}v_{ii}$. For this reason, the velocity coordinates \mathcal{E} and μ are appropriate – in Calvo *et al.* (2017), these coordinates are replaced by the equivalent coordinates v and $\lambda := v_{\perp}^2/v^2B$ for convenience. Another important difference with large aspect ratio stellarators with mirror ratios of order unity is that the correction to the lowest-order potential is $\phi_1(r, \alpha, l, t) \sim \epsilon T_i/e$ instead of $\phi_{3/2}(r, \alpha, l, t) \sim \epsilon^{3/2}T_i/e$, and as a result the radial component of the $\mathbf{E} \times \mathbf{B}$ drift contributes significantly to the radial displacement of trapped particles.

We proceed to discuss the different transport regimes predicted by the formulation of Calvo *et al.* (2017). For collision frequencies ν_{ii} larger than the characteristic frequency associated with the $\mathbf{E} \times \mathbf{B}$ drift, $|\mathbf{v}_E|/a \sim \rho_{i*}v_{ii}/a$, that is, for $\rho_{i*}/\epsilon \ll \nu_{i*} \ll \epsilon^{3/2}$, the equations derived by Calvo *et al.* (2017) predict a $1/\nu$ regime. In this regime, the particle and energy fluxes are of order

$$\Gamma_i \sim \frac{\rho_{i*}}{v_{i*}} \epsilon \rho_{i*} n_i v_{ii}, \quad Q_i \sim \frac{\rho_{i*}}{v_{i*}} \epsilon \rho_{i*} n_i T_i v_{ii}. \tag{8.2a,b}$$

Here, the characteristic length L_0 and the parameter δ in Calvo *et al.* (2017) have been replaced with a and ϵ , and our definition of ν_{i*} differs from the one used by Calvo *et al.* (2017), ν_{i*}^{Ca} , by a factor of ϵ , $\nu_{i*} = \nu_{i*}^{\text{Ca}}/\epsilon$. Note that, for large aspect stellarators with large mirror ratios, the transition to the lower collisionality transport regime happens for $\nu_{i*} \sim \rho_{i*}/\epsilon$ – in comparison, for large aspect ratio stellarators with mirror ratios close to unity, the transition happens for $\nu_{i*} \sim \rho_{i*}$. The difference is due to the typical parallel velocity of the trapped particles: for large aspect ratio stellarators with mirror ratios close to unity, trapped particles have $v_{\parallel} \sim \sqrt{\epsilon}v_{ii}$ and hence it is easy for collisions to detrap these particles, giving a large effective collision frequency ν_{ii}/ϵ for trapped particles.

The estimates for the fluxes in (8.2a,b) can also be obtained from a random walk argument similar to the one that gives the estimates in (6.2a,b). In large aspect ratio stellarators with large mirror ratios, the effective collision frequency for trapped particles is ν_{ii} and hence the random walk is composed of steps of length $\mathbf{v}_{Mi} \cdot \nabla r/v_{ii} \sim (\rho_{i*}/v_{i*})a \ll \epsilon a$ that happen with frequency ν_{ii} . The corresponding diffusion coefficient is $D \sim (\mathbf{v}_{Mi} \cdot \nabla r/v_{ii})^2 \nu_{ii} \sim \epsilon \rho_{i*}^2 a v_{ii}/v_{i*}$, where we have noted that the fraction of trapped particles is of order unity. This diffusion coefficient gives the estimates in (8.2a,b).

For $\nu_{i*} \ll \rho_{i*}/\epsilon$, the derivation of Calvo *et al.* (2017) does not assume that the component of the ∇B and curvature drifts parallel to the flux surface is smaller than the same component of the $\mathbf{E} \times \mathbf{B}$ drift. For this reason, it allows for cancellation of the average drift parallel to the flux surface for certain particles. In large aspect ratio stellarators with large mirror ratios and an electric field $|\phi'_0|$ that is much larger than T_i/eR , the $\mathbf{E} \times \mathbf{B}$ drift is larger than the magnetic drifts by $e|\phi'_0|R/T_i \gg 1$, and hence only a exponentially small number of particles with energies above $eR\phi'_0 \gg m_i v_{ii}^2$ can have vanishing drifts parallel to the flux surface. These particles can be neglected, and as a result the transport is dominated by boundary layers in velocity space of width $\Delta v_{\parallel} \ll v_{ii}$ that appear at the trapped–passing boundary or around junctures of different types of wells. In these layers, the effective collision frequency is

$$\nu_{\text{eff}} \sim \left(\frac{v_{ii}}{\Delta v_{\parallel}} \right)^2 \frac{v_{ii}}{\ln(v_{ii}/\Delta v_{\parallel})}. \tag{8.3}$$

The effective collision frequency becomes larger than v_{ii} because very small angle collisions are sufficient to change the parallel velocity by a small amount $\Delta v_{\parallel} \ll v_{ii}$. The logarithmic reduction of the effective collision frequency is due to the fact that particles in these layers have bounce points that are very close to local maxima of B , where they spend logarithmically long times. Collisions near these maxima of B are ineffective because v_{\parallel} is very close to zero, $v_{\parallel} \ll v_{ii}$, and a collision that changes the parallel velocity of a particle by Δv_{\parallel} changes its parallel kinetic energy by $m_i v_{\parallel} \Delta v_{\parallel} \ll m_i v_{ii} \Delta v_{\parallel}$ only. With such a small change in kinetic energy, the particle cannot leave the layer. The collisions that expel the particles from the boundary layer happen away from the bounce points, where $v_{\parallel} \sim v_{ii}$ and particles spend only a fraction of its bounce period of order $1/\ln(v_{ii}/\Delta v_{\parallel})$. The effective collision frequency ν_{eff} has to be comparable to the inverse of the time that it takes a particle to drift poloidally around the stellarator, $|v_E|/a \sim \rho_{i*} v_{ii}/a$, giving

$$\Delta v_{\parallel} \sim v_{ii} \sqrt{\frac{\epsilon v_{i*}}{\rho_{i*}} \frac{1}{\ln(\rho_{i*}/\epsilon v_{i*})}}. \tag{8.4}$$

The transport due to these boundary layers is of order

$$\Gamma_i \sim \sqrt{\frac{\epsilon v_{i*}}{\rho_{i*}} \ln\left(\frac{\rho_{i*}}{\epsilon v_{i*}}\right)} \epsilon^2 \rho_{i*} n_i v_{ii}, \quad Q_i \sim \sqrt{\frac{\epsilon v_{i*}}{\rho_{i*}} \ln\left(\frac{\rho_{i*}}{\epsilon v_{i*}}\right)} \epsilon^2 \rho_{i*} n_i T_i v_{ii}. \tag{8.5a,b}$$

The formulation by Calvo *et al.* (2017) does not work for $v_{ii} a/|v_E| \lesssim \epsilon^2 |\ln \epsilon|$ – this limit is derived in § 6 of Calvo *et al.* (2017) without the logarithmic correction (recall that δ must be replaced with ϵ). This means that, in normalized quantities, the \sqrt{v} regime does not extend below $\nu_{i*} \sim \epsilon |\ln \epsilon| \rho_{i*}$. We leave the study of what happens for collisionalities lower than this limit for another publication.

As other estimates given in previous sections, the estimates in (8.5a,b) can be obtained from a random walk argument. Only particles in velocity-space boundary layers with width Δv_{\parallel} , given in (8.4), participate in the transport. These particles have an effective collision frequency $\nu_{\text{eff}} \sim |v_E|/a \sim \rho_{i*} v_{ii}/a$, and hence typical radial displacements $\mathbf{v}_{Mi} \cdot \nabla r/\nu_{\text{eff}} \sim \epsilon a$. Thus, the diffusion coefficient is

$$D \sim \frac{\Delta v_{\parallel}}{v_{ii}} \ln\left(\frac{v_{ii}}{\Delta v_{\parallel}}\right) \left(\frac{\mathbf{v}_{Mi} \cdot \nabla r}{\nu_{\text{eff}}}\right)^2 \nu_{\text{eff}}, \tag{8.6}$$

where the prefactor $(\Delta v_{\parallel}/v_{ii}) \ln(v_{ii}/\Delta v_{\parallel})$ is due to the small fraction of particles that are in the boundary layers and cause transport. The logarithmic correction in $(\Delta v_{\parallel}/v_{ii}) \ln(v_{ii}/\Delta v_{\parallel})$ is due to the long bounce periods of the orbits of interest: the particles in the boundary layers spend logarithmically long times around local maxima of B , where they accumulate, giving a density large by $\ln(v_{ii}/\Delta v_{\parallel})$. The diffusion coefficient in (8.6) gives $D \sim \sqrt{(\epsilon v_{i*}/\rho_{i*}) \ln(\rho_{i*}/\epsilon v_{i*})} \epsilon^2 \rho_{i*} a v_{ii}$, and with it we can obtain the estimates in (8.5a,b).

A comparison between expressions (8.2a,b) and (8.5a,b) and expressions (6.2a,b) and (7.30a,b), sketched in figure 11, reveals that neoclassical fluxes in large aspect ratio stellarators with large mirror ratios are larger than the fluxes in equivalent large aspect ratio stellarator with mirror ratios close to unity. This is perhaps part of the reason why no stellarators with large mirror ratios have been built.

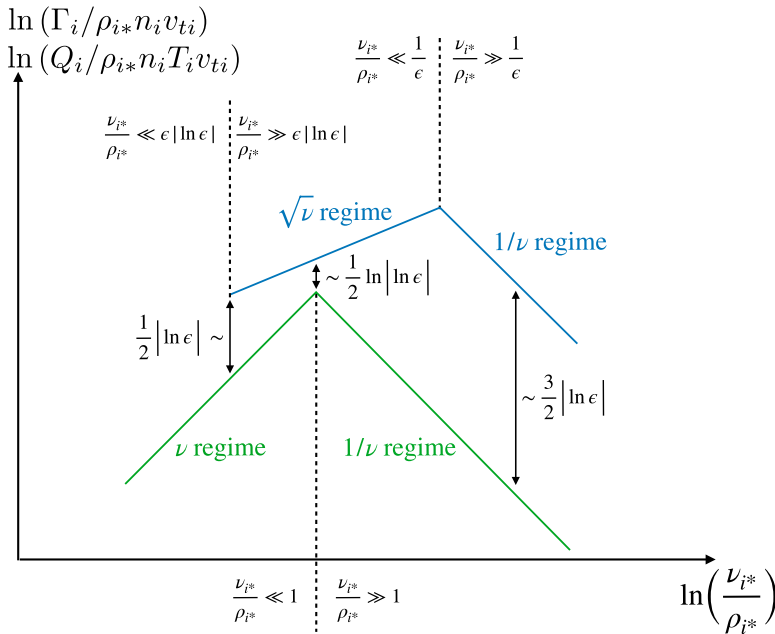


FIGURE 11. Particle and energy fluxes as a function of v_{i^*}/ρ_{i^*} for large aspect ratio stellarators with large mirror ratios (blue line) and mirror ratios close to unity (green line). Note that, for $v_{i^*} \sim \rho_{i^*}$, the difference between the fluxes of large aspect ratio stellarators with large mirror ratios and the fluxes of large aspect ratio stellarators with mirror ratios close to unity is only a factor of $\sqrt{|\ln \epsilon|}$.

8.2. Optimized large aspect ratio stellarators with mirror ratios close to unity

It is possible to optimize large aspect ratio stellarators with mirror ratios close to unity to achieve lower neoclassical transport. We treat such cases by using the expansion proposed by Calvo *et al.* (2018), that is, we consider a magnetic field that can be written as

$$B(r, \alpha, l) = B_0 + B_1^{[0]}(r, \alpha, l) + B_1^{[1]}(r, \alpha, l) + \dots \tag{8.7}$$

Here, $B_0 + B_1^{[0]}$, with $B_1^{[0]} \sim \epsilon B_0$, is omnigenous, that is,

$$\int_{l_{bL,0,W}^{[0]}}^{l_{bR,0,W}^{[0]}} \frac{\partial_\alpha B_1^{[0]}}{\sqrt{1 - \lambda B_0 - B_1^{[0]}/B_0}} dl = 0, \tag{8.8}$$

where we have used the large aspect ratio approximation to (8.1). Here, $l_{bL,0,W}^{[0]}(r, \alpha, \lambda)$ and $l_{bR,0,W}^{[0]}(r, \alpha, \lambda)$ are the bounce points that correspond to the omnigenous magnetic field, determined by the equations $B_1^{[0]}(r, \alpha, l_{bL,0,W}^{[0]})/B_0 = 1 - \lambda B_0 = B_1^{[0]}(r, \alpha, l_{bR,0,W}^{[0]})/B_0$. The correction $B_1^{[1]} \sim \delta \epsilon B_0$ is the perturbation away from omnigenicity. We assume that $\delta \ll 1$. Note that we are using the superscripts with square brackets $^{[n]}$ for the expansion in δ to distinguish it from the expansion in § 7.

We proceed to expand (5.25) in $\delta \ll 1$. We first obtain $r_{i,1,W}$ by writing

$$J_W(r, \alpha, v, \lambda) \simeq J_W^{[0]}(r, v, \lambda) + J_W^{[1]}(r, \alpha, v, \lambda) + \dots, \tag{8.9}$$

with

$$J_W^{[0]}(r, v, \lambda) := 2v \int_{l_{bL,0,W}^{[0]}}^{l_{bR,0,W}^{[0]}} \sqrt{1 - \lambda B_0 - B_1^{[0]}/B_0} dl \sim \sqrt{\epsilon} v_{ii} R \tag{8.10}$$

and

$$J_W^{[1]}(r, \alpha, v, \lambda) := -\frac{v}{B_0} \int_{l_{bL,0,W}^{[0]}}^{l_{bR,0,W}^{[0]}} \frac{B_1^{[1]}}{\sqrt{1 - \lambda B_0 - B_1^{[0]}/B_0}} dl \sim \delta \sqrt{\epsilon} v_{ii} R. \tag{8.11}$$

When we invert the relation $J_W(r, \alpha, v, \lambda)$ to obtain $\lambda_W(r, \alpha, v, J)$, we find that λ_W does not depend on α to lowest order in the expansion in $\delta \ll 1$. If we continue to next order in δ , we obtain

$$\lambda_W(r, \alpha, v, J) = \lambda_W^{[0]}(r, v, J) + \lambda_W^{[1]}(r, \alpha, v, J) + \dots, \tag{8.12}$$

where

$$\lambda_W^{[1]}(r, \alpha, v, J) := -\frac{J_W^{[1]}}{\partial_\lambda J_W^{[0]}} = \frac{2J_W^{[1]}}{v^2 B_0 \tau_W^{[0]}} \sim \frac{\delta \epsilon}{B_0}. \tag{8.13}$$

Here,

$$\tau_W^{[0]}(r, v, \lambda) := \frac{2}{v} \int_{l_{bL,0,W}^{[0]}}^{l_{bR,0,W}^{[0]}} \frac{1}{\sqrt{1 - \lambda B_0 - B_1^{[0]}/B_0}} dl \tag{8.14}$$

is the lowest-order bounce time.

With the result in (8.12), (5.25) becomes

$$\frac{c\phi'_0}{\Psi'_I} \partial_\alpha g_{i,1,W}^{[1]} - \frac{v^2 v_{ii,\perp}}{4} \partial_J (\tau_W^{[0]} J \partial_J g_{i,1,W}^{[1]}) = -\frac{c\phi'_0}{\Psi'_I} \partial_\alpha r_{i,1,W}^{[1]} \mathcal{R} f_{Mi}, \tag{8.15}$$

where $g_{i,1,W}^{[1]} \sim \delta \epsilon f_{Mi}$ and

$$r_{i,1,W}^{[1]}(r, \alpha, v, J) := \frac{m_i J_W^{[1]}}{Z_i e \phi'_0 \tau_W^{[0]}} \sim \delta \epsilon a. \tag{8.16}$$

At the junctures of several wells, we need to impose condition (5.33), which in the expansion in δ becomes

$$J_{c,I}^{[0]} \lim_{J \rightarrow J_{c,I}^{[0]}} \tau_I^{[0]} \partial_J g_{i,1,I}^{[1]} + J_{c,II}^{[0]} \lim_{J \rightarrow J_{c,II}^{[0]}} \tau_{II}^{[0]} \partial_J g_{i,1,II}^{[1]} = J_{c,III}^{[0]} \lim_{J \rightarrow J_{c,III}^{[0]}} \tau_{III}^{[0]} \partial_J g_{i,1,III}^{[1]}. \tag{8.17}$$

Importantly, since the omnigenous stellarator has a maximum value of J , $J_{om,M}$, the boundary condition for $g_{i,1,W}^{[1]}$ is not imposed at $J \rightarrow \infty$, but at $J = J_{om,M}^{[0]}$,

$$g_{i,1,W}^{[1]}(r, \alpha, v, J_{om,M}^{[0]}, t) = 0. \tag{8.18}$$

The particle and energy fluxes simplify to

$$\Gamma_i = \frac{2\pi c\phi'_0}{B_0 V'} \int_0^\infty dv \sum_W \int_0^{2\pi} d\alpha \int_{J_{m,W}}^{J_{M,W}} dJ v g_{i,1,W}^{[1]} \partial_\alpha r_{i,1,W}^{[1]} \sim \delta^2 \epsilon^{5/2} \rho_{i*} n_i v_{ii}, \tag{8.19}$$

and

$$Q_i = \frac{\pi m_i c \phi'_0}{B_0 V'} \int_0^\infty dv \sum_w \int_0^{2\pi} d\alpha \int_{J_{m,w}}^{J_{M,w}} dJ v^3 g_{i,1,W}^{[1]} \partial_\alpha r_{i,1,W}^{[1]} \sim \delta^2 \epsilon^{5/2} \rho_{i*} n_i T_i v_{ti}. \quad (8.20)$$

Note that we have used the fact that there are no transitions between wells in an omnigenous magnetic field to extend the integrals in α to the whole interval $[0, 2\pi]$.

Comparing the estimates for the fluxes in (8.19) and (8.20) with the ones for a generic large aspect ratio stellarator with mirror ratio close to unity in (5.48) and (5.50), we see that a near-omnigenous stellarator does indeed confine better than a generic one by a factor of $\delta^2 \ll 1$. This factor, however, is not uniform for all values of v_{i*} . While in the $1/v$ regime, $v_{i*}/\rho_{i*} \gg 1$, we can follow the arguments in § 6 to obtain

$$\Gamma_i \sim \delta^2 \frac{\rho_{i*}}{v_{i*}} \epsilon^{5/2} \rho_{i*} n_i v_{ti}, \quad Q_i \sim \delta^2 \frac{\rho_{i*}}{v_{i*}} \epsilon^{5/2} \rho_{i*} n_i T_i v_{ti}, \quad (8.21a,b)$$

for $v_{i*} \ll \rho_{i*}$ the scaling with δ is not a simple δ^2 , as we explain below. Before considering the regime $v_{i*} \ll \rho_{i*}$, we discuss how to obtain the estimates in (8.19), (8.20) and (8.21a,b) using a random walk argument. For $v_{i*} \gtrsim \rho_{i*}$, in between collisions, trapped particles move a radial distance $\langle \mathbf{v}_{Mi} \cdot \nabla r \rangle_{\tau,W} / \nu_{\text{eff}}$, where $\nu_{\text{eff}} \sim \nu_{ii}/\epsilon$ is the effective collision frequency for trapped particles. Since $\langle \mathbf{v}_{Mi} \cdot \nabla r \rangle_{\tau,W} \sim \delta \epsilon \rho_{i*} v_{ti}$, the typical radial displacement is $\langle \mathbf{v}_{Mi} \cdot \nabla r \rangle_{\tau,W} / \nu_{\text{eff}} \sim (\rho_{i*}/v_{i*}) \delta \epsilon a \lesssim \delta \epsilon a$. With this estimate, we obtain a random walk diffusion coefficient $D \sim \sqrt{\epsilon} (\langle \mathbf{v}_{Mi} \cdot \nabla r \rangle_{\tau,W} / \nu_{\text{eff}})^2 \nu_{\text{eff}}$, where the factor of $\sqrt{\epsilon}$ is due to the trapped-particle fraction. This diffusion coefficient $D \sim \delta^2 \epsilon^{5/2} \rho_{i*}^2 a v_{ti} / v_{i*}$ gives the estimates (8.21a,b) for $v_{i*} \gg \rho_{i*}$, and the estimates in (8.19) and (8.20) for $v_{i*} \sim \rho_{i*}$.

If we try to solve (8.15) for $v_{i*} \ll \rho_{i*}$ following the procedure in § 7, we run into problems at the largest value of J , $J_{\text{om},M}^{[0]}$, and at junctures of different types of wells. Indeed, if we neglect the collision operator in (8.15), the solution is

$$g_{i,1,W}^{[1]}(r, \alpha, v, J, t) = K_{i,W}^{[1]}(r, v, J, t) - r_{i,1,W}^{[1]} \Upsilon_{iM_i}, \quad (8.22)$$

where the function $K_{i,W}^{[1]}(r, v, J, t)$ does not depend on α . Solution (8.22) cannot vanish at $J = J_{\text{om},M}^{[0]}$ in general because $r_{i,1,W}^{[1]}$ does not vanish there. The contribution proportional to $r_{i,1,W}^{[1]}$ cannot be cancelled by a judicious choice of $K_{i,W}^{[1]}$ because $r_{i,1,W}^{[1]}$ depends on α while $K_{i,W}^{[1]}$ does not. Similarly, it is not possible to impose continuity of $g_{i,1,W}^{[1]}$ across junctures. The problem at $J = J_{\text{om},M}^{[0]}$ is a manifestation of the fact that there are orbits with second adiabatic invariants larger than $J_{\text{om},M}^{[0]}$. The perturbation $B_1^{[1]}$ will in general introduce isolated local maxima in B above those of the omnigenous magnetic field, and these maxima will not be equal to each other. The fact that there are individual maxima means that we are back to the case that we have studied in §§ 5 and 7. Thus, one needs to include $J > J_{\text{om},M}^{[0]}$. The problem at junctures of different types of wells is due to the fact that approximation (8.15) does not allow transitions between different types of wells.

We proceed to study the distribution functions and fluxes for $v_{i*} \ll \rho_{i*}$. There are two limits of interest that we consider: $\delta^2 |\ln \delta| \ll v_{i*}/\rho_{i*} \ll 1$ and $v_{i*}/\rho_{i*} \ll \delta^2 |\ln \delta|$.

8.2.1. The \sqrt{v} regime in optimized large aspect ratio stellarators with mirror ratios close to unity

We start by considering the regime $\delta^2 |\ln \delta| \ll v_{i*}/\rho_{i*} \ll 1$, where we will find the \sqrt{v} regime (Galeev *et al.* 1969; Ho & Kulsrud 1987). In the region $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ti} R$,

particles have bounce points at the new maxima introduced by the perturbation $B_1^{[1]}$ above the maxima of the omnigenous magnetic field. To understand estimate $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$, note that the particles that bounce in the new maxima of B are in an interval of λ of order $\Delta \lambda \sim \delta \epsilon / B_0$. This interval corresponds to an interval in J of order $\Delta \lambda \partial_\lambda J_W$, where $\partial_\lambda J_W \sim \epsilon^{-1/2} v_{ii} R B_0 |\ln \delta|$. The logarithm is due to the fact that the bounce points are always a small distance $\sqrt{\delta} R$ away from a local maximum of B . Particles spend logarithmically long times near these local maxima, giving $\partial_\lambda J_W \simeq -v^2 B_0 \tau_W / 2 \sim \epsilon^{-1/2} v_{ii} B_0 R |\ln \delta|$.

To determine whether the particles in the region $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$ can make up for the difference between $g_{i,1,W}^{[1]}$ and the distribution function in the passing-particle region, we estimate the size of the change in the distribution function $\Delta g_{i,1,\text{om}}$ for $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$. For $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$, we find

$$\partial_J g_{i,1,W} \sim \frac{\Delta g_{i,1,\text{om}}}{\delta |\ln \delta| \sqrt{\epsilon} v_{ii} R}. \tag{8.23}$$

Using this estimate in (5.25), we find that the dominant balance is between the collision operator and the term proportional to $\partial_\alpha r_{i,1,W}$, giving

$$\frac{v_{ii}}{\delta^2 |\ln \delta| \epsilon} \Delta g_{i,1,\text{om}} \sim \frac{c \phi'_0}{\Psi'_t} \partial_\alpha r_{i,1,W} \Upsilon f_{Mi}. \tag{8.24}$$

Here, one of the logarithms $|\ln \delta|$ due to a derivative with respect to J has cancelled with another logarithm $|\ln \delta|$ in τ_W , which is of order $R |\ln \delta| / \sqrt{\epsilon} v_{ii}$ because the bounce points of the orbits are always a small distance $\sqrt{\delta} R$ away from a local maximum of B . In (8.24), $r_{i,1,W}$ is proportional to $\lambda_W - 1/B_M \sim \delta \epsilon / B_0$ and hence of order $\delta \epsilon a$. As a result, we estimate $\Delta g_{i,1,\text{om}}$ to be

$$\Delta g_{i,1,\text{om}} \sim \delta^3 |\ln \delta| \frac{\rho_{i*}}{v_{i*}} \epsilon f_{Mi} \ll \delta \epsilon f_{Mi}. \tag{8.25}$$

This means that the solution for $g_{i,1,W}$ in the region $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$ cannot match with $g_{i,1,W}^{[1]}$ and we still have a discontinuity – see figure 12 for a sketch of the situation. A similar reasoning can be made for $g_{i,1,W}$ near a juncture.

Collisional boundary layers appear at the discontinuity between $g_{i,1,W}^{[1]}$ and the distribution function at $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$, and at the discontinuities on junctures. In these boundary layers, $g_{i,1,W}$ changes significantly over a width $\Delta J_{\sqrt{v}} \ll \sqrt{\epsilon} v_{ii} R$, giving $\partial_J g_{i,1,W} \sim g_{i,1,W} / \Delta J_{\sqrt{v}} \gg g_{i,1,W} / \sqrt{\epsilon} v_{ii} R$. This large derivatives in J make the collision operator large and comparable to the $\mathbf{E} \times \mathbf{B}$ drift (Ho & Kulsrud 1987; Calvo *et al.* 2017, 2018),

$$\frac{v_{ii}}{\epsilon} \left(\frac{\sqrt{\epsilon} v_{ii} R}{\Delta J_{\sqrt{v}}} \right)^2 \ln \left(\frac{\sqrt{\epsilon} v_{ii} R}{\Delta J_{\sqrt{v}}} \right) \sim \frac{c \phi'_0}{\Psi'_t}. \tag{8.26}$$

The logarithm in this estimate comes from the fact that the bounce points of the particles in this layer are close to local maxima of the omnigenous magnetic field. Solving (8.26)

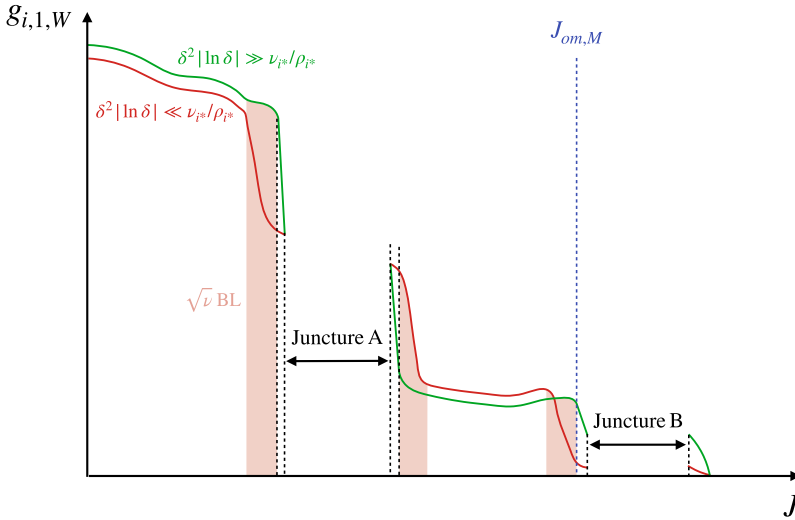


FIGURE 12. Sketch of the ion distribution function $g_{i,1,W}$ in near-omnigenous large aspect ratio stellarators with mirror ratios close to unity for $\delta^2 |\ln \delta| \ll v_{i^*} / \rho_{i^*} \ll 1$ (red line) and $v_{i^*} / \rho_{i^*} \ll \delta^2 |\ln \delta|$ (green line). The maximum value of the second adiabatic invariant in the omnigenous magnetic field $J_{om,M}^{[0]}$ is represented as a blue dashed line. We consider a case where there is a juncture in the region $J_{om,M}^{[0]} - J \gg \delta |\ln \delta| \sqrt{\epsilon} v_{ti} R$ (juncture A). Of all the junctures that appear for $J_{om,M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ti} R$, we only sketch one (juncture B). The \sqrt{v} boundary layers that appear in the case $\delta^2 |\ln \delta| \ll v_{i^*} / \rho_{i^*} \ll 1$ (red line) are highlighted as pink areas. Note that, outside of the \sqrt{v} boundary layers and the junctures, the derivative of the distribution function $g_{i,1,W}$ with respect to J remains the same as v_{i^*} / ρ_{i^*} changes.

for $\Delta J_{\sqrt{v}}$ gives

$$\frac{\Delta J_{\sqrt{v}}}{\sqrt{\epsilon} v_{ti} R} \sim \sqrt{\frac{v_{i^*}}{\rho_{i^*}} \ln \left(\frac{\rho_{i^*}}{v_{i^*}} \right)} \ll 1. \tag{8.27}$$

The distribution function in these boundary layers is not of the form in (8.22), and hence does not vanish under the integrals for the particle and energy fluxes, given in (8.19) and (8.20). Thus, the fluxes are mostly due to particles within the boundary layer, giving

$$\Gamma_i \sim \delta^2 \sqrt{\frac{v_{i^*}}{\rho_{i^*}} \ln \left(\frac{\rho_{i^*}}{v_{i^*}} \right)} \epsilon^{5/2} \rho_{i^*} n_i v_{ti}, \quad Q_i \sim \delta^2 \sqrt{\frac{v_{i^*}}{\rho_{i^*}} \ln \left(\frac{\rho_{i^*}}{v_{i^*}} \right)} \epsilon^{5/2} \rho_{i^*} n_i T_i v_{ti}. \tag{8.28a,b}$$

These are the particle and energy fluxes of the \sqrt{v} regime (Ho & Kulsrud 1987; Calvo *et al.* 2017, 2018). This regime only appears in large aspect ratio stellarators with mirror ratios close to unity when they are close to omnigenicity.

The estimates in (8.28a,b) can be obtained from a random walk argument. The particles that contribute most to radial transport are the particles in velocity-space boundary layers of width $\Delta v_{\parallel} \ll \sqrt{\epsilon} v_{ti}$. We start by estimating the width of the boundary layers. The effective collision frequency in these layers is

$$v_{\text{eff}} \sim \left(\frac{v_{ti}}{\Delta v_{\parallel}} \right)^2 \frac{v_{ti}}{\ln(\sqrt{\epsilon} v_{ti} / \Delta v_{\parallel})}. \tag{8.29}$$

We have obtained this effective collision frequency using the same arguments that we used to find (8.3). Note that the only difference between (8.3) and (8.29) is in the logarithm, where we find an extra factor of $\sqrt{\epsilon}$ in (8.29). In the layer, the effective collision frequency is of the same order as the inverse of the time that it takes for a particle to drift poloidally around the stellarator, $\nu_{\text{eff}} \sim |v_E|/a \sim \rho_{i*} v_{ti}/a$. This balance gives a boundary layer width

$$\Delta v_{\parallel} \sim \sqrt{\epsilon} v_{ti} \sqrt{\frac{v_{i*}}{\rho_{i*}} \frac{1}{\ln(\rho_{i*}/v_{i*})}}. \tag{8.30}$$

Comparing this result with (8.27), we see that the width of the layer in v_{\parallel} is smaller than in J , $\Delta J_{\sqrt{v}}/\sqrt{\epsilon} v_{ti} R \sim \ln(\rho_{i*}/v_{i*})(\Delta v_{\parallel}/\sqrt{\epsilon} v_{ti}) \gg \Delta v_{\parallel}/\sqrt{\epsilon} v_{ti}$. This is consistent with the transformation between the velocity-space variables \bar{v} and J and the cartesian velocity v . Indeed, we can take $\Delta v_{\parallel}/\sqrt{\epsilon} v_{ti} \sim B_0 \Delta \lambda/\epsilon$, with $\Delta \lambda$ the size of the layer in $\lambda := v_{\perp}^2/v^2 B$, and $\Delta J \sim \Delta \lambda \partial_{\lambda} J_W \sim \sqrt{\epsilon} v_{ti} R (B_0 \Delta \lambda/\epsilon) \ln(\epsilon/B_0 \Delta \lambda)$, where the logarithmic divergence is a result of particles spending a long time around the local maxima of B .

To obtain estimates for the fluxes, in addition to the layer width Δv_{\parallel} and the effective collision frequency ν_{eff} , we need the typical radial displacement between collisions of particles in the boundary layers, given by $\langle v_{Mi} \cdot \nabla r \rangle_{\tau,W}/\nu_{\text{eff}}$. Since $\langle v_{Mi} \cdot \nabla r \rangle_{\tau,W} \sim \delta \epsilon \rho_{i*} v_{ti}$ and $\nu_{\text{eff}} \sim \rho_{i*} v_{ti}/a$, we find $\langle v_{Mi} \cdot \nabla r \rangle_{\tau,W}/\nu_{\text{eff}} \sim \delta \epsilon a$. With all these results, we can write the random walk diffusion coefficient as

$$D \sim \frac{\Delta v_{\parallel}}{v_{ti}} \ln \left(\frac{\sqrt{\epsilon} v_{ti}}{\Delta v_{\parallel}} \right) \left(\frac{\langle v_{Mi} \cdot \nabla r \rangle_{\tau,W}}{\nu_{\text{eff}}} \right)^2 \nu_{\text{eff}}, \tag{8.31}$$

where we have used that the fraction of particles in the boundary layers is of the order of $(\Delta v_{\parallel}/v_{ti}) \ln(\sqrt{\epsilon} v_{ti}/\Delta v_{\parallel})$. The logarithmic correction in the particle fraction is due to the accumulation of slow particles around the local maxima of B . The diffusion coefficient in (8.31) gives $D \sim \delta^2 \sqrt{v_{i*} \rho_{i*}} \ln(\rho_{i*}/v_{i*}) \epsilon^{5/2} a v_{ti}$, and with it we can obtain the estimates in (8.28a,b).

8.2.2. *The v regime in optimized large aspect ratio stellarators with mirror ratios close to unity*

The \sqrt{v} regime stops being valid for $v_{i*}/\rho_{i*} \sim \delta^2 |\ln \delta|$ because the distribution function at $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ti} R$, given in (8.25), becomes comparable to $g_{i,1,W}^{[1]}$ and there is no need for a boundary layer any longer – see figure 12 for a sketch. The same happens for the discontinuities on junctures of different types of wells.

For $v_{i*}/\rho_{i*} \lesssim \delta^2 |\ln \delta|$, we need to use the procedure laid out in § 7. The resulting fluxes are given in (7.28) and (7.29). Particles with $J_{\text{om},M}^{[0]} - J \gg \delta |\ln \delta| \sqrt{\epsilon} v_{ti} R$ that are outside of the regions affected by junctures move without having to transition between different types of wells. Thus, particles with $J_{\text{om},M}^{[0]} - J \gg \delta |\ln \delta| \sqrt{\epsilon} v_{ti} R$ that are outside of regions affected by junctures are in the region of velocity space with $\partial_J K_{i,W} \neq 0$. The integral in this region has an integrand proportional to $1/\tau_W - 1/\langle \tau_W \rangle_{\alpha}$, and this difference is small because to lowest order, τ_W does not depend on α , $1/\tau_W - 1/\langle \tau_W \rangle_{\alpha} \sim \delta \sqrt{\epsilon} v_{ti}/R$. Moreover, $1/\tau_W - 1/\langle \tau_W \rangle_{\alpha} \sim \delta \sqrt{\epsilon} v_{ti}/R$ is integrated over α , and the integral vanishes to lowest order, finally giving

$$\int_0^{2\pi} \left(\frac{1}{\tau_W} - \frac{1}{\langle \tau_W \rangle_{\alpha}} \right) d\alpha \simeq \int_0^{2\pi} \frac{(\tau_W - \langle \tau_W \rangle_{\alpha})^2}{\langle \tau_W \rangle_{\alpha}^3} d\alpha \sim \frac{\delta^2 \sqrt{\epsilon} v_{ti}}{R}. \tag{8.32}$$

Then, the contribution to Γ_i and Q_i from particles with $J_{\text{om},M}^{[0]} - J \gg \delta |\ln \delta| \sqrt{\epsilon} v_{ti} R$ and outside of the regions affected by junctures are $\delta^2 \epsilon^{5/2} v_{i*} n_i v_{ti}$ and $\delta^2 \epsilon^{5/2} v_{i*} n_i T_i v_{ti}$,

respectively. Note that these are consistent with estimates that one can find using the approximate integrals (8.19) and (8.20) for the fluxes.

Conversely, in the region $J_{\text{om},M}^{[0]} - J \lesssim \delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$ or in the region where particles are affected by junctures, particles have multiple transitions, and hence $\partial_J K_{i,W} = 0$. There is then no cancellation similar to the one in (8.32). Using that the typical size of the intervals of interest in J is $\delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$ and that $\tau_W \sim R |\ln \delta| / \sqrt{\epsilon} v_{ii}$, the integrals in (7.28) and (7.29) give

$$\Gamma_i \sim \delta \epsilon^{5/2} v_{i*} n_i v_{ii}, \quad Q_i \sim \delta \epsilon^{5/2} v_{i*} n_i T_i v_{ii}. \tag{8.33a,b}$$

These contributions are larger than the ones from the particles that satisfy $J_{\text{om},M}^{[0]} - J \gg \delta |\ln \delta| \sqrt{\epsilon} v_{ii} R$ and are outside of the regions affected by junctures, and hence dominate the fluxes.

We proceed to obtain the estimates in (8.33a,b) using a random walk argument. For $v_{i*} / \rho_{i*} \ll \delta^2 |\ln \delta|$, particles remain close to the flux surface in which they started because their radial drift averages out after several poloidal turns around the stellarator. The time for several poloidal turns is of the order of $a / |v_E|$ and, in that time, particles drift radially a distance $w \sim \langle \mathbf{v}_{Mi} \cdot \nabla r \rangle_{\tau, w a} / |v_E| \sim \delta \epsilon a$. Transport is dominated by particles in thin velocity layers of width $\Delta v_{\parallel} / \sqrt{\epsilon} v_{ii} \sim \delta$ located at the trapped–passing boundary or around junctures of several wells. For particles in these layers, we can apply the formula for the effective collision frequency in a thin layer in (8.29), finding $\nu_{\text{eff}} \sim v_{ii} / \delta^2 |\ln \delta|$. With these results, we find the diffusion coefficient $D \sim \delta |\ln \delta| \sqrt{\epsilon} w^2 \nu_{\text{eff}}$, where the fraction of particles in the boundary layer is $\delta |\ln \delta| \sqrt{\epsilon}$. The logarithmic correction to the fraction of particles in the layer is due to the accumulation of particles around the local maxima of B . Using all our estimates, we finally obtain $D \sim \delta \epsilon^{5/2} a v_{ii}$, and this diffusion coefficient leads to the estimates in (8.33a,b).

In figure 13, we sketch the estimates in (8.21a,b), (8.28a,b) and (8.33a,b) for the fluxes in near-omnigeneous large aspect ratio stellarators with mirror ratios close to unity. For comparison, we sketch the corresponding estimates (6.2a,b) and (7.30a,b) for generic large aspect ratio stellarators with mirror ratios close to unity. Note that the optimization towards omnigeneity is much more effective in the $1/\nu$ regime. The reduced effectiveness of the optimization in the $\sqrt{\nu}$ and ν regimes is due to the barely trapped particles and the particles near junctures of different types of wells.

9. Conclusion

There are four results in this article that are worth emphasizing. The first one is the set of orbit-averaged equations for neoclassical transport at low collisionality in large aspect ratio stellarators with mirror ratios close to unity derived in § 5. It consists of a kinetic equation for trapped particles, given in (5.25), and the corresponding boundary conditions to be imposed around junctures of different types of wells (see (5.33)) and at the trapped–passing boundary (see (5.34)). The trapped-particle distribution function obtained from these equations can then be integrated to give the particle and energy fluxes (see (5.48) and (5.50)). To our knowledge, this is the most detailed derivation of a model for large aspect ratio stellarators with mirror ratios close to unity in the limit $v_{i*} \sim \rho_{i*}$, and in conjunction with the model described in Calvo *et al.* (2017) and Calvo *et al.* (2018) for near-omnigeneous stellarators, is the only self-consistent local model for stellarator neoclassical transport at low collisionality that we are aware of. The set of equations has been implemented in KNOSOS (Velasco *et al.* 2021), and it produces fluxes close to those calculated by DKES but with much less computational effort. The fact that the model matches DKES neoclassical fluxes is unsurprising, as we have shown in Appendix G

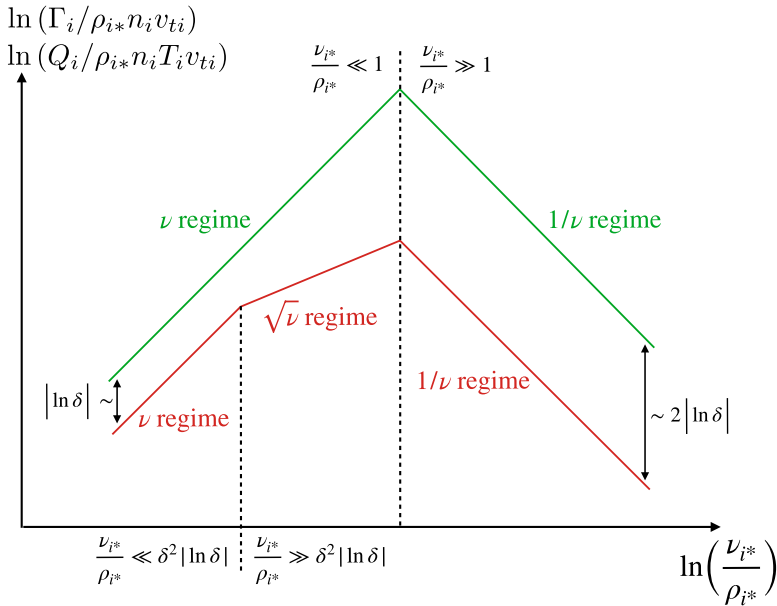


FIGURE 13. Particle and energy fluxes as a function of v_{i^*}/ρ_{i^*} for generic large aspect ratio stellarators with mirror ratios close to unity (green line) and near-omnigenous large aspect ratio stellarators with mirror ratios close to unity (red line).

that the model and DKES kinetic equations are the same to lowest order in the inverse aspect ratio. Interestingly, the derivation in Appendix G also shows that there is an $O(\sqrt{\epsilon})$ difference between our model and DKES equations, and DKES is not correct to that order. Thus, when the results from the equations in this paper differ from those of DKES, one cannot assume that DKES is correct by default. More theoretical and numerical work on these higher-order corrections to the aspect ratio expansion is needed as $\sqrt{\epsilon}$ is not very small.

The second interesting result in this article is the limit $v_{i^*} \ll \rho_{i^*}$ for generic large aspect ratio stellarators with mirror ratios close to unity, described in § 7. We have been able to show that there is no $\sqrt{\nu}$ regime when this type of stellarator is not close to omnigenity. Instead, generic large aspect ratio stellarators with mirror ratios close to unity enter the ν regime directly for $v_{i^*} \lesssim \rho_{i^*}$. We have also examined the ν regime in great detail, explaining how the distribution function behaves in this limit. The transitions from one type of well to another are a crucial aspect. While these transitions make the particle motion stochastic and diffuse particles in velocity space, there is no stochastic diffusion in real space and the neoclassical diffusion coefficient remains proportional to the collision frequency. We believe that transitions between different types of wells do not lead to stochastic real space diffusion independent of the collision frequency because ions in the regimes that we consider are confined by the electric field, and while transitions can cause jumps in the second adiabatic invariant J , they do not change the total energy of the particle. When the total energy is conserved, particles cannot move long distances ($\gg \epsilon a$) radially because they cannot overcome the barrier set by the electric potential, that is, particles remain within a distance of order ϵa of the flux surface that they started at without collisions. Collisions are needed to break conservation of total energy for each particle. While collisions conserve energy overall, in each collision, individual trapped particles can gain or lose energy. In our calculation, they mostly exchange energy with a

large population of passing particles. We should add that, even though transitions between different types of wells do not make the neoclassical diffusion independent of the collision frequency, they do enhance neoclassical diffusion, as we explain below (7.28).

Part of the outcome of the ν regime calculation is the explicit formula for the fluxes in the ν regime, given in (7.28) and (7.29). With these formulas, the fluxes can be calculated from the magnetic field of the stellarator without solving a kinetic equation. These explicit formulas might be useful for optimization routines if they can be evaluated fast. We believe that this is possible using techniques similar to those developed to perform bounce averages for KNOSOS (Velasco *et al.* 2020).

The third result that we want to emphasize is that large aspect ratio stellarators with large mirror ratios are close to omnigeneous and can be treated using the formulation developed by Calvo *et al.* (2017). These stellarators experience a $\sqrt{\nu}$ regime for $\epsilon |\ln \epsilon| \ll \nu_{i*} / \rho_{i*} \ll 1/\epsilon$. The neoclassical fluxes in large aspect ratio stellarators with large mirror ratios are significantly larger than the ones that one would obtain in an equivalent large aspect ratio stellarator with mirror ratio close to unity.

Finally, the fourth result of note is the limit of near omnigeneity for large aspect ratio stellarators with mirror ratios close to unity. When we consider this type of stellarator, we find the $\sqrt{\nu}$ regime for a range of collisionalities that depends on the deviation from omnigeneity δ : $\delta^2 |\ln \delta| \ll \nu_{i*} / \rho_{i*} \ll 1$. This interval of validity disappears when $\delta \sim 1$, explaining why we could not find a $\sqrt{\nu}$ regime in generic large aspect ratio stellarators with mirror ratios close to unity. The derivation for near-omnigeneous large aspect ratio stellarators with mirror ratios close to unity gives another interesting result, namely, optimization is less effective for $\nu_{i*} \ll \rho_{i*}$, and it is worse in the ν regime, where the fluxes are reduced only by a factor of δ instead of by a factor of δ^2 . For this reason, it is probably worth considering the use of the ν regime fluxes in (7.28) and (7.29) as optimization targets. If one can reduce effectively the fluxes in the regime that is less responsive to optimization, the fluxes in the other regimes will likely reduce even more. In particular, we expect that the optimization of the fluxes in (7.28) and (7.29) will target problematic particles at the trapped–passing boundary and in regions where there are transitions between different types of wells.

Acknowledgements

The authors would like to thank X. Liu, P.J. Catto and the two anonymous referees for their suggestions.

Editor Peter Catto thanks the referees for their advice in evaluating this article.

Funding

This work was supported by the U.S. Department of Energy under contract number DE-AC02-09CH11466. The United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Research and Training Programme (Grant Agreement No 101052200 – EUROfusion). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. This research was supported in part by Grant PGC2018-095307-B-I00, Ministerio de Ciencia, Innovación y Universidades, Spain.

Declaration of interests

The authors report no conflict of interest.

Appendix A. Discontinuities in the derivatives of the distribution function at junctures of different types of wells

In this appendix, we show how to obtain (2.36) by imposing, in the example juncture of figure 1, that the collisional flux of particles out of wells *I* and *II* enters well *III*.

For trapped particles, the collision operator in (2.15) can be written as $C_{ii}[g_{i,w}, \bar{f}_i^{(0)}] = -|v_{\parallel}|(\partial_{\mathcal{E}} F_{\mathcal{E}}[g_{i,w}, \bar{f}_i^{(0)}] + \partial_{\mu} F_{\mu}[g_{i,w}, \bar{f}_i^{(0)}])$, where

$$F_{\mathcal{E}}[g_{i,w}, \bar{f}_i^{(0)}] := -\frac{\gamma_{ii}}{|v_{\parallel}|}(H_{\mathcal{E}\mathcal{E}}[\bar{f}_i^{(0)}] \partial_{\mathcal{E}} g_{i,w} + H_{\mathcal{E}\mu}[\bar{f}_i^{(0)}] \partial_{\mu} g_{i,w} - L_{\mathcal{E}}[\bar{f}_i^{(0)}] g_{i,w}) \tag{A1}$$

is the phase-space particle flux in the \mathcal{E} direction and

$$F_{\mu}[g_{i,w}, \bar{f}_i^{(0)}] := -\frac{\gamma_{ii}}{|v_{\parallel}|}(H_{\mu\mathcal{E}}[\bar{f}_i^{(0)}] \partial_{\mathcal{E}} g_{i,w} + H_{\mu\mu}[\bar{f}_i^{(0)}] \partial_{\mu} g_{i,w} - L_{\mu}[\bar{f}_i^{(0)}] g_{i,w}) \tag{A2}$$

is the phase-space particle flux in the μ direction. We need the particle flux perpendicular to the boundary $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$. To obtain this flux, we use coordinates in which the boundary $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$ becomes trivial: instead of using $r, \alpha, \mathcal{E}, \mu$ and φ , we replace the variable \mathcal{E} by $\Delta\mathcal{E} := \mathcal{E} - \mathcal{E}_c(r, \alpha, \mu, t)$ such that $\Delta\mathcal{E} = 0$ gives the desired boundary. With these new variables, the collision operator in (2.15) becomes

$$C_{ii}[g_{i,w}, \bar{f}_i^{(0)}] = -|v_{\parallel}| \{ \partial_{\Delta\mathcal{E}}(F_{\mathcal{E}}[g_{i,w}, \bar{f}_i^{(0)}] + \partial_{\mu} \Delta\mathcal{E} F_{\mu}[g_{i,w}, \bar{f}_i^{(0)}]) + \partial_{\mu} F_{\mu}[g_{i,w}, \bar{f}_i^{(0)}] \}, \tag{A3}$$

where we have used the fact that $\partial_{\mu} \Delta\mathcal{E} = -\partial_{\mu} \mathcal{E}_c$ satisfies $\partial_{\Delta\mathcal{E}}(\partial_{\mu} \Delta\mathcal{E}) = 0$. Equation (A3) indicates that the flux of particles across the boundary $\Delta\mathcal{E} = 0$ is given by $F_{\mathcal{E}} - \partial_{\mu} \mathcal{E}_c F_{\mu}$. Integrating this flux over the gyrophase φ and along trapped orbits, we find that the phase-space flux across the boundary $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$ can finally be written as

$$-2\pi\gamma_{ii} \sum_{\sigma} \int_{l_{bL,W}}^{l_{bR,W}} [H_{\mathcal{E}\mathcal{E}}[\bar{f}_i^{(0)}] \partial_{\mathcal{E}} g_{i,w} + H_{\mathcal{E}\mu}[\bar{f}_i^{(0)}] \partial_{\mu} g_{i,w} - L_{\mathcal{E}}[\bar{f}_i^{(0)}] g_{i,w} - \partial_{\mu} \mathcal{E}_c (H_{\mu\mathcal{E}}[\bar{f}_i^{(0)}] \partial_{\mathcal{E}} g_{i,w} + H_{\mu\mu}[\bar{f}_i^{(0)}] \partial_{\mu} g_{i,w} - L_{\mu}[\bar{f}_i^{(0)}] g_{i,w})] \frac{dl}{|v_{\parallel}|}. \tag{A4}$$

Using the definition of the transit average in (2.28), the balance between the three collisional fluxes in and out of wells *I*, *II* and *III* in figure 1 is

$$\begin{aligned} & -2\pi\gamma_{ii}\tau_I [\langle H_{\mathcal{E}\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,I} \partial_{\mathcal{E}} g_{i,I} + \langle H_{\mathcal{E}\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,I} \partial_{\mu} g_{i,I} - \langle L_{\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,I} g_{i,I} \\ & - \partial_{\mu} \mathcal{E}_c (\langle H_{\mu\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,I} \partial_{\mathcal{E}} g_{i,I} + \langle H_{\mu\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,I} \partial_{\mu} g_{i,I} - \langle L_{\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,I} g_{i,I}) \\ & - 2\pi\gamma_{ii}\tau_{II} [\langle H_{\mathcal{E}\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,II} \partial_{\mathcal{E}} g_{i,II} + \langle H_{\mathcal{E}\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,II} \partial_{\mu} g_{i,II} - \langle L_{\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,II} g_{i,II} \\ & - \partial_{\mu} \mathcal{E}_c (\langle H_{\mu\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,II} \partial_{\mathcal{E}} g_{i,II} + \langle H_{\mu\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,II} \partial_{\mu} g_{i,II} - \langle L_{\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,II} g_{i,II}) \\ & = -2\pi\gamma_{ii}\tau_{III} [\langle H_{\mathcal{E}\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,III} \partial_{\mathcal{E}} g_{i,III} + \langle H_{\mathcal{E}\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,III} \partial_{\mu} g_{i,III} - \langle L_{\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,III} g_{i,III} \\ & - \partial_{\mu} \mathcal{E}_c (\langle H_{\mu\mathcal{E}}[\bar{f}_i^{(0)}] \rangle_{\tau,III} \partial_{\mathcal{E}} g_{i,III} + \langle H_{\mu\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,III} \partial_{\mu} g_{i,III} - \langle L_{\mu}[\bar{f}_i^{(0)}] \rangle_{\tau,III} g_{i,III})]. \end{aligned} \tag{A5}$$

This long expression can be simplified due to the continuity of $g_{i,w}$ and $\partial_{\mu} \mathcal{E}_c \partial_{\mathcal{E}} g_{i,w} + \partial_{\mu} g_{i,w}$ across $\mathcal{E} = \mathcal{E}_c(r, \alpha, \mu, t)$. The continuity of $g_{i,w}$ implies that $\tau_I \langle H_{pq}[\bar{f}_i^{(0)}] \rangle_{\tau,I} +$

$\tau_{II}\langle H_{pq}\bar{f}_i^{(0)}\rangle_{\tau,II} = \tau_{III}\langle H_{pq}\bar{f}_i^{(0)}\rangle_{\tau,III}$ and that $\tau_I\langle L_p\bar{f}_i^{(0)}\rangle_{\tau,I} + \tau_{II}\langle L_p\bar{f}_i^{(0)}\rangle_{\tau,II} = \tau_{III}\langle L_p\bar{f}_i^{(0)}\rangle_{\tau,III}$ for $p = \mathcal{E}, \mu$ and $q = \mathcal{E}, \mu$. Using these results and dividing by $-2\pi\gamma_{ii}$, expression (A5) simplifies to

$$\begin{aligned} &\tau_I[\langle H_{\mathcal{E}\mathcal{E}}\bar{f}_i^{(0)}\rangle_{\tau,I}(\partial_{\mathcal{E}}g_{i,I} - \partial_{\mathcal{E}}g_{i,III}) + \langle H_{\mathcal{E}\mu}\bar{f}_i^{(0)}\rangle_{\tau,I}(\partial_{\mu}g_{i,I} - \partial_{\mu}g_{i,III}) \\ &\quad - \partial_{\mu}\mathcal{E}_c(\langle H_{\mu\mathcal{E}}\bar{f}_i^{(0)}\rangle_{\tau,I}(\partial_{\mathcal{E}}g_{i,I} - \partial_{\mathcal{E}}g_{i,III}) + \langle H_{\mu\mu}\bar{f}_i^{(0)}\rangle_{\tau,I}(\partial_{\mu}g_{i,I} - \partial_{\mu}g_{i,III}))] \\ &\quad + \tau_{II}[\langle H_{\mathcal{E}\mathcal{E}}\bar{f}_i^{(0)}\rangle_{\tau,II}(\partial_{\mathcal{E}}g_{i,II} - \partial_{\mathcal{E}}g_{i,III}) + \langle H_{\mathcal{E}\mu}\bar{f}_i^{(0)}\rangle_{\tau,II}(\partial_{\mu}g_{i,II} - \partial_{\mu}g_{i,III}) \\ &\quad - \partial_{\mu}\mathcal{E}_c(\langle H_{\mu\mathcal{E}}\bar{f}_i^{(0)}\rangle_{\tau,II}(\partial_{\mathcal{E}}g_{i,II} - \partial_{\mathcal{E}}g_{i,III}) + \langle H_{\mu\mu}\bar{f}_i^{(0)}\rangle_{\tau,II}(\partial_{\mu}g_{i,II} - \partial_{\mu}g_{i,III}))] \\ &= 0. \end{aligned} \tag{A6}$$

Using the continuity of the combination of $\partial_{\mu}\mathcal{E}_c \partial_{\mathcal{E}}g_{i,W} + \partial_{\mu}g_{i,W}$, we can write $\partial_{\mu}g_{i,I} - \partial_{\mu}g_{i,III} = -\partial_{\mu}\mathcal{E}_c(\partial_{\mathcal{E}}g_{i,I} - \partial_{\mathcal{E}}g_{i,III})$ and $\partial_{\mu}g_{i,II} - \partial_{\mu}g_{i,III} = -\partial_{\mu}\mathcal{E}_c(\partial_{\mathcal{E}}g_{i,II} - \partial_{\mathcal{E}}g_{i,III})$. With these results, and employing the fact that $H_{\mathcal{E}\mu} = H_{\mu\mathcal{E}}$, expression (A6) becomes

$$\begin{aligned} &\tau_I[\langle H_{\mathcal{E}\mathcal{E}}\bar{f}_i^{(0)}\rangle_{\tau,I} - 2\langle H_{\mathcal{E}\mu}\bar{f}_i^{(0)}\rangle_{\tau,I}\partial_{\mu}\mathcal{E}_c + \langle H_{\mu\mu}\bar{f}_i^{(0)}\rangle_{\tau,I}(\partial_{\mu}\mathcal{E}_c)^2](\partial_{\mathcal{E}}g_{i,I} - \partial_{\mathcal{E}}g_{i,III}) \\ &\quad + \tau_{II}[\langle H_{\mathcal{E}\mathcal{E}}\bar{f}_i^{(0)}\rangle_{\tau,II} - 2\langle H_{\mathcal{E}\mu}\bar{f}_i^{(0)}\rangle_{\tau,II}\partial_{\mu}\mathcal{E}_c + \langle H_{\mu\mu}\bar{f}_i^{(0)}\rangle_{\tau,II}(\partial_{\mu}\mathcal{E}_c)^2](\partial_{\mathcal{E}}g_{i,II} - \partial_{\mathcal{E}}g_{i,III}) \\ &= 0. \end{aligned} \tag{A7}$$

Using the identities $\tau_I\langle H_{pq}\bar{f}_i^{(0)}\rangle_{\tau,I} + \tau_{II}\langle H_{pq}\bar{f}_i^{(0)}\rangle_{\tau,II} = \tau_{III}\langle H_{pq}\bar{f}_i^{(0)}\rangle_{\tau,III}$ again, we obtain the compact expression in (2.36).

Appendix B. Conditions on the flux surface shape imposed by the MHD equilibrium equations

In this appendix, we give the constraints that $\mathbf{x}_1(r, \alpha, l)$ must satisfy. In addition to (3.10), \mathbf{x}_1 must satisfy the first-order version of (3.3),

$$\hat{\mathbf{b}}_0 \cdot \partial_l \mathbf{x}_1 = 0, \tag{B1}$$

and a solvability constraint of the MHD force balance equations that we proceed to obtain. The second-order terms in the ϵ expansion of (3.5) and (3.6) are given by

$$\begin{aligned} &\partial_r \left(\frac{B_0 B_2}{4\pi} + \frac{B_1^2}{8\pi} - \frac{B_0^2}{4\pi} \boldsymbol{\kappa}_0 \cdot \mathbf{x}_2 - \frac{B_0}{4\pi} \frac{dB_0}{dl} \hat{\mathbf{b}}_0 \cdot \mathbf{x}_2 \right) - \frac{\partial_l(B_0 B_1)}{4\pi} \hat{\mathbf{b}}_0 \cdot \partial_r \mathbf{x}_1 \\ &\quad - \frac{B_0}{4\pi} \frac{dB_0}{dl} \partial_l \mathbf{x}_1 \cdot \partial_r \mathbf{x}_1 = \frac{B_0 B_1}{2\pi} \boldsymbol{\kappa}_0 \cdot \partial_r \mathbf{x}_1 + \frac{B_0^2}{4\pi} \partial_{ll}^2 \mathbf{x}_1 \cdot \partial_r \mathbf{x}_1, \end{aligned} \tag{B2}$$

and

$$\begin{aligned} &\partial_{\alpha} \left(\frac{B_0 B_2}{4\pi} + \frac{B_1^2}{8\pi} - \frac{B_0^2}{4\pi} \boldsymbol{\kappa}_0 \cdot \mathbf{x}_2 - \frac{B_0}{4\pi} \frac{dB_0}{dl} \hat{\mathbf{b}}_0 \cdot \mathbf{x}_2 \right) - \frac{\partial_l(B_0 B_1)}{4\pi} \hat{\mathbf{b}}_0 \cdot \partial_{\alpha} \mathbf{x}_1 \\ &\quad - \frac{B_0}{4\pi} \frac{dB_0}{dl} \partial_l \mathbf{x}_1 \cdot \partial_{\alpha} \mathbf{x}_1 = \frac{B_0 B_1}{2\pi} \boldsymbol{\kappa}_0 \cdot \partial_{\alpha} \mathbf{x}_1 + \frac{B_0^2}{4\pi} \partial_{ll}^2 \mathbf{x}_1 \cdot \partial_{\alpha} \mathbf{x}_1. \end{aligned} \tag{B3}$$

To eliminate B_2 and \mathbf{x}_2 from these equations, we differentiate (B2) with respect to α and (B3) with respect to r , and we subtract these two derivatives from each other to obtain

$$\begin{aligned} &\frac{B_0}{4\pi} \partial_l [B_0(\partial_{\alpha}^2 \mathbf{x}_1 \cdot \partial_r \mathbf{x}_1 - \partial_{lr}^2 \mathbf{x}_1 \cdot \partial_{\alpha} \mathbf{x}_1 + (\boldsymbol{\kappa}_0 \cdot \partial_{\alpha} \mathbf{x}_1)(\hat{\mathbf{b}}_0 \cdot \partial_r \mathbf{x}_1) - (\boldsymbol{\kappa}_0 \cdot \partial_r \mathbf{x}_1)(\hat{\mathbf{b}}_0 \cdot \partial_{\alpha} \mathbf{x}_1))] \\ &= -2P' \boldsymbol{\kappa}_0 \cdot \partial_{\alpha} \mathbf{x}_1, \end{aligned} \tag{B4}$$

where we have used (3.14) to write B_1 in terms of x_1 and (B1) to obtain $\hat{\mathbf{b}}_0 \cdot \partial_{l_r}^2 x_1 = 0$ and $\hat{\mathbf{b}}_0 \cdot \partial_{l_\alpha}^2 x_1 = 0$. Note that, if x_1 is a solution to (3.10), (B1) and (B4), the function $x_1(r, \alpha, l) + \lambda(r, \alpha)\hat{\mathbf{b}}_0(l)$ is also a solution. Here, $\lambda(r, \alpha)$ can be any function of r and α . This set of equivalent solutions arises from the fact that one is free to choose where l vanishes.

Appendix C. Lowest-order ion distribution function in large aspect ratio stellarators with mirror ratios close to unity

In this appendix, we first solve for the lowest-order trapped-particle distribution function $g_{i,0,W}$ in Appendix C.1 and we then obtain the lowest-order passing-particle distribution function $h_{i,0}$ in Appendix C.2. We finish with a discussion of the distribution function in the barely-passing-particle region in Appendix C.3.

C.1. Lowest-order trapped-particle distribution function

We proceed to rewrite (2.27) using the coordinates \bar{v} and J . We neglect the time derivatives and the source S_i employing the estimates in (5.1). Thus, (2.27) becomes

$$\begin{aligned} & \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla \alpha \rangle_{\tau,W} (\partial_\alpha g_{i,W} + \partial_\alpha J_W \partial_J g_{i,W}) \\ & + \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} (\partial_r g_{i,W} + \partial_r \bar{v} \partial_{\bar{v}} g_{i,W} + \partial_r J_W \partial_J g_{i,W}) \\ & \simeq \langle C_{ii}[g_{i,W}, \bar{f}_i^{(0)}] \rangle_{\tau,W}, \end{aligned} \tag{C1}$$

where the derivatives of $g_{i,W}$ with respect to r and α are performed holding \bar{v} and J fixed, whereas the derivatives of \bar{v} and J_W with respect to the same variables are performed holding \mathcal{E} and μ constant. Employing (5.17) and (5.18), to lowest order in ϵ , (C1) can be rewritten as

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,0,W} - \langle C_{ii}[g_{i,0,W}, \bar{f}_i^{(0)}] \rangle_{\tau,W} = 0. \tag{C2}$$

The terms proportional to $\partial_{\bar{v}} g_{i,0,W}$ and $\partial_r g_{i,0,W}$ have been neglected because we assume that the derivatives of $g_{i,0,W}$ with respect to \bar{v} and r satisfy $\partial_{\bar{v}} \ln g_{i,0,W} \sim v_{ii}^{-1}$ and $\partial_r \ln g_{i,0,W} \sim a^{-1}$. Recall that the estimate $\partial_r \ln g_{i,0,W} \sim a^{-1} \ll (\epsilon a)^{-1}$ is valid because the derivative is performed holding \bar{v} and J constant.

The fact that the derivative of $g_{i,0,W}$ with respect to J is larger than its derivative with respect to \bar{v} (compare (4.18) and (4.19)) simplifies significantly the collision operator in (C2) because the term with two derivatives with respect to J becomes the largest,

$$\begin{aligned} C_{ii}[g_{i,0,W}, \bar{f}_i^{(0)}] &= \gamma_{ii} \nabla_v J_W \cdot \partial_J (\nabla_v \nabla_v H[h_{i,0}] \cdot \nabla_v J_W \partial_J g_{i,0,W}) [1 + O(\epsilon^{1/2})] \\ &\sim \frac{v_{ii}}{\epsilon} g_{i,0,W}. \end{aligned} \tag{C3}$$

Here, the Rosenbluth potential $H[h_{i,0}]$ is evaluated using only $h_{i,0}$ because the contribution to the integral from $g_{i,0,W}$ is smaller by $\sqrt{\epsilon}$ due to the smallness of the region of velocity space where $g_{i,0,W}$ is defined. Since $\nabla_v J_W \simeq \tau_w v_{\parallel} \hat{\mathbf{b}}$ for trapped particles (see (4.12)), we need the component $\hat{\mathbf{b}} \cdot \nabla_v \nabla_v H[h_{i,0}] \cdot \hat{\mathbf{b}}$ of $\nabla_v \nabla_v H[h_{i,0}]$ for the lowest-order version of the collision operator. We find

$$\hat{\mathbf{b}} \cdot \nabla_v \nabla_v H[h_{i,0}] \cdot \hat{\mathbf{b}} = \int \frac{|\mathbf{v}_\perp - \mathbf{v}'_\perp|^2 h_{i,0}(\mathbf{v}')}{|\mathbf{v} - \mathbf{v}'|^3} d^3 v'. \tag{C4}$$

Note that, for trapped particles, the parallel component of the velocity is small in ϵ , giving $\mathbf{v} \simeq \bar{v} \hat{\mathbf{e}}_\perp(\mathbf{x}, \varphi)$, where $\hat{\mathbf{e}}_\perp$ is defined in (2.3). Using this result and writing

$\mathbf{v}' \simeq \bar{v}' \xi' \hat{\mathbf{b}} + \bar{v}' \sqrt{1 - \xi'^2} (\cos \varphi' \hat{\mathbf{e}}_{\perp} + \sin \varphi' \hat{\mathbf{b}} \times \hat{\mathbf{e}}_{\perp})$, (C4) can be written as

$$\hat{\mathbf{b}} \cdot \nabla_v \nabla_v H[h_{i,0}] \cdot \hat{\mathbf{b}} \simeq H_{bb}[h_{i,0}](\bar{v}), \tag{C5}$$

where we have defined the new functional

$$H_{bb}[h_{i,0}](\bar{v}) := \int \frac{\bar{v}^2 - 2\bar{v} \bar{v}' \sqrt{1 - \xi'^2} \cos \varphi' + \bar{v}'^2 (1 - \xi'^2)}{(\bar{v}^2 - 2\bar{v} \bar{v}' \sqrt{1 - \xi'^2} \cos \varphi' + \bar{v}'^2)^{3/2}} h_{i,0}(\bar{v}', \xi') \bar{v}'^2 d\bar{v}' d\xi' d\varphi'. \tag{C6}$$

This result demonstrates that $\hat{\mathbf{b}} \cdot \nabla_v \nabla_v H[h_{i,0}] \cdot \hat{\mathbf{b}}$ depends only on \bar{v} to lowest order in ϵ . With all these results, (C3) becomes

$$C_{ii}[g_{i,0,W}, \bar{f}_i^{(0)}] = \gamma_{ii} H_{bb}[h_{i,0}](\bar{v}) \tau_W v_{\parallel} \partial_J (\tau_W v_{\parallel} \partial_J g_{i,0,W}) [1 + O(\epsilon^{1/2})] \sim \frac{v_{ii}}{\epsilon} g_{i,0,W}. \tag{C7}$$

Transit averaging equation (C7), we obtain

$$\langle C_{ii}[g_{i,0,W}, \bar{f}_i^{(0)}] \rangle_{\tau,W} = \gamma_{ii} H_{bb}[h_{i,0}](\bar{v}) \partial_J (\tau_W J \partial_J g_{i,0,W}) [1 + O(\epsilon^{1/2})] \sim \frac{v_{ii}}{\epsilon} g_{i,0,W}. \tag{C8}$$

At the junctures of several wells, we need to impose continuity of $g_{i,0,W}$ and use condition (2.36) to relate the discontinuous derivatives on different sides of the well. We proceed to rewrite (2.36) in the new coordinates \bar{v} and J . The derivative of $g_{i,0,W}$ with respect to \mathcal{E} is $\partial_{\mathcal{E}} g_{i,0,W} = \partial_{\mathcal{E}} \bar{v} \partial_{\bar{v}} g_{i,0,W} + \partial_{\mathcal{E}} J_W \partial_J g_{i,0,W}$. Noting that $\partial_{\mathcal{E}} \bar{v} = 1/\bar{v} \sim 1/v_{ii}$, $\partial_{\mathcal{E}} J_W = \tau_W \simeq R/\sqrt{\epsilon} v_{ii}$, $\partial_{\bar{v}} g_{i,0,W} \sim g_{i,0,W}/v_{ii}$ and $\partial_J g_{i,0,W} \sim g_{i,0,W}/\sqrt{\epsilon} v_{ii} R$, the derivative with respect to \mathcal{E} simplifies to $\partial_{\mathcal{E}} g_{i,0,W} \simeq \tau_W \partial_J g_{i,0,W}$. Equation (2.36) can be simplified further by noting that $\mathcal{E}_c(r, \alpha, \mu, t)$ is given by (5.29) and hence $\partial_{\mu} \mathcal{E}_c \simeq B_{iM}$. With these results, (2.36) becomes

$$\begin{aligned} & \lim_{J \rightarrow J_{c,I}} \tau_I \partial_J g_{i,0,I} \sum_{\sigma} \int_{l_{bL,I}}^{l_{bR,I}} (H_{\mathcal{E}\mathcal{E}}[h_{i,0}] - 2B_{iM} H_{\mathcal{E}\mu}[h_{i,0}] + B_{iM}^2 H_{\mu\mu}[h_{i,0}]) \frac{dl}{|v_{\parallel}|} \\ & + \lim_{J \rightarrow J_{c,II}} \tau_{II} \partial_J g_{i,0,II} \sum_{\sigma} \int_{l_{bL,II}}^{l_{bR,II}} (H_{\mathcal{E}\mathcal{E}}[h_{i,0}] - 2B_{iM} H_{\mathcal{E}\mu}[h_{i,0}] + B_{iM}^2 H_{\mu\mu}[h_{i,0}]) \frac{dl}{|v_{\parallel}|} \\ & = \lim_{J \rightarrow J_{c,III}} \tau_{III} \partial_J g_{i,0,III} \sum_{\sigma} \int_{l_{bL,III}}^{l_{bR,III}} (H_{\mathcal{E}\mathcal{E}}[h_{i,0}] - 2B_{iM} H_{\mathcal{E}\mu}[h_{i,0}] + B_{iM}^2 H_{\mu\mu}[h_{i,0}]) \frac{dl}{|v_{\parallel}|}. \end{aligned} \tag{C9}$$

Note that, $H_{\mathcal{E}\mathcal{E}} = \nabla_v \mathcal{E} \cdot \nabla_v \nabla_v H \cdot \nabla_v \mathcal{E} = \mathbf{v} \cdot \nabla_v \nabla_v H \cdot \mathbf{v}$, $H_{\mathcal{E}\mu} = \nabla_v \mathcal{E} \cdot \nabla_v \nabla_v H \cdot \nabla_v \mu = \mathbf{v} \cdot \nabla_v \nabla_v H \cdot \mathbf{v}_{\perp}/B$ and $H_{\mu\mu} = \nabla_v \mu \cdot \nabla_v \nabla_v H \cdot \nabla_v \mu = \mathbf{v}_{\perp} \cdot \nabla_v \nabla_v H \cdot \mathbf{v}_{\perp}/B^2$. Thus,

$$\begin{aligned} & H_{\mathcal{E}\mathcal{E}}[h_{i,0}] - 2B_{iM} H_{\mathcal{E}\mu}[h_{i,0}] + B_{iM}^2 H_{\mu\mu}[h_{i,0}] \\ & = \left(v_{\parallel} \hat{\mathbf{b}} + \mathbf{v}_{\perp} \left(1 - \frac{B_{iM}}{B} \right) \right) \cdot \nabla_v \nabla_v H[h_{i,0}] \cdot \left(v_{\parallel} \hat{\mathbf{b}} + \mathbf{v}_{\perp} \left(1 - \frac{B_{iM}}{B} \right) \right). \end{aligned} \tag{C10}$$

Since $1 - B_{iM}/B \sim \epsilon$ but $v_{\parallel} \sim \sqrt{\epsilon} v_{ii}$, $H_{\mathcal{E}\mathcal{E}}[h_{i,0}] - 2B_{iM} H_{\mathcal{E}\mu}[h_{i,0}] + B_{iM}^2 H_{\mu\mu}[h_{i,0}] = v_{\parallel}^2 H_{bb}[h_{i,0}] (1 + O(\epsilon^{1/2}))$. Then, by dividing (C9) by $H_{bb}[h_{i,0}](\bar{v})$, we find

$$J_{c,I} \lim_{J \rightarrow J_{c,I}} \tau_I \partial_J g_{i,0,I} + J_{c,II} \lim_{J \rightarrow J_{c,II}} \tau_{II} \partial_J g_{i,0,II} = J_{c,III} \lim_{J \rightarrow J_{c,III}} \tau_{III} \partial_J g_{i,0,III}. \tag{C11}$$

Equation (C11) could have also been obtained by following a procedure similar to the one in Appendix A. Note that (C8) implies that the phase-space particle flux across the boundary $J = J_{c,W}(r, \alpha, \bar{v}, t)$ is $-2\pi \gamma_{ii} H_{bb}[h_{i,0}](\bar{v}) J_{c,W} \tau_W \partial_J g_{i,0,W}$.

Using (C2) and (C8), we find that the equation for $g_{i,0,W}$ is

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,0,W} - \gamma_{ii} H_{bb}[h_{i,0}](\bar{v}) \partial_J (\tau_W J \partial_J g_{i,0,W}) = 0. \tag{C12}$$

We need to impose the continuity condition (2.24) that implies that $g_{i,0,W}$ for $J \rightarrow \infty$ must be equal to $h_{i,0}$ at the trapped–passing boundary, where $\xi \sim \sqrt{\epsilon} \ll 1$. Then,

$$\lim_{J \rightarrow \infty} g_{i,0,W_{bt}}(r, \alpha, \bar{v}, J, t) = h_{i,0}(r, \bar{v}, 0, t). \tag{C13}$$

Note that we have approximated the trapped–passing boundary by $\xi = 0$. We discuss this approximation in detail in Appendix C.3.

To find the solution to (C12), we first need to show that

$$\lim_{J \rightarrow \infty} \tau_{W_{bt}} J \partial_J g_{i,0,W_{bt}}(r, \alpha, \bar{v}, J, t) = 0, \tag{C14}$$

that is, there is no collisional flux from the trapped region into the passing region. Property (C14) can be shown to be true following the same procedure that we employed to find property (5.36). With property (C14), we can find the solution $g_{i,0,W}$ given in (5.4). We multiply (C12) by $g_{i,0,W}$ to write

$$\begin{aligned} & \frac{c\phi'_0}{\Psi'_t} \partial_\alpha \left(\frac{g_{i,0,W}^2}{2} \right) + \gamma_{ii} H_{bb}[h_{i,0}](\bar{v}) \tau_W J (\partial_J g_{i,0,W})^2 \\ & - \gamma_{ii} H_{bb}[h_{i,0}](\bar{v}) \partial_J (g_{i,0,W} \tau_W J \partial_J g_{i,0,W}) = 0. \end{aligned} \tag{C15}$$

We integrate this equation over J and α and we sum over W . We use (C11) at the junctures of different wells, and we employ property (C14) to show that a boundary term at $J \rightarrow \infty$ vanishes. After all these calculations, we find

$$\sum_W \int_{\alpha_{L,W}}^{\alpha_{R,W}} d\alpha \int_{J_{m,W}}^{J_{M,W}} dJ \tau_W J (\partial_J g_{i,0,W})^2 = 0. \tag{C16}$$

This equation implies that $g_{i,0,W}$ is independent of J . Using this result in (C12), we also find that $g_{i,0,W}$ is independent of α . We obtain solution (5.4) by applying the boundary condition (C13).

C.2. *Lowest-order passing-particle distribution function*

To lowest order in ϵ , according to (5.1), we can neglect the time derivatives and the source in (2.30), finding

$$\left\langle \frac{B}{|v_\parallel|} C_{ii}[h_i, h_i] \right\rangle_{fs} \simeq 0. \tag{C17}$$

For most passing particles, $v_\parallel \simeq \sigma \sqrt{2(\mathcal{E} - \mu B_0 - Z_i e \phi_0(r, t)/m_i)}$ is very close to being independent of α and l . Thus, for most values of ξ , (C17) becomes

$$C_{ii}[h_{i,0}, h_{i,0}] = 0. \tag{C18}$$

We need to impose boundary condition (2.25) that requires that the derivatives of $g_{i,W}$ and h_i with respect to velocity are continuous across the trapped–passing boundary. The continuity of the derivatives imposes that the collisional flux across the trapped–passing

boundary be continuous. This collisional flux is dominated by pitch-angle scattering events, and the pitch-angle scattering flux is controlled by the derivative of $g_{i,W}$ with respect to J in the trapped region of velocity space. Since we have shown that $\partial_J g_{i,0,W} = 0$, only the derivatives of $g_{i,1,W}$ play a role in the collisional flux across the trapped–passing boundary. Due to the smallness of $|v_{\parallel}| \sim \sqrt{\epsilon} v_{ii}$ in the trapped region, $\nabla_v g_{i,1,W} \sim \epsilon^{-1/2} g_{i,1,W}/v_{ii} \sim \epsilon^{1/2} h_{i,0}/v_{ii}$, where we have used $g_{i,1,W} \sim \epsilon h_{i,0}$. Hence, the collisional flux across the trapped–passing boundary is small in ϵ because $\nabla_v g_{i,1,W}$ is small in ϵ . The lack of collisional flux across the trapped–passing boundary gives the boundary condition

$$\partial_{\xi} h_{i,0}(r, \bar{v}, 0, t) = 0, \tag{C19}$$

for $h_{i,0}$. A more rigorous derivation of this boundary condition can be found in Appendix C.3.

To obtain boundary condition (C19), the fact that there is no piece of $g_{i,W}$ of order $\epsilon^{1/2} g_{i,0,W}$ is important. Using (4.5), (4.6) and (5.18), we find that the next-order corrections to (C12) are

$$\begin{aligned} & \frac{c\phi'_0}{\Psi'_t} \partial_{\alpha} g_{i,1,W} - \gamma_{ii} H_{bb}[h_{i,0}](\bar{v}) \partial_J (\tau_W J \partial_J g_{i,1,W}) = \langle C_{ii}[g_{i,0,W}, h_{i,0}] \rangle_{\tau,W} \\ & + \frac{m_i c \phi'_0 \bar{v}^2}{2Z_i e B_0 \Psi'_t} \langle \partial_{\alpha} B_1 \rangle_{\tau,W} \left(\partial_r g_{i,0,W} - \frac{Z_i e \phi'_0}{m_i \bar{v}} \partial_{\bar{v}} g_{i,0,W} \right). \end{aligned} \tag{C20}$$

To obtain this result, we have used the fact that $\langle C_{ii}[g_{i,0,W}, h_{i,0}] \rangle_{\tau,W}$ is of order $v_{ii} g_{i,0,W}$ instead of $v_{ii} g_{i,0,W}/\epsilon$ because $g_{i,0,W}$ only depends on \bar{v} . Thus, there is no correction to $g_{i,0,W}$ of order $\epsilon^{1/2} g_{i,0,W}$, and we can use boundary condition (C19).

The only possible solution to (C18) with condition (C19) is a stationary Maxwellian with density and temperature that are flux functions, as shown in (5.5).

C.3. Lowest-order barely-passing-particle distribution function

The coordinate ξ is not a constant of the motion for barely passing particles with $|\xi| \sim \sqrt{\epsilon} \ll 1$. In this small region of velocity space, it is more convenient to use as a coordinate the value of ξ at the location of the maximum of U ,

$$\begin{aligned} \bar{\xi}(r, \mathcal{E}, \mu, \sigma, t) & := \sigma \sqrt{\frac{\mathcal{E} - U_M(r, \mu, t)}{\mathcal{E} - Z_i e \phi_{U_M}(r, \mu, t)/m_i}} \\ & = \frac{\sigma \sqrt{2(\mathcal{E} - \mu B_M(r) - Z_i e \phi_0(r, t)/m_i)}}{\bar{v}} [1 + O(\epsilon^{1/2})], \end{aligned} \tag{C21}$$

where ϕ_{U_M} is the value of the potential at the location where U is maximum, and $B_M(r)$ is the maximum of B on flux surface r .

Using \bar{v} and $\bar{\xi}$ as coordinates, the parallel velocity can be written as

$$v_{\parallel} = \frac{\bar{\xi}}{|\bar{\xi}|} \bar{v} \sqrt{\bar{\xi}^2 + \frac{B_{1,M}(r) - B_1(r, \alpha, l)}{B_0}} [1 + O(\epsilon^{1/2})], \tag{C22}$$

where $B_{1,M}(r)$ is the maximum value of B_1 on flux surface r . According to this approximation, v_{\parallel} and hence ξ are constant for $|\bar{\xi}| \gg \sqrt{\epsilon}$. Moreover, $\bar{\xi} \simeq \xi$ for $|\bar{\xi}| \gg \sqrt{\epsilon}$. Unfortunately, we need to consider $|\bar{\xi}| \sim \sqrt{\epsilon}$ because it is in this region that we find the trapped–passing boundary: at the trapped–passing boundary, $\bar{\xi} = 0$.

In the barely-passing-particle region $|\bar{\xi}| \sim \sqrt{\epsilon}$, we can write the distribution function as

$$\begin{aligned}
 & h_i(r, \bar{v}, \bar{\xi}, t) + \Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t) \\
 & \simeq h_i(r, \bar{v}, \xi, t) + (\bar{\xi} - \xi) \partial_{\xi} h_i(r, \bar{v}, \xi, t) + \Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t), \tag{C23}
 \end{aligned}$$

where $h_i(r, \bar{v}, \bar{\xi}, t)$ is the function $h_i(r, \bar{v}, \xi, t)$ but with ξ replaced by $\bar{\xi}$, and $h_i(r, \bar{v}, \xi, t) \simeq h_{i,0}(r, \bar{v}, \xi, t)$ is the passing particle distribution function discussed in Appendix C.2. The passing-particle distribution function $h_i(r, \bar{v}, \xi, t) \simeq h_{i,0}(r, \bar{v}, \xi, t)$ is determined by (C18) and satisfies

$$\partial_{\xi} h_i(r, \bar{v}, \xi, t) \lesssim \sqrt{\epsilon} h_{i,0}, \tag{C24}$$

for $|\xi| \sim \sqrt{\epsilon}$, as we will see at the end of this appendix. The piece $\Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t)$ is small in ϵ and vanishes for $|\bar{\xi}| \gg \sqrt{\epsilon}$, leaving the solution $h_i(r, \bar{v}, \xi, t)$ for freely passing particles with $|\xi| \sim 1$. We proceed to show that $\Delta h_{i,\text{bp}}$ simply ensures that there is continuity in the collisional fluxes across the barely-passing-particle region.

In Appendix C.1, we show that $\partial_J g_{i,0,W} = 0$. With this result, we can give a bound for the size of $\partial_{\bar{\xi}} \Delta h_{i,\text{bp}}$. Using the continuity condition (2.24) to write

$$\lim_{J \rightarrow \infty} g_{i,W_{\text{bt}}}(r, \alpha, \bar{v}, J, t) = \lim_{\bar{\xi} \rightarrow 0} [h_i(r, \bar{v}, \bar{\xi}, t) + \Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t)], \tag{C25}$$

we can deduce that

$$\lim_{J \rightarrow \infty} \partial_{\bar{v}} g_{i,W_{\text{bt}}}(r, \alpha, \bar{v}, J, t) = \lim_{\bar{\xi} \rightarrow 0} [\partial_{\bar{v}} h_i(r, \bar{v}, \bar{\xi}, t) + \partial_{\bar{v}} \Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t)]. \tag{C26}$$

Applying this result to (2.25), where $\partial_{\mathcal{E}} = \partial_{\mathcal{E}} \bar{v} \partial_{\bar{v}} + \partial_{\mathcal{E}} J_W \partial_J$ for trapped particles and $\partial_{\mathcal{E}} = \partial_{\mathcal{E}} \bar{v} \partial_{\bar{v}} + \partial_{\mathcal{E}} \bar{\xi} \partial_{\bar{\xi}}$ for passing particles, we find

$$\lim_{J \rightarrow \infty} \partial_{\mathcal{E}} J_{W_{\text{bt}}} \partial_J g_{i,W_{\text{bt}}}(r, \alpha, \bar{v}, J, t) = \lim_{\bar{\xi} \rightarrow 0} \partial_{\mathcal{E}} \bar{\xi} [\partial_{\bar{\xi}} h_i(r, \bar{v}, \bar{\xi}, t) + \partial_{\bar{\xi}} \Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t)]. \tag{C27}$$

Using $\partial_{\mathcal{E}} J_W = \tau_W$, $\partial_{\mathcal{E}} \bar{\xi} = (1 - \bar{\xi}^2)/\bar{v}^2 \bar{\xi}$ and $\partial_J g_{i,0,W} = 0$, this condition becomes

$$\lim_{\bar{\xi} \rightarrow 0} \frac{\partial_{\bar{\xi}} h_i(r, \bar{v}, \bar{\xi}, t) + \partial_{\bar{\xi}} \Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t)}{\bar{\xi}} = \bar{v}^2 \lim_{J \rightarrow \infty} \tau_{W_{\text{bt}}} \partial_J g_{i,1,W_{\text{bt}}}(r, \alpha, \bar{v}, J, t). \tag{C28}$$

Using $|\bar{\xi}| \sim \sqrt{\epsilon}$, $\tau_W \sim \epsilon^{-1/2} R/v_{ii}$, $\partial_J \sim \epsilon^{-1/2}/v_{ii} R$ and (C24), we find the estimate

$$\partial_{\bar{\xi}} \Delta h_{i,\text{bp}} \sim \sqrt{\epsilon} h_{i,0}. \tag{C29}$$

To calculate $\Delta h_{i,\text{bp}}$, we need to integrate $\partial_{\bar{\xi}} \Delta h_{i,\text{bp}} \sim \sqrt{\epsilon} h_{i,0}$ over $\bar{\xi}$. Since the region of interest is $|\bar{\xi}| \sim \sqrt{\epsilon}$, the integral is over an interval of $\bar{\xi}$ of order $\sqrt{\epsilon}$, and hence $\Delta h_{i,\text{bp}} \sim \sqrt{\epsilon} \partial_{\bar{\xi}} \Delta h_{i,\text{bp}} \sim \epsilon h_{i,0}$, justifying the neglect of $\Delta h_{i,\text{bp}}$ in (C25) to obtain (C13). In Appendix D we will see that $\Delta h_{i,\text{bp}} \sim \epsilon h_{i,0}$ is in fact only a bound as $\Delta h_{i,\text{bp}}$ is smaller than $\epsilon h_{i,0}$.

We proceed to solve (C17) in the barely-passing-particle region. Using $B \simeq B_0$ and (C18), we find that (C17) becomes, to lowest order in ϵ ,

$$B_0 \langle |v_{\parallel}|^{-1} C_{ii} [(\bar{\xi} - \xi) \partial_{\xi} h_i(r, \bar{v}, \xi, t) + \Delta h_{i,\text{bp}}, h_{i,0}] \rangle_{\text{fs}} = 0, \tag{C30}$$

where we have neglected the contributions of the difference $\bar{\xi} - \xi$ and of $\Delta h_{i,\text{bp}}$ to the Rosenbluth potentials because they are small in ϵ . The derivatives of $\bar{\xi} - \xi$ and $\Delta h_{i,\text{bp}}$ with

respect to $\bar{\xi}$ in the barely-passing-particle region are large, $\partial_{\bar{\xi}} \ln(\bar{\xi} - \xi) \sim \partial_{\bar{\xi}} \ln \Delta h_{i,\text{bp}} \sim \epsilon^{-1/2}$. Thus, in the collision operator applied to $\bar{\xi} - \xi$ and $\Delta h_{i,\text{bp}}$, the term that contains two derivatives with respect to $\bar{\xi}$ dominates,

$$C_{ii} [(\bar{\xi} - \xi) \partial_{\xi} h_i(r, \bar{v}, \xi, t) + \Delta h_{i,\text{bp}}, h_{i,0}] \simeq \gamma_{ii} \nabla_v \bar{\xi} \cdot \partial_{\bar{\xi}} [\nabla_v \nabla_v H[h_{i,0}] \cdot \nabla_v \bar{\xi} \partial_{\bar{\xi}} ((\bar{\xi} - \xi) \partial_{\xi} h_i(r, \bar{v}, \xi, t) + \Delta h_{i,\text{bp}})] \sim \nu_{ii} h_{i,0}. \tag{C31}$$

Using $\nabla_v \bar{\xi} \simeq (\xi/\bar{\xi}) \hat{\mathbf{b}}$, $\partial_{\bar{\xi}} \bar{\xi} \simeq \bar{\xi}/\xi$ and (C5), and employing the fact that the derivative of $\partial_{\xi} h_i(r, \bar{v}, \xi, t)$ with respect to $\bar{\xi}$ is small compared with the derivatives of $\bar{\xi} - \xi$ and $\Delta h_{i,\text{bp}}$ with respect to $\bar{\xi}$, the collision operator simplifies to

$$C_{ii} [(\bar{\xi} - \xi) \partial_{\xi} h_i(r, \bar{v}, 0, t) + \Delta h_{i,\text{bp}}, h_{i,0}] \simeq \frac{\gamma_{ii} H_{bb}[h_{i,0}]}{\bar{v}^2 \bar{\xi}} v_{\parallel} \partial_{\bar{\xi}} \left[\frac{v_{\parallel}}{\bar{\xi}} \left(\partial_{\bar{\xi}} \Delta h_{i,\text{bp}} + \left(1 - \frac{\bar{v} \bar{\xi}}{v_{\parallel}} \right) \partial_{\xi} h_i(r, \bar{v}, \xi, t) \right) \right] \sim \nu_{ii} h_{i,0}. \tag{C32}$$

Employing this result, (C30) becomes

$$\frac{\gamma_{ii} H_{bb}[h_{i,0}](\bar{v})}{\bar{v}^2 |\bar{\xi}|} \partial_{\bar{\xi}} \left[\frac{\langle |v_{\parallel}| \rangle_{\text{fs}}}{|\bar{\xi}|} \partial_{\bar{\xi}} \Delta h_{i,\text{bp}} + \left(\frac{\langle |v_{\parallel}| \rangle_{\text{fs}}}{|\bar{\xi}|} - \bar{v} \right) \partial_{\xi} h_i(r, \bar{v}, \xi, t) \right] = 0. \tag{C33}$$

We solve for $\partial_{\bar{\xi}} \Delta h_{i,\text{bp}}$ by realizing that $\partial_{\bar{\xi}} \Delta h_{i,\text{bp}}$ is small for $|\bar{\xi}| \gg \sqrt{\epsilon}$ because $\Delta h_{i,\text{bp}}$ vanishes for $|\bar{\xi}| \gg \sqrt{\epsilon}$, finding

$$\partial_{\bar{\xi}} \Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t) = \left(\frac{|\bar{\xi}|}{\left\langle \sqrt{\bar{\xi}^2 + (B_{1,M} - B_1)/B_0} \right\rangle_{\text{fs}}} - 1 \right) \partial_{\xi} h_i(r, \bar{v}, \xi, t), \tag{C34}$$

where we have used the lowest-order version of expression (C22) for the parallel velocity. We can integrate once more to obtain $\Delta h_{i,\text{bp}}$,

$$\Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t) = \partial_{\xi} h_i(r, \bar{v}, \xi, t) \frac{\bar{\xi}}{|\bar{\xi}|} \int_{|\bar{\xi}|}^{\infty} \left(1 - \frac{\xi'}{\left\langle \sqrt{\xi'^2 + (B_{1,M} - B_1)/B_0} \right\rangle_{\text{fs}}} \right) d\xi', \tag{C35}$$

where we have used that $\Delta h_{i,\text{bp}}$ vanishes for $|\bar{\xi}| \gg \sqrt{\epsilon}$.

Substituting (C34) into (C28), we obtain

$$\begin{aligned} \lim_{\xi \rightarrow 0^+} \partial_{\xi} h_i(r, \bar{v}, \xi, t) &= - \lim_{\xi \rightarrow 0^-} \partial_{\xi} h_i(r, \bar{v}, \xi, t) \\ &= \left\langle \sqrt{\frac{B_{1,M} - B_1}{B_0}} \right\rangle_{\text{fs}} \bar{v}^2 \lim_{J \rightarrow \infty} \tau_{W_{\text{bt}}} \partial_J g_{i,1,W_{\text{bt}}}(r, \alpha, \bar{v}, J, t) \sim \sqrt{\epsilon} h_{i,0}. \end{aligned} \tag{C36}$$

Since the derivatives of $h_{i,0}$ with respect to ξ at $\xi = 0$ are small in ϵ , they have to be set to zero to lowest order, giving condition (C19).

Appendix D. Barely passing particles

In this appendix, we show that the correction to the distribution function due to the barely passing particles, $\Delta h_{i,\text{bp}}$, is of order $\epsilon^2 f_{Mi}$ and hence several boundary conditions that we have used in this article are valid. To discuss the barely passing particles, we use the formalism developed in Appendix C.3. In Appendix C.3, we state that $\Delta h_{i,\text{bp}}$ is at most of order $\epsilon h_{i,0}$, but in reality it is smaller. In § 5.2 and Appendix F, we show that $\tau_w \partial_J g_{i,1,w}$ and $\tau_w \partial_J g_{i,3/2,w}$ vanish for $J \rightarrow \infty$. These results mean that (C28) must be modified to give

$$\lim_{\bar{\xi} \rightarrow 0} \frac{\partial_{\bar{\xi}} h_{i,3/2}(r, \bar{v}, \bar{\xi}, t) + \partial_{\bar{\xi}} \Delta h_{i,\text{bp}}(r, \bar{v}, \bar{\xi}, t)}{\bar{\xi}} = \bar{v}^2 \lim_{J \rightarrow \infty} \tau_{w_{\text{bt}}} \partial_J g_{i,2,w_{\text{bt}}}(r, \alpha, \bar{v}, J, t). \tag{D1}$$

This equation gives the estimate

$$\partial_{\bar{\xi}} \Delta h_{i,\text{bp}} \sim \epsilon^{3/2} f_{Mi}. \tag{D2}$$

By integrating $\partial_{\bar{\xi}} \Delta h_{i,\text{bp}}$ over the barely-passing-particle region, of width $\Delta \bar{\xi} \sim \sqrt{\epsilon}$, we obtain the size that we announced at the start of this appendix,

$$\Delta h_{i,\text{bp}} \sim \epsilon^2 f_{Mi}. \tag{D3}$$

The small size of this contribution justifies neglecting $\Delta h_{i,\text{bp}}$ in the continuity condition (C25) to obtain the boundary conditions (5.34) and (F8).

Solutions (C34) and (C35) are valid in the barely-passing-particle region even at the high order that we are considering. Following the same procedure that led to (C36) but using (D1) instead of (C28), we obtain the boundary condition

$$\begin{aligned} \lim_{\xi \rightarrow 0^+} \partial_{\xi} h_{i,3/2}(r, \bar{v}, \xi, t) &= - \lim_{\xi \rightarrow 0^-} \partial_{\xi} h_{i,3/2}(r, \bar{v}, \xi, t) \\ &= \left\langle \sqrt{\frac{B_{1,M} - B_1}{B_0}} \right\rangle_{\text{fs}} \bar{v}^2 \lim_{J \rightarrow \infty} \tau_{w_{\text{bt}}} \partial_J g_{i,2,w_{\text{bt}}}(r, \alpha, \bar{v}, J, t). \end{aligned} \tag{D4}$$

This is the boundary condition that connects $g_{i,2,w}$ and $h_{i,3/2}$.

Appendix E. Example of a problematic well juncture

In § 5.2.2, we state that wells end at locations $\alpha_{L,w}$ and $\alpha_{R,w}$ where $J_{m,w} = J_{M,w}$. There are cases in which this property is not obvious. We show such an example in figure 14(a), in which well *I* disappears at $\alpha = \alpha_{R,I}$ because $J_{m,I} = J_{M,I} = 0$ at that value of α . The problem with this configuration arises for $\alpha > \alpha_{R,I}$, where well *II* and *III* merge into a new larger well. There are different ways to address this problem, and we comment on two: one could terminate wells *II* and *III* at $\alpha = \alpha_{R,I}$ and start a new well *IV* at that value of α , or one could terminate well *III* and declare that the merged well beyond $\alpha_{R,I}$ is well *II*. In either of these options, a well disappears without $J_{m,w}$ being equal to $J_{M,w}$. Moreover, if we were to choose the second option and declare that the merged well is well *II*, the function $J_{c,II}$ that determines the location of the juncture in velocity space becomes discontinuous in α . Both of these features are undesirable, and instead we choose to add a ‘false’ two-well juncture for $\alpha > \alpha_{R,I}$, represented by a red dashed line in figure 14(a). We can locate the ‘false’ juncture $J_{c,II} = J_{c,III}$ that artificially separates particles into wells *II* and *III* wherever we wish. The only condition that we impose for convenience is that the extended functions $J_{c,II}$ and $J_{c,III}$ be continuous, as shown in the example of figure 14(b).

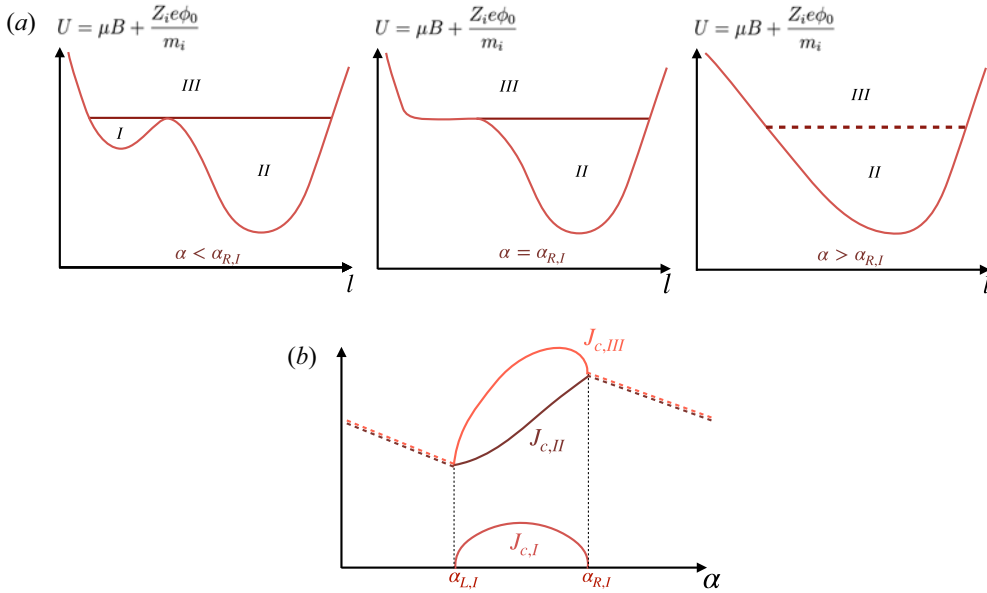


FIGURE 14. (a) Example of limit $\alpha_{R,I}$ in which well *I* closes and well *II* and *III* become indistinguishable. There are several ways to treat such a case. We choose to define a ‘false’ juncture $J_{c,II} = J_{c,III}$ (represented by a dashed line) that splits particles into well *II* and well *III* arbitrarily for $\alpha > \alpha_{R,I}$. (b) Sketch of the functions $J_{c,I}$, $J_{c,II}$ and $J_{c,III}$ vs α . The sketch is consistent with the $\alpha_{R,I}$ sketched in figure (a). The ‘false’ juncture $J_{c,II} = J_{c,III}$ is represented by dashed lines. The shape of $J_{c,II} = J_{c,III}$ in the region where well *I* does not exist is arbitrary. We choose it such that $J_{c,II}(r, \alpha, \bar{v}, t)$ and $J_{c,III}(r, \alpha, \bar{v}, t)$ are continuous in α .

Note that such a ‘false’ juncture does not affect the final result because junctures impose continuity of the distribution function and continuity of the collisional velocity-space flux, and hence continuity of the distribution function derivatives. With these ‘false’ junctures, we ensure that wells end only at values of α where $J_{m,W} = J_{M,W}$.

Appendix F. Correction $g_{i,3/2,W}$ to the trapped-particle distribution function

In this appendix, we show that $g_{i,3/2,W}$ does not depend on α or J and hence $\tau_W J \partial_J g_{i,3/2,W} = 0$ for all J and not only at the trapped–passing boundary.

We start by discussing the contributions to $g_{i,3/2,W}$ due to the collision operator. One might think that the differential piece of the linearized collision operator applied on $g_{i,1,W}$ has pieces of order $\epsilon^{1/2} v_{iJ} f_{Mi}$ that would have to be included in an equation for $g_{i,3/2,W}$, but after careful manipulations using $\nabla_v \bar{v} = \mathbf{v}/\bar{v}$ and (4.12) for $\nabla_v J_W$, we find that this is not the case. Indeed, since $\nabla_v = \nabla_v \bar{v} \partial_{\bar{v}} + \nabla_v J_W \partial_J$ and

$$\begin{aligned} \nabla_v \cdot (\dots) &= \frac{\tau_W \langle B \rangle_{\tau,W} |v_{\parallel}|}{\bar{v}} \left[\partial_{\bar{v}} \left(\frac{\bar{v}}{\tau_W \langle B \rangle_{\tau,W} |v_{\parallel}|} (\dots) \cdot \nabla_v \bar{v} \right) \right. \\ &\quad \left. + \partial_J \left(\frac{\bar{v}}{\tau_W \langle B \rangle_{\tau,W} |v_{\parallel}|} (\dots) \cdot \nabla_v J_W \right) \right], \end{aligned} \tag{F1}$$

we obtain

$$\begin{aligned}
 & \langle C_{ii,D}^\ell [g_{i,1,W}] \rangle_{\tau,W} - \frac{\bar{v}^2 v_{ii,\perp}(r, \bar{v}, t)}{4} \partial_J (\tau_W J \partial_J g_{i,1,W}) \\
 &= - \frac{\bar{v}^2 \tau_W J v_{ii,\perp}(r, \bar{v}, t)}{4} \partial_J g_{i,1,W} \partial_J \ln \langle B \rangle_{\tau,W} \\
 &+ \langle B \rangle_{\tau,W} \partial_J \left[\frac{\tau_W^2}{\langle B \rangle_{\tau,W}} \left\langle v_{\parallel}^2 \left(\frac{v^2 v_{ii,\perp}(r, v, t)}{4} - \frac{\bar{v}^2 v_{ii,\perp}(r, \bar{v}, t)}{4} \right) \right\rangle_{\tau,W} \partial_J g_{i,1,W} \right. \\
 &+ \frac{2\mu \tau_W^2}{\langle B \rangle_{\tau,W}} \left\langle \frac{v^2 B v_{ii,\perp}}{4} \left(1 - \frac{\langle B \rangle_{\tau,W}}{B} \right)^2 \right\rangle_{\tau,W} \partial_J g_{i,1,W} \\
 &+ \frac{\tau_W^2}{\langle B \rangle_{\tau,W}} \left\langle \left(\frac{v_{i,\parallel}}{2} - \frac{v_{i,\perp}}{4} \right) (v_{\parallel}^2 + 2\mu (B - \langle B \rangle_{\tau,W}))^2 \right\rangle_{\tau,W} \partial_J g_{i,1,W} \\
 &+ \left. \frac{\tau_W}{\bar{v} \langle B \rangle_{\tau,W}} \left\langle \frac{v^2 v_{i,\parallel}}{2} (v_{\parallel}^2 + 2\mu (B - \langle B \rangle_{\tau,W})) \right\rangle_{\tau,W} \partial_{\bar{v}} g_{i,1,W} \right] \\
 &+ \frac{\langle B \rangle_{\tau,W}}{\bar{v}} \partial_{\bar{v}} \left(\frac{\langle v^4 v_{i,\parallel} \rangle_{\tau,W}}{2\bar{v} \langle B \rangle_{\tau,W}} \partial_{\bar{v}} g_{i,1,W} \right) \sim v_{ii} g_{i,1,W} \sim \epsilon^2 \rho_{i*} \frac{v_{iJ} f_{Mi}}{a}. \tag{F2}
 \end{aligned}$$

The contributions of the integral part of the linearized collision operator to the trapped particle equation can be neglected to the order that determines $g_{i,3/2,W}$ because their size is given by $C_{ii,I}^\ell [h_{i,3/2}] \sim v_{ii} h_{i,3/2} \sim \epsilon^{5/2} \rho_{i*} v_{iJ} f_{Mi}/a$ and $C_{ii,I}^\ell [g_{i,1,W}] \sim \epsilon^{1/2} v_{ii} g_{i,1,W} \sim \epsilon^{5/2} \rho_{i*} v_{iJ} f_{Mi}/a$. Here, the integral operator applied on $g_{i,1,W}$ has an extra factor of $\sqrt{\epsilon}$ because the region of velocity space where $g_{i,1,W}$ is defined is small in $\sqrt{\epsilon}$.

We proceed to consider the contributions to the equation for $g_{i,3/2,W}$ due to the $\mathbf{E} \times \mathbf{B}$ drift $(c/B) \hat{\mathbf{b}} \times \nabla \phi_{3/2}$ (see (4.5) and (5.18)). This $\mathbf{E} \times \mathbf{B}$ contribution has already been accounted for in (5.25) in the term proportional to $(c\phi'_0/\Psi'_t) \partial_\alpha r_{i,1,W}$. Keeping higher-order corrections in (5.22), we find

$$J_W(r, \alpha, \bar{v}, \bar{\lambda}, t) \simeq 2\bar{v} \int_{l_{bL,W}}^{l_{bR,W}} \sqrt{1 - \bar{\lambda} B_0 - \frac{B_1(r, \alpha, l)}{B_0} - \frac{2Z_i e \phi_{3/2}(r, \alpha, l, t)}{m_i \bar{v}^2}} dl. \tag{F3}$$

We calculate $\partial_\alpha \bar{\lambda}_W$ by differentiating equation (F3) with respect to α holding r, \bar{v} and J constant. We find

$$\left\langle \frac{\bar{v}^2}{2B_0} \partial_\alpha B_1 + \frac{Z_i e}{m_i} \partial_\alpha \phi_{3/2} \right\rangle_{\tau,W} \simeq - \frac{\bar{v}^2 B_0}{2} \partial_\alpha \bar{\lambda}_W. \tag{F4}$$

This result shows that the $\mathbf{E} \times \mathbf{B}$ drift due to $\phi_{3/2}$ is included naturally in the term $(c\phi'_0/\Psi'_t) \partial_\alpha r_{i,1,W}$ because, using (4.5), we can write the radial drifts as

$$\begin{aligned}
 \langle (\mathbf{v}_E + \mathbf{v}_{Mi}) \cdot \nabla r \rangle_{\tau,W} &= - \frac{m_i c}{Z_i e \Psi'_t} \left\langle \frac{\bar{v}^2}{2B_0} \partial_\alpha B_1 + \frac{Z_i e}{m_i} \partial_\alpha \phi_{3/2} \right\rangle_{\tau,W} [1 + O(\epsilon)] \\
 &= \frac{c\phi'_0}{\Psi'_t} \partial_\alpha r_{i,1,W} [1 + O(\epsilon)]. \tag{F5}
 \end{aligned}$$

Due to all the considerations listed above, the equation for $g_{i,3/2,W}$ is

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,3/2,W} - \frac{\bar{v}^2 v_{ii,\perp}}{4} \partial_J (\tau_W J \partial_J g_{i,3/2,W}) = 0. \tag{F6}$$

At this order, the equation for the junctures of different wells is still the simple result

$$J_{c,I} \lim_{J \rightarrow J_{c,I}} \tau_I \partial_J g_{i,3/2,I} + J_{c,II} \lim_{J \rightarrow J_{c,II}} \tau_{II} \partial_J g_{i,3/2,II} = J_{c,III} \lim_{J \rightarrow J_{c,III}} \tau_{III} \partial_J g_{i,3/2,III}. \tag{F7}$$

Equations (F6) and (F7) have to be solved in conjunction with the boundary condition

$$\lim_{J \rightarrow \infty} g_{i,3/2,W_{\text{bt}}}(r, \alpha, \bar{v}, J, t) = h_{i,3/2}(r, \bar{v}, 0, t) \tag{F8}$$

that we obtain from condition (2.24) (see Appendix D for more detail). Using the same methodology that led to solution (5.4), we can show that the solution to (F6), (F7) and (F8) is given by (5.40) and hence $g_{i,3/2,W}$ does not depend on α or J .

Appendix G. Derivation of (5.25), (5.33) and (5.34) from the DKES kinetic equation

In this appendix, we derive (5.25), (5.33) and (5.34) from the DKES kinetic equation. The DKES kinetic equation assumes that the ion distribution function is the Maxwellian distribution function f_{Mi} corrected by the small piece \hat{f}_i ,

$$f_i(r, \alpha, l, v, \gamma, t) = f_{Mi}(r, v, t) + \hat{f}_i(r, \alpha, l, v, \gamma, t). \tag{G1}$$

The velocity-space coordinates are the magnitude of the velocity v and the angle γ between the velocity \mathbf{v} and the magnetic field, $\gamma := \arccos(\hat{\mathbf{b}} \cdot \mathbf{v}/v)$. The DKES kinetic equation in these variables is

$$\begin{aligned} & \left[v \cos \gamma + \frac{c\phi'_0}{\langle B^2 \rangle_{\text{fs}}} (\mathbf{B} \times \nabla r) \cdot \nabla l \right] \partial_l \hat{f}_i + \frac{c\phi'_0}{\Psi'_t} \frac{B^2}{\langle B^2 \rangle_{\text{fs}}} \partial_\alpha \hat{f}_i + \frac{1}{2B} v \sin \gamma \partial_l B \partial_\gamma \hat{f}_i \\ & - \mathcal{L}_{ii}[\hat{f}_i] = -\mathbf{v}_{Mi} \cdot \nabla r \Upsilon f_{Mi}, \end{aligned} \tag{G2}$$

where

$$\mathcal{L}_{ii}[\hat{f}_i] := \frac{v_{ii,\perp}(v)}{4 \sin \gamma} \partial_\gamma (\sin \gamma \partial_\gamma \hat{f}_i) \tag{G3}$$

is a model pitch-angle scattering operator. This equation has to be solved with magnetic fields that satisfy the MHD constraint $(\nabla \times \mathbf{B}) \cdot \nabla r = 0$. In the $\{r, \alpha, l\}$ coordinates, this constraint is

$$\partial_\alpha \left(\frac{B}{\Psi'_t} \right) + \partial_l \left[(\hat{\mathbf{b}} \times \nabla r) \cdot \nabla l \right] = 0. \tag{G4}$$

To obtain (5.25), (5.33) and (5.34) from (G2), we need to expand in $\rho_{i*} \sim v_{i*} \ll 1$ and in $\epsilon \ll 1$. We first expand in $\rho_{i*} \sim v_{i*} \ll 1$,

$$\hat{f}_i = \hat{f}_i^{(0)} + \hat{f}_i^{(1)} + \dots, \tag{G5}$$

where $\hat{f}_i^{(n)} \sim \rho_{i*}^n \hat{f}_i$. For the expansion $\rho_{i*} \sim v_{i*} \ll 1$, we need to use the velocity-space coordinates $v, \lambda = \sin^2 \gamma / B$ and σ . In these coordinates, (G2) becomes

$$\begin{aligned} & \left[v_{\parallel} + \frac{c\phi'_0}{\langle B^2 \rangle_{\text{fs}}} (\mathbf{B} \times \nabla r) \cdot \nabla l \right] \partial \hat{f}_i + \frac{c\phi'_0}{\Psi'_t} \frac{B^2}{\langle B^2 \rangle_{\text{fs}}} \partial_{\alpha} \hat{f}_i \\ & - \frac{\lambda c\phi'_0}{\langle B^2 \rangle_{\text{fs}}} \left[(\hat{\mathbf{b}} \times \nabla r) \cdot \nabla l \partial_l B + \frac{B}{\Psi'_t} \partial_{\alpha} B \right] \partial_{\lambda} \hat{f}_i - \mathcal{L}_{ii}[\hat{f}_i] = -\mathbf{v}_{Mi} \cdot \nabla r \Upsilon_{i f_{Mi}}, \end{aligned} \tag{G6}$$

where $v_{\parallel} = v \cos \gamma = \sigma v \sqrt{1 - \lambda B}$ and

$$\mathcal{L}_{ii}[\hat{f}_i] = \frac{v_{i\perp}(v)}{v^2 B} v_{\parallel} \partial_{\lambda} (v_{\parallel} \lambda \partial_{\lambda} \hat{f}_i). \tag{G7}$$

To lowest order in $\rho_{i*} \ll 1$, (G6) simplifies to $v_{\parallel} \partial \hat{f}_i^{(0)} = 0$. Following the same procedure that we used in § 2, we split $\hat{f}_i^{(0)}$ into a trapped-particle distribution function $\hat{g}_{i,W}(r, \alpha, v, \lambda, t)$ and a passing-particle distribution function $\hat{h}_i(r, v, \lambda, \sigma, t)$. To next order in $\rho_{i*} \ll 1$, (G6) becomes

$$\begin{aligned} & v_{\parallel} \partial \hat{f}_i^{(1)} + \frac{c\phi'_0}{\Psi'_t} \frac{B^2}{\langle B^2 \rangle_{\text{fs}}} \partial_{\alpha} \hat{f}_i^{(0)} - \frac{\lambda c\phi'_0}{\langle B^2 \rangle_{\text{fs}}} \left[(\hat{\mathbf{b}} \times \nabla r) \cdot \nabla l \partial_l B + \frac{B}{\Psi'_t} \partial_{\alpha} B \right] \partial_{\lambda} \hat{f}_i^{(0)} \\ & - \mathcal{L}_{ii}[\hat{f}_i^{(0)}] = -\mathbf{v}_{Mi} \cdot \nabla r \Upsilon_{i f_{Mi}}. \end{aligned} \tag{G8}$$

To eliminate the first term in this equation, we transit average in the trapped-particle region, and we multiply by $B/|v_{\parallel}|$ and flux surface average in the passing-particle region. To perform these operations, we use

$$-\frac{\lambda}{|v_{\parallel}|} \left[(\hat{\mathbf{b}} \times \nabla r) \cdot \nabla l \partial_l B + \frac{B}{\Psi'_t} \partial_{\alpha} B \right] = \frac{2}{v^2} \left[(\hat{\mathbf{b}} \times \nabla r) \cdot \nabla l \partial_l |v_{\parallel}| + \frac{B}{\Psi'_t} \partial_{\alpha} |v_{\parallel}| \right], \tag{G9}$$

and (G4) to write

$$-\left\langle \frac{\lambda B c\phi'_0}{|v_{\parallel}| \langle B^2 \rangle_{\text{fs}}} \left[(\hat{\mathbf{b}} \times \nabla r) \cdot \nabla l \partial_l B + \frac{B}{\Psi'_t} \partial_{\alpha} B \right] \right\rangle_{\text{fs}} = 0 \tag{G10}$$

and

$$-\left\langle \frac{\lambda c\phi'_0}{\langle B^2 \rangle_{\text{fs}}} \left[(\hat{\mathbf{b}} \times \nabla r) \cdot \nabla l \partial_l B + \frac{B}{\Psi'_t} \partial_{\alpha} B \right] \right\rangle_{\tau, W} = \frac{2c\phi'_0}{v^2 \tau_W \Psi'_t \langle B^2 \rangle_{\text{fs}}} \partial_{\alpha} \hat{J}_W, \tag{G11}$$

where \hat{J}_W is defined in (5.52). With these results, the transit average of (G8) becomes

$$\begin{aligned} & \frac{c\phi'_0}{\Psi'_t} \frac{\langle B^2 \rangle_{\tau, W}}{\langle B^2 \rangle_{\text{fs}}} \partial_{\alpha} \hat{g}_{i,W} + \frac{2c\phi'_0}{v^2 \tau_W \Psi'_t \langle B^2 \rangle_{\text{fs}}} \partial_{\alpha} \hat{J}_W \partial_{\lambda} \hat{g}_{i,W} \\ & - \langle \mathcal{L}_{ii}[\hat{g}_{i,W}] \rangle_{\tau, W} = -\langle \mathbf{v}_{Mi} \cdot \nabla r \rangle_{\tau, W} \Upsilon_{i f_{Mi}}, \end{aligned} \tag{G12}$$

for trapped particles. The flux surface average of (G8) multiplied by $B/|v_{\parallel}|$ gives

$$\left\langle \frac{B}{|v_{\parallel}|} \mathcal{L}_{ii}[\hat{h}_i] \right\rangle_{\text{fs}} = 0, \tag{G13}$$

for passing particles.

Equation (G12) can be simplified even further for trapped particles by noting that

$$\partial_\lambda \hat{J}_W = -\frac{v^2 \tau_W}{2} \langle B^2 \rangle_{\tau,W}. \tag{G14}$$

Using this expression, we can rewrite (G12) as

$$\begin{aligned} & \frac{2c\phi'_0}{v^2 \tau_W \Psi'_t \langle B^2 \rangle_{fs}} (-\partial_\lambda \hat{J}_W \partial_\alpha \hat{g}_{i,W} + \partial_\alpha \hat{J}_W \partial_\lambda \hat{g}_{i,W}) \\ & - \langle \mathcal{L}_{ii}[\hat{g}_{i,W}] \rangle_{\tau,W} = - \langle \mathbf{v}_{Mi} \cdot \nabla r \rangle_{\tau,W} \Upsilon f_{Mi}. \end{aligned} \tag{G15}$$

This form of the equation shows that \hat{J}_W is a constant of the motion in DKES. Indeed, if instead of using v and λ as velocity-space coordinates, we use v and \hat{J} , (G15) can be written as

$$\frac{c\phi'_0 \langle B^2 \rangle_{\tau,W}}{\Psi'_t \langle B^2 \rangle_{fs}} \partial_\alpha \hat{g}_{i,W} - \langle \mathcal{L}_{ii}[\hat{g}_{i,W}] \rangle_{\tau,W} = - \langle \mathbf{v}_{Mi} \cdot \nabla r \rangle_{\tau,W} \Upsilon f_{Mi}, \tag{G16}$$

where

$$\langle \mathcal{L}_{ii}[\hat{g}_{i,W}] \rangle_{\tau,W} = \frac{v^2 v_{i,\perp}(v) \langle B^2 \rangle_{\tau,W}}{4} \partial_j \left(\lambda \tau_W^2 \langle B^2 \rangle_{\tau,W} \left\langle \frac{v_{\parallel}^2}{B} \right\rangle_{\tau,W} \partial_j \hat{g}_{i,W} \right). \tag{G17}$$

In the DKES kinetic equation, there will also be junctures of different wells such as the one shown in figure 1. These junctures are determined by values of \hat{J} that depend on r , α , v and the well index W , $\hat{J} = \hat{J}_{c,W}(r, \alpha, v)$. These values of \hat{J} satisfy

$$\hat{J}_{c,I}(r, \alpha, v) + \hat{J}_{c,II}(r, \alpha, v) = \hat{J}_{c,III}(r, \alpha, v), \tag{G18}$$

and imposing that the collisional flux across these junctures is continuous gives the condition

$$\begin{aligned} & \lim_{\hat{J} \rightarrow \hat{J}_{c,I}} \tau_I^2 \langle B^2 \rangle_{\tau,I} \left\langle \frac{v_{\parallel}^2}{B} \right\rangle_{\tau,I} \partial_j \hat{g}_{i,I} + \lim_{\hat{J} \rightarrow \hat{J}_{c,II}} \tau_{II}^2 \langle B^2 \rangle_{\tau,II} \left\langle \frac{v_{\parallel}^2}{B} \right\rangle_{\tau,II} \partial_j \hat{g}_{i,II} \\ & = \lim_{\hat{J} \rightarrow \hat{J}_{c,III}} \tau_{III}^2 \langle B^2 \rangle_{\tau,III} \left\langle \frac{v_{\parallel}^2}{B} \right\rangle_{\tau,III} \partial_j \hat{g}_{i,III}. \end{aligned} \tag{G19}$$

Equations (G13), (G16) and (G19) can be further simplified using the expansion in $\epsilon \ll 1$. Using a procedure similar to the one we follow in § 5.2, we find

$$\hat{g}_{i,W} = \hat{g}_{i,1,W} + \hat{g}_{i,3/2,W} + \hat{g}_{i,2,W} + \dots, \tag{G20}$$

and

$$\hat{h}_i = \hat{h}_{i,3/2} + \dots, \tag{G21}$$

where $\hat{g}_{i,n,W} \sim \epsilon^n f_{Mi}$ and $\hat{h}_{i,n} \sim \epsilon^n f_{Mi}$. Using the fact that $B \simeq B_0$ to lowest order, we obtain that \hat{J} and J are approximately the same coordinate, $\hat{J} \simeq B_0 J$. As a result, (G16) and (G19) become (5.25) and (5.33) to lowest order. As explained in § 5.2, the continuity of derivatives across the trapped–passing boundary implies that the size of \hat{h}_i is $\epsilon^{3/2} f_{Mi}$, and as a result we obtain boundary condition (5.34) for $\hat{g}_{i,1,W}$.

Appendix H. Boundary layer at the juncture of several magnetic wells

In junctures where particles leave two of the wells and enter a third one, a collisional boundary layer appears in the region

$$\frac{|J - J_{c,W}|}{\sqrt{\epsilon} v_{ii} R} \sim \frac{v_{i*}}{\rho_{i*}} \ln \left(\frac{\rho_{i*}}{v_{i*}} \right) \ll 1 \tag{H1}$$

(we justify this estimate below). The distribution function $g_{i,1,W}$ in this boundary layer can be written as

$$g_{i,1,W} = K_{i,junct,W}(r, \alpha, v, \Delta J_{junct}, t) - r_{i,1,W} \mathcal{R} f_{Mi}, \tag{H2}$$

where $\Delta J_{junct} := J - J_{c,W}(r, \alpha, v)$. The function $K_{i,junct,W}(r, \alpha, v, \Delta J_{junct}, t)$ changes rapidly in ΔJ_{junct} , and slowly in α .

We can find the equation for $K_{i,junct,W}$ by substituting (H2) into (5.25) and neglecting small terms

$$-\frac{c\phi'_0}{\Psi'_t} \partial_\alpha J_{c,W} \partial_{\Delta J_{junct}} K_{i,junct,W} - \frac{v^2 v_{ii,\perp} J_{c,W}}{4} \partial_{\Delta J_{junct}} (\tau_{c,W} \partial_{\Delta J_{junct}} K_{i,junct,W}) = 0, \tag{H3}$$

where the logarithmically diverging function

$$\tau_{c,W}(r, \alpha, v, \Delta J_{junct}) := -\hat{\tau}_{c,W,\log}(r, \alpha, v) \ln |\Delta J_{junct}| + \hat{\tau}_{c,W,0}(r, \alpha, v) \tag{H4}$$

is the asymptotic approximation for τ_W near $J = J_{c,W}$. We have neglected the derivative of $K_{i,junct,W}$ with respect to α because it is small compared with the first term in (H3). We can obtain estimate (H1) by balancing the two terms in (H3),

$$\frac{c\phi'_0}{\Psi'_t} \frac{\partial_\alpha J_{c,W}}{\Delta J_{junct}} \sim \frac{v^2 v_{ii,\perp} J_{c,W} \hat{\tau}_{c,W,\log}}{(\Delta J_{junct})^2} \ln |\Delta J_{junct}|. \tag{H5}$$

Indeed, this equation gives

$$|\Delta J_{junct}| \sim \left| \frac{\Psi'_t v^2 v_{ii,\perp} \hat{\tau}_{c,W,\log}}{c\phi'_0 \partial_\alpha \ln J_{c,W}} \right| \ln \left| \frac{c\phi'_0 \partial_\alpha \ln J_{c,W}}{\Psi'_t v^2 v_{ii,\perp} \hat{\tau}_{c,W,\log}} \right| \sim \frac{v_{i*}}{\rho_{i*}} \ln \left(\frac{\rho_{i*}}{v_{i*}} \right). \tag{H6}$$

For the order of magnitude estimate, we have used $\phi'_0 \sim T_i/ea$, $\Psi'_t \sim aB_0$, $J_{c,W} \sim \sqrt{\epsilon} v_{ii} R$ and $\hat{\tau}_{c,W,\log} \sim R/\sqrt{\epsilon} v_{ii}$. With (H6), we can simplify $\tau_{c,W}$ to

$$\tau_{c,W} \simeq \hat{\tau}_{c,W,\log} \ln \left| \frac{c\phi'_0 \partial_\alpha \ln J_{c,W}}{\Psi'_t v^2 v_{ii,\perp} \hat{\tau}_{c,W,\log}} \right|. \tag{H7}$$

From here on, we use this approximation that makes $\tau_{c,W}$ independent of ΔJ_{junct} .

Equation (H3) has to be solved in conjunction with continuity of $K_{i,\text{junct},W}$ across the juncture and along with condition (5.33), which in the layer becomes

$$\begin{aligned}
 & J_{c,I} \tau_{c,I} \lim_{\Delta J_{\text{junct}} \rightarrow 0} \partial_{\Delta J_{\text{junct}}} K_{i,\text{junct},I} + J_{c,II} \tau_{c,II} \lim_{\Delta J_{\text{junct}} \rightarrow 0} \partial_{\Delta J_{\text{junct}}} K_{i,\text{junct},II} \\
 & = J_{c,III} \tau_{c,III} \lim_{\Delta J_{\text{junct}} \rightarrow 0} \partial_{\Delta J_{\text{junct}}} K_{i,\text{junct},III}.
 \end{aligned}
 \tag{H8}$$

For large ΔJ_{junct} , the function $K_{i,\text{junct},W}(r, \alpha, v, \Delta J_{\text{junct}}, t)$ must tend to the solutions outside of the layer,

$$\lim_{|\Delta J_{\text{junct}}| \rightarrow \infty} K_{i,\text{junct},W}(r, \alpha, v, \Delta J_{\text{junct}}, t) = K_{i,W}(r, \alpha, v, J_{c,W}, t).
 \tag{H9}$$

With boundary conditions (H9), the solutions to (H3) are

$$\begin{aligned}
 & K_{i,\text{junct},W}(r, \alpha, v, \Delta J_{\text{junct}}, t) = K_{i,W}(r, \alpha, v, J_{c,W}, t) \\
 & + A_W(r, \alpha, v, t) \exp\left(-\frac{4c\phi'_0 \partial_\alpha \ln J_{c,W}}{\Psi'_t v^2 v_{ii,\perp} \tau_{c,W}} \Delta J_{\text{junct}}\right),
 \end{aligned}
 \tag{H10}$$

where the function $A_W(r, \alpha, v, t)$ is a constant of integration.

We proceed to discuss how to use solution (H10) to solve the problem. We consider a juncture like the one in figure 1 such that particles leave well I and II to enter well III, that is, $-(c\phi'_0/\Psi'_t) \partial_\alpha J_{c,W} > 0$ for $W = I, II, III$. This means that the exponential in solution (H10) diverges for well III, where $\Delta J_{\text{junct}} > 0$, and thus the solution in well III can tend to $K_{i,III}(r, \alpha, v, J_{c,III}, t)$ only if $A_{III}(r, \alpha, v, t) = 0$. As a result, the derivative of $K_{i,\text{junct},III}$ with respect to ΔJ_{junct} vanishes, and condition (H8) simply gives

$$J_{c,I} \tau_{c,I} \lim_{\Delta J_{\text{junct}} \rightarrow 0} \partial_{\Delta J_{\text{junct}}} K_{i,\text{junct},I} + J_{c,II} \tau_{c,II} \lim_{\Delta J_{\text{junct}} \rightarrow 0} \partial_{\Delta J_{\text{junct}}} K_{i,\text{junct},II} = 0.
 \tag{H11}$$

Imposing this condition and the fact that $K_{i,\text{junct},I}$ and $K_{i,\text{junct},II}$ have to be equal to each other at $\Delta J_{\text{junct}} = 0$ to be continuous with $K_{i,\text{junct},III}$, we find the functions A_I and A_{II} . The final solutions are

$$\begin{aligned}
 & K_{i,\text{junct},I}(r, \alpha, v, \Delta J_{\text{junct}}, t) = K_{i,I}(r, \alpha, v, J_{c,I}, t) \\
 & + \frac{\partial_\alpha J_{c,II} [K_{i,II}(r, \alpha, v, J_{c,II}, t) - K_{i,I}(r, \alpha, v, J_{c,I}, t)]}{\partial_\alpha J_{c,I} + \partial_\alpha J_{c,II}} \\
 & \times \exp\left(-\frac{4c\phi'_0 \partial_\alpha \ln J_{c,I}}{\Psi'_t v^2 v_{ii,\perp} \tau_{c,I}} \Delta J_{\text{junct}}\right),
 \end{aligned}
 \tag{H12}$$

and

$$\begin{aligned}
 & K_{i,\text{junct},II}(r, \alpha, v, \Delta J_{\text{junct}}, t) = K_{i,II}(r, \alpha, v, J_{c,II}, t) \\
 & + \frac{\partial_\alpha J_{c,I} [K_{i,I}(r, \alpha, v, J_{c,I}, t) - K_{i,II}(r, \alpha, v, J_{c,II}, t)]}{\partial_\alpha J_{c,I} + \partial_\alpha J_{c,II}} \\
 & \times \exp\left(-\frac{4c\phi'_0 \partial_\alpha \ln J_{c,II}}{\Psi'_t v^2 v_{ii,\perp} \tau_{c,II}} \Delta J_{\text{junct}}\right).
 \end{aligned}
 \tag{H13}$$

With these solutions, we can find $K_{i,III}(r, \alpha, v, J_{c,III}, t) = K_{i,\text{junct},III}(r, \alpha, v, 0, t)$ because $K_{i,\text{junct},III}(r, \alpha, v, 0, t)$ has to be equal to the values $K_{i,\text{junct},I}$ and $K_{i,\text{junct},II}$ at $\Delta J_{\text{junct}} = 0$. The

solution for $K_{i,III}(r, \alpha, v, J_{c,III}, t)$ is the same one that we found by applying conservation of particles, given in (7.8).

It is worth noting that if we apply this method to a hypothetical boundary layer in a juncture where particles leave one well to enter two other wells, we find that the exponential in solution (H10) diverges in two of the wells (the two that the particles enter), and as a result, the functions A_W have to vanish in those wells. One ends up finding that the solution is constant in all the wells, recovering the result that $K_{i,W}$ is constant across this kind of juncture.

Appendix I. Boundary layers around discontinuities in $\partial_J K_{i,W}$

There are two types of boundary layers that form around discontinuities in $\partial_J K_{i,W}$: the layers on the boundaries of the regions where $K_{i,W}$ is independent of J , and the layer on the trapped–passing boundary. We describe both types of layers in this appendix.

1.1. Boundary layers on the boundaries of the regions where $\partial_J K_{i,W} = 0$

To discuss the boundary layers that appear on the boundaries of the regions where $K_{i,W}$ does not depend on J , we use the example of the layer that arises around $J = J_{c,III,M}$ in the example discussed in § 7.1.2. We use the variable $\Delta J_{bl} := J - J_{c,III,M}$. The typical size of ΔJ_{bl} is

$$\frac{\Delta J_{bl}}{\sqrt{\epsilon v_{ii} R}} \sim \sqrt{\frac{v_{i*}}{\rho_{i*}}} \ll 1, \tag{I1}$$

as we will show below.

Particles behave differently for ΔJ_{bl} positive or negative. For $\Delta J_{bl} > 0$, we can write the distribution function in the boundary layer region as

$$g_{i,1,III} = K_{i,III} - r_{i,1,III} \Upsilon f_{Mi} + g_{i,1,III}^{(bl)}(r, \alpha, v, \Delta J_{bl}, t) + O\left(\frac{v_{i*}}{\rho_{i*}} \epsilon f_{Mi}\right), \tag{I2}$$

where

$$\frac{g_{i,1,III}^{(bl)}}{\epsilon f_{Mi}} \sim \sqrt{\frac{v_{i*}}{\rho_{i*}}} \ll 1. \tag{I3}$$

Substituting (I2) into (5.25), we obtain the equation for the boundary layer for $\Delta J_{bl} > 0$,

$$\frac{c\phi'_0}{\Psi'_t} \partial_\alpha g_{i,1,III}^{(bl)} - \frac{v^2 v_{ii,\perp} J_{c,III,M} \tau_{III}}{4} \partial_{\Delta J_{bl} \Delta J_{bl}}^2 g_{i,1,III}^{(bl)} = 0, \tag{I4}$$

where τ_{III} is evaluated at $J = J_{c,III,M}$. For $J > J_{c,III,M}$, particles move uninterruptedly along α . Thus, we impose periodic boundary conditions in α for $g_{i,1,III}^{(bl)}$. We also impose

$$\lim_{\Delta J_{bl} \rightarrow \infty} g_{i,1,III}^{(bl)}(r, \alpha, v, \Delta J_{bl}, t) = 0 \tag{I5}$$

and continuity and continuity of derivatives with $g_{i,1,III}$ for $J < J_{c,III,M}$.

The distribution function in the region $J < J_{c,III,M}$ is simply $g_{i,1,III} = K_{i,III} - r_{i,1,III} \Upsilon f_{Mi} + g_{i,1,III}^{(1)}$, where $g_{i,1,III}^{(1)}$ is determined by (7.23). In other words, unlike for $J > J_{c,III,M}$, we do not distinguish between $g_{i,1,III}^{(1)}$ and the boundary layer piece of the

distribution function because the equation for $g_{i,1,III}^{(1)}$ contains collisions. As a result, the continuity of $g_{i,1,III}$ across $J = J_{c,III,M}$ gives

$$g_{i,1,III}^{(bl)}(r, \alpha, v, 0, t) = g_{i,1,III}^{(1)}(r, \alpha, v, J_{c,III,M}, t), \tag{16}$$

and the continuity of the derivative with respect to J gives

$$\partial_J K_{i,III}(r, v, J_{c,III,M}^+, t) + \partial_{\Delta J_{bl}} g_{i,1,III}^{(bl)}(r, \alpha, v, 0, t) = \partial_J g_{i,1,III}^{(1)}(r, \alpha, v, J_{c,III,M}, t), \tag{17}$$

where $\partial_J K_{i,III}(r, v, J_{c,III,M}^+, t)$ is the derivative of $K_{i,III}$ with respect to J evaluated at a value of J slightly above $J = J_{c,III,M}$. Recall that $\partial_J K_{i,III}$ vanishes for $J < J_{c,III,M}$.

By balancing the different terms in (14), we obtain estimate (II) for the boundary layer width. The estimate for the size of $g_{i,1,III}^{(bl)}$ in (I3) is obtained from the fact that $g_{i,1,III}^{(bl)}$ is the integral of $\partial_{\Delta J_{bl}} g_{i,1,III}^{(bl)}$ over an interval of ΔJ_{bl} of the size given in (II). The size of $\partial_{\Delta J_{bl}} g_{i,1,III}^{(bl)}$, $\partial_{\Delta J_{bl}} g_{i,1,III}^{(bl)} \sim \sqrt{\epsilon} f_{Mi} / v_{ii} R$, is determined by boundary condition (I7).

Equations (16) and (17) are also the boundary conditions for $g_{i,1,III}^{(1)}$ in the region where $\partial_J K_{i,III}$ vanishes. Note that this means that $g_{i,1,III}^{(1)}$ is larger than expected by a factor of $\sqrt{\rho_{i*} / v_{i*}}$ near $J = J_{c,III,M}$, but it becomes of order $(v_{i*} / \rho_{i*}) \epsilon f_{Mi}$ for $(J_{c,III,M} - J) / \sqrt{\epsilon} v_{ii} R \gg \sqrt{v_{i*} / \rho_{i*}}$, except for some regions of small phase-space volume that we proceed to discuss.

In the example that we are considering (see figures 6 and 7), $g_{i,1,III}^{(1)}$ at $J = J_{c,III,M}$ for α slightly smaller than π ($\alpha = \pi^-$) is different from $g_{i,1,III}^{(1)}$ at $J = J_{c,III,M}$ for α slightly larger than π ($\alpha = \pi^+$). The particles with $J = J_{c,III,M}$ at $\alpha = \pi^+$ are particles that used to have a second adiabatic invariant close to $J = J_{c,II,M}$ and hence their number is determined by the value of $g_{i,1,II}^{(1)}$ around $J = J_{c,II,M}$ and $\alpha = \pi^-$. The particles with $J = J_{c,III,M}$ at $\alpha = \pi^-$ transition into well I where they have a second adiabatic invariant very similar to $J_{c,I,M}$. This means that $g_{i,1,I}^{(1)}$ around $J = J_{c,I,M}$ and $\alpha = \pi^+$ is determined by $g_{i,1,III}^{(1)}$ at $J = J_{c,III,M}$ and $\alpha = \pi^-$, and hence is large by a factor of $\sqrt{\rho_{i*} / v_{i*}}$ in a region of width $\sqrt{\epsilon} v_{i*} / \rho_{i*} v_{ii} R$ around $J_{c,I,M}$. The particles in this region around $J_{c,I,M}$, in turn, transition back into well III at some value of α between π and 2π . This means that there is another region around another value of J that connects to $J_{c,I,M}$ where $g_{i,1,III}^{(1)}$ is larger than expected. These regions of large $g_{i,1,W}^{(1)}$ go on until $g_{i,1,W}^{(1)}$ becomes sufficiently small due to diffusion in J .

Boundary condition (I7) gives the velocity-space flux conservation condition (7.16). Indeed, integrating (7.23) in α and in J for the juncture sketched in figures 6 and 7, we find

$$\begin{aligned} & -\frac{v^2 v_{ii,\perp}}{4} \left[J_{c,III,M} \lim_{J \rightarrow J_{c,III,M}^+} \left\langle \tau_{III} \partial_J g_{i,1,III}^{(1)} \right\rangle_\alpha - J_{c,I,M} \lim_{J \rightarrow J_{c,I,M}^-} \left\langle \tau_I \partial_J g_{i,1,I}^{(1)} \right\rangle_\alpha \right. \\ & \left. - J_{c,II,M} \lim_{J \rightarrow J_{c,II,M}^-} \left\langle \tau_{II} \partial_J g_{i,1,II}^{(1)} \right\rangle_\alpha \right] = \frac{m_i v^2 v_{ii,\perp}}{4 Z_i e \phi'_0} (J_{c,III,M} - J_{c,I,M} - J_{c,II,M}) \Upsilon_i f_{Mi}. \end{aligned} \tag{18}$$

Using boundary condition (I7), the definition of $r_{i,1,W}$ in (5.26) and $\partial_J \lambda_W \simeq -2/v^2 B_0 \tau_W$, (18) becomes

$$\begin{aligned} & J_{c,III,M} \lim_{J \rightarrow J_{c,III,M}^+} \left\langle \tau_{III} (\partial_J K_{i,III} - \partial_J r_{i,1,III} \Upsilon_i f_{Mi}) \right\rangle_\alpha \\ & - J_{c,I,M} \lim_{J \rightarrow J_{c,I,M}^-} \left\langle \tau_I (\partial_J K_{i,I} - \partial_J r_{i,1,III} \Upsilon_i f_{Mi}) \right\rangle_\alpha \end{aligned}$$

$$\begin{aligned}
 & - J_{c,II,m} \lim_{J \rightarrow J_{c,II,m}} \left\langle \tau_{II} \left(\partial_J K_{i,II} - \partial_J r_{i,1,II} \Upsilon_{fMi} \right) \right\rangle_\alpha \\
 & = - J_{c,III,M} \lim_{\Delta J_{bl} \rightarrow 0} \left\langle \tau_{III} \partial_{\Delta J_{bl}} g_{i,1,III}^{\{bl\}} \right\rangle_\alpha + J_{c,I,m} \lim_{\Delta J_{bl} \rightarrow 0} \left\langle \tau_I \partial_{\Delta J_{bl}} g_{i,1,I}^{\{bl\}} \right\rangle_\alpha \\
 & \quad + J_{c,II,m} \lim_{\Delta J_{bl} \rightarrow 0} \left\langle \tau_{II} \partial_{\Delta J_{bl}} g_{i,1,II}^{\{bl\}} \right\rangle_\alpha.
 \end{aligned} \tag{19}$$

We obtain the velocity-space flux conservation condition (7.16) from this expression by employing the fact that (14) and (15) and the equivalent equations for the boundary layers at $J = J_{c,I,m}$ and $J_{c,II,m}$ give

$$\lim_{\Delta J_{bl} \rightarrow 0} \langle \tau_W \partial_{\Delta J_{bl}} g_{i,1,W}^{\{bl\}} \rangle_\alpha = 0. \tag{110}$$

Incidentally, τ_{III} in (14) diverges logarithmically at $\alpha = \pi$, where $J_{c,III} = J_{c,III,M}$. This divergence is integrable and does not affect the previous discussion.

1.2. Boundary layer at the trapped–passing boundary

The collisional layer that appears at the trapped–passing boundary is a result of the lowest-order solution (7.19) not satisfying property (5.36) at the trapped–passing boundary. This layer is best described in the coordinate $\Delta\lambda := \lambda - 1/B_M \simeq \lambda - 1/B_0 + B_{1,M}/B_0^2$, which is of order

$$\frac{|\Delta\lambda|B_0}{\epsilon} \sim \sqrt{\frac{v_{i*}}{\rho_{i*}} \frac{1}{\ln(\rho_{i*}/v_{i*})}} \ll 1 \tag{111}$$

(we will argue that this estimate is correct below). We consider $g_{i,1,W}^{\{1\}}$, determined by (7.23), as a function of $\Delta\lambda$. We find below that in the region of velocity space that satisfies (111), $g_{i,1,W}^{\{1\}}$ is not of order $(v_{i*}/\rho_{i*})\epsilon_{fMi}$, but much larger,

$$\frac{g_{i,1,W}^{\{1\}}}{\epsilon_{fMi}} \sim \sqrt{\frac{v_{i*}}{\rho_{i*}} \frac{1}{\ln(\rho_{i*}/v_{i*})}} \ll 1. \tag{112}$$

Using estimates (111) and (112) in (7.23), we find

$$\frac{c\phi'_0}{\Psi'_I} (\partial_\alpha g_{i,1,W}^{\{1\}} + \partial_\alpha \lambda_W \partial_{\Delta\lambda} g_{i,1,W}^{\{1\}}) - \frac{v^2 v_{ii,\perp}}{4} \tau_W J_W (\partial_J \lambda_W)^2 \partial_{\Delta\lambda}^2 g_{i,1,W}^{\{1\}} = 0. \tag{113}$$

Here, we can use $\partial_J \lambda_W \simeq -2/v^2 B_0 \tau_W$ and $\partial_\alpha \lambda_W \simeq -B_0^{-2} \langle \partial_\alpha B_1 \rangle_{\tau,W}$ to simplify (113) to

$$\frac{c\phi'_0}{\Psi'_I} \left(\partial_\alpha g_{i,1,W}^{\{1\}} - \frac{\langle \partial_\alpha B_1 \rangle_{\tau,W}}{B_0^2} \partial_{\Delta\lambda} g_{i,1,W}^{\{1\}} \right) - \frac{v_{ii,\perp} J_W}{v^2 B_0^2 \tau_W} \partial_{\Delta\lambda}^2 g_{i,1,W}^{\{1\}} = 0. \tag{114}$$

We need to evaluate the different coefficients in this equation for small $\Delta\lambda$. We will use the property that, for $\Delta\lambda \rightarrow 0$, the particle trajectory covers the whole flux surface in such a way that the average of any function over the length of the line is related to the flux

surface average of the function by

$$\lim_{\Delta\lambda \rightarrow 0} \frac{\int_{l_{bL,W}}^{l_{bR,W}} (\dots) dl}{l_{bR,W} - l_{bL,W}} = \frac{\langle B(\dots) \rangle_{fs}}{\langle B \rangle_{fs}}. \tag{I15}$$

Then, we can write

$$\frac{J_W}{\tau_W} \simeq \frac{\langle B|v_{\parallel, tp}| \rangle_{fs}}{\langle B|v_{\parallel, tp}|^{-1} \rangle_{fs}} \simeq \frac{\langle |v_{\parallel, tp}| \rangle_{fs}}{\langle |v_{\parallel, tp}|^{-1} \rangle_{fs}}, \tag{I16}$$

where $v_{\parallel, tp}$ is the parallel velocity of the particles at the trapped–passing boundary. Note that the average $\langle |v_{\parallel, tp}|^{-1} \rangle_{fs}$ does not diverge logarithmically despite $|v_{\parallel, tp}|$ depending linearly on l near the maximum of B_1 because the integral over α eliminates divergence. With this result, (I14) simplifies to

$$\frac{c\phi'_0}{\Psi'_t} \left(\partial_\alpha g_{i,1,W}^{\{1\}} - \frac{\langle \partial_\alpha B_1 \rangle_{\tau,W}}{B_0^2} \partial_{\Delta\lambda} g_{i,1,W}^{\{1\}} \right) - \frac{v_{ii,\perp} \langle |v_{\parallel, tp}| \rangle_{fs}}{v^2 B_0^2 \langle |v_{\parallel, tp}|^{-1} \rangle_{fs}} \partial_{\Delta\lambda}^2 g_{i,1,W}^{\{1\}} = 0. \tag{I17}$$

To obtain an estimate for $\langle \partial_\alpha B_1 \rangle_{\tau,W}$, we need to consider the integral

$$\langle \partial_\alpha B_1 \rangle_{\tau,W} \simeq \frac{2}{v\tau_W} \int_{l_{bL,0,W}}^{l_{bR,0,W}} \frac{\partial_\alpha B_1(l)}{\sqrt{(B_{1,M} - B_1(l))/B_0 - \Delta\lambda B_0}} dl. \tag{I18}$$

We first note that, for $\Delta\lambda \rightarrow 0$, this integral vanishes,

$$\begin{aligned} \lim_{\Delta\lambda \rightarrow 0} \langle \partial_\alpha B_1 \rangle_{\tau,W} &\simeq \frac{2}{v \int_0^{2\pi} d\alpha \int_0^{L_0} dl |v_{\parallel, tp}|^{-1}} \int_0^{2\pi} d\alpha \int_0^{L_0} dl \frac{\partial_\alpha B_1(l)}{\sqrt{(B_{1,M} - B_1(l))/B_0}} \\ &= -\frac{4B_0}{v \int_0^{2\pi} d\alpha \int_0^{L_0} dl |v_{\parallel, tp}|^{-1}} \int_0^{2\pi} d\alpha \partial_\alpha \left(\int_0^{L_0} \sqrt{(B_{1,M} - B_1(l))/B_0} dl \right) = 0. \end{aligned} \tag{I19}$$

Here, we have used the fact that the length L , defined below (2.31), is independent of α to lowest order in the aspect ratio expansion, $L(r, \alpha) \simeq L_0(r)$. We proceed to determine how integral (I18) goes to zero for small $\Delta\lambda$. A sketch of the integral path is shown in figure 15: the path of integration is composed of an almost surface-covering piece and a piece of short length in the region $\alpha_a < \alpha < \alpha_b$, where α_a and α_b are the values of α for which the maximum of B on the line α is equal to $1/\lambda$. For most values of α and $\Delta\lambda$, the bulk of the integral is the almost surface-covering piece, and its size can be estimated by replacing the line integral by an integral over the surface that the line covers, shaded in figure 15,

$$\begin{aligned} \langle \partial_\alpha B_1 \rangle_{\tau,W} &\sim \frac{1}{v \int_0^{2\pi} d\alpha \int_0^{L_0} dl |v_{\parallel, tp}|^{-1}} \left[\int_0^{\alpha_a} d\alpha \int_0^{L_0} dl \frac{\partial_\alpha B_1(l)}{\sqrt{(B_{1,M} - B_1(l))/B_0 - \Delta\lambda B_0}} \right. \\ &\quad \left. + \int_{\alpha_b}^{2\pi} d\alpha \int_0^{L_0} dl \frac{\partial_\alpha B_1(l)}{\sqrt{(B_{1,M} - B_1(l))/B_0 - \Delta\lambda B_0}} \right]. \end{aligned} \tag{I20}$$

The value of this surface integral is dominated by the region near the maximum of B_1 because the rest of surface integral vanishes as shown by (I19). To avoid cluttering

notation, we choose α and l such that $\alpha = 0$ and $l = 0$ at the maximum of B_1 , that is, $B_{1,M} = B_1(r, 0, 0)$. Then, around the maximum,

$$B_1(r, \alpha, l) \simeq B_{1,M} + \frac{1}{2}l^2 \partial_{ll}^2 B_1(r, 0, 0) + \alpha l \partial_{\alpha l}^2 B_1(r, 0, 0) + \frac{1}{2}\alpha^2 \partial_{\alpha\alpha}^2 B_1(r, 0, 0), \tag{I21}$$

with $\partial_{ll}^2 B_1(r, 0, 0) < 0$, $\partial_{\alpha\alpha}^2 B_1(r, 0, 0) < 0$ and

$$\partial_{ll}^2 B_1(r, 0, 0) \partial_{\alpha\alpha}^2 B_1(r, 0, 0) - [\partial_{\alpha l}^2 B_1(r, 0, 0)]^2 > 0. \tag{I22}$$

This Taylor expansion around the maximum $l = 0$ and $\alpha = 0$ implies that $\alpha_a \sim \alpha_b \sim \sqrt{\epsilon^{-1} \Delta \lambda B_0}$, and that the integrals over l in (I20) are of order

$$\int_0^L \frac{\partial_\alpha B_1(l)}{\sqrt{(B_{1,M} - B_1(l))/B_0 - \Delta \lambda B_0}} dl \sim \sqrt{\epsilon} B_0 R \alpha |\ln(\alpha^2 + \epsilon^{-1} \Delta \lambda B_0)| \tag{I23}$$

for small α . Thus, we find

$$\langle \partial_\alpha B_1 \rangle_{\tau, W} \sim B_0^2 \Delta \lambda \ln \left(\frac{\epsilon}{B_0 \Delta \lambda} \right). \tag{I24}$$

This is only an estimate. The real value and sign of $\langle \partial_\alpha B_1 \rangle_{\tau, W}$ depends strongly on the particular orbit under consideration, determined by its α . Before we proceed, we point out that it is important not to confuse the α shown in figure 15, where a single orbit samples several values of α , with the value of α that we assign to the orbit (for example, we can assign to an orbit the value of α where it has its left bounce point). Going back to the dependence of $\langle \partial_\alpha B_1 \rangle_{\tau, W}$ on the particular orbit, consider what happens as the bounce points move from α_a to α_b . When a bounce point approaches α_a , the part of the path near the bounce, which we ignored in our estimate (I24), becomes dominant due to a logarithmic divergence at the bounce point. As a result, $\langle \partial_\alpha B_1 \rangle_{\tau, W}$ is the value of $\partial_\alpha B_1$ at the bounce point, which is of order $\alpha_a \partial_{\alpha\alpha}^2 B_1(r, 0, 0) \sim \sqrt{\epsilon \Delta \lambda B_0^{3/2}}$ – importantly, note that the logarithmic divergence can only dominate when the bounce point is very close to α_b , that is, close by $\exp(-R/(l_{bR,0,W} - l_{bL,0,W})) \ll 1$, and hence our estimate (I24) is valid for most orbits. The situation for a bounce point in α_b is similar, but in this case the sign of $\partial_\alpha B_1$ at the bounce point is the opposite to the one that we consider above, showing that $\langle \partial_\alpha B_1 \rangle_{\tau, W}$ changes sign.

Using estimate (I24) in (I17), we find the estimate (I11) for the width of the layer. The estimate for the size of $g_{i,1,W}^{(1)}$ in (I12) is obtained from the estimate for the width of the layer and the fact that the derivative of $g_{i,1,W}^{(1)}$ with respect to $\Delta \lambda$ must be such that property (5.36) is satisfied. Note, however, that property (5.36) is not imposed on the equation as a boundary condition. The boundary condition at $\Delta \lambda = 0$ is

$$g_{i,1,W}^{(1)}(r, \alpha, v, 0, t) = 0. \tag{I25}$$

The phase-space flows along α and across J at large values of $\Delta \lambda$ are the ones that ensure that property (5.36) is satisfied. These phase-space flows, determined by (7.23) and boundary conditions (I6) and (I7), are the boundary condition that we need to impose for large $\Delta \lambda$ to solve (I17).

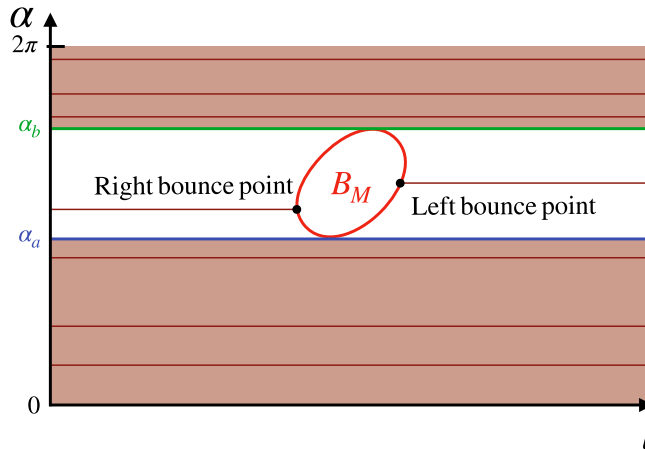


FIGURE 15. Path of integration of the integral in (I18) in the (l, α) plane (thin dark pink line), and area of integration for the integral in (I20) (shaded region). For large J , λ must be close to $1/B_M$, and hence it must have bounce points at a value of B close to B_M . In the figure, we sketch the contour $B = 1/\lambda \approx B_M$ as a red line (we have assumed that there is only one maximum of B). The lines α_a and α_b are marked in blue and green, respectively.

REFERENCES

- BEIDLER, C.D., ALLMAIER, K., ISAEV, M.Y., KASILOV, S.V., KERNBLICHLER, W., LEITOLD, G.O., MAAßBERG, H., MIKKELSEN, D.R., MURAKAMI, S., SCHIMDT, M., *et al.* 2011 Benchmarking of the mono-energetic transport coefficients – results from the International Collaboration on Neoclassical Transport in Stellarators (ICNTS). *Nucl. Fusion* **51**, 076001.
- BEIDLER, C.D., HITCHON, W.N.G. & SHOHEIT, J.L. 1987 “Hybrid” Monte Carlo simulation of ripple transport in stellarators. *J. Comput. Phys.* **72**, 220.
- BOOZER, A.H. & GARDNER, H.J. 1990 The bootstrap current in stellarators. *Phys. Fluids B* **2**, 2408.
- BRAGINSKII, S.I. 1958 Transport phenomena in a completely ionized two-temperature plasma. *Sov. Phys. JETP* **6**, 358.
- CALVO, I., PARRA, F.I., ALONSO, J.A. & VELASCO, J.L. 2014 Optimizing stellarators for large flows. *Plasma Phys. Control. Fusion* **56**, 094003.
- 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. Control. Fusion* **59**, 055014.
- 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**, 905840407.
- CARY, J.R., ESCANDE, D.F. & TENNYSON, J.L. 1986 Adiabatic-invariant change due to separatrix crossing. *Phys. Rev. A* **34**, 4256.
- CARY, J.R. & SHASHARINA, S.G. 1997a Helical plasma confinement devices with good confinement properties. *Phys. Rev. Lett.* **78**, 674.
- CARY, J.R. & SHASHARINA, S.G. 1997b Omnigenity and quasihelicity in helical plasma confinement systems. *Phys. Plasmas* **4**, 3323.
- CATTO, P.J. 2019 Collisional alpha transport in a weakly non-quasisymmetric stellarator magnetic field. *J. Plasma Phys.* **85**, 905850213.
- DINKLAGE, A., BEIDLER, C.D., HELANDER, P., FUCHERT, G., MAAßBERG, H., RAHBARNIA, K., PEDERSEN, T.S., TURKIN, Y., WOLF, R.C., ALONSO, A., *et al.* 2018 Magnetic configuration effects on the Wendelstein 7-X stellarator. *Nat. Phys.* **14**, 855.
- DINKLAGE, A., YOKOYAMA, M., TANAKA, K., VELASCO, J.L., LÓPEZ-BRUNA, D., BEIDLER, C.D., SATAKE, S., ASCASÍBAR, E., ARÉVALO, J., BALDZUHN, J., *et al.* 2013 Inter-machine validation

- study of neoclassical transport modelling in medium- to high-density stellarator-heliotron plasmas. *Nucl. Fusion* **53**, 063022.
- DOBROTT, D. & GREENE, J.M. 1971 Probability of trapping-state transitions in a toroidal device. *Phys. Fluids* **14**, 1525.
- GALIEEV, A.A., SADGEEV, R.Z., FURTH, H.P. & ROSENBLUTH, M.N. 1969 Plasma diffusion in a toroidal stellarator. *Phys. Rev. Lett.* **22**, 511.
- GARREN, D.A. & BOOZER, A.H. 1991 Existence of quasihelically symmetric stellarators. *Phys. Fluids B* **3**, 2822.
- GOLDSTON, R.J., WHITE, R.B. & BOOZER, A.H. 1981 Confinement of high-energy trapped particles in tokamaks. *Phys. Rev. Lett.* **47**, 647.
- HAZELTINE, R.D. 1973 Recursive derivation of drift-kinetic equation. *Plasma Phys.* **15**, 77.
- HAZELTINE, R.D. & CATTO, P.J. 1981 Bumpy torus transport in the low collision frequency limit. *Phys. Fluids* **24**, 290.
- HINTON, F.L. & HAZELTINE, R.D. 1976 Theory of plasma transport in toroidal confinement systems. *Rev. Mod. Phys.* **48**, 239.
- HIRSHMAN, S.P., SHANG, K.C., VAN RIJ, W.I. JR., BEASLEY, C.O. JR. & CRUME, E.C. 1986 Plasma transport coefficients for nonsymmetric toroidal confinement systems. *Phys. Fluids* **29**, 2951.
- HO, D.D.M. & KULSRUD, R.M. 1987 Neoclassical transport in stellarators. *Phys. Fluids* **30**, 442.
- HUANG, B., SATAKE, S., KANNO, R., SUGAMA, H. & MATSUOKA, S. 2017 Benchmark of the local drift-kinetic models for neoclassical transport simulation in helical plasmas. *Phys. Plasmas* **24**, 022503.
- KOVRIZHNYKH, L.M. 1984 Neoclassical theory of transport processes in toroidal magnetic confinement systems, with emphasis on non-axisymmetric configurations. *Nucl. Fusion* **24**, 851.
- MATSUOKA, S., SATAKE, S., KANNO, R. & SUGAMA, H. 2015 Effects of magnetic drift tangential to magnetic surfaces on neoclassical transport in non-axisymmetric plasmas. *Phys. Plasmas* **22**, 072511.
- MYNICK, H.E. 1983 Effect of collisionless detrapping on nonaxisymmetric transport in a stellarator with radial electric field. *Phys. Fluids* **26**, 2609.
- NEMOV, V.V., KASILOV, S.V., KERNBICHLER, W. & HEYN, M.F. 1999 Evaluation of $1/\nu$ neoclassical transport in stellarators. *Phys. Plasmas* **6**, 4622.
- PARRA, F.I., CALVO, I., HELANDER, P. & LANDREMAN, M. 2015 Less constrained omnigenous stellarators. *Nucl. Fusion* **55**, 033005.
- PAUL, E.J., LANDREMAN, M., POLI, F.M., SPONG, D.A., SMITH, H.M. & DORLAND, W. 2017 Rotation and neoclassical ripple transport in ITER. *Nucl. Fusion* **57**, 116044.
- ROSENBLUTH, M.N., MACDONALD, W.M. & JUDD, D.L. 1957 Fokker-Planck equation for an inverse-square force. *Phys. Rev.* **107**, 1.
- SATAKE, S., OKAMOTO, M., NAKAJIMA, N., SUGAMA, H. & YOKOYAMA, M. 2006 Non-local simulation of the formation of neoclassical ambipolar electric field in non-axisymmetric configurations. *Plasma Fusion Res.* **1**, 002.
- SUGAMA, H., MATSUOKA, S., SATAKE, S. & KANNO, R. 2016 Radially local approximation of the drift kinetic equation. *Phys. Plasmas* **23**, 042502.
- VELASCO, J.L., CALVO, I., PARRA, F.I., D'HERBEMONT, V., SMITH, H.M., CARRALERO, D. Estrada, T. & THE W7-X TEAM 2021 Fast simulations for large aspect ratio stellarators with the neoclassical code KNOSOS. *Nucl. Fusion* **61**, 116013.
- VELASCO, J.L., CALVO, I., PARRA, F.I. & GARCÍA-REGAÑA, J.M. 2020 KNOSOS: a fast orbit-averaging neoclassical code for stellarator geometry. *J. Comput. Phys.* **418**, 109512.