Angular momentum and rotational energy of mean flows in toroidal magnetic fields

OPEN ACCESS
IOP Publishing | International Atomic Energy Agency
Nuclear Fusion
Nucl. Fusion **60** (2020) 096018 (26pp)
https://doi.org/10.1088/1741-4326/ab9fa8

Angular momentum and rotational energy of mean flows in toroidal magnetic fields

M. Wiesenberger[1] **and M. Held**[2]

[1] Department of Physics, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
[2] Department of Space, Earth and Environment, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

E-mail: mattwi@fysik.dtu.dk

Received 5 March 2020, revised 15 June 2020
Accepted for publication 24 June 2020
Published 10 August 2020

Abstract

We derive the balance equation for the Favre averaged angular momentum in toroidal not necessarily axisymmetric magnetic field equilibria. We find that the components of angular momentum are given by the covariant poloidal and toroidal components of $\boldsymbol{E}\times\boldsymbol{B}$ and parallel flow velocities and we separately identify all relevant stress tensors, torques and source terms for each of these components. Our results feature the Favre stress generalisations of previously found Reynolds stresses like the diamagnetic or parallel $\boldsymbol{E}\times\boldsymbol{B}$ stress, as well as the density gradient drive term. Further, we identify the magnetic shear as a source of poloidal $\boldsymbol{E}\times\boldsymbol{B}$ angular momentum and discuss the mirror and the Lorentz force. Here, we find that the geodesic transfer term, the Stringer-Winsor spin-up term and the ion-orbit loss term are all part of the Lorentz force and are in fact one and the same term.

Discussing the relation to angular velocity we build the inertia tensor with the help of the first fundamental form of a flux-surface. In turn, the inertia tensor is used to construct a flux-surface averaged rotational energy for $\boldsymbol{E}\times\boldsymbol{B}$ surface flows of the plasma. The evolution of this rotational energy features a correction of previous results due to the inertia tensor. In particular, this correction suggests that density sources on the high-field side contribute much more to zonal flow energy generation than on the low field side.

Our derivation is based on a full-F, electromagnetic, gyro-kinetic model in a long-wavelength limit. The results can be applied to gyro-kinetic as well as gyro-fluid theories and can also be compared to drift-kinetic and drift-fluid models. Simplified cases for the magnetic field geometry including the axisymmetric purely toroidal and purely poloidal magnetic fields are discussed, as are the angular momentum balance of the electromagnetic fields, the ion-orbit loss mechanism and the parallel acceleration.

Keywords: Reynolds stress, Favre stress, zonal flows, ion orbit loss, parallel acceleration, gyro-kinetic, flux coordinates

(Some figures may apper in colour only in the online journal)

1. Introduction

The double periodicity of a toroidal magnetic field configuration can be associated with two rotational degrees of freedom: toroidal and poloidal rotation. In a toroidally confined plasma both toroidal and poloidal rotation are observed and subject to intensive research.

Studies of toroidal rotation favour the toroidally symmetric tokamak case, where the symmetry leads to the exact conservation of the collective[3] canonical angular momentum [1–3]. Of particular interest is the so-called intrinsic rotation, which refers to the ability of the plasma to spontaneously rotate without application of an external torque like neutral beam injection [4–6]. This is an important topic because toroidal rotation stabilizes the plasma against instabilities like the resistive wall mode.

The ideal toroidal symmetry of a tokamak is broken in stellarators and in reality also in tokamaks due to magnetic ripple effects from external field coils spacing. In fact, stellarator physics is different from tokamaks in some important aspects [7]. Neoclassical transport levels are much higher in a stellarator than in a tokamak even though stellarator optimization aims at reducing these levels down or below turbulent transport levels. More importantly however, the exact invariance of toroidal angular momentum is lost in a stellarator due to the lack of axial symmetry[4]. It is argued that in this case it is impossible for the plasma to rotate as fast as in (quasi-) axisymmetric devices [8, 9] since the radial electric field is restricted by the ambipolarity condition but that zonal flows may still develop.

Poloidal angular momentum, just as toroidal angular momentum, has two components in a general magnetic field, one stemming from the parallel velocity projected to the poloidal direction $u_\parallel \hat{\boldsymbol{b}} \cdot \boldsymbol{e}_\vartheta$, the other from the drifts perpendicular to the magnetic field $\boldsymbol{u}_\perp \cdot \boldsymbol{e}_\vartheta$ (toroidal momentum analogously with $\boldsymbol{e}_\varphi$). Here, $u_\parallel \equiv \boldsymbol{u} \cdot \hat{\boldsymbol{b}}$ is the parallel flow velocity, $\boldsymbol{u}_\perp \equiv \hat{\boldsymbol{b}} \times (\boldsymbol{u} \times \hat{\boldsymbol{b}})$ is the perpendicular flow velocity, $\hat{\boldsymbol{b}}$ is the magnetic unit vector and $\boldsymbol{e}_\vartheta$ and $\boldsymbol{e}_\varphi$ are the covariant poloidal and toroidal base vectors. In reverse this means that both parallel velocity as well as the perpendicular drifts contribute to both toroidal as well as poloidal rotation. This is simply the geometrical observation that parallel and perpendicular directions versus poloidal and toroidal directions are different basis vectors for a flux-surface. This being said, the poloidal component of $\boldsymbol{E} \times \boldsymbol{B}$ velocity $u_{E,\vartheta} \equiv \boldsymbol{u}_E \cdot \boldsymbol{e}_\vartheta$ gains significant interest because of its role in the formation of a transport barrier during the L-H transition [10–12]. The high confinement mode is accompanied by a narrow potential well just inside the separatrix of a diverted magnetic field geometry. The associated radial electric field drives a strongly sheared and flux-aligned $\boldsymbol{E} \times \boldsymbol{B}$ mean flow, which suppresses turbulence and thus reduces the radial flow of particles and heat out of the confined region. This $\boldsymbol{E} \times \boldsymbol{B}$ shear flow is believed to emerge out of turbulent fluctuations via the Reynolds stress, yet other mechanisms like the ion-orbit loss mechanism [13–15] or the Favre stress and background density gradient drive [16] are currently under discussion as well. Recent results suggest that the latter significantly alter the generation mechanism of $\boldsymbol{E} \times \boldsymbol{B}$ zonal flows for high density fluctuation amplitudes and steep density gradients [16, 17].

It is instructive to introduce rotation also from a purely mechanical perspective. Consider a particle of mass m confined to a toroidal surface. Its Lagrangian reads $L_p = m(R^2\dot{\varphi}^2 + a^2\dot{\vartheta}^2)/2$ with the geometrical toroidal angle φ and poloidal angle ϑ. In an ideal torus the distance from the major axis $R(\vartheta) = R_0 + a\cos\vartheta$, with R_0 the major radius, is independent of the geometric toroidal angle φ. The distance from the minor axis $a = a_0$ remains the minor radius a_0. The Euler–Lagrange equations directly yield the conservation of toroidal angular momentum $\dot{L}_\varphi = 0$ with $L_\varphi = mR^2\dot{\varphi}$. This is a consequence of the independence of R and a of the toroidal angle φ. We then have $\dot{\varphi} = L_\varphi/m(R_0 + a\cos\vartheta)^2$. Notice that the angular frequency $\dot{\varphi}$ is higher on the torus inside $\vartheta = \pi$ than on the outside $\vartheta = 0$, which we intuitively expect. In contrast, the equation for the poloidal angle is given by the non-linear differential equation $\dot{\vartheta} = -L_\varphi^2 \sin\vartheta/m^2a^4(R_0/a + \cos\vartheta)^3$. We observe that the poloidal angular momentum is not a conserved quantity for $L_\varphi \neq 0$. Furthermore, on a generally shaped toroidal flux-surface like that of a stellarator R as well as a depend on both φ and θ. There, neither toroidal nor poloidal angular momenta are conserved and ϑ and φ obey a coupled set of non-linear differential equations.

In this contribution we calculate the toroidal and poloidal angular momentum balance separately for both the $\boldsymbol{E} \times \boldsymbol{B}$ and the parallel velocity part. Previous work is mostly restricted to toroidal symmetry [1–3], simplified magnetic field geometry [16, 18–20] or delta-f modelling [18–20]. Here, we are interested in how the angular momentum anchors to the background magnetic field in the absence of a symmetry, what components appear in the complete stress tensor beside the ever present Reynolds stress and the impact of high fluctuation amplitudes and small gradient length scales.

Our derivation rests upon two pillars: (i) a full-F gyrokinetic formalism, where finite Larmor radius and polarization effects are taken in the long-wavelength limit and (ii) a drift ordering of the resulting energy-momentum balance itself. The long-wavelength limit is a way to obtain closed expressions in the energy-momentum balance. The main effect of the full-F formalism is the appearance of the density inside flux-surface averages. In order to present the main nonlinearities in a convenient form we introduce the Favre average - a density weighted flux-surface average [16]. As a natural consequence, the Favre stress emerges, which generalizes the conventional Reynolds stress. The drift ordering is necessary to neglect geometric correction terms that would otherwise clutter the resulting expressions and to easily identify fluid moments from velocity space integrals. However, our momentum balance equations are valid only up to order three within this ordering.

The magnetic field geometry is arbitrary and we in particular do not invoke a toroidal symmetry. Thus, as long as the orderings hold, our results are applicable to various

[3] After species and particle summation - individual particles exchange momentum through fluctuating electromagnetic fields.

[4] Axisymmery, axial symmetry and toroidal symmetry are used interchangeably throughout this manuscript.

devices, including tokamaks and stellarators, the reversed field pinch and field-reversed configurations. Further, we make no assumptions on the form of the gyro-kinetic distribution function and our results thus apply to gyro-kinetic as well as gyro-fluid models. At the same time we allow a direct comparison to drift-reduced fluid equations due to the applied drift ordering.

We carefully recall the definition of angular momentum from the underlying particle Lagrangian in suitable coordinates and construct the inertia tensor with the help of the first fundamental form of general flux-surfaces. This enables us to then construct and discuss a rotational energy balance. Within the energy balance equations we keep terms up to order four in the drift ordering.

This manuscript is divided into the following parts. In section 2 we review the magnetic field representation via flux-coordinates in order to setup suitable poloidal and toroidal angle coordinates. Our main derivation then proceeds with the definition of the gyro-kinetic action in section 3, which encompasses our assumptions on the model, specifically the long-wavelength limit. The drift ordering scheme is presented in section 4. The latter enables us to then derive the poloidal and toroidal angular momentum balance up to order three within this ordering and in particular replace gyro-fluid with regular fluid moments in the result. In section 5 we apply the previously proposed Favre decomposition [16] in order to identify the signature of relative density fluctuations in both known and novel components of the stress tensor. In section 6 we derive the relation between angular momentum and angular velocity and identify the inertia tensor. Furthermore, we find the time evolution for the rotational energy using the previously derived momentum balance. Finally, we discuss the significance of our results on various topics discussed in the literature in section 7, including the electromagnetic field momentum, drift-fluid models, the ion orbit loss mechanism and the transition to simplified geometries. A provides a formulary intended as a quick reference list of the most often used relations and notations.

2. Preliminary: the magnetic field in flux-coordinates

A toroidal magnetic field equilibrium can be represented by so-called flux-coordinates ρ, ϑ, φ ([21] calls them magnetic coordinates) where the magnetic field lines appear straight

$$\mathcal{B}^2 = \mathrm{d}\psi_p \wedge \mathrm{d}\varphi + \mathrm{d}\psi_t \wedge \mathrm{d}\vartheta \tag{1}$$

Here, $\psi_p(\rho)$ is the poloidal flux and $\psi_t(\rho)$ is the toroidal flux and we have $\mathrm{d}\psi_p = \iota(\rho)\mathrm{d}\psi_t$ where we introduced the rotational transform ι. Further, ρ is any radially increasing flux label, ϑ is the poloidal flux angle and φ is the toroidal flux angle coordinate. Note that ϑ increases in the counter-clockwise direction in the poloidal plane while φ increases clockwise if viewed from above to get a right-handed coordinate system. We emphasize that in general φ and ϑ are different from the geometric angles. In this manuscript we always refer to flux angles when speaking of the toroidal and poloidal angles or directions and will highlight when these angles coincide with the geometric angles.

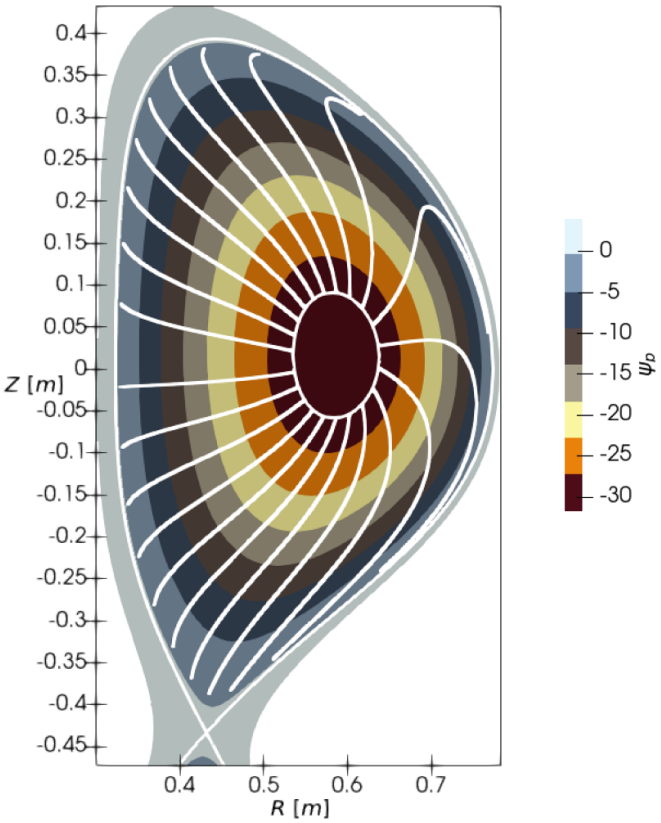

Figure 1. Numerically integrated [23] flux coordinates for an axisymmetric equilibrium. The contours of the poloidal flux label ψ_p are given in colour with white markers for the separatrix and the starting contour for ϑ integration, while the lines of constant poloidal flux angle ϑ are given in white as well.

There are many different toroidal flux coordinate systems, notably Hamada and Boozer coordinates [21, 22]. In figure 1 we show an example of a numerically integrated [23] flux-coordinate system for an axisymmetric tokamak magnetic field. Here, we show the lines of constant ψ_p in colour and the lines of constant poloidal flux angle ϑ in white. The toroidal flux angle φ coincides with the geometric angle.

The magnetic field $\mathcal{B}^2 = \mathrm{d}\mathcal{A}^1$ can be written as a total differential of the magnetic potential

$$\mathcal{A}^1 = \psi_p \mathrm{d}\varphi + \psi_t \mathrm{d}\vartheta \tag{2}$$

which notably identifies $A_\varphi = \psi_p(\rho)$ and $A_\vartheta = \psi_t(\rho)$. At the same time $\mathrm{d}\mathcal{B}^2 = \mathrm{d}\circ\mathrm{d}\mathcal{A}^1 = 0$ immediately as $\mathrm{d}\circ\mathrm{d} = 0$ for the exterior derivative d. This is the coordinate-free expression of vanishing divergence.

We formulate equations (1) and (2) in terms of differential forms, which we here introduce because the gyro-kinetic theory heavily relies on them (for an excellent introduction to

differential geometry for physicists see Frankel's text [24]). An interesting (if somewhat aloof) property of using differential forms is that they (and therefore the magnetic field) can be defined without the existence of a metric tensor. Recall for example that the 1-form $\mathrm{d}\vartheta$ symbolizes the planes that are constructed by keeping ϑ constant and varying ρ and φ, which is a purely topological operation. In contrast, the gradient basis vector $\nabla\vartheta$ is the vector that is perpendicular to the planes of constant ϑ, which requires a metric to define.

We are of course aware of the practicality that the physicist's notation of Equation (1) provides

$$\boldsymbol{B} = \nabla\psi_p \times \nabla\varphi + \nabla\psi_t \times \nabla\vartheta \tag{3}$$

We are here able to identify the poloidal $\boldsymbol{B}_p := \nabla\psi_p \times \nabla\varphi$ and toroidal $\boldsymbol{B}_t := \nabla\psi_t \times \nabla\vartheta$ parts of the magnetic field vector $\boldsymbol{B}$. With the choice of signs in equation (3) and assuming $\nabla\psi_p$ points radially outwards, we get a left-handed field-line winding when going in the positive φ direction since $\boldsymbol{B}_p \sim -\boldsymbol{e}_\vartheta$. Furthermore, notice the useful properties

$$\nabla\psi_p = \boldsymbol{e}_\varphi \times \boldsymbol{B} \tag{4}$$

$$\nabla\psi_t = \boldsymbol{e}_\vartheta \times \boldsymbol{B} \tag{5}$$

where $\boldsymbol{e}_\varphi$ and $\boldsymbol{e}_\vartheta$ are the covariant basis vectors, that is the vectors that generate the directional derivatives along φ and ϑ, or in other words, $\boldsymbol{e}_\varphi$ is the tangent vector to the line that we get when keeping ρ and ϑ constant and varying φ ($\boldsymbol{e}_\vartheta$ analogous). We emphasize that we mean these two vectors when we speak of toroidal $\boldsymbol{e}_\varphi$ and poloidal $\boldsymbol{e}_\vartheta$ directions in contrast to the $\nabla\varphi$ and $\nabla\vartheta$ directions. For example, in figure 1 $\boldsymbol{e}_\varphi$ points perpendicularly out of the plane while $\boldsymbol{e}_\vartheta$ is tangent to the contours of ψ_p (!) and in particular does not point in the same direction as $\nabla\vartheta$, which has component out of the flux-surface as well.

When we deal with a symmetric field independent of the geometric toroidal angle, we will choose φ as the geometric toroidal angle and keep ϑ as a flux-coordinate with $\nabla\vartheta \cdot \nabla\varphi = \nabla\rho \cdot \nabla\varphi = 0$ as we do in figure 1. This type of coordinates is known as symmetry flux or PEST coordinates [25]. Notice that we do not use the geometric poloidal angle since we want to keep the form equation (1). A useful property of this type of coordinate is that $qR^2/\sqrt{g} \equiv I(\rho)$ is a flux function, which allows us to write

$$\boldsymbol{B} = I(\rho)\nabla\varphi + \nabla\psi_p \times \nabla\varphi \tag{6}$$

Last, note that all flux coordinates are problematic when an X-point with $\nabla\psi_p = 0$ is present in or close to the domain of interest. In fact, any coordinate system with a flux label as the first coordinate is problematic when an X-point is present [26]. On the one hand the poloidal flux ψ_p is continuous and well-defined across the separatrix. However, the toroidal flux ψ_t as well as the poloidal flux angle ϑ are only well-defined up to but not including or across the separatrix and furthermore ι^{-1} diverges on the separatrix. This is expected since the poloidal component of $\boldsymbol{B}$ vanishes at the X-point. In practice, the divergence manifests for example in figure 1 where the coordinate lines for ϑ are distorted when getting close to the separatrix on the low field side of the tokamak.

Last, we introduce the flux surface average (see for example [22]) as an average over a small volume - a differential shell centered around the flux-surface. We define

$$\langle f \rangle(\psi_p) := \frac{\partial}{\partial v}\int_\Omega \mathrm{dV} f = \int_{\psi_p} \frac{f(\boldsymbol{x})}{|\nabla v|}\,\mathrm{dA} \tag{7}$$

where we define $v(\psi_p) := \int_0^{\psi_p} \mathrm{dV}$ as the volume flux label. In flux coordinates we have $\mathrm{dA} = \sqrt{g}|\nabla\rho|\mathrm{d}\vartheta\mathrm{d}\varphi$. The average fulfills the identity

$$\langle \nabla \cdot \boldsymbol{j} \rangle = \frac{\partial}{\partial v}\langle \boldsymbol{j} \cdot \nabla v \rangle \tag{8}$$

Also note that for any divergence free vector field $\nabla \cdot \boldsymbol{j} = 0$ and a flux function $f(\psi_p)$ we have

$$\langle \nabla \cdot (\boldsymbol{j} f) \rangle = 0 \tag{9}$$

which is proven straightforwardly.

In summary, using flux coordinates for the following derivation defines suitable angle coordinates as well as poloidal and toroidal directions. We expect the resulting expressions to be valid for any flux coordinate system within the closed field-line region up to the separatrix. We remark that the numerical issues of flux coordinates close to the separatrix do not affect the theoretical results presented here.

3. Fundamentals of Hamiltonian dynamics

3.1. Model definition

In this section we define our gyro-kinetic model and discuss the approximations that go into it. Our goal is to set up a model suitable for edge and scrape-off layer conditions. Literature on the derivation of gyro-kinetic models based on Lie-transform perturbation theory include the rather technical review [27] and references therein. A friendlier tutorial can be found in [28] or the more recent [29]. Here, we start directly with the gyro-centre Poincaré 1-form expressed in the transformed phase-space coordinates $\boldsymbol{Z} := \{\boldsymbol{X}, w_\parallel, \mu, \theta\}$, with gyro-centre coordinate $\boldsymbol{X}$, parallel canonical moment $w_\parallel$, magnetic moment μ, gyro-angle θ

$$\gamma := \left(q\boldsymbol{A} + m w_\parallel \hat{\boldsymbol{b}}\right) \cdot \mathrm{d}\boldsymbol{X} + \frac{m}{q}\mu\mathrm{d}\theta \tag{10}$$

with species mass m and charge q and we omit the species label. We have the magnetic background potential $\boldsymbol{A} \cdot \mathrm{d}\boldsymbol{X} \equiv \mathcal{A}^1$ from equation (2) and the background magnetic field unit vector $\hat{\boldsymbol{b}} := \boldsymbol{B}/B$. In flux-coordinates equation (10) explicitly reads

$$\gamma = (q\psi_t + m w_\parallel b_\vartheta)\mathrm{d}\vartheta + (q\psi_p + m w_\parallel b_\varphi)\mathrm{d}\varphi + \frac{m}{q}\mu\mathrm{d}\theta \tag{11}$$

We remark that this 1-form is already enlightening because it immediately identifies

$$\gamma_\varphi = q\psi_p + mw_\parallel b_\varphi \tag{12}$$

as the toroidal angular momentum and

$$\gamma_\vartheta = q\psi_t + mw_\parallel b_\vartheta \tag{13}$$

as the poloidal angular momentum. Recall here that angular momentum is defined as the canonically conjugate momentum to the angle coordinate. In anticipation of the following discussion we here remark that $q\psi_p$ and $q\psi_t$ will lead to the toroidal and poloidal components of the $\boldsymbol{E}\times\boldsymbol{B}$ velocity contribution. The parallel velocity contribution is given by the two components of the magnetic field unit vector b_ϑ and b_φ as expected. Unfortunately however, the definitions for toroidal and poloidal angular momentum in equations (12) and (13) are *not* coordinate invariant and therefore care must be taken when comparing results from different coordinate systems. This is evident since the value of b_φ and b_ϑ depend on the choice of coordinates. Physically, we attribute this to different reference points/axes for the rotation that different angle coordinates entail.

The symplectic 2-form, defined by the Poincaré 1-form, $w := \mathrm{d}\gamma$, defines the geometry of phase-space much the same way the metric tensor g defines the geometry of ordinary space. The difference is that ω defines areas instead of distances and is skew-symmetric instead of symmetric (see [24]). In 6-dimensional phase-space coordinates we have

$$\begin{aligned}\omega_{ij} &= \frac{\partial\gamma_j}{\partial Z^i} - \frac{\partial\gamma_i}{\partial Z^j}\\ \omega &= \begin{pmatrix} -q(\boldsymbol{B}^*\times) & -m\hat{\boldsymbol{b}} & 0 & 0\\ m\hat{\boldsymbol{b}}^{\mathrm{T}} & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{m}{q}\\ 0 & 0 & -\frac{m}{q} & 0 \end{pmatrix}\end{aligned} \tag{14}$$

$$q\boldsymbol{B}^* := q\boldsymbol{B} + mw_\parallel \boldsymbol{\nabla}\times\hat{\boldsymbol{b}} \tag{15}$$

$$B^*_\parallel = \boldsymbol{B}^*\cdot\hat{\boldsymbol{b}} = B + \frac{mw_\parallel}{q}(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})_\parallel \tag{16}$$

Note the covariant vector components b_i (with $\hat{\boldsymbol{b}}^{\mathrm{T}} := (b_1, b_2, b_3)$) and the appearance of the determinant of the metric tensor g in the definition of the cross-product $(\boldsymbol{B}^*\times)_{ij} := \sqrt{g}\varepsilon_{ikj}B^{*k}$ with contravariant components B^{*k}.

The phase space volume $\mathrm{vol} := \omega\wedge\omega\wedge\omega = \sqrt{\det(\omega)}\mathrm{d}^6Z$ reads

$$\sqrt{\det(\omega)}\mathrm{d}^3X\mathrm{d}w_\parallel\mathrm{d}\mu\mathrm{d}\theta = m^2\sqrt{g}|B^*_\parallel|\mathrm{d}^3X\mathrm{d}w_\parallel\mathrm{d}\mu\mathrm{d}\theta \tag{17}$$

Notice that the volume form is proportional to $|B^*_\parallel|$ not just $B^*_\parallel$ as often noted since it needs to remain positive. More importantly, it is apparent that the coordinate system possesses a (coordinate) singularity at $mw_\parallel = -qB/(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})_\parallel$, where $B^*_\parallel = 0$ and thus $\det(\omega) = 0$. This destroys the symplecticity of the 2-form ω, the volume form Eq (17) vanishes and the inverse of ω diverges (and thus the equations of motion). It is questionable how we can deal with this singularity especially when we later integrate over the phase-space volume to form the field Equations. Furthermore, when deriving gyro-fluid models terms $\propto (B^*_\parallel)^{-1}$ prevent identifying velocity space moments that involve $B^*_\parallel$ in the volume element. This problem is often ignored in the literature or circumvented by requiring $(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})_\parallel = 0$ and we will follow this approach in this work. For a low-β stellarator $\boldsymbol{\nabla}\times\boldsymbol{B} = 0$, however for general tokamak magnetic fields the requirement is only approximately fulfilled. As we will show in section 7.1 the problem is also resolved by simplifying the magnetic field to purely toroidal or poloidal. Interestingly, the requirement $(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})_\parallel = 0$ relates to the integrability condition for vector fields perpendicular to the magnetic field. The Frobenius theorem [24] states that planes perpendicular to $\hat{\boldsymbol{b}}(\boldsymbol{x})$ everywhere exist in the sense that there exist functions $\lambda(\boldsymbol{x})$ and $f(\boldsymbol{x})$ such that $\lambda(\boldsymbol{x})\hat{\boldsymbol{b}}(\boldsymbol{x}) = \boldsymbol{\nabla}f$ if and only if $(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})\cdot\hat{\boldsymbol{b}} = 0$. In other words we surmise that the existence of drift-planes is a prerequisite for gyro-kinetic and -fluid models. Our Hamiltonian reads

$$\begin{aligned} H :&= \frac{\left(mw_\parallel - q\mathcal{A}_{1,\parallel}\right)^2}{2\,\mathrm{m}} + \mu B + q\Psi\\ &\equiv \frac{1}{2}mw^2_\parallel + \mu B + H_f \end{aligned} \tag{18}$$

with the effective gyro-centre potentials

$$q\mathcal{A}_{1,\parallel} := qA_{1,\parallel} + \frac{m\mu}{2qB}\Delta_\perp A_{1,\parallel} \tag{19}$$

$$q\Psi := q\phi + \frac{m\mu}{2qB}\Delta_\perp\phi - \frac{1}{2}m\left(\frac{\boldsymbol{\nabla}_\perp\phi}{B}\right)^2 \tag{20}$$

where we define the field Hamiltonian $H_f := q\Psi - qw_\parallel\mathcal{A}_{1,\parallel} + q^2\mathcal{A}^2_{1,\parallel}/2\,m$ to contain all terms dependent on the electromagnetic field perturbations ϕ and $A_{1,\parallel}$. The potential ϕ is in fact a first order term where the zeroth order ϕ_0 has been neglected. The first order perturbation $\mathcal{A}_{1,\parallel}$ is not to be confused with the zeroth order magnetic field potential $\mathcal{A}^1$. Finally, see table A1 in A for definitions of $\boldsymbol{\nabla}_\perp$ and $\Delta_\perp$. Here, we follow [1, 2] and use the Hamiltonian formulation with $mw_\parallel := mv_\parallel + q\mathcal{A}_{1,\parallel}$ such that the electromagnetic field variations appear in the Hamiltonian only and do not disturb the symplectic geometry (10). We note that we

(a) neglect all terms $k^3_\perp\rho^3_0$ with gyro-radius $\rho_0 := \sqrt{2B\mu m}/eB$ and higher in the Hamiltonian (this especially neglects the second order guiding centre contributions, which according to [30] leads to guiding centre drifts in the polarization equation). In particular, both the polarization contribution (the last term in equation (20)) as well as the finite Larmor radius effects are taken in the long-wavelength limit [31].

(b) neglect compressional Alfvén waves entering through $\boldsymbol{A}_{\mathbf{1},\perp}$ [32]
(c) neglect all terms non-linear in the magnetic potential $A_{1,\parallel}$ (except in the parallel kinetic energy). This approximation implies the absence of $A_{1,\parallel}$ terms in the polarization and of ϕ terms in the parallel Ampère law [32] Equation and vice versa ϕ terms in the parallel Ampère law and therefore decouples the two equations, which is numerically desirable [5]

Our model is comparable to Reference [33] with the difference that we additionally take the long-wavelength limit in the gyro-average operator. We note that with our approximations the Hamiltonian formulation with $w_\parallel$ is entirely equivalent to the symplectic formulation using $v_\parallel$ in the sense that the resulting equations are the same. The Hamiltonian formulation is more convenient here since γ is time-independent. We also remark that the gyro-average and polarization corrections in our gyro-kinetic model equation (18) resemble the second order guiding centre transformation terms in guiding-centre models [34, 35]. However, since we logically start with and approximate a gyro-kinetic model we will keep referring to our model as gyro-kinetic.

We introduce the gyro-kinetic particle distribution function $F(\boldsymbol{Z},t) \equiv F(\boldsymbol{X}, w_\parallel, \mu, t)$ (independent of gyro-angle θ, which is averaged out). The Vlasov equation states

$$\frac{\mathrm{d}}{\mathrm{d}t}F = \frac{\partial F}{\partial t} + \dot{Z}^i \frac{\partial F}{\partial Z^i} = S \tag{21}$$

Here, S is a general kinetic source term in gyro-centre phase-space $S(\boldsymbol{X}, w_\parallel, \mu, t)$. With the kinetic source function S we formally represent effects like for example non-elastic collisions, plasma-neutral interactions, heating of the plasma, or plasma fuelling and bear in mind that detailed expressions for S are not part of this manuscript. We call S a source understanding that it can act as a sink as well.

Next, with the 1-form γ in equation (10) and the Hamiltonian H in equation (18) we can define a particle Lagrangian

$$L_p := \gamma_i \dot{Z}^i - H \tag{22}$$

Together with the volume form in equation (17) and the phase space distribution function F we can then define the system Lagrangian $\mathcal{L}_p := \sum_s \int \mathrm{vol}(\boldsymbol{Z}) F(\boldsymbol{Z},t) L_p(\boldsymbol{Z}, \dot{\boldsymbol{Z}}, t)$, where we sum over species. Finally, we close the system with a field Lagrangian and propose the action integral

$$\begin{aligned}\mathcal{S} := &\sum_s \int \mathrm{d}t \int \mathrm{d}V \mathrm{d}w_\parallel \mathrm{d}\mu \mathrm{d}\theta m^2 B F(\gamma_i \dot{Z}^i - H) \\ &- \int \mathrm{d}t \int \mathrm{d}V \frac{(\boldsymbol{\nabla}_\perp A_{1,\parallel})^2}{2\mu_0}\end{aligned} \tag{23}$$

[5] Desirable might be an understatement. We are not aware of any successful attempts to numerically solve the completely coupled set of equations in a turbulence simulation.

pagebreak where $\mathrm{d}V := \sqrt{g}\mathrm{d}^3X$ is the spatial volume form. The action in equation (23) plus the Vlasov equation (21) are the central relations in every gyro-kinetic model. They completely define the system that we investigate. In particular this means that $\mathcal{S}$ contains all approximations to our model and that the following calculations are exact. We remark that

(a) the neglect of the electric energy $\boldsymbol{E}^2$ in the field part of equation (23) leads to quasineutrality (that is a vanishing right hand side in equation (42))
(b) the magnetic field energy in (23) neglects the $A_{1,\parallel}(\boldsymbol{\nabla} \times \boldsymbol{B})_\parallel$ contribution from the background magnetic field. This leads to the omission of the background equilibrium current in the Ampère equation (43). This approximation is in line with $(\boldsymbol{\nabla} \times \hat{\boldsymbol{b}})_\parallel = 0$.

3.2. The Vlasov-Maxwell equations

In the Lagrangian picture [36] the equations of motion can be retrieved from the action equation (23) by expressing $\boldsymbol{Z} = \boldsymbol{Z}(\boldsymbol{Z}_0, t_0; t)$, using $F(Z,t) = F_0(Z_0, t_0)$ by the Vlasov equation (21), taking the integration to the initial positions and time[6] and then varying $\delta\mathcal{S}/\delta\boldsymbol{Z} = 0$. This indeed recovers the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_p}{\partial \dot{Z}^i} - \frac{\partial L_p}{\partial Z^i} = 0 \tag{24}$$

The application of the Euler Lagrange equations (24) yields the Hamilton equations of motion (using $\mathrm{d}\gamma_i/\mathrm{d}t = \dot{Z}^j \partial\gamma_i/\partial Z^j$)

$$Z^i_H \omega_{ij} = -\partial_j H \;\leftrightarrow\; i_Z \omega = -\mathrm{d}H \tag{25}$$

where we define $Z^i_H \equiv \dot{Z}^i$ as the components of the Hamiltonian vector field on phase space. Here, i_Z is the inner product with the vector field Z_H and d is the total differential. The particle trajectories are given by the streamlines of Z_H (with $J := \omega^{-1}$)

$$\frac{\mathrm{d}Z^i}{\mathrm{d}t} = Z^i_H = J^{ij}\frac{\partial H}{\partial Z^j} \tag{26}$$

The time-derivative of any phase-space function along the trajectory is then given by

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{Z},t) = \frac{\partial f}{\partial t} + \dot{Z}^i \frac{\partial f}{\partial Z^i} = \frac{\partial f}{\partial t} + Z^i_H \frac{\partial f}{\partial Z^i} \tag{27}$$

Here and in the following we use $\dot{Z}$ synonymously with Z_H. In particular, the derivative of the Hamiltonian gives

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\boldsymbol{Z},t) = \frac{\partial H}{\partial t} + Z^i_H \frac{\partial H}{\partial Z^i} = \frac{\partial H}{\partial t} \tag{28}$$

where we use equation (26) and the antisymmetry of J.

[6] Technically, here we also need to know that the volume form is conserved in time $B(z)\mathrm{d}^6Z = B(z_0)\mathrm{d}^6z_0$, something that we will need to show explicitly.

Explicit expressions for the inverse of the symplectic 2-form equation (14) and the gradient of the Hamiltonian (18) are

$$J = \begin{pmatrix} \frac{1}{qB}(\hat{\boldsymbol{b}}\times) & \frac{1}{mB}\boldsymbol{B}^* & 0 & 0 \\ -\frac{1}{mB}\boldsymbol{B}^{*\mathrm{T}} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{q}{m} \\ 0 & 0 & \frac{q}{m} & 0 \end{pmatrix} \tag{29}$$

$$(\partial H)^{\mathrm{T}} = \begin{pmatrix} \mu B\boldsymbol{\nabla}\ln B + \boldsymbol{\nabla}H_f & mv_\parallel & B^\# & 0 \end{pmatrix} \tag{30}$$

with $(\hat{\boldsymbol{b}}\times)^{ij} := \sqrt{g}^{-1}\epsilon^{ikj}b_k$ and $\boldsymbol{\nabla}H_f = -v_\parallel q\boldsymbol{\nabla}\mathcal{A}_{1,\parallel} + q\boldsymbol{\nabla}\Psi$. The μ component of ∂H contains corrections due to the fluctuating electric field $B^\# := B + m\Delta_\perp\phi/2q^2B$. An explicit expression for the components of Z_H (or $\dot{Z}$) can now be formed given equations (29) and (30) (with $mv_\parallel \equiv mw_\parallel - q\mathcal{A}_{1,\parallel}$)

$$\begin{aligned}\dot{\boldsymbol{X}} =& \frac{1}{B}\left(\boldsymbol{B}^* v_\parallel + \frac{1}{q}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}H\right) \\ =& \frac{1}{B}\left(\boldsymbol{B}v_\parallel + \frac{mv_\parallel^2}{q}\boldsymbol{\nabla}\times\hat{\boldsymbol{b}} + \frac{\mu B}{q}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B\right. \\ &\left. + v_\parallel\boldsymbol{\nabla}\times\mathcal{A}_{1,\parallel}\hat{\boldsymbol{b}} + \hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\Psi\right)\end{aligned} \tag{31}$$

$$\begin{aligned}m\dot{w}_\parallel =& -\frac{\boldsymbol{B}^*}{B}\cdot\boldsymbol{\nabla}H \\ =& -\frac{1}{B}\left(\boldsymbol{B} + \frac{mv_\parallel}{q}\boldsymbol{\nabla}\times\hat{\boldsymbol{b}} + \boldsymbol{\nabla}\times\mathcal{A}_{1,\parallel}\hat{\boldsymbol{b}}\right) \\ &\cdot(\mu B\boldsymbol{\nabla}\ln B + q\boldsymbol{\nabla}\Psi) + q\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}\mathcal{A}_{1,\parallel}\end{aligned} \tag{32}$$

$$\dot{\mu} = 0 \tag{33}$$

$$\dot{\theta} = \frac{qB}{m} + \frac{\Delta_\perp\phi}{2B} \tag{34}$$

The phase space volume $\mathrm{vol} = \omega\wedge\omega\wedge\omega$ is conserved along the particle trajectories

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{vol} = \mathcal{L}_Z\mathrm{vol} = (\mathrm{d}i_Z\omega)\wedge\omega\wedge\omega = 0 \tag{35}$$

where the Lie derivative on differential forms is given by Cartan's formula [24] $\mathcal{L}_Z\alpha = \mathrm{d}(i_Z\alpha) + i_Z\mathrm{d}\alpha$ and per definition $\mathrm{d}i_Z\omega = -\mathrm{d}\circ\mathrm{d}H = 0$. In coordinates $\mathrm{d}(i_Z\mathrm{vol}) = 0$ reads

$$\frac{1}{\sqrt{\det(\omega)}}\partial_i\left(\sqrt{\det(\omega)}Z_H^i\right) = 0 \tag{36}$$

The conservation of volume thus translates to a vanishing divergence of the Hamiltonian vector field in phase space

$$\boldsymbol{\nabla}\cdot(B\dot{\boldsymbol{X}}) + \frac{\partial}{\partial w_\parallel}(B\dot{w}_\parallel) = 0 \tag{37}$$

Notice that volume conservation does not mean that B is conserved along particle trajectories, we rather have $\dot{B} = \dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}B$.

The conservation of the particle distribution function $F(\boldsymbol{X}, v_\parallel, \mu, t)$ is expressed by the gyro-kinetic Vlasov equation $\mathrm{d}F/\mathrm{d}t = S$, which together with phase space volume conservation (37) reads in conservative form

$$\frac{\partial(BF)}{\partial t} + \boldsymbol{\nabla}\cdot\left(BF\dot{\boldsymbol{X}}\right) + \frac{\partial\left(BF\dot{w}_\parallel\right)}{\partial w_\parallel} = BS \tag{38}$$

The Vlasov-equation equation (38) together with the equations of motion equation (31)-(33) forms the first half of the Vlasov-Maxwell system.

In order to derive the Maxwell equations we first define the velocity space moment operator [37]

$$\|\zeta\| := \int \mathrm{d}w_\parallel\mathrm{d}\mu\mathrm{d}\theta m^2BF\zeta \tag{39}$$

where $\zeta(\boldsymbol{X}, w_\parallel, \mu, t)$ is any function defined on phase-space and the integration encompasses the entire velocity space. Notice that we name the first few fluid moments $N := \|1\|$, $NW_\parallel := \|w_\parallel\|$ and $P_\perp := \|\mu B\|$ and give a comprehensive list in A.2.

We also define the moment operator for the source function S analogous to the velocity space moment operator for the gyro-kinetic distribution function F in equation (39)

$$\|\zeta\|_S := \int \mathrm{d}w_\parallel\mathrm{d}\mu\mathrm{d}\theta m^2BS\zeta \tag{40}$$

Analogous to the moments of F we name the source moments $S_N := \|1\|_S$, $S_{P_\perp} = \|\mu B\|_S$, etc

Using equation (38) together with the fact that $\partial/\partial t$ and $\boldsymbol{\nabla}$ commute with the velocity integral and F vanishes for $w_\parallel = \pm\infty$ we find the important identity [37]

$$\frac{\partial}{\partial t}\|\zeta\| + \boldsymbol{\nabla}\cdot\|\zeta\dot{\boldsymbol{X}}\| = \left\|\frac{\mathrm{d}\zeta}{\mathrm{d}t}\right\| + \|\zeta\|_S \tag{41}$$

The variation of the action (23) with respect to $\phi(\boldsymbol{x})$ yields the quasi-neutrality equation

$$\frac{\delta\mathcal{S}}{\delta\phi} = \frac{\delta}{\delta\phi(\boldsymbol{x})}\sum_s\int \mathrm{d}V\mathrm{d}w_\parallel\mathrm{d}\mu\mathrm{d}\theta m^2BFH = 0 \tag{42}$$

and with respect to $A_{1,\parallel}$ the parallel Ampère law

$$\begin{aligned}\frac{\delta\mathcal{S}}{\delta A_{1,\parallel}} =& \frac{\delta}{\delta A_{1,\parallel}(\boldsymbol{x})}\sum_s\int \mathrm{d}V\mathrm{d}w_\parallel\mathrm{d}\mu\mathrm{d}\theta m^2BFH \\ &+ \frac{\delta}{\delta A_{1,\parallel}(\boldsymbol{x})}\int \mathrm{d}V\,\frac{(\boldsymbol{\nabla}_\perp A_{1,\parallel})^2}{2\mu_0} = 0\end{aligned} \tag{43}$$

where we used that γ does not depend on either ϕ or $A_\parallel$. Now, recall the variational derivative. For each $\zeta \in \{\phi, A_{1,\parallel}\}$ and $H = H(\zeta, \boldsymbol{\nabla}_\perp\zeta, \Delta_\perp\zeta)$ we have

$$\frac{\delta}{\delta\zeta(\boldsymbol{x})}\int_{R^3} \mathrm{d}^3x' FBH = FB\frac{\partial H}{\partial\zeta} - \boldsymbol{\nabla}\cdot\left(hFB\frac{\partial H}{\partial\boldsymbol{\nabla}_\perp\zeta}\right) + \Delta_\perp\left(FB\frac{\partial H}{\partial\Delta_\perp\zeta}\right) \tag{44}$$

Notice the appearance of FB inside the divergence/Laplace operators. Carrying out the variations with the help of equation (44) in the polarization and Ampère equations (42) and (43) and identifying the velocity space moments (39) yields[7]

$$\sum_\mathrm{s} qN - \boldsymbol{\nabla}\cdot\boldsymbol{P}_\mathrm{gy} = 0 \tag{45}$$

$$\sum_\mathrm{s} qNU_\parallel + \boldsymbol{\nabla}\cdot(\boldsymbol{M}^\mathrm{gy}\times\hat{\boldsymbol{b}}) = -\frac{1}{\mu_0}\Delta_\perp A_\parallel \tag{46}$$

with $mU_\parallel \equiv \left(mW_\parallel - q\mathcal{A}_{1,\parallel}\right)$, $j_{\mathrm{mag},\parallel} := \boldsymbol{\nabla}\cdot(\boldsymbol{M}^\mathrm{gy}_\perp\times\hat{\boldsymbol{b}}) = (\boldsymbol{\nabla}\times\boldsymbol{M}^\mathrm{gy})\cdot\hat{\boldsymbol{b}} - (\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})\cdot\boldsymbol{M}^\mathrm{gy}$ and the gyro-kinetic polarization and magnetization densities

$$\boldsymbol{P}_\mathrm{gy} := -\sum_\mathrm{s}\left[\boldsymbol{\nabla}_\perp\left(\frac{mP_\perp}{2qB^2}\right) + \frac{mN\boldsymbol{\nabla}_\perp\phi}{B^2}\right] \tag{47}$$

$$\boldsymbol{M}^\mathrm{gy}_\perp := \sum_\mathrm{s}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\left(\frac{m(Q_\parallel + U_\parallel P_\perp)}{2qB^2}\right) \tag{48}$$

Note that the parallel component of the polarization current $\boldsymbol{j}_\mathrm{pol}\cdot\hat{\boldsymbol{b}} = \partial\boldsymbol{P}_\mathrm{gy}\cdot\hat{\boldsymbol{b}}/\partial t = 0$ vanishes in equation (46). Also, the parallel part of the magnetization density $\boldsymbol{M}^\mathrm{gy}_\parallel := -\|\mu\|\hat{\boldsymbol{b}} \equiv -P_\perp\hat{\boldsymbol{b}}/B$ does not contribute to the parallel magnetization current.

In total, we now explicitly derived the equations of the Vlasov-Maxwell system in equation (38),(45) and (46) together with the equations of motion in (31)-(33).

3.3. Interlude: relation between gyro-fluid and fluid moments

Gyro-fluid quantities like $\|1\| = N(\boldsymbol{X},t)$ or $\|v_\parallel\| = U_\parallel(\boldsymbol{X},t)$ are given in gyro-centre coordinates $\boldsymbol{X}$ and are thus not directly comparable to the physical fluid quantities, which we denote with lower case letters $n(\boldsymbol{x},t) := \int \mathrm{d}^3v f(\boldsymbol{x},\boldsymbol{v},t)$, $u_\parallel(\boldsymbol{x},t) := \int \mathrm{d}^3v v_\parallel f(\boldsymbol{x},\boldsymbol{v},t)$ ..., where $f(\boldsymbol{x},\boldsymbol{v},t)$ is the distribution function in particle phase-space (and we here overburden the use of v as the velocity on top of the volume flux-label). We need to use the gyro-kinetic phase-space coordinate transformations to transform between particle and gyro-kinetic phase-space moments. Helpfully, Reference [27] relates the coordinate transformation to the variational derivative of the action. With our action (23) we obtain

$$||\xi||_v = \|\zeta\| + \Delta_\perp\left(\frac{m\|\mu B\zeta\|}{2qB^2}\right) + \boldsymbol{\nabla}\cdot\left(\frac{m\|\zeta\|\boldsymbol{\nabla}_\perp\phi}{B^2}\right) \tag{49}$$

where ξ is ζ transformed to particle coordinates and $\|\xi\|_v := \int \mathrm{d}^3v\xi f$ is the particle phase-space moment operator. Thus, $||\xi||_v$ is the physical fluid moment corresponding to $\|\zeta\|$. In equation (49) we immediately see that the actual fluid moment $\|\xi\|_v$ equals the gyro-fluid moment $\|\zeta\|$ up to an order $\mathcal{O}(\rho_0^2k_\perp^2)$ correction. For example the density transforms as

$$n = N + \Delta_\perp\left(\frac{mP_\perp}{2q^2B^2}\right) + \boldsymbol{\nabla}\cdot\left(\frac{mN}{qB^2}\boldsymbol{\nabla}_\perp\phi\right) \tag{50}$$

The right hand side terms appear exactly in the polarization equation (45), which we obtained from the variational principle. This shows that equation (45) is the gyrofluid version of quasineutrality $\sum_\mathrm{s} qn = 0$.

It is possible to invert the relation between gyro-fluid and fluid quantities. We follow [39, 40] and explicitly express the first two gyro-fluid quantities N and $U_\parallel$ in terms of the true fluid quantities n and $u_\parallel$ in the long-wavelength limit up to order $(\rho_0k_\perp)^2$.

$$N = n - \Delta_\perp\left(\frac{mnt_\perp}{2q^2B^2}\right) - \boldsymbol{\nabla}\cdot\left(\frac{mn}{qB^2}\boldsymbol{\nabla}_\perp\phi\right) \tag{51}$$

$$NU_\parallel = nu_\parallel - \Delta_\perp\left(\frac{m(q_\parallel + u_\parallel p_\perp)}{2q^2B^2}\right) \tag{52}$$

Note that we neglect the potential part in equation (52) since we miss the corresponding term in the Hamiltonian.

The moments of S transform back to particle phase space analogous to equation (49). This is because the coordinate transformation works for any phase-space function, not just the distribution function F. For example, we have

$$S_n = S_N + \Delta_\perp\left(\frac{mS_P}{2q^2B^2}\right) + \boldsymbol{\nabla}\cdot\left(\frac{mS_N}{qB^2}\boldsymbol{\nabla}_\perp\phi\right) \tag{53}$$

where S_n is the true fluid particle source term. We are now able to formulate the only constraint we have for the source term namely that it should conserve the total electric charge via

$$\sum_\mathrm{s} qS_n = \sum_\mathrm{s} qS_N - \boldsymbol{\nabla}\cdot\boldsymbol{S}_P = 0 \tag{54}$$

where we define the polarization source

[7] The attentive reader will notice that equations (45) and (46) are only semi-elliptic since the projection tensor h is only positive semi-definite. Concerns about existence and uniqueness of solutions are dealt with under "degenerate partial differential equations" in the mathematical literature. In particular, the field of stochastic differential equations contains a solution to the Dirichlet problem, see for example Reference [38].

$$\boldsymbol{S}_P = -\sum_{\mathrm{s}} \left[\boldsymbol{\nabla}_\perp \left(\frac{m S_{P_\perp}}{2qB^2} \right) + \frac{m S_N \boldsymbol{\nabla}_\perp \phi}{B^2} \right] \tag{55}$$

Note that equation (54) is completely analogous to equation (45).

4. The poloidal, toroidal and parallel momentum balance

4.1. Poloidal and toroidal $\boldsymbol{E}\times\boldsymbol{B}$ momentum

With the model developed in Sections 2 and 3 we are now ready to start the derivation of the balance equations for the angular momentum density. Keep in mind that we do not assume a toroidal symmetry here. This prohibits us from using Noether's theorem to derive an exact angular momentum balance from the action equation (23)[1, 2]. Instead, we begin by computing the time derivative of $qA_\varphi = q\psi_p$, which is the first part of the toroidal angular momentum (12)

$$\begin{aligned} q\frac{\mathrm{d}\psi_p}{\mathrm{d}t} &= q\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}\psi_p \\ &= \left(\mu B \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B}{B} + m v_\parallel^2 \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\kappa}}{B} \right)\cdot\boldsymbol{\nabla}\psi_p \\ &\quad + qA_{1,\parallel} v_\parallel \boldsymbol{K}_\kappa \cdot \boldsymbol{\nabla}\psi_p - \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\psi_p}{B}\cdot\boldsymbol{\nabla}H_f \end{aligned} \tag{56}$$

where we separated the field Hamiltonian H_f. Now, to simplify the right hand side of equation (56) we need to relate the variational derivative to ordinary derivatives. Consider a generic Hamiltonian dependence $H(\zeta, \boldsymbol{\nabla}_\perp\zeta, \Delta_\perp\zeta)$ and $\boldsymbol{\eta} := \hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\psi_p/B$

$$\begin{aligned} \boldsymbol{\eta}\cdot\boldsymbol{\nabla}H = \sum_{\zeta\in\{\phi, A_{1,\parallel}\}} \Bigg\{ &\frac{\partial H}{\partial\zeta}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\zeta + \frac{\partial H}{\partial\boldsymbol{\nabla}_\perp\zeta}\cdot\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}_\perp\zeta \\ &+ \frac{\partial H}{\partial\Delta_\perp\zeta}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\Delta_\perp\zeta \Bigg\} \end{aligned}$$

In order to proceed we need to commute $\boldsymbol{\eta}\cdot\boldsymbol{\nabla}$ with $\boldsymbol{\nabla}_\perp$ and $\Delta_\perp$. To avoid tedious geometrical correction terms we now introduce a drift ordering [20, 41], where we order

(a) the frequency of turbulent fluctuations compared to the ion gyro-frequency as small $\omega/\Omega_i \sim \delta^2 \ll 1$, where $\Omega_i = eB/m_i$
(b) the derivatives ∇_k of the dynamical fields as $\rho_i|\nabla_k \ln F| \sim \rho_i|\nabla_k \ln\phi| \sim \rho_i|\nabla_k \ln A_{1,\parallel}| \sim \rho_i k_\perp \sim \delta$ with ion thermal gyro-radius $\rho_i = \sqrt{m_i T_i}/q_i B$. This in particular orders the $\boldsymbol{E}\times\boldsymbol{B}$ velocity compared to the ion thermal velocity as $u_E/c_{s,i} \sim \delta$ where $c_{s,i} = \sqrt{T_i/m_i}$
(c) all derivatives on the magnetic field (vectors) as $L_B^{-1} \sim |\boldsymbol{\nabla}\ln B| \sim 1/R$, where R is the major radius and take $\rho_i/L_B \sim \delta^3$.
(d) $\rho_i|\nabla_\parallel \ln\phi| \sim \rho_i|\nabla_\parallel \ln A_{1,\parallel}| \sim \rho_i|\nabla_\parallel \ln F| \sim \rho_i k_\parallel \sim \delta^3$ that is parallel derivatives on the magnetic field variation scale. This implies $k_\parallel/k_\perp \sim \delta^2$

Note that Reference [29] orders $\rho_i/L_B \sim \delta^4$. However, this would completely neglect all curvature terms in our scheme. In our ordering the Hamiltonian H_f (18) appears to be second order.

We now neglect all terms of order δ^4 on the right hand side of Equation (56). With the above orderings we directly have that $\eta^i\boldsymbol{\nabla}_i h^{kl} \sim \delta^3$ and $h^{kl}\boldsymbol{\nabla}_l\eta^i \sim \delta^3$. With this and $\partial H/\partial\boldsymbol{\nabla}_\perp\phi = m\boldsymbol{\nabla}_\perp\phi/B^2$ we can order $m\boldsymbol{\nabla}_\perp\phi\cdot\eta^i\boldsymbol{\nabla}_i\boldsymbol{\nabla}_\perp\phi = m\boldsymbol{\nabla}_\perp\phi\cdot\boldsymbol{\nabla}_\perp(\eta^i\boldsymbol{\nabla}_i\phi) + \mathcal{O}(\delta^5)$. With similar arguments we can order $\eta^i\boldsymbol{\nabla}_i\Delta_\perp\phi = \Delta_\perp(\eta^i\boldsymbol{\nabla}_i\phi) + \mathcal{O}(\delta^5)$. Then we have [1]

$$\begin{aligned} F\boldsymbol{\eta}\cdot\boldsymbol{\nabla}H = \sum_\zeta \Bigg\{ &\frac{\delta}{\delta\zeta(\boldsymbol{x})}\left(\int_{R^3}\mathrm{d}^3x' FH\right)\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\zeta \\ &+ \boldsymbol{\nabla}\cdot\left[F\frac{\partial H}{\partial\boldsymbol{\nabla}_\perp\zeta}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\zeta + \boldsymbol{\nabla}_\perp\left(\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\zeta F\frac{\partial H}{\partial\Delta_\perp\zeta}\right)\right. \\ &\left. -2\boldsymbol{\nabla}_\perp\left(F\frac{\partial H}{\partial\Delta_\perp\zeta}\right)\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\zeta\right] \Bigg\} \end{aligned} \tag{57}$$

This equation is a useful identity and in fact holds for any vector field that commutes with $\boldsymbol{\nabla}_\perp$ and $\Delta_\perp$. It links the ordinary derivative on H to the variational derivative and correction terms that appear as exact divergences.

Summing over all species, integrating over velocity space and inserting our Hamiltonian from equation (18) we get

$$\begin{aligned} \sum_{\mathrm{s}} \|\boldsymbol{\eta}\cdot\boldsymbol{\nabla}H\| = &\frac{1}{\mu_0}\Delta_\perp A_{1,\parallel}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}A_{1,\parallel} \\ &+ \sum_{\mathrm{s}}\boldsymbol{\nabla}\cdot\left[-\frac{mN\boldsymbol{\nabla}_\perp\phi}{B^2}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\phi + \boldsymbol{\nabla}_\perp\left(\frac{m\|\mu B\|}{2qB^2}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\phi\right)\right. \\ &-\boldsymbol{\nabla}_\perp\left(\frac{m\|\mu B\|}{qB^2}\right)\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\phi - \boldsymbol{\nabla}_\perp\left(\frac{m\|\mu B v_\parallel\|}{2qB^2}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}A_{1,\parallel}\right) \\ &\left.+\boldsymbol{\nabla}_\perp\left(\frac{m\|\mu B v_\parallel\|}{qB^2}\right)\boldsymbol{\eta}\cdot\boldsymbol{\nabla}A_{1,\parallel}\right] \end{aligned} \tag{58}$$

Now, we focus on the term $q\dot{\psi}_p = q\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}\psi_p$ on the left hand side of equation (56). First we insert q into the velocity space moment equation (41). We find $\partial_t\|q\| + \boldsymbol{\nabla}\cdot\|q\dot{\boldsymbol{X}}\| = \|q\|_S$. Under species summation we see that we can identify the polarization equation (45) $\sum_{\mathrm{s}} qN = \boldsymbol{\nabla}\cdot\boldsymbol{P}_{\mathrm{gy}}$ and analogously the quasineutrality for the sources equation (54) $\sum_{\mathrm{s}} qS_N = \boldsymbol{\nabla}\cdot\boldsymbol{S}_P$. The next step is to apply the flux-surface average equation (7) to obtain $\partial_v\left\{\partial_t\langle\boldsymbol{P}_{\mathrm{gy}}\cdot\boldsymbol{\nabla}v\rangle + \sum_{\mathrm{s}}\left\langle\|q\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}v\|\right\rangle - \langle\boldsymbol{S}_P\cdot\boldsymbol{\nabla}v\rangle\right\} = 0$. After volume integration $\int_0^v \mathrm{d}v$ (the inner integration boundary vanishes) and multiplying with $\mathrm{d}\psi_p/\mathrm{d}v$ we obtain

$$\sum_{\mathrm{s}}\left\langle\|q\dot{\psi}_p\|\right\rangle = -\frac{\partial}{\partial t}\langle\boldsymbol{P}_{\mathrm{gy}}\cdot\boldsymbol{\nabla}\psi_p\rangle + \langle\boldsymbol{S}_P\cdot\boldsymbol{\nabla}\psi_p\rangle \tag{59}$$

which recovers the radial part of the polarization current $\boldsymbol{j}_{\rm pol} \equiv \partial \boldsymbol{P}_{\rm gy}/\partial t$. We stress that equation (59) is an important identity [1]. It links the derivative of the poloidal flux or in fact the first part of the toroidal angular momentum of particles to the polarization current and sources.

Now, we further investigate the terms appearing from equation (59) by explicitly inserting our polarization density (47) $-\frac{\partial}{\partial t}\langle \boldsymbol{P}_{\rm gy}\cdot\boldsymbol{\nabla}\psi_p\rangle = \sum_{\rm s}\frac{\partial}{\partial t}\left[mN\boldsymbol{\nabla}_\perp\phi\cdot\boldsymbol{\nabla}\psi_p/B^2 + m\boldsymbol{\nabla}_\perp(\|\mu B\|/2qB^2)\cdot\boldsymbol{\nabla}\psi_p\right]$. The second term can be simplified using the dynamical pressure equation $\mathrm{d}(\mu B)/\mathrm{d}t = \mu B\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}\ln B$ in equation (41) yielding $\frac{\partial}{\partial t}\|\mu B\| = -\boldsymbol{\nabla}\cdot\|\mu B\dot{\boldsymbol{X}}\| + \|\mu B\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}\ln B\| + \|\mu B\|_S$. We get the useful identity

$$\begin{aligned}&\frac{\partial}{\partial t}\left\langle\boldsymbol{\nabla}_\perp\left(\frac{m\|\mu B\|}{2qB^2}\right)\cdot\boldsymbol{\nabla}\psi_p\right\rangle - \left\langle\boldsymbol{\nabla}_\perp\left(\frac{m\|\mu B\|_S}{2qB^2}\right)\cdot\boldsymbol{\nabla}\psi_p\right\rangle\\ &= \frac{\partial}{\partial v}\left\langle\boldsymbol{\nabla}v\cdot\boldsymbol{\nabla}\left(\frac{m\|\mu B\|}{2qB^2}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}\phi - \frac{m\|\mu B v_\parallel\|}{2qB^2}\boldsymbol{\eta}\cdot\boldsymbol{\nabla}A_{1,\parallel}\right)\right\rangle\end{aligned} \tag{60}$$

One key ingredient for equation (60) is to use $(\boldsymbol{a}\cdot\boldsymbol{\nabla})\boldsymbol{b} = \boldsymbol{\nabla}(\boldsymbol{a}\cdot\boldsymbol{b}) - \boldsymbol{a}\times(\boldsymbol{\nabla}\times\boldsymbol{b}) - \boldsymbol{b}\times(\boldsymbol{\nabla}\times\boldsymbol{a}) - (\boldsymbol{b}\cdot\boldsymbol{\nabla})\boldsymbol{a}$ to show $\langle\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\psi_p\cdot\boldsymbol{\nabla}(\lambda\boldsymbol{u}_\perp))\rangle = \partial_v\langle\boldsymbol{\nabla}v\cdot\boldsymbol{\nabla}(\lambda\boldsymbol{u}_\perp\cdot\boldsymbol{\nabla}\psi_p)\rangle + \mathcal{O}(\delta^4)$ in our ordering. Now, we add the terms equation (58) and (60) and use our ordering to eliminate the magnetic field derivatives to get

$$\begin{aligned}&\frac{\partial}{\partial t}\left\langle\boldsymbol{\nabla}_\perp\left(\frac{m\|\mu B\|}{2qB^2}\right)\cdot\boldsymbol{\nabla}\psi_p\right\rangle + \langle\|\nabla_\eta H\|\rangle\\ &-\left\langle\boldsymbol{\nabla}_\perp\left(\frac{mS_{P_\perp}}{2qB^2}\right)\cdot\boldsymbol{\nabla}\psi_p\right\rangle\\ =&\frac{\partial}{\partial v}\left\langle\frac{\nabla^v A_{1,\parallel}\nabla_\eta A_{1,\parallel}}{\mu_0} - \frac{mN\nabla^v\phi\nabla_\eta\phi}{B^2}\right\rangle\\ &+\frac{\partial}{\partial v}\left\langle\frac{m\nabla^v A_{1,\parallel}\nabla_\eta\|\mu B v_\parallel\|}{qB^2} - \frac{m\nabla^v\phi\nabla_\eta\|\mu B\|}{qB^2}\right\rangle\end{aligned}$$

where we use the abbreviation $\nabla^v = \boldsymbol{\nabla}v\cdot\boldsymbol{\nabla}$ and $\nabla_\eta = \boldsymbol{\eta}\cdot\boldsymbol{\nabla}$ and imply species summation to present this intermediate result. We also used that $\boldsymbol{\eta}$ commutes with $\boldsymbol{\nabla}_\perp$ in our ordering and that the flux-surface average of $\boldsymbol{\nabla}\cdot(\boldsymbol{\eta}h)$ vanishes exactly.

Furthermore, we replace the gyro-centre quantities by their particle analogons, which is possible in our ordering since the correction terms are of higher order (see equations (51)). Finally, the term $\|\mathcal{A}_{1,\parallel}v_\parallel\|\boldsymbol{K}_\kappa\cdot\boldsymbol{\nabla}\psi_p$ in equation (56) vanishes under species summation and the parallel Ampère law to lowest order. With the help of equation (4) we then finally arrive at

$$\begin{aligned}&\frac{\partial}{\partial t}\sum_{\rm s}\langle mnu_{E,\varphi}\rangle + \frac{\partial}{\partial v}\sum_{\rm s}\langle mnu_{E,\varphi}(u^v_E+u^v_D)\rangle\\ &-\frac{\partial}{\partial v}\left\langle B_{1,\perp,\varphi}\left(\frac{1}{\mu_0}B^v_{1,\perp} - M^{{\rm em},v}_\perp\right)\right\rangle\\ &= -\left\langle\left(\boldsymbol{j}_f\times\boldsymbol{B}\right)_\varphi\right\rangle + \sum_{\rm s}\langle mS_n u_{E,\varphi}\rangle\end{aligned} \tag{61}$$

Here, we define the $\boldsymbol{E}\times\boldsymbol{B}$ drift $\boldsymbol{u}_E$, the grad-B drift $\boldsymbol{u}_{\boldsymbol{\nabla}B}$, the diamagnetic drift $\boldsymbol{u}_D$, the curvature drift $\boldsymbol{u}_\kappa$, the first order magnetic fluctuations $\boldsymbol{B}_{1,\perp}$ and the electromagnetic magnetization density $M^{\rm em}_\perp$ (different from $\boldsymbol{M}^{\rm gy}$ by a factor 2 and fluid instead of gyro-fluid quantities)

$$\boldsymbol{u}_E := \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\phi}{B} \qquad \boldsymbol{u}_{\boldsymbol{\nabla}B} := t_\perp\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B}{qB} \tag{62}$$

$$\boldsymbol{u}_D := \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}p_\perp}{qnB} \qquad \boldsymbol{u}_\kappa := (t_\parallel + mu^2_\parallel)\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\kappa}}{qB} \tag{63}$$

$$\boldsymbol{B}_{1,\perp} := \boldsymbol{\nabla}A_{1,\parallel}\times\hat{\boldsymbol{b}} \qquad \boldsymbol{M}^{\rm em}_\perp := \sum_{\rm s}\frac{m\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}(q_\parallel + p_\perp u_\parallel)}{qB^2} \tag{64}$$

and $\boldsymbol{b}_{1,\perp} := \boldsymbol{B}_{1,\perp}/B$. Equation (61) describes the evolution of the toroidal $\boldsymbol{E}\times\boldsymbol{B}$ angular momentum density and is the first result of this paper. The second term on the left side is the average over the convective acceleration term with radial velocity $u^v_E + u^v_D$, the sum of $\boldsymbol{E}\times\boldsymbol{B}$ and diamagnetic velocity. In section 5 we will show that this term can be split into an advective part and components of the turbulent stress tensor. Note that the appearance of the diamagnetic velocity in the gyro-kinetic momentum balance is a consequence of the pressure gradient in the polarization density (47) and thus ultimately a gyro-averaging effect. This contrasts to a drift-fluid model where diamagnetic velocity appears as a fluid-drift. The remaining terms on the left hand side are two stress terms stemming from magnetic fluctuations. On the right hand side the Lorentz force originating from the "free" current $\boldsymbol{j}_f := \sum_{\rm s} qn(\boldsymbol{u}_\kappa + \boldsymbol{u}_{\boldsymbol{\nabla}B})$ appears and we obtain a momentum source term proportional to the $\boldsymbol{E}\times\boldsymbol{B}$ velocity and the density source S_n.

Another point we note is that the poloidal analogue of equation (61) follows immediately. Recall equation (5) together with $\iota\boldsymbol{\nabla}\psi_t = \boldsymbol{\nabla}\psi_p$ in flux coordinates. This yields $u_{E,\vartheta} = \boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\psi_t/B^2 = \iota^{-1}u_{E,\varphi}$ and thus from equation (61) directly follows the equation for the *poloidal* $\boldsymbol{E}\times\boldsymbol{B}$ angular momentum density

$$\begin{aligned}&\frac{\partial}{\partial t}\sum_{\rm s}\langle mnu_{E,\vartheta}\rangle + \frac{\partial}{\partial v}\langle mnu_{E,\vartheta}(u^v_E+u^v_D)\rangle\\ &-\frac{\partial}{\partial v}\left\langle B_{1,\perp,\vartheta}\left(\frac{1}{\mu_0}B^v_{1,\perp} - M^{{\rm em},v}_\perp\right)\right\rangle\\ &+\left[\sum_{\rm s}\langle mnu_{E,\vartheta}(u^v_E+u^v_D)\rangle\right.\\ &\left.-\left\langle B_{1,\perp,\vartheta}\left(\frac{1}{\mu_0}B^v_{1,\perp} - M^{{\rm em},v}_\perp\right)\right\rangle\right]\frac{\partial}{\partial v}\ln\iota\\ &= -\left\langle\left(\boldsymbol{j}_f\times\boldsymbol{B}\right)_\vartheta\right\rangle + \sum_{\rm s}\langle mS_n u_{E,\vartheta}\rangle\end{aligned} \tag{65}$$

Equation (65) exhibits a similar structure as equation (61) with the additional appearance of the magnetic shear $\sigma := \partial_v\ln\iota$ [22]. Depending on its sign the shear term can both

dampen and generate poloidal $\boldsymbol{E}\times\boldsymbol{B}$ angular momentum. However, we emphasize that the shear appears as a purely geometrical correction to the poloidal momentum balance. Physically, equations (65) and (61) contain the same information since the two components of the $\boldsymbol{E}\times\boldsymbol{B}$ drift are related.

Before we continue with the identification of the various stress terms in section 5 and a more detailed interpretation of our results, we first derive the equations for the remaining angular momentum components, namely the poloidal and toroidal angular momentum components stemming from $u_\parallel$. As it turns out we will get the full parallel momentum balance as a by-product. Finally, recall that both $u_{E,\varphi} = \boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\psi_p/B^2$ and $u_{E,\vartheta} = \boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\psi_t/B^2$ are related to the radial electric field, a fact that will lead to the identification of the electromagnetic field angular momentum density in section 7.2.

4.2. Parallel (angular) momentum

We now turn to the parallel terms in the toroidal canonical momentum $\gamma_\varphi = q\psi_p + mw_\parallel b_\varphi$ as well as the poloidal canonical momentum $\gamma_\vartheta = q\psi_t + mw_\parallel b_\vartheta$. Repeating the ordering scheme from the previous section one could assume that $\partial u_\parallel/\partial t \sim \delta^2$ and argue that therefore only $\mathcal{O}(\delta^2)$ terms should be kept in our ordering. However, we note that the ions accelerate very slowly. This is because $n\dot{u}_{\parallel,i} \sim \nabla_\parallel p_i \sim \delta^3$. Note that this requires $\boldsymbol{\nabla}_\perp u_{\parallel,i}$ to be small as well. In contrast, the electron velocity is mainly determined by parallel Ohm's law $nu_{\parallel,e} \sim \eta_\parallel^{-1}(\nabla_\parallel p_e + \nabla_\parallel\phi) \sim \mathcal{O}(1)$ with $\eta_\parallel$ being the parallel resistivity. In order to reflect these considerations we order (in line with reference [41])

$$\partial_t u_{\parallel,i}/\Omega_{c,i}c_s \sim \delta^3 \tag{66}$$

This ordering mandates that the terms $\partial_t m_i u_{\parallel,i} \sim \partial_t m_i u_{E,\varphi} \sim \delta^3$ are similar in size. The parallel ion velocity itself is larger than the $\boldsymbol{E}\times\boldsymbol{B}$ velocity but we order its time derivative smaller by the same factor. In total, we again do not assume toroidal symmetry but we do use the drift ordering and keep terms up to $\mathcal{O}(\delta^3)$.

We start with (for $\eta\in\{\varphi,\vartheta\}$)

$$m\frac{\mathrm{d}}{\mathrm{d}t}(w_\parallel b_\eta) = mw_\parallel\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}b_\eta + m\dot{w}_\parallel b_\eta \tag{67}$$

With the vector triple product rule applied to $(\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}H)\times\boldsymbol{B}^*$ we see $m\dot{w}_\parallel\hat{\boldsymbol{b}} = q(\dot{\boldsymbol{X}}\times\boldsymbol{B}) + mw_\parallel\dot{\boldsymbol{X}}\times(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}}) - \mu B\boldsymbol{\nabla}\ln B - \boldsymbol{\nabla}H_f$. Next, we note $\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}b_\eta = -(\dot{\boldsymbol{X}}\times(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}}))_\eta + \dot{X}^i\partial_\eta b_i$ (Notice that we do not use the covariant derivative here since $b_\eta \equiv \hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{e}}_\eta$ is a scalar quantity and $(\boldsymbol{a}\cdot\boldsymbol{\nabla}\boldsymbol{b})_\eta \neq \boldsymbol{a}\cdot\boldsymbol{\nabla}b_\eta$; the first is the component of a covariant derivative while the second is the directional derivative of b_η). Finally, we have $q(\dot{\boldsymbol{X}}\times\boldsymbol{B}) = mv_\parallel w_\parallel\boldsymbol{K}_\kappa\times\boldsymbol{B} + \mu B\boldsymbol{K}_{\boldsymbol{\nabla}B}\times\boldsymbol{B} + (\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}H_f)\times\boldsymbol{B}/B$. We thus have

$$\begin{aligned} m\frac{\mathrm{d}}{\mathrm{d}t}(w_\parallel b_\eta) = {}& mv_\parallel w_\parallel(\boldsymbol{K}_\kappa\times\boldsymbol{B})_\eta + \mu B(\boldsymbol{K}_{\boldsymbol{\nabla}B}\times\boldsymbol{B})_\eta \\ & - b_\eta\nabla_\parallel H_f - \mu B\partial_\eta(\ln B) + mw_\parallel\dot{X}^i\partial_\eta b_i \end{aligned} \tag{68}$$

$$m\frac{\mathrm{d}}{\mathrm{d}t}w_\parallel = -\mu B\nabla_\parallel\ln B - \nabla_\parallel H_f - \frac{m}{q}w_\parallel\boldsymbol{K}_\kappa\cdot\boldsymbol{\nabla}H \tag{69}$$

The second identity follows immediately from the equations of motion (32). In the drift ordering (and under species summation to make $qv_\parallel A_{1,\parallel}$ vanish) we can write $mw_\parallel\dot{X}^i\partial_\eta b_i = mv_\parallel^2 b^i\partial_\eta b_i + \mathcal{O}(\delta^4)$, which we interpret as a generalized curvature contribution. With similar arguments as in the previous section we can recover the variational derivatives in $\sum_\mathrm{s}\|b_\eta\nabla_\parallel H_f\|$ using equation (57). However, the remaining terms are all $\mathcal{O}(\delta^5)$ and can be safely neglected in our ordering. Taking the velocity space moment we arrive at

$$\begin{aligned} & m\frac{\partial}{\partial t}\|w_\parallel\|b_\eta + m\boldsymbol{\nabla}\cdot(\|w_\parallel\dot{\boldsymbol{X}}\|b_\eta) + \|\mu B\|\partial_\eta\ln B \\ & = \|mv_\parallel^2(\boldsymbol{K}_\kappa\times\boldsymbol{B})_\eta + \mu B(\boldsymbol{K}_{\boldsymbol{\nabla}B}\times\boldsymbol{B})_\eta\| \\ & \quad + \|mv_\parallel^2\|B^i\partial_\eta b_i/B + m\|w_\parallel\|_S b_\eta \end{aligned} \tag{70}$$

$$m\frac{\partial}{\partial t}\|w_\parallel\| + m\boldsymbol{\nabla}\cdot(\|\dot{\boldsymbol{X}}w_\parallel\|) + \|\mu B\|\nabla_\parallel\ln B = m\|w_\parallel\|_S \tag{71}$$

We note that $m\frac{\partial}{\partial t}\|w_\parallel\| = m\frac{\partial}{\partial t}\|v_\parallel\| + q\frac{\partial}{\partial t}\|\mathcal{A}_{1,\parallel}\|$ and

$$\begin{aligned} & \|(mw_\parallel - q\mathcal{A}_{1,\parallel})\dot{\boldsymbol{X}}\| - m\|v_\parallel\|_S \\ & = \|mv_\parallel^2\|\left(\hat{\boldsymbol{b}} + \frac{\boldsymbol{\nabla}\times\mathcal{A}_{1,\parallel}\hat{\boldsymbol{b}}}{B}\right) + mNU_\parallel\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\psi}{B} \\ & \quad + \frac{m\|mv_\parallel^3\|}{q}\frac{\boldsymbol{\nabla}\times\hat{\boldsymbol{b}}}{B} + \frac{m\|\mu Bv_\parallel\|}{q}\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B}{B} \end{aligned}$$

Note here that the curvature terms vanish under the divergence in our ordering. As a final step we again apply the flux-surface average and note that with implied species summation the term $\partial\|q\mathcal{A}_{1,\parallel}\|/\partial t$ vanishes using the polarization equation. Then we have

for $\eta\in\{\varphi,\vartheta\}$

$$\begin{aligned} & \frac{\partial}{\partial t}\sum_\mathrm{s}\langle mnu_\parallel b_\eta\rangle + \sum_\mathrm{s}\frac{\partial}{\partial v}\left\langle mnu_\parallel b_\eta u_E^v + (p_\parallel + mnu_\parallel^2)b_\eta b_{1,\perp}^v\right\rangle \\ & = \sum_\mathrm{s} -\left\langle p_\perp\frac{\partial\ln B}{\partial\eta} + (p_\parallel + mnu_\parallel^2)b^i\frac{\partial b_i}{\partial\eta}\right\rangle \\ & \quad + \langle(\boldsymbol{j}_f\times\boldsymbol{B})_\eta\rangle + \sum_\mathrm{s} m\langle S_{nu_\parallel}b_\eta\rangle \end{aligned} \tag{72}$$

while the average parallel momentum reads

$$\sum_{\mathrm{s}}\left\{\frac{\partial}{\partial t}\left\langle mnu_{\parallel}\right\rangle+\frac{\partial}{\partial v}\left\langle mnu_{\parallel}u_E^v+(p_{\parallel}+mnu_{\parallel}^2)b_{1,\perp}^v\right\rangle\right\}$$
$$=\sum_{\mathrm{s}}\left\{\left\langle -p_{\perp}\nabla_{\parallel}\ln B\right\rangle+m\left\langle S_{nu_{\parallel}}\right\rangle\right\} \tag{73}$$

The two components of equation (72) complement the previously derived $\boldsymbol{E}\times\boldsymbol{B}$ velocity components in equations (65) and (61).

In total, Eqs (61), (65), (72) and (73) form the basis of our discussion for the remainder of the manuscript.

5. Favre averaged momentum equations

In order to discuss the effect of turbulent fluctuations on flux-surface averaged quantities a Reynolds decomposition is traditionally used to rewrite nonlinearities in the averaged evolution equations. For any function $h(\boldsymbol{x})$ we have

$$h\equiv\langle h\rangle+\widetilde{h} \tag{74}$$

Unfortunately, as we point out in [16], the Reynolds decomposition technique does not lead to well-behaved terms when the absolute density n appears in the non-linear terms in the sense that (i) absolute density fluctuations $\widetilde{n}$ appear instead of relative density fluctuations $\widetilde{n}/\langle n\rangle$, (ii) the radial advective part is not correctly recovered and (iii) effects from the density gradient $\partial_v\ln\langle n\rangle$ are not evident. We will thus follow [16] and introduce the so-called Favre decomposition.

Consider a term of the form $\langle nh\rangle$. If we multiply and divide by $\langle n\rangle$, we can write $\langle nh\rangle\equiv\langle n\rangle[\![h]\!]$. Here, we introduce the so-called Favre average

$$[\![h]\!]:=\frac{\langle nh\rangle}{\langle n\rangle} \tag{75}$$

which can be understood as a density weighted Reynolds average. We note that this definition is species dependent through the dependence on the species density n. The Favre average then allows the definition of the Favre decomposition

$$h\equiv[\![h]\!]+\widehat{h} \tag{76}$$

The Favre average reduces to the Reynolds average for small fluctuation amplitudes or if the density is a flux-function $[\![h]\!]=\langle h\rangle+\left\langle\widetilde{n}\widetilde{h}\right\rangle/\langle n\rangle\approx\langle h\rangle$. Reference [16] also reported $[\![u_\vartheta]\!]\approx\langle u_\vartheta\rangle$ within a few percent since $[\![\widehat{u_\vartheta}]\!]\approx 0$ in gyro-fluid simulations. We emphasize that the Favre average is a technique to present an equation in a way that can be easily interpreted physically. While it changes the appearance of an equation it does not change its content.

5.1. Favre averaged covariant $\boldsymbol{E}\times\boldsymbol{B}$ velocity

We first apply the Favre average technique to the continuity equation $\partial_t n+\boldsymbol{\nabla}\cdot(n\boldsymbol{u})=S_n$ to get

$$\frac{\partial}{\partial t}\langle n\rangle+\frac{\partial}{\partial v}\left(\langle n\rangle\mathcal{U}^v\right)=\langle S_n\rangle \tag{77}$$

where we define the average radial velocity

$$\mathcal{U}^v:=[\![u_E^v+u_{\parallel}b_{1,\perp}^v]\!] \tag{78}$$

and we use $u^v=u_E^v+u_{\parallel}b_{1,\perp}^v+\mathcal{O}(\delta^3)$. If we now replace all terms of the form $\langle nh\rangle$ with $\langle n\rangle[\![h]\!]$ in equation (61), then insert the continuity equation (77) and use $[\![gh]\!]=[\![g]\!][\![h]\!]+\left[\!\left[\widehat{g}\widehat{h}\right]\!\right]$ (equation (A10)) and $[\![u_D^v]\!]=\mathcal{O}(\delta^3)$, we get

$$\sum_{\mathrm{s}}\left\{m\langle n\rangle\left(\frac{\partial}{\partial t}+\mathcal{U}^v\frac{\partial}{\partial v}\right)[\![u_{E,\varphi}]\!]\right\}$$
$$=-\frac{\partial}{\partial v}\mathcal{T}_{\perp,\varphi}^v-\left\langle\left(\boldsymbol{j}_f\times\boldsymbol{B}\right)_{\varphi}\right\rangle+\sum_{\mathrm{s}}m\mathcal{S}_{u_{E,\varphi}} \tag{79}$$

and similarly in equation (65) we get

$$\sum_{\mathrm{s}}\left\{m\langle n\rangle\left(\frac{\partial}{\partial t}+\mathcal{U}^v\frac{\partial}{\partial v}\right)[\![u_{E,\vartheta}]\!]\right\}$$
$$=-\frac{\partial}{\partial v}\mathcal{T}_{\perp,\vartheta}^v-\left\langle\left(\boldsymbol{j}_f\times\boldsymbol{B}\right)_{\vartheta}\right\rangle+\sum_{\mathrm{s}}m\mathcal{S}_{u_{E,\vartheta}}$$
$$-\left(\sum_{\mathrm{s}}m\langle n\rangle[\![u_{E,\vartheta}]\!]\mathcal{U}^v+\Theta_{\vartheta}^v\right)\frac{\partial}{\partial v}\ln\iota \tag{80}$$

where we identify with $\left\langle B_{1,\perp}^v\right\rangle=\mathcal{O}(\delta^3)$ and $\langle M_{\perp}^{\mathrm{em},v}\rangle=\mathcal{O}(\delta^3)$ in the drift ordering

$$\mathcal{T}_{\perp,\eta}^v:=\sum_{\mathrm{s}}^{\text{for}\eta\in\{\varphi,\vartheta\}}m\langle n\rangle\mathcal{F}_{\perp,\vartheta}^v+\mathcal{M}_{\vartheta}^v \tag{81}$$

$$\mathcal{F}_{\perp,\eta}^v:=\underbrace{\left[\!\left[\widehat{u_{E,\eta}}\widehat{u_E^v}\right]\!\right]}_{\mathcal{F}_{E,\eta}^v}+\underbrace{\left[\!\left[\widehat{u_{E,\eta}}\widehat{u_D^v}\right]\!\right]}_{\mathcal{F}_{D,\eta}^v}\underbrace{-[\![u_{E,\eta}]\!][\![u_{\parallel}b_{1,\perp}^v]\!]}_{\mathcal{F}_{F,\eta}^v} \tag{82}$$

$$\mathcal{M}_{\eta}^v:=\underbrace{-\frac{1}{\mu_0}\left\langle\widetilde{B_{1,\perp,\eta}}\widetilde{B_{1,\perp}^v}\right\rangle}_{\mathcal{M}_{B,\eta}^v}+\underbrace{\left\langle\widetilde{B_{1,\perp,\eta}}\widetilde{M_{\perp}^{\mathrm{em},v}}\right\rangle}_{\mathcal{M}_{M,\eta}^v} \tag{83}$$

$$\mathcal{S}_{u_{E,\eta}}:=\left\langle\widetilde{S_n}\widetilde{u_{E,\eta}}\right\rangle+\langle S_n\rangle\left(\langle u_{E,\eta}\rangle-[\![u_{E,\eta}]\!]\right) \tag{84}$$

Note that the density $\langle n\rangle$ is species dependent and therefore we cannot divide equations (79) and (80) by $\langle n\rangle$. What is usually possible is to neglect the electron mass, which reduces the sum equations (79) and (80) to a sum over all ion species.

Equations (79) and (80) describe the evolution of the Favre averaged covariant components of the $\boldsymbol{E}\times\boldsymbol{B}$ velocity in general, not necessarily axisymmetric magnetic field geometry up to third order in the drift ordering. On the left hand side we find a radial advection term proportional to $\mathcal{U}^v$ [16]. The first term on the right hand side is the total perpendicular stress $\mathcal{T}^v_{\perp,\eta}$, which consists of the perpendicular Favre stress $\mathcal{F}^v_{\perp,\eta}$ and the Maxwell stress $\mathcal{M}^v_\eta$. We note here that we define the Favre stress as a kinematic stress (''stress divided by mass density'') with units $\mathrm{m}^2/\mathrm{s}^2$ as opposed to the Maxwell stress which has units of stress $\mathrm{N/m}^2$.

The kinematic Favre stress $\mathcal{F}^v_{\perp,\eta}$ contains the $\boldsymbol{E}\times\boldsymbol{B}$ Favre stress $\mathcal{F}^v_{E,\eta}$. As Reference [16] points out the $\boldsymbol{E}\times\boldsymbol{B}$ Favre stress $\mathcal{F}^v_{E,\eta}$ can be written as

$$\mathcal{F}^v_{E,\eta} = \mathcal{R}^v_{E,\eta} - \llbracket\widetilde{u_{E,\eta}}\rrbracket\left\llbracket\widetilde{u^v_E}\right\rrbracket + \left\langle\widetilde{nu_{E,\eta}}\widetilde{u^v_E}\right\rangle/\langle n\rangle \tag{85}$$

where the $\boldsymbol{E}\times\boldsymbol{B}$ Reynolds stress is $\mathcal{R}^v_{E,\eta} := \left\langle\widetilde{u_{E,\eta}}\widetilde{u^v_E}\right\rangle$ [42] and the often neglected [5] triple term appears on the right-hand side of equation (85). An advantage of the Favre decomposition is that the density fluctuations are automatically contained as relative fluctuation levels as is evident in the triple term in equation (85). An analogous identity to equation (85) holds for the diamagnetic Favre stress $\mathcal{F}^v_{D,\eta}$

$$\mathcal{F}^v_{D,\eta} = \mathcal{R}^v_{D,\eta} - \llbracket\widetilde{u_{E,\eta}}\rrbracket\left\llbracket\widetilde{u^v_D}\right\rrbracket + \left\langle\widetilde{nu_{E,\eta}}\widetilde{u^v_D}\right\rangle/\langle n\rangle\,, \tag{86}$$

which encompasses the diamagnetic Reynolds stress $\mathcal{R}^v_{D,\eta} = \left\langle\widetilde{u_{E,\eta}}\widetilde{u^v_D}\right\rangle$ [20]. Note that the diamagnetic Favre stress is asymmetric in contrast to $\boldsymbol{E}\times\boldsymbol{B}$ Favre stress. In this form the diamagnetic stress consist of the radial component of the diamagnetic velocity together with the η component of the $\boldsymbol{E}\times\boldsymbol{B}$ velocity. This is a consequence of using the pressure equation to evaluate the time-derivative of the diamagnetic velocity [20], which we have done using equation (60). Otherwise the transpose of the diamagnetic stress consisting of the radial $\boldsymbol{E}\times\boldsymbol{B}$ component and the η component of the diamagnetic velocity appears [43]. We elaborate further on different interpretations of the $\boldsymbol{E}\times\boldsymbol{B}$ angular momentum density in section 7.2. In addition to $\mathcal{F}^v_{E,\eta}$ and $\mathcal{F}^v_{D,\eta}$ we find the stress term $\mathcal{F}^v_{F,\eta}$ that appears for fluctuating magnetic field $\widetilde{b^v_{1,\perp}}$. This term is in fact a remainder of the actual magnetic flutter Favre stress term $m\left\llbracket\widehat{u_{E,\eta}u_\parallel b^v_{1,\perp}}\right\rrbracket$ that would appear, had we not neglected the $A_{1,\parallel}$ nonlinearities in the Hamiltonian (18) (through the variaton in equation (57)). We expect $\mathcal{F}^v_{F,\eta}$ to vanish for small relative density fluctuations and to only play a role for $\mathcal{O}(\widetilde{n}/\langle n\rangle)\sim 1$ fluctuation amplitudes, due to the similar dependence as the second term in the Favre stresses [16].

The Maxwell stress $\mathcal{M}^v_\eta$ consists of the symmetric vacuum field contribution $\mathcal{M}^v_{B,\eta}$ and the asymmetric magnetization stress term $\mathcal{M}^v_{M,\eta}$. The role of the vacuum Maxwell stress $\mathcal{M}^v_{B,\eta}$ on the generation of sheared $\boldsymbol{E}\times\boldsymbol{B}$ flows was highlighted previously in for example [44, 45]. The novel asymmetric magnetization stress $\mathcal{M}^v_{M,\eta}$ appears in its present form analogously to the diamagnetic stress as a consequence of using the pressure equation (60). In section 7.2 we will encounter its transpose in the full electromagnetic field stress tensor. It notably contains a contribution from the parallel heat flux $q_\parallel + p_\perp u_\parallel$ and physically originates in the magnetization term in parallel Ampère's law equation (46).

As was highlighted in [16] the density gradient $\partial_v \ln\langle n\rangle$ contributes to the evolution of $\boldsymbol{E}\times\boldsymbol{B}$ shear flow. Consider

$$\frac{\partial}{\partial v}\mathcal{T}^v_{\perp,\eta} = \sum_{\mathrm{s}}\left\{m\langle n\rangle\left(\frac{\partial}{\partial v}\mathcal{F}^v_{\perp,\eta} + \mathcal{F}^v_{\perp\eta}\frac{\partial}{\partial v}\ln\langle n\rangle\right)\right\} + \frac{\partial}{\partial v}\mathcal{M}^v_\eta \tag{87}$$

We emphasize that both $\boldsymbol{E}\times\boldsymbol{B}$ and diamagnetic Favre stresses appear in the density gradient drive term $m\langle n\rangle\mathcal{F}^v_{\perp,\eta}\partial_v\ln\langle n\rangle$ and that this term is non-zero even if $\partial_v\mathcal{F}^v_{\perp\eta}$ vanishes. This is particularly interesting for the steep density gradient that develops during the transition to H-mode.

Contrary to the toroidal angular momentum density, the poloidal $\boldsymbol{E}\times\boldsymbol{B}$ angular momentum density in equation (80) is influenced by a gradient in the rotational transform profile or magnetic shear $\sigma = \partial_v\ln\iota$. The magnetic shear is known to influence the $\boldsymbol{E}\times\boldsymbol{B}$ shear flow evolution [46, 47]. In particular the shear dampens drift-wave turbulence and leads to narrow zonal flows [47]. Furthermore, it dampens the Kelvin–Helmholtz instability, which would otherwise be driven by the $\boldsymbol{E}\times\boldsymbol{B}$ velocity shear [46]. In equation (80), we explicitly identify two magnetic shear contributions. The first shear term $m\langle n\rangle\mathcal{U}^v\llbracket u_{E,\vartheta}\rrbracket\sigma$ corresponds to roughly exponential growth or damping of poloidal flows, assuming that the average radial velocity is constant (which is a good estimate since it is approximately the $\boldsymbol{E}\times\boldsymbol{B}$ radial particle transport). The second shear term appears analogous to the density gradient $\partial_v\ln n$ term and contributes even if $\langle n\rangle\mathcal{F}^v_{\perp,\vartheta}$ and $\mathcal{M}^v_\vartheta$ are ''radially'' homogeneous (no volume derivative).

On the right hand side of equation (79) and (80) we further find the components of the Lorentz force originating from the radial curvature drift current $\boldsymbol{j}_f$ defined in equation (A21). The grad-B induced current part of this term is the Stringer-Winsor spin-up term [48–50]. In order to see this recall that $(\boldsymbol{j}_f\times\boldsymbol{B})_\varphi = \boldsymbol{j}_f\cdot\boldsymbol{\nabla}\psi_p \sim p_\perp\mathcal{K}(\psi_p)$ that is the radial component of the free current (see A1 for the definition of the curvature operator $\mathcal{K}$). The same term was found in δF drift-fluid models [18, 19, 45] and was there called the geodesic transfer term. In any case the term is known to excite geodesic acoustic modes and to both dampen or drive zonal flows depending on the parameter regime [18, 50]. We further discuss this term in relation to the ion orbit loss mechanism in section 7.4.

Finally, on the right side of equation (79) and (80) we find source related terms contained in $\mathcal{S}_{u_{E,\eta}}$ defined in equation (84). The term $\left\langle\widetilde{S_n}\widetilde{u_{E,\vartheta}}\right\rangle$ in equation (80) describes the poloidal spin-up mechanism for poloidally asymmetric particle sources described in [49]. In equation (79) we find an equivalent term also for the toroidal $\boldsymbol{E}\times\boldsymbol{B}$ velocity. A pol-

oidally (or toroidally) asymmetric particle source can generate or dampen toroidal $\boldsymbol{E}\times\boldsymbol{B}$ velocity. This should be contrasted with Reference [51], which finds angular momentum generation susceptible to the poloidal location of neutrals through viscosity and heat flux effects. In this contribution collisional effects are treated only indirectly subsuming the collision operator into the kinetic "source" term S in the Vlasov equation (21). The second source term is proportional to the difference between Reynolds and Favre averaged $\boldsymbol{E}\times\boldsymbol{B}$ velocity $\langle u_{E,\eta}\rangle - [\![u_{E,\eta}]\!]$. For small density fluctuations we thus expect this term to vanish and only contribute for large fluctuation amplitudes.

5.2. Favre averaged parallel velocity

For the parallel angular momentum components (72) we have

$$\begin{aligned}&\text{for } \eta\in\{\varphi,\vartheta\}\\ &\sum_s\left\{m\langle n\rangle\left(\frac{\partial}{\partial t}[\![u_\parallel b_\eta]\!]+\mathcal{U}^v\frac{\partial}{\partial v}[\![u_\parallel b_\eta]\!]\right)\right\}\\ &=-\frac{\partial}{\partial v}\mathcal{T}^v_{\parallel,\eta}+\sum_s\left\langle -p_\perp\frac{\partial \ln B}{\partial\eta}-(p_\parallel+mnu_\parallel^2)b^i\frac{\partial b_i}{\partial\eta}\right\rangle\\ &\quad+\langle(\boldsymbol{j}_f\times\boldsymbol{B})_\eta\rangle+\sum_s m\mathcal{S}_{u_\parallel b_\eta}\end{aligned}\tag{88}$$

where $\mathcal{U}^v$ is given in equation (78) and we identWithify

$$\mathcal{T}^v_{\parallel,\eta}:=\sum_s\left\{m\langle n\rangle\mathcal{F}^v_{\parallel,\eta}+\mathcal{K}^v_{\parallel,\eta}\right\}\tag{89}$$

$$\mathcal{K}^v_{\parallel,\eta}:=\left\langle\widetilde{p_\parallel}\,\widetilde{b_\eta}\,\widetilde{b^v_{1,\perp}}\right\rangle\tag{90}$$

$$\mathcal{F}^v_{\parallel,\eta}:=\underbrace{\left[\!\left[\widehat{u_\parallel b_\eta}\,\widehat{u^v_E}\right]\!\right]}_{\mathcal{F}^v_{E,\parallel,\eta}}+\underbrace{\left[\!\left[\widehat{u_\parallel b_\eta}\,\widehat{u_\parallel b^v_{1,\perp}}\right]\!\right]}_{\mathcal{F}^v_{F,\parallel,\eta}}\tag{91}$$

$$\mathcal{S}_{u_\parallel b_\eta}:=\langle S_{nu_\parallel}b_\eta\rangle-\langle S_n\rangle[\![u_\parallel b_\eta]\!]\tag{92}$$

the Favre average we re-write equation (73) into

$$\begin{aligned}&\sum_s\left\{m\langle n\rangle\left(\frac{\partial}{\partial t}[\![u_\parallel]\!]+\mathcal{U}^v\frac{\partial}{\partial v}[\![u_\parallel]\!]\right)\right\}\\ &=-\frac{\partial}{\partial v}\mathcal{T}^v_\parallel+\sum_s\left\{m\mathcal{S}_{u_\parallel}-\langle p_\perp\nabla_\parallel\ln B\rangle\right\}\end{aligned}\tag{93}$$

where $\mathcal{U}^v$ is given in equation (78) and we identify

$$\mathcal{T}^v_\parallel:=\sum_s\left\{m\langle n\rangle\mathcal{F}^v_\parallel+\mathcal{K}^v_\parallel\right\}\tag{94}$$

$$\mathcal{K}^v_\parallel:=\left\langle\widetilde{p_\parallel}\,\widetilde{b^v_{1,\perp}}\right\rangle\tag{95}$$

$$\mathcal{F}^v_\parallel:=\underbrace{\left[\!\left[\widehat{u_\parallel}\,\widehat{u^v_E}\right]\!\right]}_{\mathcal{F}^v_{E,\parallel}}+\underbrace{\left[\!\left[\widehat{u_\parallel}\,\widehat{u_\parallel b^v_{1,\perp}}\right]\!\right]}_{\mathcal{F}^v_{F,\parallel}}\tag{96}$$

$$\mathcal{S}_{u_\parallel}:=\langle S_{nu_\parallel}\rangle-\langle S_n\rangle[\![u_\parallel]\!]\tag{97}$$

Analogous to equation (79) in equations (93) and (88) we find a radial advection term of momentum by $\mathcal{U}^v$ followed by various stress terms contained in $\mathcal{T}^v_\parallel$ respectively $\mathcal{T}^v_{\parallel,\eta}$. Again, we define the Favre stress as a kinematic stress and analogous relations to equation (85) hold for the parallel Favre stress components. The parallel $\boldsymbol{E}\times\boldsymbol{B}$ Favre stress $\mathcal{F}^v_{E,\parallel}$ respectively $\mathcal{F}^v_{E,\parallel,\eta}$ now depends on fluctuations in the parallel velocity instead of $\boldsymbol{E}\times\boldsymbol{B}$ velocity. The Reynolds stress analogue of $\mathcal{F}^v_{E,\parallel}$ is well-known in the literature on intrinsic toroidal rotation (see e.g. [5]), however we point out here that $\mathcal{F}^v_{E,\parallel,\eta}$ is the actual component that drives angular momentum $[\![u_\parallel b_\eta]\!]$ instead of just parallel momentum $[\![u_\parallel]\!]$. The parallel magnetic flutter Favre stress term $\mathcal{F}^v_{F,\parallel}$ respectively $\mathcal{F}^v_{F,\parallel,\eta}$ is a transfer term appearing for magnetic fluctuations $b_{1,\perp}$. The kinetic stress term $\mathcal{K}^v_\parallel$ respectively $\mathcal{K}^v_{\parallel,\eta}$ is related to the kinetic dynamo mechanism as for example discussed for the reversed field pinch in Reference [52, 53]. On the right hand side we find the mirror force term $-\langle p_\perp\nabla_\parallel\ln B\rangle$ respectively $-\langle b_\eta p_\perp\nabla_\parallel\ln B\rangle$. In the equation for the parallel angular momentum equation (88) we find an additional geometrical correction to the mirror force. Finally, the momentum source term $\mathcal{S}_{u_\parallel}$ respectively $\mathcal{S}_{u_\parallel b_\eta}$ represents angular momentum generation by external sources. Note that with the definition of a velocity source $S_{u_\parallel}$ via $S_{nu_\parallel}:=nS_{u_\parallel}+u_\parallel S_n$ we can write

$$\mathcal{S}_{u_\parallel}=\langle n\rangle[\![S_{u_\parallel}]\!]+\left\langle\widetilde{u_\parallel}\widetilde{S_n}\right\rangle+\langle S_n\rangle\left(\langle u_\parallel\rangle-[\![u_\parallel]\!]\right)\tag{98}$$

and analogous for $\mathcal{S}_{u_\parallel b_\eta}$. Equation (98) now consists of the Favre averaged velocity source plus a contribution from a poloidally asymmetric source term analogous to equation (84).

We comment here on the appearance of the Lorentz force in the equation for the parallel angular momentum equation (88). The Lorentz force acts perpendicularly to the magnetic field line and should not contribute to the parallel momentum at all. Indeed, we can further simplify the right hand side of equation (88) to

$$\begin{aligned}&\sum_s\left\langle -p_\perp\frac{\partial\ln B}{\partial\eta}-(p_\parallel+mnu_\parallel^2)b^i\frac{\partial b_i}{\partial\eta}\right\rangle+\langle(\boldsymbol{j}_f\times\boldsymbol{B})_\eta\rangle\\ &\quad=-\sum_s\left\langle p_\perp b_\eta\nabla_\parallel\ln B+(p_\parallel+mnu_\parallel^2)\tilde{\kappa}_\eta\right\rangle\end{aligned}\tag{99}$$

where we define $\tilde{\kappa}_\eta:=\boldsymbol{e_\eta}\cdot\boldsymbol{\kappa}-b^i\partial_\eta b_i$ with the curvature $\boldsymbol{\kappa}:=\hat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\hat{\boldsymbol{b}}$. Now, only the component of the mirror force

$-p_\perp b_\eta \nabla_\parallel \ln B$ and a geometric correction term appear. To see the mirror force recall the sign of the magnetic moment vector $\boldsymbol{\mu} = -\mu\hat{\boldsymbol{b}}$ and the guiding centre parallel magnetization density [27, 30] $M^{\text{gy}}_\parallel = \|\boldsymbol{\mu}\| = -\|\mu\|\hat{\boldsymbol{b}} = -P_\perp\hat{\boldsymbol{b}}/B$. The force acting on magnetic dipoles is [54] $\boldsymbol{f}_d = \boldsymbol{\nabla}(\boldsymbol{\mu}\cdot\boldsymbol{B}) = -\mu\boldsymbol{\nabla}B$. Taking the velocity space moment we get $\boldsymbol{F}_d = \|\boldsymbol{f}_d\| = M_\parallel \boldsymbol{\nabla}B = -P_\perp\boldsymbol{\nabla}\ln B = -P_\perp\hat{\boldsymbol{b}}\nabla_\parallel \ln B - P_\perp\boldsymbol{\nabla}_\perp \ln B$. The parallel part reads $F_{d,\parallel} = -P_\perp\nabla_\parallel \ln B$, which is what appears in equation (73), while $F_{d,\eta} = -P_\perp\partial_\eta \ln B$ appears in (72). The perpendicular part gives rise to the $\boldsymbol{\nabla}B$ drift. Last, notice that $\nabla_\parallel \ln B = -\boldsymbol{\nabla}\cdot\hat{\boldsymbol{b}}$ such that

$$-\left\langle p_\perp \nabla_\parallel \ln B\right\rangle = \left\langle \nabla_\parallel \widetilde{p_\perp}\right\rangle \tag{100}$$

$$-\left\langle b_\eta p_\perp \nabla_\parallel \ln B\right\rangle = \left\langle \nabla_\parallel \widetilde{(b_\eta p_\perp)}\right\rangle \tag{101}$$

Pressure fluctuations are required to affect the the angular momentum generation via the mirror force.

5.3. Total angular momentum density

The velocity equations (79)/(80) and (88) can be easily cast back into conservative form using the continuity equation (77) and $\langle n\rangle[\![h]\!] = \langle nh\rangle$ for any h. Summing up the results, we finally find the evolution of the total average poloidal and toroidal angular momentum density

$$\begin{aligned}
&\text{for } \eta \in \{\varphi, \vartheta\}\\
&\sum_{\text{s}}\left\{\frac{\partial}{\partial t} m\left\langle n(u_\parallel b_\eta + u_{E,\eta})\right\rangle\right.\\
&\quad\left. + \frac{\partial}{\partial v} m\left\langle n(u_\parallel b_\eta + u_{E,\eta})\right\rangle \mathcal{U}^v\right\} + \frac{\partial}{\partial v}\left(\mathcal{T}^v_{\perp,\eta} + \mathcal{T}^v_{\parallel,\eta}\right)\\
&= -\delta_{\eta\vartheta}\left(\sum_{\text{s}} m\left\langle n u_{E,\vartheta}\right\rangle \mathcal{U}^v + \mathcal{T}^v_{\perp,\vartheta} + \mathcal{T}^v_{\parallel,\vartheta}\right)\frac{\partial}{\partial v}\ln\iota\\
&-\sum_{\text{s}}\left\langle p_\perp \frac{\partial \ln B}{\partial \eta} + (p_\parallel + mnu_\parallel^2) b^i \frac{\partial b_i}{\partial \eta} - m\left(S_{nu_\parallel} b_\eta + S_n u_{E,\eta}\right)\right\rangle
\end{aligned} \tag{102}$$

where δ is the Kronecker delta. The magnetic shear term only contributes to the poloidal angular momentum. The convective term proportional to $\mathcal{U}^v$ vanishes under volume integration up to a surface contribution as does the total stress term $\mathcal{T}^v_{\perp,\eta} + \mathcal{T}^v_{\parallel\eta}$. In equation (102) we further find that the momentum transfer to the background magnetic field is mediated by the mirror force and the generalized curvature force term on the right hand side. Clearly, the Lorentz force term cancels in the total angular momentum density evolution. Finally, we recover the external source terms on the right hand side.

We see that the total angular momentum in equation (102) is given by the covariant components of the $\boldsymbol{E}\times\boldsymbol{B}$ and parallel velocities. Comparing this to the total advection velocity $\boldsymbol{u} := \|\dot{\boldsymbol{X}}\| = \boldsymbol{u}_E + \boldsymbol{u}_\kappa + \boldsymbol{u}_{\boldsymbol{\nabla}B} + u_\parallel\hat{\boldsymbol{b}} + u_\parallel\boldsymbol{b}_{1,\perp}$ that appears in the continuity equation $\partial_t n + \boldsymbol{\nabla}\cdot(n\boldsymbol{u}) = S_n$ we see that the curvature, grad-B and magnetic flutter velocities do not appear in the angular momentum (102) even though we at least expected the magnetic flutter term as an order $\mathcal{O}(\delta)$ term. At this point recall equation (59), which identifies the radial polarization current with the macroscopic expression for the angular momentum density (except $u_\parallel b_\eta$). The polarization density $\boldsymbol{P}_{\text{gy}}$ is directly connected to the definition of the Hamiltonian (18) through the variational principle. Since we neglected the second order guiding center corrections we accordingly miss the guiding center polarization density [2, 30] and thus the corresponding curvature terms in our angular momentum density. On the other hand we also neglected the non-linear terms in $A_{1,\parallel}$ in the Hamiltonian, which accounts for the missing magnetic flutter velocity $mv_\parallel b_{1,\perp}$ in the polarization [27] and thus angular momentum density (102).

6. The rotational energy

6.1. Angular momentum and angular velocity

In section 5 we have derived equations for the covariant components of the $\boldsymbol{E}\times\boldsymbol{B}$ and parallel velocity, which add up to the total angular momentum density in equation (102). We now focus on the angular momentum as a vector quantity. We define

$$\boldsymbol{u}_L := \boldsymbol{u}_E + u_\parallel\hat{\boldsymbol{b}} = u_{L,\varphi}\boldsymbol{\nabla}\varphi + u_{L,\vartheta}\boldsymbol{\nabla}\vartheta + u_{L,v}\boldsymbol{\nabla}v \tag{103}$$

We are now interested only in the part of the flow that stays within a given flux-surface, because this flow can be constructed from the covariant φ and ϑ components of $\boldsymbol{u}_L$ that we have available. To see this, we formulate the projection tensor onto the flux surfaces

$$h_S := 1 - \hat{\boldsymbol{\rho}}\hat{\boldsymbol{\rho}} \tag{104}$$

with the contravariant radial unit vector $\hat{\boldsymbol{\rho}} := \boldsymbol{\nabla}v/|\boldsymbol{\nabla}v|$. With this we can split the flow velocity according to $\boldsymbol{u}_L = \boldsymbol{u}_L|_{\psi_p} + u_L^{\hat{\rho}}\hat{\boldsymbol{\rho}}$ where we define the surface or rotational velocity

$$\begin{aligned}
\boldsymbol{L} \equiv \boldsymbol{u}_L|_{\psi_p} := h_S\cdot\boldsymbol{u}_L &= u_{L,\vartheta}\boldsymbol{\nabla}_S\vartheta + u_{L,\varphi}\boldsymbol{\nabla}_S\varphi\\
&= u_L^\vartheta\boldsymbol{e}_{\boldsymbol{\vartheta}} + u_L^\varphi\boldsymbol{e}_{\boldsymbol{\varphi}}
\end{aligned} \tag{105}$$

where we follow [25] and introduce the surface operator $\boldsymbol{\nabla}_S := h_S\cdot\boldsymbol{\nabla}$. We thus have $L_i = u_{E,i} + u_\parallel b_i$ for $i\in\{\varphi,\vartheta,\rho\}$. As expected we do not need the radial component of $\boldsymbol{u}_L$ to construct the surface flow in equation (105).

It is now important to see that $\boldsymbol{\nabla}_S\varphi$ and $\boldsymbol{\nabla}_S\vartheta$ form the contravariant basis of the flux surface as a stand-alone manifold and analogous $\boldsymbol{e}_{\boldsymbol{\varphi}}$ and $\boldsymbol{e}_{\boldsymbol{\vartheta}}$ are its covariant basis vectors In fact, explicitly writing h_S into components we realize that all components $h_{S,\rho k}$ vanish for $k\in\{\varphi,\vartheta,\rho\}$. We thus define $\mathcal{I}$ as the two-dimensional tensor consisting of the non-zero components of h_S, that is

$$\mathcal{I} := \begin{pmatrix} g_{\vartheta\vartheta} & g_{\vartheta\varphi}\\ g_{\varphi\vartheta} & g_{\varphi\varphi}\end{pmatrix} \tag{106}$$

The interested reader will recognize $\mathcal{I}$ as the the first fundamental form of flux surfaces parameterized with ϑ and φ. The first fundamental form $\mathcal{I}$ can be interpreted as the two-dimensional metric tensor of the flux-surface thought as a standalone structure and is thus an intrinsic structure of the magnetic flux surfaces (and in particular has a well-defined expression in every coordinate system). Unfortunately, the flux-surface average is not an intrinsic surface operation since it requires the knowledge of the volume form $\sqrt{g}$ to compute. Also, note that the components of $\mathcal{I}$ and its inverse are given by $\mathcal{I}_{\varphi\varphi} = \boldsymbol{e}_{\boldsymbol{\varphi}} \cdot \boldsymbol{e}_{\boldsymbol{\varphi}}$, $\mathcal{I}_{\vartheta\varphi} = \boldsymbol{e}_{\boldsymbol{\vartheta}} \cdot \boldsymbol{e}_{\boldsymbol{\varphi}}$, $\mathcal{I}_{\vartheta\vartheta} = \boldsymbol{e}_{\boldsymbol{\vartheta}} \cdot \boldsymbol{e}_{\boldsymbol{\vartheta}}$ and $\mathcal{I}^{\varphi\varphi} = \boldsymbol{\nabla}_S\varphi \cdot \boldsymbol{\nabla}_S\varphi$, $\mathcal{I}^{\vartheta\varphi} = \boldsymbol{\nabla}_S\vartheta \cdot \boldsymbol{\nabla}_S\varphi$, $\mathcal{I}^{\vartheta\vartheta} = \boldsymbol{\nabla}_S\vartheta \cdot \boldsymbol{\nabla}_S\vartheta$ respectively.

Now, the fundamental form $\mathcal{I}$ has another interpretation, namely as the inertia tensor of rotations in ϑ and φ. To see this recall that the *contravariant* components of the surface velocity $\boldsymbol{L}$, L^φ and L^ϑ are actually the *angular velocities* with units s^{-1}, because the particle trajectory is given by $\dot{\varphi} = L^\varphi$ and $\dot{\vartheta} = L^\vartheta$. In contrast, the *covariant* components $L_i = \mathcal{I}_{ij}L^j$ for $i,j \in \{\vartheta, \varphi\}$ form the *angular momentum* as it results in equation (102) that is mL_i has units kg $\mathrm{m}^2\mathrm{s}^{-1}$. This leaves $m\mathcal{I}$ as the (kinematic) inertia tensor that connects the angular velocity and angular momentum of a fluid element rotating on a flux-surface.

6.2. Mean and fluctuating angular momentum

Consider now the mean surface velocity field generated by Favre averaged covariant φ and ϑ velocity components

$$\boldsymbol{L}_m := [\![L_\vartheta]\!]\,\boldsymbol{\nabla}_S\vartheta + [\![L_\varphi]\!]\,\boldsymbol{\nabla}_S\varphi \tag{107}$$

The time evolution of $m\langle n\rangle \boldsymbol{L}_m$ is directly given by equation (102). First, we emphasize that the corresponding *angular velocity* components of $\boldsymbol{L}_m$, $L^i = \mathcal{I}^{ij}[\![L_i]\!]$ are *not* flux functions since the inertia tensor does not commute with the flux-surface average and thus $\boldsymbol{L}_m \neq [\![L^\vartheta]\!]\,\boldsymbol{e}_\vartheta + [\![L^\varphi]\!]\,\boldsymbol{e}_\varphi$ or in other words, if angular momentum is a flux-function then angular velocity cannot be at the same time. In fact, we perform the splitting $L_i = [\![L_i]\!] + \widehat{L}_i$ expecting that the relative fluctuations $\widehat{L}_i$ are small and that u_i is well-described by its Favre average $[\![L_i]\!]$. *A priori*, these arguments of course also hold the other way, if angular velocity were a flux-function then angular momentum cannot be at the same time and we should split the angular velocities.

At this point recall the discussion in the introduction. When angular momentum is conserved, a particle moves faster closer to the axis (for example on the high field side in figure 1). We take this as an indication that angular velocities are not well-described by flux-surface averages, while angular momenta are. Furthermore, in the equations in section 5 (for example equation (79)) we see that the average *angular momentum* is fed by turbulent fluctuations through the stress tensor, which we interpret as an indication that fluctuations $\widehat{L}_i$ and not $\widehat{L}^i$ become small.

6.3. Total energy evolution

Before we construct a zonal or mean flow rotational energy we first focus on the total energy evolution of our system. We follow Reference [39] and derive the pressure Equations (the thermal energy) for $p_\perp$ and $p_\parallel$ directly from the moment evolution equation (41). We point out that we need to keep terms one order higher in the energy conservation law than in the momentum conservation law, that is $\mathcal{O}(\delta^4)$ in our ordering. This is due to the fundamental property of the gyro-kinetic system [27] that a higher order Hamiltonian needs to be kept in the system to obtain polarization effects and an exact energy invariant. If we thus neglect all terms of order $\mathcal{O}(\delta^5)$, use parallel Ampère's law (46) and apply the species summation we get

$$\frac{\partial}{\partial t}\left\langle \sum_{\mathrm{s}} \left\{ p_\perp + \frac{1}{2}p_\parallel + \frac{1}{2}mnu_\parallel^2 \right\} + \frac{(\boldsymbol{\nabla}_\perp A_{1\parallel})^2}{2\mu_0} \right\rangle + \frac{\partial}{\partial v}\left\langle j^v_{\mathcal{E},p} \right\rangle = \left\langle \boldsymbol{j}_f \cdot \boldsymbol{E}_\perp + j_\parallel E_\parallel \right\rangle + \sum_{\mathrm{s}} \left\langle S_{p_\perp} + \frac{1}{2}S_{p_\parallel + mnu_\parallel^2} \right\rangle \tag{108}$$

where $\boldsymbol{E}_\perp = -\boldsymbol{\nabla}_\perp\phi$ and $E_\parallel = -(\nabla_\parallel\phi + \boldsymbol{b}_{1,\perp}\cdot\boldsymbol{\nabla}\phi)$ and $j_\parallel := \sum_{\mathrm{s}} qnu_\parallel$. We formally summarize all total divergences into the term $j^v_{\mathcal{E},p}$. An interesting side-remark here is to view the energy conservation equation (108) to lowest order, which leaves Bernoulli's identity $\left\langle p_\perp + p_\parallel/2 + mu_\parallel^2/2 \right\rangle = const$ along fluid trajectories. On the right side of equation (108) appears the energy exchange term $\left\langle \boldsymbol{j}_f \cdot \boldsymbol{E}_\perp + j_\parallel E_\parallel \right\rangle$ as well as the pressure source terms (heating).

On the other side using the definition of Ψ in equation (20) and the polarization equation (45) we find

$$\sum_{\mathrm{s}} \|q\Psi\| = \sum_{\mathrm{s}} \boldsymbol{\nabla}\cdot\left(\frac{m\|\mu B\|}{2qB^2}\boldsymbol{\nabla}_\perp\phi - \phi\boldsymbol{\nabla}_\perp\frac{m\|\mu B\|}{2qB^2} \right) - \boldsymbol{\nabla}\left(\phi\frac{mN\boldsymbol{\nabla}_\perp\phi}{B^2} \right) + \frac{1}{2}mN\frac{(\boldsymbol{\nabla}_\perp\phi)^2}{B^2} \tag{109}$$

which recovers the $\boldsymbol{E}\times\boldsymbol{B}$ kinetic energy density in the last term on the right hand side. Interestingly, a completely analogous relation holds for the term $\|q\Psi\|_S$ (by replacing $\|\mu B\|$ with $\|\mu B\|_S$ and N with S_N in equation (109)) since we require the sources to preserve quasineutrality in equation (54). Applying equation (41) to $q\dot{\Psi} = q\partial_t\Psi + q\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}\Psi$ and using (109) and (58) for Ψ under species summation and neglecting again terms of order $\mathcal{O}(\delta^5)$ the result is given by

$$\frac{\partial}{\partial t}\left\langle \frac{1}{2}\rho_M u_E^2 \right\rangle + \frac{\partial}{\partial v}\left\langle j^v_{\mathcal{E},\psi} \right\rangle = -\left\langle \boldsymbol{j}_f\cdot\boldsymbol{E}_\perp + j_\parallel E_\parallel \right\rangle + \sum_{\mathrm{s}} \frac{1}{2}m\left\langle S_n u_E^2 \right\rangle \tag{110}$$

where we identify the total mass density $\rho_M := \sum_{\mathrm{s}} mn$ since the $\boldsymbol{E}\times\boldsymbol{B}$ velocity is species independent and again

summarize all divergences into the formal $j^v_{\mathcal{E},\psi}$ term. The density source S_n either generates or destroys kinetic $\boldsymbol{E}\times\boldsymbol{B}$ energy depending on its sign. The term appears analogous to the momentum source in equation (61). The sum of equations (108) and (110) recovers the conservation of the flux-surface averaged total energy of our model since the energy exchange term $\left\langle \boldsymbol{j}_f\cdot\boldsymbol{E}_\perp + j_\parallel E_\parallel\right\rangle$ cancels.

6.4. Mean rotational energy evolution

The direct approach to a rotational energy density is the kinetic energy of the surface flow velocity $\boldsymbol{L}$

$$E_{\text{rot}} := \sum_{\text{s}} \frac{1}{2} m \left\langle n\boldsymbol{L}\cdot\mathcal{I}\boldsymbol{L}\right\rangle = \left\langle \frac{1}{2}\rho_M \boldsymbol{u}_E^2|_{\psi_p}\right\rangle + \sum_{\text{s}}\left\langle \frac{1}{2} mnu_\parallel^2\right\rangle \tag{111}$$

This energy is equivalent to subtracting the radial $\boldsymbol{E}\times\boldsymbol{B}$ energy $\langle \rho_M u_{E,v} u_E^v/2\rangle$ from the total kinetic energy density $\sum_{\text{s}}\left\langle mn(\boldsymbol{u}_E^2+u_\parallel^2)/2\right\rangle$. It is now important to realize that contrary to the parallel kinetic energy the $\boldsymbol{E}\times\boldsymbol{B}$ rotational energy density can be related to the (species summed) angular momentum evolution. This is because the $\boldsymbol{E}\times\boldsymbol{B}$ drift velocity is equal for all species. We can write

$$\frac{1}{2}\left\langle \rho_M \boldsymbol{u}_E|^2_{\psi_p}\right\rangle = E_{\text{zonal}} + E_{\text{fluc}} \tag{112}$$

where we define

$$E_{\text{zonal}} := \frac{1}{2}\left\langle\rho_M\right\rangle \left[\!\left[\mathcal{I}^{ij}\right]\!\right]_M \left[\!\left[u_{E,i}\right]\!\right]_M \left[\!\left[u_{E,j}\right]\!\right]_M \tag{113}$$

$$\begin{aligned} E_{\text{fluc}} :=& \left\langle\rho_M\right\rangle \left[\!\left[\mathcal{I}^{ij}\left[\!\left[u_{E,i}\right]\!\right]_M \widehat{u_{E,j}}\right]\!\right]_M \\ &+ \frac{1}{2}\left\langle\rho_M\right\rangle \left[\!\left[\mathcal{I}^{ij}\right]\!\right]_M \left[\!\left[\widehat{u_{E,i}}\right]\!\right]_M \widehat{u_{E,j}} \end{aligned} \tag{114}$$

and here introduce the total mass density in the Favre averages

$$\left[\!\left[h\right]\!\right]_M := \left\langle\rho_M\right\rangle^{-1} \sum_{\text{s}} m\left\langle nh\right\rangle \tag{115}$$

for any (possibly species dependent) function h. If h is species independent equation (115) simplifies to $\left[\!\left[h\right]\!\right]_M = \left\langle\rho_M h\right\rangle / \left\langle\rho_M\right\rangle$. With $u_{E,\vartheta} = \iota^{-1} u_{E,\varphi}$ we can simplify further

$$\begin{aligned} E_{\text{zonal}} &= \frac{1}{2}\left\langle\rho_M\right\rangle \left[\!\left[\iota^{-2}\mathcal{I}^{\vartheta\vartheta} + 2\iota^{-1}\mathcal{I}^{\vartheta\varphi} + \mathcal{I}^{\varphi\varphi}\right]\!\right]_M \left[\!\left[u_{E,\varphi}\right]\!\right]_M^2 \\ &\equiv \frac{1}{2}\left\langle\rho_M\right\rangle \left[\!\left[u_{E,\varphi}\right]\!\right]_M^2 \left[\!\left[\mathcal{I}_0\right]\!\right]_M \end{aligned} \tag{116}$$

Here, we introduce the inertia factor $\mathcal{I}_0 := (\iota^{-1},1)\mathcal{I}^{-1}(\iota^{-1},1)^{\text{T}}$. For a purely toroidal magnetic field we have $\mathcal{I}_0 = R^{-2}$ as expected. For symmetry flux coordinates we have $g_{\vartheta\vartheta} = R^2(\nabla\psi_p)^2/I^2\iota^2$, $g_{\varphi\vartheta}=0$ and $g_{\varphi\varphi}=R^2$ and thus $\mathcal{I}_0 = R^{-2}(1+I^2/|\nabla\psi_p|^2) = B^2/|\nabla\psi_p|^2$. The inertia factor vanishes for a slab magnetic field. In this case our zonal flow energy agrees with [16] and in the case of small density fluctuations also with its δF analogue [19, 45]. Since $\mathcal{I}_0$ is time-independent we can use the evolution equations for the density equation (77) and angular momentum (79) to get

$$\begin{aligned} &\frac{\partial}{\partial t}E_{\text{zonal}} + \frac{\partial}{\partial v}\left(E_{\text{zonal}}\left[\!\left[u^v\right]\!\right]_M\right) \\ &= -\left[\!\left[\mathcal{I}_0\right]\!\right]_M \left[\!\left[u_{E,\varphi}\right]\!\right]_M \left(\frac{\partial}{\partial v}\mathcal{T}^v_{\perp,\varphi} + \left\langle(\boldsymbol{j}_f\times\boldsymbol{B})_\varphi\right\rangle\right) \\ &\quad - \frac{1}{2}\left[\!\left[u_{E,\varphi}\right]\!\right]_M^2 \frac{\partial}{\partial v}\left(\left\langle\rho_M\right\rangle\left[\!\left[\widehat{\mathcal{I}_0}\widehat{u^v}\right]\!\right]_M\right) + \mathcal{S}_{\text{zonal}} \end{aligned} \tag{117}$$

where we neglected the term $\langle n\boldsymbol{u}\cdot\nabla\mathcal{I}_0\rangle$ in the continuity equation as small in our ordering and we have

$$\mathcal{S}_{\text{zonal}} := \left[\!\left[\mathcal{I}_0\right]\!\right]_M \left[\!\left[u_{E,\varphi}\right]\!\right]_M \mathcal{S}_{u_{E,\varphi}} + \frac{1}{2}\left[\!\left[u_{E,\varphi}\right]\!\right]_M^2 \sum_{\text{s}}\left\langle mS_n\mathcal{I}_0\right\rangle \tag{118}$$

In equation (117) we find the term $\left[\!\left[u^v\right]\!\right]_M$ as the convective velocity for the zonal flow energy. On the right side the derivative of the total perpendicular stress $\mathcal{T}^v_{\perp,\varphi}$ given by equation (87) appears. Thus, the Favre and Maxwell stress given by fluctuating velocities and the fluctuating magnetic field in equations (82) and (83) respectively together with a gradient in the density $\partial_v \ln\langle n\rangle$ can appear as sources for zonal flow energy. The $\boldsymbol{E}\times\boldsymbol{B}$ Favre stress $\mathcal{F}^v_{E,\varphi}$ was already identified as a source for zonal flow energy in a slab geometry in [16]. The vacuum field Maxwell stress $\mathcal{M}^v_{B,\varphi}$ and the $\boldsymbol{E}\times\boldsymbol{B}$ Reynolds stress $\mathcal{R}^v_{E,\varphi}$ (contained in our $\mathcal{F}^v_{\perp,\varphi}$ according to equation (85)) appear in similar form in δF models [19, 45]. Compared to these previous findings we find the additional appearance of the diamagnetic Favre stress $\mathcal{F}^v_{D,\varphi}$ contained in $\mathcal{F}^v_{\perp,\eta}$ and the magnetization stress $\mathcal{M}^v_{M,\varphi}$ contained in $\mathcal{M}^v_\varphi$. In addition, we find the inertia correction factor $\left[\!\left[\mathcal{I}_0\right]\!\right]_M$ that vanishes only in the simple slab geometry. On the right hand side of equation (117) we further find the Lorentz force term, which includes the geodesic transfer term. This term represents an energy transfer to the internal energy density equation (108) since we know the Lorentz force to transfer angular momentum to the parallel angular momentum density. Disregarding the inertia correction factor $\mathcal{I}_0$ this term was also identified earlier to transfer energy to the zonal flow [19, 45, 50].

The second term on the right hand side is a novel term that appears for fluctuating radial velocity $\widehat{u^v}$ and the inertia factor $\widehat{\mathcal{I}_0}$. In order to estimate the importance of the inertia factor we plot $\mathcal{I}_0$ for an exemplary tokamak equilibrium in figure 2. We immediately see that the inertia factor is not a flux function and is much smaller on the low-field side than on the high-field side. Furthermore, it diverges at the X-point and the O-point. At the same time the toroidal component of the $\boldsymbol{E}\times\boldsymbol{B}$ velocity $u_{E,\varphi}$ is zero at these points since the magnetic field is purely toroidal (and thus the zonal flow energy remains finite). Further, the divergence at the X-point is an integrable singularity as shown in figure 3, where we plot the flux-surface average $\langle\mathcal{I}_0\rangle$. Here, we mainly see that there appear gradients close to the separatrix and in the core of the domain.

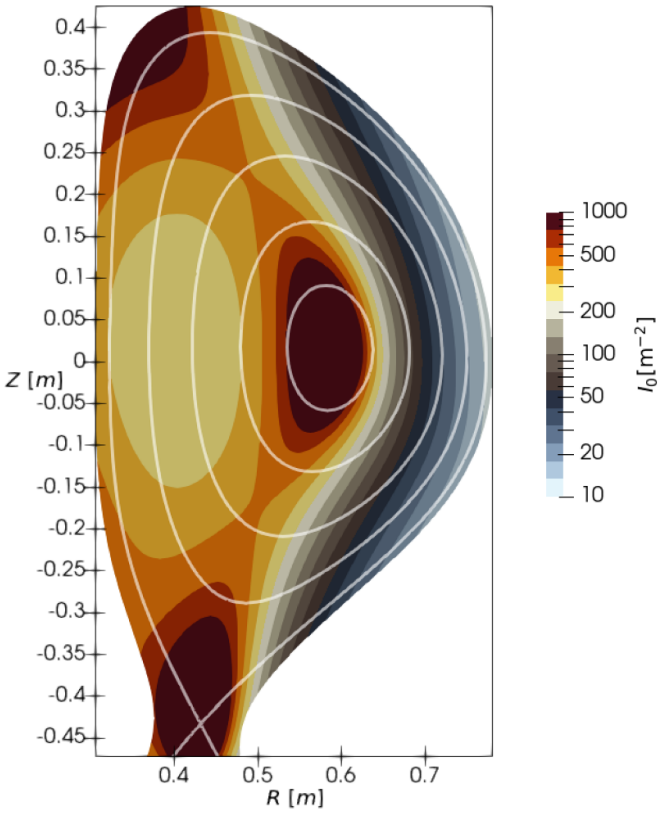

Figure 2. The $\mathcal{I}_0$ factor on an example tokamak equilibrium. Note the logarithmic colour scale, which is cut at the top at 1000 due to the divergence at the X-point and the O-point. The contour lines are given at $\rho_t = 0.2, 0.4, 0.6, 0.8, 1.0$ with the toroidal flux label $\rho_t := \sqrt{\psi_t/\psi_{t,\text{sep}}}$.

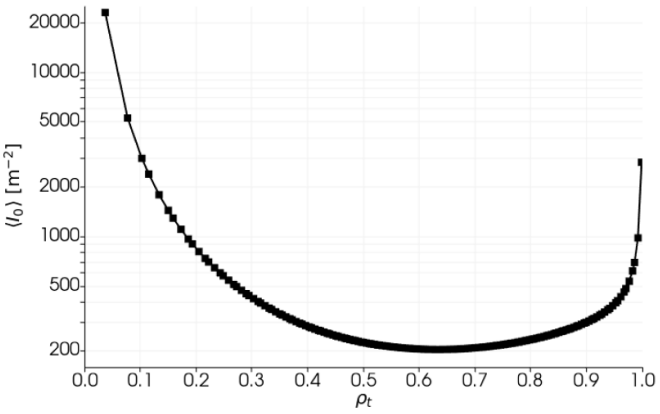

Figure 3. The $\langle \mathcal{I}_0 \rangle$ factor on an example tokamak equilibrium as a function of the toroidal flux label $\rho_t := \sqrt{\psi_t/\psi_{t,\text{sep}}}$. Note the logarithmic scale of the y-axis.

Finally, on the right hand side of equation (116) we find the source term $\mathcal{S}_{\text{zonal}}$. This term contains a contribution from the density source S_n proportional to the inertia factor and the square toroidal $\boldsymbol{E}\times\boldsymbol{B}$ velocity. The sign of this contribution depends only on the sign of S_n. Comparing to figure 2 we see that the inertia factor is almost 2 orders of magnitude higher on the high field side than on the low field side. A particle source on the tokamak high field side is a far more effective source for zonal flow energy than on the low field side. This supports experimental evidence that H-mode access is favored by fueling plasma on the inboard side of a tokamak (for example in MAST [55]). A second contributor is the angular momentum source defined in equation (84), which we already discussed to be pronounced for poloidally asymmetric particle sources.

7. Discussion

7.1. Simplified magnetic field geometries

It is common in the existing literature to reduce the full three-dimensional magnetic field geometry to simplify expressions. The general magnetic field in equation (3) with both toroidal and poloidal components reduces to a purely toroidal magnetic field for $\psi_p = 0$ and the purely poloidal field for $\psi_t = 0$. All our results so far hold for the general magnetic field without axisymmetry. We thus first discuss the poloidal and toroidal fields without assuming axisymmetry. A glance at the gyrokinetic 1-form equation (11) convinces us that in each of these cases both the poloidal and toroidal angular momentum have a single component. In a poloidal field the poloidal angular momentum contains only the parallel velocity $u_\parallel b_\vartheta$ while the toroidal angular momentum consists only of the $\boldsymbol{E}\times\boldsymbol{B}$ flow $u_{E,\varphi}$ and vice versa for the purely toroidal magnetic field geometry.

For the poloidal field the resulting evolution equations are actually already available and we do not need to compute anything further. The relevant equations are equation (79) and the ϑ component of (72). For the purely toroidal magnetic field the parallel momentum balance is given by the φ component of (72), however the $\boldsymbol{E}\times\boldsymbol{B}$ momentum is problematic since ι is zero and thus equation (80) does not hold. Furthermore, since ψ_p vanishes the flux-surface average needs to be redefined with the help of ψ_t.

In the following we will discuss the axisymmetric case for the general, the purely toroidal and the purely poloidal magnetic fields, which allows further simplifications.

7.1.1. General axisymmetric magnetic field. An axisymmetric magnetic field can be written as in equation (6) and is a general feature of the tokamak configuration. It is well known that in this case the toroidal angular momentum density is a conserved quantity [1, 2]. In our derivation axisymmetry leads to the full toroidal angular momentum conservation (up to external sources) in the φ component of equation (102). The φ derivatives in the first two terms on the right hand side vanish and the magnetic shear does not contribute. Comparing this result obtained in the drift ordering to the exact result obtained using Noether's theorem [1] we find a difference of half of the diamagnetic drift. The factor one half is difficult to interpret physically. In our derivation we used the pressure equation to evaluate this term and obtain the full diamagnetic drift. At the same time there is a freedom in how this term is treated in that we could equally cast the diamagnetic drift completely under

the time derivative instead of the right hand side. We comment more on this feature in section 7.2.

7.1.2. Purely toroidal, axisymmetric magnetic field. In the axisymmetric case we discuss here we can write (with cylindrical coordinates R, Z and toroidal angle φ).

$$\psi_t = \int^R B_0(R')\mathrm{d}R', \quad \psi_p = 0 \tag{119}$$

$$\boldsymbol{B}(R) = B_0 \frac{R_0}{R}\hat{\boldsymbol{e}}_\varphi \tag{120}$$

$$\langle h\rangle_Z := \frac{1}{2\pi L_Z}\iint h(R,Z,\varphi)\mathrm{d}Z\mathrm{d}\varphi \tag{121}$$

The gyro-kinetic 1-form equation (11) becomes $\gamma = (q\psi_t \mathrm{d}Z + mw_\parallel b_\varphi \mathrm{d}\varphi + m\mu\mathrm{d}\theta/q$ and now has symmetry in both the R and Z-direction, which makes both $q\psi_t(R)$ and $mw_\parallel R$ conserved quantities separately. This is in fact an important point to emphasize. The purely toroidal magnetic field has two symmetries and thus two exactly conserved quantities instead of just one in the general axisymmetric geometry. In the derivation of the poloidal $\boldsymbol{E}\times\boldsymbol{B}$ momentum in section 4, all we have to do is replace $\boldsymbol{\nabla}\psi_p$ with $\boldsymbol{\nabla}\psi_t(R) = B(R)\hat{\boldsymbol{e}}_R$, which defines $\boldsymbol{\eta} := \hat{\boldsymbol{e}}_\varphi \times \boldsymbol{\nabla}\psi_t/B(R) \equiv \hat{\boldsymbol{e}}_Z$. Equation (80) thus reads (with zero magnetic shear)

$$\begin{aligned}&\sum_s \left\{ m\langle n\rangle_Z \left(\frac{\partial}{\partial t} + \mathcal{U}^R\frac{\partial}{\partial R}\right)[\![u_{E,Z}]\!]_Z\right\} + \frac{\partial}{\partial R}\mathcal{M}^R_Z \\ &= \sum_s \left\{ -m\langle n\rangle_Z \left(\frac{\partial}{\partial R}\mathcal{F}^R_{\perp,Z} + \mathcal{F}^R_{\perp,Z}\frac{\partial}{\partial R}\ln\langle n\rangle\right) + m\mathcal{S}_{u_{E,Z}}\right\}\end{aligned} \tag{122}$$

In the limit $B_0(R) = B_0$ and without $A_{1,\parallel}$ and finite Larmor radius effects this equation agrees with [16]. The parallel angular momentum balance equation (88) now reduces to

$$\begin{aligned}&\sum_s \left\{ m\langle n\rangle_Z \left(\frac{\partial}{\partial t} + \mathcal{U}^R\frac{\partial}{\partial R}\right)[\![u_\parallel R]\!]_Z + \frac{\partial}{\partial R}\mathcal{K}^R_{\parallel,Z}\right\} \\ &= \sum_s \left\{ -m\langle n\rangle_Z \left(\frac{\partial}{\partial R}\mathcal{F}^R_{\parallel,Z} + \mathcal{F}^R_{\parallel,Z}\frac{\partial}{\partial R}\ln\langle n\rangle_Z\right) + m\mathcal{S}_{u_\parallel R}\right\}\end{aligned} \tag{123}$$

Due to the symmetry in R and Z neither the Lorentz force, nor the mirror force appears in equations (122) and (123). Further note that the continuity equation $\partial_t\langle n\rangle_Z + \partial_R(\mathcal{U}^R\langle n\rangle_Z) = \langle S_n\rangle_Z$ can be used to cast these equations into conservative form.

7.1.3. Purely poloidal, axisymmetric magnetic field. The poloidal field approximation with $\psi_t = 0$ is potentially interesting for the field-reversed configuration [56], provided that our orderings in section 3 and 4 hold. We will here investigate the axisymmetric case since, as discussed before, the non-axisymmetric case is already covered. The Poincaré 1-form equation (11) reads $\gamma = q\psi_p\mathrm{d}\varphi + mw_\parallel\sqrt{g_{\vartheta\vartheta}}^{-1}\mathrm{d}\vartheta + \frac{m}{q}\mu\mathrm{d}\theta$. This results in $B^*_\parallel = \boldsymbol{B}^*\cdot\hat{\boldsymbol{e}}_\vartheta = \boldsymbol{B}\cdot\hat{\boldsymbol{e}}_\vartheta \equiv B_p$ with $\hat{\boldsymbol{e}}_\vartheta := \boldsymbol{e}_\vartheta/|\boldsymbol{e}_\vartheta|$. The approximation clearly breaks at the X-point where $B_p = 0$, however this point might be redundant since flux coordinates themselves do not exist on the last closed flux-surface where ι^{-1} diverges as we discussed in section 2.

It is interesting to note that toroidal symmetry now leads to the exact conservation of $\gamma_\varphi = qA_\varphi = q\psi_p$ since $\hat{\boldsymbol{e}}_\vartheta$ has no component in $\mathrm{d}\varphi$ in a symmetric situation. The toroidal angular momentum conservation in the poloidal field approximation thus contains only the toroidal component of the $\boldsymbol{E}\times\boldsymbol{B}$ motion. In this case we can write (note that equation (4) still holds) $\boldsymbol{\eta} := \hat{\boldsymbol{e}}_\vartheta \times \boldsymbol{\nabla}\psi_p/B_p = \boldsymbol{e}_{\boldsymbol{\varphi}}$ which is possible with equation (4), $\boldsymbol{B} = B_p\hat{\boldsymbol{e}}_\vartheta$ and $\boldsymbol{e}_{\boldsymbol{\vartheta}}\cdot\boldsymbol{e}_{\boldsymbol{\varphi}} = g_{\vartheta\varphi} = 0$. The vector $\boldsymbol{\eta}$ thus points in the actual toroidal direction and does not have a poloidal component. We further have $u^v_E = -\frac{\mathrm{d}v}{\mathrm{d}\psi_p}\frac{\partial\phi}{\partial\varphi}/B_p$ In comparison, we have that $u_{E,\vartheta} = 0$, that is in the poloidal field approximation $\boldsymbol{u}_E = \hat{\boldsymbol{e}}_\vartheta \times \nabla\phi/B_p$ has no poloidal component. The non-zero part of the momentum fluxes is thus

$$\begin{aligned}&\sum_s \left\{ m\langle n\rangle \left(\frac{\partial}{\partial t} + \mathcal{U}^v\frac{\partial}{\partial v}\right)[\![u_{E,\varphi}]\!]\right\} + \frac{\partial}{\partial v}\mathcal{M}^v_\varphi \\ &= \sum_s \left\{ -m\langle n\rangle \left(\frac{\partial}{\partial v}\mathcal{F}^v_{\perp,\varphi} + \mathcal{F}^v_{\perp,\varphi}\frac{\partial}{\partial v}\ln\langle n\rangle\right) + m\mathcal{S}_{u_{E,\varphi}}\right\}\end{aligned} \tag{124}$$

where we used that the φ component of $\boldsymbol{j}_f\times\boldsymbol{B}$ vanishes with $((\hat{\boldsymbol{e}}_\vartheta\times\boldsymbol{\nabla}\ln B_p)\times\hat{\boldsymbol{e}}_\vartheta)_\varphi = ((\hat{\boldsymbol{e}}_\vartheta\times\boldsymbol{\kappa})\times\hat{\boldsymbol{e}}_\vartheta)_\varphi = 0$ due to this means that in the poloidal field approximation there is no transfer term between $\boldsymbol{E}\times\boldsymbol{B}$ motion and parallel momentum just as in the purely toroidal magnetic field in equation (122).

In contrast the equation for the parallel momentum in toroidally symmetric cases becomes (with $b_\varphi = 0$ and $b_\vartheta = \sqrt{g_{\vartheta\vartheta}}$)

$$\begin{aligned}&\sum_s \left\{ m\langle n\rangle \left(\frac{\partial}{\partial t} + \mathcal{U}^v\frac{\partial}{\partial v}\right)[\![u_\parallel b_\vartheta]\!] + \frac{\partial}{\partial v}\mathcal{K}^v_{\parallel,\vartheta}\right\} \\ &= \sum_s \left\{ -m\langle n\rangle \left(\frac{\partial}{\partial v}\mathcal{F}^v_{\parallel,\vartheta} + \mathcal{F}^v_{\parallel,\vartheta}\frac{\partial}{\partial v}\ln\langle n\rangle\right) + m\mathcal{S}_{u_\parallel b_\vartheta}\right. \\ &\qquad \left. - \left\langle p_\perp\frac{\partial\ln B}{\partial\vartheta} + (p_\parallel + mnu^2_\parallel)b^\vartheta\frac{\partial b_\vartheta}{\partial\vartheta}\right\rangle\right\}\end{aligned} \tag{125}$$

In contrast to the purely toroidal magnetic field here we find the mirror force and the geometric correction in the poloidal direction on the right hand side. This means that the

background magnetic field acts as a source/sink of parallel momentum.

Again, we note that the continuity equation $\partial_t \langle n\rangle + \partial_v(\mathcal{U}^v \langle n\rangle) = \langle S_n\rangle$ can be used to cast equations (124) and (125) into conservative form.

7.2. The momentum of electromagnetic fields in matter

We now note that we have the possibility to rewrite Equation (61) using identity equation (60) to cast the diamagnetic drift under the time derivative (using $u_E^v = -\boldsymbol{\nabla}_\perp\phi\cdot \boldsymbol{e}_\varphi \mathrm{d}v/\mathrm{d}\psi_p$)

$$\begin{aligned}&\frac{\partial}{\partial t}\left\langle (\boldsymbol{P}_{\mathrm{em}}\times\boldsymbol{B})_\varphi\right\rangle\\ &\quad-\frac{\partial}{\partial v}\left\langle E_\varphi P^v_{\mathrm{em}} + \left(\frac{1}{\mu_0}B_{1,\perp,\varphi} - M^{\mathrm{em}}_{\perp,\varphi}\right)B^v_{1,\perp}\right\rangle\\ &\quad= -\left\langle (\boldsymbol{j}_f\times\boldsymbol{B})_\varphi\right\rangle + S_{\mathrm{em},\varphi}\end{aligned} \tag{126}$$

with

$$\boldsymbol{E} := -\boldsymbol{\nabla}_\perp\phi \tag{127}$$

$$\boldsymbol{P}_{\mathrm{em}} := -\sum_{\mathrm{s}} mn\left(\frac{\boldsymbol{\nabla}_\perp\phi}{B^2} + \frac{\boldsymbol{\nabla}_\perp p_\perp}{qnB^2}\right) \tag{128}$$

$$\boldsymbol{M}^{\mathrm{em}}_\perp := \sum_{\mathrm{s}} m\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}(q_\parallel + u_\parallel p_\perp)}{qB^2} \tag{129}$$

$$\boldsymbol{S}_{\mathrm{em}} := \sum_{\mathrm{s}} mS_n\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\phi}{B} + \frac{m\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}S_{p_\perp}}{qB} \tag{130}$$

Equation (126) is the evolution equation for the electromagnetic momentum flux $\boldsymbol{g} := \boldsymbol{D}\times\boldsymbol{B}$. The electric part in the displacement field $\boldsymbol{D} := \epsilon_0\boldsymbol{E} + \boldsymbol{P}_{\mathrm{em}}$ vanishes because we neglected the corresponding field part of the action (23) and have quasineutrality. The momentum tensor has the form $T^v_\varphi := -E_\varphi D^v - H_\varphi B^v_{1,\perp}$ with the magnetizing field $\boldsymbol{H} := \boldsymbol{B}_{1,\perp}/\mu_0 - \boldsymbol{M}^{\mathrm{em}}_\perp$. The momentum flux $\boldsymbol{g}$ and tensor T correspond to the ones given in Reference [57]. With the identification of the Lorentz force density $\boldsymbol{f}_L = \boldsymbol{j}_f\times\boldsymbol{B}$ on the right hand side we can write equation (126) as

$$\frac{\partial}{\partial t}\langle g_\varphi\rangle + \frac{\partial}{\partial v}\langle T^v_\varphi\rangle = -\langle f_{L,\varphi}\rangle + \langle S_{\mathrm{em},\varphi}\rangle \tag{131}$$

Notice the minus in the Lorentz force, which is a signature that g_φ is indeed the momentum flux for the electromagnetic field rather than for the plasma itself. Furthermore, the form of the Lorentz force motivates the identification of $\boldsymbol{j}_f$ as the *free* current as opposed to the bound polarization current. The ϑ component of the momentum flux follows by multiplying equation (131) with ι^{-1}

$$\frac{\partial}{\partial t}\langle g_\vartheta\rangle + \frac{\partial}{\partial v}\langle T^v_\vartheta\rangle = -\langle f_{L,\vartheta}\rangle - \langle T^v_\vartheta\rangle\frac{\partial}{\partial v}\ln\iota + \langle S_{\mathrm{em},\vartheta}\rangle \tag{132}$$

Here, notably a contribution from the magnetic shear appears on the right hand side as a coupling term to the external magnetic field.

In equation (128) we define the electromagnetic polarization charge $\boldsymbol{P}_{\mathrm{em}}$ analogous to the magnetization $\boldsymbol{M}^{\mathrm{em}}_\perp$ (129) (which we repeat here for convenience) and different from the gyro-centre polarization charge $\boldsymbol{P}_{\mathrm{gy}}$ by half the diamagnetic drift. We remark that neither of these quantities is uniquely defined. The form $\boldsymbol{P}_{\mathrm{em}}$ and $\boldsymbol{M}^{\mathrm{em}}_\perp$ highlights the physical origin of polarization and magnetization in gyro-kinetic models. Here, we can view the plasma as a collection of charged discs that can be magnetized and polarized. The disc polarization $\boldsymbol{\pi} := m\hat{\boldsymbol{b}}\times\dot{\boldsymbol{X}}/B$ stems from the drift velocities and reflects that due to the drifts the gyro-orbits are no longer closed [27, 30]. Macroscopically, in our model we have $\boldsymbol{P}_{\mathrm{em}} = mn\hat{\boldsymbol{b}}\times(\boldsymbol{u}_E + \boldsymbol{u}_D)/B$. On the other side, the magnetization $\boldsymbol{M}^{\mathrm{em}}_\perp$ contains the moving electric dipole contribution. An electric dipole $\boldsymbol{\pi}$ that moves with velocity $v_\parallel\hat{\boldsymbol{b}}$ along the magnetic field lines induces a magnetic moment $\boldsymbol{\mu} = \boldsymbol{\pi}\times v_\parallel\hat{\boldsymbol{b}}$. However, we only find the diamagnetic part to the moving dipole contribution. We are missing the contribution $mnu_\parallel\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\phi/B^2$ since we neglected the corresponding non-linear coupling terms in the Hamiltonian (18).

7.3. Comparison to drift-fluid models

We note that equation (126) can also be viewed as a relation for the radial force density $\langle(\boldsymbol{P}_{\mathrm{em}}\times\boldsymbol{B})_\varphi\rangle = \sum_{\mathrm{s}}\left\langle\frac{m}{qB^2}(qn\boldsymbol{\nabla}_\perp\phi + \boldsymbol{\nabla}_\perp p)\cdot\boldsymbol{\nabla}\psi_p\right\rangle$ where the force density $-qn\boldsymbol{E}_\perp + \boldsymbol{\nabla}_\perp p$ appears inside the bracket on the right hand side. If the right hand side of equation (126) is zero, the radial pressure gradient and the radial electric field strength balance each other. Alternatively, we can rewrite equation (126) as

$$\begin{aligned}&\frac{\partial}{\partial t}\sum_{\mathrm{s}} m\langle n(u_{E,\varphi} + u_{D,\varphi})\rangle\\ &\quad+\frac{\partial}{\partial v}\left[\sum_{\mathrm{s}} m\langle n(u_{E,\varphi} + u_{D,\varphi})\rangle[\![u^v_E]\!]\right]\\ &\quad+\frac{\partial}{\partial v}\left[\sum_{\mathrm{s}} m\langle n\rangle\left(\mathcal{F}^v_{E,\varphi} + \mathcal{F}^{\mathrm{T}v}_{D,\varphi}\right) + \mathcal{M}^{\mathrm{T}v}_\varphi\right]\\ &= -\left\langle(\boldsymbol{j}_f\times\boldsymbol{B})_\varphi\right\rangle\\ &\quad+\sum_{\mathrm{s}} m\left\langle S_n u_{E,\varphi} + \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}S_{p_\perp}}{qB}\cdot\boldsymbol{e}_\varphi\right\rangle\end{aligned} \tag{133}$$

where the sum of $\boldsymbol{E}\times\boldsymbol{B}$ and diamagnetic drift appear under the time derivative. We point out that equation (133) compares to equation (79) and is distinguished by the appearance of the transpose of the diamagnetic and Maxwell stresses and the additional appearance of a pressure source on the right hand side in a form analogous to a diamagnetic drift term. This latter term appears through the use of the pressure equation in bringing the diamagnetic drift under the time derivative. The ϑ-component of equation (133) is obtained by multiplying with ι^{-1}. Equation (133) is also the form closest to the

drift-fluid (generalized) vorticity equation [41, 43]. To compare one needs to take the flux-surface average over the generalized vorticity equation and then integrate over the volume. This immediately allows the interpretation of equation (133) as the volume integrated Equation for a divergence free current or a closed current loop. The Favre decomposition needs to be introduced in order to recover our stress terms. Further, the momentum balance that results from integration of the ensemble averaged kinetic Vlasov equation also has a similar form to equation (133) as seen for example in Reference [9]. The difference is that we only recover the lower order $\boldsymbol{E}\times\boldsymbol{B}$ and diamagnetic velocities instead of the full plasma velocity.

We point out that the pressure source (heating) on the right hand side of equation (133) is not present in the drift-fluid generalized vorticity equation with plasma-neutral interactions [58]. Further, our source terms disagree with Reference [41], where a momentum source instead of a density or pressure source is presented. The cause for these differences should be clarified in future work. In the present formulation the momentum source term in equation (133) reflects (i) the presence of a formal kinetic source S on the right hand side of the gyro-kinetic Valsov equation (21) that is (ii) quasi-neutral under species summation equation (54) and is (iii) transformed according to the gyro-centre transformation rules equation (53). On the other hand the Stringer-Winsor spin-up term agrees with our results.

Finally, we emphasize that the evolution equation for the $\boldsymbol{E}\times\boldsymbol{B}$ flow equation (61), the evolution for the electromagnetic field momentum equation (126) and the interpretation as a radial force density or the sum of $\boldsymbol{E}\times\boldsymbol{B}$ and diamagnetic drifts in equation (133) are completely equivalent views of the same result. In particular, physical arguments made with one of the three equations immediately translate into the other two.

7.4. Relation to the ion orbit loss mechanism

The ion orbit loss mechanism [3, 6, 13–15, 59] refers to the idea that ion orbits close to the X-point end on the divertor target or the vessel wall and are thus lost to the confined plasma region. It is thought that the poloidal magnetic field close to the X-point is small such that the grad-B curvature drift velocity dominates over the parallel velocity making ions drift across the separatrix. This then generates a net flux of positive charge out of the confined region. In particle phase space the ions that are on a loss orbit are situated on a 'loss-cone' encompassing ions with small parallel velocity and large perpendicular velocity / magnetic moment. It is reported that the perpendicular kinetic energy of the loss cone reaches down to thermal energies [14].

The ion orbit loss is often invoked in models explaining the L-H transition [13–15], where it is thought that the outward current leaves a small region inside the separatrix negatively charged, which generates a strong radial electric field. This field in turn drives a strong poloidal shear $\boldsymbol{E}\times\boldsymbol{B}$ flow that then forms the transport barrier typical for the high confinement mode. On the other side the same idea is used to explain intrinsic toroidal rotation [3, 6, 59], the observation that the plasma rotates toroidally without controlled external sources like NBI. The main ingredient here is to assume that the rate by which ions enter loss orbits depends on the direction of their parallel velocity. This then generates an asymmetry between losses of so-called co- and counter-current ions. Since ions carry toroidal momentum, the preferential loss in one direction accelerates the plasma in the other.

Since our derivation of poloidal and toroidal angular momentum balance makes no assumption on the form of the distribution function F (in particular it does not assume that F is Maxwellian) and the particle orbits are retained via equations (31) the ion orbit loss mechanism must consequentially be contained in our results. Here, we want to identify the relevant terms for both poloidal and toroidal rotation.

The net surface integrated current $\int_{\psi_p}\boldsymbol{j}\cdot\mathbf{dA} = \langle\boldsymbol{j}\cdot\boldsymbol{\nabla}v\rangle$[8] flowing through a flux-surface ψ_p, in particular the separatrix, by magnetic drifts is given by

$$\begin{aligned}\sum_s \langle qn(\boldsymbol{u}_{\boldsymbol{\nabla}B}+\boldsymbol{u}_\kappa)\cdot\boldsymbol{\nabla}v\rangle &= \left\langle \boldsymbol{j}_f\cdot\boldsymbol{\nabla}v\right\rangle\\ &= \sum_s\left\langle\left(\|\mu B\|\boldsymbol{K}_{\boldsymbol{\nabla}B}+\|mv_\|^2\|\boldsymbol{K}_\kappa\right)\cdot\boldsymbol{\nabla}v\right\rangle\\ &= \frac{\mathrm{d}v}{\mathrm{d}\psi_p}\sum_s\left\langle(p_\perp\boldsymbol{K}_{\boldsymbol{\nabla}B}+(p_\|+mnu_\|^2)\boldsymbol{K}_\kappa)\cdot\boldsymbol{\nabla}\psi_p\right\rangle\end{aligned}\tag{134}$$

where we inserted the definition of curvature and grad-B drifts equations (A19) and (A18) and the velocity space moments to emphasize the origin of $\boldsymbol{j}_f$ as particle drifts. At this point recall again that $\langle\boldsymbol{j}_f\cdot\boldsymbol{\nabla}v\rangle \equiv \langle(\boldsymbol{j}_f\times\boldsymbol{B})_\varphi\rangle\,\mathrm{d}v/\mathrm{d}\psi_p \equiv \langle(\boldsymbol{j}_f\times\boldsymbol{B})_\vartheta\rangle\,\mathrm{d}v/\mathrm{d}\psi_t$ by virtue of equations (4) and (5). The term described in equation (134) is nothing but the Lorentz force term that appears in our momentum equations in section 5 and which we already identified as the Stringer-Winsor spin-up or geodesic transfer term. The ion orbit loss contribution must be contained in the first term on the right side of equation (134) since it was argued that ions with large μ and small $v_\|$ fall on loss orbits. A signature of ion orbit loss would be if the ion term in equation (134) is larger than the electron contribution at or close to the separatrix.

At this point we notice that for favourable curvature drift direction the curvature vectors counter-align with $\boldsymbol{\nabla}\psi_p$ ($\mathcal{K}(\psi_p)<0$, decelerate) on the top and align ($\mathcal{K}(\psi_p)>0$, accelerate) on the bottom of the tokamak. In order for the flux-surface average in equation (134) to yield a non-vanishing result we therefore need an up-down asymmetry of the pressure in the flux-surface. Furthermore, we notice that for our example tokamak equilibrium in figure 1 we have $\langle\boldsymbol{K}_\kappa\cdot\boldsymbol{\nabla}\psi_p\rangle = 0$. Indeed, more generally we find $\boldsymbol{\nabla}\cdot\boldsymbol{K}_{\boldsymbol{\nabla}B} = -\boldsymbol{\nabla}\cdot\boldsymbol{K}_\kappa = -\boldsymbol{K}_\kappa\cdot\boldsymbol{\nabla}\ln B \sim \mathcal{O}(\delta^6)$, which results in $\partial_v\langle K_\kappa\rangle = \partial_v(\langle\boldsymbol{K}_\kappa\cdot\boldsymbol{\nabla}\psi_p\rangle\,\mathrm{d}v/\mathrm{d}\psi_p)\sim\mathcal{O}(\delta^6)$ and thus $\langle\boldsymbol{K}_\kappa\cdot\boldsymbol{\nabla}\psi_p\rangle\approx 0$.

[8] Recall the definition of the flux-surface average equation (7) to see that this is indeed the area integral.

This means that only the fluctuations in $p_\perp$, $p_\parallel$ and $nu_\parallel^2$ contribute and we can write[9]

$$\left\langle \boldsymbol{j}_f \cdot \boldsymbol{\nabla}\psi_p \right\rangle = \sum_s \left\langle (\widetilde{p_\perp}\boldsymbol{K}_{\boldsymbol{\nabla}B} + (\widetilde{p_\parallel} + \widetilde{mnu_\parallel^2})\boldsymbol{K}_\kappa) \cdot \boldsymbol{\nabla}\psi_p \right\rangle \tag{135}$$

7.4.1. Poloidal $\boldsymbol{E}\times\boldsymbol{B}$ flow. Even though, as argued in equation (102) in section 5, the Lorentz force does not generate net poloidal momentum, it does generate $\boldsymbol{E}\times\boldsymbol{B}$ momentum, respectively a radial electric field $u_{E,\vartheta} \sim \boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\psi_t$. We thus conclude that the ion-orbit loss mechanism may indeed contribute to the radial electric field through the Lorentz force.

On the other hand, we emphasize that the Lorentz force is not the only candidate that contributes to the poloidal $\boldsymbol{E}\times\boldsymbol{B}$ flow generation. Any other term in equation (80) could be equally important. Besides the $\boldsymbol{E}\times\boldsymbol{B}$ Favre stress we identified for example the diamagnetic stress $\mathcal{F}^v_{D,\eta}$ or the density gradient and magnetic shear related terms as additional candidates that may be equally relevant for the L-H transition.

7.4.2. Intrinsic toroidal rotation. The ion loss mechanism is through the Lorentz force indeed contained in the toroidal angular momentum conservation for $\boldsymbol{E}\times\boldsymbol{B}$ (79) and parallel (88) angular momentum. However, as we discussed in equation (102) in section 5 the Lorentz force does not actually generate net angular momentum, neither poloidal nor toroidal. A loss of ions through the separatrix does thus not generate toroidal angular momentum. As is shown in equation (102) for an axisymmetric equilibrium the only sources for toroidal angular momentum are the actual source terms $\mathcal{S}_{nu_\parallel}$ and $\mathcal{S}_n$ on the right hand side. In order to explain an intrinsic rotation profile in this case we thus need to focus on the radial advection and stress terms, which describe the radial in-/outflow of momentum through the boundary flux-surface. This requires a description of the turbulent fluctuations entering the stress terms, which is difficult to acquire short of a full-scale simulation of the model equations. The literature thus often invokes phenomenological models, for example the asymmetric turbulent diffusion [3] where a preferential loss of co- or counter-current ions through the separatrix generates a net momentum gain for the remaining plasma inside the separatrix.

7.5. Comparison to parallel acceleration

The argument was made [60–63] that in experimental measurements the parallel velocity $u_\parallel$ is measured and not the parallel momentum density $nu_\parallel$. It was concluded that therefore $u_\parallel$ respectively $\langle u_\parallel \rangle$ should be the quantity that theoretical work should focus on when discussing intrinsic rotation. In our view, neither premise nor conclusion of this hypothesis holds. First, the velocity *can* be measured at the same position and time as the density with for example velocity space tomography [64] (and it should be noted that it is the velocity with respect to the line of sight rather than the parallel velocity that is actually measured in charge exchange diagnostics). Second, $u_\parallel$ is *not* the angular momentum; $u_\parallel b_\varphi \approx u_\parallel R$ is and only part of it at that. Also, recall that even though it is not technically wrong to compute $\langle u_\parallel \rangle$ the flux-surface average (7) is a volume average and should be taken over density like quantities (like $nu_\parallel$). Finally, what comes out of a gyro-kinetic moment expansion (as performed in [60–63]) is the *gyro-fluid* parallel velocity $U_\parallel$, not the actually measured fluid velocity $u_\parallel$. As we discuss in section 3.3 care must be taken when comparing gyro-fluid quantities like $U_\parallel$ to the actually physically measured fluid quantity $u_\parallel$ due to the involved coordinate transformation of equation (49), which for $U_\parallel$ is given in equation (52). The time evolution equation for $u_\parallel$ reads in our ordering (keeping terms up to $\mathcal{O}(\delta^3)$)

$$\begin{aligned}\frac{\partial u_\parallel}{\partial t} &+ (\hat{\boldsymbol{b}} + \widetilde{\boldsymbol{b}_\perp})\cdot\boldsymbol{\nabla}u_\parallel^2/2 + \boldsymbol{u}_E\cdot\boldsymbol{\nabla}u_\parallel \\ &+ \frac{1}{mn}\boldsymbol{\nabla}\cdot\left((\hat{\boldsymbol{b}} + \widetilde{\boldsymbol{b}_\perp})p_\parallel\right) + \frac{1}{m}t_\perp\nabla_\parallel \ln B \\ &+ \frac{q}{m}\partial_t A_{1,\parallel} + \frac{q}{m}(\hat{\boldsymbol{b}} + \widetilde{\boldsymbol{b}_\perp})\cdot\boldsymbol{\nabla}\phi = S_{u_\parallel}\end{aligned} \tag{136}$$

The terms that appear beside the time derivative are in order the parallel advection term, the $\boldsymbol{E}\times\boldsymbol{B}$ advection term, the parallel pressure gradient term, the mirror force term and the last two terms form the parallel electric field. In equation (136) we see the local parallel acceleration of a single (ion) species. However, working with accelerations instead of force densities as in equation (93) does not reveal that after species summation and flux-surface averaging all net internal forces vanish and only external forces remain. As collectively generated, internal forces neither the pressure gradient nor the electric field can be the source of an intrinsic rotation profile. We point out here that the only external force that is able to make a contribution, the mirror force term $\langle p_\perp \nabla_\parallel \ln B\rangle$, was neglected in [60–63].

8. Conclusions

Our main results are the Favre averaged covariant poloidal and toroidal velocity evolution equations (79), (80), and (88) applicable in arbitrary magnetic field geometry including tokamaks, the reversed field pinch, the field-reversed configuration and stellarators. The $\boldsymbol{E}\times\boldsymbol{B}$ equations (79), (80) and the parallel components in equation (88) sum to the total angular momentum in equation (102).

The usefulness of the Favre-average formalism mainly stems from the identification of the Favre stress as the mediator between turbulent fluctuations and flux-surface averaged profiles. In our full-F gyro-kinetic formulation the perpendicular Favre stress $\mathcal{F}^v_{\perp,\eta}$ appears in the $\boldsymbol{E}\times\boldsymbol{B}$ part of the angular momentum as a natural extension of the Reynolds stress through the density weighted flux-surface average - the Favre average [16]. The perpendicular Favre stress consists

[9] If we assume $p_\parallel = p_\perp = p$, we can further simplify $\langle \boldsymbol{j}_f\cdot\boldsymbol{\nabla}\psi_p\rangle = \sum_s \left\langle \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\tilde{p}}{B}\cdot\boldsymbol{\nabla}\psi_p\right\rangle + \left\langle \widetilde{mnu_\parallel^2}\boldsymbol{K}_\kappa\cdot\boldsymbol{\nabla}\psi_p\right\rangle$ where we use that $\boldsymbol{K}_{\boldsymbol{\nabla}B} + \boldsymbol{K}_\kappa = \boldsymbol{K}$ and $\boldsymbol{\nabla}\cdot\boldsymbol{K} = 0$ (see A1). Then we find the radial component of the diamagnetic drift $\langle u_D^v\rangle\, \mathrm{d}\psi_p/\mathrm{d}v$ in the first term on the right hand side.

of the previously found $\boldsymbol{E}\times\boldsymbol{B}$ contribution $\mathcal{F}^v_{E,\eta}$, but also of the novel diamagnetic $\mathcal{F}^v_{D,\eta}$ and magnetic flutter $\mathcal{F}^v_{F,\eta}$ contributions defined in equation (82). Besides the Favre stress, the vacuum Maxwell stress $\mathcal{M}^v_{B,\eta}$ and magnetization stresses $\mathcal{M}^v_{M,\eta}$ defined in equation (83) appear. We highlight the relation to the general density gradient drive term in equation (87). Furthermore, the Lorentz force originating from the curvature and grad-B drift induced currents represents a source for $\boldsymbol{E}\times\boldsymbol{B}$ angular momentum density. Finally, poloidally asymmetric density sources equation (84) contribute to angular momentum generation.

Analogous to the $\boldsymbol{E}\times\boldsymbol{B}$ part, the parallel component of the angular momentum density equation (88) is generated by the parallel Favre stress $\mathcal{F}^v_{\parallel,\eta}$ in equation (91) as well as the kinetic stress $\mathcal{K}^v_{\parallel,\eta}$ equation (90) stemming from magnetic fluctuations. The Lorentz force appears with an opposite sign as in the $\boldsymbol{E}\times\boldsymbol{B}$ equation thus vanishing in the summed total angular momentum density in equation (102), both toroidally as well as poloidally. In addition, in equation (88) the mirror force appears as a source of parallel momentum density.

We construct the inertia tensor from the first fundamental form in equation (106). The relevant discussion is based on the mean flow generated by the covariant, Favre averaged velocity components that we investigate in the first part of the paper. From there we construct the rotational energy in equation (111). The $\boldsymbol{E}\times\boldsymbol{B}$ part of this energy can be split into a mean 'zonal' and fluctuating part and we present the evolution of the mean in equation (117) using the previously derived evolution equations for angular momentum. The main finding compared to a simplified geometry is the appearance of a correction factor due to the inertia tensor, which in particular modifies the effect of the density source on the right hand side. A density source on the high field side is a more effective source of zonal flow energy than on the low field side.

We show that we recover previous results obtained in simplified geometries. Interestingly, the purely toroidal magnetic field leads to the exact conservation of both the poloidal $\boldsymbol{E}\times\boldsymbol{B}$ velocity as well as the parallel angular momentum density. This is because an additional symmetry is introduced by this geometry. We also point out that the ion orbit loss mechanism as outlined in the literature is identical to the 'geodesic transfer term' and the 'Stringer-Winsor spin-up mechanism' and is contained in our results in the Lorentz force term on the right hand side of the poloidal $\boldsymbol{E}\times\boldsymbol{B}$ angular momentum equation (80). Finally, we clarify several misconceptions in connection with 'parallel acceleration' relating previous findings to our results.

The main drawback of our derivation is the long-wavelength limit in the gyro-kinetic action equation (23), which effectively reduces our model to a drift-kinetic model and misses higher order finite Larmor radius and polarization effects that could play a role for the L-H transition. We mainly perform this limit in order to avoid an infinite sum in the relation between the ordinary and the variational derivative equation (57) and to avoid the introduction of a fluid closure of the infinite expansions in the polarization and gyro-averages [31]. The drift ordering in section 4 avoids geometrical correction factors stemming from for example perpendicular derivatives on the magnetic field unit vector in section 4 and allows to recover fluid (as opposed to gyro-fluid) moments in our equations and to compare to existing drift-fluid models via equation (133). However, our momentum balance equations are only valid up to order three and the energy balance equations up to order four within this ordering. Future work could address the above issues.

Our results can be used to verify simulation results. The application of these results within full-F gyro-fluid models is subject of ongoing research. However, as previously stated, the available equations in this work are by no means restricted to gyro-fluid models since the derivation contains no assumption on the form of the distribution function. Thus the presented results are relevant also for other frameworks beyond gyro-fluid models like for example gyro-kinetic or drift-fluid models.

The experimental validation of our results may be challenging due to the number of terms that appear in the evolution equations (79), (80), and (88) that in particular require the measurement of plasma potential, parallel velocity, density, pressure and possibly magnetic field fluctuations at the same time and positions. Further, a problematic operation is the flux-surface average. The argument that a time average over the measurement interval equates the flux-surface average only holds if the measured quantity is a flux-function in the first place. On the other hand we provide the theoretical foundation for a discussion of the dominant physical mechanisms that generate poloidal and toroidal angular momentum density and rotational energy in any toroidal magnetic field configuration.

Acknowledgments

We acknowledge fruitful discussions with N. Tronko, P. Strand, V. Naulin, and J.J. Rasmussen. The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 713683 (COFUNDfellowsDTU). This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under Grant Agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Appendix A. Formulary

A.1. Flux surface and Favre average

The flux-surface average (see for example [22]) is an average over a small volume - a differential shell centered around the flux-surface. We define

Table A1. Definitions of geometric operators with b^i the contra-variant components of $\hat{\boldsymbol{b}}$ and g^{ij} the contra-variant elements of the metric tensor. We assume $(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})_\parallel = 0$.

Name	Symbol	Definition
Projection tensor	h	$h^{ij} := g^{ij} - b^i b^j$ Note$h^2 = h$
Perpendicular gradient	$\boldsymbol{\nabla}_\perp$	$\boldsymbol{\nabla}_\perp f := \hat{\boldsymbol{b}}\times(\boldsymbol{\nabla}f\times\hat{\boldsymbol{b}}) \equiv h\cdot\boldsymbol{\nabla}f$
Perpendicular divergence	$\boldsymbol{\nabla}_\perp^\dagger$	$\boldsymbol{\nabla}_\perp^\dagger\cdot\boldsymbol{v} := -\boldsymbol{\nabla}\cdot(h\cdot\boldsymbol{v}) = -\boldsymbol{\nabla}\cdot\boldsymbol{v}_\perp$
Perpendicular laplacian	$\Delta_\perp$	$\Delta_\perp f := \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}_\perp f) = \boldsymbol{\nabla}\cdot(h\cdot\boldsymbol{\nabla}f) \equiv -\boldsymbol{\nabla}_\perp^\dagger\cdot\boldsymbol{\nabla}_\perp$
Curl-b curvature operator	$\mathcal{K}_\kappa$	$\mathcal{K}_\kappa(f) := \boldsymbol{K}_{\boldsymbol{\kappa}}\cdot\boldsymbol{\nabla}f = \frac{1}{B}(\hat{\boldsymbol{b}}\times\boldsymbol{\kappa})\cdot\boldsymbol{\nabla}f$ with $\boldsymbol{\kappa} := \hat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\hat{\boldsymbol{b}}$
Grad-B curvature operator	$\mathcal{K}_{\boldsymbol{\nabla}B}$	$\mathcal{K}_{\boldsymbol{\nabla}B}(f) := \boldsymbol{K}_{\boldsymbol{\nabla}\boldsymbol{B}}\cdot\boldsymbol{\nabla}f = \frac{1}{B}(\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B)\cdot\boldsymbol{\nabla}f$
Curvature operator	$\mathcal{K}$	$\mathcal{K}(f) := \boldsymbol{K}\cdot\boldsymbol{\nabla}f = \boldsymbol{\nabla}\cdot\left(\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}f}{B}\right) = \boldsymbol{\nabla}\times\frac{\hat{\boldsymbol{b}}}{B}\cdot\boldsymbol{\nabla}f,$
Parallel derivative	$\nabla_\parallel$	$\nabla_\parallel f := \boldsymbol{B}\cdot\boldsymbol{\nabla}f/B$ Notice $\boldsymbol{\nabla}\cdot\hat{\boldsymbol{b}} = -\nabla_\parallel\ln B$

$$\langle f\rangle(\psi_p) := \frac{\partial}{\partial v}\int_\Omega \mathrm{dV} f = \int_{\psi_p}\frac{f(\boldsymbol{x})}{|\boldsymbol{\nabla}v|}\mathrm{dA} \tag{A1}$$

where we define $v(\psi_p) := \int_0^{\psi_p}\mathrm{dV}$ as the volume flux label and for the second identity, recall the co-area formula

$$\int_\Omega f(\boldsymbol{x})\mathrm{dV} = \int_0^\rho \mathrm{d}\rho'\left(\int_{\rho'=\mathrm{const}}\frac{f(\boldsymbol{x})}{|\boldsymbol{\nabla}\rho|}\mathrm{dA}\right) \tag{A2}$$

where $\rho(\psi_p)$ is any flux label and Ω is the volume enclosed by the contour $\rho = \mathrm{const}$. In flux coordinates we have $\mathrm{dA} = \sqrt{g}|\boldsymbol{\nabla}\rho|\mathrm{d}\vartheta\mathrm{d}\varphi$. The co-area formula can be viewed as a change of variables in the volume integral. The average fulfills the identities (with scalars λ and μ)

$$\langle\mu f + \lambda g\rangle = \mu\langle f\rangle + \lambda\langle g\rangle \tag{A3}$$

$$\langle f(\psi_p)g(\boldsymbol{x})\rangle = f(\psi_p)\langle g(\boldsymbol{x})\rangle \tag{A4}$$

$$\begin{aligned}\langle\boldsymbol{\nabla}\cdot\boldsymbol{j}\rangle &= \frac{\partial}{\partial v}\langle\boldsymbol{j}\cdot\boldsymbol{\nabla}v\rangle \\ &= \left(\frac{\mathrm{d}v}{\mathrm{d}\rho}\right)^{-1}\frac{\partial}{\partial\rho}\left(\frac{\mathrm{d}v}{\mathrm{d}\rho}\langle\boldsymbol{j}\cdot\boldsymbol{\nabla}\rho\rangle\right)\end{aligned} \tag{A5}$$

Also note that for any divergence free vector field $\boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$ and a flux function $f(\psi_p)$ we have

$$\langle\boldsymbol{\nabla}\cdot(\boldsymbol{j}f)\rangle = 0 \tag{A6}$$

which is proven straightforwardly.

We note the Reynolds decomposition for any function $h(\boldsymbol{x})$

$$h \equiv \langle h\rangle + \widetilde{h} \tag{A7}$$

and its generalization, the Favre average and decomposition

$$[\![h]\!] := \frac{\langle nh\rangle}{\langle n\rangle} \tag{A8}$$

$$h \equiv [\![h]\!] + \widehat{h} \tag{A9}$$

Table A2. List of the first few gyro-fluid moments: gyro-fluid density N, parallel canonical velocity $W_\parallel$ the perpendicular and parallel pressure ($P_\perp$ and $P_\parallel$)/ temperature $T_\perp$ and $T_\parallel$ as well the parallel heat flux $Q_\parallel$.

N	$\|\|1\|\|$	$P_\perp \equiv NT_\perp$	$\|\|\mu B\|\|$
$NW_\parallel$	$\|\|w_\parallel\|\|$	$P_\parallel \equiv NT_\parallel$	$\|\|m(w_\parallel - W_\parallel)^2\|\|$
$N\phi$	$\|\|\phi\|\|$	$Q_\parallel$	$\|\|\mu B(w_\parallel - W_\parallel)\|\|$
$NA_{1,\parallel}$	$\|\|A_{1,\parallel}\|\|$		

where n is the particle density, which makes the Favre average species dependent. The Favre average fulfills

$$[\![gh]\!] = [\![g]\!][\![h]\!] + \left[\!\left[\hat{g}\hat{h}\right]\!\right] \tag{A10}$$

It is sometimes useful to define the Favre average using the total mass density $\rho_M := \sum_\mathrm{s} mn$ as

$$[\![h]\!]_M := \langle\rho_M\rangle^{-1}\sum_\mathrm{s} m\langle nh\rangle \tag{A11}$$

for any (possibly species dependent) function h. If h is species independent this definition simplifies to $[\![h]\!]_M = \langle\rho_M h\rangle / \langle\rho_M\rangle$.

A.2. Fluid moments

The velocity space moments equations (39) and (40) read

$$\|\zeta\| := \int \mathrm{d}w_\parallel\mathrm{d}\mu\mathrm{m}^2 BF\zeta$$

$$\|\zeta\|_S := \int \mathrm{d}w_\parallel\mathrm{d}\mu\mathrm{d}\theta m^2 BS\zeta$$

In table (A2) we name the first few velocity space moments of the gyro-kinetic distribution function F. The moments over the gyro-kinetic source function S are named analogous as S_N, $S_{NW_\parallel}$, $S_{P_\perp}$, $S_{P_\parallel}$ and $S_{Q_\parallel}$. We can identify $\|mv_\parallel\| = mNW_\parallel - qN\mathcal{A}_{f,\parallel} \equiv mNU_\parallel$ and

$$\|mv_\parallel^2\| = P_\parallel + mNU_\parallel^2, \quad \|\mu Bv_\parallel\| = Q_\parallel + P_\perp U_\parallel, \tag{A12}$$

The relation between gyro-fluid quantities $N(\boldsymbol{X},t)$, $U_\parallel(\boldsymbol{X},t)$, ... given in gyro-centre coordinates $\boldsymbol{X}$ and the physical fluid quantities, which we denote with lower case letters

$n(\boldsymbol{x},t) := \int \mathrm{d}^3 v f(\boldsymbol{x},\boldsymbol{v},t)$, $u_\parallel(\boldsymbol{x},t) := \int \mathrm{d}^3 v v_\parallel f(\boldsymbol{x},\boldsymbol{v},t)$ ..., where $f(\boldsymbol{x},\boldsymbol{v},t)$ is the distribution function in particle phase-space (and we here overburden the use of v as the velocity instead of the volume flux-label) is given by equation (49)

$$\|\xi\|_v = \|\zeta\| + \Delta_\perp\left(\frac{m\|\mu B\zeta\|}{2qB^2}\right) + \boldsymbol{\nabla}\cdot\left(\frac{m\|\zeta\|\boldsymbol{\nabla}_\perp\phi}{B^2}\right) \tag{A13}$$

This relation can be inverted up to order δ^2 as for example in equations (51) and (52)

$$N = n - \Delta_\perp\left(\frac{mnt_\perp}{2q^2B^2}\right) - \boldsymbol{\nabla}\cdot\left(\frac{mn}{qB^2}\boldsymbol{\nabla}_\perp\phi\right) \tag{A14}$$

$$NU_\parallel = nu_\parallel - \Delta_\perp\left(\frac{m(q_\parallel + u_\parallel p_\perp)}{2q^2B^2}\right) \tag{A15}$$

Analogous relations hold for the moments of the gyro-kinetic source function S_N and $S_{NU_\parallel}$.

A.3. Fluid velocities

We introduce for any vector $\boldsymbol{u}$

$$u^v := \boldsymbol{u}\cdot\boldsymbol{\nabla}v \qquad \boldsymbol{\nabla}v := \frac{\mathrm{d}v}{\mathrm{d}\psi_p}\boldsymbol{\nabla}\psi_p \tag{A16}$$

$$u_\varphi := \boldsymbol{u}\cdot\boldsymbol{e}_\varphi \qquad u_\vartheta := \boldsymbol{u}\cdot\boldsymbol{e}_\vartheta \qquad u_\parallel := \boldsymbol{u}\cdot\hat{\boldsymbol{b}} \tag{A17}$$

We define the $\boldsymbol{E}\times\boldsymbol{B}$ drift $\boldsymbol{u}_E$, the grad-B drift $\boldsymbol{u}_{\nabla B}$, the diamagnetic drift $\boldsymbol{u}_D$, the curvature drift $\boldsymbol{u}_\kappa$, the first order magnetic fluctuations $\boldsymbol{B}_{1,\perp}$ and the electromagnetic magnetization density $M_\perp^{\mathrm{em}}$

$$\boldsymbol{u}_E := \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\phi}{B} \qquad \boldsymbol{u}_{\nabla B} := t_\perp\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B}{qB} \tag{A18}$$

$$\boldsymbol{u}_D := \frac{\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}p_\perp}{qnB} \qquad \boldsymbol{u}_\kappa := (t_\parallel + mu_\parallel^2)\frac{\hat{\boldsymbol{b}}\times\boldsymbol{\kappa}}{qB} \tag{A19}$$

$$\boldsymbol{B}_{1,\perp} := \boldsymbol{\nabla}A_{1,\parallel}\times\hat{\boldsymbol{b}} \qquad \boldsymbol{M}_\perp^{\mathrm{em}} := \sum_{\mathrm{s}}\frac{m\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}(q_\parallel + p_\perp u_\parallel)}{qB^2} \tag{A20}$$

Note that $\left\langle\widetilde{B_{1,\perp}^v}\right\rangle = \mathcal{O}(\delta^3)$ in the drift ordering and $\boldsymbol{b}_{1,\perp} := \boldsymbol{B}_{1,\perp}/B$. Finally, we have the free current

$$\boldsymbol{j}_f := \sum_{\mathrm{s}} qn(\boldsymbol{u}_\kappa + \boldsymbol{u}_{\nabla B}) \tag{A21}$$

originating in the particle curvature and grad-B drifts.

ORCID iDs

M. Wiesenberger https://orcid.org/0000-0002-5921-0163
M. Held https://orcid.org/0000-0001-8171-8038

References

[1] Scott B. and Smirnov J. 2010 *Phys. Plasmas* **17** 112302
[2] Brizard A.J. and Tronko N. 2011 *Phys. Plasmas* **18** 082307
[3] Stoltzfus-Dueck T. 2012 *Phys. Plasmas* **19** 055908
[4] Rice J.E. *et al* 2007 *Nucl. Fusion* **47** 1618–24
[5] Diamond P.H.*et al* 2013 *Nucl. Fusion* **53** 104019
[6] Stoltzfus-Dueck T. 2019 *Plasma Phys. Control. Fusion* **61** 124003
[7] Helander P. *et al* 2012 *Plasma Phys. Control. Fusion* **54** 124009
[8] Helander P. and Simakov A.N. 2008 *Phys. Rev. Lett.* **101** 145003
[9] Sugama H., Watanabe T.H., Nunami M. and Nishimura S. 2011 *Plasma Phys. Control. Fusion* **53** 024004
[10] Diamond P.H., Itoh S.I., Itoh K. and Hahm T.S. 2005 *Plasma Phys. Control. Fusion* **47** R35–R161
[11] Fujisawa A. 2009 *Nucl. Fusion* **49** 013001
[12] Gürcan O.D. and Diamond P.H. 2015 *J. Phys. A: Math. Theor.* **48** 293001
[13] Connor J.W. and Wilson H.R. 2000 *Plasma Phys. Control. Fusion* **42** R1
[14] Chang C.S., Kue S. and Weitzner H. 2002 *Phys. Plasmas* **9** 3884
[15] Ku S. *et al* 2018 *Phys. Plasmas* **25** 056107
[16] Held M., Wiesenberger M., Kube R. and Kendl A. 2018 *Nucl. Fusion* **58** 104001
[17] Held M., Wiesenberger M. and Kendl A. 2019 *Nucl. Fusion* **59** 026015
[18] Scott B. 2003 *Phys. Lett.* A **320** 53–62
[19] Naulin V., Kendl A., Garcia O.E., Nielsen A.H. and Rasmussen J.J. 2005 *Phys. Plasmas* **12** 052515
[20] Madsen J., Rasmussen J.J., Naulin V. and Nielsen A.H. 2017 *Phys. Plasmas* **24** 062309
[21] Boozer A.H. 2005 *Rev. Mod. Phys.* **76** 1071–141
[22] D'haeseleer W., Hitchon W., Callen J. and Shohet J. 1991 *Flux Coordinates and Magnetic Field Structure* Springer Series in Computational Physics (Berlin: Springer)
[23] Wiesenberger M., Held M. and Einkemmer L. 2017 *J. Comput. Phys.* **340** 435–50
[24] Frankel T. 2004 *The Geometry of Physics: an Introduction* 2nd ed (Cambridge: Cambridge University Press)
[25] Grimm R., Dewar R.L. and Manickam J. 1983 *J. Comput. Phys.* **49** 94–117
[26] Wiesenberger M., Held M., Einkemmer L. and Kendl A. 2018 *J. Comput. Phys.* **373** 370–84
[27] Brizard A.J. and Hahm T.S. 2007 *Rev. Mod. Phys.* **79** 421
[28] Krommes J.A. 2012 *Annu. Rev. Fluid Mech.* **44** 175–201
[29] Tronko N. and Chandre C. 2018 *J. Plasma Phys.* **84** 925840301
[30] Brizard A.J. 2013 *Phys. Plasmas* **20** 092309
[31] Held M., Wiesenberger M. and Kendl A. 2020 *Nucl. Fusion* **60** 066014
[32] T S Hahm L.W. and Madsen J. 2009 *Phys. Plasmas* **16** 022305
[33] Madsen J. 2013 *Phys. Plasmas* **20** 072301
[34] Madsen J. 2010 *Phys. Plasmas* **17** 082107
[35] Jorge R., Ricci P. and Loureiro N.F. 2017 *J. Plasma Phys.* **83** 905830606
[36] Sugama H. 2000 *Phys. Plasmas* **7** 466
[37] Brizard A.J. 1992 *Physics of Fluids B: Plasma Physics* **4** 1213–28
[38] Øksendal B. 2013 *Stochastic Differential Equations: An Introduction With Applications* 6th ed (Berlin: Springer)

[39] Madsen J., Rasmussen J.J., Naulin V., Nielsen A.H. and Treue F. 2015 *Plasma Phys. Control. Fusion* **57** 054016
[40] Held M., Wiesenberger M., Madsen J. and Kendl A. 2016 *Nucl. Fusion* **56** 126005
[41] Simakov A.N. and Catto P.J. 2003 *Phys. Plasmas* **10** 4744
[42] Diamond P.H. and Kim Y. 1991 *Phys. Fluids* B **3** 1626–33
[43] Smolyakov A.I., Diamond P.H. and Medvedev M.V. 2000 *Phys. Plasmas* **7** 3987–92
[44] Craddock G.G. and Diamond P.H. 1991 *Phys. Rev. Lett.* **67** 1535–8
[45] Scott B. 2005 *New J. Phys.* **7** 92
[46] Burrell K.H. 1997 *Phys. Plasmas* **4** 1499
[47] Kendl A. 2003 *Phys. Rev. Lett.* **90** 035006–1
[48] Hassam A.H. and Drake J.F. 1993 *Phys. Fluids* B **5** 4022
[49] Hassam A.H. and Antonsen T.M. 1994 *Phys. Plasmas* **1** 337
[50] Hallatschek K. and Biskamp D. 2000 *Phys. Rev. Lett.* **86** 1223
[51] Helander P., Fulop T. and Catto P. 2003 *Phys. Plasmas* **10** 4396–404
[52] Prager S.C. 1999 *Plasma Phys. Control. Fusion* **41** A129
[53] Ding W.X., Lin L., Brower D.L., Almagri A.F., Chapman B.E., Fiksel G., Den Hartog D.J. and Sarff J.S. 2013 *Phys. Rev. Lett.* **110** 065008
[54] Jackson J.D. 1999 *Classical Electrodynamics* 3rd ed (New York: Wiley)
[55] Akers R. *et al* 2002 *Phys. Plasmas* **9** 3919–29
[56] Binderbauer M.W. *et al* 2015 *Phys. Plasmas* **22** 056110
[57] Medina R. and Stephany J. 2017 The energy-momentum tensor of electromagnetic fields in matter arXiv:1703.02109
[58] Thrysoe A., Loiten M., Madsen J., Naulin V., Nielsen A. and Rasmussen J.J. 2018 *Phys. Plasmas* **25** 032307
[59] deGrassie J., Groebner R., Burrell K.H. and Solomon W. 2009 *Nucl. Fusion* **49** 085020
[60] Wang L. and Diamond P.H. 2013 *Phys. Rev. Lett.* **110** 265006
[61] Peng S. and Wang L. 2017 *Phys. Plasmas* **24** 012304
[62] Peng S., Wang L. and Pan Y. 2017 *Nucl. Fusion* **57** 036003
[63] Wang L., Peng S. and Diamond P.H. 2018 *Plasma Sci. Technol.* **20** 074004
[64] Salewski M. *et al* 2018 *Nucl. Fusion* **58** 036017