Quasisymmetry: A hidden symmetry of magnetic fields

Matt Landreman

August 12, 2019

1 Introduction

Quasisymmetry is a beautiful and remarkable concept in plasma physics. It can be summarized as follows: if the magnitude of the magnetic field $B = |\boldsymbol{B}|$ has continuous symmetry in certain special coordinate systems, in the sense that B is independent of one of the coordinates, then the guiding center particle trajectories behave exactly as if they were in a truly symmetric magnetic field. The full vector magnetic field $\boldsymbol{B}$ need not be symmetric. One reason this idea is exciting is that it means it may be possible to combine the good confinement properties of a tokamak with the stability and steady-state capability of a stellarator. The ∇B and curvature drifts mean that particle trajectories in nonuniform magnetic fields are complicated; quasisymmetry gives a way to guarantee that the particle orbits will be confined even in lights of these cross-field drifts.

Quasisymmetry is closely related to Boozer coordinates, which are a special poloidal angle θ and toroidal angle ζ in a toroidal magnetohydrodynamic (MHD) equilibrium. Symmetry of B in Boozer coordinates is significant, but symmetry of B with respect to other choices of the poloidal and toroidal angles does not generally have the same consequences.

For a first definition of quasisymmetry and its varieties, we can take

$$\begin{aligned}
\text{Quasi-axisymmetry:} \quad & B = B(s,\theta), \\
\text{Quasi-helical symmetry:} \quad & B = B(s, M\theta - N\zeta), \\
\text{Quasi-poloidal symmetry:} \quad & B = B(s,\zeta), \\
\text{Quasisymmetry:} \quad & \text{Any of the above,}
\end{aligned} \tag{1}$$

where s is any coordinate that labels the flux surfaces, and M and N are nonzero integers. These three classes of quasisymmetry are illustrated in figure 1. We will see that there are other equivalent definitions of quasisymmetry.

Quasisymmetric magnetic fields are often loosely described in the following way: the particle orbits and neoclassical transport in a quasisymmetric field are analogous to those in a truly axisymmetric field. ('Neoclassical' refers to the fluxes due to guiding-center motion and collisions, in the absence of turbulence.) There are several precise statements along these lines that can be made:

- The equations of motion for a guiding center, when expressed in Boozer coordinates, depend on position only through s and B.
- The gyro-averaged Lagrangian, i.e. the Lagrangian for guiding center motion, when expressed in Boozer coordinates, depends on position only through s and B.
- In quasisymmetric magnetic fields, there is a conserved quantity analogous to canonical angular momentum that can guarantee confinement of charged particles.

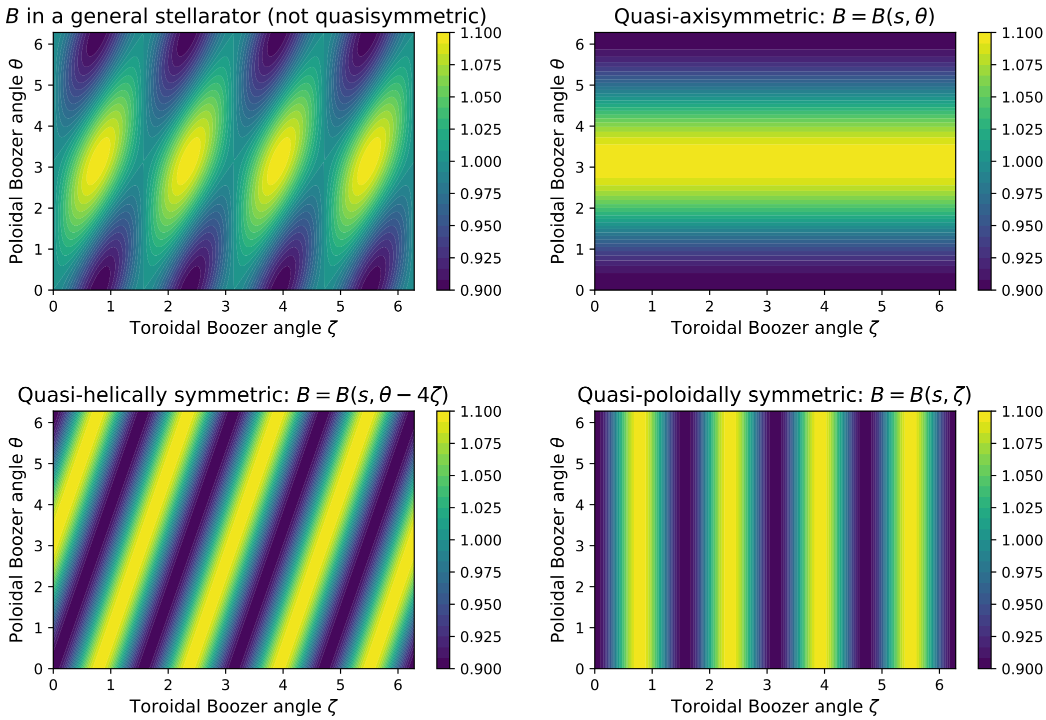

Figure 1: Magnetic field strength B on a typical flux surface in a non-quasisymmetric stellarator and in the three types of quasisymmetric stellarator.

- When B is expressed as a function of Boozer coordinates, neoclassical radial fluxes and parallel flows depend on the magnetic field geometry only through B and functions of s.

All of these points will be discussed in detail in this document. Quasisymmetry also has implications that will not be discussed here, but which can be found in other works:

- In quasisymmetric fields, the radial electric field is not determined by radial neoclassical transport, as it is in a general stellarator [1, 2].
- Larger plasma flows are likely to be allowed in a quasisymmetric stellarator compared to a general stellarator [3], although flow as large as the ion thermal speed is still unlikely to be allowed except in true axisymmetry [4, 5].

1.1 History and further reading

The concept of quasisymmetry has been developed over the years by several authors, and work has been done in the areas of theory, numerics, and experiment. It was first recognized theoretically that neoclassical transport and guiding center orbits depended on position only through (s, B) when expressed in Boozer coordinates, and so two plasmas that differed in Cartesian coordinates could in principle have 'isomorphic' orbits and neoclassical transport [6, 7]. Several years later, it was discovered numerically that quasi-helical configurations could actually be found [8]. Both refs [7, 8] also pointed out the conserved canonical angular momentum. This discovery of numerical quasi-helically symmetric configurations led to the design and construction of the Helically Symmetric Experiment, HSX [9]. Quasi-axisymmetric configurations were discovered numerically thereafter

[10, 11]. The difficulty of achieving quasisymmetry exactly throughout a volume was clarified in [12, 13], where it was pointed out that the relevant system of equations becomes overdetermined at third order in the inverse aspect ratio; however this argument does not rigorously disprove the existence of solutions. The relationship of quasisymmetry to the more general condition of omnigenity was elucidated in [14, 15], and further discussed in [16]. Experimental evidence from HSX for the improved confinement associated with quasisymmetry was reported in [17]. The National Compact Stellarator eXperiment (NCSX) [18], based on a quasi-axisymmetric configuration, was partially constructed, but cancelled before completion. The Chinese First Quasi-axisymmetric Stellarator (CFQS) [19, 20], presently under construction, is likely to be the first quasi-axisymmetric experiment that is fully built. Other quasisymmetric configurations obtained by numerical optimization can be found in [21, 22, 23, 24, 25, 26, 27]. In recent years, the theory of flows, electric fields, and transport in quasisymmetric or nearly-quasisymmetric plasmas has been further developed [1, 2, 4, 5, 3, 28, 29, 30, 31]. Quasisymmetry was discussed in the language of differential forms in [32]. Very recently, new methods of generating quasisymmetric configurations without optimization have been reported [33, 34]. Further theoretical discussion of quasisymmery can be found in the review of Boozer [35] and in the more general review of stellarator theory by Helander [36].

The remaining discussion is structured as follows. First, in section 2 we will discuss standard axisymmetry (continuous rotational symmetry), showing that there is an associated conservation of the so-called canonical angular momentum, and this conservation law can guarantee confinement of particle orbits. However, axisymmetry requires an electric current inside the confinement region, motivating us to seek nonaxisymmetric configurations with equally good confinement. Boozer coordinates are then reviewed in section 3. In section 4 we demonstrate that the trajectory equations for guiding centers in Boozer coordinates depend on position only through s and B. The Lagrangian for guiding-center motion is discussed in section 5, where it is shown that in Boozer coordinates, this Lagrangian depends on position only through s and B. The conserved canonical angular momentum associated with quasisymmetry is discussed in section 6. Neoclassical transport and its relation to quasisymmetry is covered in section 7. Other definitions of quasisymmetry, equivalent to the one given above, are explored in section 8. Some considerations of quasisymmetry near the magnetic axis are discussed in section 9. In section 10 we discuss the accuracy to which quasisymmetry has and can be achieved, and we conclude with some open questions in section 11.

2 Axisymmetry, conservation laws, and confinement

The ∇B and curvature drifts mean that particle trajectories in nonuniform magnetic fields are complicated. How can we guarantee that the particle orbits will be confined even in lights of these cross-field drifts? It turns out that at least in one case, axisymmetry, the particle motion is strongly constrained due to a conservation law. By 'axisymmetry', we mean continuous rotational symmetry. The confinement of orbits in axisymmetric fields is sometimes called Tamm's Theorem, after one of the inventors of the tokamak.

2.1 Lagrangian mechanics and Noether's theorem

We will use Lagrangian mechanics here and in later sections, so let us review the key concepts of Lagrangian mechanics. A Lagrangian is a function $L\left(x^1,\ldots,x^N,\dot{x}^1,\ldots,\dot{x}^N,t\right)$, where t is time, time derivatives are denoted with over-dots, the variables x^j are a set of N coordinates, and the time derivatives $\dot{x}^j$ of each coordinate are treated as independent variables. The Lagrangian for a system is independent of the choice of coordinates. Given a Lagrangian and particular choice of coordinates, the "Euler-Lagrange" equations of motion in those coordinates are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right) = \frac{\partial L}{\partial x^j} \tag{2}$$

for each j. From this equation we can see that if $\partial L/\partial x^j = 0$ for a certain j, then $\partial L/\partial \dot{x}^j$ for that j is independent of time (i.e. conserved.) This fundamental connection between continuous symmetries and conservation laws is known as Noether's Theorem, after the mathematician Emmy Noether. Note that a conservation law is only implied by symmetries that are continuous; discrete symmetry does not count.

2.2 Canonical angular momentum

For a nonrelativistic charged particle in a prescribed static magnetic field, the Lagrangian is [37, 38]

$$L(\boldsymbol{x}, \dot{\boldsymbol{x}}) = q\boldsymbol{A}(\boldsymbol{x}) \cdot \dot{\boldsymbol{x}} + \frac{m}{2}|\dot{\boldsymbol{x}}|^2, \tag{3}$$

where q is the particle charge, m is the particle mass, and $\boldsymbol{A}$ is the vector potential, satisfying $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. We will neglect electric fields. To justify (3), it is sufficient to verify that (2) applied to (3) gives

$$m\ddot{\boldsymbol{x}} = q\dot{\boldsymbol{x}} \times \boldsymbol{B}, \tag{4}$$

which corresponds to Newton's law $\boldsymbol{F} = m\boldsymbol{a}$ with the usual Lorentz force.

Now suppose we have cylindrical coordinates (R, ϕ, Z) and let $(\boldsymbol{e}_R,\ \boldsymbol{e}_\phi,\ \boldsymbol{e}_Z)$ denote the associated unit vectors. Noting that $d\boldsymbol{e}_R/d\phi = \boldsymbol{e}_\phi$ and that the position vector is

$$\boldsymbol{x} = R\boldsymbol{e}_R + Z\boldsymbol{e}_Z, \tag{5}$$

then

$$\dot{\boldsymbol{x}} = \dot{R}\boldsymbol{e}_R + R\dot{\phi}\boldsymbol{e}_\phi + \dot{Z}\boldsymbol{e}_Z. \tag{6}$$

Therefore the Lagrangian can be written

$$L\left(R, \phi, Z, \dot{R}, \dot{\phi}, \dot{Z}\right) = q\left(A_R\dot{R} + A_\phi R\dot{\phi} + A_Z\dot{Z}\right) + \frac{m}{2}\left(\dot{R}^2 + R^2\dot{\phi}^2 + \dot{Z}^2\right) \tag{7}$$

where $A_R = \boldsymbol{A} \cdot \boldsymbol{e}_R$, $A_\phi = \boldsymbol{A} \cdot \boldsymbol{e}_\phi$, and $A_Z = \boldsymbol{A} \cdot \boldsymbol{e}_Z$. Now suppose there is some gauge such that

$$\frac{\partial A_R}{\partial \phi} = 0, \qquad \frac{\partial A_\phi}{\partial \phi} = 0, \qquad \frac{\partial A_Z}{\partial \phi} = 0, \tag{8}$$

i.e. axisymmetry. Then $\partial L/\partial\phi = 0$, so we have a conserved quantity that we will denote p_ϕ:

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = qA_\phi R + mR^2\dot{\phi} = qA_\phi R + mRv_\phi \tag{9}$$

where $v_\phi = R\dot{\phi}$ is the azimuthal component of the particle velocity. The quantity p_ϕ is called the canonical angular momentum. If there is no magnetic field, p_ϕ is the usual kinetic angular momentum. The last term in (9) is bounded in practice, because v_ϕ is limited due to energy conservation. Furthermore, the A_ϕ term can be made much larger than the v_ϕ term by making the magnetic field sufficiently strong. Indeed, writing $A_\phi \sim BL$ where B is a characteristic field strength and L is a characteristic length scale, the ratio of the terms in p_ϕ is

$$\frac{mRv_\phi}{qA_\phi R} \sim \frac{mv_\phi}{qBL} \sim \frac{\rho}{L}, \tag{10}$$

the ratio of the gyroradius ρ to the length scale L, and this ratio can be made $\ll 1$. In this case, particles approximately conserve $A_\phi R$, meaning they are confined to surfaces of constant $A_\phi R$. Even if we do not neglect the v_ϕ term, a particle's departure from a given $A_\phi R$ surface is bounded due to energy conservation. Therefore, if we generate bounded $A_\phi R$ surfaces such as in figure (2), and make the magnitude of the field strong enough (so the departures from surfaces associated with the v_ϕ term are not as large as the confinement region), particles will be confined.

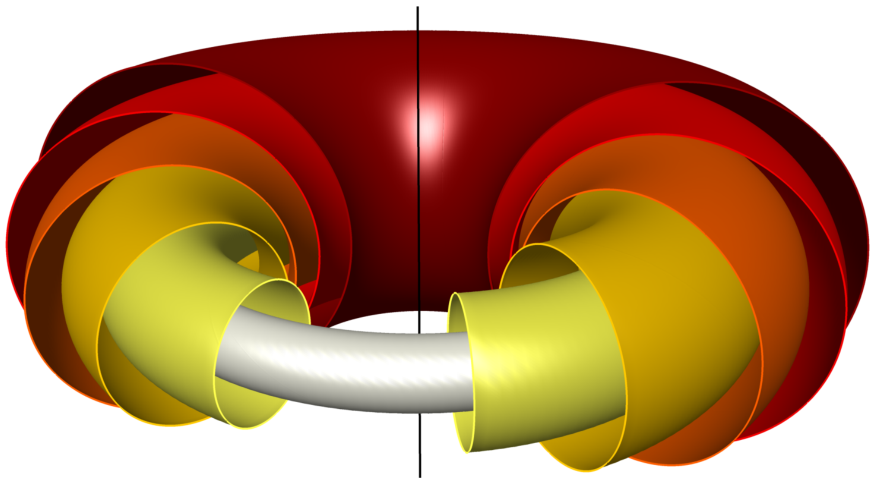

Figure 2: Robust confinement is possible in axisymmetry if the surfaces of $A_\phi R$ are bounded, like the surfaces here. The black vertical line indicates the symmetry axis.

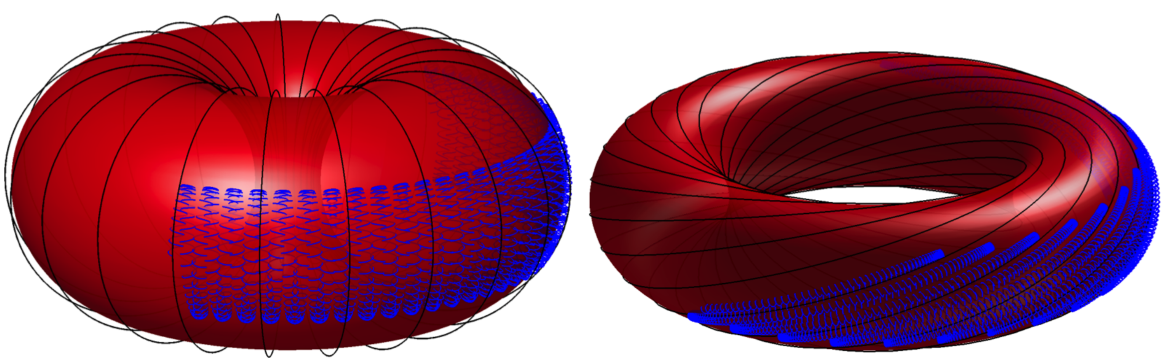

Figure 3: Blue curves denote particle trajectories, black lines denote magnetic field lines. A dipole-like field is shown at left, and a tokamak-like field is shown at right.

Figure 3 shows calculations of a typical trapped particle trajectory in two axisymmetric magnetic fields. In each case, the particle trajectory is computed by direct integration of (4). The trajectories are quite complicated, showing a combination of gyration perpendicular to $\boldsymbol{B}$, mirror-reflection, and cross-field drift. Despite this complexity, the orbits remain confined close to a constant-$A_\phi R$ surface, shown in red. The left case shows an example in which $B_\phi = 0$, as in the dipole fields surrounding some planets. Energetic particles in the Earth's Van Allen belts undergo orbits like the one shown. The right case in figure 3 has a large B_ϕ, as in a tokamak. In both cases, the cross-field drifts cause the particle orbits to precess in the ϕ direction.

2.3 Comments

If one tried to make axisymmetric $A_\phi R$ surfaces that had the topology of nested spheres instead of nested tori, $A_\phi R$ would need to vary with Z along the $R = 0$ axis. But then $B_R = R^{-1}\partial(A_\phi R)/\partial Z$ would diverge along this axis. Regularity of B_R on the symmetry axis therefore leads one to consider nested tori.

It is straightforward to show that (8) is equivalent to the more natural definition of axisymmetry in terms of $\boldsymbol{B}$ rather than $\boldsymbol{A}$:

$$\frac{\partial B_R}{\partial \phi} = 0, \qquad \frac{\partial B_\phi}{\partial \phi} = 0, \qquad \frac{\partial B_Z}{\partial \phi} = 0. \tag{11}$$

That (8) implies (11) follows immediately from $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. Conversely, if $\boldsymbol{B}$ satisfies the axisymmetry conditions (11), a vector potential can always be constructed that satisfies (8), with one example being

$$A_R(R,Z) = \int_0^Z dZ'\, B_\phi(R,Z'), \qquad A_\phi(R,Z) = \frac{1}{R}\int_0^R dR'\, R'\, B_Z(R',Z), \qquad A_Z = 0. \tag{12}$$

You can verify that the curl of (12) gives $\boldsymbol{B}$ using the fact that $\nabla \cdot \boldsymbol{B} = 0$ implies

$$\frac{1}{R}\frac{\partial(RB_R)}{\partial R} = -\frac{\partial B_Z}{\partial Z}. \tag{13}$$

It turns out that the $A_\phi R$ surfaces happen to be magnetic surfaces, also known as flux surfaces, meaning surfaces on which the magnetic field lines lie. To prove this fact, we can note

$$B_R = -\frac{\partial A_\phi}{\partial Z}, \qquad B_Z = \frac{1}{R}\frac{\partial(A_\phi R)}{\partial R}, \tag{14}$$

from which it follows that

$$\boldsymbol{B} \cdot \nabla(A_\phi R) = B_R \frac{\partial(A_\phi R)}{\partial R} + B_Z \frac{\partial(A_\phi R)}{\partial Z} = 0. \tag{15}$$

Therefore the field lines of $\boldsymbol{B}$ are tangent to the constant-$A_\phi R$ surfaces.

2.4 Problems with confinement using axisymmetry

Unfortunately, in order to form toroidal nested $A_\phi R$ surfaces in axisymmetry, there must be an electric current inside the surfaces, i.e. inside the confinement region. This fact can be proved from the integral form of Ampere's Law,

$$\mu_0 I = \oint \boldsymbol{B} \cdot d\boldsymbol{l}, \tag{16}$$

where I is the current through the curve used for the integration. We apply this equation to a constant-ϕ curve lying on one of the toroidal $A_\phi R$ surfaces, with the curve linking the surface the short way around ('poloidally'). Since the length increment $\boldsymbol{dl}$ is perpendicular to both $\boldsymbol{e}_\phi$ and to $\nabla(A_\phi R)$,

$$\boldsymbol{dl} \mid\mid \nabla(A_\phi R) \times \boldsymbol{e}_\phi = -\boldsymbol{e}_R \frac{\partial(A_\phi R)}{\partial Z} + \boldsymbol{e}_Z \frac{\partial(A_\phi R)}{\partial R}. \tag{17}$$

We therefore have

$$\boldsymbol{dl} = \left[\left(\frac{\partial(A_\phi R)}{\partial R} \right)^2 + \left(\frac{\partial(A_\phi R)}{\partial Z} \right)^2 \right]^{-1/2} \left[-\boldsymbol{e}_R \frac{\partial(A_\phi R)}{\partial Z} + \boldsymbol{e}_Z \frac{\partial(A_\phi R)}{\partial R} \right] d\ell, \tag{18}$$

where $d\ell = |\boldsymbol{dl}|$. Using (14), the integrand of (16) is

$$\boldsymbol{B} \cdot \boldsymbol{dl} = \frac{1}{R} \left[\left(\frac{\partial(A_\phi R)}{\partial R} \right)^2 + \left(\frac{\partial(A_\phi R)}{\partial Z} \right)^2 \right]^{1/2} d\ell = \frac{|\nabla(A_\phi R)|}{R} d\ell, \tag{19}$$

which is never negative, and which must be positive (except perhaps at isolated points) to have nested toroidal surfaces. Therefore, the enclosed current I in (16) must be nonzero.

The current inside the confinement region is a problem for several reasons. First, the current is hard to drive. A transformer effect can be used to drive the current for a short time, but not for long durations. Injection of particles or radio-frequency waves can be used to drive the current, but this takes significant energy, and would require using a substantial fraction of the electricity generated in a fusion reactor. Furthermore, however the current is produced, the current tends to be unstable. The current represents a non-equilibrium situation, and hence it is a reservoir of free energy which instabilities can feed off of. If the current disrupts due to an instability, the nested $A_\phi R$ surfaces disappear and confinement is lost.

We are therefore motivated to look for ways to achieve confinement without the axisymmetry condition (11). Without axisymmetry, there is no particular reason for particles to stay in a confined region as they undergo the ∇B and curvature drifts. However, it turns out that condition of quasisymmetry will ensure good confinement even in the absence of axisymmetry.

3 Boozer coordinates

Before proceeding to discuss quasisymmetry in detail, it is valuable to understand Boozer coordinates. Boozer coordinates are (s, θ, ζ) where s is any flux surface label, and θ and ζ are poloidal and toroidal angles defined in a particular way. For concreteness, we will use $s = \psi$, where $2\pi\psi$ is the toroidal flux, though this choice is not necessary. The three coordinates (s, θ, ζ) are generally not orthogonal, so let us review some general characteristics of non-orthogonal coordinates in $\mathbb{R}^3$. Further discussion of general coordinate systems can be found in [39] or sections 2.10-2.11 of [40].

3.1 General non-orthogonal coordinates

For non-orthogonal coordinates, there are two natural sets of basis vectors that can be used to represent general vector quantities. One set of basis vectors is $(\nabla\psi, \nabla\theta, \nabla\zeta)$, where

$$\nabla\psi = \boldsymbol{e}_x \frac{\partial\psi}{\partial x} + \boldsymbol{e}_y \frac{\partial\psi}{\partial y} + \boldsymbol{e}_z \frac{\partial\psi}{\partial z} \tag{20}$$

with analogous definitions for $\nabla\theta$ and $\nabla\zeta$. Here, y and z are held fixed in the $\partial/\partial x$ differentiation, the analogous statements hold for $\partial/\partial y$ and $\partial/\partial z$, and $\{\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z\}$ are the Cartesian unit basis vectors. The other set of basis vectors is $(\partial\boldsymbol{x}/\partial\psi,\ \partial\boldsymbol{x}/\partial\theta,\ \partial\boldsymbol{x}/\partial\zeta)$, where $\boldsymbol{x}$ is the position vector, and

$$\frac{\partial\boldsymbol{x}}{\partial\psi}=\boldsymbol{e}_x\frac{\partial x}{\partial\psi}+\boldsymbol{e}_y\frac{\partial y}{\partial\psi}+\boldsymbol{e}_z\frac{\partial z}{\partial\psi}, \tag{21}$$

with analogous definitions for $\partial\boldsymbol{x}/\partial\theta$ and $\partial\boldsymbol{x}/\partial\zeta$. Here, in contrast to (20), θ and ζ are held fixed in the $\partial/\partial\psi$ differentiation, and analogous statements hold for $\partial/\partial\theta$ and $\partial/\partial\zeta$. If the coordinates were orthogonal, then $\nabla\psi$ would be parallel to $\partial\boldsymbol{x}/\partial\psi$, and similarly for the other two coordinates. However, for non-orthogonal coordinates these directions are generally different. This important difference between orthogonal vs non-orthogonal coordinates is illustrated in figure 4.

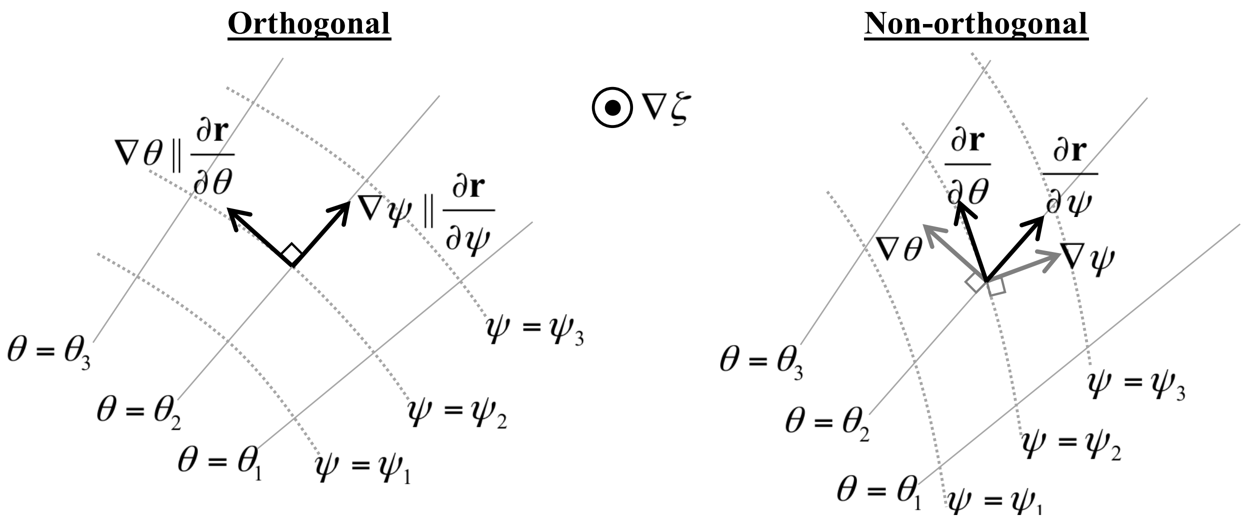

Figure 4: Orthogonal vs non-orthogonal coordinates.

For any vector $\boldsymbol{B}$, (not only the magnetic field), we can decompose the vector in either basis, writing

$$\boldsymbol{B}=B_\psi\nabla\psi+B_\theta\nabla\theta+B_\zeta\nabla\zeta \tag{22}$$

and

$$\boldsymbol{B}=B^\psi\frac{\partial\boldsymbol{x}}{\partial\psi}+B^\theta\frac{\partial\boldsymbol{x}}{\partial\theta}+B^\zeta\frac{\partial\boldsymbol{x}}{\partial\zeta}. \tag{23}$$

We can relate the two sets of basis vectors as follows. First, note that the vector $\partial\boldsymbol{x}/\partial\psi$ by definition points in a direction along which θ and ζ do not increase. Thus,

$$\frac{\partial\boldsymbol{x}}{\partial\psi}\cdot\nabla\theta=0 \quad\text{and}\quad \frac{\partial\boldsymbol{x}}{\partial\psi}\cdot\nabla\zeta=0. \tag{24}$$

It follows that

$$\frac{\partial\boldsymbol{x}}{\partial\psi}=J\nabla\theta\times\nabla\zeta \tag{25}$$

for some coefficient J. To determine this coefficient, consider a step $d\boldsymbol{x}$ at fixed θ and ζ: $d\psi=d\boldsymbol{x}\cdot\nabla\psi$, so $(\partial\boldsymbol{x}/\partial\psi)\cdot\nabla\psi=1$. (This same result can also be seen by forming the dot product of (20) with (21) and recognizing the result as the chain rule applied to $\partial\psi(\boldsymbol{x})/\partial\psi=1$, where $\boldsymbol{x}=\boldsymbol{x}(x,y,z)$.) We therefore have

$$\frac{\partial\boldsymbol{x}}{\partial x^j}\cdot\nabla x^i=\delta^i_j. \tag{26}$$

The dot product of $\nabla\psi$ with (25) now tells us

$$J = \frac{1}{\nabla\psi \cdot \nabla\theta \times \nabla\zeta}. \tag{27}$$

By repeating the same analysis with cyclic permutations of the coordinates, we arrive at the so-called 'dual relations':

$$\frac{\partial \boldsymbol{x}}{\partial \psi} = J\nabla\theta \times \nabla\zeta, \qquad \frac{\partial \boldsymbol{x}}{\partial \theta} = J\nabla\zeta \times \nabla\psi, \qquad \frac{\partial \boldsymbol{x}}{\partial \zeta} = J\nabla\psi \times \nabla\theta. \tag{28}$$

Therefore, an equivalent expression to (23) is

$$\boldsymbol{B} = \frac{1}{J}\left(B^{\psi}\nabla\theta \times \nabla\zeta + B^{\theta}\nabla\zeta \times \nabla\psi + B^{\zeta}\nabla\psi \times \nabla\theta\right). \tag{29}$$

Now let us specialize to the case in which $\boldsymbol{B}$ is the magnetic field. If good magnetic surfaces exist, then $\boldsymbol{B} \cdot \nabla\psi = 0$. Applying $\cdot\nabla\psi$ to (29), then, we find $B^{\psi} = 0$. There are consequently 5 nonzero components of $\boldsymbol{B}$ between the two representations: $(B_{\psi}, B_{\theta}, B_{\zeta}, B^{\theta}, B^{\zeta})$.

3.2 Boozer coordinates

There are many choices which can be made for θ and ζ. One choice for ζ is the standard cylindrical azimuthal angle ϕ, and one choice for θ is $\text{atan2}(z - z_0,\ R - R_0)$ where atan2 is the arctangent with appropriate sign, $R = \sqrt{x^2 + y^2}$ is the cylindrical radius (i.e. major radius), and $(z_0(\zeta),\ R_0(\zeta))$ are the location of the magnetic axis. However, we can add any single-valued functions to one set of angles to get another set of angle coordinates:

$$\theta' = \theta + f(\boldsymbol{x}), \qquad \zeta' = \zeta + g(\boldsymbol{x}) \tag{30}$$

It can be shown that this freedom can be used to make two of the components of $\boldsymbol{B}$ simplify:

$$B^{\theta}/J = \iota, \qquad B^{\zeta}/J = 1, \tag{31}$$

where $\iota(\psi)$ is the rotational transform, which is constant on a magnetic surface. The details of this construction will not be given here, but they can be found in section 2.2 of [36]. Then using the result

$$\frac{d\psi_p}{d\psi} = \iota \tag{32}$$

where $2\pi\psi_p$ is the poloidal flux, then (29) becomes

$$\begin{aligned} \boldsymbol{B} &= \nabla\psi \times \nabla\theta + \iota\nabla\zeta \times \nabla\psi \\ &= \nabla\psi \times \nabla\theta + \nabla\zeta \times \nabla\psi_p. \end{aligned} \tag{33}$$

Coordinates in which (33) is satisfied are "straight field line coordinates". Forming the dot product of (33) with $\nabla\zeta$ or $\nabla\theta$ yields the following useful results:

$$\boldsymbol{B} \cdot \nabla\zeta = \nabla\psi \cdot \nabla\theta \times \nabla\zeta = 1/J, \tag{34}$$

$$\boldsymbol{B} \cdot \nabla\theta = \iota\nabla\psi \cdot \nabla\theta \times \nabla\zeta = \iota\boldsymbol{B} \cdot \nabla\zeta = \iota/J. \tag{35}$$

There is still freedom left in the transformation (30), and this freedom can be used to simplify other components of $\boldsymbol{B}$. Boozer coordinates are constructed by using the transformation (30) to make B_{θ} and B_{ζ} constant on magnetic surfaces, with these flux functions usually denoted I and G:

$$B_{\theta} = I(\psi), \qquad B_{\zeta} = G(\psi). \tag{36}$$

We will not give the construction here; it is given in section 2.5 of [36]. The remaining coefficient B_ψ still generally depends on all three coordinates, and it is often denoted β since it turns out to be related to the plasma pressure (as we will show in the next section). Hence,

$$\boldsymbol{B} = \beta(\psi, \theta, \zeta)\nabla\psi + I(\psi)\nabla\theta + G(\psi)\nabla\zeta. \tag{37}$$

Taking the dot product of (33) with (37), we obtain a useful expression:

$$\nabla\psi \cdot \nabla\theta \times \nabla\zeta = \frac{B^2}{G(\psi) + \iota I(\psi)}. \tag{38}$$

This quantity is the (inverse) Jacobian of the transformation between Cartesian and Boozer coordinates. For understanding quasisymmetry, it is noteworthy that this Jacobian varies on a magnetic surface only through B.

3.3 The radial Boozer component

A result we will need below is the following: a symmetry of B in Boozer angles implies that β has the same symmetry. This result is obtained by considering the MHD equilibrium relation

$$\boldsymbol{j} \times \boldsymbol{B} = \nabla p, \tag{39}$$

where $\mu_0 \boldsymbol{j} = \nabla \times \boldsymbol{B}$, $\boldsymbol{j}$ is the current density, and $p = p(\psi)$ is the plasma pressure. We evaluate $\boldsymbol{j}$ from the curl of (37):

$$\nabla \times \boldsymbol{B} = \nabla\beta \times \nabla\psi + \frac{dI}{d\psi}\nabla\psi \times \nabla\theta + \frac{dG}{d\psi}\nabla\psi \times \nabla\zeta. \tag{40}$$

In the first right-hand side term, note that

$$\nabla\beta = \frac{\partial\beta}{\partial\psi}\nabla\psi + \frac{\partial\beta}{\partial\theta}\nabla\theta + \frac{\partial\beta}{\partial\zeta}\nabla\zeta. \tag{41}$$

Forming (40) $\times$ (33), one finds the MHD equilibrium equation (39) becomes

$$\left(\frac{\partial\beta}{\partial\zeta} + \iota\frac{\partial\beta}{\partial\theta} - \frac{dG}{d\psi} - \iota\frac{dI}{d\psi}\right)(\nabla\psi \cdot \nabla\theta \times \nabla\zeta)\,\nabla\psi = \mu_0 \frac{dp}{d\psi}\nabla\psi. \tag{42}$$

The directions of the right and left sides of this equation are always parallel, so we can remove the $\nabla\psi$ vectors from each side. Applying (38),

$$\frac{\partial\beta}{\partial\zeta} + \iota\frac{\partial\beta}{\partial\theta} - \frac{dG}{d\psi} - \iota\frac{dI}{d\psi} = \mu_0 \frac{G + \iota I}{B^2}\frac{dp}{d\psi}. \tag{43}$$

Writing Fourier expansions for $1/B^2$ and β,

$$\frac{1}{B^2} = \sum_{m,n} \mathcal{B}_{m,n} e^{im\theta - in\zeta}, \tag{44}$$
$$\beta = \sum_{m,n} \beta_{m,n} e^{im\theta - in\zeta},$$

then (43) gives

$$\beta_{m,n} = \frac{\mu_0 \mathcal{B}_{m,n}(G + \iota I)}{i(\iota m - n)}\frac{dp}{d\psi}. \tag{45}$$

(Note that the $m = n = 0$ term of β is not determined by (43).) Therefore, if B is independent of ζ, so all the $n \neq 0$ terms in $\mathcal{B}_{m,n}$ vanish, then all the $n \neq 0$ terms of $\beta_{m,n}$ will vanish as well, so β will be independent of ζ.

More generally, suppose B depends on θ and ζ only through the helical combination

$$\chi = M\theta - N\zeta \tag{46}$$

where M and N and fixed integers. Then this same property is true of $1/B^2$. Hence, $\mathcal{B}_{m,n}$ vanishes unless $m = kM$ and $n = kN$ for some rational number k. By (45), then $\beta_{m,n}$ vanishes unless $m = kM$ and $n = kN$. Hence, β depends on θ and ζ only through the combination χ, exactly like B.

4 Direct demonstration that orbits in quasisymmetric fields are isomorphic to those in a tokamak

We are now in a position to demonstrate one of the key properties of quasisymmetric fields: the equations of motion for guiding centers, when expressed in Boozer coordinates, depend on position only through ψ and B.

We will proceed by finding equations for the time derivative of the coordinates (ψ, θ, ζ). These equations will turn out to depend on position only through ψ and B. Therefore, if we have a magnetic field configuration in which B depends on position only through (ψ, θ), independent of ζ, then the guiding-center dynamics will be the same as if the particle were in an axisymmetric field (with the same profile functions $p(\psi)$, $\iota(\psi)$, $B_\theta(\psi)$, and $B_\zeta(\psi)$.) Since the guiding-center trajectories are confined in the axisymmetric analogous field, the are confined in the original field as well.

As a first step, it is useful to write the magnetic drifts together in a convenient form as follows. Observe that in a general MHD equilibrium, defining $\boldsymbol{b} = \boldsymbol{B}/B$ to be the unit vector along the field,

$$\begin{aligned} \boldsymbol{b}\cdot\nabla\boldsymbol{b} &= (\nabla\times\boldsymbol{b})\times\boldsymbol{b} + \underbrace{\nabla\left(\frac{\boldsymbol{b}\cdot\boldsymbol{b}}{2}\right)}_{=0} = -\frac{(\nabla B\times\boldsymbol{B})\times\boldsymbol{B}}{B^3} + \frac{(\nabla\times\boldsymbol{B})\times\boldsymbol{B}}{B^2} \\ &= -\frac{\boldsymbol{B}\boldsymbol{B}\cdot\nabla B}{B^3} + \frac{\nabla B}{B} + \frac{\mu_0\nabla p}{B^2}. \end{aligned} \tag{47}$$

Therefore the ∇B and curvature drifts

$$\boldsymbol{v}_{\nabla B} = \frac{mv_\perp^2}{2qB^3}\boldsymbol{B}\times\nabla B, \qquad \boldsymbol{v}_\kappa = \frac{mv_{||}^2}{qB^2}\boldsymbol{B}\times(\boldsymbol{b}\cdot\nabla\boldsymbol{b}), \tag{48}$$

where m and q are the particle mass and charge, can be written together as

$$\boldsymbol{v}_d = \boldsymbol{v}_{\nabla B} + \boldsymbol{v}_\kappa = \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\boldsymbol{B}\times\nabla B + \frac{mv_{||}^2\mu_0}{qB^4}\frac{dp}{d\psi}\boldsymbol{B}\times\nabla\psi. \tag{49}$$

Evidently the ∇B and curvature drifts are in the same direction when there is no plasma pressure.

4.1 Vacuum fields

The demonstration of isomorphic orbits for vacuum fields requires less algebra than the case of general magnetic fields, so let us show the vacuum case first. The general case is treated in the next section.

A few results for vacuum fields that we will use are

$$\boldsymbol{B} = G\nabla\zeta, \qquad \boldsymbol{B}\cdot\nabla\zeta = \nabla\psi\cdot\nabla\theta\times\nabla\zeta = B^2/G, \tag{50}$$

We have set $\beta = 0$ since $\beta \propto dp/d\psi$ from (45). (The $m = n = 0$ Fourier mode of β can be set to zero by suitable choice of the origin of θ or ζ.) We have also used the fact that $I(\psi)$ is $\mu_0/(2\pi)$ times the toroidal current inside the surface ψ, which vanishes in vacuum, hence $I = 0$. Since $G(\psi)$ is $\mu_0/(2\pi)$ times the poloidal current outside the surface ψ, then G in a vacuum field is a constant (independent of ψ).

To get the rate of change of any of the coordinates x^j, we take the dot product of the total guiding-center velocity $v_{||}\boldsymbol{b} + \boldsymbol{v}_d$ into the gradient of that coordinate. For instance, again using over-dots to denote time derivatives,

$$\begin{aligned}
\dot{\psi} &= (v_{||}\boldsymbol{b} + \boldsymbol{v}_d)\cdot\nabla\psi = \boldsymbol{v}_d\cdot\nabla\psi = \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\boldsymbol{B}\times\nabla B\cdot\nabla\psi \\
&= \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)(G\nabla\zeta)\times\left(\frac{\partial B}{\partial\theta}\nabla\theta\right)\cdot\nabla\psi \\
&= -\frac{m}{qB}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\frac{\partial B}{\partial\theta}.
\end{aligned} \tag{51}$$

Next,

$$\dot{\zeta} = (v_{||}\boldsymbol{b} + \boldsymbol{v}_d)\cdot\nabla\zeta = \frac{v_{||}}{B}\boldsymbol{B}\cdot\nabla\zeta + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\underbrace{\boldsymbol{B}\times\nabla B\cdot\nabla\zeta}_{=0} = \frac{v_{||}B}{G}. \tag{52}$$

Finally,

$$\begin{aligned}
\dot{\theta} &= (v_{||}\boldsymbol{b} + \boldsymbol{v}_d)\cdot\nabla\theta = \frac{v_{||}}{B}\boldsymbol{B}\cdot\nabla\theta + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\boldsymbol{B}\times\nabla B\cdot\nabla\theta \\
&= \frac{v_{||}}{B}\frac{\iota B^2}{G} + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)(G\nabla\zeta)\times\left(\frac{\partial B}{\partial\psi}\nabla\psi\right)\cdot\nabla\theta \\
&= \frac{v_{||}\iota B}{G} + \frac{m}{qB}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\frac{\partial B}{\partial\psi}.
\end{aligned} \tag{53}$$

Note that the kinetic energy $v^2/2$ and magnetic moment $\mu = v_\perp^2/(2B)$ are constants of the guiding-center motion, so $v_\perp$ and $v_{||}$ do vary with position, but only through B:

$$v_\perp^2 = 2\mu B, \qquad v_{||} = \pm\sqrt{v^2 - 2\mu B}. \tag{54}$$

Therefore, (51)-(53) show that the equations for the guiding-center orbits in Boozer coordinates depend on position only through $\iota(\psi)$ and B.

4.2 General MHD equilibrium

Now we will re-calculate $\dot{\psi}$, $\dot{\theta}$, and $\dot{\zeta}$ for the case of MHD equilibrium rather than a vacuum field. Using (49), we find

$$\begin{aligned}
\dot{\psi} &= (v_{||}\boldsymbol{b} + \boldsymbol{v}_d) \cdot \nabla\psi = \boldsymbol{v}_d \cdot \nabla\psi = \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right) \boldsymbol{B} \times \nabla B \cdot \nabla\psi \\
&= \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\left[(G\nabla\zeta) \times \left(\frac{\partial B}{\partial\theta}\nabla\theta\right) \cdot \nabla\psi + (I\nabla\theta) \times \left(\frac{\partial B}{\partial\zeta}\nabla\zeta\right) \cdot \nabla\psi\right] \\
&= \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\left(-G\frac{\partial B}{\partial\theta} + I\frac{\partial B}{\partial\zeta}\right)\frac{B^2}{G + \iota I},
\end{aligned} \tag{55}$$

$$\begin{aligned}
\dot{\theta} &= (v_{||}\boldsymbol{b} + \boldsymbol{v}_d) \cdot \nabla\theta = \frac{v_{||}}{B}\boldsymbol{B} \cdot \nabla\theta + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right) \boldsymbol{B} \times \nabla B \cdot \nabla\theta + \frac{m v_{||}^2 \mu_0}{qB^4}\frac{dp}{d\psi}\boldsymbol{B} \times \nabla\psi \cdot \nabla\theta \\
&= \frac{v_{||}\iota}{B}\frac{B^2}{G + \iota I} + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\left[(G\nabla\zeta) \times \left(\frac{\partial B}{\partial\psi}\nabla\psi\right) \cdot \nabla\theta + (\beta\nabla\psi) \times \left(\frac{\partial B}{\partial\zeta}\nabla\zeta\right) \cdot \nabla\theta\right] \\
&\qquad + \frac{m v_{||}^2 \mu_0}{qB^4}\frac{dp}{d\psi} G\nabla\zeta \times \nabla\psi \cdot \nabla\theta \\
&= \frac{v_{||}\iota B}{G + \iota I} + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\left(G\frac{\partial B}{\partial\psi} - \beta\frac{\partial B}{\partial\zeta}\right)\frac{B^2}{G + \iota I} + \frac{m v_{||}^2 \mu_0}{qB^4}\frac{dp}{d\psi}\frac{GB^2}{G + \iota I},
\end{aligned} \tag{56}$$

$$\begin{aligned}
\dot{\zeta} &= (v_{||}\boldsymbol{b} + \boldsymbol{v}_d) \cdot \nabla\zeta = \frac{v_{||}}{B}\boldsymbol{B} \cdot \nabla\zeta + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right) \boldsymbol{B} \times \nabla B \cdot \nabla\zeta + \frac{m v_{||}^2 \mu_0}{qB^4}\frac{dp}{d\psi}\boldsymbol{B} \times \nabla\psi \cdot \nabla\zeta \\
&= \frac{v_{||}}{B}\frac{B^2}{G + \iota I} + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\left[(I\nabla\theta) \times \left(\frac{\partial B}{\partial\psi}\nabla\psi\right) \cdot \nabla\zeta + (\beta\nabla\psi) \times \left(\frac{\partial B}{\partial\theta}\nabla\theta\right) \cdot \nabla\zeta\right] \\
&\qquad + \frac{m v_{||}^2 \mu_0}{qB^4}\frac{dp}{d\psi} I\nabla\theta \times \nabla\psi \cdot \nabla\zeta \\
&= \frac{v_{||} B}{G + \iota I} + \frac{m}{qB^3}\left(\frac{v_\perp^2}{2} + v_{||}^2\right)\left(\beta\frac{\partial B}{\partial\theta} - I\frac{\partial B}{\partial\psi}\right)\frac{B^2}{G + \iota I} - \frac{m v_{||}^2 \mu_0}{qB^4}\frac{dp}{d\psi}\frac{IB^2}{G + \iota I}.
\end{aligned} \tag{57}$$

Noting that β is determined by B via (45), it can be seen that these equations for $\dot{\psi}$, $\dot{\theta}$, and $\dot{\zeta}$ depend on position only through B and functions of ψ.

5 The guiding-center Lagrangian

Another perspective on quasisymmetry comes from the Lagrangian for guiding center motion. We will show that in Boozer coordinates, this Lagrangian depends on position only through the radial (flux surface label) coordinate and through B. There are two important consequences. First, the guiding-center dynamics of two plasmas with the same B and radial profiles are isomorphic. Second, if B can be made independent of one of the Boozer angles (or a linear combination thereof), there will be a conserved quantity by Noether's theorem. This conserved quantity ensures good confinement of the guiding centers (Tamm's theorem), just as it did in axisymmetry.

The Lagrangian for guiding center motion has been derived by Littlejohn [41], and is

$$L(\boldsymbol{x}, v_{||}, \mu, \varphi, \dot{\boldsymbol{x}}, \dot{v}_{||}, \dot{\mu}, \dot{\varphi}, t) = q\boldsymbol{A}\cdot\dot{\boldsymbol{x}} + mv_{||}\boldsymbol{b}\cdot\dot{\boldsymbol{x}} + \frac{m}{q}\mu\dot{\varphi} - \frac{mv_{||}^2}{2} - \mu B - q\Phi, \tag{58}$$

where a dot on top of a symbol denotes a time derivative. Here, $\boldsymbol{b} = \boldsymbol{B}/B$ is a unit vector, μ is the magnetic moment, φ is the gyrophase, $\boldsymbol{A}$ is the vector potential, and Φ is the electrostatic potential. Forming the Euler-Lagrange equations for this Lagrangian, one finds the following equations of motion:

$$\dot{\boldsymbol{x}} = v_{||}\boldsymbol{b} + \frac{1}{qB_{||}^*}\boldsymbol{b}\times\left(mv_{||}^2\boldsymbol{b}\cdot\nabla\boldsymbol{b} + \mu\nabla\boldsymbol{b} - q\boldsymbol{E}^*\right), \tag{59}$$

$$\dot{v}_{||} = -\frac{1}{m}\boldsymbol{b}\cdot(\mu\nabla B - q\boldsymbol{E}) + v_{||}\boldsymbol{b}\cdot\nabla\boldsymbol{b}\cdot\dot{\boldsymbol{x}}, \tag{60}$$

$$\dot{\varphi} = qB/m, \tag{61}$$

$$\dot{\mu} = 0, \tag{62}$$

where $\boldsymbol{E}^* = -\nabla\Phi - \partial\boldsymbol{A}^*/\partial t$, $B_{||}^* = \boldsymbol{B}^*\cdot\boldsymbol{b}$, $\boldsymbol{B}^* = \nabla\times\boldsymbol{A}^*$, and $\boldsymbol{A}^* = \boldsymbol{A} + m(q)^{-1}v_{||}\boldsymbol{b}$. To leading order in $\rho_* = \rho/L \sim v/(\Omega L)$, where $\Omega = qB/m$ and L is a typical scale length, then $B_{||}^* \approx B$. Note also that the last term in (60) is small compared to the first right-hand-side term by a factor of ρ_*, since $\boldsymbol{b}\cdot\nabla\boldsymbol{b}\cdot\boldsymbol{b} = 0$. Hence, (59)-(62) give the usual guiding center equations of motion. The Lagrangian for guiding-center motion is also discussed in section 6.3 of [42].

Let us now write the Lagrangian (58) in Boozer coordinates. We will assume $\boldsymbol{B}$ is time-independent for simplicity. We need an expression for $\boldsymbol{A}$ in some gauge, and a suitable expression is

$$\boldsymbol{A} = \psi\nabla\theta - \psi_p\nabla\zeta, \tag{63}$$

where again $2\pi\psi_p$ is the poloidal flux. It can be seen that the curl of this expression is (33). The Lagrangian (58) also includes $\dot{\boldsymbol{x}}$, which can be written

$$\dot{\boldsymbol{x}} = \dot{\psi}\frac{\partial\boldsymbol{x}}{\partial\psi} + \dot{\theta}\frac{\partial\boldsymbol{x}}{\partial\theta} + \dot{\zeta}\frac{\partial\boldsymbol{x}}{\partial\zeta}. \tag{64}$$

Substituting (63)-(64) into the Lagrangian (58), and using (26), we obtain

$$L(\boldsymbol{x}, v_{||}, \mu, \varphi, \dot{\boldsymbol{x}}, \dot{v}_{||}, \dot{\mu}, \dot{\varphi}) = q\psi\dot{\theta} - q\psi_p\dot{\zeta} + \frac{mv_{||}}{B}\left(\dot{\psi}\beta + \dot{\theta}I + \dot{\zeta}G\right) + \frac{m}{q}\mu\dot{\varphi} - \frac{mv_{||}^2}{2} - \mu B - q\Phi. \tag{65}$$

It can be seen that this Lagrangian depends on position only through (1) the flux functions ψ, ψ_p, I, and G, (2) the field magnitude B, and (3) the electrostatic potential Φ. Thus if B and Φ are independent of ζ, the particle dynamics will be exactly analogous to the dynamics in a tokamak. If B is independent of ζ, the mean Φ will tend to be independent of ζ as well since Φ is a density moment of the guiding-center distribution function. Hence, the requirement that Φ have symmetry is not a major restriction.

An ignorable coordinate can also arise in the case of helical symmetry. To define this kind of symmetry, we introduce

$$\chi = M\theta - N\zeta \tag{66}$$

where M and N and fixed integers. Noting $\dot{\chi} = M\dot{\theta} - N\dot{\zeta}$, (65) is equivalent to

$$L = \frac{q\psi}{M}\left(\dot{\chi} + N\dot{\zeta}\right) - q\psi_p\dot{\zeta} + \frac{mv_{||}}{B}\left[\dot{\psi}\beta + \frac{I}{M}\left(\dot{\chi} + N\dot{\zeta}\right) + \dot{\zeta}G\right] + \frac{m}{q}\mu\dot{\varphi} - \frac{mv_{||}^2}{2} - \mu B - q\Phi. \tag{67}$$

If $B = B(\psi, \chi)$, i.e. if B depends on θ and ζ only through the combination χ, then (67) looks identical to the Lagrangian (65) for an axisymmetric magnetic field $B(\psi, \theta)$, but with certain substitutions made.

6 Conserved canonical angular momentum and Tamm's theorem

6.1 Noether's theorem

Now that we have the guiding-center Lagrangian, let us return to Noether's theorem. Again, this theorem is the observation that for any coordinate x^j and Lagrangian L, if $\partial L/\partial x^j = 0$ (i.e. x^j is an ignorable coordinate), then the associated Euler-Lagrange equation is $(d/dt)\partial L/\partial \dot{x}^j = 0$, so $\partial L/\partial \dot{x}^j$ is a conserved quantity. From (65), we see that in the case of quasi-axisymmetry $B = B(\psi, \theta)$, the conserved quantity is

$$p_\zeta = \left(\frac{\partial L}{\partial \dot{\zeta}}\right)_{\theta, \dot{\theta}, \zeta} = -q\psi_p + \frac{m v_{||} G}{B}. \tag{68}$$

Similarly, in the case of quasi-poloidal symmetry, $B = B(\psi, \zeta)$, the conserved quantity from (65) is

$$p_\theta = \left(\frac{\partial L}{\partial \dot{\theta}}\right)_{\zeta, \dot{\zeta}, \theta} = q\psi + \frac{m v_{||} I}{B}. \tag{69}$$

For helical symmetry, where the Lagrangian is independent of ζ at fixed χ, the conserved quantity obtained from (67) is

$$p_\chi = \left(\frac{\partial L}{\partial \dot{\zeta}}\right)_{\chi, \dot{\chi}, \zeta} = q\frac{N}{M}\psi - q\psi_p + \frac{m v_{||}}{B}\left(G + \frac{N}{M} I\right). \tag{70}$$

6.2 Comparison with true axisymmetry

These conserved quantities for quasisymmetric magnetic fields, p_ζ, p_θ, and p_χ, resemble the canonical angular momentum p_ϕ in (9) that is conserved in a truly axisymmetric magnetic field. Note that ϕ differs from the toroidal Boozer angle ζ even in axisymmetry. Eq (9) is the conserved quantity associated with Noether's Theorem for the *full* single-particle Lagrangian, as opposed to the guiding center Lagrangian.

To understand the correspondence in detail, we need to use the result that truly axisymmetric magnetic fields can be represented

$$\boldsymbol{B} = \nabla\phi \times \nabla\psi_p + G(\psi)\nabla\phi, \tag{71}$$

where $\nabla\phi = R^{-1}\boldsymbol{e}_\phi$, and $G(\psi)$ is the same quantity appearing in the Boozer coordinate representation (37). The fact that the same quantity G appears in both (37) and (71) can be seen by applying Ampere's Law to a loop in the toroidal direction. At the same time, from the standard formula for the curl in cylindrical coordinates, we find that in any gauge in which $\boldsymbol{A}$ is axisymmetric (so $\partial A_R/\partial\phi = 0$ and $\partial A_Z/\partial\phi = 0$),

$$\boldsymbol{B} = -\boldsymbol{e}_R \frac{1}{R}\frac{\partial(R A_\phi)}{\partial Z} + \boldsymbol{e}_\phi\left(\frac{\partial A_R}{\partial Z} - \frac{\partial A_Z}{\partial R}\right) + \boldsymbol{e}_Z \frac{1}{R}\frac{\partial(R A_\phi)}{\partial R}. \tag{72}$$

Comparing this expression with (71), we find the R and Z components imply $RA_\phi = -\psi_p+$ constant. Thus the $\boldsymbol{A}$ term in (9) is identical to the ψ_p term in (68) (up to an unimportant constant.)

The correspondence between the two conservation laws is still not obviously complete, since (9) involves v_ϕ whereas (68) involves $v_{||}$. This difference can be understood in terms of gyroaveraging, as follows. Squaring (71) yields $B^2 = (|\nabla\psi_p|^2 + G^2)/R^2$. Applying the result to (71)$\times\nabla\psi_p$, one finds

$$\boldsymbol{B}\times\nabla\psi_p = G\boldsymbol{B} - R^2B^2\nabla\phi. \tag{73}$$

Therefore, the velocity term in (9) can be written

$$Rm\boldsymbol{v}\cdot\boldsymbol{e}_\phi = R^2m\boldsymbol{v}\cdot\nabla\phi = \frac{m}{B}\boldsymbol{v}\cdot(G\boldsymbol{b} - \boldsymbol{b}\times\nabla\psi_p). \tag{74}$$

The $\boldsymbol{v}\cdot\boldsymbol{b}$ term in (74) gives the $v_{||}$ term in (68), while the $\boldsymbol{v}\cdot\boldsymbol{b}\times\nabla\psi_p$ term in (74) averages to zero over the Larmor gyration. Thus, (68) corresponds precisely to the gyroaverage of (9) (up to a constant.) This correspondence make sense, since the former applies to the gyrocenter while the latter applies to the exact particle position.

6.3 Tamm's theorem, revisited

Using the conserved quantities (68)-(70), the argument of Tamm's theorem from (9)-(10) follows in quasisymmetric fields just as it did for axisymmetric fields. Each of the possible conserved quantities (68)-(70) has a term involving $q\psi$ and/or $q\psi_p$, which can be estimated as $\sim qBL^2$, where B is a typical field strength and L is a typical length scale of the field. Also, each possible conserved quantity has a term involving G or I, each of which is $\sim BL$. The ratio of terms is, for the quasi-axisymmetric case of p_ζ,

$$\frac{mv_{||}G/B}{q\psi_p} \sim \frac{mvBL/B}{qBL^2} \sim \frac{v}{\Omega L} \sim \frac{\rho}{L} = \rho_*, \tag{75}$$

where ρ_* is typically $\ll 1$ for magnetized laboratory plasmas. In the same way, the ratio of terms is ρ_* as well for p_θ and p_χ. Therefore, to a good approximation, particles in quasisymmetric fields are confined to surfaces of constant ψ, or equivalently, to surfaces of constant ϕ_p. Even if the small $v_{||}$ terms are not neglected, these terms are bounded due to energy conservation. Thus, despite the complexity of the guiding-center drifts in nonuniform magnetic fields, particle trajectories in quasisymmetric fields are confined close to flux surfaces.

6.4 Direct calculation of conservation law

The conservation of canonical angular momentum in quasisymmetry can also be shown in a direct way, without reference to a Lagrangian. Such a proof for the case of general helicity is given in the appendix of [8] and in appendix A of [28]. Here we consider just the quasi-axisymmetry case for simplicity. Our goal is to show that (68) is constant along the guiding center trajectory, using $\partial B/\partial\zeta = 0$ but not assuming true axisymmetry. More precisely, the equation we wish to derive is

$$\left(v_{||}\boldsymbol{b} + \boldsymbol{v}_d\right)\cdot\nabla\left(-q\psi_p + \frac{mv_{||}G}{B}\right) = 0, \tag{76}$$

where the gradient is taken at fixed μ and total energy

$$W = \frac{mv_{||}^2}{2} + \mu B + \frac{q}{m}\Phi, \tag{77}$$

where $\boldsymbol{v}_d$ is the cross-field drift. We use the following convenient expression for this drift:

$$\hat{\boldsymbol{v}}_d = \frac{mv_{||}}{qB}\nabla \times (v_{||}\boldsymbol{b}), \tag{78}$$

where again the gradient is performed at fixed μ and W. Evaluating the gradient in (78) using

$$\nabla \times \boldsymbol{b} = \boldsymbol{b} \times [(\nabla \times \boldsymbol{b}) \times \boldsymbol{b}] + \boldsymbol{b}\boldsymbol{b} \cdot \nabla \boldsymbol{b} = \boldsymbol{b} \times (\boldsymbol{b} \cdot \nabla \boldsymbol{b}) + \boldsymbol{b}\boldsymbol{b} \cdot \nabla \boldsymbol{b} \tag{79}$$

and

$$0 = \nabla W = mv_{||}\nabla v_{||} + \mu \nabla B + \frac{q}{m}\nabla \Phi, \tag{80}$$

we find

$$\hat{\boldsymbol{v}}_d = \frac{-\nabla \Phi \times \boldsymbol{B}}{B^2} + \frac{mv_\perp^2}{2qB^3}\boldsymbol{B} \times \nabla B + \frac{mv_{||}^2}{qB}\boldsymbol{b} \times (\boldsymbol{b} \cdot \nabla \boldsymbol{b}) + \frac{mv_{||}^2}{qB}(\boldsymbol{b} \cdot \nabla \times \boldsymbol{b})\boldsymbol{b}. \tag{81}$$

If the kinetic energy were held fixed rather than W, the same result would be obtained except without the first right-hand-side term. Here, we recognize the $\boldsymbol{E} \times \boldsymbol{B}$ drift, ∇B drift, and curvature drift. The last term is unphysical (though there is a physical drift of the same form with $v_{||}^2 \to v_\perp^2/2$) and it would be compensated by a factor of $B/B_{||}^*$ used in a more accurate expression for the drifts. However, since this term is parallel to the much larger 'streaming' term $v_{||}\boldsymbol{b}$ and does not move particles off a field line, it is of no consequence. Hence, we will take $\boldsymbol{v}_d = \hat{\boldsymbol{v}}_d$ going forward.

It can be seen that there are four terms in (76). One term vanishes by itself:

$$v_{||}\boldsymbol{b} \cdot \nabla(-q\psi_p) = 0. \tag{82}$$

Another term is

$$\begin{aligned}
\boldsymbol{v}_d \cdot \nabla(-q\psi_p) =& -q\frac{mv_{||}}{qB}\nabla \times (v_{||}\boldsymbol{b}) \cdot \nabla \psi_p \\
=& -\frac{mv_{||}\iota}{B}\nabla \times \left[\frac{v_{||}}{B}(\beta \nabla \psi + I \nabla \theta + G \nabla \zeta)\right] \cdot \nabla \psi \\
=& -\frac{mv_{||}\iota}{B}\left[\nabla\left(\frac{v_{||}\beta}{B}\right) \times \nabla \psi \cdot \nabla \psi + \nabla\left(\frac{v_{||}I}{B}\right) \times \nabla \theta \cdot \nabla \psi + \nabla\left(\frac{v_{||}G}{B}\right) \times \nabla \zeta \cdot \nabla \psi\right] \\
=& -\frac{mv_{||}\iota}{B}\left[(\nabla \zeta \times \nabla \theta \cdot \nabla \psi) I \frac{\partial}{\partial \zeta}\left(\frac{v_{||}}{B}\right) + (\nabla \theta \times \nabla \zeta \cdot \nabla \psi) G \frac{\partial}{\partial \theta}\left(\frac{v_{||}}{B}\right)\right]
\end{aligned} \tag{83}$$

where we have applied (37) and (78). In the last line, the quantity $\partial(v_{||}/B)/\partial \zeta$ (at fixed W and μ) vanishes in quasi-axisymmetry, since $\partial B/\partial \zeta = 0$, and since $v_{||}$ depends on position only through B in light of (77). Then noting (35), (83) reduces to

$$\begin{aligned}
\boldsymbol{v}_d \cdot \nabla(-q\psi_p) =& -\frac{mv_{||}\iota}{B}(\nabla \theta \times \nabla \zeta \cdot \nabla \psi) G \frac{\partial}{\partial \theta}\left(\frac{v_{||}}{B}\right) \\
=& -\frac{mv_{||}}{B}(\boldsymbol{B} \cdot \nabla \theta)\frac{\partial}{\partial \theta}\left(\frac{v_{||}G}{B}\right) = -v_{||}\boldsymbol{b} \cdot \nabla\left(\frac{mv_{||}G}{B}\right).
\end{aligned} \tag{84}$$

The last term is equal and opposite to one of the other terms in (76). Thus, we have shown

$$(v_{||}\boldsymbol{b} + \boldsymbol{v}_d) \cdot \nabla\left(-q\psi_p + \frac{mv_{||}G}{B}\right) = \boldsymbol{v}_d \cdot \nabla\left(\frac{mv_{||}G}{B}\right). \tag{85}$$

Since the remaining term on the right-hand side is one order smaller in ρ_* than (83), we could stop at this point, content that the conservation law holds within the accuracy to which we know $\boldsymbol{v}_d$. In fact, it can be shown that even the remaining term in (85) vanishes, which can be done as follows. Applying (78),

$$\begin{aligned}
\boldsymbol{v}_d \cdot \nabla \left(\frac{m v_{||} G}{B} \right) =& \frac{m^2 v_{||}}{qB} \nabla \times \left(\frac{v_{||}}{B} \boldsymbol{B} \right) \cdot \nabla \left(\frac{v_{||} G}{B} \right) \\
=& \frac{m^2 v_{||}}{qB} \left[\left(\nabla \frac{v_{||}}{B} \right) \times \boldsymbol{B} + \frac{v_{||}}{B} \nabla \times \boldsymbol{B} \right] \cdot \left[G \nabla \frac{v_{||}}{B} + \frac{v_{||}}{B} \frac{dG}{d\psi} \nabla \psi \right] \\
=& \frac{m^2 v_{||}}{qB} \left[\left(\nabla \frac{v_{||}}{B} \right) \times \boldsymbol{B} \cdot \left(\frac{v_{||}}{B} \frac{dG}{d\psi} \nabla \psi \right) + \frac{v_{||}}{B} \nabla \times \boldsymbol{B} \cdot \left(G \nabla \frac{v_{||}}{B} \right) \right] \\
=& \frac{m^2 v_{||}}{qB} \left[\frac{\partial}{\partial \theta} \left(\frac{v_{||}}{B} \right) \right] \frac{v_{||}}{B} \left[\frac{dG}{d\psi} \nabla \theta \times \boldsymbol{B} \cdot \nabla \psi + G \nabla \times \boldsymbol{B} \cdot \nabla \theta \right] \\
=& \frac{m^2 v_{||}}{qB} \left[\frac{\partial}{\partial \theta} \left(\frac{v_{||}}{B} \right) \right] \frac{v_{||}}{B} \left\{ \frac{dG}{d\psi} G \nabla \theta \times \nabla \zeta \cdot \nabla \psi + G \nabla \times \boldsymbol{B} \cdot \nabla \theta \right\}. \qquad (86)
\end{aligned}$$

In these steps, we have used the MHD equilibrium relation $\nabla \times \boldsymbol{B} \cdot \nabla \psi = 0$. In the last term of (86) we substitute (40) and note that $\partial \beta / \partial \zeta = 0$ due to the analysis of section 3.3. Therefore, the terms in curly braces in the last line of (86) sum to 0, and so the right-hand side of (85) vanishes, proving the conservation equation (76).

The direct proof of the conserved helical momentum in the case of helical symmetry is exactly analogous.

7 Neoclassical transport

Another perspective on quasisymmetry comes from considering the calculation of neoclassical transport (radial fluxes and parallel flows, including the parallel current.) We will see that in a quasisymmetric magnetic field, this calculation yields identical results to the calculation in a tokamak, up to a few substitutions. Consequently, the large $1/\nu$ transport regime which is problematic for general stellarators is absent.

Neoclassical transport is calculated by solving the drift-kinetic equation

$$v_{||} \boldsymbol{b} \cdot \nabla f_1 + (\boldsymbol{v}_d \cdot \nabla \psi) \frac{\partial f_M}{\partial \psi} = C[f_1] \qquad (87)$$

where

$$f_M = n(\psi) \left[\frac{m}{2\pi T(\psi)} \right]^{3/2} \exp \left(-\frac{m v^2}{2T(\psi)} \right) \qquad (88)$$

is the leading-order Maxwellian flux function, f_1 is the correction to the Maxwellian, C is the linearized collision operator, and gradients are performed at fixed μ and W. The potential Φ can be taken to be a flux function to leading order, so v is constant on magnetic surfaces. Once the drift-kinetic equation is solved for f_1, the particle fluxes are computed from

$$\Gamma = \left\langle \int d^3 v \, f_1 \boldsymbol{v}_d \cdot \nabla \psi \right\rangle, \qquad (89)$$

where $\langle \ldots \rangle$ denotes a flux surface average, the same integral with an extra factor of energy gives the energy flux, and the parallel flows are computed from

$$n V_{||} = \int d^3 v \, f_1 v_{||}. \qquad (90)$$

Now let us consider each term in (87)-(90) in Boozer coordinates, showing that each depends on the geometry only through ψ and B. This is already clear for $v_{||}$ due to (77). The rest of the streaming term in (87) is

$$
\begin{aligned}
\boldsymbol{b}\cdot\nabla f_1 =& \frac{1}{B}\left[(\boldsymbol{B}\cdot\nabla\theta)\frac{\partial f_1}{\partial\theta}+(\boldsymbol{B}\cdot\nabla\zeta)\frac{\partial f_1}{\partial\zeta}\right]\\
=& \frac{1}{B}(\nabla\psi\cdot\nabla\theta\times\nabla\zeta)\left[\iota\frac{\partial f_1}{\partial\theta}+\frac{\partial f_1}{\partial\zeta}\right]\\
=& \frac{B}{G+\iota I}\left[\iota\frac{\partial f_1}{\partial\theta}+\frac{\partial f_1}{\partial\zeta}\right],
\end{aligned}
\tag{91}
$$

where we have used (38). Thus, the streaming term introduces dependence on geometry only through ψ and B.

Next, the radial drift term, from (83) and (38), is

$$
\boldsymbol{v}_d\cdot\nabla\psi=\frac{mv_{||}}{qB}\frac{B^2}{G+\iota I}\left[G\frac{\partial}{\partial\theta}\left(\frac{v_{||}}{B}\right)-I\frac{\partial}{\partial\zeta}\left(\frac{v_{||}}{B}\right)\right].
\tag{92}
$$

In this form it can be seen that this term depends on position only through ψ and B.

Finally, the linearized collision operator C can be expressed in terms of $v_{||}$ and $v_\perp$, so it introduces no additional geometry dependence. Thus, if B varies on a ψ surface only through θ, i.e. it is independent of ζ as in true axisymmetry, the solution f_1 will have the same property (at fixed μ). Or in the case of helical symmetry, where B varies on a ψ surface only through χ, then f_1 will do so as well.

The last step is to show that the integrals (89)-(90) introduce no additional geometry dependence. For the latter, it is sufficient to observe

$$
\int d^3v = 2\pi\sum_\sigma\int_0^{v^2/(2B)}d\mu\int_0^\infty dv\frac{Bv}{|v_{||}|}
\tag{93}
$$

where $\sigma=\mathrm{sgn}(v_{||})$, which depends on geometry only through ψ and B. To show that (89) does not introduce extra geometry dependence, we note that the flux surface average of any quantity Q is

$$
\langle Q\rangle=\frac{\int_0^{2\pi}d\theta\int_0^{2\pi}d\zeta\frac{Q}{\nabla\psi\cdot\nabla\theta\times\nabla\zeta}}{\int_0^{2\pi}d\theta\int_0^{2\pi}d\zeta\frac{1}{\nabla\psi\cdot\nabla\theta\times\nabla\zeta}}=\frac{\int_0^{2\pi}d\theta\int_0^{2\pi}d\zeta\frac{Q}{B^2}}{\int_0^{2\pi}d\theta\int_0^{2\pi}d\zeta\frac{1}{B^2}}.
\tag{94}
$$

Hence the flux surface average does not introduce geometry dependence beyond ψ and B. Therefore the same is true of the integral in (89).

Thus, in all the steps for calculating neoclassical phenomena, the equations 'know' about the geometry only through the radial coordinate and through B. A symmetry of B therefore causes the calculation to be isomorphic to the calculation in a tokamak.

8 Equivalent conditions for quasisymmetry

There are in fact several equivalent conditions for quasisymmetry. Assuming $\boldsymbol{j}\times\boldsymbol{B}=\nabla p$, and assuming that good flux surfaces exist, the following statements are equivalent:

(a) $B=B(\psi,M\theta-N\zeta)$ for Boozer angles θ and ζ.

(b) $B = B(\psi, M\theta' - N\zeta')$ for angles θ' and ζ' in any other coordinate system for which the Jacobian $(\nabla\psi \cdot \nabla\theta' \times \nabla\zeta')^{-1}$ depends on position only through ψ and B.

(c) $\boldsymbol{B} \times \nabla\psi \cdot \nabla B = F(\psi)\boldsymbol{B} \cdot \nabla B$ for some flux function $F(\psi)$.

(d) $\nabla\psi \times \nabla B \cdot \nabla(\boldsymbol{B} \cdot \nabla B) = 0$.

(e) There exists a vector field $\boldsymbol{u}$ which is nonzero almost everywhere and which satisfies the following conditions:

$$\boldsymbol{u} \cdot \nabla B = 0, \tag{95}$$

$$\nabla \times (\boldsymbol{u} \times \boldsymbol{B}) = 0, \tag{96}$$

$$(\nabla \times \boldsymbol{B}) \times \boldsymbol{u} + \nabla(\boldsymbol{u} \cdot \boldsymbol{B}) = 0. \tag{97}$$

(f) There exists a vector field $\boldsymbol{u}$ which is nonzero almost everywhere and which satisfies the following conditions:

$$\nabla \cdot \boldsymbol{u} = 0, \tag{98}$$

$$\nabla \times (\boldsymbol{u} \times \boldsymbol{B}) = 0, \tag{99}$$

$$(\nabla \times \boldsymbol{B}) \times \boldsymbol{u} + \nabla(\boldsymbol{u} \cdot \boldsymbol{B}) = 0. \tag{100}$$

The only common coordinate system to which condition (b) applies is Hamada coordinates. The equivalence of (a) and (b) is shown in appendix A of [43], appendix B of [16], and page 20 of [36].

8.1 Equivalence of (a) and (c)

To see that (a) implies (c), we can write $B = B(\psi, \chi)$ where $\chi = M\theta - N\zeta$, and compute

$$\begin{aligned}\frac{\boldsymbol{B} \times \nabla\psi \cdot \nabla B}{\boldsymbol{B} \cdot \nabla B} &= \frac{G(\nabla\zeta \times \nabla\psi \cdot \nabla\theta)\frac{\partial B}{\partial \theta} + I(\nabla\theta \times \nabla\psi \cdot \nabla\zeta)\frac{\partial B}{\partial \zeta}}{(\boldsymbol{B} \cdot \nabla\zeta)\frac{\partial B}{\partial \zeta} + (\boldsymbol{B} \cdot \nabla\theta)\frac{\partial B}{\partial \theta}} = \frac{G\frac{\partial B}{\partial \theta} - I\frac{\partial B}{\partial \zeta}}{\frac{\partial B}{\partial \zeta} + \iota\frac{\partial B}{\partial \theta}} \\ &= \frac{MG\frac{\partial B}{\partial \chi} + NI\frac{\partial B}{\partial \chi}}{-N\frac{\partial B}{\partial \chi} + M\iota\frac{\partial B}{\partial \chi}} = \frac{MG + NI}{\iota M - N}.\end{aligned} \tag{101}$$

The final expression is a flux function, and so we can identify it as the function $F(\psi)$ that satisfies (c).

Conversely, suppose (c) is satisfied. Then

$$G\frac{\partial B}{\partial \theta} - I\frac{\partial B}{\partial \zeta} = F(\psi)\left(\frac{\partial B}{\partial \zeta} + \iota\frac{\partial B}{\partial \theta}\right). \tag{102}$$

Introducing the double Fourier series

$$B(\psi, \theta, \zeta) = \sum_{m,n} B_{mn}(\psi)\exp(im\theta - in\zeta), \tag{103}$$

we obtain

$$(mG + nI + nF - \iota mF)B_{mn} = 0. \tag{104}$$

Either B_{mn} or the expression in parentheses must vanish for each (m, n) pair. The mode B_{00} can be nonzero, while if m is nonzero, B_{mn} can be nonzero only if

$$\frac{n}{m} = \frac{\iota F - G}{I + F}, \tag{105}$$

and if n is nonzero, $B_{m,n}$ can be nonzero only if the reciprocal of (105) is satisfied. Thus, the nonzero B_{mn} modes must all have the same ratio of m to n, and so (a) is satisfied. Thus (a) and (c) are equivalent.

8.2 Equivalence of (c) and (d)

To show the equivalence of (c) and (d), we first derive a vector identity by evaluating $[(\nabla\psi \times \nabla B) \times \boldsymbol{B}] \times \boldsymbol{B}$. If the BAC-CAB rule is applied first to the quantities within square brackets,

$$[(\nabla\psi \times \nabla B) \times \boldsymbol{B}] \times \boldsymbol{B} = -(\boldsymbol{B} \cdot \nabla B)\nabla\psi \times \boldsymbol{B}. \tag{106}$$

On the other hand, if the BAC-CAB rule is applied treating the quantity in parentheses as a unit, we obtain

$$[(\nabla\psi \times \nabla B) \times \boldsymbol{B}] \times \boldsymbol{B} = \boldsymbol{B}(\boldsymbol{B} \cdot \nabla\psi \times \nabla B) - (\nabla\psi \times \nabla B)B^2. \tag{107}$$

Equating (106) to (107) and dividing by $\boldsymbol{B} \cdot \nabla B$,

$$\boldsymbol{B} \times \nabla\psi = \boldsymbol{B}\frac{\boldsymbol{B} \times \nabla\psi \cdot \nabla B}{\boldsymbol{B} \cdot \nabla B} - \nabla\psi \times \nabla B \frac{B^2}{\boldsymbol{B} \cdot \nabla B}. \tag{108}$$

We take the divergence, noting $\nabla \cdot (\boldsymbol{B} \times \nabla\psi) = 0$ on the left hand side due to MHD equilibrium, leaving

$$\boldsymbol{B} \cdot \nabla \left(\frac{\boldsymbol{B} \times \nabla\psi \cdot \nabla B}{\boldsymbol{B} \cdot \nabla B} \right) = -\frac{B^2}{(\boldsymbol{B} \cdot \nabla B)^2} \nabla\psi \times \nabla B \cdot \nabla(\boldsymbol{B} \cdot \nabla B). \tag{109}$$

The equivalence of conditions (c) and (d) follows immediately.

8.3 Equivalence of (e) and previous conditions

Let us first prove that (a) implies (e). We will assume here $M \neq 0$; the case $M = 0$ can be handled by reversing the roles of ζ and θ in the argument that follows. We will take the independent variables for partial derivatives to the (ψ, θ, ζ) except where a subscript indicates $\chi = M\theta - N\zeta$ is fixed, so $(\partial B/\partial\zeta)_\chi = 0$. We will show that the choice

$$\boldsymbol{u} = \left(\frac{\partial \boldsymbol{x}}{\partial \zeta} \right)_\chi \tag{110}$$

satisfies (95)-(97). First,

$$\boldsymbol{u} \cdot \nabla B = \left(\frac{\partial \boldsymbol{x}}{\partial \zeta} \right)_\chi \cdot \left[\left(\frac{\partial B}{\partial \psi} \right)_\chi \nabla\psi + \left(\frac{\partial B}{\partial \chi} \right)_\psi \nabla\chi \right] = 0 \tag{111}$$

from (26). Next, applying the dual relations to (110),

$$\boldsymbol{u}\times\boldsymbol{B}=\frac{\nabla\psi\times\nabla\chi}{\nabla\psi\cdot\nabla\chi\times\nabla\zeta}\times\boldsymbol{B}=\frac{\nabla\psi\times\nabla\chi}{\nabla\psi\cdot\nabla\chi\times\nabla\zeta}\times\left(\nabla\psi\times\nabla\theta+\iota\nabla\zeta\times\nabla\psi\right) \tag{112}$$

$$=\frac{M\nabla\psi\times\nabla\theta-N\nabla\psi\times\nabla\zeta}{M\nabla\psi\cdot\nabla\theta\times\nabla\zeta}\times\left(\nabla\psi\times\nabla\theta+\iota\nabla\zeta\times\nabla\psi\right) \tag{113}$$

$$=\frac{N-\iota M}{M}\nabla\psi.$$

Therefore (96) is satisfied.

Finally, to check (97), we first write

$$\boldsymbol{u}\cdot\boldsymbol{B}=\left(\frac{\partial\boldsymbol{x}}{\partial\zeta}\right)_{\chi}\cdot\left(\beta\nabla\psi+I\nabla\theta+G\nabla\zeta\right) \tag{114}$$

$$=\left(\frac{\partial\boldsymbol{x}}{\partial\zeta}\right)_{\chi}\cdot\left(\beta\nabla\psi+\frac{I}{M}\nabla\chi+\frac{NI+MG}{M}\nabla\zeta\right)$$

$$=\frac{NI+MG}{M}.$$

Another quantity we need is

$$(\nabla\times\mathbf{B})\times\mathbf{u}=\left(\nabla\beta\times\nabla\psi+\nabla I\times\nabla\theta+\nabla G\times\nabla\zeta\right)\times\left(\frac{\partial\mathbf{r}}{\partial\zeta}\right)_{\chi} \tag{115}$$

$$=\left(\frac{\partial\beta}{\partial\theta}\nabla\theta\times\nabla\psi+\frac{\partial\beta}{\partial\zeta}\nabla\zeta\times\nabla\psi+\frac{dI}{d\psi}\nabla\psi\times\nabla\theta+\frac{dG}{d\psi}\nabla\psi\times\nabla\zeta\right)\times\frac{\nabla\psi\times\nabla\chi}{\nabla\psi\cdot\nabla\chi\times\nabla\zeta}$$

$$=\left(\frac{\partial\beta}{\partial\theta}\nabla\theta\times\nabla\psi+\frac{\partial\beta}{\partial\zeta}\nabla\zeta\times\nabla\psi+\frac{dI}{d\psi}\nabla\psi\times\nabla\theta+\frac{dG}{d\psi}\nabla\psi\times\nabla\zeta\right)\times\frac{M\nabla\psi\times\nabla\theta-N\nabla\psi\times\nabla\zeta}{M\nabla\psi\cdot\nabla\theta\times\nabla\zeta}$$

$$=\left(\frac{\partial\beta}{\partial\zeta}\nabla\zeta\times\nabla\psi+\frac{dG}{d\psi}\nabla\psi\times\nabla\zeta\right)\times\frac{M\nabla\psi\times\nabla\theta}{M\nabla\psi\cdot\nabla\theta\times\nabla\zeta}$$

$$+\left(\frac{\partial\beta}{\partial\theta}\nabla\theta\times\nabla\psi+\frac{dI}{d\psi}\nabla\psi\times\nabla\theta\right)\times\frac{-N\nabla\psi\times\nabla\zeta}{M\nabla\psi\cdot\nabla\theta\times\nabla\zeta}$$

$$=\frac{1}{M}\left[M\frac{\partial\beta}{\partial\zeta}-M\frac{dG}{d\psi}+N\frac{\partial\beta}{\partial\theta}-N\frac{dI}{d\psi}\right]\nabla\psi.$$

However we know from section 3.3 that $B=B\left(\psi,\chi\right)$ implies $\beta=\beta\left(\psi,\chi\right)$, so in the last line of (115),

$$M\frac{\partial\beta}{\partial\zeta}+N\frac{\partial\beta}{\partial\theta}=M\left(-N\right)\frac{\partial\beta}{\partial\chi}+NM\frac{\partial\beta}{\partial\chi}=0. \tag{116}$$

Combining (114)-(116), the quantity in (97) is

$$(\nabla\times\mathbf{B})\times\mathbf{u}+\nabla\left(\mathbf{u}\cdot\mathbf{B}\right)=\frac{1}{M}\left[-M\frac{dG}{d\psi}-N\frac{dI}{d\psi}\right]\nabla\psi+\nabla\left(\frac{NI+MG}{M}\right)=0. \tag{117}$$

Hence, (a) implies (e).

Conversely, we will now show that (e) implies (c). From (99),

$$\boldsymbol{u}\times\boldsymbol{B}=\nabla\gamma \tag{118}$$

for some γ. The $\boldsymbol{B}$ component gives $\boldsymbol{B}\cdot\nabla\gamma = 0$. On any irrational surface, then γ must be constant. If there is any magnetic shear, γ must be constant on every surface by continuity. We will not consider the case of a rational ι with no magnetic shear, since such a magnetic field would be extremely sensitive to breakup of surfaces by resonant magnetic perturbations. We can therefore assume $\gamma = \gamma(\psi)$. From $\boldsymbol{B}\times$(118),

$$\boldsymbol{u} = \frac{\gamma' \boldsymbol{B}\times\nabla\psi}{B^2} + u_{||}\boldsymbol{B} \tag{119}$$

for some $u_{||}$, where $\gamma' = d\gamma/d\psi$. Eq (95) gives

$$u_{||} = -\frac{\gamma' \boldsymbol{B}\times\nabla\psi\cdot\nabla B}{B^2 \boldsymbol{B}\cdot\nabla B}. \tag{120}$$

Note that γ' must be nonzero except perhaps at isolated values of ψ, or else $\boldsymbol{u}$ would vanish everywhere throughout a finite volume, contrary to the assumption of condition (e). We next consider the $\boldsymbol{B}$ component of (97). The first term vanishes:

$$\boldsymbol{B}\cdot(\nabla\times\boldsymbol{B})\times\boldsymbol{u} = -\mu_0\nabla p\cdot\boldsymbol{u} = 0. \tag{121}$$

We are left with

$$0 = \boldsymbol{B}\cdot\nabla\left(\frac{\gamma' \boldsymbol{B}\cdot\nabla\psi\cdot\nabla B}{\boldsymbol{B}\cdot\nabla B}\right), \tag{122}$$

which is equivalent to (c).

8.4 Equivalence of (e) and (f)

Observe that (99) can be written

$$\boldsymbol{B}\cdot\nabla\boldsymbol{u} - \boldsymbol{u}\cdot\nabla\boldsymbol{B} - (\nabla\cdot\boldsymbol{u})\boldsymbol{B} = 0, \tag{123}$$

while (100) can be written

$$0 = -(\nabla\boldsymbol{B})\cdot\boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{B} + (\nabla\boldsymbol{u})\cdot\boldsymbol{B} + (\nabla\boldsymbol{B})\cdot\boldsymbol{u} = \boldsymbol{u}\cdot\nabla\boldsymbol{B} + (\nabla\boldsymbol{u})\cdot\boldsymbol{B}. \tag{124}$$

Subtracting (123) from (124),

$$0 = 2\boldsymbol{u}\cdot\nabla\boldsymbol{B} + (\nabla\boldsymbol{u})\cdot\boldsymbol{B} - \boldsymbol{B}\cdot\nabla\boldsymbol{u} + (\nabla\cdot\boldsymbol{u})\boldsymbol{B}. \tag{125}$$

The dot product with $\boldsymbol{B}$ then gives

$$0 = 2\boldsymbol{u}\cdot(\nabla\boldsymbol{B})\cdot\boldsymbol{B} + (\nabla\cdot\boldsymbol{u})B^2 = 2B\boldsymbol{u}\cdot\nabla B + (\nabla\cdot\boldsymbol{u})B^2. \tag{126}$$

From this last expression and $B \neq 0$, it is clear that (e) and (f) are equivalent.

9 Considerations near the magnetic axis

Several insights into quasisymmetric fields can be gained by considering the region near the magnetic axis. The magnetic axis is the innermost field line in the confinement region, a sort of degenerate flux surface that has zero area.

First, observe that $\nabla p = 0$ on the axis. Therefore (47) can be written

$$\nabla_\perp B = \kappa B \boldsymbol{n}, \tag{127}$$

where $\nabla_\perp = \nabla - \boldsymbol{b}\boldsymbol{b}\cdot\nabla$ is the projection of the gradient into the plane perpendicular to $\boldsymbol{B}$, $\kappa = |\boldsymbol{b}\cdot\nabla\boldsymbol{b}|$ is the curvature of the field lines, and $\boldsymbol{n} = \kappa^{-1}\boldsymbol{b}\cdot\nabla\boldsymbol{b}$ is a unit vector in the direction of curvature. If the magnetic axis is to close on itself, κ must be nonzero for some of its length. Hence the right-hand side is nonzero in these regions, so $\nabla_\perp B$ is nonzero. At any toroidal angle corresponding to these points on the axis, the $\propto \cos\theta$ Fourier mode of B will therefore have a nonzero amplitude. This means that only $M = 1$ quasisymmetry is possible near the axis. Quasi-poloidal symmetry, which corresponds to $M = 0$, is not allowed. It is still possible to find magnetic fields with rough quasi-poloidal symmetry [44, 45], but the level of symmetry-breaking near the axis will necessarily be larger than for quasi-axisymmetry or quasi-helical symmetry.

On the magnetic axis, B must be independent of θ. Therefore for both quasi-axisymmetry and quasi-helical symmetry, B on axis must also be independent of ζ. This result, that B must be constant along the axis, is an easy-to-state necessary condition for quasisymmetry.

Another fact that can be learned from near-axis analysis, in the usual case where there is no current on the axis, is that the axis must have nonzero torsion in some region, meaning the axis is not confined to a plane. Otherwise, ι would vanish on axis, which is not acceptable. This result is proved in section 5.3 of [46]. Stated another way, in order to have nonzero ι on axis when the axis has zero torsion everywhere, the flux surfaces must have rotating (ζ-dependent) elongation that breaks quasisymmetry.

Much more can be said about quasisymmetry in the vicinity of the magnetic axis. Further details can be found in [12, 13, 46, 34, 47].

10 How accurately has quasisymmetry been achieved?

In principle, quasisymmetry can be achieved exactly, for axisymmetric fields are quasisymmetric. Whether or not quasisymmetry can be achieved exactly without axisymmetry is unknown. Several magnetic configurations have been found numerically that are quasisymmetric in some approximation, including HSX, NCSX, ESTELL, CFQS, and others. However, all of these configurations in fact have significant deviations from quasisymmetry. For example, $B(\theta,\zeta)$ for the outer surfaces of HSX and NCSX is shown in the top panels of figure 5. While there is a general tendency for the B contours to close helically or toroidally respectively, the contours that do close this way are not straight, and many contours close with different topologies (such as those near local minina or maxima of B). Contours of B look somewhat more symmetric on interior surfaces, but the deviations from symmetry remain apparent throughout the volume. These departures from quasisymmetry can be conceptually divided into two sources: ripple due to the choice of discrete coils used, and breaking of symmetry that would remain even if a continuous sheet current were used to produce the field. Ripple from the 48 modular coils accounts for much of the symmetry-breaking in HSX, apparent in the 3D view in figure 5. Nonetheless, significant symmetry-breaking has always been observed even in "fixed boundary" numerical solutions, i.e. solutions of $\boldsymbol{j}\times\boldsymbol{B} = \nabla p$ within a toroidal domain that exist regardless of the feasibility of supporting the configuration with realistic electromagnetic coils.

It is known that quasisymmetry can be achieved more accurately in domains with high aspect ratio A, defined as some averaged major radius divided by some averaged minor radius. For example, Garren and Boozer showed [12, 13] that the equations of MHD equilibrium and quasisymmetry are fewer than the available degrees of freedom through the first two orders in an expansion in

$A \gg 1$, but the equations outnumber the degrees of freedom beginning at the third order. Put another way, it should at least be possible to make symmetry-breaking decrease $\propto 1/A^3$ for high-A configurations. As a result, if one wishes to have symmetry-breaking no larger than some threshold value, quasisymmetry to this degree can be achieved in configurations with sufficiently large A. This effect is shown in the bottom panels of figure figure 5. Here, numerical fixed-boundary configurations are shown [48] for which the symmetry-breaking in the $B(\theta, \zeta)$ contours of the boundary is essentially invisible by eye. The symmetry-breaking mode amplitudes in these configurations are all smaller than the Earth's magnetic field of ~ 0.5 Tesla. The aspect ratio of these configurations (80) is likely too large to be of experimental interest. Important practical questions are how accurately quasisymmetry can be achieved (without axisymmetry) at lower A, and what level of symmetry-breaking is tolerable.

11 Open questions

A number of open questions remain in the theory of quasisymmetry:

- Does any example exist of a nonaxisymmetric field with exact quasisymmetry throughout a finite volume? (Refs [12, 13] are not a disproof, for these works merely show the system of equations is overdetermined; overdetermined systems can still have solutions.)
- Should we actually be striving for quasisymmetry, or is the weaker condition of omnigenity sufficient?
- Without using optimization, is there a way to directly construct magnetic configurations that have exact quasisymmetry at a single magnetic surface off of the magnetic axis?
- Flows as fast as the ion thermal speed do not seem to be allowed in quasisymmetric fields [4, 5], yet quasisymmetry does seem to allow faster flows than those allowed in a non-quasisymmetric stellarator [3]. Can the allowed magnitude of the flow be clarified theoretically?
- Are there analogies to quasisymmetry in other physical systems?
- To what accuracy should we strive to achieve quasisymmetry? How close is close enough?

References

[1] P Helander and A N Simakov. *Phys. Rev. Lett.*, 101:145003, 2008.

[2] H Sugama, T H Watanabe, M Nunami, and S Nishimura. *Plasma Phys. Controlled Fusion*, 53:024004, 2011.

[3] P Helander. *Phys. Plasmas*, 14:104501, 2007.

[4] A N Simakov and P Helander. *Plasma Phys. Controlled Fusion*, 53:024005, 2011.

[5] H Sugama, T H Watanabe, M Nunami, and S Nishimura. *Phys. Plasmas*, 18:082505, 2011.

[6] Pytte and A H Boozer. *Phys. Fluids*, 24:88, 1981.

[7] A H Boozer. *Phys. Fluids*, 26:496, 1983.

[8] J Nührenberg and R Zille. *Phys. Lett. A*, 129:113, 1988.

[9] F S B Anderson, A F Almagri, D T Anderson, P G Matthews, J N Talmadge, and J L Shohet. The helically symmetric experiment (HSX): Goals, design, and status. *Fusion Tech.*, 27:273, 1995.

[10] J Nührenberg, W Lotz, and S Gori. Quasi-axisymmetric tokamaks. In *Proceedings of the Joint Varenna-Lausanne International Workshop on Theory of Fusion Plasmas (Bologna: Editrice Compositori)*, page 3, 1994.

[11] P R Garabedian. Stellarators with the magnetic symmetry of a tokamak. *Phys. Plasmas*, 3:2483, 1996.

[12] D A Garren and A H Boozer. *Phys. Fluids B*, 3:2805, 1991.

[13] D A Garren and A H Boozer. *Phys. Fluids B*, 3:2822, 1991.

[14] J R Cary and S G Shasharina. Helical plasma confinement devices with good confinement properties. *Phys. Rev. Lett.*, 78:674, 1997.

[15] J R Cary and S G Shasharina. Omnigenity and quasihelicity in helical plasma confinement systems. *Phys. Plasmas*, 4:3323, 1997.

[16] M Landreman and P J Catto. Omnigenity as generalized quasisymmetry. *Phys. Plasmas*, 19:056103, 2012.

[17] J. M. Canik, D. T. Anderson, F. S. B. Anderson, K. M. Likin, J. N. Talmadge, and K. Zhai. Experimental demonstration of improved neoclassical transport with quasihelical symmetry. *Phys. Rev. Lett.*, 98:085002, 2007.

[18] M C Zarnstorff, L A Berry, A Brooks, E Fredrickson, G-Y Fu, S Hirshman, S Hudson, L-P Ku, E Lazarus, D Mikkelsen, D Monticello, G H Neilson, N Pomphrey, A Reiman, D Spong, D Strickler, A Boozer, W A Cooper, R Goldston, R Hatcher, M Isaev, C Kessel, J Lewandowski, J F Lyon, P Merkel, H Mynick, B E Nelson, C Nuehrenberg, M Redi, W Reiersen, P Rutherford, R Sanchez, J Schmidt, and R B White. *Plasma Phys. Controlled Fusion*, 43, 2001.

[19] H Liu, A Shimizu, M Isobe, S Okamura, S Nishimura, C Suzuki, Y Xu, X Zhang, B Liu, J Huang, X Wang, H Liu, C Tang, D Yin, Y Wan, and the CFQS team. Magnetic configuration and modular coil design for the chinese first quasi-axisymmetric stellarator. 13:3405067, 2018.

[20] A Shimizu, H Liu, M Isobe, S Okamura, S Nishimura, C Suzuki, Y Xu, X Zhang, B Liu, J Huang, X Wang, H Liu, C Tang, and the CFQS team. Configuration property of the chinese first quasi-axisymmetric stellarator. 13:3403123, 2018.

[21] P R Garabedian. Three-dimensional analysis of tokamaks and stellarators. 105:13716, 2008.

[22] P R Garabedian and G B McFadden. Design of the DEMO fusion reactor following ITER. *J. Res. Natl. Inst. Stand. Technol.*, 114:229, 2009.

[23] L P Ku and A H Boozer. Modular coils and plasma configurations for quasi-axisymmetric stellarators. *Nucl. Fusion*, 50:125005, 2010.

[24] L P Ku and A H Boozer. New classes of quasi-helically symmetric stellarators. *Nucl. Fusion*, 51:013004, 2011.

[25] M Drevlak, F Brochard, P Helander, J Kisslinger, M Mikhailov, C Nührenberg, J Nührenberg, and Y Turkin. ESTELL: A quasi-toroidally symmetric stellarator. *Contrib. Plasma Phys.*, 53:459, 2013.

[26] S. A. Henneberg, M. Drevlak, C. Nührenberg, C. D. Beidler, Y. Turkin, J. Loizu, and P. Helander. Properties of a new quasi-axisymmetric configuration. *Nucl. Fusion*, 59:026014, 2019.

[27] M Drevlak, C D Beidler, J Geiger, P Helander, and Y Turkin. Optimisation of stellarator equilibria with ROSE. *Nucl. Fusion*, 59:016010, 2019.

[28] M Landreman and P J Catto. Effects of the radial electric field in a quasisymmetric stellarator. *Plasma Phys. Controlled Fusion*, 53:015004, 2011.

[29] I Calvo, F I Parra, J L Velasco, and J A Alonso. Stellarators close to quasisymmetry. *Plasma Phys. Controlled Fusion*, 55:125014, 2013.

[30] I Calvo, F I Parra, J A Alonso, and J L Velasco. Optimizing stellarators for large flows. *Plasma Phys. Controlled Fusion*, 56:094003, 2014.

[31] I Calvo, F I Parra, J L Velasco, and J A Alonso. Flow damping in stellarators close to quasisymmetry. *Plasma Phys. Controlled Fusion*, 57:014014, 2015.

[32] J W Burby and H Qin. Toroidal precession as a geometric phase. *Phys. Plasmas*, 20:012511, 2013.

[33] G G Plunk and P Helander. Quasi-axisymmetric magnetic fields: weakly non-axisymmetric case in a vacuum. *J. Plasma Phys.*, 84:905840205, 2018.

[34] M Landreman, W Sengupta, and G G Plunk. Direct construction of optimized stellarator shapes. II. Numerical quasisymmetric solutions. *J. Plasma Phys.*, 85:905850103, 2018.

[35] A H Boozer. Quasi-helical symmetry in stellarators. *Plasma Phys. Controlled Fusion*, 37:A103, 1995.

[36] P Helander. *Rep. Prog. Phys.*, 77:087001, 2014.

[37] H Goldstein, C Poole, and J Safko. *Classical mechanics, 3rd ed.* Addison Wesley, 2001.

[38] J D Jackson. *Classical electrodynamics, 3rd ed.* Wiley, 1999.

[39] W. D. D'haeseleer, W. N. G. Hitchon, J. D. Callen, , and J. L. Shohet. *Flux coordinates and magnetic field structure.* Springer-Verlag, 1991.

[40] G B Arfken and H J Weber. *Mathematical Methods for Physicists.* Academic Press, 2005.

[41] R G Littlejohn. *J. Plasma Phys.*, 29:111, 1983.

[42] P Helander and D J Sigmar. *Collisional Transport in Magnetized Plasmas.* Cambridge University Press, 2002.

[43] H Sugama and S Nishimura. *Phys. Plasmas*, 9:4637, 2002.

[44] D A Spong, S P Hirshman, L A Berry, J F Lyon, R H Fowler, D J Strickler, M J Cole, B N Nelson, D E Williamson, A S Ware, D Alban, R Sanchez, G Y Fu, D A Monticello, W H Miner, and P M Valanju. Physics issues of compact drift optimized stellarators. *Nucl. Fusion*, 41:711, 2001.

[45] D J Strickler, S P Hirshman, D A Spong, M J Cole, J F Lyon, B E Nelson, and D E Williamson. Development of a robust quasi-poloidal compact stellarator. *Fusion Sci. Tech.*, 45:15, 2004.

[46] M Landreman and W Sengupta. Direct construction of optimized stellarator shapes. I. Theory in cylindrical coordinates. *J. Plasma Phys.*, 84:905840616, 2018.

[47] M Landreman. Optimized quasisymmetric stellarators are consistent with the Garren-Boozer construction. *Plasma Phys. Controlled Fusion*, 61:075001, 2019.

[48] M Landreman and W Sengupta. Analytic description and construction of quasisymmetric stellarators with triangularity. *In preparation*, 2019.

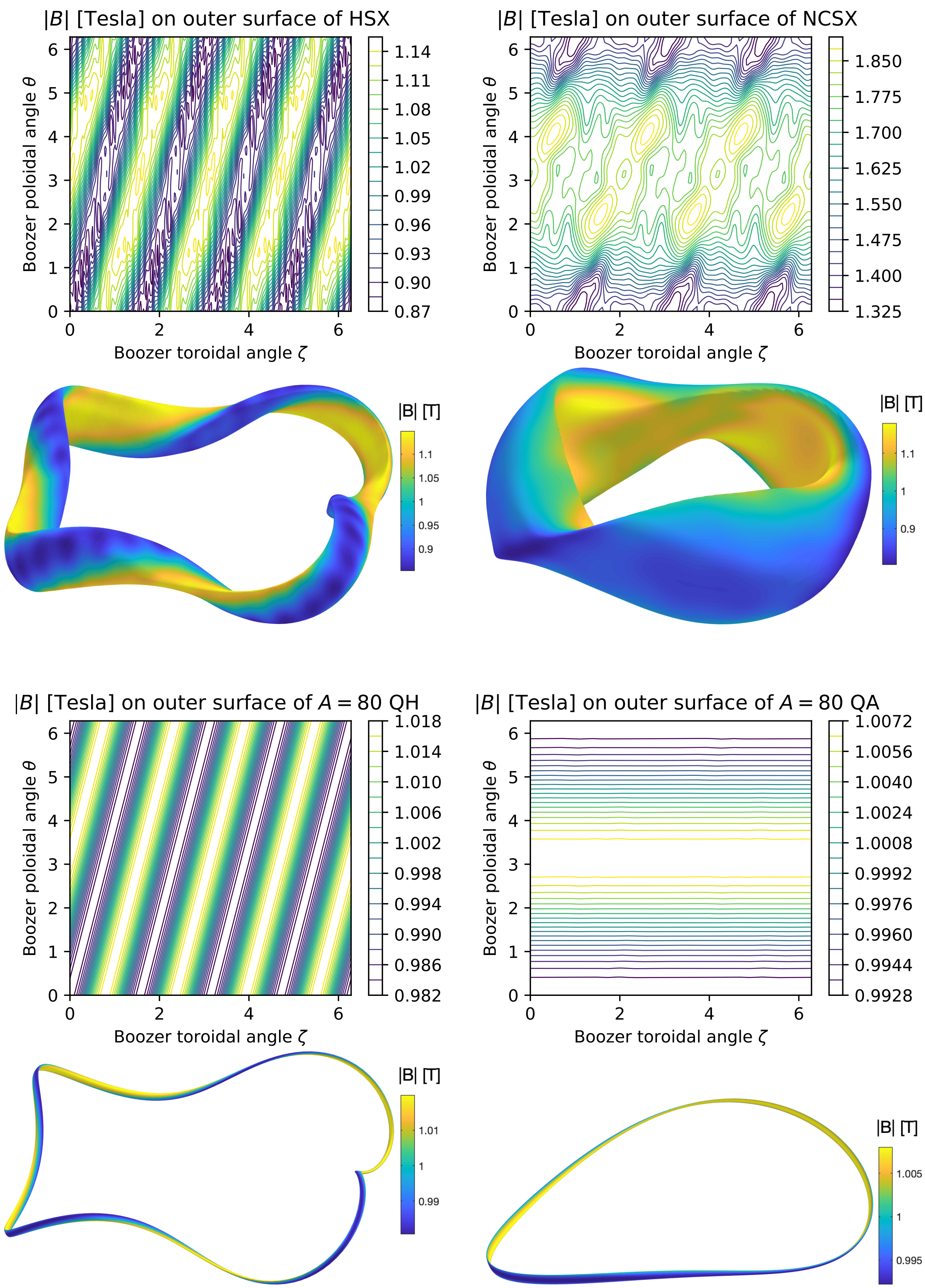

Figure 5: Significant deviations from quasisymmetry exist in experimental designs to date with real coils, such as HSX and NCSX. At sufficiently high aspect ratio, and if errors associated with realistic coils are not included, quasisymmetry can be achieved to arbitrarily high accuracy, as in the aspect ratio 80 configurations shown in the bottom panels, generated by the method of [48].