

Theory of plasma confinement in non-axisymmetric magnetic fields

Per Helander

Max Planck Institute for Plasma Physics,
Wendelsteinstraße 1, 17491 Greifswald, Germany

To be published in Reports on Progress in Physics

“In the history of controlled thermonuclear fusion, there have been no ideas comparable in beauty and conceptual significance with that of the stellarator.”

(V.D. Shafranov, 1980)

“There is only one really original idea in the whole of magnetic fusion and it is the Stellarator.”

(J.B. Taylor, 2007)

1 Introduction

Stellarator research began on a definite date, 25 March 1951. It was triggered by an article on the front page of the New York Times reporting that the Argentinian dictator Juan Péron had announced that his country had achieved energy production by controlled nuclear fusion. The claims were quickly dismissed by most leading physicists of the day, but Lyman Spitzer Jr, head of the Princeton Astrophysics Department, started thinking seriously about the problem of how to confine an ionised plasma, and summarised his ideas in a classified report “A Proposed Stellarator” on 23 July [1]. It is a visionary document that anticipates the development of fusion research in several important ways. Of course, the problem turned out to be far more difficult than originally thought, but most of Spitzer’s key ideas have stood the test of time.

In the intervening years, many concepts for achieving fusion energy production have been proposed and many types of experimental devices have been built, but only a few have proved to be practical. The most prominent one is the tokamak, also invented in 1951 by Tamm and Sakharov but independently conceived also by several others. Since the 1970’s most of the world fusion energy effort has been on tokamaks, but it has also become increasingly clear that the stellarator has important advantages. In both types of devices, the fusion plasma is confined by a twisted toroidal magnetic field, but the twist is produced in different ways. In tokamaks, it is created by a toroidal plasma current, and in stellarators by non-axisymmetric coils. Tokamaks are axisymmetric and therefore automatically confine all collisionless particle orbits, which results in relatively good plasma confinement, but the necessary current is usually induced by transformer action, making these devices vulnerable to current-driven instabilities and difficult to operate in steady state. There are other ways of driving the plasma current, but these tend to be expensive and restrict the achievable plasma parameters.

Stellarators, on the other hand, are necessarily non-axisymmetric and do not, in general, confine the plasma as well. But they are inherently able to operate in steady state, and can have confinement properties comparable to that of tokamaks, if the magnetic field is properly optimised.

It is not easy for a newcomer to learn stellarator physics, because most of the results in the field are scattered over six decades of journal and conference publications. Only one book [2] and a few review articles [3, 4, 5, 6, 7, 8, 9] have been published on the subject, but these either do not describe the most recent developments or do not go into mathematical details. The aim of the present paper is to give a mathematical overview of the theory of stellarator magnetic fields and their simplest physical properties. It is not meant to be comprehensive, but rather to serve as an introduction to scientists entering the field. For this reason, several elementary concepts are explained and mathematical manipulations are given in detail. No attempt has been made to review all the literature, and the attention has been restricted to basic theory rather than experimental results. And within the realm of theory, the emphasis lies on equilibrium and confinement, whereas the theory of macroscopic and microscopic stability is not treated. On the other hand, certain fundamental questions are ignored, particularly those that have not yet been fully resolved, such as the existence of magnetic surfaces in fields without a continuous symmetry.

The reader is assumed to have a basic knowledge of plasma physics, including its underpinnings such as electrodynamics, analytical mechanics, hydrodynamics and kinetic theory, but need not be a specialist in the field of fusion physics. The material is divided into three chapters in a logical progression. The first one describes magnetohydrodynamic issues, such as the structure of the magnetic field, the plasma current, magnetic coordinates, etc. The following chapter treats particle motion and the problem how the magnetic field may be optimised for good orbit confinement. Finally, the concluding chapter gives an introduction into the kinetic theory of equilibrium and transport in non-axisymmetric plasmas, in particular the basics of neoclassical theory.

2 Magnetic field

The first step in constructing a theory of plasma confinement in stellarators must necessarily be a mathematical analysis of the confining, non-axisymmetric magnetic field. What is required of such a field to confine a plasma, and how is it most conveniently described mathematically? In

the present section, the focus lies on the magnetic field itself and on the conditions imposed upon it by magnetohydrodynamics (MHD).

2.1 Force balance

The aim of a stellarator or a tokamak is to confine a plasma with a certain pressure p by means of a magnetic field \mathbf{B} . This is possible since the pressure force of the plasma can be balanced by a magnetic force produced by a plasma current. If the plasma current is denoted by \mathbf{J} , it is related to the magnetic field by Ampère’s law,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad (1)$$

in steady state. Most of the magnetic field is usually created by external coils, and this “vacuum field” is modified by currents within the plasma. The amount of plasma current necessary for confinement is determined from the MHD equation of motion,

$$\rho \left(\frac{\partial \mathbf{V}}{\partial t} + \mathbf{V} \cdot \nabla \mathbf{V} \right) = \mathbf{J} \times \mathbf{B} - \nabla p - \nabla \cdot \boldsymbol{\pi}, \quad (2)$$

where ρ denotes the density, \mathbf{V} the plasma flow velocity, and the viscosity $\boldsymbol{\pi}$ is much smaller than the pressure p if the plasma is in a state close to local thermodynamic equilibrium. In a steady state (mechanical equilibrium) without flows,

$$\frac{\partial}{\partial t} = 0, \quad \mathbf{V} = 0,$$

the equation of motion thus becomes

$$\mathbf{J} \times \mathbf{B} = \nabla p. \quad (3)$$

It follows from this equation that the vectors \mathbf{B} and \mathbf{J} lie in surfaces of constant pressure,

$$\mathbf{B} \cdot \nabla p = 0, \quad \mathbf{J} \cdot \nabla p = 0.$$

According to a famous theorem in topology, the Poincaré index (or “hairy ball”) theorem, a compact surface which is everywhere tangential to a non-vanishing vector field free of singularities must have the topology of a torus (see e.g. Thurston [10]). The theorem implies, for instance, that you cannot comb the hair on a coconut in a uniform way, and that there is always one place on Earth where the wind is not blowing horizontally. In the present context, it says that surfaces of constant pressure in a magnetically confined

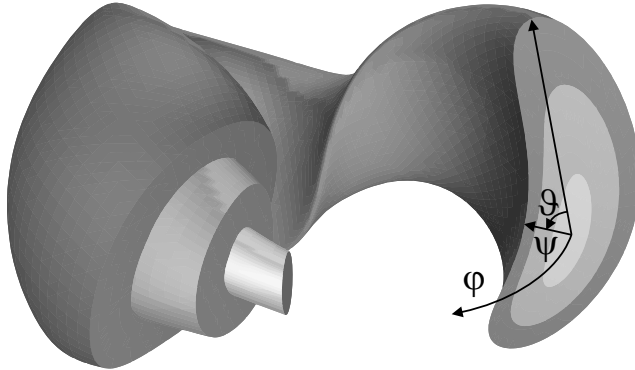


Figure 1: *Toroidally nested surfaces of constant pressure.*

plasma must have toroidal topology. On each such surface, the stream lines of \mathbf{B} and \mathbf{J} are wound around the torus, intersecting each other at a finite angle according to Eq. (3).

The inertial term $\rho \mathbf{V} \cdot \nabla \mathbf{V}$ on the left-hand-side of Eq. (2) contains the centrifugal and Coriolis forces and is comparable to the pressure gradient only if the flow velocity is of the order of the ion thermal speed $v_T \sim (p/\rho)^{1/2}$. It will be shown from kinetic theory below that such a large flow velocity is unattainable in most magnetic field configurations as long as the ion gyro-radius is much smaller than the equilibrium scale length.

2.2 Magnetic coordinates

The most desirable situation for plasma confinement is when the surfaces of constant pressure are *nested*, i.e., lie inside each other. The innermost surface is then just a line, the so-called magnetic axis. To describe such an equilibrium, it is useful to introduce *magnetic coordinates*, where one coordinate is constant on the constant-pressure surfaces and the field lines are straight in terms of the other coordinates [11]. To this end, first let ϑ and φ be arbitrary poloidal and toroidal angles. That is, we let (p, ϑ, φ) be an arbitrary coordinate system such that ϑ increases by 2π when the torus is traversed the short way around (the poloidal direction), and φ increases by 2π the long way around the torus (the toroidal direction). Since $\mathbf{B} \cdot \nabla p = 0$, the magnetic field can then be written as

$$\mathbf{B} = B_1(p, \vartheta, \varphi) \nabla p \times \nabla \vartheta + B_2(p, \vartheta, \varphi) \nabla \varphi \times \nabla p, \quad (4)$$

and since $\nabla \cdot \mathbf{B} = 0$ we have

$$0 = \nabla B_1 \cdot (\nabla p \times \nabla \vartheta) + \nabla B_2 \cdot (\nabla \varphi \times \nabla p) = \left(\frac{\partial B_1}{\partial \varphi} + \frac{\partial B_2}{\partial \vartheta} \right) \nabla p \cdot (\nabla \vartheta \times \nabla \varphi).$$

Hence

$$\frac{\partial}{\partial \varphi} \int_0^{2\pi} B_1 d\vartheta = 0$$

for all values of p and φ , so that

$$\int_0^{2\pi} B_1 d\vartheta = g(p)$$

for some function $g(p)$. This implies that B_1 and B_2 can be written in the form

$$B_1 = \frac{\partial f}{\partial \vartheta} + \frac{g(p)}{2\pi} \quad (5)$$

$$B_2 = -\frac{\partial f}{\partial \varphi} + \frac{h(p)}{2\pi}, \quad (6)$$

for some functions $f(p, \vartheta, \varphi)$ and $h(p)$. By renaming the functions f , g , and h , writing $\psi'(p) = g(p)/2\pi$, $\chi'(p) = h(p)/2\pi$, and $\lambda(p, \vartheta, \varphi) = f(p, \vartheta, \varphi)/\psi'(p)$, where primes refer to derivatives, we thus obtain

$$B_1 = \psi'(p) \left(1 + \frac{\partial \lambda}{\partial \vartheta} \right),$$

$$B_2 = \chi'(p) - \psi'(p) \frac{\partial \lambda}{\partial \varphi}.$$

Then $\mathbf{B} = B_1 \nabla p \times \nabla \vartheta + B_2 \nabla \varphi \times \nabla p$ becomes

$$\mathbf{B} = \left(1 + \frac{\partial \lambda}{\partial \vartheta} \right) \nabla \psi \times \nabla \vartheta + \nabla \varphi \times \nabla \chi - \frac{\partial \lambda}{\partial \varphi} \nabla \varphi \times \nabla \psi,$$

or, finally,

$$\mathbf{B} = \nabla \psi \times \nabla \theta + \nabla \varphi \times \nabla \chi, \quad (7)$$

where $\theta = \vartheta + \lambda$. This is the representation of \mathbf{B} in magnetic coordinates, and naturally leads to a decomposition of \mathbf{B} into toroidal and poloidal components (given by the first and second terms, respectively).

It is rewarding to interpret the result physically. We first note that ψ and χ are constant on surfaces of constant pressure, and we can choose both these functions to vanish on the magnetic axis. Let us calculate the magnetic flux that passes through a poloidal cross section ($\varphi = \text{constant}$)

between the magnetic axis and some surface $\psi = \text{constant}$, see Fig. 2. This flux is given by the surface integral

$$\int \mathbf{B} \cdot d\mathbf{S} = \int \mathbf{B} \cdot \hat{\mathbf{n}} dS,$$

where $\hat{\mathbf{n}} = \nabla\varphi/|\nabla\varphi|$. The surface element dS is related to the volume element

$$dV = \sqrt{g} d\psi d\theta d\varphi,$$

by $dV = dS d\varphi/|\nabla\varphi|$, and the Jacobian of the coordinates (ψ, θ, φ) , taken to be a right-handed system, is given by

$$\frac{1}{\sqrt{g}} = (\nabla\psi \times \nabla\theta) \cdot \nabla\varphi = \mathbf{B} \cdot \nabla\varphi. \quad (8)$$

The toroidal magnetic flux is thus

$$\int \mathbf{B} \cdot d\mathbf{S} = \int_0^\psi d\psi' \int_0^{2\pi} d\theta \frac{\mathbf{B} \cdot \nabla\varphi}{\nabla\psi' \cdot (\nabla\theta \times \nabla\varphi)} = 2\pi\psi. \quad (9)$$

Here the prime in ψ' does not denote a derivative but is used to distinguish the integration variable from the end point of the integral, $0 < \psi' < \psi$.

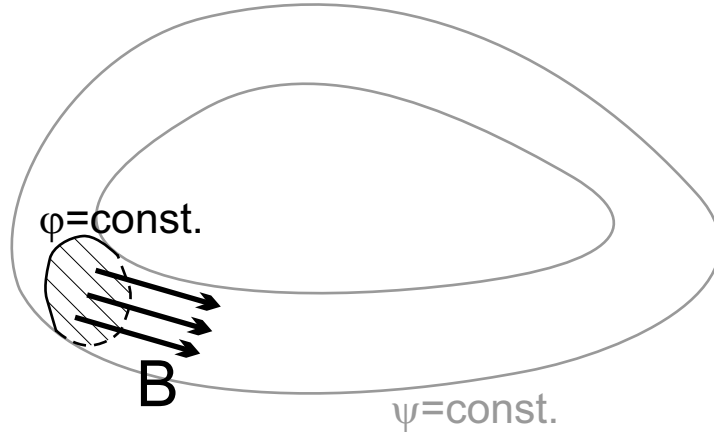


Figure 2: *The magnetic flux passing through a surface of constant φ is equal to $2\pi\psi$.*

Similarly, we can evaluate the poloidal flux, i.e., the amount of magnetic flux that passes through a surface $\theta = \text{constant}$ between the magnetic axis and a given flux surface ψ , as in Fig. 3. This flux is

$$\int \mathbf{B} \cdot d\mathbf{S} = \int_0^\psi d\psi' \int_0^{2\pi} d\varphi \frac{\mathbf{B} \cdot \nabla\theta}{\nabla\psi' \cdot (\nabla\theta \times \nabla\varphi)} = 2\pi\chi \quad (10)$$

Since the toroidal and poloidal magnetic fluxes are constant on surfaces of constant pressure, these surfaces are called flux surfaces. Functions that are constant on such surfaces, i.e., only depend on ψ and are independent of θ and φ , are called flux functions.

If χ is regarded as a function of ψ , the derivative

$$\iota(\psi) = \frac{d\chi}{d\psi} \quad (11)$$

is called the *rotational transform*¹, and $q = 1/\iota$ is called the *safety factor*. The rotational transform indicates how many poloidal turns a field line makes during each toroidal turn around the flux surface. This conclusion follows from the observation that, along a field line, θ and φ vary in proportion,

$$\frac{d\theta}{d\varphi} = \frac{\mathbf{B} \cdot \nabla\theta}{\mathbf{B} \cdot \nabla\varphi} = \frac{(\nabla\varphi \times \nabla\chi) \cdot \nabla\theta}{(\nabla\psi \times \nabla\theta) \cdot \nabla\varphi} = \iota.$$

This equation also indicates that the magnetic field lines are straight in the (θ, φ) plane, which is the defining property of magnetic coordinates. In fact, if we introduce $\alpha = \theta - \iota\varphi$, then

$$\mathbf{B} = \nabla\psi \times \nabla\alpha, \quad (12)$$

so that $\mathbf{B} \cdot \nabla\alpha = 0$ and α is constant along the magnetic field. Equation (12) is called the *Clebsch* representation of the magnetic field and is a general way of (locally) writing divergence-free vector fields, including those that do not possess flux surfaces.

In order to explore the properties of magnetic coordinates further, we first note that since $\mathbf{J} \cdot \nabla p = 0$, we can write

$$\mu_0 \mathbf{J} = J_1 \nabla\psi \times \nabla\theta + J_2 \nabla\varphi \times \nabla\psi \quad (13)$$

in analogy with (4). Furthermore, since $\nabla \cdot \mathbf{J} = 0$ because of Eq. (1) the same reasoning as that leading up to Eqs. (5) and (6) implies J_1 and J_2 can be written as

$$J_1 = I'(\psi) - \frac{\partial K}{\partial \theta},$$

$$J_2 = -G'(\psi) + \frac{\partial K}{\partial \varphi},$$

for some functions $I(\psi)$, $G(\psi)$ and $K(\psi, \theta, \varphi)$. Since then

$$\mu_0 \mathbf{J} = \nabla \times (I \nabla\theta + G \nabla\varphi + K \nabla\psi) = \nabla \times \mathbf{B}, \quad (14)$$

¹This quantity is sometimes denoted by $\iota/2\pi$ or t .

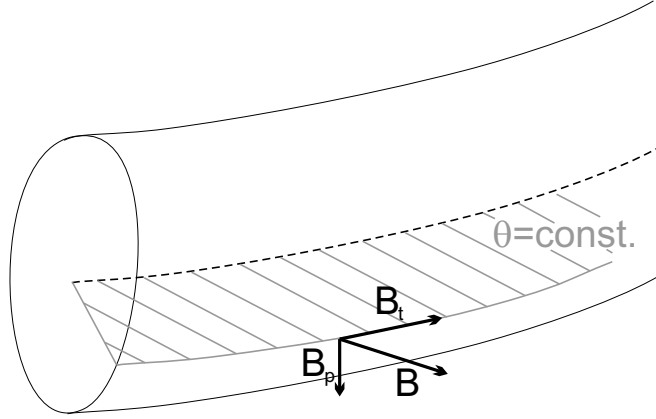


Figure 3: The poloidal component, B_p , of the magnetic field contributes to the magnetic flux passing through a ribbon of constant θ . This flux is equal to $2\pi\chi$.

it follows that it must be possible to write the magnetic field as

$$\mathbf{B} = I\nabla\theta + G\nabla\varphi + K\nabla\psi + \nabla H, \quad (15)$$

where the function $H(\psi, \theta, \varphi)$ plays the role of an integration constant. The flux functions $I(\psi)$ and $G(\psi)$ have a simple physical interpretation. The integral of the magnetic field around a contour of constant ψ and φ becomes according to Eq. (15)

$$\oint_{\varphi=\text{constant}} \mathbf{B} \cdot d\mathbf{r} = 2\pi I(\psi). \quad (16)$$

According to Ampère's law, this integral is equal to μ_0 times the toroidal current inside the flux surface ψ . Of course, the integral could be taken along any poloidally closed contour without affecting the result. Similarly, a toroidal contour integral of \mathbf{B} implies that

$$\oint_{\theta=\text{constant}} \mathbf{B} \cdot d\mathbf{r} = 2\pi G(\psi)$$

is equal to μ_0 times the poloidal current between the flux surface labelled by ψ and infinity. In both tokamaks and stellarators, most of this current tends to flow in the toroidal field coils. In a stellarator, I tends to be much smaller than G .

2.3 The flux-surface average

The volume enclosed by a flux surface is equal to

$$V(\psi) = \int_0^\psi d\psi' \int_0^{2\pi} d\theta \int_0^{2\pi} \sqrt{g} d\varphi, \quad (17)$$

where the Jacobian was given in Eq. (8). The *flux-surface average* of an arbitrary function $f(\psi, \theta, \varphi)$ is defined as the volume average of this function over the volume between two neighbouring flux surfaces

$$\begin{aligned} \langle f \rangle &= \lim_{\Delta\psi \rightarrow 0} \frac{1}{V(\psi + \Delta\psi) - V(\psi)} \int_\psi^{\psi + \Delta\psi} d\psi \int_0^{2\pi} d\theta \int_0^{2\pi} f(\psi, \theta, \varphi) \sqrt{g} d\varphi \\ &= \frac{1}{V'(\psi)} \int_0^{2\pi} d\theta \int_0^{2\pi} f(\psi, \theta, \varphi) \sqrt{g} d\varphi \end{aligned} \quad (18)$$

Since the volume element can be written $dV = dS d\psi / |\nabla\psi|$, where

$$dS = \frac{d\theta d\varphi}{(\nabla\theta \times \nabla\varphi) \cdot \mathbf{n}}$$

is the area element on a flux surface whose normal vector is $\mathbf{n} = \nabla\psi / |\nabla\psi|$, the flux-surface average can be written as

$$\langle f \rangle = \int \frac{f dS}{|\nabla\psi|} / \int \frac{dS}{|\nabla\psi|}.$$

If we regard the Jacobian \sqrt{g} as a function not of (ψ, θ, φ) but of the Clebsch coordinates (ψ, α, φ) , then

$$V'(\psi) = \int_0^{2\pi} d\alpha \int_0^{2\pi} \sqrt{g(\psi, \alpha, \varphi)} d\varphi$$

The φ -integral is here taken at constant ψ and α , and can thus be interpreted as an integral along the magnetic field. The line element along \mathbf{B} is

$$dl = \frac{B d\varphi}{\mathbf{B} \cdot \nabla\varphi} = \sqrt{g} B d\varphi,$$

so

$$V'(\psi) = \int_0^{2\pi} d\alpha \int \frac{dl}{B}, \quad (19)$$

where the l -integral is to be taken over the length of the field line once around the torus. Similarly, the flux-surface average becomes

$$\langle f \rangle = \frac{1}{V'(\psi)} \int_0^{2\pi} d\alpha \int \frac{f dl}{B}.$$

The divergence of a vector \mathbf{F} is given by

$$\nabla \cdot \mathbf{F} = \frac{1}{\sqrt{g}} \sum_i \frac{\partial}{\partial x^i} (\sqrt{g} \mathbf{F} \cdot \nabla x^i),$$

in arbitrary coordinates x^i . Choosing these to be our magnetic coordinates (ψ, θ, φ) , we find the useful relation

$$\langle \nabla \cdot \mathbf{F} \rangle = \frac{1}{V'(\psi)} \frac{\partial}{\partial \psi} \langle V'(\psi) \mathbf{F} \cdot \nabla \psi \rangle. \quad (20)$$

An immediate consequence is that

$$\langle \mathbf{B} \cdot \nabla f \rangle = 0 \quad (21)$$

for any single-valued function f . Another corollary is that a so-called *magnetic differential equation*

$$\mathbf{B} \cdot \nabla f = S,$$

only has a solution if $\langle S \rangle = 0$.

2.4 Magnetic field lines

Each magnetic field line is identified by two coordinates (ψ, α) , where $\alpha = \theta - \iota\varphi$ is the Clebsch angle. We have already noted that, along a field line, we have $d\theta/d\varphi = \iota$ in magnetic coordinates. Thus, if we follow a field line once around the torus toroidally the poloidal angle changes by

$$\theta \rightarrow \theta + 2\pi\iota.$$

We thus conclude that (ψ, α) and $(\psi, \alpha + 2\pi\iota)$ label the same field line. There is a qualitative difference between flux surfaces where rotational transform ι is a rational number and where it is irrational. In the former case, $\iota = n/m$, the field line returns to where it started – it bites its tail, so to speak – after m toroidal turns around the torus. If ι is irrational, it never does so and instead traces out the entire flux surface. Not only does it come arbitrarily close to each point on the surface, but in addition it covers the surface evenly in a certain sense. This follows from “Weyl’s lemma” [12], which asserts that if ι is an irrational number and $f(\alpha)$ an arbitrary, continuous, 2π -periodic function, then

$$\frac{1}{2\pi} \int_0^{2\pi} f(\alpha) d\alpha = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m f(2\pi k\iota).$$

This result is intuitively plausible from the fact that if ι is not irrational but a high-order rational ($\iota = n/m$ with large m and n), then the points with the poloidal angle $\theta = 2\pi k\iota$ are approximately evenly distributed over the interval $[0, 2\pi)$ modulo 2π . Indeed, if $\iota = n/m$ then

$$\frac{1}{2\pi} \int_0^{2\pi} f(\alpha) d\alpha \simeq \frac{1}{m} \sum_{k=1}^m f(2\pi k\iota)$$

if $m \gg 1$.

Weyl's lemma enables us to interpret the expressions for the volume (17) enclosed by a flux surface and the flux-surface average (18) in an interesting way. If $\iota \simeq n/m$ we have

$$V'(\psi) \simeq \frac{2\pi}{m} \sum_{k=1}^m \int \frac{dl}{B} \Big|_{\alpha=2\pi k\iota}.$$

But the right-hand-side of this expression is simply the integral along a single field line m toroidal times around the torus, i.e., until the field line closes on itself, which we write as

$$V'(\psi) \simeq \frac{2\pi}{m} \oint \frac{dl}{B}.$$

The flux-surface-average (18) can thus be expressed as

$$\langle f \rangle \simeq \oint \frac{f dl}{B} / \oint \frac{dl}{B}, \quad (22)$$

in terms of integrals around a closed field line on a high-order rational magnetic surface. Since the rotational transform may be approximated to any degree of accuracy by a sufficiently high-order rational number, this allows us to interpret the flux-surface average for any continuous function f in this way.

This result is readily understood in terms of volume averages in flux tubes. A slender flux tube along the magnetic field has a cross section of area

$$dS = \frac{d\Phi}{B},$$

where $d\Phi$ is the magnetic flux inside the tube. Its volume is thus

$$dV = \int dl dS = d\Phi \int \frac{dl}{B},$$

and it is clear that the flux-surface average (22) is an average over the volume of a long flux tube that is wound many turns around the torus.

There is an important application of these concepts to the theory of MHD stability of stellarators. Although this subject is beyond the scope of the present review, it is worth mentioning that a magnetic configuration with $V''(\psi) < 0$ is said to possess a *magnetic well*. This is beneficial for stability, as can be understood intuitively from the following argument. Consider an interchange instability causing a certain flux tube to move radially outward. In ideal MHD, the magnetic field is frozen into the plasma, so the tube carries both plasma and magnetic field with it. Now, the quantity $V'(\psi)$ can be understood as the “specific volume” of the plasma: $dV = V'(\psi)d\psi$ is the volume available to a bit of plasma threaded by the flux $d\psi$. If $V''(\psi)$ is positive, the specific volume increases with radius, so that a flux tube that moves radially outward will tend to expand. If the pressure decreases with radius, $p'(\psi) < 0$, this liberates thermal energy stored in the plasma, and we conclude that plasma equilibria with $V''(\psi) > 0$ are particularly vulnerable to pressure-driven instabilities.

The *curvature* of a magnetic field line is defined by

$$\boldsymbol{\kappa} = \mathbf{b} \cdot \nabla \mathbf{b}, \quad (23)$$

where $\mathbf{b} = \mathbf{B}/B$ is the unit vector in the direction of the field. Since $\mathbf{b} \cdot \nabla \mathbf{b} \cdot \mathbf{b} = 0$, the curvature vector is orthogonal to the direction of the field line. If the latter is locally approximated by a circle as closely as possible, the curvature vector $\boldsymbol{\kappa}$ points in the direction from the field line toward the centre of this circle, and its magnitude $\kappa = 1/R$ is the inverse of its radius. The component of $\boldsymbol{\kappa}$ that is orthogonal to the flux surface is called the *normal curvature*

$$\kappa_n = \mathbf{n} \cdot \boldsymbol{\kappa},$$

where $\mathbf{n} = \nabla\psi/|\nabla\psi|$. The normal curvature is important in stability theory since, according to the ideal MHD energy principle [13], negative normal curvature is destabilising if the pressure decreases outward, $dp/d\psi < 0$. The component of the curvature vector that is tangential to the surface,

$$\kappa_g = (\mathbf{n} \times \mathbf{b}) \cdot \boldsymbol{\kappa},$$

is called the *geodesic curvature*. This quantity can be defined for any curve with tangent vector \mathbf{b} and vanishes for geodesics, which are defined as curves of shortest length: a geodesic between two points (not too far from each other) on a surface is the shortest curve between these two points lying

entirely within the surface. That κ_g vanishes for such a curve is intuitively clear from the fact that a curve of shortest length should be “as straight as possible”. Since

$$\boldsymbol{\kappa} = -\mathbf{b} \times \left(\nabla \times \frac{\mathbf{B}}{B} \right) = -\mathbf{b} \times \frac{\mu_0 \mathbf{J} + \mathbf{b} \times \nabla B}{B},$$

the curvature can be written as

$$\boldsymbol{\kappa} = \frac{\mu_0 \nabla p}{B^2} + \frac{\nabla_{\perp} B}{B}, \quad (24)$$

where $\nabla_{\perp} B = \nabla B - (\mathbf{b} \cdot \nabla B) \mathbf{b}$ is the component of ∇B that is perpendicular to the field. We note that the first term in Eq. (24) is small in a low-pressure plasma (in the sense $p \ll B^2/2\mu_0$) and never contributes to the geodesic curvature, which can be written as

$$\kappa_g = (\mathbf{n} \times \mathbf{b}) \cdot \nabla \ln B. \quad (25)$$

2.5 Hamada and Boozer coordinates

In our construction of magnetic coordinates, we started with completely arbitrary poloidal and toroidal angles, i.e., ϑ and φ were allowed to represent any functions with the right periodicity properties. The poloidal angle was then modified so as to make the magnetic field lines straight, but the toroidal angle was left intact. For each choice of toroidal angle φ there is thus a corresponding choice of poloidal angle θ leading to the representation (7). There are, therefore, many sets of magnetic coordinates (θ, φ) , and the freedom to choose φ arbitrarily can be used to achieve further simplifications. Two particularly useful sets of magnetic coordinates were identified by Hamada [14] and Boozer [15], respectively.

If we write

$$\begin{aligned} \theta &= \theta' + \iota\omega, \\ \varphi &= \varphi' + \omega \end{aligned}$$

where $\omega(\psi, \theta, \varphi)$ is well-behaved and periodic in the poloidal and toroidal directions, then Eqs. (7) and (15) become

$$\mathbf{B} = \nabla\psi \times \nabla\theta' + \nabla\varphi' \times \nabla\chi = I\nabla\theta' + G\nabla\varphi' + K'\nabla\psi + \nabla H' \quad (26)$$

with

$$H' = H + (\iota I + G)\omega,$$

$$K' = K - \omega \frac{d}{d\psi}(\iota I + G).$$

For *Boozer coordinates* ω is chosen so that H' vanishes,

$$\omega = -\frac{H}{\iota I + G}$$

i.e.,

$$\mathbf{B} = I(\psi)\nabla\theta + G(\psi)\nabla\varphi + K(\psi, \theta, \varphi)\nabla\psi, \quad (27)$$

where we have dropped primes. The Jacobian for these coordinates is found from the scalar product of the two expressions (26) for \mathbf{B} , giving

$$\frac{1}{\sqrt{g}} = (\nabla\psi \times \nabla\theta) \cdot \nabla\varphi = \frac{B^2}{\iota I + G} \quad (28)$$

The basic advantage of Boozer coordinates is that the magnetic field has a particularly simple covariant representation (27). It may seem surprising that this representation has a component proportional to $\nabla\psi$ although the magnetic field is tangential to the flux surfaces. The reason is that the coordinate system is non-orthogonal, so that $\nabla\theta$ and $\nabla\varphi$ have components that are perpendicular to the flux surface that need to be cancelled by the final term in Eq. (27). Geometrically, Boozer coordinates have the property that lines that are perpendicular to \mathbf{B} and tangential to flux surfaces (the streamlines of $\mathbf{B} \times \nabla\psi$) are straight in these coordinates.

For *Hamada coordinates*, ω is instead chosen so that K' vanishes, i.e.,

$$\mathbf{B} = I\nabla\theta + G\nabla\varphi + \nabla H,$$

where $I = I(\psi)$, $G = G(\psi)$ and $H = H(\psi, \theta, \varphi)$. The current becomes

$$\mu_0\mathbf{J} = \nabla \times \mathbf{B} = \nabla I \times \nabla\theta + \nabla G \times \nabla\varphi \quad (29)$$

In other words, the streamlines of both \mathbf{B} and \mathbf{J} are straight in Hamada coordinates. They both wind around the torus, but do so with different rotational transforms. By analogy with Eqs.(7) and (11), it follows from Eq. (29) that the rotational transform of the current streamlines is equal to

$$\iota_J = -\frac{G'}{I'} = -\frac{dG}{dI}. \quad (30)$$

Finally, we note that the Hamada Jacobian is a flux function, as follows from Eq. (31) below with $K = 0$.

It is important to note that in axisymmetric configurations, the toroidal Boozer and Hamada angles are not equal to the geometric angle used in cylindrical coordinates around the axis of symmetry. The vector $\nabla\varphi$ does not point purely in the toroidal direction if φ denotes the toroidal angle in Boozer or Hamada coordinates.

2.6 More on force balance

We are now in a position to reconsider the force balance relation (3). Substituting Eqs.(7) and (14) gives

$$\mathbf{J} \times \mathbf{B} = \frac{1}{\mu_0 \sqrt{g}} \left(\iota \frac{\partial K}{\partial \theta} + \frac{\partial K}{\partial \varphi} - \iota I' - G' \right) \nabla \psi = p'(\psi) \nabla \psi, \quad (31)$$

which we divide by $\nabla\psi$ and flux-surface average to obtain

$$\iota I' + G' = -\frac{\mu_0}{4\pi^2} p' V'.$$

This is the basic requirement that the toroidal and poloidal currents, I and G respectively, need to satisfy in order to produce a magnetic force that balances the pressure gradient. It can be expressed as a condition on the rotational transforms of the magnetic field lines (11) and the current lines (30),

$$\iota - \iota_J = -\frac{\mu_0}{4\pi^2} \frac{p' V'}{I'}.$$

Let us consider two magnetic field lines intersecting two current lines in the points (P, Q) and (P', Q') , respectively, as in Fig. 4. These four lines thus bound a “rectangle” on the flux surface, with current flowing across the segments PQ and $P'Q'$. If we consider the volume between this rectangle and the corresponding one on the neighbouring flux surface $\psi + d\psi$, the total current across the boundary PQ is

$$I_{PQ} = \int \mathbf{J} \cdot d\mathbf{S},$$

where the surface element is

$$d\mathbf{S} = \frac{\nabla\alpha}{(\nabla\psi \times \nabla\alpha) \cdot \nabla\varphi} d\psi d\varphi,$$

in the coordinates (ψ, α, φ) , or

$$d\mathbf{S} = \frac{\nabla\alpha}{B} d\psi dl,$$

in terms of the arc length $dl = d\varphi/(\mathbf{b} \cdot \nabla\varphi)$ along the field. Note that the Clebsch angle $\alpha = \theta - \iota\varphi$ is constant along the magnetic field and thus over the surface of the integral.

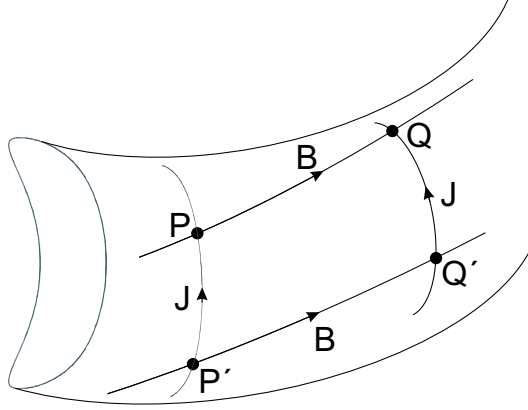


Figure 4: Area bounded by two magnetic field lines and two streamlines of the current. The currents flowing across the segments PQ and $P'Q'$ are equal.

Since

$$\mathbf{J} \times \mathbf{B} = \mathbf{J} \times (\nabla\psi \times \nabla\alpha) = (\mathbf{J} \cdot \nabla\alpha)\nabla\psi,$$

we have $\mathbf{J} \cdot \nabla\alpha = p'(\psi)$ and

$$I_{PQ} = (p'd\psi) \int_P^Q \frac{dl}{B}.$$

Since the current flowing across the segment PQ must equal that across $P'Q'$, we conclude that

$$p'(\psi) \int_P^Q \frac{dl}{B} = p'(\psi) \int_{P'}^{Q'} \frac{dl}{B},$$

as long as the end points in the two integrals lie on the same stream lines of the current. In the special case of a rational magnetic surface, we can take the integral along a closed field line, so that P and Q coincide. We then conclude that the resulting integral, is the same for all field lines on the same flux surface, unless $p'(\psi)$ vanishes,

$$p'(\psi) \frac{\partial}{\partial\alpha} \oint \frac{dl}{B} = 0. \quad (32)$$

In other words, unless the pressure gradient vanishes, the so-called *Hamada condition* needs to be satisfied: all flux tubes carrying the same magnetic flux on a given surface must have the same volume. This condition hints at a paradox in low-pressure plasmas: it must hold in any configuration with $\mathbf{J} \times \mathbf{B} = \nabla p \neq 0$, no matter how small the pressure gradient may be, but is not generally true in the absence of pressure. The reason for this difficulty is the assumption that the magnetic field traces out perfectly nested flux surfaces, which, as we shall see, in general requires a singular current on surfaces with rational rotational transform.

2.7 Plasma current

The magnetic field in a stellarator is primarily created by the magnetic field coils, but it is modified by the plasma current that arises to maintain force balance (3). This current generally has components both perpendicular and parallel to the magnetic field, $\mathbf{J} = J_{\parallel} \mathbf{b} + \mathbf{J}_{\perp}$. Only the perpendicular component of the current is necessary for producing a magnetic force $\mathbf{J} \times \mathbf{B} = \nabla p$, namely,

$$\mathbf{J}_{\perp} = \frac{\mathbf{B} \times \nabla p}{B^2}, \quad (33)$$

but in addition it is necessary to satisfy $\nabla \cdot \mathbf{J} = 0$, which follows from $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$ and usually requires a parallel (so-called Pfirsch-Schlüter) current since

$$\nabla \cdot \mathbf{J}_{\perp} = (\mathbf{B} \times \nabla p) \cdot \nabla \left(\frac{1}{B^2} \right)$$

in general does not vanish. Since $\nabla \cdot (J_{\parallel} \mathbf{B}/B) = \mathbf{B} \cdot \nabla (J_{\parallel}/B)$, the parallel current density is given by

$$J_{\parallel} = u(\psi, \theta, \varphi) p'(\psi) B + \frac{\langle J_{\parallel} B \rangle B}{\langle B^2 \rangle}, \quad (34)$$

where the function $u(\psi, \theta, \varphi)$ satisfies the magnetic differential equation

$$\mathbf{B} \cdot \nabla u = -(\mathbf{B} \times \nabla \psi) \cdot \nabla \left(\frac{1}{B^2} \right) \quad (35)$$

and we have fixed the integration constant by requiring $\langle u B^2 \rangle = 0$. The Pfirsch-Schlüter current, which we define as the first term on the right-hand side of Eq. (34),

$$J_{\parallel}^{PS} = u p' B, \quad (36)$$

vanishes in configurations with zero geodesic curvature (25), since

$$\mathbf{b} \cdot \nabla u = -\frac{2\kappa_g |\nabla\psi|}{B^2},$$

but otherwise always arises as soon as there is a non-zero pressure gradient. The second term in Eq. (34) describes the Ohmic current, the bootstrap current (see below), and any non-inductively driven current. In a plasma that satisfies a conventional Ohm's law,

$$\eta J_{\parallel} = E_{\parallel} = -\nabla_{\parallel}\phi - \frac{\partial A_{\parallel}}{\partial t},$$

where η is the resistivity, and ϕ and \mathbf{A} the electromagnetic potentials, there is no contribution to $\langle J_{\parallel} B \rangle$ from the electrostatic field $-\nabla_{\parallel}\phi$, because of Eq. (21).

The full current vector is given by the sum of the diamagnetic current (33) and the parallel current (34),

$$\mathbf{J} = \frac{\mathbf{B} \times \nabla p}{B^2} + \left(u(\psi, \theta, \varphi) p'(\psi) + \frac{\langle J_{\parallel} B \rangle}{\langle B^2 \rangle} \right) \mathbf{B}. \quad (37)$$

In keeping with convention, we have defined the first term in the brackets as the Pfirsch-Schlüter current, but one should note that this term is not divergence-free and should therefore not be treated as an entity separate from the diamagnetic current. A physically meaningful question, however, is whether these two terms contribute to the total toroidal current flowing around the torus? That is, does the pressure gradient alone create a net toroidal current in an Ohmic plasma?

The net toroidal current flowing in the volume V bounded by the flux surface ψ is equal to the surface integral of the current density across an arbitrary surface of constant φ ,

$$I_{\text{tor}}(\psi) = \int_0^{\psi} d\psi \int_0^{2\pi} (\mathbf{J} \cdot \nabla\varphi) \sqrt{g} d\theta \Big|_{\varphi=\text{const.}},$$

where $d\mathbf{S} = (\nabla\varphi) \sqrt{g} d\psi d\theta$ is the surface element. Since this current cannot depend on the toroidal location of the surface over which the integral is taken, we can write

$$I_{\text{tor}}(\psi) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^{\psi} d\psi \int_0^{2\pi} (\mathbf{J} \cdot \nabla\varphi) \sqrt{g} d\theta = \frac{1}{2\pi} \int_V \mathbf{J} \cdot \nabla\varphi dV.$$

The last term in Eq. (37), which represents the Ohmic and non-inductively driven currents, contributes the amount

$$\frac{1}{2\pi} \int_V \frac{\langle J_{\parallel} B \rangle}{\langle B^2 \rangle} \mathbf{B} \cdot \nabla \varphi dV = \frac{1}{2\pi} \int_V \frac{\langle J_{\parallel} B \rangle}{\iota I + G} dV$$

to the net toroidal current, where we have used Boozer coordinates with the Jacobian (28). As already mentioned, the first two terms in Eq. (37) should, strictly speaking, be considered together since they are not individually divergence-free, but the second term makes no contribution,

$$\frac{1}{2\pi} \int_V u p' (\mathbf{B} \cdot \nabla \varphi) dV = 0,$$

because of the condition $\langle u B^2 \rangle = 0$. The diamagnetic current, finally, i.e., the first term on the right-hand side of Eq. (37), makes the contribution

$$\frac{1}{2\pi} \int_V \frac{dp}{d\psi} \frac{(I \nabla \theta \times \nabla \psi) \cdot \nabla \varphi}{B^2} dV = - \int \frac{I p'}{\iota I + G} dV,$$

which can be non-zero if $I \neq 0$. More precisely, the pressure gradient produces a net toroidal current on a given flux surface only if there is a net toroidal current inside this surface. In other words, if a toroidal current is driven by other means – for instance non-inductively or Ohmically, as in a tokamak – then the diamagnetic and Pfirsch-Schlüter currents cause an additional net current to flow in the toroidal direction. In the absence of other currents (which is the usual situation in a stellarator), there is no net diamagnetic plus Pfirsch-Schlüter current toroidally.

There is a simple geometric reason for this conclusion. Consider the line integral

$$\oint_C \mathbf{B} \cdot d\mathbf{r},$$

taken along a contour C that

- (i) lies entirely within a flux surface,
- (ii) first runs perpendicular to \mathbf{B} poloidally once around the torus until it encounters the field line on which it started,
- (iii) and then follows the magnetic field back to the starting point,

as in Fig. 5. Only the last portion of C will contribute to the integral. On the other hand, according to Ampère's law the integral is equal to the enclosed toroidal current (16),

$$\oint_C \mathbf{B} \cdot d\mathbf{r} = 2\pi I(\psi).$$

Thus, if $I(\psi) = 0$ the curves perpendicular to \mathbf{B} are poloidally closed. But these curves are stream lines of the diamagnetic current, and it is thus obvious that this current will only give rise to a net toroidal current if $I(\psi) \neq 0$.

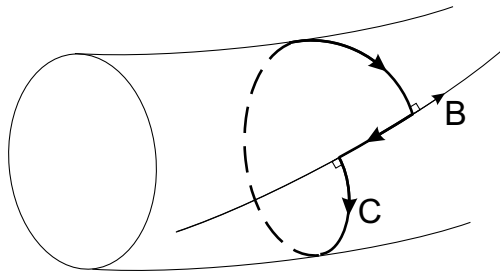


Figure 5: *The integration contour C runs perpendicular to \mathbf{B} poloidally once around the torus, and then along \mathbf{B} back to the starting point.*

Like any current does, the Pfirsch-Schlüter current creates a magnetic field and thus changes the magnetic equilibrium, making the latter dependent on the plasma pressure. If this pressure is increased from zero, the Pfirsch-Schlüter current builds up and changes the shape and position of the magnetic surfaces. One of the most noticeable changes is an outward (in major radius) shift of the central flux surfaces, the so-called *Shafranov shift*. It arises for essentially the same reason that a rubber tyre expands radially when it is being pumped up. In a classical stellarator, with inverse aspect ratio $\epsilon \ll 1$, the Shafranov shift becomes large enough to be comparable to the minor radius at an average plasma pressure of about

$$\beta = \frac{2\mu_0 p}{B^2} \sim \frac{\epsilon^2}{2},$$

which sets an equilibrium limit on the achievable pressure.

2.8 Magnetic shear

In general, the cross section of a flux tube changes shape along the tube. If, for instance, the cross section is square at some location, further along the tube it will typically be rectangular or trapezoidal, having been compressed in one direction and stretched in the other; one speaks of *magnetic shear*.

Magnetic shear arises if the rotational transform varies from flux surface to surface, $dt/d\psi \neq 0$, since the field lines on neighbouring flux surfaces then do not rotate equally fast around the magnetic axis and flux tubes therefore

get sheared in the poloidal direction. For this reason, the *global magnetic shear* is in the literature defined as $\iota'(\psi)$, or as some quantity proportional to $\iota'(\psi)$. (In the tokamak literature it is usually defined as $s = rq'(r)/q$, where $q = 1/\iota$ and r is an appropriately chosen flux function that reduces to the minor radius when the flux surfaces are circular.)

The *local magnetic shear* is traditionally defined as [17]

$$S = -\mathbf{X} \cdot (\nabla \times \mathbf{X}),$$

with

$$\mathbf{X} = \frac{\mathbf{B} \times \nabla\psi}{|\nabla\psi|^2}, \quad (38)$$

but the nomenclature is not entirely consistent in the literature. Sometimes ψ is replaced by some other flux function $F(\psi)$, and some authors replace \mathbf{X} by $\mathbf{b} \times \nabla\psi/|\nabla\psi|$. In any case, using $\mathbf{B} = \nabla\psi \times \nabla\alpha$ in the definition (38) gives

$$\mathbf{X} = \nabla\alpha - \frac{\nabla\psi \cdot \nabla\alpha}{|\nabla\psi|^2} \nabla\psi$$

and

$$S = \mathbf{X} \cdot \left[\nabla \left(\frac{\nabla\psi \cdot \nabla\alpha}{|\nabla\psi|^2} \right) \times \nabla\psi \right] = \mathbf{B} \cdot \nabla \left(\frac{\nabla\psi \cdot \nabla\alpha}{|\nabla\psi|^2} \right).$$

If we denote the angle between $\nabla\psi$ and $\nabla\alpha$ by λ , so that $\nabla\psi \cdot \nabla\alpha = |\nabla\psi||\nabla\alpha|\cos\lambda$, it follows that the line integral of S/B along a field line between two points, P and Q , is equal to

$$\int_P^Q \frac{Sdl}{B} = \left[\frac{|\nabla\alpha|}{|\nabla\psi|} \cos\lambda \right]_P^Q.$$

Hence it is clear that S is a measure of the growth of λ along a field line.

The flux-surface average of S does not vanish, although it is of the form (21), because the last term in

$$S = \mathbf{B} \cdot \nabla \left(\frac{\nabla\psi \cdot (\nabla\theta - \iota\nabla\varphi)}{|\nabla\psi|^2} - \iota'(\psi)\varphi \right).$$

is not single-valued. Instead

$$\langle S \rangle = -4\pi^2 \frac{d\iota}{dV},$$

where $V(\psi)$ is the volume (17) inside the flux surface. On a flux-surface average, the local shear is thus equal to (a multiple of) the global shear.

2.9 Expansion around the magnetic axis

Spitzer was the first to realise that it is possible to create a rotational transform without a toroidal current [1]. This may appear surprising in view of Ampère's law (16), which seems to state that a net poloidal field can only arise if the toroidal current $I(\psi)$ is non-zero. But the fact that the average poloidal field vanishes in the sense expressed by this equation does not imply $\iota = 0$, because ι measures a different kind of poloidal-field average. In terms of mathematics, the point is that the expression

$$\iota = \frac{\langle \mathbf{B} \cdot \nabla \theta \rangle}{\langle \mathbf{B} \cdot \nabla \varphi \rangle}$$

is different from

$$\frac{I}{G} = \frac{\langle \mathbf{B} \cdot (\nabla \varphi \times \nabla \psi) \rangle}{\langle \mathbf{B} \cdot (\nabla \psi \times \nabla \theta) \rangle},$$

which follows from Eqs. (15) and (18). In fact, *most* non-axisymmetric MHD equilibria have non-zero rotational transform even in the absence of a net toroidal current. An example from tokamaks is furnished by magnetic islands. A magnetic island need not enclose any plasma current but nevertheless consists of flux surfaces traced out by field lines. A stellarator example is shown in Figure 6, which depicts two poloidal cuts (at constant toroidal geometric angle) of a magnetic surface in Wendelstein 7-X, with the poloidal field indicated by arrows. There is no net toroidal current and the line integral (16) thus vanishes.

A Spitzer realised, there are three ways of twisting the magnetic field in a large-aspect-ratio torus [18]. The first is to use a toroidal current, the second is to make the magnetic axis not lie in a plane, and the third is to make the poloidal cross section of the flux surfaces (which are elliptical close to the magnetic axis) rotate poloidally as one goes around the torus in the toroidal direction. Most stellarators use a combination of these methods.

To see how a poloidal magnetic field and a rotational transform arises mathematically, one needs to solve Maxwell's equations, $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$. We follow Mercier [19] and expand these equations in the vicinity of the magnetic axis, which is denoted by $\mathbf{r}_0(l)$ when parameterised by the arc length l . Then $\hat{\mathbf{e}}_1 = d\mathbf{r}_0/dl$ is a unit tangent vector along the axis (which we have earlier denoted by \mathbf{b}) and

$$\hat{\mathbf{e}}_2 = \frac{1}{\kappa} \frac{d\hat{\mathbf{e}}_1}{dl}$$

is a unit vector in the direction of the curvature (23). A third unit vector can be defined by $\hat{\mathbf{e}}_3 = \hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2$, and since $d(\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2)/dl = 0$ we have $\hat{\mathbf{e}}_1 \cdot (d\hat{\mathbf{e}}_2/dl) =$

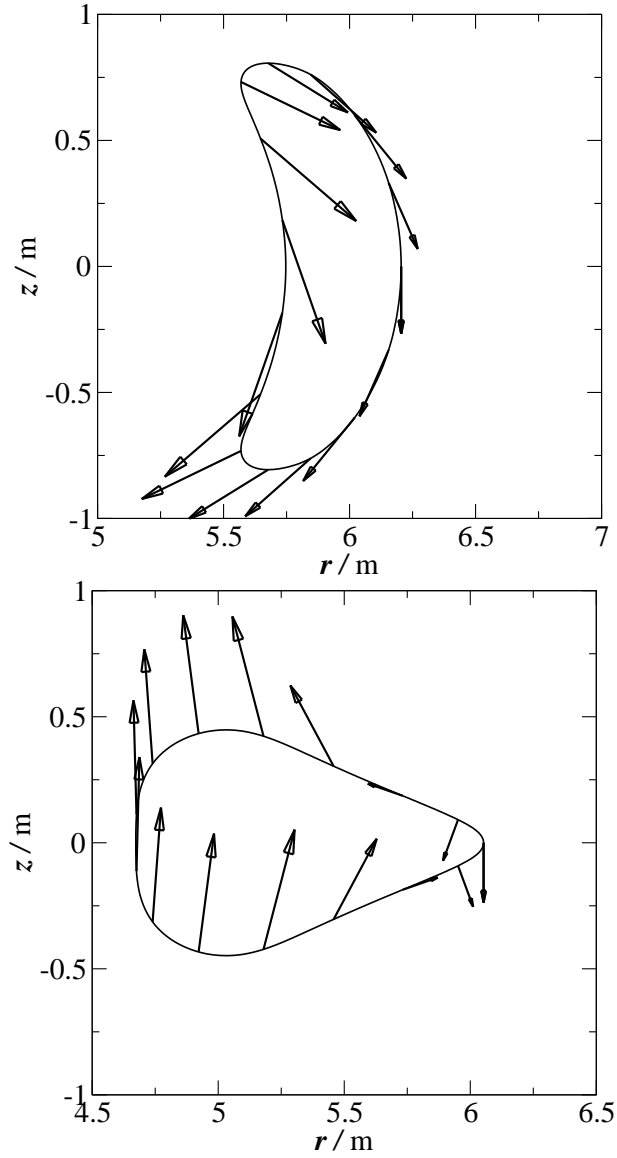


Figure 6: *The two up-down symmetric cross sections of a flux surface in the Wendelstein 7-X stellarator. The arrows indicate the strength and direction of the magnetic field projected onto these surfaces. There is no net toroidal current, and the integral (16) thus vanishes.*

$-\kappa$. Because also $\hat{\mathbf{e}}_2 \cdot (d\hat{\mathbf{e}}_2/dl) = 0$ it must be possible to write

$$\frac{d\hat{\mathbf{e}}_2}{dl} = -\kappa\hat{\mathbf{e}}_1 + \tau\hat{\mathbf{e}}_3$$

for some number τ , which is called the *torsion*. In a similar way, one concludes from $d(\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_3)/dl = d(\hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_3)/dl = 0$ that

$$\frac{d\hat{\mathbf{e}}_3}{dl} = -\tau\hat{\mathbf{e}}_2.$$

If the magnetic axis lies in a plane, then $d\hat{\mathbf{e}}_3/dl = 0$, but more generally the torsion measures how much it departs locally from a plane curve. These so-called Frenet-Serret relations play an important role in the differential geometry of curves.

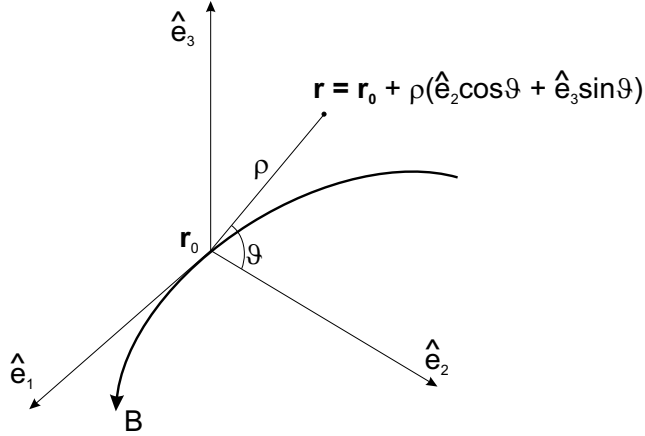


Figure 7: *Cylindrical coordinates locally aligned with the magnetic axis.*

We now introduce local cylindrical coordinates (ρ, ϑ, l) in the vicinity of the magnetic axis,

$$\mathbf{r}(\rho, \vartheta, l) = \mathbf{r}_0(l) + \hat{\mathbf{e}}_2(l)\rho \cos \vartheta + \hat{\mathbf{e}}_3(l)\rho \sin \vartheta,$$

as in Fig. 7. Then

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial \rho} &= \hat{\mathbf{e}}_2 \cos \vartheta + \hat{\mathbf{e}}_3 \sin \vartheta, \\ \frac{\partial \mathbf{r}}{\partial \vartheta} &= \rho(-\hat{\mathbf{e}}_2 \sin \vartheta + \hat{\mathbf{e}}_3 \cos \vartheta), \\ \frac{\partial \mathbf{r}}{\partial l} &= (1 - \kappa\rho \cos \vartheta)\hat{\mathbf{e}}_1 + \tau\rho(-\hat{\mathbf{e}}_2 \sin \vartheta + \hat{\mathbf{e}}_3 \cos \vartheta), \end{aligned}$$

and the distance between two neighbouring points is given by

$$ds^2 = d\mathbf{r} \cdot d\mathbf{r} = d\rho^2 + \rho^2(d\vartheta + \tau dl)^2 + (1 - \kappa\rho \cos \vartheta)^2 dl^2.$$

Hence it is clear that the coordinate system is not orthogonal, but can be made so by introducing the modified cylindrical angle

$$\omega = \vartheta + \int_0^l \tau(l') dl',$$

where the lower limit of integration is arbitrary. Then

$$ds^2 = d\rho^2 + \rho^2 d\omega^2 + (1 - \kappa\rho \cos \vartheta)^2 dl^2,$$

and close to the magnetic axis, where the term $\kappa\rho \cos \vartheta$ can be neglected, the variables (ρ, ω, l) are equivalent to ordinary cylindrical coordinates locally aligned with the axis. We thus express the magnetic field as

$$\mathbf{B} = B_\rho \hat{\mathbf{e}}_\rho + B_\omega \hat{\mathbf{e}}_\omega + B_l \hat{\mathbf{e}}_l,$$

where

$$\hat{\mathbf{e}}_\alpha = \frac{\partial \mathbf{r}}{\partial \alpha} \bigg/ \left| \frac{\partial \mathbf{r}}{\partial \alpha} \right|, \quad \alpha = (\rho, \omega, l)$$

denote the coordinate unit vectors. The components of \mathbf{B} can be expanded as

$$\begin{aligned} B_\rho(\rho, \omega, l) &= a_1(\omega, l)\rho + O(\rho^2), \\ B_\omega(\rho, \omega, l) &= b_1(\omega, l)\rho + O(\rho^2), \\ B_l(\rho, \omega, l) &= B_0(l) + O(\rho), \end{aligned}$$

where numerical indices indicate the order in ρ . The current density on the axis becomes

$$\mu_0 J = \hat{\mathbf{e}}_l \cdot (\nabla \times \mathbf{B}) = \frac{1}{\rho} \frac{\partial(\rho B_\omega)}{\partial \rho} - \frac{1}{\rho} \frac{\partial B_\rho}{\partial \omega} = 2b_1 - \frac{\partial a_1}{\partial \omega}, \quad (39)$$

where we have used the expression for the curl in cylindrical coordinates.

Our remaining task is to relate this representation of the magnetic field to that using magnetic coordinates (12), which close to the axis reduces to $\mathbf{B} = \nabla\psi \times \nabla\alpha$ with

$$\psi = \psi_2(\omega, l)\rho^2 + O(\rho^3),$$

so that, to the requisite accuracy,

$$a_1 = \frac{\partial\psi_2}{\partial\omega} \frac{\partial\alpha}{\partial l} - \frac{\partial\psi_2}{\partial l} \frac{\partial\alpha}{\partial\omega}, \quad (40)$$

$$b_1 = -2\psi_2 \frac{\partial\alpha}{\partial l}, \quad (41)$$

$$B_0 = 2\psi_2 \frac{\partial\alpha}{\partial\omega}. \quad (42)$$

Hence

$$\frac{\partial}{\partial\omega} \left(\frac{b_1}{\psi_2} \right) + \frac{\partial}{\partial l} \left(\frac{B_0}{\psi_2} \right) = 0$$

and

$$\frac{\partial a_1}{\partial\omega} = -\frac{1}{2} \left[\frac{b_1}{\psi_2} \frac{\partial^2\psi_2}{\partial\omega^2} - \frac{\partial}{\partial l} \left(\frac{B_0}{\psi_2} \right) \frac{\partial\psi_2}{\partial\omega} + \frac{\partial}{\partial\omega} \left(\frac{B_0}{\psi_2} \frac{\partial\psi_2}{\partial l} \right) \right],$$

enabling us to write Eq. (39) as

$$b_1 = \frac{\mu_0 J}{2} - \frac{1}{4\psi_2} \left(b_1 \frac{\partial^2\psi_2}{\partial\omega^2} - B_0' \frac{\partial\psi_2}{\partial\omega} + B_0 \frac{\partial^2\psi_2}{\partial\omega\partial l} \right). \quad (43)$$

We proceed by noting that the level curves of ψ at fixed l must be elliptical near the magnetic axis, since the function ψ vanishes on the axis and is of order $O(\rho^2)$ in its vicinity. The general equation for an ellipse is

$$x = r e^{-\eta/2} \cos v,$$

$$y = r e^{\eta/2} \sin v,$$

where v varies between 0 and 2π , e^η denotes the ratio of the axes pointing in the y - and x -directions, respectively, and the area is πr^2 . If we let $d(l)$ denote the angle between the x -axis and the curvature vector \hat{e}_2 , as in Fig. (8), then

$$x = \rho \cos u,$$

$$y = \rho \sin u,$$

with

$$u = \vartheta + d(l) = \omega + d(l) - \int_0^l \tau dl',$$

and it follows that $y/x = \tan u = e^\eta \tan v$ as well as $r^2 = \rho^2(e^\eta \cos^2 u + e^{-\eta} \sin^2 u)$.

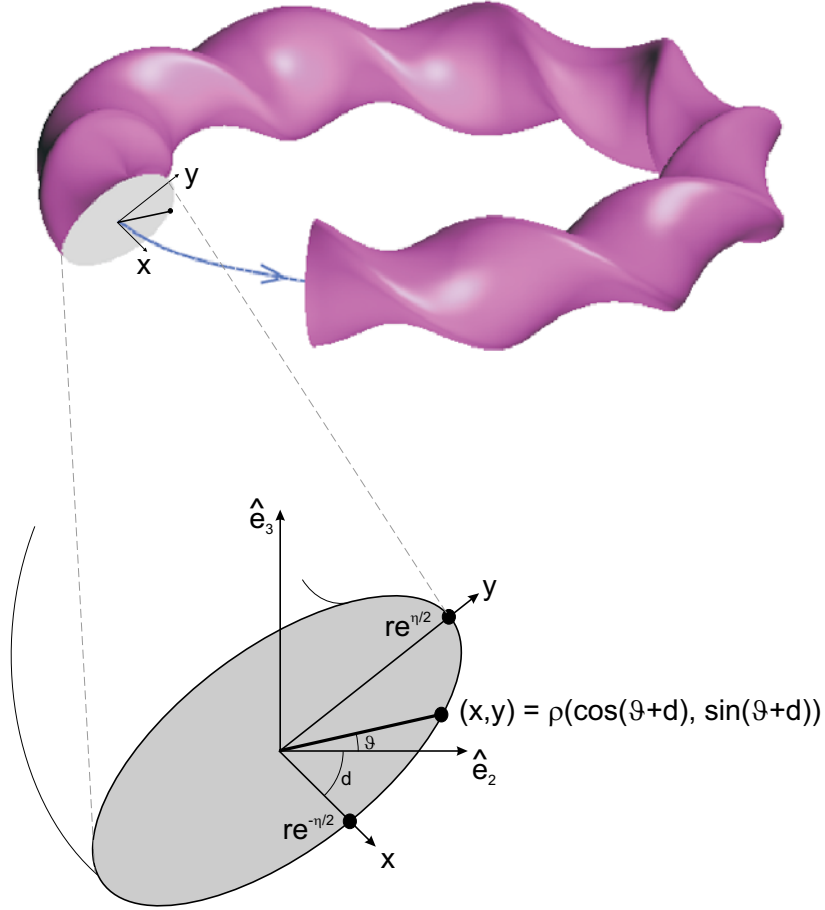


Figure 8: *Close to the magnetic axis, the flux surfaces have elliptical cross section. The angle between one of the semi-axes of the ellipse and the curvature vector is denoted by d . In general, this angle rotates as one moves along the magnetic axis around the torus, and this contributes to the rotational transform.*

From Eq. (9) we know that the toroidal magnetic flux is equal to $2\pi\psi \simeq \pi r^2 B_0(l)$, so we obtain the following expression for ψ close to the magnetic axis

$$\psi = \frac{B_0 r^2}{2} = \frac{B_0 \rho^2}{2} (\cosh \eta + \sinh \eta \cos 2u) .$$

It follows that

$$\begin{aligned}\frac{\partial\psi_2}{\partial\omega} &= -B_0 \sinh \eta \sin 2u, \\ \frac{\partial^2\psi_2}{\partial\omega^2} &= -2B_0 \sinh \eta \cos 2u,\end{aligned}$$

and Eq. (43) reduces to

$$\frac{b_1}{2\psi_2} = \frac{1}{\cosh \eta} \left[\frac{\mu_0 J}{2B_0} + \frac{1}{4\psi_2} \left(\frac{B'_0}{B_0} \frac{\partial\psi_2}{\partial\omega} - \frac{\partial^2\psi_2}{\partial\omega\partial l} \right) \right],$$

which in turn can be written as

$$\frac{b_1}{2\psi_2} = \frac{1}{\cosh \eta} \left(\frac{\mu_0 J}{2B_0} + d' - \tau \right) + \frac{(\eta'/2) \sin 2u - d' + \tau}{\cosh \eta + \sinh \eta \cos 2u},$$

and be used to calculate the magnetic angle α from Eq. (41),

$$\alpha(\rho, \omega, l) = \alpha_0 - \int_0^l \frac{b_1}{2\psi_2} dl'.$$

To do the integral, it is useful to recall that $v = \arctan(e^{-\eta} \tan u)$ and therefore

$$\left(\frac{\partial v}{\partial l} \right)_{\rho, \omega} = \frac{\partial v}{\partial \eta} \eta' + \frac{\partial v}{\partial u} (d' - \tau) = \frac{-(\eta'/2) \sin 2u + d' - \tau}{\cosh \eta + \sinh \eta \cos 2u}.$$

It thus follows that the Clebsch angle is equal to

$$\alpha(\rho, \omega, l) = \alpha_0 + v - \int_0^l \left(\frac{\mu_0 J}{2B_0} + d' - \tau \right) \frac{dl'}{\cosh \eta},$$

and we are now finally in a position to calculate the rotational transform ι .

To find out how many times a field line encircles the magnetic axis poloidally during one toroidal turn, we consider the difference in poloidal angle between the field line and the curve $\vartheta = 0$, which we shall call C . In a device with a circular magnetic axis, such as a tokamak, the curve C describes a circle on the inboard side of the flux surface. More generally, in any configuration with a plane magnetic axis, C lies in the same plane and does not encircle the axis poloidally. Still more generally, if the magnetic axis does not “wobble” too much, C will still not encircle the axis poloidally, but may do so an integer number of times, N , if the magnetic axis is very

twisted. If we go around the torus once toroidally, the angle d , which measures the rotation of the elliptical flux surface cross section must increase by an amount

$$\int_0^L d'(l)dl = m\pi,$$

where m is an integer and L the length of the magnetic axis. Thus, if we follow C once around the torus, the angle $u = \theta + d$ also increases by $m\pi$, and so does therefore v . The Clebsch angle $\alpha = \theta - \iota\varphi$ thus increases by the amount

$$\Delta\alpha = \alpha(L) - \alpha(0) = m\pi - \int_0^L \left(\frac{\mu_0 J}{2B_0} + d' - \tau \right) \frac{dl}{\cosh \eta},$$

when following C once around the torus. Relative to C , a magnetic field line has a poloidal twist equal to $-\Delta\alpha$, and since C itself rotates N times poloidally we conclude that the rotational transform is equal to

$$\iota = \frac{1}{2\pi} \int_0^L \left[\frac{\mu_0 J}{2B_0} - (\cosh \eta - 1)d' - \tau \right] \frac{dl}{\cosh \eta} - N. \quad (44)$$

Looking at this equation, we conclude that a rotational transform can be produced in three different ways: through an electric current, by rotating the poloidal cross section of the flux tube, and by torsion of the magnetic axis. Tokamaks and reversed field pinches use the first method, and stellarators usually rely on the latter two methods. The last method alone was employed by the first stellarator built in Princeton, which had circular cross section and the magnetic axis bent into the form of a figure eight. This method, which relies on torsion alone, is mathematically analogous to the famous ‘‘Berry phase’’ in quantum mechanics [20].

To understand why a rotational transform can be produced without a toroidal current, it is perhaps helpful to think of water flowing through a pipe. This is, in fact, more than a mental picture, because Euler’s equation for the steady flow of an incompressible fluid,

$$\mathbf{v} \cdot \nabla \mathbf{v} = -\nabla P,$$

where P denotes the pressure divided by the density, can be written as

$$(\nabla \times \mathbf{v}) \times \mathbf{v} = -\nabla Q,$$

with $Q = P + v^2/2$, and is thus similar to the MHD equilibrium condition (3), if we replace \mathbf{v} by \mathbf{B} . Every steady, incompressible flow thus corresponds

to an MHD equilibrium. Thus consider water flowing in a toroidally closed pipe, corresponding to a flux tube surrounding the magnetic axis. The rotational transform is equal to the average number of poloidal turns taken by a streamline during one toroidal revolution. Such a poloidally twisting flow can arise even in the absence of vorticity, $\nabla \times \mathbf{v} = 0$, if the cross section of the pipe rotates. This will cause the streamlines to rotate similarly, albeit with some slippage [21].

2.10 Variational principle

It is possible, and computationally advantageous, to formulate the equilibrium force-balance problem, $\mathbf{J} \times \mathbf{B} = \nabla p$, as a variational principle. That is, one can show that solutions to this equation are states of minimum MHD energy

$$W = \int \left(\frac{p}{\gamma - 1} + \frac{B^2}{2\mu_0} \right) dV, \quad (45)$$

subject to various constraints. The first term represents the thermal energy of the plasma (γ is the adiabatic index, which is 5/3 for a mono-atomic gas) and the second term is the magnetic energy. To understand why there should be such a variational principle at all, we cite the original source [11]:

“Suppose that everywhere in a given rigid toroidal tube T with perfectly conducting walls there is a viscous perfectly conducting fluid with an adiabatic equation of state, and also a magnetic field tangent to the tube walls. Suppose that any heat generated by the viscosity is somehow magically removed, so that each element of fluid is isentropic. The system can then lose energy but not gain it, since there can be no energy flux through the walls.

Let the fluid be initially at rest. In general, it will not be in equilibrium and will start to move. As long as it moves it loses energy, so it must eventually come to rest in a state of less energy than its initial state. Clearly an initially resting state of minimum energy cannot start moving at all, and so must be in equilibrium, i.e., satisfy the magnetostatic equations.”

Of course, the minimum of W is $B = p = 0$, but this is not the state we are interested in. We must remember that during the motion, there are quantities that remain constant, and we should therefore seek the minimum energy state subject to the corresponding constraints. Because the fluid is assumed to be perfectly conducting, the magnetic field is frozen into it and cannot change its topology. Flux surfaces can change shape but remain otherwise intact. The total toroidal flux ψ enclosed by each flux surface is a conserved quantity, and so is the rotational transform $\iota(\psi)$. And because the

fluid is assumed to move adiabatically, the entropy density p/ρ^γ is conserved, so the total mass inside each flux surface $\psi = \psi_0$ is proportional to

$$\int_{\psi < \psi_0} p^{1/\gamma} dV$$

and remains constant.

We now formulate the variational principle, following Ref. [11] closely. Consider all functions $p(\mathbf{r})$, $\mathbf{B}(\mathbf{r})$ and $\psi(\mathbf{r})$ defined in a given toroidal (but not necessarily axisymmetric) domain T and satisfying the constraints

- (a) ψ has nested toroidal level surfaces ranging from the innermost one, $\psi = 0$, to $\psi = \psi_{\max}$ on the boundary;
- (b) $\nabla \cdot \mathbf{B} = 0$;
- (c) $\mathbf{B} \cdot \nabla \psi = 0$;
- (d) The toroidal magnetic flux (9) inside the toroid defined by each level curve $\psi = \psi_0 \in [0, \psi_{\max}]$ is equal to $2\pi\psi_0$;
- (e) The corresponding poloidal flux (10) is equal to $2\pi\chi(\psi)$;
- (f)

$$\int_{\psi < \psi_0} p^{1/\gamma} dV = M(\psi_0);$$

where $\chi(\psi)$ and $M(\psi)$ are fixed, given functions. Among all such triplets (p, \mathbf{B}, ψ) , the energy functional (45) is made stationary if, and only if, p is a function of ψ alone and $\mathbf{J} \times \mathbf{B} = \nabla p$, where $\mathbf{J} = \nabla \times \mathbf{B}/\mu_0$.

To prove this statement, we start by varying the pressure p . Since $dV = d\psi dS/|\nabla\psi|$, where dS is the area element on surfaces of constant ψ , the constraint (f) can be written as

$$M'(\psi_0) = \int_{\psi=\psi_0} \frac{p^{1/\gamma} dS}{|\nabla\psi|} \quad (46)$$

and implies that only variations satisfying

$$\int_{\psi=\psi_0} \frac{p^{1/\gamma-1} \delta p}{|\nabla\psi|} dS = 0$$

are allowed. For instance, if p_0 and p_1 are the pressures at two points \mathbf{r}_0 and \mathbf{r}_1 on the same ψ -surface, then

$$\delta p = [p_0^{1-1/\gamma} \delta(\mathbf{r} - \mathbf{r}_0) - p_1^{1-1/\gamma} \delta(\mathbf{r} - \mathbf{r}_1)] |\nabla\psi|$$

is an allowable variation. But then

$$\delta W = \int_0^{\psi_{\max}} d\psi \int \frac{\delta p}{\gamma - 1} \frac{dS}{|\nabla\psi|}$$

only vanishes if $p_0 = p_1$, and we conclude that p must be a flux function, $\mathbf{B} \cdot \nabla p = 0$.

To vary \mathbf{B} , we introduce arbitrary poloidal and toroidal angles (ϑ, φ) , and note that according to Eq. (7) it is possible to write

$$\mathbf{B} = \nabla\psi \times (\nabla\vartheta + \nabla\lambda - \iota\nabla\varphi),$$

where $\lambda(\psi, \vartheta, \varphi)$ is a single-valued function and $\iota = d\chi/d\psi$. This is a general representation of a magnetic field satisfying the constraint (c)-(e), and we can thus vary \mathbf{B} by varying λ and ψ separately. Varying λ gives

$$\begin{aligned} \mu_0\delta W &= \int \mathbf{B} \cdot (\nabla\psi \times \nabla\delta\lambda) dV = \int \nabla(\delta\lambda) \cdot (\mathbf{B} \times \nabla\psi) dV \\ &= - \int \delta\lambda \nabla\psi \cdot (\nabla \times \mathbf{B}) dV = 0, \end{aligned}$$

and implies $\mathbf{J} \cdot \nabla\psi = 0$.

Finally, we need to vary ψ , remembering the constraint (f). When ψ is varied the shape of the flux surface changes. The volume of the flux surface $\psi = c$ is equal to

$$V(c) = \int_0^c d\psi \int \frac{dS}{|\nabla\psi|},$$

and its variation is

$$\delta V(c) = - \int_{\psi=c} \frac{\delta\psi}{|\nabla\psi|} dS. \quad (47)$$

According to Eq. (46), the pressure can be written as

$$p(\psi) = \left(\frac{M'(\psi)}{V'(\psi)} \right)^\gamma,$$

and its variation at constant ψ is thus

$$\frac{\delta p(\psi)}{p(\psi)} = - \frac{\gamma \delta V'(\psi)}{V'(\psi)},$$

so that

$$\int \delta p dV = -\gamma \int \frac{p \delta V'(\psi)}{V'(\psi)} dV = -\gamma \int p \delta V' d\psi.$$

Integrating by parts and using Eq. (47) gives

$$\int \delta p \, dV = \gamma \int p'(\psi) \delta V \, d\psi = -\gamma \int p' \delta\psi \, dV.$$

The pressure variation at a constant point in space is equal to the sum of $\delta p(\psi)$ and $p'(\psi)\delta\psi$, where the second term arises because the surfaces of constant ψ change shape. The variation of W with respect to ψ thus becomes

$$\delta W = \int \left(\frac{p' \delta\psi + \delta p}{\gamma - 1} + \frac{\mathbf{B} \cdot \delta \mathbf{B}}{\mu_0} \right) dV = \int \left(-p' \delta\psi + \frac{\mathbf{B} \cdot (\nabla \delta\psi \times \nabla \alpha)}{\mu_0} \right) dV$$

where $\alpha = \vartheta + \lambda - \iota\varphi$. Integrating this expression by parts remembering that $\delta\psi = 0$ on the boundary gives

$$\delta W = \int \delta\psi \left[\frac{\nabla \cdot (\mathbf{B} \times \nabla \alpha)}{\mu_0} - p' \right] dV,$$

and it follows that $\delta W = 0$ implies

$$(\nabla \times \mathbf{B}) \cdot \nabla \alpha = \mu_0 p'(\psi).$$

Since $\mathbf{J} \times \mathbf{B} = (\mathbf{J} \cdot \nabla \alpha) \nabla \psi$ we finally conclude that $\mathbf{J} \times \mathbf{B} = \nabla p$. The proof of the variational principle is thus complete, and we conclude that a three-dimensional magnetic equilibrium is completely determined by the shape of the outermost flux surface and two radial profile functions, which were here chosen as $\chi(\psi)$ and $M(\psi)$, but could also be chosen differently, e.g., $\iota(\psi)$ and $p(\psi)$.

2.11 Rational surfaces, magnetic islands and chaotic field lines

The variational principle we have just proved ensures the existence of MHD equilibria [22], but does not guarantee that the pressure profile and the magnetic field are continuous or that the current is free from singularities. The existence of well-behaved equilibria is an open and notoriously difficult question since many decades [23]. In fact, if one insists on perfect magnetic surfaces, in many cases the current density turns out to be infinite on every magnetic surface where ι is a rational number.

To see where the troubles begin, we try solving Eq. (35) for the parallel current (34) by Fourier expanding the functions $1/B^2$ and u ,

$$u(\psi, \theta, \varphi) = \sum_{m,n} u_{mn}(\psi) e^{i(m\theta - n\varphi)},$$

$$\frac{1}{B^2} = \sum_{m,n} h_{mn}(\psi) e^{i(m\theta - n\varphi)},$$

using Boozer coordinates. This gives

$$(m\iota - n)u_{mn}(\psi) = -(nI + mG)h_{mn}(\psi),$$

which implies

$$u_{mn}(\psi) = -\frac{nI + mG}{m\iota - n} h_{mn}(\psi) + \Delta_{mn}\delta(\psi - \psi_{mn}), \quad (48)$$

where ψ_{mn} denotes the resonant surface where $\iota = n/m$, and the coefficient Δ_{mn} multiplying the delta function is arbitrary. Note the mathematical origin of this delta function: the solution to the equation $xf(x) = 0$ in the space of generalised functions $f(x)$ is not $f(x) = 0$ but $f(x) = c\delta(x)$, where c is an arbitrary constant.

There are apparently two possible kinds of singularity on each rational surface, a surface current represented by the delta function, to which we return later, and a divergent Pfirsch-Schlüter current,

$$J_{\parallel} \sim \frac{h_{mn}}{\psi - \psi_{mn}} \frac{dp}{d\psi}. \quad (49)$$

At least one of the quantities $p'(\psi)$ or $h_{mn}(\psi)$ must therefore vanish on the rational surface lest there be an infinite Pfirsch-Schlüter current flowing in either direction. That the product $p'(\psi)h_{mn}(\psi)$ indeed vanishes on the rational surface can be seen from the relation

$$\begin{aligned} h_{mn} &= \frac{1}{4\pi^2} \int_0^{2\pi} d\theta \int_0^{2\pi} e^{-i(m\theta - n\varphi)} \frac{d\varphi}{B^2} \\ &= \frac{1}{4\pi^2} \int_0^{2\pi} e^{-im\alpha} d\alpha \int_0^{2\pi} e^{-i(m\iota - n)\varphi} \frac{d\varphi}{B^2}, \end{aligned}$$

where the last integral can be written as

$$\int_0^{2\pi} e^{-i(m\iota - n)\varphi} \frac{d\varphi}{B^2} = \frac{1}{\iota I + G} \int e^{-i(m\iota - n)\varphi} \frac{dl}{B}$$

and thus becomes independent of α on the resonant surface unless $p'(\psi_{mn}) = 0$ according to Eq. (32). Another way of stating this result is to say that magnetic fields with $h_{mn}(\psi_{mn}) \neq 0$ are unable to support a finite pressure gradient at $\psi = \psi_{mn}$.

One physical reason why they are unable to do so is that the *transport* becomes infinite if the Pfirsch-Schlüter current diverges, because the dissipation associated with this current due to finite resistivity then becomes infinite [15]. For instance, if the current satisfies the simple Ohm's law

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = \eta \mathbf{J},$$

then scalar multiplication by \mathbf{J} and flux-surface averaging gives

$$\langle \mathbf{v} \cdot \nabla p \rangle = - \langle \eta J^2 \rangle,$$

where we have used $\mathbf{J} \times \mathbf{B} = \nabla p$ and $\langle \mathbf{J} \cdot \mathbf{E} \rangle = - \langle \mathbf{J} \cdot \nabla \phi \rangle = 0$ in steady state. Hence and from Eq. (49) it follows that the classical transport (associated with classical transport from Coulomb collisions and gyromotion alone) becomes infinite, with a diffusion coefficient diverging as

$$D \sim \left| \frac{h_{mn}}{\psi - \psi_{mn}} \right|^2.$$

The pressure gradient is thus proportional to $(\psi - \psi_{mn})^2$ in the neighbourhood of the rational surface unless $h_{mn}(\psi_{mn}) = 0$.

Another way of removing the singularities in Eq. (48) is to relinquish the requirement of nested flux surfaces, and allow magnetic islands to form at rational surfaces. These are widely discussed in the literature, but for completeness we give a brief account of the essentials. We first note that there is a general representation of the magnetic field that has the same form as Eq. (7) but does not require the existence of magnetic surfaces. If (r, θ, φ) are general coordinates, and

$$\mathbf{A} = A_r \nabla r + A_\theta \nabla \theta + A_\varphi \nabla \varphi,$$

is the covariant representation of the magnetic potential, we introduce

$$g(r, \theta, \varphi) = \int_{r_0}^r A_r(r', \theta, \varphi) dr',$$

$$\psi = A_\theta - \frac{\partial g}{\partial \theta},$$

$$\chi = -A_\varphi + \frac{\partial g}{\partial \varphi},$$

where r_0 is arbitrary. Then

$$\mathbf{A} = \psi \nabla \theta - \chi \nabla \varphi + \nabla g, \tag{50}$$

and

$$\mathbf{B} = \nabla \times \mathbf{A} = \nabla\psi \times \nabla\theta + \nabla\varphi \times \nabla\chi \quad (51)$$

becomes superficially equal to Eq. (7). If χ can be written as a function of ψ alone, then nested flux surfaces exist (since then $\mathbf{B} \cdot \nabla\psi = 0$) and Eq. (51) is entirely equivalent to Eq. (7). However, the representation (51) is completely general and holds even if magnetic surfaces do not exist. It implies that the equations for the field lines are Hamiltonian,

$$\begin{aligned} \frac{d\psi}{d\varphi} &= \frac{\mathbf{B} \cdot \nabla\psi}{\mathbf{B} \cdot \nabla\varphi} = -\frac{\partial\chi}{\partial\theta}, \\ \frac{d\theta}{d\varphi} &= \frac{\mathbf{B} \cdot \nabla\theta}{\mathbf{B} \cdot \nabla\varphi} = \frac{\partial\chi}{\partial\psi}. \end{aligned}$$

The fact that magnetic field lines describe a Hamiltonian system can also be realised from the fact that they satisfy a variational principle. This follows from the circumstance that a charged particle moves in such a way as to make the action

$$S = \int L_p dt$$

stationary, where in the Lagrangian

$$L_p = \frac{mv^2}{2} + Ze\mathbf{A} \cdot \mathbf{v} - Ze\phi, \quad (52)$$

Ze denotes the charge, m the mass, v the velocity and ϕ the electrostatic potential. In the limit $m/Ze \rightarrow 0$ and $\phi = 0$, the particle exactly follows the magnetic field, and the action reduces to

$$S \rightarrow Ze \int \mathbf{A} \cdot d\mathbf{r} = Ze \int (\psi d\theta - \chi d\varphi), \quad (53)$$

where we have used Eq. (50), ignoring the term ∇g , which only adds an unimportant constant to S and does not contribute to its variation δS , if the end points of the integral are fixed. Equation (53) is similar to

$$S = \int (pdq - Hdt),$$

so it is not surprising that the equations for the field lines are Hamiltonian.

Hamiltonian systems are generally chaotic, and so are, therefore, magnetic fields if the ‘‘Hamiltonian’’ $\chi(\psi, \theta, \varphi)$ depends on all three coordinates. If the Hamiltonian is Fourier decomposed,

$$\chi(\psi, \theta, \varphi) = \chi_0(\psi) + \sum_{m,n \neq 0} \chi_{mn}(\psi) e^{i(m\theta - n\varphi)},$$

the field-line equations become

$$\frac{d\theta}{d\varphi} = \iota(\psi) + \sum_{m,n \neq 0} \chi'_{mn}(\psi) e^{i(m\theta - n\varphi)},$$

$$\frac{d\psi}{d\varphi} = -i \sum_{m,n \neq 0} m \chi_{mn}(\psi) e^{i(m\theta - n\varphi)},$$

with $\iota = d\chi_0/d\psi$. If $\chi_0 \gg \chi_{mn}$, so that the magnetic field “almost” possesses good flux surfaces, then a straightforward expansion of the equations give the field lines as

$$\theta = \theta_0 + \iota(\psi_0)\varphi,$$

$$\psi = \psi_0 - \sum_{mn} \frac{m\chi_{mn}(\psi_0)}{m\iota(\psi_0) - n} e^{i[m\theta_0 + (m\iota(\psi_0) - n)\varphi]}.$$

This treatment obviously fails near rational surfaces, and also indicates that the Fourier components with $n/m = \iota(\psi)$ play a dominant role there. Discarding all other components gives the Hamiltonian

$$\chi(\psi, \theta, \varphi) = \chi_0(\psi) + f(\psi, \theta, \varphi),$$

where

$$f(\psi, \theta, \varphi) = \sum_{n/m = \iota(\psi_0)} \chi_{mn}(\psi) e^{i(m\theta - n\varphi)}.$$

If we instead of (ψ, θ) use ψ and $\alpha = \theta - \iota(\psi_0)\varphi$ as canonical coordinates, we need to replace the Hamiltonian by

$$H(\psi, \alpha, \varphi) = \chi(\psi, \theta, \varphi) - \iota(\psi_0)(\psi - \psi_0) = \chi_0(\psi) + f(\psi, \alpha) - \iota(\psi_0)(\psi - \psi_0),$$

where we have recognised that the function $f(\psi, \theta, \varphi)$ can be written as a function of only two variables, ψ and α . Apart from an unimportant constant, this Hamiltonian is approximately equal to

$$H(\psi, \alpha, \varphi) = \frac{\chi_0''(\psi_0)(\psi - \psi_0)^2}{2} + f(\psi_0, \alpha).$$

Being independent of “time”, φ , this Hamiltonian is clearly integrable and describes the formation of magnetic islands around the resonant surface in question. The shape of these islands depends on the shape of the “potential well”, $f(\psi_0, \alpha)$, and becomes sinusoidal if only a single pair of harmonics is kept in the Fourier series,

$$H(\psi, \alpha, \varphi) = \frac{\chi_0''(\psi_0)(\psi - \psi_0)^2}{2} + 2\chi_{mn} \cos m\alpha.$$

The system is then mathematically equivalent to an ordinary pendulum. Different values of H correspond to different field lines, and the separatrix of the magnetic island corresponds to $H = 2\chi_{mn}$, so that

$$\psi - \psi_0 = \pm \sqrt{\frac{4\chi_{mn}(1 - \cos m\alpha)}{\iota'(\psi_0)}}$$

on the separatrix. The width of the island becomes

$$\Delta\psi = \sqrt{\frac{32\chi_{mn}}{\iota'(\psi_0)}}.$$

If this width is much smaller than the distance to the next island, our neglect of the non-resonant Fourier components is justified. If the islands are wide enough to overlap, the treatment breaks down and the magnetic field is generally chaotic.

If the aspect ratio is large and the poloidal cross section circular, then $\psi'(r) = rB_\varphi$, $\iota'(\psi) = \iota s/r^2 B$, where s is the magnetic shear, $B_r = \delta B_r \sin m\alpha$, with

$$\delta B_r = \frac{2m\chi_{mn}}{rR},$$

and the radial width of the island becomes

$$\Delta r = \frac{\Delta\psi}{rB_\varphi} = 4\sqrt{\frac{rR}{ns} \frac{\delta B_r}{B_\varphi}}.$$

If a chain of well-defined islands exists, these form little “stellarators” in their own right, consisting of nested flux surfaces, each having a certain rotational transform (44). If a current were to be added in each island, with a magnitude chosen so as cancel the rotational transform and thus “unwind” the twisted magnetic field within the island, the latter would disappear. This is the origin of the delta-function singularity on the right of Eq. (48), which indicates what current is necessary to produce a magnetic field free from islands.

As already mentioned, magnetic fields that do not possess any (continuous) symmetry are, in general, chaotic, and do not possess nested flux surfaces everywhere in the plasma. A long-standing question is under what conditions such surfaces do exist, at least approximately, and how they can be computed numerically. The most expeditious way to do so is to use the variational principle from the previous subsection, which is the strategy employed by the widely used VMEC code [24]. If flux surfaces do not exist

throughout the plasma, the problem is much more complicated, but numerical codes have been constructed to calculate the magnetic field [25, 26]. Some of the complications are of deeply conceptual nature [23, 27], but notable progress has been made in recent years [28, 29].

Finally, it should be pointed out that the actual small-scale properties of the magnetic field is not a question of ideal MHD alone. On small enough scales, $\mathbf{J} \times \mathbf{B}$ does not necessarily equal ∇p , since other terms in the force balance relation (2) can become important. For instance, neoclassical transport can force the plasma flow velocity to change abruptly across a magnetic island. As we shall see below, the requirement that the cross-field transport should be ambipolar determines the radial electric field in most non-axisymmetric plasma equilibria. The $E \times B$ flow therefore changes rapidly across the separatrix of a magnetic island, making the viscosity appreciable in this region. This can lead to a back-reaction on the magnetic island, causing the island to shrink and the magnetic field to “heal” [30]. The extent to which the field has islands or chaotic regions can therefore depend decisively on non-ideal MHD effects.

3 Particle orbits

Having developed the necessary mathematical apparatus for describing plasma-confining, non-axisymmetric magnetic fields, we now turn to the question of how charged particles move in such fields. Since the mean free path of the plasma constituents is very long in a fusion plasma, a necessary condition for good confinement is that most collisionless orbits of particles starting within the plasma volume should remain in there. In the present section, we analyse this requirement mathematically.

3.1 Guiding-centre Lagrangian

Particle orbits are of paramount importance for plasma confinement. A *magnetised* plasma is characterised by the condition

$$\delta = \frac{\rho}{L} \ll 1, \quad (54)$$

for the largest gyroradius ρ (usually that of the ions) and smallest macroscopic scale length L of interest. When this condition is satisfied, particle orbits can be described in terms of the motion of guiding centres. The guiding centre of a charged particle is located at

$$\mathbf{R} = \mathbf{r} - \frac{\mathbf{b} \times \mathbf{v}}{\Omega},$$

where \mathbf{r} is the particle position, \mathbf{v} its velocity and $\Omega = ZeB/m$ its gyrofrequency. Guiding centres move along the magnetic field and drift relatively slowly across it, i.e.,

$$\dot{\mathbf{R}} = v_{\parallel} \mathbf{b} + \mathbf{v}_d,$$

where $|\mathbf{v}_d/v| \sim \delta$. On the particle level, the point of magnetic plasma confinement is to prevent the plasma particles from moving across the flux surfaces, and in a magnetised plasma this goal is thus ensured to zeroth order in δ .

To first order in δ , the guiding-centre motion is described by the gyro-average of the Lagrangian (52), which is equal to [31, 32]

$$L = \frac{m(\mathbf{b} \cdot \dot{\mathbf{R}})^2}{2} + Ze\mathbf{A} \cdot \dot{\mathbf{R}} - \mu B - Ze\phi, \quad (55)$$

where \mathbf{A} and ϕ are the electromagnetic potentials and $\mu = mv_{\perp}^2/2B$ is the magnetic moment, which is regarded as a constant of the motion. In reality it is not perfectly, but very nearly, conserved. More precisely, one can show that an “improved” invariant μ exists which is equal to $mv_{\perp}^2/2B$ in lowest order and remains constant to *all* orders in δ [33]; its variation is

$$\frac{\dot{\mu}}{\Omega\mu} \sim e^{-c_0/\delta},$$

where c_0 is a constant of order unity. The magnetic moment is an example of an adiabatic invariant, a quantity that is approximately conserved because of a separation of scales. In the present case, this separation of scales occurs if $\delta \ll 1$ and the time scale on which guiding-centre motion is considered is longer than the inverse cyclotron frequency.

The Lagrangian (55) is simply the gyro-average of the particle Lagrangian (a simple proof is given in Ref. [32]), and all space-dependent quantities are thus to be evaluated at the position \mathbf{R} . It provides an accurate description of the motion under three conditions: the plasma should be magnetised, $\delta \ll 1$, the electric field should be relatively small,

$$\frac{E}{vB} \sim \delta,$$

and the fields should vary slowly in time, $\partial/\partial t \sim \delta\Omega$. The Euler-Lagrange equations,

$$\frac{d\mathbf{P}}{dt} = \frac{\partial L}{\partial \mathbf{R}},$$

where $\mathbf{P} = \partial L / \partial \dot{\mathbf{R}}$, yield the usual drift equations found in every textbook on plasma physics. The Hamiltonian is defined as

$$H = \mathbf{P} \cdot \dot{\mathbf{R}} - L = \frac{mv_{\parallel}^2}{2} + \mu B + Ze\phi, \quad (56)$$

with $v_{\parallel} = \mathbf{b} \cdot \dot{\mathbf{R}}$, and is, as usual, equal to the energy. It is a conserved quantity for motion in steady-state fields.

3.2 Guiding-centre motion in a stationary field

The Lagrangian formulation makes it straightforward to calculate the guiding-centre motion in arbitrary coordinates. To describe the motion in a steady-state toroidal equilibrium, we use Clebsch coordinates (ψ, α, l) , where $\mathbf{B} = \nabla\psi \times \nabla\alpha$ and l is the arc length along the field. The choice of the origin is somewhat arbitrary; we may for example measure the length l along each field line starting from some surface that cuts the magnetic field at a non-zero angle, e.g., at a toroidal location $\varphi = \varphi_0$. It is clear that this construction only works locally, since after one turn around the torus the points would be multiply labelled.

Because $\partial \mathbf{R} / \partial l = \mathbf{b}$, the parallel velocity becomes

$$\mathbf{b} \cdot \dot{\mathbf{R}} = \mathbf{b} \cdot \left(\frac{\partial \mathbf{R}}{\partial l} \dot{l} + \frac{\partial \mathbf{R}}{\partial \psi} \dot{\psi} + \frac{\partial \mathbf{R}}{\partial \alpha} \dot{\alpha} \right) \simeq \dot{l},$$

since $\dot{\psi}$ and $\dot{\alpha}$ vanish to lowest order in δ . The Lagrangian (55) thus becomes

$$L = \frac{m\dot{l}^2}{2} - Ze\alpha\dot{\psi} - \mu B - Ze\phi,$$

where we have chosen $\mathbf{A} = -\alpha \nabla\psi$. Since L does not contain $\dot{\alpha}$, the canonical momentum conjugate to α vanishes, $p_{\alpha} = \partial L / \partial \dot{\alpha} = 0$, which implies $\dot{p}_{\alpha} = \partial L / \partial \alpha = 0$, or

$$\dot{\psi} = -\frac{\mu}{Ze} \frac{\partial B}{\partial \alpha} - \frac{\partial \phi}{\partial \alpha}.$$

We are interested in motion in a stationary field, so it is logical to use Eq. (56) to express v_{\parallel} as a function of (H, μ, \mathbf{R}) ,

$$v_{\parallel} = \pm \sqrt{\frac{2(H - \mu B - Ze\phi)}{m}}, \quad (57)$$

and we conclude that

$$\dot{\psi} = \frac{mv_{\parallel}}{Ze} \left(\frac{\partial v_{\parallel}}{\partial \alpha} \right)_{H, \mu, \psi, l}. \quad (58)$$

Similarly, from

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = -Ze\alpha,$$

and $\dot{p}_\psi = \partial L / \partial \psi$ we find

$$\dot{\alpha} = -\frac{mv_\parallel}{Ze} \left(\frac{\partial v_\parallel}{\partial \psi} \right)_{H,\mu,\alpha,l}. \quad (59)$$

These relations describe the first-order drift away from a field line (ψ, α) and become particularly useful if instead of the instantaneous motion we consider the accumulated drift over a finite time interval,

$$\Delta\psi = \int \dot{\psi} dt = \int \frac{\dot{\psi} dl}{v_\parallel}.$$

Here, we are treating the radial drift as a small perturbation on the zeroth-order motion along the magnetic field. If we consider the change in ψ accumulated along an unperturbed orbit from l_1 to l_2 , we obtain

$$\Delta\psi = \frac{1}{Ze} \frac{\partial \mathcal{J}}{\partial \alpha}, \quad (60)$$

$$\Delta\alpha = -\frac{1}{Ze} \frac{\partial \mathcal{J}}{\partial \psi}, \quad (61)$$

where

$$\mathcal{J}(H, \mu, \psi, \alpha) = \int_{l_1}^{l_2} mv_\parallel dl, \quad (62)$$

and the end points are held fixed when the derivatives are taken. The symmetry displayed by Eqs. (60) and (61) indicate that the time-averaged motion can be expressed in Hamiltonian form. This can, for instance, be accomplished by taking the appropriate average of the Lagrangian (55) and then forming the Hamiltonian for the average motion. There is a simple expression in terms of \mathcal{J} for the time t_{12} required for the particle to travel from l_1 to l_2 . Since $\partial v_\parallel / \partial H = 1/mv_\parallel$ we have

$$t_{12} = \left(\frac{\partial \mathcal{J}}{\partial H} \right)_{\mu,\psi,\alpha,l_1,l_2}.$$

The quantity \mathcal{J} plays a particularly important role for trapped particles. When the cross-field drift is neglected, these particles bounce back and forth along the magnetic field between points where $v_\parallel(H, \mu, \mathbf{R}) = 0$.

Superimposed on the bounce motion, there is a slow cross-field drift, so that the field-line coordinates (ψ, α) change slowly with time. But if we take t_1 and t_2 to be the bounce points, Eqs. (60) and (61) imply that $\mathcal{J}(H, \mu, \psi, \alpha)$ remains constant, since the amount it changes by over one period of the bounce motion vanishes,

$$\Delta\mathcal{J} = \frac{\partial\mathcal{J}}{\partial\psi}\Delta\psi + \frac{\partial\mathcal{J}}{\partial\alpha}\Delta\alpha = 0.$$

\mathcal{J} is usually called the parallel, or second, adiabatic invariant (μ being the first), and is approximately conserved also in non-stationary fields that evolve slowly. In fact, life on Earth would be difficult without this property. Cosmic rays consisting of highly energetic ions continuously bombard the magnetosphere and are trapped in its magnetic field. They undergo bounce motion and precession around the Earth, and the requirement that μ and \mathcal{J} should be constant prevents them from hitting the Earth's surface.

3.3 Orbit confinement in a torus

We are now in a position to discuss the confinement of particle orbits to first order in our expansion parameter δ . Our main concern is to determine whether the magnetic field provides confinement in the radial direction, i.e., whether $\Delta\psi$ is small.

In zeroth order, the particle orbits strictly follow the magnetic field lines and are either trapped or untrapped (circulating) depending on the magnetic moment. Trapped particles bounce back and forth between two points of equal magnetic field strength whilst circulating particles simply follow a field line around the torus. In addition to this motion there is the first-order drift, and this is far less harmful for circulating orbits than for trapped ones. In fact, it follows from Eq. (60) that the untrapped orbits are well confined to first order in δ . Consider, for instance, a circulating particle on a high-order rational surface, $\iota = n/m$, so that the zeroth-order orbit comes back to the starting point after n poloidal and m toroidal revolutions. An identical orbit is obtained by shifting the starting point poloidally by any multiple of $2\pi/n$. If l_1 and l_2 are chosen to be equal to the initial and final point (which coincide) in Eq. (62), it follows that

$$\frac{\Delta\mathcal{J}}{\Delta\alpha} = 0,$$

if $\Delta\alpha = 2\pi/n$. For large n , this should be a very good approximation to $\partial\mathcal{J}/\partial\alpha$, leading us to conclude that the radial displacement (60) must be

small. This can also be seen from the following expression for the guiding-centre drift in a stationary field,

$$\mathbf{v}_d = \frac{v_{\parallel}}{B} \nabla \times \left(\frac{v_{\parallel} \mathbf{B}}{\Omega} \right), \quad (63)$$

where $\Omega = ZeB/m$ and the differentiation is taken at constant H and μ . (This can either be proved from the Lagrangian formalism above [32] or be verified directly by comparing the resulting expression for $\mathbf{v} \cdot \nabla \psi$ and $\mathbf{v} \cdot \nabla \alpha$ with Eqs. (58) and (59) above.) We use Eq. (63) to evaluate

$$\Delta \psi = \int (\mathbf{v}_d \cdot \nabla \psi) \frac{dl}{v_{\parallel}} = \int \nabla \cdot \left(\frac{v_{\parallel} \mathbf{B} \times \nabla \psi}{\Omega} \right) \frac{dl}{B},$$

along a zeroth-order orbit many turns around the torus. It follows from Eq. (22) that this integral is proportional to the flux surface average

$$\left\langle \nabla \cdot \left(\frac{v_{\parallel} \mathbf{B} \times \nabla \psi}{\Omega} \right) \right\rangle = 0,$$

which vanishes according to Eq. (20). The radial drift of a circulating particle thus vanishes when averaged over a zeroth-order orbit.

Trapped orbits are, however, not so easily confined and constitute the Achilles heel of stellarators. The net radial drift of a trapped orbit between two consecutive bounce points, l_1 and l_2 , is given by Eq. (60), which is non-zero in general. When the integral (62) is taken between two bounce points, the result usually depends on α . A stellarator will only confine trapped orbits if special care is taken.

The simplest way to avoid the problem (mathematically speaking) would be to eliminate trapped orbits altogether. If we assume that the electrostatic potential is a flux function, which we shall later prove to the accuracy required, trapped orbits are absent if, and only if, the magnetic field is constant on each flux surface,

$$|\mathbf{B}| = B(\psi). \quad (64)$$

This is however a very restrictive condition that is generally impossible to satisfy. For instance, near the magnetic axis the pressure gradient vanishes and Eq. (24) thus implies $\nabla_{\perp} B = B\boldsymbol{\kappa}$. The magnetic field strength thus increases in the direction of the curvature vector and cannot therefore be independent of the poloidal angle unless $\boldsymbol{\kappa} = 0$. But the curvature must be finite *somewhere* along the magnetic axis, and it follows that the field cannot satisfy the condition (64) in its vicinity. For the same reason, quasipoloidal symmetry (defined below) is impossible.

Trapped orbits must thus exist, but would be perfectly confined if the cross-field drift vanishes in each point,

$$\mathbf{v}_d \cdot \nabla \psi = \frac{v_\perp^2/2 + v_\parallel^2}{\Omega} (\mathbf{b} \times \nabla \ln B) \cdot \nabla \psi = 0.$$

Such magnetic configurations are called *isodynamic* [34] and require the geodesic curvature (25) to vanish everywhere, so that all field lines are geodesics. Unfortunately, the class of such fields is very restricted in toroidal geometry. The proof is complicated and will not be reproduced here. It is, however, possible to try to reduce the geodesic curvature as much as possible, by making B almost constant in the direction $\mathbf{b} \times \nabla \psi$. Besides reducing the radial drift, this has the additional advantage of minimising the Pfirsch-Schlüter current (36) and thus making the Shafranov shift smaller. This is an important element of stellarator optimisation.

3.4 Quasisymmetry

The second most straightforward way to confine the trapped orbits is to use a magnetic field whose strength is independent of α , i.e.,

$$|\mathbf{B}| = B(\psi, l). \quad (65)$$

If the variation of $B = |\mathbf{B}|$ along the field is the same for all field lines on each flux surface, then the integral (62) with v_\parallel given by Eq. (57) will clearly be independent of α . (We are still assuming that the electrostatic potential does not vary within the flux surface.) Such a magnetic field is called *quasisymmetric* [35, 36]. In a quasisymmetric field, a guiding centre does not “know” what field line it is on, within a given flux surface; for a guiding centre “feels” the magnetic field strength, which regulates v_\perp through the constancy of $\mu = mv_\perp^2/2B$, and its gradient along \mathbf{B} , which controls the mirror force, $F_\parallel = -\mu \nabla_\parallel B$. If it drifts to a point on another field line with the same local values of B and $\nabla_\parallel B$, it cannot “tell the difference” as far as its motion along \mathbf{B} is concerned.

The use of the condition (65) requires some caution. We have already mentioned that the arc length l can only be used as a local coordinate, and we therefore do not expect to be able to write the magnetic field strength in the form (65) globally. What we require is rather that Eq. (65) holds locally to ensure that the integral (62) becomes independent of the field line along which it is taken. This is the case if the parallel variation of B is a function only of ψ and B ,

$$\mathbf{b} \cdot \nabla B = f(\psi, B),$$

which requires

$$(\nabla\psi \times \nabla B) \cdot \nabla(\mathbf{b} \cdot \nabla B) = 0. \quad (66)$$

To explore the implication of this requirement, we note that the vector $\mathbf{B} \times \nabla\psi$ is tangential to flux surfaces and can therefore be written as

$$\mathbf{B} \times \nabla\psi = E\nabla\psi \times \nabla B + F\mathbf{B}, \quad (67)$$

for some functions E and F . Taking the scalar product of this equation with ∇B gives

$$F = \frac{(\mathbf{B} \times \nabla\psi) \cdot \nabla B}{\mathbf{B} \cdot \nabla B},$$

while instead taking the vector product with \mathbf{B} yields $E = -B^2/\mathbf{B} \cdot \nabla B$. Finally, evaluating the divergence of Eq. (67) and using $\mathbf{J} \cdot \nabla\psi = 0$ gives

$$\mathbf{B} \cdot \nabla F = (\nabla\psi \times \nabla B) \cdot \nabla \left(\frac{B^2}{\mathbf{B} \cdot \nabla B} \right).$$

Hence and from Eq. (66) it follows that $\mathbf{B} \cdot \nabla F = 0$, so that F is constant along each field line. It must therefore be constant on flux surfaces with irrational rotational transform, since each field line on such a surface covers it densely. If ι is rational this is not the case, but F must, by continuity, still be constant since it is so on neighbouring irrational surfaces. Quasisymmetry thus means that the quantity

$$\frac{(\mathbf{B} \times \nabla\psi) \cdot \nabla B}{\mathbf{B} \cdot \nabla B} = F(\psi) \quad (68)$$

must be a flux function (if $\mathbf{J} \cdot \nabla\psi = 0$).

It is instructive to explore the implication of this condition in Boozer coordinates (ψ, θ, φ) . In terms of these coordinates, the magnetic field is given by (27) and Eq. (68) becomes

$$G \frac{\partial B}{\partial \theta} - I \frac{\partial B}{\partial \varphi} = F(\psi) \left(\iota \frac{\partial B}{\partial \theta} + \frac{\partial B}{\partial \varphi} \right). \quad (69)$$

If B is Fourier transformed,

$$B(\psi, \theta, \varphi) = \sum_{m,n} B_{m,n}(\psi) e^{i(m\theta - n\varphi)},$$

this requires that the condition

$$[mG + nI - F(m\iota - n)] B_{m,n} = 0$$

should be satisfied for all (m, n) , which requires either $B_{m,n} = 0$ or

$$F(\psi) = \frac{(m/n)G(\psi) + I(\psi)}{(m/n)\iota(\psi) - 1}.$$

Since the left-hand side is independent of m/n , this relation can only be satisfied for one particular value of this ratio, M/N say. The magnetic field variation over the flux surface then only contains the corresponding helicity

$$B = \sum_k B_{kM,kN}(\psi)e^{ik(M\theta - N\varphi)},$$

and thus can be written as $B = B(\psi, M\theta - N\varphi)$.

If $N = 0$, the field is *quasi-axisymmetric* [37], if $M = 0$ it is *quasipoloidally symmetric*, and if neither M nor N vanishes it is *quasihelically symmetric* [35]. In either of these cases, the magnetic field strength depends only on ψ and a particular linear combination of the Boozer angles [36, 38]. It is, in this sense, independent of one of the two angles. It should be emphasized that it is only the *magnitude* of \mathbf{B} that has this property; its *direction* will still in general depend on all three coordinates, see Fig. 3.4.

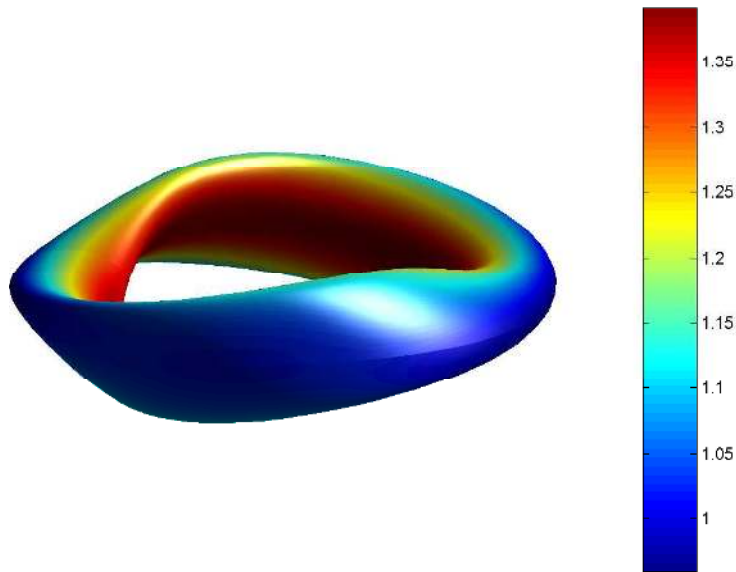


Figure 9: A flux surface in the quasi-axisymmetric stellarator design ESTELL [39]. The colours refer to the magnetic field strength.

The fact that the magnetic field strength only depends on one of the Boozer angles leads to a simple explanation why quasisymmetric configura-

rations confine guiding-centre orbits. In ordinary Boozer coordinates, the guiding-centre Lagrangian (55) is

$$L = \frac{m}{2B^2} (I\dot{\theta} + G\dot{\varphi})^2 + Ze(\psi\dot{\theta} - \chi\dot{\varphi}) - \mu B - Ze\phi, \quad (70)$$

where we have chosen $\mathbf{A} = \psi\nabla\theta - \chi\nabla\varphi$, see Eq. (7). If we replace the poloidal coordinate by $\vartheta = \theta - N\varphi/M$ we instead obtain

$$L = \frac{m}{2B^2} [I\dot{\vartheta} + (G + NI/M)\dot{\varphi}]^2 + Ze[\psi\dot{\vartheta} - (\chi - N\psi/M)\dot{\varphi}] - \mu B - Ze\phi. \quad (71)$$

Now, if B and ϕ are independent of φ , the canonical momentum conjugate to this angle,

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = \frac{m(G + NI/M)v_{\parallel}}{B} - Ze(\chi - N\psi/M) \quad (72)$$

is a constant of the motion,

$$\dot{p}_\varphi = \frac{\partial L}{\partial \varphi} = 0.$$

The point is the following. Guiding-centre motion in a magnetised plasma usually possesses two invariants, the energy H and the magnetic moment μ . A third constant of the motion is in general required for confining the orbits, and in a quasisymmetric field this is provided by the canonical momentum p_φ conjugate to the ignorable coordinate φ . The key is that the guiding-centre Lagrangian, when written in suitable coordinates, only depends on the strength of the magnetic field and not on its direction.

Since the term in Eq. (72) that is proportional to v_{\parallel} is a small correction (of order gyroradius over scale length) to the last term, the conservation of p_φ implies that a particle can only move a small distance radially. This conclusion is the basis for confinement in axisymmetric fields and is known as Tamm's theorem.

In a quasisymmetric field, the level contours of the field strength $B(\psi, M\theta - N\varphi)$ are, on each flux surface, straight lines in Boozer coordinates. Since the magnetic field lines are also straight, it is obvious that all field lines are equivalent in the sense expressed by Eq. (65). The level contours of B are actually straight not only in Boozer coordinates, but in a large class of other magnetic coordinates too, including Hamada coordinates. In fact, whenever a magnetic coordinate system (ψ, θ, φ) has a Jacobian (8) that is a function of ψ and B only,

$$(\nabla\psi \times \nabla B) \cdot \nabla\sqrt{g} = 0, \quad (73)$$

the quasisymmetry condition (68) implies that B only contains a single helicity in these coordinates, $B = B(\psi, M\theta - N\varphi)$. This can be seen as follows. Multiplying the co- and contravariant representations of \mathbf{B} , Eqs. (7) and (15), with each other gives

$$B^2 = \frac{1}{\sqrt{g}} \left(\iota I + G + \iota \frac{\partial H}{\partial \theta} + \frac{\partial H}{\partial \varphi} \right),$$

and Eq. (73) thus implies

$$(\nabla\psi \times \nabla B) \cdot \nabla H' = 0,$$

where we have denoted the derivative along \mathbf{B} by a prime. The function $H' = \mathbf{b} \cdot \nabla H$ therefore depends only on ψ and B , a property it shares with the function $B' = \mathbf{b} \cdot \nabla B$ according to Eq. (66). If we regard H as a function of the coordinates (ψ, α, B) , so that

$$H' = \frac{\partial H}{\partial B} B',$$

we thus have

$$0 = (\nabla\psi \times \nabla B) \cdot \nabla H' = -\frac{\partial^2 H}{\partial \alpha \partial B} B'^2.$$

At points where $B' \neq 0$ we thus require $\partial^2 H / \partial \alpha \partial B = 0$, which has the general solution

$$H(\psi, \alpha, B) = H_1(\psi, B) + H_2(\psi, \alpha), \quad (74)$$

where H_1 and H_2 are arbitrary functions. The function $H_2(\psi, \alpha)$ can be discarded on the grounds that it is constant along field lines and therefore on entire flux surfaces if the rotational transform is irrational (and, by continuity, also on rational surfaces). We thus conclude that

$$(\nabla\psi \times \nabla B) \cdot \nabla H = 0,$$

which in our magnetic coordinates (ψ, θ, φ) can be written as

$$\frac{\partial H}{\partial \theta} \frac{\partial B}{\partial \varphi} - \frac{\partial H}{\partial \varphi} \frac{\partial B}{\partial \theta} = 0.$$

But when expressing the quasisymmetry condition (68) in these coordinates, using

$$(\mathbf{B} \times \nabla\psi) \cdot \nabla B = \frac{1}{\sqrt{g}} \left(G \frac{\partial B}{\partial \theta} - I \frac{\partial B}{\partial \varphi} - \frac{\partial H}{\partial \theta} \frac{\partial B}{\partial \varphi} + \frac{\partial H}{\partial \varphi} \frac{\partial B}{\partial \theta} \right),$$

we then obtain exactly the same result (69) as when using Boozer coordinates, which implies the single-helicity condition, $B = B(\psi, M\theta - N\varphi)$. This relation thus holds not only in Boozer coordinates but in any magnetic coordinate system with the property (73).

Finally, we note that quasisymmetry can be formulated simply and succinctly as a condition that the magnetic field strength should be periodic,

$$B(\psi, \alpha, l + L) = B(\psi, \alpha, l),$$

with a period, $L = L(\psi)$, that is constant on each flux surface. To see this, consider a series of points on the same field line, a distance L apart from one another, i.e., the points

$$(\psi, \alpha, l_0 + nL), \quad n = 1, 2, 3, \dots,$$

The angular separation (in Boozer coordinates) between consecutive points is

$$\Delta\varphi = \int_{l_0}^{l_0+L} (\mathbf{B} \cdot \nabla\varphi) \frac{dl}{B} = \frac{1}{\iota I + G} \int_{l_0}^{l_0+L} B dl, \quad (75)$$

$$\Delta\theta = \iota \Delta\varphi,$$

and is independent of l_0 , because of the periodicity of $B(l)$. Hence it follows that the contours of constant field strength are straight lines in Boozer coordinates (if ι is irrational), which we have seen is equivalent to quasisymmetry.

3.5 Omnigenous fields

Quasisymmetry is a sufficient condition for the confinement of particle orbits, but it is not necessary. There is a much wider class of magnetic fields that have the property that the time-averaged radial drift vanishes,

$$\int (\mathbf{v}_d \cdot \nabla\psi) \frac{dl}{v_{\parallel}} = 0, \quad (76)$$

where the integral is taken between two bounce points for trapped particles. Such fields are called *omnigenous* [40], and quasisymmetry is thus a special case of omnigenicity (or omnigenity).

As already mentioned, the orbits that circulate around the torus automatically have the property that their time-averaged radial drift vanishes in lowest order. In an omnigenous field this is also the case for trapped orbits, which therefore, on the slow time scale of the bounce-averaged motion, *precess* from one field line to the next on the flux surface in question. As we

shall see in the next chapter, the electrostatic potential is usually approximately a flux function, $\phi \simeq \phi(\psi)$. In this case the bounce points are points of equal magnetic field strength, and the topology of the precession is then determined by the topology of the constant- B contours on the flux surface.

Since the net radial drift between two consecutive bounces of a trapped particle vanishes in an omnigenous field, Eq. (60) implies that $\partial\mathcal{J}/\partial\alpha = 0$, and the contours of constant \mathcal{J} must therefore coincide with flux surfaces. It is understood, of course, that the end points of the integration in the definition (62) of \mathcal{J} are chosen to coincide with the bounce points. When the \mathcal{J} -contours and the flux surfaces thus coincide, there are clearly two possibilities: the magnetic axis is either a maximum or a minimum of \mathcal{J} . Depending on the sign of $\partial\mathcal{J}/\partial\psi < 0$, the precession frequency

$$\omega_\alpha = \frac{\Delta\alpha}{\Delta t} = -\frac{1}{Ze\tau_b} \frac{\partial\mathcal{J}}{\partial\psi} \quad (77)$$

is either positive or negative. Here we have used Eq. (61) and denoted the bounce time by

$$\tau_b = \int_{l_1}^{l_2} \frac{dl}{v_{\parallel}}.$$

To characterise the direction of the precession, it is useful to compare it with that of the diamagnetic rotation $\omega_{dia} = \mathbf{v}_{dia} \cdot \nabla\alpha$, where $\mathbf{v}_{dia} = \mathbf{b} \times \nabla p_a / Zn_a e B$ is the diamagnetic velocity of the species a . (The diamagnetic velocity refers, of course, to the *fluid* and not to the particles of the species in question.) Hence

$$\omega_{dia} = \frac{p'_a(\psi)}{n_a Z e},$$

where the pressure gradient p'_a of the species in question is usually negative, and it follows that the precessional and diamagnetic frequencies have the same sign if $\partial\mathcal{J}/\partial\psi > 0$, and opposite signs otherwise.

Omnigenous configurations have a number of important properties [41]. One such property concerns the local minima of $B = |\mathbf{B}|$ along the magnetic field. A particle trapped in the immediate vicinity of such a minimum has a very small parallel velocity, and therefore $\mathcal{J} = 0$. Since \mathcal{J} is conserved as the particle drifts to neighbouring field lines, it continues to be trapped at the bottom of the magnetic well ($B = B_{\min}$) on each field line it reaches. Since the flux surface is an isopotential (by assumption), both the kinetic energy $E = mv^2/2$ and the magnetic moment $mv_{\perp}^2/2B$ are conserved, and since $v_{\parallel} = 0$, they are related by $E = \mu B$. It thus follows that B remains constant along the drift trajectory of the particle. Since the latter follows

the contour $B = B_{\min} = \text{const.}$ on the flux surface, it follows that all the B-field minima it visits are equal. The contours of minimum magnetic field strength are therefore poloidally, helically, or toroidally closed.

If the other contours of constant B are similarly closed, all trapped particles will precess around the torus according to these three topological possibilities, since the magnetic field strength at the bounce points E/μ is conserved during the precession. An omnigenous field where the contours close poloidally, rather than toroidally (as in the tokamak) or helically, is called *quasi-isodynamic* [42].

If the various constant- B contours close in topologically different ways, for instance, if the global maximum of B on the flux surface is a point rather than a curve, there will generally be particles that make a transition between trapped and circulating trajectories. When a trapped particle precesses to a field line where the maximum field strength equals the field strength at the bounce point, it will spend a very long time in the vicinity of this point and its bounce time formally becomes infinitely long, $\tau_b \rightarrow \infty$. The perturbative calculation of the precession (61) then fails, since it depends on an integration along a field line rather than an exact trajectory. The adiabatic invariance of \mathcal{J} is then broken, and it turns out that this quantity receives a quasi-random addition as the topology of the orbit changes. The addition is quasi-random because it depends on the phase of the particle along its orbit, and is thus not available within the bounce-averaged description. Thus, after many such orbit transitions (several precessions around the torus) the particle has essentially undergone a radial random walk. By definition, this is not allowed in an omnigenous field, where all constant- B contours must therefore be closed topologically in the same way.

Now let us consider the constancy of the normalised parallel invariant

$$\hat{\mathcal{J}}(\psi, B_b) = \frac{\mathcal{J}}{mv} \sqrt{\frac{\mu}{E}} = \int \sqrt{B_b - B(l)} dl$$

in greater mathematical detail [41]. The integral is taken between two consecutive bounce points and we have denoted the magnetic field strength at these points by $B_b = E/\mu$. For each value of the field strength B between B_{\min} and B_b there are (at least) two points l_k of equal field strength between the bounce points, so if the integration variable is changed from l to B we obtain

$$\hat{\mathcal{J}}(\psi, B_b) = \int_{B_{\min}}^{B_b} \sqrt{B_b - B} f(\psi, \alpha, B) dB,$$

with

$$f(\psi, \alpha, B) = \sum_k \left(\frac{1}{|\nabla_{\parallel} B|} \right)_{l=l_k}.$$

In an omnigenous field,

$$\frac{\partial \hat{J}}{\partial \alpha} = \int_{B_{\min}}^{B_b} \sqrt{B_b - B} \frac{\partial f}{\partial \alpha} dB = 0 \quad (78)$$

for all values of $B_b \in [B_{\min}, B_{\max}]$. To see what this implies for f , we multiply Eq. (78) by $1/\sqrt{X - B_b}$ and integrate over B_b from B_{\min} to some arbitrary field strength X between B_{\min} and B_{\max} . Interchanging the order of integration and using

$$\int_B^X \sqrt{\frac{B_b - B}{X - B_b}} dB_b = \frac{\pi(X - B)}{2}$$

gives

$$\int_{B_{\min}}^X \frac{\partial f}{\partial \alpha} (X - B) dB = 0.$$

Differentiating twice with respect to X finally gives

$$\frac{\partial f}{\partial \alpha} = \frac{\partial}{\partial \alpha} \sum_k \frac{\sigma_k}{\nabla_{\parallel} B(l_k)} = 0, \quad (79)$$

where $\sigma_k = \text{sign } \nabla_{\parallel} B(l_k)$. This result can be pictured as follows. Consider the variation of B along a field line. The field strength will typically equal a given value $B \in [B_{\min}, B_{\max}]$ in two points, l_- and l_+ , one at either sign of the minimum, see Fig. 3.5. The field strength will thus lie in the interval $[B, B + dB]$ in the regions $[l_-, l_- - dl_-]$ and $[l_+, l_+ + dl_+]$. The differential lengths dl_- and dl_+ are in general unequal, and are different on different field lines, but their sum is the same for all field lines on the same flux surface,

$$\frac{\partial(dl_- + dl_+)}{\partial \alpha} = 0,$$

since $\partial f/\partial \alpha = 0$.

This property implies that every integral of the form

$$\int_{B < B_b} F(B) dl \quad (80)$$

is independent of α , where $F(B)$ is an arbitrary function depending only on B and the integral is taken between two bounce points with fixed field

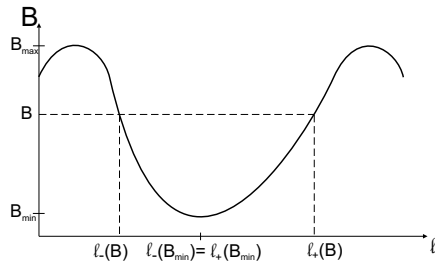


Figure 10: *The variation of magnetic field strength along a field line.*

strength B_b . A corollary is that *the field lines are of equal length*. i.e., the arc length measured along the field between two points of equal field strength, B_b , on either side of B_{\min} is the same for all field lines on a flux surface. So is, in fact, also the toroidal angular separation in Boozer coordinates (75). Specifically, taking $B_b = B_{\max}$ we find that the toroidal separation between two successive field maxima is independent of α . Since the corresponding poloidal separation is $\Delta\theta = \iota\Delta\varphi$, we conclude that the level contour $B = B_{\max}$ (at fixed ψ) is a straight line in Boozer coordinates (and in other magnetic coordinates whose Jacobian only depends on ψ and B). In quasisymmetric configurations all contours of B are straight, while in omnigenous ones only the maximum- B contours are in general straight.

3.6 Quasi-isodynamic fields

By definition, a quasi-isodynamic field is omnigenous and has poloidally, but not toroidally, closed contours of the magnetic field strength B . If the net toroidal current inside a particular flux surface ψ vanishes, we have already noted that the perpendiculars to \mathbf{B} also close poloidally on themselves, see Fig. 5. Since both these curves and the B_{\max} -contours are straight lines in Boozer coordinates, they must coincide. Therefore, within each flux surface, the field strength reaches its maximum on curves that are perpendicular to \mathbf{B} if the toroidal current vanishes.

In any quasi-isodynamic configuration, it is appropriate to use (ψ, α, B) as independent coordinates within each period [43, 44]. Of course, each B then corresponds to two points in each period, l_- and l_+ , which we shall refer to as the two *branches* of the field. The Jacobian is

$$(\nabla\psi \times \nabla\alpha) \cdot \nabla B = \mathbf{B} \cdot \nabla B, \quad (81)$$

and the covariant representation for the magnetic field is

$$\mathbf{B} = B_\psi \nabla \psi + B_\alpha \nabla \alpha + B_B \nabla B.$$

The components are obtained by taking the scalar product with the basis vectors, e.g.,

$$\mathbf{e}_\alpha = \left(\frac{\partial \mathbf{r}}{\partial \alpha} \right)_{\psi, B} = \frac{\nabla B \times \nabla \psi}{\mathbf{B} \cdot \nabla B}, \quad (82)$$

$$\mathbf{e}_B = \left(\frac{\partial \mathbf{r}}{\partial B} \right)_{\psi, \alpha} = \frac{\mathbf{B}}{\mathbf{B} \cdot \nabla B}, \quad (83)$$

so that

$$B_\alpha = \mathbf{B} \cdot \mathbf{e}_\alpha = -\frac{(\mathbf{B} \times \nabla \psi) \cdot \nabla B}{\mathbf{B} \cdot \nabla B}, \quad (84)$$

$$B_B = \frac{B^2}{\mathbf{B} \cdot \nabla B}.$$

Hence and from Eq. (68) we conclude that B_α is a flux function in quasisymmetric configurations.

In general omnigenous fields this is however not the case, but from the fact that the line integral of the magnetic field once around the torus in the poloidal direction is equal to μ_0 times the enclosed toroidal current (16)

$$\oint \mathbf{B} \cdot d\mathbf{r} = \oint \mathbf{B} \cdot \frac{\partial \mathbf{r}}{\partial \alpha} d\alpha = 2\pi I(\psi),$$

it follows that there must exist a function $h(\psi, \alpha, B)$ such that

$$B_\alpha = I(\psi) + \frac{\partial h}{\partial \alpha}, \quad (85)$$

The function h is thus defined only up to an additive constant, which could depend on the branch but which can be chosen so as to make h branch-independent,

$$h(\psi, \alpha, B)|_{l_-} = h(\psi, \alpha, B)|_{l_+}.$$

To see this, we note that the requirement that no current should cross the magnetic surface, $\nabla \psi \cdot (\nabla \times \mathbf{B}) = 0$, implies

$$\frac{\partial B_\alpha}{\partial B} - \frac{\partial B_B}{\partial \alpha} = 0,$$

which can be written as

$$\frac{\partial}{\partial \alpha} \left(\frac{B^2}{\mathbf{B} \cdot \nabla B} \right) = \frac{\partial^2 h}{\partial \alpha \partial B}.$$

Summing over the two branches and applying the result (79) thus gives

$$\sum_k \sigma_k \frac{\partial^2 h}{\partial \alpha \partial B} = 0,$$

which can be integrated in B to yield

$$\sum_k \sigma_k \frac{\partial h}{\partial \alpha} = 0,$$

and it follows that h can be chosen to be branch-independent. So is, then, also the function B_α according to Eq. (85).

This conclusion has an important consequence for the Pfirsch-Schlüter current (36). If we regard the function u entering in the latter as a function of ψ , α and B , we have according to Eqs. (35) and (84)

$$\frac{\partial u}{\partial B} = - \frac{(\mathbf{B} \times \nabla \psi) \cdot \nabla (1/B^2)}{\mathbf{B} \cdot \nabla B} = - \frac{2B_\alpha}{B^3}.$$

The difference in u measured at two points on the same field line where B has maximum and a minimum, respectively, is thus

$$u(\psi, \alpha, B_{\max}) - u(\psi, \alpha, B_{\min}) = - \int_{B_{\min}}^{B_{\max}} \frac{2B_\alpha dB}{B^3}.$$

Since B_α is branch-independent, it follows that u has the same value at two consecutive maxima. In other words, referring to Fig. 3.5, we can write

$$u(l_-(B_{\max})) = u(l_+(B_{\max})).$$

Thus, when following a field line around the torus, the function u always has the same value at the field maxima, and since these form poloidally closed curves, u is constant on these curves. Moreover, u must vanish on these curves since the Pfirsch-Schlüter plus diamagnetic currents do not carry any net toroidal current. (We assume that the net toroidal current inside the flux surface under consideration vanishes.) Finally, since the B_{\max} -curves are perpendicular to \mathbf{B} , the diamagnetic current (33) is parallel to these curves, so no current crosses them. In other words, *the streamlines of the current close upon themselves within each period of a quasi-isodynamic stellarator* [45]. This is very different from a tokamak, where the Pfirsch-Schlüter current has a dipolar structure, flowing in one toroidal direction (co-current) on the outboard side and in the other (counter-current) direction on the inboard side of the torus. It is this dipolar current that produces

most of the Shafranov shift, and quasi-isodynamic stellarators therefore have relatively small such shift. This is advantageous since it is the Shafranov shift that normally sets the equilibrium beta limit in stellarators.

The covariant component B_α of the field given by Eq. (84) is related to the radial orbit excursion. By using Eqs. (63) and (84), the radial drift can be written as

$$\mathbf{v}_d \cdot \nabla \psi = v_{\parallel} (\mathbf{b} \times \nabla \psi) \cdot \nabla \left(\frac{v_{\parallel}}{\Omega} \right) = -B_\alpha v_{\parallel} \nabla_{\parallel} \left(\frac{v_{\parallel}}{\Omega} \right). \quad (86)$$

Hence it follows that the net radial displacement of a particle moving along the field from l_0 to l is equal to

$$\Delta \psi = - \int_{l_0}^l B_\alpha \nabla_{\parallel} \left(\frac{v_{\parallel}}{\Omega} \right) dl'.$$

This result holds generally, but in a quasi-isodynamic field it is useful to employ Eqs. (84) and (85) to write

$$\Delta \psi = - \frac{I}{\Omega} [v_{\parallel}(l) - v_{\parallel}(l_0)] - \int_{B(l_0)}^{B(l)} \frac{\partial h}{\partial \alpha} \frac{\partial}{\partial B} \left(\frac{v_{\parallel}}{\Omega} \right) dB. \quad (87)$$

The corresponding equation in a tokamak field, $\mathbf{B} = F(\psi) \nabla \varphi + \iota \nabla \varphi \times \nabla \psi$, is

$$\Delta \psi = \frac{F}{\iota \Omega} [v_{\parallel}(l) - v_{\parallel}(l_0)],$$

as follows from the constancy of

$$p_\varphi = mR^2 \dot{\varphi} + ZeRA_\varphi = \frac{mFv_{\parallel}}{B} - Ze\chi.$$

We thus note that the first term on the right-hand side of Eq.(87) is ‘‘tokamak-like’’ whilst the second term is peculiar to stellarators, depending, as it were, on a non-zero α -derivative, i.e., on the non-equivalence of different field lines on the same flux surface. This is the only term present if the enclosed current vanishes, $I(\psi) = 0$, which is the normal situation in a stellarator without much bootstrap current.

An important difference between tokamaks and quasi-isodynamic stellarators follows from these relations. Let us consider trapped orbits and choose the reference point l_0 to be the bounce point, so that

$$\Delta \psi = - \frac{\partial}{\partial \alpha} \int_{B(l_0)}^{B(l)} h \frac{\partial}{\partial B} \left(\frac{v_{\parallel}}{\Omega} \right) dB$$

in a current-free quasi-isodynamic stellarator and

$$\Delta\psi = \frac{Fv_{\parallel}}{i\Omega}$$

in the tokamak. In the latter case, there is a relation between the sign of the parallel velocity and the radial displacement: $\Delta\psi > 0$ if a particle is moving in the direction where it makes a positive contribution to the current, and vice versa if $\Delta\psi < 0$. In a quasi-isodynamic stellarator, there is no such correlation, and a trapped particle with given $\sigma = v_{\parallel}/|v_{\parallel}|$ is just as likely to be displaced radially inward as outward (as measured from the location of the bounce point). As we shall see, this implies that the bootstrap current vanishes at low collisionality in a quasi-isodynamic stellarator.

3.7 Maximum- J configurations

We have already remarked that the sign of $\partial\mathcal{J}/\partial\psi$ determines the direction of the precessional drift. In so-called *maximum- J* configurations, where \mathcal{J} peaks on the axis and $\partial\mathcal{J}/\partial\psi < 0$, this drift is in the opposite direction to the diamagnetic drift for each species. This property is known to be beneficial for microstability [46]. For instance, the collisionless trapped-electron mode arises because of a resonance between the precession of trapped electrons and the electron drift wave, which is absent if these are in the opposite directions [47, 48].

We now demonstrate a simple property of maximum- J configurations. Consider the precession of deeply trapped particles in an omnigenous field. Substituting

$$\nabla B = \frac{\partial B}{\partial\psi}\nabla\psi + \frac{\partial B}{\partial\alpha}\nabla\alpha + \frac{\partial B}{\partial l}\nabla l,$$

where the last term vanishes at $B = B_{\min}$, in the expression for the magnetic drift gives

$$\mathbf{v}_d \cdot \nabla\alpha = \frac{v_{\perp}^2}{2\Omega B^2}(\nabla B \times \nabla\alpha) \cdot \mathbf{B} = \frac{v_{\perp}^2}{2\Omega} \frac{\partial B}{\partial\psi}.$$

Comparing with Eq. (77), we conclude that that the magnetic field strength needs to increase with radius,

$$\frac{\partial B}{\partial\psi} > 0,$$

in a maximum- J configuration, in order for the deeply trapped particles to precess in the correct direction. Since B also increases in the direction

along the field away from the minimum, $B = B_{\max}$, we conclude that B must have a local minimum on the magnetic axis, away from which the field strength increases in all directions. At the point of this minimum, Eq. (24) implies that the magnetic axis must have vanishing curvature, since the pressure gradient vanishes. Thanks to the diamagnetic effect of the plasma, it is much easier to achieve a minimum in the magnetic field strength, and the maximum- J property, in a plasma of finite pressure than in a vacuum magnetic field.

Another circumstance worth noting is that the maximum- J property of a magnetic configuration is not entirely independent of whether it possesses a magnetic well. The latter, it will be recalled, arises if the volume $V(\psi)$ inside the flux surface ψ has negative second derivative, $V''(\psi) < 0$, which is beneficial for ideal MHD stability. But the sign of $V''(\psi)$ coincides with that of the average precession, in the sense that the average is taken over velocity space and over the entire magnetic surface. Because of Eq. (63), the drift in the direction of $\nabla\alpha$ is

$$\mathbf{v}_d \cdot \nabla\alpha = \frac{v_{\parallel}}{\Omega} \nabla \cdot (v_{\parallel} \mathbf{b} \times \nabla\alpha),$$

and its average over the pitch angle $\xi = v_{\parallel}/v$ and the flux surface becomes

$$\left\langle \int_0^1 \mathbf{v}_d \cdot \nabla\alpha \, d\xi \right\rangle = \left\langle \int_0^{1/B} \frac{v_{\parallel}}{\Omega} \nabla \cdot (v_{\parallel} \mathbf{b} \times \nabla\alpha) \frac{B d\lambda}{2\sqrt{1-\lambda B}} \right\rangle,$$

where $\lambda = (1 - \xi^2)/B$. Hence

$$\begin{aligned} \left\langle \int_0^1 \mathbf{v}_d \cdot \nabla\alpha \, d\xi \right\rangle &= \frac{v^2 B}{3\Omega} \left\langle \nabla \cdot \left(\frac{\mathbf{B} \times \nabla\alpha}{B^2} \right) \right\rangle \\ &= \frac{v^2 B}{3\Omega V'} \frac{d}{d\psi} V' \left\langle \frac{\mathbf{B} \times \nabla\alpha}{B^2} \cdot \nabla\psi \right\rangle = -\frac{v^2 B}{3\Omega} \frac{V''}{V'}, \end{aligned}$$

where we have used Eq. (20), and it follows that the average of $\mathbf{v}_d \cdot \nabla\alpha$ and $V''(\psi)$ have opposite signs. In maximum- J configurations, $\mathbf{v}_d \cdot \nabla\alpha$ is positive for all *trapped* particles on a bounce-average; in magnetic-well configurations its average over *all* particles is positive. The question whether a magnetic field satisfies the maximum- J condition is thus related (but not identical) to the question whether it has a magnetic well.

3.8 Negative results

We have saved the bad news for last in this section: neither quasisymmetric nor omnigenous magnetic fields exist. Garren and Boozer [49] showed that it

is possible to satisfy the requirements of quasisymmetry on one flux surface but that neighbouring ones will in general not be quasisymmetric. This negative result was established by an expansion in the distance from the magnetic axis and comparing the number of free parameters with the number of constraints that need to be satisfied. It is possible to satisfy these up to second order in the inverse aspect ratio, but quasisymmetry is broken in third order. However, since this happens only in third order and stellarators tend to have large aspect ratio, this result is probably not very serious in practice. Very good approximations to quasisymmetry throughout an entire toroidal volume have been constructed numerically, e.g., by enforcing quasihelically symmetry at the plasma edge [50]. One quasihelical device, HSX, has also been built and found to confirm several theoretical predictions concerning the effect of quasi-helical symmetry on plasma behaviour [51].

The second negative result concerns quasi-isodynamic magnetic fields, which can be shown not to exist unless they are quasi-poloidally symmetric. For a simple proof, we recall that it is possible to construct magnetic coordinates for any choice of the toroidal angle φ . For a quasi-isodynamic field, it is convenient to choose this angle so that B is independent of the poloidal angle [52],

$$|\mathbf{B}| = B(\psi, \varphi).$$

Recalling the Jacobian (8) and the co- and contravariant representations (7) and (15), we quickly find that

$$B^2 = (\iota I + G)\mathbf{B} \cdot \nabla\varphi + \mathbf{B} \cdot \nabla H,$$

which we can use to evaluate integrals of the form (80),

$$I(\alpha) = \int_{\varphi_0}^{\varphi_1} F(B) \frac{B d\varphi}{\mathbf{B} \cdot \nabla\varphi},$$

taken along a field line labelled by $\alpha = \theta - \iota\varphi$ between two points with the same field strength, $B(\varphi_0) = B(\varphi_1)$. We have already proved that any such integral is independent of α , but

$$I(\alpha) = \int_{\varphi_0}^{\varphi_1} \frac{F(B)}{B} \left(\iota I + G + \frac{\mathbf{B} \cdot \nabla H}{\mathbf{B} \cdot \nabla\varphi} \right) d\varphi,$$

where

$$\frac{\mathbf{B} \cdot \nabla H}{\mathbf{B} \cdot \nabla\varphi} = \left(\frac{\partial H}{\partial \varphi} \right)_{\psi, \alpha},$$

and it follows that

$$0 = \frac{dI}{d\alpha} = \int_{\varphi_0}^{\varphi_1} \frac{F(B)}{B} \frac{\partial^2 H}{\partial \varphi \partial \alpha} d\varphi.$$

This condition can only hold for all $F(B)$ if $\partial^2 H / \partial \varphi \partial \alpha = 0$, i.e., if H is of the form (74),

$$H(\psi, \alpha, \varphi) = H_1(\psi, \varphi) + H_2(\psi, \alpha).$$

By the same argument as that following Eq. (74) it follows that the field must be quasi-poloidally symmetric. Only such fields are thus quasi-isodynamic, and, more generally, only quasisymmetric fields are omnigenous. Fortunately this “negative” result seems not to be very serious. As argued already by Cary and Shasharina [41], who first established it, it appears possible to construct fields that are arbitrarily close to being omnigenous but are still very far from quasisymmetric. That it is indeed possible to construct fields that are omnigenous to a high degree of approximation is seen from Fig. 3.8, taken from Ref. [45], which shows contour levels of \mathcal{J} for a range of different values of the pitch-angle variable $B_b = E/\mu$ measuring the magnetic field strength at the bounce points. These values are chosen as $B_b = B_{\min} + n(B_{\max} - B_{\min})$ with n ranging from 1 to 6.

4 Kinetic theory

We now turn to the kinetic theory of non-axisymmetric plasma equilibria. We are only interested in length scales exceeding the ion gyroradius and therefore consider the drift kinetic equation for the distribution function f_a of each species a . The form of this equation depends decisively on the magnitude of the electric field. In the so-called drift ordering,

$$E/B = O(\delta v_{Ta}),$$

where $v_{Ta} = (2T_a/m_a)^{1/2}$ denotes the thermal velocity, the drift kinetic equation has the familiar appearance

$$\frac{\partial f_a}{\partial t} + (v_{\parallel} \mathbf{b} + \mathbf{v}_{da}) \cdot \nabla f_a = C_a(f_a), \quad (88)$$

where the independent variables are the guiding-centre position, the energy (Hamiltonian) $H = m_a v^2/2 + e_a \phi$, and the magnetic moment $\mu = m_a v_{\perp}^2/2B$. In the MHD ordering, however,

$$E/B = O(v_{Ta}),$$

the drift kinetic equation is more complicated. Our first task is therefore to determine which ordering is appropriate. As already mentioned following Eq. (3), it turns out that the MHD ordering is not possible in stellarators.

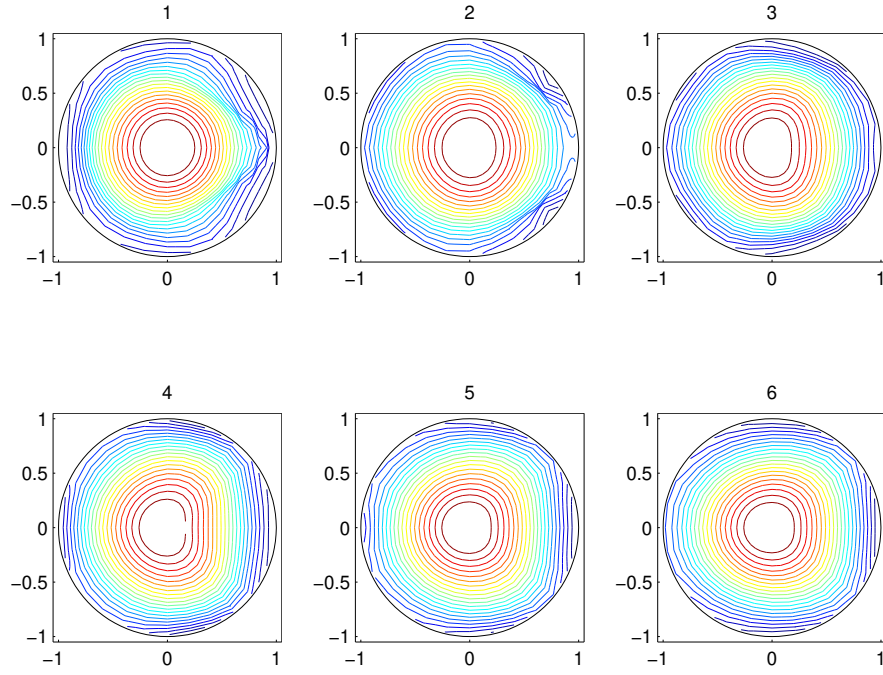


Figure 11: *Curves of constant \mathcal{J} for a quasi-isodynamically optimised stellarator [45] in a polar coordinate representation (\sqrt{s}, θ) , where s is the toroidal flux normalised to its value at the plasma edge and θ is the poloidal Boozer angle. The \mathcal{J} -contours are approximately aligned with the flux surfaces, which are circular in this representation. The six figures refer to different values of $n = (E/\mu - B_{\min})/(B_{\max} - B_{\min})$ ranging from 1 to 6.*

4.1 Impossibility of rapid plasma rotation

To prove this assertion [53], we explore the consequences of the MHD ordering, assuming that the $E \times B$ flow is comparable to thermal speed of the bulk ions. Their Larmor orbits then do not close in the the laboratory rest frame and the magnetic moment $m_a v_\perp^2 / 2B$ is not conserved. Moreover, it is straightforward to show that the plasma fluid generally has a lowest-order flow speed \mathbf{V} comparable to the ion thermal speed, approximately satisfying

$$\mathbf{E} + \mathbf{V} \times \mathbf{B} = 0. \quad (89)$$

This conclusion follows from the Fokker-Planck equation,

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \frac{e}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f}{\partial \mathbf{v}} = C(f),$$

using the shifted velocity $\mathbf{u} = \mathbf{v} - \mathbf{V}$ as an independent variable,

$$\frac{\partial f}{\partial t} + (\mathbf{V} + \mathbf{u}) \cdot \nabla f + \frac{e}{m} \left(\mathbf{E}' + \mathbf{u} \times \mathbf{B} - \frac{\partial \mathbf{V}}{\partial t} - (\mathbf{V} + \mathbf{u}) \cdot \nabla \mathbf{V} \right) \cdot \frac{\partial f}{\partial \mathbf{u}} = C(f), \quad (90)$$

where \mathbf{V} is arbitrary at the moment, and $\mathbf{E}' = \mathbf{E} + \mathbf{V} \times \mathbf{B}$ is the electric field in the moving frame, see, e.g., p. 77 in Ref. [32]. If the distribution function and the electric field are expanded in δ , we obtain in zeroth order

$$\frac{e}{m} (\mathbf{E}'_0 + \mathbf{u} \times \mathbf{B}) \cdot \frac{\partial f_0}{\partial \mathbf{u}} = 0. \quad (91)$$

This equation implies that f_0 is constant along characteristics in velocity space. In the direction of \mathbf{B} , these extend to infinity unless $\mathbf{b} \cdot \mathbf{E}'_0 = 0$, and we must therefore require $E_\parallel = 0$ in order that f_0 should vanish for infinitely large u_\parallel . Since we are interested in plasma equilibria ($\partial/\partial t = 0$), the electric field is at least approximately electrostatic, $\mathbf{E}_0 = -\nabla\phi_0$, and we conclude the potential is a flux function in lowest order, $\phi_0 = \phi_0(\psi)$, a result we have already used repeatedly in the previous chapter. We can thus choose \mathbf{V} so that $\mathbf{E}'_0 = 0$, namely by taking the perpendicular component to be $\mathbf{V}_\perp = \mathbf{B} \times \nabla\phi_0 / B^2$. Equation (91) then implies that f_0 is independent of the gyro-angle and that the plasma flow velocity approximately equals \mathbf{V} in the direction perpendicular to \mathbf{B} . \mathbf{V} can thus be chosen to be equal to the plasma flow velocity in lowest order.

The drift kinetic equation is derived from Eq. (90) by averaging over the gyro-angle [54, 55, 56, 57]. If the velocity space coordinates are taken to be $w = mu^2/2$ and $\mu = mu_\perp^2/2B$, the result is in lowest order

$$\frac{\partial f_0}{\partial t} + (u_\parallel \mathbf{b} + \mathbf{V}) \cdot \nabla f_0 + w \frac{\partial f_0}{\partial w} + \dot{\mu} \frac{\partial f_0}{\partial \mu} = C(f_0), \quad (92)$$

where $\dot{\mu} = 0$ and

$$\dot{w} = eE_{\parallel}u_{\parallel} - mu_{\parallel}\mathbf{V} \cdot \nabla\mathbf{V} \cdot \mathbf{b} - mu_{\parallel}^2\mathbf{b} \cdot \nabla\mathbf{V} \cdot \mathbf{b} + \mu B\mathbf{V} \cdot \nabla \ln B. \quad (93)$$

Thus, when defined relative to the moving frame, the magnetic moment μ is conserved, but the kinetic energy w varies in the complicated way given by Eq. (93).

We now look for stationary solutions of the drift kinetic equation (92). Of course, we expect these to be Maxwellian, which is proved by multiplying the equation by $\ln f_0$ and integrating over velocity space, which gives

$$\nabla \cdot \mathbf{Y} = - \int \ln f_0 C(f_0) 2\pi v_{\perp} dv_{\perp} dv_{\parallel}, \quad (94)$$

where

$$\mathbf{Y} = - \int (\mathbf{V} + v_{\parallel}\mathbf{b}) f_0 (\ln f_0 - 1) 2\pi v_{\perp} dv_{\perp} dv_{\parallel}$$

is the entropy flux. The left-hand side of Eq. (94) is annihilated by a flux-surface average, and it follows from the H-theorem [32] that f_0 must be a Maxwellian, whose density n and temperature T may however vary over each flux surface.

The next step is to substitute this Maxwellian into Eq. (92) without the time-derivative, which gives an equation that can only be satisfied if the following five relations are satisfied [56, 57]:

$$\begin{aligned} \left(\nabla \ln n + \frac{e\nabla\phi_1}{T} + \frac{m}{T}\mathbf{V} \cdot \nabla\mathbf{V} \right) \cdot \mathbf{b} &= 0, \\ \mathbf{b} \cdot \nabla T &= 0, \\ \mathbf{V} \cdot \nabla \left(\ln n - \frac{3}{2} \ln T \right) &= 0, \\ \nabla \cdot (n\mathbf{V}) &= 0, \\ \mathbf{b} \cdot \nabla\mathbf{V} \cdot \mathbf{b} - \frac{1}{3} \nabla \cdot \mathbf{V} &= 0. \end{aligned} \quad (95)$$

The second of these equations implies that irrational flux surfaces (and, by continuity, also rational ones) are isothermal. Since $\mathbf{V} \cdot \nabla\psi = 0$, the third equation thus implies $\mathbf{V} \cdot \nabla n = 0$. This reduces the fourth equation to an incompressibility condition,

$$\nabla \cdot \mathbf{V} = 0, \quad (96)$$

and the fifth one to

$$\mathbf{b} \cdot \nabla\mathbf{V} \cdot \mathbf{b} = 0. \quad (97)$$

We now recall Eq. (89) and note that

$$0 = \nabla \times (\mathbf{V} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \mathbf{V} - \mathbf{V} \cdot \nabla \mathbf{B}, \quad (98)$$

which combined with Eq. (97) leads to $\mathbf{V} \cdot \nabla \mathbf{B} \cdot \mathbf{B} = 0$. Since $(\nabla \mathbf{B}) \cdot \mathbf{B} = B \nabla B$ we thus conclude that

$$\mathbf{V} \cdot \nabla B = 0. \quad (99)$$

In other words, the streamlines of the flow are given by the intersection between flux surfaces and surfaces of constant B . This means that the velocity field can be written as

$$\mathbf{V}(\mathbf{r}) = g(\mathbf{r}) \nabla \psi \times \nabla B$$

for some function $g(\mathbf{r})$ of the spatial coordinates \mathbf{r} . The parallel component of the flow is thus

$$V_{\parallel} \mathbf{b} = g(\mathbf{r}) \nabla \psi \times \nabla B - \frac{d\phi_0}{d\psi} \frac{\mathbf{b} \times \nabla \psi}{B},$$

where $\mathbf{E}_0 = -\nabla \phi_0$. Taking the scalar product of this equation with $\mathbf{b} \times \nabla \psi$ gives an expression for g ,

$$g \mathbf{b} \cdot \nabla B + \frac{1}{B} \frac{d\phi_0}{d\psi} = 0,$$

and thus enables us to write down an explicit expression for the lowest-order flow velocity,

$$\mathbf{V} = -\frac{d\phi_0}{d\psi} \frac{\nabla \psi \times \nabla B}{\mathbf{B} \cdot \nabla B}. \quad (100)$$

The requirement (96) that this flow field should be incompressible now implies a constraint on the spatial variation of the magnetic field strength,

$$(\nabla \psi \times \nabla B) \cdot \nabla (\mathbf{B} \cdot \nabla B) = 0, \quad (101)$$

which is identical to Eq. (66) and therefore implies quasisymmetry. One should perhaps be slightly careful with this choice of words, because in the MHD ordering the equilibrium is described not by Eq. (3) but by

$$\rho \mathbf{V} \cdot \nabla \mathbf{V} = \mathbf{J} \times \mathbf{B} - \nabla p.$$

Since we defined Boozer coordinates assuming Eq. (3), we cannot take quasisymmetry to mean that B should be independent of one of the Boozer angles; instead we note that the large flow ordering implies quasisymmetry

in the sense defined by Eq. (68). In any case, it is clear that most magnetic fields cannot support rapid rotation.

It is remarkable that this conclusion follows from the drift kinetic equation already in zeroth order of the gyroradius expansion. The result is therefore robust against any effects of turbulence, regardless of its nature, as long as the turbulent fluctuations are small. For instance, the gyrokinetic equation, which is commonly used to describe turbulence, treats first-order (in δ) fluctuations. Physically, the point is that, in zeroth order, the transport is infinitely much faster within flux surfaces than across them. This implies that such surfaces must be isothermal and isopotential, and the parallel viscosity forces the flow velocity to be small (in comparison with the ion thermal speed) unless the magnetic field is quasisymmetric.

Moreover, rapid plasma rotation may be impossible even if the field *does* satisfy the quasisymmetry requirement (68). The ion drift kinetic equation then has a lowest-order Maxwellian solution, but the electrons also need to be considered and quasineutrality should be satisfied. The density distribution over the flux surface is governed by Eq. (95), where the last term represents the centrifugal force associated with the plasma rotation. This term can be rewritten using the fact that $\mathbf{V} \cdot \mathbf{B}$ is a flux function because of Eqs. (68) and (100),

$$\mathbf{V} \cdot \nabla \mathbf{V} \cdot \mathbf{B} = -\mathbf{V} \cdot \nabla \mathbf{B} \cdot \mathbf{V} \stackrel{(98)}{=} -\mathbf{B} \cdot \nabla \mathbf{V} \cdot \mathbf{V},$$

so that

$$\mathbf{b} \cdot \nabla \left(\ln n + \frac{e\phi_1}{T} - \frac{mV^2}{2T} \right) = 0.$$

For electrons, whose charge is $-e$ and thermal velocity exceeds V , this expression reduces to a Boltzmann relation, $n_e \propto \exp(e\phi_1/T_e)$, while for the ions

$$n_i \propto \exp \left(\frac{m_i V^2}{T_i} - \frac{e\phi_1}{T_i} \right).$$

Quasineutrality, $n_i = n_e$ serves to determine the potential variation ϕ_1 within the flux surface, and we find that the density becomes

$$n_e = n_i = N(\psi) \exp \left(\frac{m_i V^2}{2(T_e + T_i)} \right). \quad (102)$$

This expression is analogous to that in a toroidally spinning tokamak [56], where the density bulges out on the outboard side due to the centrifugal

force. In a quasisymmetric stellarator, it severely constrains the magnetic-field geometry.

To see why, let us again employ $\mathbf{z} = (\psi, \alpha, B)$ as our coordinates and recall the covariant basis vectors (82)-(83). The metric elements are defined as $g_{mn} = \mathbf{e}_m \cdot \mathbf{e}_n$. According to Eqs. (100) and (101), the vector \mathbf{e}_α points in the symmetry direction in which neither B nor $\mathbf{b} \cdot \nabla B$ varies, and it follows that

$$\frac{\partial g_{BB}}{\partial \alpha} = \frac{\partial}{\partial \alpha} \left(\frac{B}{\mathbf{b} \cdot \nabla B} \right) = 0.$$

Furthermore, since \mathbf{V} points in the direction of \mathbf{e}_α and $\nabla \cdot (n\mathbf{V}) = \mathbf{V} \cdot \nabla n = 0$, Eq. (102) implies $\partial V^2 / \partial \alpha = 0$ and thus

$$\frac{\partial g_{\alpha\alpha}}{\partial \alpha} = 0,$$

because of Eq. (100). Also, because

$$g_{\alpha B} = \mathbf{e}_\alpha \cdot \mathbf{e}_B = -\frac{(\mathbf{B} \times \nabla \psi) \cdot \nabla B}{(\mathbf{B} \cdot \nabla B)^2},$$

we have

$$\frac{\partial g_{\alpha B}}{\partial \alpha} = 0,$$

and we conclude that the metric elements g_{BB} , $g_{\alpha B}$ and $g_{\alpha\alpha}$ governing the Riemannian geometry within the magnetic surface are all independent of α . Using techniques from differential geometry, Sugama et al. [58] have demonstrated that, at least close to the magnetic axis, this implies that the surface is not only quasisymmetric, but also (axi)symmetric in the usual sense. The conclusion is that stellarator plasmas cannot rotate at velocities comparable to the ion thermal speed.

4.2 Neoclassical equilibrium and transport

We are thus free to use the drift kinetic equation (88) in the small-flow ordering. As we shall see, the transport is of second order in δ , so if we restrict our attention to plasma equilibria (time derivatives of order $\delta^2 v_T / L$), these should be described by the steady-state equation

$$(v_{\parallel} \mathbf{b} + \mathbf{v}_{da}) \cdot \nabla f_a = C_a(f_a).$$

Solving this equation is the aim of neoclassical transport theory [59, 60, 61, 62, 63], which we now outline.

4.2.1 Moderate collisionality

If the distribution function is expanded in $\delta \ll 1$ as $f_a = f_{a0} + f_{a1} + \dots$, then we obtain in lowest order

$$v_{\parallel} \nabla_{\parallel} f_{a0} = C_a(f_{a0}).$$

(Plasma regimes with sufficiently small collision frequency cannot be described by this procedure and will be discussed later.) We multiply this equation by $\ln f_{a0}$ and integrate over velocity space, using

$$d^3v = \sum_{\sigma} \frac{2\pi B dH d\mu}{m^2 |v_{\parallel}|}, \quad (103)$$

where the sum is taken over both signs, $\sigma = \pm 1$, of the parallel velocity. Because of the property (21), the left-hand side vanishes upon flux-surface averaging, and we obtain

$$\left\langle \int C_a(f_{a0}) \ln f_{a0} d^3v \right\rangle = 0,$$

The physical interpretation is that collisions should not produce entropy from the equilibrium distribution. For an electron-ion plasma, it follows that the electron and ion distribution functions are Maxwellian, possibly with different temperatures (because of the smallness of the collisional energy transfer between electrons and ions when $m_e \ll m_i$),

$$f_{a0} = N_a(\psi) \left(\frac{m_a}{2\pi T_a} \right)^{3/2} \exp \left(-\frac{m_a v^2/2 + e_a \phi}{T_a} \right),$$

and quasineutrality, $n_e = n_i$, with $n_a = N_a \exp(-e_a \phi/T_a)$, implies that the electrostatic potential is a flux function, $\phi = \phi_0(\psi)$, in lowest order.

The correction to the Maxwellian is determined from the first-order kinetic equation

$$v_{\parallel} \nabla_{\parallel} f_{a1} - C_a(f_{a1}) = -\mathbf{v}_{da} \cdot \nabla f_{a0}. \quad (104)$$

The solution of this equation can be accomplished analytically in certain asymptotic collisionality regimes, but frequently the solution needs to be found numerically. Whatever method is used, the distribution function f_{a1} is a linear combination of the driving terms on the right,

$$-\mathbf{v}_{da} \cdot \nabla f_{a0} = (\mathbf{v}_{da} \cdot \nabla \psi) \left[A_{1a} + A_{2a} \left(\frac{m_a v^2}{2T_a} - \frac{5}{2} \right) \right] f_{a0}, \quad (105)$$

where

$$A_{1a} = -\frac{d\phi_0}{d\psi} - \frac{dp_a}{d\psi},$$

$$A_{2a} = -\frac{dT_a}{d\psi},$$

are the so-called thermodynamic forces. The function f_{a1} is thus uniquely determined by the local density and temperatures, their gradients, and the radial electric field, $\phi'_0(\psi)$. From f_{a1} most equilibrium properties of interest can be calculated, such as the parallel flow velocity of each species

$$V_{a\parallel} = \frac{1}{n_a} \int v_{\parallel} f_{a1} d^3v, \quad (106)$$

the parallel current

$$J_{\parallel} = \sum_a n_a e_a V_{a\parallel},$$

and the neoclassical cross-field fluxes

$$\langle \mathbf{\Gamma}_a \cdot \nabla \psi \rangle = \left\langle \int f_{a1} \mathbf{v}_{da} \cdot \nabla \psi d^3v \right\rangle, \quad (107)$$

$$\langle \mathbf{q}_a \cdot \nabla \psi \rangle = \left\langle \int \left(\frac{m_a v^2}{2} - \frac{5T_a}{2} \right) f_{a1} \mathbf{v}_{da} \cdot \nabla \psi d^3v \right\rangle. \quad (108)$$

In the present ordering, all these quantities will thus be linear functions of the thermodynamic forces.

We do not endeavour to solve the first-order drift kinetic equation (104), but give a brief outline of the various collisionality regimes where this can be accomplished analytically. At very high collisionality, in the so-called Pfirsch-Schlüter regime, the collision frequency ν_a exceeds the transit frequency $\omega_{ta} = v_{Ta}/L$, where L is the macroscopic length, making the collisional mean free path $\lambda = v_{Ta}/\nu_a$ shorter than L . By performing a *subsidiary* expansion of f_{a1} in the small parameter $\omega_{ta}/\nu_a = \lambda/L$ [64]

$$f_{a1} = f_{a1}^{(-1)} + f_{a1}^{(0)} + f_{a1}^{(1)} + \dots, \quad (109)$$

one finds a series of ‘‘Spitzer problems’’,

$$C_a \left(f_{a1}^{(-1)} \right) = 0,$$

$$C_a \left(f_{a1}^{(0)} \right) = v_{\parallel} \nabla_{\parallel} f_{a1}^{(-1)},$$

$$C_a(f_{a1}^{(1)}) = v_{\parallel} \nabla_{\parallel} f_{a1}^{(0)} + \mathbf{v}_{da} \cdot \nabla f_{a0}.$$

The reason why the expansion (109) starts at the -1st order is that the notation is such that each term $f_{am}^{(n)}$ is order $\delta^m (\lambda/L)^n f_{a0}$, and the first term turns out to be of order $\delta (\lambda/L)^{-1} f_{a0}$. When the equations have been solved, the neoclassical transport can be calculated by inserting $f_{a1}^{(-1)}$ in Eqs. (107) and (108). The cross-field fluxes thus scale as $\delta^2 L/\lambda$ and are proportional to the collision frequency.

The mathematical treatment of the opposite, low-collisionality limit, $\nu_a/\omega_{ta} \ll 1$ depends on whether the magnetic configuration is omnigenous. If the radial drift velocity vanishes upon a bounce average (76), it can be written as a derivative, taken at constant energy and magnetic moment,

$$\mathbf{v}_{da} \cdot \nabla \psi = v_{\parallel} \nabla_{\parallel} \Delta \psi,$$

of some function $\Delta \psi$ of the phase-space coordinates. Equation (104) then becomes

$$v_{\parallel} \nabla_{\parallel} \left(f_{a1} + \Delta \psi \frac{\partial f_{a0}}{\partial \psi} \right) = C_a(f_{a1}),$$

and is solved by taking the left-hand side to vanish in lowest order, and then considering the orbit-average of the next-order equation. This is the route taken in neoclassical transport theory for tokamaks [32].

If the magnetic field is not omnigenous, the correct subsidiary expansion of the distribution function in the small parameter $\nu_a/\omega_{ta} \ll 1$ is

$$f_{a1} = f_{a1}^{(-1)} + f_{a1}^{(0)} + f_{a1}^{(1)} + \dots,$$

where

$$\begin{aligned} v_{\parallel} \nabla_{\parallel} f_{a1}^{(-1)} &= 0, \\ v_{\parallel} \nabla_{\parallel} f_{a1}^{(0)} - C_a(f_{a1}^{(-1)}) &= -\mathbf{v}_{da} \cdot \nabla f_{a0}. \end{aligned} \quad (110)$$

The first of these equations implies that $f_{a1}^{(-1)}$ is a function of constants of the motion, H , μ , and ψ . In the second equation, the first term represents a time derivative along the orbit, and can therefore be annihilated by taking an orbit average. This is done by dividing the by v_{\parallel} and integrating along the field. For circulating particles, the integral can be extended indefinitely and converted to a flux-surface average (22), giving

$$\left\langle \frac{B}{|v_{\parallel}|} \left[\mathbf{v}_{da} \cdot \nabla f_{a0} - C_a(f_{a1}^{(-1)}) \right] \right\rangle = 0,$$

where the first term vanishes because of the argument given in connection with Eq. (63). In the trapped domain of velocity space, we instead multiply Eq. (110) by $dl/|v_{\parallel}|$, sum over $\sigma = v_{\parallel}/|v_{\parallel}|$, and integrate over l between two consecutive bounce points, yielding

$$B \sum_{\sigma} \int_{l_1}^{l_2} \left[\sigma \frac{\partial f_{a1}}{\partial l} + \frac{\mathbf{v}_{da} \cdot \nabla f_{a0} - C_a(f_{a1}^{(0)})}{|v_{\parallel}|} \right] dl = 0. \quad (111)$$

The first term vanishes because we impose the boundary condition that the number of co-moving particles at each bounce point should equal the number of counter-moving ones, $f(l_j, \sigma > 0) = f(l_j, \sigma < 0)$. The remaining equation is solved for $f_{a1}^{(-1)}$, which apparently becomes of order $f_{a1}^{(-1)} \sim v_{da} f_{a0} / (\nu_a L)$, where ν_a denotes the collision frequency. The transport,

$$\langle \mathbf{\Gamma}_a \cdot \nabla \psi \rangle = \left\langle \int f_{a1}^{(-1)} \mathbf{v}_{da} \cdot \nabla \psi d^3v \right\rangle,$$

thus scales as $v_{da}^2 / \nu_a \propto T_a^{7/2}$ since $v_{da} \propto v^2$ and $\nu_a \propto v^{-3}$ for fast particles (which make the dominant contribution to the transport). This so-called $1/\nu$ -transport can be interpreted in terms of a random walk of trapped particles taking radial steps of length $v_{da} \Delta t$ at time intervals Δt equal to the inverse effective collision frequency for scattering in and out of the trapping regions. If these comprise a fraction f_t of velocity space, the effective scattering frequency is ν_a / f_t^2 and we expect a diffusion coefficient

$$D_a \sim f_t \frac{\Delta r^2}{\Delta t} \sim \frac{f_t^3 v_{da}^2}{\nu_a}. \quad (112)$$

This type of transport scales very unfavourably with temperature, $D_a \sim T_a^{7/2}$, and can lead to prohibitively large energy losses in unoptimised stellarators.

4.2.2 Low collisionality

Whereas the treatment above is adequate in all collisionality regimes for axisymmetric systems, it needs to be modified at low collisionality in stellarators with unconfined orbits. For instance, in the limit of vanishing collision frequency, some particles will simply drift out of the device, leaving a ‘‘hole’’ in velocity space, and the distribution function will not even approximately be Maxwellian. More importantly, at slightly higher collisionalities, there are experimentally important regimes of transport that cannot be captured by the ordering employed above.

To treat these, we again expand $f_a = f_{a0} + f_{a1} + \dots$ but now require in lowest order

$$v_{\parallel} \nabla_{\parallel} f_{a0} = 0,$$

thus ruling out the possibility of short mean free path. In next order, then,

$$v_{\parallel} \nabla_{\parallel} f_{a1} + \mathbf{v}_{da} \cdot \nabla f_{a0} = C_a(f_{a0}). \quad (113)$$

In order to annihilate the first term, we again take an orbit average, just like in the previous subsection. For circulating particles, we have already seen that the average of the driving term, $\mathbf{v}_{da} \cdot \nabla f_{a0}$, vanishes, so that

$$\left\langle \frac{BC_a(f_{a0})}{|v_{\parallel}|} \right\rangle = 0, \quad (114)$$

and for trapped particles we obtain

$$\int_{l_1}^{l_2} [\mathbf{v}_{da} \cdot \nabla f_{a0} - C_a(f_{a0})] \frac{dl}{v_{\parallel}} = 0. \quad (115)$$

Because of the appearance of the first term, containing the drift \mathbf{v}_{da} , there is no guarantee that f_{a0} must be Maxwellian. Only if the time scale for radial drift motion is longer than the collision time for some reason, or the fraction of escaping trapped particles is small, can we expect the plasma to be in local thermodynamic equilibrium. In practice, this is usually the case for thermal ions unless the temperature is too high. But above a critical temperature (which is high in configurations that are optimised for good confinement), the radial drift is not negligible in Eq. (115) and the neoclassical transport becomes non-local in nature, because the radial steps in the random walk are no longer short in comparison with the gradient length scale [65]. This is a qualitative difference between stellarators and tokamaks, where the drift orbits are closed and the random-walk step size is limited by their width.

If the drift term is small enough, we can make a subsidiary expansion

$$f_{a0} = F_{a0} + F_{a1} + \dots$$

and obtain in zeroth order

$$\int [\mathbf{v}_{da}^{\alpha} \cdot \nabla F_{a0} - C_a(F_{a0})] \frac{dl}{v_{\parallel}} = 0, \quad (116)$$

where the integral is taken once around the trapped orbit (back and forth along the field). Here we have split the drift velocity into its components within and across the magnetic surface,

$$\mathbf{v}_{da} = \mathbf{v}_{da}^{\alpha} + \mathbf{v}_{da}^{\psi}.$$

The most important contribution to the drift within the magnetic surface typically comes from the $E \times B$ drift, $\mathbf{b} \times \nabla\phi_0(\psi)/B$, which is principally in the poloidal direction. Then multiplying Eqs. (114) and (116) by $\ln F_{a0}$, integrating over H and μ , and again appealing to the H-theorem shows that F_{a0} must be Maxwellian and ϕ a lowest-order flux function. Note that $\mathbf{v}_{da}^\alpha \cdot \nabla\psi = 0$ by assumption, so that the drift term in Eq. (111) vanishes for a Maxwellian flux function.²

The correction F_{a1} to the Maxwellian is obtained from the remainder of the bounce-averaged drift kinetic equation (115),

$$\int [\mathbf{v}_{da} \cdot \nabla F_{a0} + \mathbf{v}_{da}^\alpha \cdot \nabla F_{a1} - C_a(F_{a1})] \frac{dl}{v_{\parallel}} = 0. \quad (117)$$

As in the case of moderate collisionality considered earlier, the function F_{a1} depends only on the local plasma parameters and their gradients. However, the dependence on $\phi'(\psi)$ is no longer linear because this quantity appears not only in the thermodynamic force A_{1a} from $\mathbf{v}_{da} \cdot \nabla F_{a0}$ but also in the drift term $\mathbf{v}_{da}^\alpha \cdot \nabla F_{a1}$. Note that the first term in the integral (117) vanishes upon integration in omnigenous configurations, which can thus be treated by the moderate-collisionality expansion considered earlier. In configurations that are not omnigenous, there is enhanced transport, whose character depends on which of the two terms involving F_{a1} dominates in Eq. (117). If the last term does so, then

$$\int [\mathbf{v}_{da} \cdot \nabla F_{a0} - C_a(F_{a1})] \frac{dl}{v_{\parallel}} = 0,$$

which is identical to the equation (111) that was solved in the previous subsection when treating the $1/\nu$ -regime of transport. Our two expansion schemes thus overlap and yield similar results in this regime.

At lower collisionality, the first term containing F_{a1} in Eq. (117) instead dominates, so that

$$\int (\mathbf{v}_{da} \cdot \nabla F_{a0} + \mathbf{v}_{da}^\alpha \cdot \nabla F_{a1}) \frac{dl}{v_{\parallel}} = 0, \quad (118)$$

approximately holds in the trapped region of phase space, whereas F_{a1} still vanishes in the circulating region. Its derivative is discontinuous across the trapped-passing boundary in this approximation, indicating that collisions

²There is actually a complication having to do with the the fact that $(B/v_{\parallel})\mathbf{v}_{da}^\alpha$ is not necessarily divergence-free, but usually it is so to a good approximation. In the DKES code, this is achieved by approximating the $E \times B$ velocity by $\mathbf{B} \times \nabla\phi_0/\langle B^2 \rangle$ [66].

must be taken into account in a boundary layer. Before considering the latter, we note that Eq. (118) indicates that

$$\frac{F_{a1}}{F_{a0}} \sim \frac{\overline{v_{da}^\psi}}{v_{da}^\alpha},$$

where we have indicated the orbit average by an overbar. Our expansion scheme is only consistent if this ratio is small. Note that this is purely a matter of magnetic-field geometry. The ratio is not small in the gyroradius or collisionality expansions; it is only small if the magnetic and electric drifts for some reason have the property that their poloidal components dominate over the radial ones on a bounce average. This is typically the case at large aspect ratio, where the $E \times B$ drift scales inversely with the minor radius and the magnetic drifts with the major radius. What then happens is that the drift \mathbf{v}_{da}^α convects the trapped particles poloidally around the torus before they have had time to drift far radially, and the radial excursion of the bounce-averaged orbits becomes of order $\Delta r \sim v_{da}^\psi/\omega_\alpha \ll r$, where $\omega_\alpha \sim \mathbf{v}_{da}^\alpha \cdot \nabla\alpha$ is the frequency of the poloidal drift. The collisionless orbits are thus confined, but collisions will scatter the particles in and out of the local trapping regions and thereby make them undergo a random walk with the step size Δr . The effective collision frequency for such scattering depends quadratically on the distance $\Delta\xi$ (in terms of pitch angle) to the trapping boundary in velocity space, $\nu_{\text{eff}} \sim \nu_a/\Delta\xi^2$. Multiplying $\nu_{\text{eff}}\Delta r^2$ by the fraction of participating particles ($\sim \Delta\xi$) gives the diffusion coefficient estimate

$$D_a \sim \frac{\nu_a}{\Delta\xi} \left(\frac{v_{da}}{\omega_\alpha} \right)^2,$$

which diverges as $\Delta\xi \rightarrow 0$, indicating that the most important role is played by particles close to the trapping boundary. The width of this boundary layer is limited from below by the assumption $\nu_{\text{eff}} < \omega_\alpha$ made in deriving Eq. (118), which implies $\Delta\xi > (\nu/\omega_\alpha)^{1/2}$ and results in the diffusion coefficient

$$D_a \sim \frac{\nu_a^{1/2} v_{da}^2}{\omega_\alpha^{3/2}}, \quad (119)$$

which is proportional to the square root of the collision frequency.

If the local magnetic field maxima, B_{max} , on each flux surface are different from each other, so that B_{max} varies from one field line to another, the precessional drift in Eq. (118) will also cause collisionless detrapping and retrapping. As mentioned in Sec. 3.5 above this leads to transport even

in the limit of vanishing collision frequency, since each de- or re-trapping event is associated with a radial displacement. This type of transport can be important for high-energy ions [67].

In practice, the collisionality regimes discussed above are usually not well separated. For this reason, and because of the complexity of the full Coulomb collision operator, it is necessary to solve the drift kinetic equation numerically in order to obtain accurate transport coefficients. A comprehensive overview of efforts in this direction has recently been given by Beidler et al. [63]. Figure 4.2.2 shows an example of such a numerical calculation, with the various asymptotic regimes indicated by dashed lines.

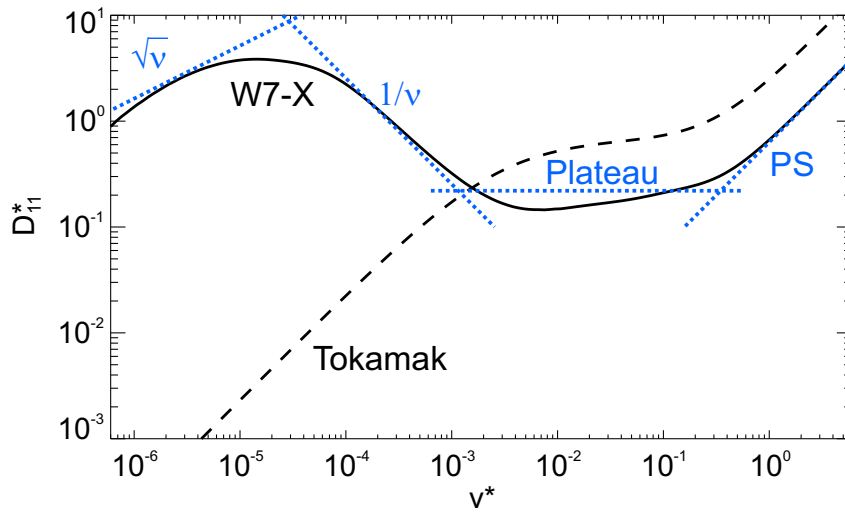


Figure 12: The so-called “mono-energetic” diffusion coefficient (see Ref. [63] for details) vs collisionality, $\nu^* = \nu R/v$, where ν is the mono-energetic pitch-angle-scattering frequency, R the major radius and v the speed of the particles, in the standard configuration of W7-X (bold) and a tokamak (dashed) with similar aspect ratio ($r/R = 0.255/5.527$) and an elongation of 1.5. The asymptotic regimes are indicated by dotted straight lines. In the order of increasing collisionality: the $\sqrt{\nu}$ -regime, the $1/\nu$ -regime, the plateau regime and the Pfirsch-Schlüter regime. At very low collisionality (below the range shown) the transport again becomes proportional to ν . The diffusivity has been normalised to the plateau value in a circular tokamak, and the radial electric field has been chosen as $E_r/vB = 3 \cdot 10^{-5}$, where B is the magnetic field strength. If the electric field is made larger, the transition from the $\sqrt{\nu}$ -regime to the $1/\nu$ -regime occurs at higher collisionality. From Ref. [48].

4.3 Fluctuations and turbulent transport

In the treatment just given, we focussed on the equilibrium properties of the plasma, treating the time derivative as $O(\delta^2 v_T/L)$. This is sufficient for calculating the collisional (neoclassical) transport but fails to capture turbulent fluctuations and transport. To do so, we need to elevate the time derivative to order $O(\delta v_T/L)$ and also allow f_{a1} to vary on the length scale of the gyro-radius. If it is assumed that the fluctuating electric and magnetic fields, $\delta\mathbf{E} = -\nabla\delta\phi - \partial\delta\mathbf{A}/\partial t$ and $\delta\mathbf{B} = \nabla \times \delta\mathbf{A}$, are small and the wave numbers are ordered as

$$k_{\parallel}L \sim k_{\perp}\rho_i \sim 1, \quad (120)$$

the result is the famous gyrokinetic equation

$$\frac{\partial g_a}{\partial t} + (v_{\parallel}\mathbf{b} + \mathbf{v}_{da} + \delta\mathbf{v}_{da}) \cdot \nabla(f_{a0} + g_a) - \langle C_a(g_a) \rangle_{\mathbf{R}} = \frac{e_a f_{a0}}{T_a} \frac{\partial \langle \chi \rangle_{\mathbf{R}}}{\partial t}, \quad (121)$$

where the distribution function has been written as

$$f_{a1} = -\frac{e_a \delta\phi(\mathbf{r}, t)}{T_a} f_{a0} + g_a(\mathbf{R}, H, \mu, t),$$

and where $\chi = \delta\phi - \mathbf{v} \cdot \delta\mathbf{A}$ is the gyrokinetic potential. Here, the gyro-average at fixed guiding-centre position is denoted by $\langle \dots \rangle_{\mathbf{R}}$, and the perturbation of the drift velocity is given by

$$\delta\mathbf{v}_{da} = \frac{\mathbf{b} \times \nabla \langle \chi \rangle_{\mathbf{R}}}{B}. \quad (122)$$

According to Eq. (120) perturbations are assumed to vary much more rapidly across the field than along it. The physical reason for this ordering is that unless the parallel phase velocity exceeds the ion thermal speed,

$$\frac{\omega}{k_{\parallel}} \gg v_{Ti},$$

there is strong ion Landau damping. Since the frequency for drift waves is of order $\omega_* \sim k_{\perp}\rho_i v_{Ti}/L$, it follows that the parallel wavelength must be of order L if $k_{\perp}\rho_i = O(1)$ to avoid Landau damping. For each Fourier component of the fluctuations we then have

$$\langle \chi \rangle_{\mathbf{R}, \mathbf{k}} = J_0 \left(\delta\phi_{\mathbf{k}} - v_{\parallel} \delta A_{\parallel \mathbf{k}} \right) + \frac{J_1 v_{\perp}}{k_{\perp}} \delta B_{\parallel \mathbf{k}}, \quad (123)$$

where the argument of the Bessel functions is $k_{\perp} v_{\perp} / \Omega_a$, $\delta B_{\parallel} = \mathbf{b} \cdot \delta\mathbf{B}$, and we have adopted the Coulomb gauge, $\nabla \cdot \delta\mathbf{A} = 0$.

The gyrokinetic field equations determining $\delta\phi$, δA_{\parallel} and δB_{\parallel} are

$$\begin{aligned}\sum_a \frac{n_a e_a^2}{T_a} \delta\phi &= \sum_a e_a \int g_a J_0 d^3v, \\ \delta A_{\parallel} &= \frac{\mu_0}{k_{\perp}^2} \sum_a e_a \int v_{\parallel} g_a J_0 d^3v, \\ \delta B_{\parallel} &= -\frac{\mu_0}{k_{\perp}} \sum_a e_a \int v_{\perp} g_a J_1 d^3v,\end{aligned}\tag{124}$$

where the volume element in velocity space is given by Eq. (103). The gyrokinetic particle and heat fluxes are

$$\begin{pmatrix} \delta\mathbf{\Gamma}_a \cdot \nabla\psi \\ \delta\mathbf{q}_a \cdot \nabla\psi \end{pmatrix} = \int \left(\frac{1}{\frac{m_a v^2}{2} - \frac{5T_a}{2}} \right) g_a \delta\mathbf{v}_d \cdot \nabla\psi d^3v,$$

and are thus of order δ^2 in our basic gyroradius expansion (54). This is the same order as the neoclassical transport, and we thus expect that the two transport channels should be comparable, at least generally speaking. In practice, turbulent transport tends to dominate except in low-collisionality plasmas with axisymmetry.

4.4 Ambipolarity and plasma rotation

There is an important difference between neoclassical and turbulent transport concerning ambipolarity. It follows from Eqs. (122), (123) and (124) that the turbulent transport is automatically ambipolar,

$$\langle \delta\mathbf{J} \cdot \nabla\psi \rangle = \sum_a e_a \langle \delta\mathbf{\Gamma}_a \cdot \nabla\psi \rangle = 0,$$

in leading order, regardless of the magnitude of the radial electric field. But, as we shall see, neoclassical transport is in general not ambipolar unless the electric field assumes a particular value. Since the total transport must be ambipolar (on the transport time scale $\partial/\partial t \sim \delta^2 v_{T_a}/L$), the radial electric field must therefore adjust so as to make the neoclassical channel ambipolar (unless the field is quasisymmetric). This fixes the perpendicular flow velocity of each species,

$$\mathbf{V}_{a\perp} = \frac{\mathbf{b} \times (\nabla\phi - \nabla p_a/n_a e_a)}{B},$$

and the parallel flow through Eq. (106). The rotation of the plasma within each flux surface is thus determined by the local density and temperature gradients of all species.

If, on the other hand, the field is axisymmetric or quasisymmetric, the neoclassical transport turns out to be automatically ambipolar and there is no constraint on the electric field in this order. The plasma is, so to say, free to rotate as it wishes, and the rotation is only governed by slower (next-order in δ) processes – the transport of angular momentum in an axisymmetric device. The point is that, if the field is axi- or quasisymmetric, then the momentum in the symmetry direction is a conserved quantity that cannot be created or destroyed, only moved elsewhere, just like mass or energy. The profile of the electric field is then set by momentum transport, just like the density profile is determined by particle transport. Non-quasisymmetric devices are different: there is no direction in which plasma momentum is conserved, and the rotation is clamped at a level determined by the density and temperature gradients.

To prove our assertion that only quasisymmetric fields exhibit automatic ambipolarity [68, 69], we multiply the first-order drift kinetic equation (104) by f_{a1}/f_{a0} , integrate over velocity space and take the flux-surface average. This gives an equation for the entropy balance

$$-\left\langle \mathbf{\Gamma}_a \cdot \left(\frac{\nabla p_a}{p_a} + \frac{e_a \nabla \phi}{T_a} \right) + \frac{\mathbf{q}_a \cdot \nabla T_a}{T_a^2} \right\rangle = - \left\langle \int \frac{f_{a1} C_a(f_{a1})}{f_{a0}} d^3 v \right\rangle. \quad (125)$$

If the collisionality is very low, we instead multiply Eq. (117) by $(F_{a1}/F_{a0})\pi v^3 dv d\lambda$ and integrate, which gives an entropy-balance equation similar to Eq. (125).

Now, if the transport is to be automatically ambipolar, it must be so for all gradients, and we are thus free to choose the density and temperature gradients to vanish. Multiplying Eq. (125) by T_a and summing over all species then gives

$$-\langle \mathbf{J} \cdot \nabla \phi \rangle = - \sum_a T_a \left\langle \int \frac{f_{a1} C_a(f_{a1})}{f_{a0}} d^3 v \right\rangle \geq 0, \quad (126)$$

where we have invoked the H-theorem on the right. A similar equation, but involving F_{a1} instead of f_{a1} , can be obtained in the case of low collisionality from Eq. (117). By definition, the transport is intrinsically ambipolar if, and only if, this current vanishes, regardless of the value of $\phi'_0(\psi)$ (as long as it is small enough to satisfy the linear approximation). According to the H-theorem the current only vanishes if f_{a1} is of the form

$$f_{a1} = \left[\frac{\Delta n_a}{n_a} + \frac{m_a v_{\parallel} V_{\parallel}}{T_a} + \left(\frac{m_a v^2}{2T_a} - \frac{5}{2} \right) \frac{\Delta T_a}{T_a} \right] f_{a0},$$

where Δn_a , V_{\parallel} and ΔT_a are density, parallel flow velocity and temperature contained in f_{a1} . Since the collision operator vanishes whenever the current (126) does so, it follows from the drift kinetic equation (104) that the part of f_{a1} that is caused by the radial electric field is odd in v_{\parallel} , so that Δn_a and ΔT_a vanish. But then

$$v_{\parallel} \nabla_{\parallel} f_{a1} = \frac{m_a v^2}{T_a} \left[\left(1 - \frac{\lambda B}{2}\right) \nabla_{\parallel} V_{\parallel} - \frac{\lambda}{2} \nabla_{\parallel} (V_{\parallel} B) \right] f_{a0},$$

where the derivative is taken at fixed $\lambda = 2\mu/m_a v^2$ and v , must equal

$$-\mathbf{v}_d \cdot \nabla f_{a0} = \frac{m_a v^2 f_{a0}}{T_a B^3} \frac{d\phi}{d\psi} \left(1 - \frac{\lambda B}{2}\right) (\mathbf{B} \times \nabla B) \cdot \nabla \psi,$$

for all λ , where we have recalled Eq. (105). This can only be the case if the following two conditions are satisfied:

$$\begin{aligned} \nabla_{\parallel} V_{\parallel} &= -\frac{d\phi}{d\psi} \frac{(\mathbf{B} \times \nabla B) \cdot \nabla \psi}{B^3}, \\ \nabla_{\parallel} (V_{\parallel} B) &= 0. \end{aligned}$$

The second of these conditions implies that the parallel velocity is of the form $V_{\parallel} = U(\psi)/B$, where $U(\psi)$ is a flux function. Substituting this result in the first condition gives

$$U(\psi) = -\frac{d\phi}{d\psi} \frac{(\mathbf{B} \times \nabla \psi) \cdot \nabla B}{\mathbf{B} \cdot \nabla B},$$

which implies that the expression (68) must be a flux function. The neoclassical transport is thus automatically ambipolar only if the magnetic field is quasisymmetric. The total plasma flow velocity is obtained by adding the $\mathbf{E} \times \mathbf{B}$ drift,

$$\mathbf{V} = \frac{U(\psi)\mathbf{b}}{B} + \frac{d\phi}{d\psi} \frac{\mathbf{b} \times \nabla \psi}{B},$$

and has the property $\mathbf{V} \cdot \nabla B = 0$, so that the plasma flows in the symmetry direction.

As already emphasised, the radial electric field in a non-quasisymmetric stellarator is determined by ambipolarity,

$$\langle \mathbf{J} \cdot \nabla \psi \rangle = \sum_a e_a \langle \mathbf{\Gamma}_a \cdot \nabla \psi \rangle = 0,$$

where the current is generally a nonlinear function of the electric field. Because of this nonlinearity, the equation may have several roots. Generally

speaking, a large negative (inward pointing) radial electric field reduces the ion current and enhances the electron current, and vice versa for a positive field. Since the transport coefficients tend to be larger for ions than for electrons, because of the scalings (112) and (119), the radial electric field normally points inward, and the so-called ion root is realised. But if the electrons are heated strongly and preferentially, their diffusion coefficient can exceed that of the ions and the electron root, corresponding to positive radial electric field, is instead attained. If some flux surfaces exhibit the ion root and others the electron root, there will be an intermediate region where the radial electric field varies rapidly, giving rise to strongly sheared $E \times B$ rotation. This is indeed observed in experiments.

4.5 Quasisymmetric and quasi-isodynamic configurations

The neoclassical transport in a perfectly quasisymmetric stellarator is very similar to that in a tokamak. We have already seen in Eq. (71) that the guiding-centre Lagrangian is identical to that in a tokamak if one makes the replacements

$$\begin{aligned}\theta &\rightarrow \theta - (N/M)\varphi, \\ G &\rightarrow G + NI/M, \\ \chi &\rightarrow \chi - N\psi/M.\end{aligned}$$

With these replacements, the orbits are thus identical, and so is therefore the first-order drift kinetic equation (104). There is a one-to-one mapping, an isomorphism [70], between the drift kinetic equation in a tokamak and a quasisymmetric stellarator. If one has computed the neoclassical transport in a tokamak, the corresponding transport in a quasisymmetric stellarator can thus be obtained immediately by a suitable replacement of variables.

To see this explicitly, we assume that the magnetic field strength depends on the coordinates as $B = B(\psi, \vartheta)$, where $\vartheta = \theta - N\varphi/M$, and note that then

$$\mathbf{b} \cdot \nabla f = \left(\iota - \frac{N}{M} \right) \frac{v_{\parallel} B}{\iota I + G} \left(\frac{\partial f}{\partial \vartheta} \right)_{\psi, \varphi},$$

and

$$\mathbf{v}_{da} \cdot \nabla \psi = \frac{v_{\parallel}}{B} (G \nabla \varphi \times \nabla \psi) \cdot \nabla \left(\frac{v_{\parallel}}{\Omega_a} \right) = \left(G + \frac{NI}{M} \right) \frac{v_{\parallel} B}{\iota I + G} \frac{\partial}{\partial \vartheta} \left(\frac{v_{\parallel}}{\Omega_a} \right)_{\psi, \varphi},$$

so that the first-order drift kinetic equation (104) becomes

$$\frac{v_{\parallel} B}{\iota I + G} \frac{\partial}{\partial \vartheta} \left(f_{a1} + \frac{G + NI/M}{\iota - N/M} \frac{v_{\parallel}}{\Omega_a} \frac{\partial f_{a0}}{\partial \psi} \right) = \frac{C_a(f_{a1})}{\iota - N/M}. \quad (127)$$

For an axisymmetric device $M = 1$ and $N = 0$, and for a quasisymmetric stellarator $N \neq 0$. It is clear that if we compare two devices with the same magnetic field strength dependence $B = B(\psi, \theta - N\varphi/M)$ but different N/M , the distribution function will be same if the collision frequency ν and the radial gradients are adjusted so as to make

$$\frac{\nu}{\iota - N/M} \quad \text{and} \quad \frac{G + NI/M}{\iota - N/M} \frac{\partial f_{a0}}{\partial \psi} \quad (128)$$

the same in the two devices.

It follows that in perfectly quasisymmetric stellarators there are no collisionality regimes where the transport scales as Eq. (112) or (119). In practice, however, it is difficult to achieve quasisymmetry to the accuracy required to avoid enhanced neoclassical losses, as compared with tokamaks. Neoclassical transport tends always to be important in stellarators, at least if the temperature is sufficiently high. There is, in all stellarators so far, however quasisymmetric or omnigenous, always a $1/\nu$ -regime of enhanced transport, as can be seen, for instance, from several figures in Ref. [63].

In quasi-isodynamic magnetic fields, the mathematical problem of calculating the neoclassical transport can also be reduced to that in a tokamak, if the collisionality is sufficiently low [43, 44]. This is readily seen by using Eq. (87) in the drift kinetic equation (104), written as

$$v_{\parallel} \nabla_{\parallel} \left(f_{a1} + \Delta\psi \frac{\partial f_{a0}}{\partial \psi} \right) = C_a(f_{a1}),$$

where the radial displacement $\Delta\psi$ is split into a ‘‘tokamak’’ part and a ‘‘stellarator’’ part,

$$\Delta\psi = \Delta_t\psi + \Delta_s\psi,$$

defined by

$$\Delta_t\psi = -\frac{I}{\Omega} [v_{\parallel}(l) - v_{\parallel}(l_0)], \quad (129)$$

$$\Delta_s\psi = -\frac{\partial}{\partial \alpha} \int_{B(l_0)}^{B(l)} h \frac{\partial}{\partial B} \left(\frac{v_{\parallel}}{\Omega} \right) dB. \quad (130)$$

The distribution function can thus be split accordingly,

$$f_a = f_{a0} + f_{at} + f_{as}.$$

where the tokamak part, f_{at} , -solves the kinetic equation

$$v_{\parallel} \nabla_{\parallel} \left(f_{at} + \Delta_t\psi \frac{\partial f_{a0}}{\partial \psi} \right) = C_a(f_{at}),$$

which is identical to that solved in the neoclassical theory of axisymmetric systems. The stellarator part can be written as

$$f_{as} = g_a - \Delta_s \psi \frac{\partial f_{a0}}{\partial \psi},$$

where g_a solves the equation

$$v_{\parallel} \nabla_{\parallel} g_a = C_a \left(g_a - \Delta_s \psi \frac{\partial f_{a0}}{\partial \psi} \right).$$

At low collisionality, where the bounce frequency exceeds the collision frequency, it is appropriate to expand the solution, $g_a = g_{a0} + g_{a1} + \dots$, and require

$$\begin{aligned} v_{\parallel} \nabla_{\parallel} g_{a0} &= 0, \\ v_{\parallel} \nabla_{\parallel} g_{a1} &= C_a \left(g_{a0} - \Delta_s \psi \frac{\partial f_{a0}}{\partial \psi} \right). \end{aligned}$$

The orbit average of the second of these equations annihilates the left-hand side and determines g_{a0} . Since the displacement $\Delta_s \psi$ is odd in v_{\parallel} , there is no net drive in the trapped part of velocity space, where thus $g_{a0} = 0$. In the circulating region, we multiply the equation by B/v_{\parallel} and take the flux-surface average, which is equivalent to an orbit average many turns around the torus, as seen from

$$\left\langle \frac{B}{v_{\parallel}} (\dots) \right\rangle = \oint d\alpha \int \frac{B(\dots) dl}{v_{\parallel} B} / \oint d\alpha \int \frac{dl}{B},$$

where the l -integral is taken along the field over one period of the device. Since

$$\oint d\alpha \int \frac{\Delta_s \psi dl}{v_{\parallel}}$$

vanishes because of the α -derivative in Eq. (130), we conclude that g_{a0} vanishes also in the circulating region of phase space. Thus, at low collisionality, the first-order distribution function is equal to

$$f_{a1} = f_{at} - \Delta_s \psi \frac{\partial f_{a0}}{\partial \psi}. \quad (131)$$

4.6 Bootstrap current

One of the most important predictions of the kinetic theory of plasma equilibrium is the so-called bootstrap current. We have seen in Eq. (34) that the quantity $\langle J_{\parallel} B \rangle$ is not determined by the requirement of MHD equilibrium alone and thus needs to be calculated kinetically, by solving the drift kinetic equation (104) or (113). In general stellarator geometry, this is a difficult task, and the bootstrap current has only been calculated analytically in the limit of vanishing collision frequency, using a simplified collision operator [71]. As in tokamaks, the bootstrap current is independent of the collision frequency in this limit. When the collisionality is finite but small, the first correction to the bootstrap current is proportional to the square root of the collision frequency and tends therefore to be important in practice [72]. At higher collisionality, the current, and indeed all neoclassical transport coefficients, need to be computed numerically [63, 73]. There are, however, two limits in which an analytical calculation is possible.

The first such limit is that of a perfectly quasisymmetric stellarator, where the calculation is similar to that in a tokamak according to isomorphism discussed in the previous section. In accordance with Eq. (127) and (128), the bootstrap current is proportional to

$$J_{BS} \propto \frac{G + NI/M}{\iota - N/M} \frac{\partial f_{a0}}{\partial \psi}, \quad (132)$$

where the term NI/M in the numerator can usually be neglected, being negligible in practically all stellarators. This physical reason for the scaling (132) can be understood from the particle orbits in a quasisymmetric field. Since the canonical momentum (72) is conserved, the radial excursion, $\Delta\psi$, of a particle as it moves along its orbit can be determined from $\Delta p_{\varphi} = 0$, which implies

$$\Delta\psi = \frac{G + NI/M}{\iota - N/M} \cdot \frac{\Delta v_{\parallel}}{Ze}$$

for small $\Delta\psi$, so that $\Delta\chi = \iota\Delta\psi$. As a trapped particle moves around its orbit, it is thus shifted outward or inward depending on the sign of v_{\parallel} times the same multiplier as in Eq. (132). For a density gradient with the usual sign, the bootstrap current is positive (in the sense that it increases the rotational transform) in tokamaks and quasi-axisymmetric configurations. According to Eq. (132), it is negative in quasi-helically symmetric ones if $N/M > \iota$, and then acts to “unwind” the magnetic field.

The second tractable limit is that of quasi-isodynamic geometry, where the distribution function is given by Eq. (131). Hence the parallel current

carried by each species on a particular flux surface consists of two terms,

$$\langle J_{a\parallel} B \rangle = \left\langle e_a B \int v_{\parallel} \left(f_{at} - \Delta_s \psi \frac{\partial f_{a0}}{\partial \psi} \right) d^3 v \right\rangle,$$

where the first, tokamak-like, term vanishes if the total current enclosed by the flux surface vanishes, $I(\psi) = 0$, so that Eq. (129) and f_{at} vanish. The second term, which is specific to quasi-isodynamic stellarators,

$$- \left\langle e_a B \int \Delta_s \psi \frac{\partial f_{a0}}{\partial \psi} d^3 v \right\rangle = - \left\langle 2\pi e_a B^2 \int \Delta_s \psi \frac{\partial f_{a0}}{\partial \psi} v^3 dv d\lambda \right\rangle,$$

also vanishes when the flux-surface average is performed, because of the α -derivative in Eq. (130). One therefore concludes that the total current vanishes (for any collision operator) in quasi-isodynamic stellarators.

The importance of the bootstrap current is that it changes the rotational transform of the magnetic field. This can improve plasma confinement (in a tokamak because the banana width decreases with increasing ι), and the bootstrap current is indeed a relatively “cheap” way of producing rotational transform, which is otherwise created by complicated and expensive coils in stellarators. A bootstrap current of either sign will affect the radial profile of the rotational transform, $\iota(\psi)$, and may cause undesired resonances, $\iota = n/m$, and magnetic islands to form within the plasma. Furthermore, if the total bootstrap current is non-zero, the rotational transform at the plasma edge is changed, which may interfere with divertor operation.

5 Further reading

In the present review, an attempt has been made to collect and explain basic results that are fundamental to the understanding of three-dimensionally shaped plasma equilibria. The selection of material is inevitably somewhat arbitrary, but has, first and foremost, been done with the student in mind. The theory of three-dimensionally shaped plasmas is not an orderly documented field of theoretical physics, and simple derivations of useful results can be difficult to find in the literature.

Several important topics have been omitted in this paper, and the greatest omission is the theory of MHD stability. This is an old and well developed area of plasma physics – indeed, the energy principle of ideal MHD [13] originally arose out of the need to assess the stability of early stellarator experiments – but this topic is already explained pedagogically in several other texts [2, 16, 17, 74]. Within the area of ideal MHD, other omissions

are the question of the mathematical existence of equilibria devoid of any continuous symmetry [23, 27] and the various numerical techniques that have been developed for calculating such magnetic fields [24, 28, 29]. For a general introduction to the theory of three-dimensional magnetic equilibria, the early review by Solovév and Shafranov can be recommended [3].

Regarding other topics, readers specifically interested in neoclassical transport in stellarators are encouraged to consult Refs. [61, 62, 63]. MHD equilibrium and stability, magnetic coordinates, particle orbits, and many other topics are treated in the review by Boozer [75]. General introductions to stellarator physics that do not go into mathematical details are found in Refs. [6, 7, 8, 9]. They are valuable both as introductory overviews and as a source for references to the wider research literature.

Acknowledgement

The author is grateful to Craig Beidler, Joaquim Loizu, Jürgen Nührenberg, Josefine Proll, Alessandro Zocco and two anonymous referees for reading the manuscript and suggesting valuable improvements. Craig Beidler, Michael Drevlak, Andrea Kleiber and Jürgen Nührenberg have kindly supplied the figures.

References

- [1] L. Spitzer Jr, Project Matterhorn Report PM-S-1 NYO-993, Princeton University (1951).
- [2] M. Wakatani, *Stellarator and heliotron devices* (Oxford University Press, 1998).
- [3] L.S. Solovév and V.D. Shafranov, in *Reviews of Plasma Physics*, Vol 5 (Consultants Bureau, New York, 1970).
- [4] K. Miyamoto, Nucl. Fusion **18**, 243 (1978).
- [5] V.D. Pustovitov and V.D. Shafranov, in *Reviews of Plasma Physics*, Vol 15 (Consultants Bureau, New York, 1990).
- [6] F. Wagner, Trans. Fusion Techn. **29**, 407 (1996).
- [7] A.H. Boozer, Phys. Fluids **5**, 1647 (1998).
- [8] H. Wobig, Plasma Phys. Contr. Fusion **41**, A159 (1999).
- [9] C. Beidler et al, in *Fusion Physics*, Chapter 5 (IAEA, Vienna, 2012).
- [10] W.P. Thurston, *Three-dimensional geometry and topology*, Volume 1 (Princeton University Press, Princeton, 1997).
- [11] M.D. Kruskal and R.M. Kulsrud, Phys. Fluids **1** (1958).
- [12] H. Weyl, Mathematische Annalen **77**, 21 (1916).
- [13] K. Hain, R. Lüst and A. Schlüter, Z. Naturf. **12A**, 833 (1957); I.B. Bernstein, E.A. Frieman, M.D. Kruskal and R.M. Kulsrud, Proc. Roy. Soc. A**244**, 17 (1958).
- [14] S. Hamada, Nucl. Fusion **2**, 23 (1962).
- [15] A.H. Boozer, Phys. Fluids **24**, 1999 (1981).
- [16] J.P. Freidberg, *Ideal magnetohydrodynamics* (Plenum Press, New York and London, 1987).
- [17] J.M. Greene and J.L. Johnson, Plasma Phys. **10**, 729 (1968).
- [18] L. Spitzer Phys. Fluids **1**, 253 (1958).

- [19] C. Mercier, Nucl. Fusion **4**, 213 (1964).
- [20] A. Bhattacharjee, G. M. Schreiber, and J.B. Taylor Phys. Fluids B**4**, 2737 (1992).
- [21] V.D. Shafranov, Nucl. Fusion **20**, 1075 (1980).
- [22] H.K. Moffatt, J. Fluid Mech. **159**, 359 (1985).
- [23] H. Grad, Phys. Fluids **10**, 137 (1967).
- [24] S.P. Hirshman, W.I. van Rij and P. Merkel, Comput. Phys. Comm. **43**, 143 (1986).
- [25] A.H. Reiman and H.S. Greenside, Comput. Phys. Commun. **43**, 157 (1986).
- [26] Y. Suzuki, N. Nakajima, K. Watanabe, Y. Nakamura and T. Hayashi, Nucl. Fusion **46**, L19 (2006).
- [27] S.R. Hudson and N. Nakajima, Phys. Plasmas **17**, 052511 (2010).
- [28] S.P. Hirshman, R. Sanchez and C.R. Cook, Phys. Plasmas **18**, 062504 (2011).
- [29] S.R. Hudson, R.L. Dewar, G. Dennis, M.J. Hole, M. McGann et al., Phys. Plasmas **19**, 112502 (2012).
- [30] C.C. Hegna, Phys. Plasmas **19**, 056101 (2012).
- [31] J.B. Taylor, Phys. Fluids **7**, 767 (1964).
- [32] P. Helander and D.J. Sigmar, *Collisional transport in magnetized plasmas* (Cambridge University Press, 2002).
- [33] M.D. Kruskal, J. Math. Phys. **3**, 806 (1962).
- [34] D. Palumbo, Nuovo Cimento B **53**, 507 (1968).
- [35] J. Nührenberg and R. Zille, Phys. Lett. A **129**, 113 (1988).
- [36] A.H. Boozer, Plasma Phys. Control. Fusion **37**, A103 (1995).
- [37] J. Nührenberg, W. Lotz, S. Gori, in *Theory of Fusion Plasmas* (Editrice Compositori, Bologna, 1994), p. 3.
- [38] A.H. Boozer, Phys. Fluids **26**, 496 (1983).

- [39] M. Drevlak, F. Brochard, P. Helander, J. Kisslinger, M. Mikhailov, C. Nührenberg, J. Nührenberg, Y. Turkin, *Contrib. Plasma Phys.* **53**, 459 (2013).
- [40] L.S. Hall and B. McNamara, *Phys. Fluids* **18**, 552 (1975).
- [41] J.R. Cary and S. Shasharina, *Phys. Plasmas* **4**, 3323 (1997).
- [42] S. Gori, W. Lotz, and J. J. Nührenberg, in *Theory of Fusion Plasmas* (Editrice Compositori, Bologna, 1996), p. 335.
- [43] P. Helander and J. Nührenberg, *Plasma Phys. Control. Fusion* **51**, 055004 (2009).
- [44] M. Landreman and P.J. Catto, *Phys. Plasmas* **19**, 056103 (2012).
- [45] A.A. Subbotin, M.I. Mikhailov, V.D. Shafranov, M. Yu. Isaev, C. Nührenberg, J. Nührenberg, R. Zille, V.V. Nemov, S.V. Kasilov, V.N. Kaluzhnyj, and W.A. Cooper, *Nucl. Fusion* **46**, 921 (2006).
- [46] M.N. Rosenbluth, *Phys. Fluids* **11**, 869 (1968).
- [47] J.H.E. Proll, P. Helander, J.W. Connor and G.G. Plunk, *Phys. Rev. Lett.* **108**, 245002 (2012).
- [48] P. Helander, C.D. Beidler, T.M. Bird, M. Drevlak, Y. Feng, R. Hatzky, F. Jenko, R. Kleiber, J.H.E. Proll, Yu. Turkin, and P. Xanthopoulos, *Plasma Phys. Contr. Fusion* **54**, 124009 (2012).
- [49] D.A. Garren and A.H. Boozer, *Phys. Fluids B3*, 2822 (1991).
- [50] C. Nührenberg, M.I. Mikhailov, J. Nührenberg, V.D. Shafranov, *Plasma Phys. Rep.* **36**, 558 (2010).
- [51] J.M. Canik, D.T. Anderson, F.S. Anderson, K.M. Likin, J.N. Talmadge, and K.Zhai, *Phys. Rev. Lett.* **98**, 085002 (2007).
- [52] A.A. Skovoroda, *Plasma Phys. Control. Fusion* **47**, 1911 (2005).
- [53] P. Helander, *Phys. Plasmas* **14**, 104501 (2007).
- [54] R.D. Hazeltine and A.A. Ware, *Plasma Phys. Contr. Fusion* **20**, 673 (1978).
- [55] I.B. Bernstein and P.J. Catto, *Phys. Fluids* **28**, 1342 (1985).

- [56] F.L. Hinton and S.K. Wong, *Phys. Fluids* **28**, 3082 (1985).
- [57] R. Hazeltine and F.L. Hinton, *Phys. Plasmas* **12**, 102506 (2005).
- [58] H. Sugama, T.-H. Watanabe, M. Nunami, and S. Nishimura, *Phys. Plasmas* **18**, 082505 (2011).
- [59] A.A. Galeev and R.Z. Sagdeev, in *Reviews of Plasma Physics*, edited by M.A. Leontovich (Consultants Bureau, New York, 1979) Vol 7, p. 307.
- [60] L.M. Kovrizhnykh, *Nucl. Fusion* **24**, 435 (1984).
- [61] D.D-M. Ho and R.M. Kulsrud, *Phys. Fluids* **30**, 442 (1987).
- [62] H.E. Mynick, *Phys. Plasmas* **13**, 058102 (2006).
- [63] C.D. Beidler, K. Allmaier, M.Yu. Isaev, S.V. Kasilov, W. Kernbichler, G.O. Leitold, H. Maaßberg, D.R. Mikkelsen, S. Murakami, M. Schmidt, D.A. Spong, V. Tribaldos, and A. Wakasa, *Nucl. Fusion* **51**, 076001 (2011).
- [64] R.D. Hazeltine and F.L. Hinton, *Phys. Fluids* **16**, 1883 (1973).
- [65] V. Tribaldos and J. Guasp, *Plasma Phys. Control. Fusion* **47**, 545 (2005).
- [66] S.P. Hirshman, K.C. Shaing, W.I. van Rij, C.O. Beasley, and E.C. Crume, *Phys. Fluids* **29**, 2951 (1986).
- [67] C.D. Beidler, Ya.I. Kolesnichenko, V.S. Marchenko, I.N. Sidorenko, and H. Wobig, *Phys. Plasmas* **8**, 2731 (2001).
- [68] P. Helander and A.N. Simakov, *Phys. Rev. Lett.* **101**, 145003 (2008).
- [69] I. Calvo, F.I. Parra, J.L. Velasco and J.A. Alonso, *Plasma Phys. Contr. Fusion* **55**, 125014 (2013).
- [70] A.H. Boozer, *Phys. Fluids* **26**, 496 (1983).
- [71] K.C. Shaing and J.D. Callen, *Phys. Fluids* **26**, 3315 (1983).
- [72] P. Helander, J. Geiger and H. Maaßberg, *Phys. Plasmas* **18**, 092505 (2011).

- [73] M. Landreman, H. Smith, A. Mollén and P. Helander, to appear in Phys. Plasmas (2014).
- [74] J. Nührenberg, P. Merkel, C. Schwab, U. Schwenn, A. Cooper, T. Hayashi, S.P. Hirshman, J.L. Johnson, D.A. Monticello and A.H. Reiman, Plasma Phys. Control. Fusion **35**, B115 (1993).
- [75] A.H. Boozer, Rev. Mod. Phys. **76**, 1071 (2004).