

# An asymptotic Grad–Shafranov equation for quasisymmetric stellarators

Nikita Nikulsin <sup>1,†</sup>, Wrick Sengupta <sup>1</sup>, Rogério Jorge <sup>2</sup> and Amitava Bhattacharjee <sup>1</sup>

<sup>1</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

<sup>2</sup>Department of Physics, University of Wisconsin-Madison, Madison, WI 53706, USA

(Received 13 March 2024; revised 15 June 2024; accepted 21 June 2024)

A first-order model is derived for quasisymmetric stellarators where the vacuum field due to coils is dominant, but plasma-current-induced terms are not negligible and can contribute to magnetic differential equations, with  $\beta$  of the order of the ratio induced to vacuum fields. Under these assumptions, it is proven that the aspect ratio must be large and a simple expression can be obtained for the lowest-order vacuum field. The first-order correction, which involves both vacuum and current-driven fields, is governed by a Grad–Shafranov equation and the requirement that flux surfaces exist. These two equations are not always consistent, and so this model is generally overconstrained, but special solutions exist that satisfy both equations simultaneously. One family of such solutions is the set of first-order near-axis solutions. Thus, the first-order near-axis model is a subset of the model presented here. Several other solutions outside the scope of the near-axis model are also found. A case study comparing one such solution to a VMEC-generated solution shows good agreement.

**Keywords:** fusion plasma, plasma confinement, plasma devices

## 1. Introduction

Quasisymmetry, along with quasi-isodynamicity, is one of the two main approaches to confining particle orbits inside a stellarator (Boozer 1998; Helander 2014). Quasisymmetry is usually defined as a stellarator configuration where the magnitude of the magnetic field satisfies  $|\mathbf{B}| = B(\psi, M\theta - N\phi)$  for some integers  $M, N$ , where  $(\psi, \theta, \phi)$  is a straight field line coordinate system (Helander 2014). Thus, the magnetic field magnitude is two-dimensional, but the full vector can still be three-dimensional. Such a device can be modelled using the near-axis expansion (Garren & Boozer 1991*a,b*; Landreman & Sengupta 2018), which consists of Taylor-expanding all quantities of interest in either the distance from the magnetic axis (direct coordinates) or the square root of toroidal flux (inverse coordinates) (Jorge, Sengupta & Landreman 2020), and balancing the coefficients in the magnetohydrodynamics (MHD) equilibrium equations at each order in

† Email address for correspondence: [nnikulsin@princeton.edu](mailto:nnikulsin@princeton.edu)

the Taylor expansion, resulting in a system of coupled ordinary differential and algebraic equations. However, the imposition of the quasisymmetry constraint, in addition to force balance, results in the system being overconstrained at all orders beyond the first in the near-axis expansion (Garren & Boozer 1991a). At second order, this can be rectified by not imposing quasisymmetry directly, but rather optimizing for it. Unlike traditional optimization, second-order near-axis optimization is much faster due to the significantly reduced parameter space (Landreman 2022). Beyond second order, there is evidence that the Taylor series of the near-axis model begins to diverge (Rodriguez 2022).

Magnetic shear appears at third order in the near-axis expansion, and thus depends on third-order quantities. While, in some cases, setting the third-order quantities to zero and evaluating the shear based on just lower-order quantities provides a reasonable estimate, in other cases, such an approach produces large deviations from the shear calculated in numerical solutions (Rodriguez 2022). This presents a challenge for the modelling of quasisymmetric devices, as shear is an important quantity that affects ballooning stability of equilibria, among other things (Connor, Hastie & Taylor 1978).

A complementary model for quasiaxisymmetric stellarators near axisymmetry that can incorporate shear has been derived in a previous paper (Sengupta *et al.* 2024a). In this paper, we propose a more general model that is applicable to both quasiaxisymmetric and quasi-helically symmetric stellarators far from axisymmetry. In both cases, the difficulties related to shear are avoided by not using Taylor expansions like the near-axis model does. Instead, we perform an asymptotic expansion around a vacuum magnetic field  $\nabla\chi$ , where  $\chi$  is the magnetic scalar potential. The dominance of the vacuum field is a natural assumption, given the design principles of stellarators, which aim to confine the plasma without having a large plasma current (Freidberg 2014). Our approach is inspired by Strauss' derivation of reduced MHD (Strauss 1997), who also expanded around a vacuum field, with the main difference being that Strauss assumes  $p = O(\epsilon^2)$ , whereas we allow for a higher  $\beta$  with  $p = O(\epsilon)$ , where  $\epsilon$  is a small parameter. We will show that such an ordering requires  $\chi$  to have a specific form. In the limit of low  $\beta$ , our model will reduce to the equilibrium limit of Strauss' equations under the assumption of quasisymmetry with our choice of  $\chi$ . The derivation will closely follow that of the Freidberg high- $\beta$  stellarator model (Freidberg 2014), except that Freidberg expands around a purely toroidal field  $B_0\hat{\phi}$ , whereas we allow for a more general zeroth-order field. Once the model is derived, we will present several solutions and a numerical validation of one of those solutions.

## 2. Derivation

To begin, we expand the magnetic field around a vacuum magnetic field  $\nabla\chi$ , where  $\nabla^2\chi = O(\epsilon^2)$ :

$$\mathbf{B} = \left(1 + \frac{B_1}{B_v}\right) \nabla\chi + \mathbf{B}_{\perp 1} + O(\epsilon^2). \quad (2.1)$$

We will order relative to the  $\nabla\chi$  term:  $B_v = |\nabla\chi| = O(1)$ ,  $B_1 = O(\epsilon)$  and  $\mathbf{B}_{\perp 1} = O(\epsilon)$ , where  $\epsilon = \max|\mathbf{B} - \nabla\chi|/B_v \ll 1$  and  $\mathbf{B}_{\perp 1} \perp \nabla\chi$ . In addition, the derivative along the vacuum field must be ordered as  $\nabla\chi \cdot \nabla = O(\epsilon)$ , whereas  $\nabla_{\perp} = \nabla - B_v^{-2}\nabla\chi\nabla\chi \cdot \nabla = O(1)$ . The derivative ordering can be justified as follows. The the perpendicular length scale will be  $\sim r_0$ , whereas the length scale along  $\nabla\chi$  will be  $\sim 2\pi R_0$ , where  $r_0$  and  $R_0$  are the minor and major radii, respectively. Thus, the ratio of the length scales must be less than  $1/2\pi$ , and no more than 0.1 for any realistic aspect ratio.

Taking the divergence of (2.1), we have  $\nabla \cdot \mathbf{B} = \nabla\chi \cdot \nabla(B_1/B_v) + \nabla \cdot \mathbf{B}_{\perp 1} = \nabla_{\perp} \cdot \mathbf{B}_{\perp 1} + O(\epsilon^2)$ . Since  $\nabla_{\perp}$  is two-dimensional, a stream function  $A = O(\epsilon)$  can be

introduced:  $\mathbf{B}_{\perp 1} = \nabla A \times \nabla \chi$ . Note that Strauss uses the symbol  $\psi$  for his stream function, but we will use  $A$  instead to avoid confusion with the flux surface label.

Following Strauss, we can write the current density as follows:

$$\mathbf{j} = \frac{\mathbf{B} \times \nabla p}{B^2} + j_{\parallel} \frac{\mathbf{B}}{B} = \frac{\nabla \chi \times \nabla p}{B_v^2} - \frac{\mathbf{B}}{\mu_0} \Delta^* A + O(\epsilon^2), \tag{2.2}$$

where  $\Delta^* = B_v^{-2} \nabla \cdot (B_v^2 \nabla_{\perp})$ , as defined by Strauss (1997), and the expression for  $j_{\parallel}$  is easily obtained by dotting the curl of (2.1) with  $\nabla \chi$  and using the identity  $\nabla f \cdot \nabla \times \mathbf{U} = -\nabla \cdot (\nabla f \times \mathbf{U})$ . An alternate expression can be obtained by taking the curl of (2.1):

$$\mathbf{j} = \frac{1}{\mu_0} \nabla \left( \frac{B_1}{B_v} \right) \times \nabla \chi - \frac{\nabla \chi}{\mu_0} \nabla^2 A + \frac{1}{\mu_0} (\nabla \chi \cdot \nabla) \nabla A - \frac{1}{\mu_0} (\nabla A \cdot \nabla) \nabla \chi. \tag{2.3}$$

The last two terms are both  $O(\epsilon^2)$ . This becomes more obvious for the last term if we write it as  $(\nabla A \cdot \nabla) \nabla \chi = \nabla (\nabla \chi \cdot \nabla A) - (\nabla \chi \cdot \nabla) \nabla A$ . Equating the perpendicular components of the current at lowest order, we obtain

$$\left[ \nabla \left( \frac{B_1}{\mu_0 B_v} \right) + \frac{1}{B_v^2} \nabla p \right] \times \nabla \chi = 0. \tag{2.4}$$

Unless the second term is a gradient, both terms are linearly independent and must be individually zero at order  $\epsilon$  (see Appendix A for details). If the second term is a gradient, we must have either  $B_v = B_v(p)$  or  $B_v = B_0 + O(\epsilon)$  where  $B_0 = \text{const}$ ; however, the former cannot be true as  $|\mathbf{B}|$  can only be a flux function in axisymmetry (Schief 2003), and since  $|\mathbf{B}| = B_v + O(\epsilon^2)$ ,  $B_v$  cannot be a flux function either. Thus, to satisfy the equation, we must either ensure that both terms are individually zero, i.e.  $p = O(\epsilon^2)$  and  $B_1 = 0$ , or let  $B_v = B_0 + O(\epsilon)$ . The former corresponds to the equilibrium limit of the Strauss equations, and so we focus on the latter. The equation then yields:

$$\frac{B_0 B_1}{\mu_0} + p = 0. \tag{2.5}$$

This is simply the lowest-order radial pressure balance condition; similar expressions appear in most reduced MHD models that assume  $p = O(\epsilon)$  (Freidberg 2014; Zocco, Helander & Weitzner 2020; Kruger, Hegna & Callen 1998). Taking the divergence of (2.2) then gives the generalized version of Strauss’ (26) in the equilibrium limit:

$$\mathbf{B} \cdot \nabla \left( \frac{j_{\parallel}}{B} \right) + \nabla \cdot \left( \frac{\mathbf{B} \times \nabla p}{B^2} \right) = \frac{-1}{\mu_0} \mathbf{B} \cdot \nabla \Delta^* A + \nabla \cdot \left( \frac{1}{B_v^2} \right) \cdot (\nabla \chi \times \nabla p) + O(\epsilon^3) = 0. \tag{2.6}$$

The only other equation of Strauss’ reduced MHD that is non-trivial in the equilibrium limit can be written simply as  $\mathbf{B} \cdot \nabla p = 0$ .

Equation (2.6) can be further simplified by exploiting the two-term quasisymmetry condition, the imposition of which is equivalent to demanding that  $|\mathbf{B}| = B(\psi, M\theta - N\phi)$  (Helander 2014):  $(\mathbf{B} \times \nabla \psi) \cdot \nabla B = F(\psi) \mathbf{B} \cdot \nabla B$ , where  $\psi$  is the toroidal flux. Applying the ordering, this condition can be written in one of two equivalent ways:

$$(\nabla \chi \times \nabla \psi) \cdot \nabla B_v = \mathbf{B} \cdot \nabla B_v, \tag{2.7a}$$

$$(\nabla B_v \times \nabla \chi) \cdot \nabla (\psi + A) = \nabla \chi \cdot \nabla B_v, \tag{2.7b}$$

where an alternate flux surface label  $\Psi = \int d\psi / F(\psi)$  is defined to absorb the  $F(\psi)$  factor. Note that  $F(\psi)$  is  $O(\epsilon^{-1})$  (Helander 2014), so  $\Psi = O(\epsilon)$ . Using (2.7a), the second term in

(2.6) can be rewritten as  $p'(\nabla\chi \times \nabla\Psi) \cdot \nabla B_v^{-2} = p'\mathbf{B} \cdot \nabla B_v^{-2}$ , which results in both terms in the equation having  $\mathbf{B} \cdot \nabla$  acting on something. Since  $p'$  commutes with  $\mathbf{B} \cdot \nabla$ , we can remove the  $\mathbf{B} \cdot \nabla$  operator, obtaining a Grad–Shafranov-like equation:

$$\Delta^*A - \mu_0 \frac{dp}{d\Psi} \left( \frac{1}{B_v^2} - \frac{1}{B_0^2} \right) = H(\Psi), \quad (2.8)$$

where  $H$  and  $\mu_0 p'/B_0^2$  are functions of  $\Psi$  that appear due to the integration. Although  $H$  is arbitrary, the  $B_0$  term cannot be absorbed into it as it is necessary to cancel the  $O(1)$  part of  $1/B_v^2$  so that all terms are of the same order. Further, note that  $H$  is a degree of freedom that allows us to specify the toroidal current profile; its axisymmetric equivalent is the  $FF'$  term in the standard Grad–Shafranov equation.

To make further progress,  $\chi$  has to be specified explicitly. The constraint  $B_v = B_0 + O(\epsilon)$ , which, as we have seen, arises from the high- $\beta$  assumption, requires that  $\chi = B_0 l$ . This can be seen by writing  $B_v^2$  in orthogonal Mercier coordinates  $(\rho, \omega, l)$  aligned to a magnetic axis with curvature  $\kappa$  and torsion  $\tau$  (Solov'ev & Shafranov 1970; Jorge *et al.* 2020):

$$B_v^2 = \left( \frac{\partial\chi}{\partial\rho} \right)^2 + \frac{1}{\rho^2} \left( \frac{\partial\chi}{\partial\omega} \right)^2 + \frac{1}{h^2} \left( \frac{\partial\chi}{\partial l} \right)^2 = B_0^2 + O(\epsilon), \quad (2.9)$$

where  $h = 1 - \kappa\rho \cos\theta$ ,  $\theta = \omega - \int_0^l \tau dl$ . On the axis, we must have  $\partial\chi/\partial\rho|_{\rho=0} = \lim_{\rho \rightarrow 0} \rho^{-1} \partial\chi/\partial\omega = 0$ , since the axis is a field line. Thus, the constant  $B_0^2$  can only come from the last term. It then follows that  $\kappa\rho = O(\epsilon)$ , so that  $1/h^2 = 1 + O(\epsilon)$ ; otherwise,  $\chi$  would have depended on  $\rho$  and  $\omega$  at the lowest order, which is a contradiction. With that in mind, we can integrate at the zeroth order, obtaining  $\chi = B_0 l + O(\epsilon)$ . We can ignore the  $O(\epsilon)$  correction to  $\chi$  without loss of generality; as will be shown at the end of the section, it can be absorbed into the  $\nabla A \times \nabla\chi$  term. Note that we can now give  $\epsilon$  a more intuitive meaning: given that  $\rho \sim r_0$  and  $\kappa, \tau \sim 1/R_0$ , we should order  $r_0 = O(1)$  and  $R_0 = O(\epsilon)$ , and hence we have a large aspect ratio expansion with  $\epsilon$  acting as the inverse aspect ratio. Finally, it should be emphasized that the only arbitrary assumption made in this derivation is that  $p = O(\epsilon)$ . The rest of the ordering, while not rigorous, is still justified by heuristic arguments.

As a sanity check, we will demonstrate that  $\chi = B_0 l$  only breaks the divergence-free condition at second order, which is acceptable, since this is a first-order model. Using the Laplace operator in orthogonal Mercier coordinates (see (3.4) of Sengupta *et al.* (2024b)), we have

$$\nabla^2\chi = \frac{1}{\rho h} \frac{\partial}{\partial\rho} \left( \rho h \frac{\partial\chi}{\partial\rho} \right) + \frac{1}{\rho^2 h} \frac{\partial}{\partial\omega} \left( h \frac{\partial\chi}{\partial\omega} \right) + \frac{1}{h} \frac{\partial}{\partial l} \left( \frac{1}{h} \frac{\partial\chi}{\partial l} \right) = -\frac{B_0}{h^3} \frac{\partial h}{\partial l} = O(\epsilon^2), \quad (2.10)$$

since  $h = 1 + O(\epsilon)$  and  $\partial/\partial l = B_0^{-1} h^2 \nabla\chi \cdot \nabla = O(\epsilon)$ .

Throughout the rest of the paper, we will use non-orthogonal Mercier coordinates  $(x, y, l)$ , where  $x = \rho \cos\theta$  and  $y = \rho \sin\theta$ . The metric tensor is as follows (Solov'ev & Shafranov 1970):

$$g^{ik} = \frac{1}{h^2} \begin{pmatrix} h^2 + \tau^2 y^2 & -\tau^2 xy & \tau y \\ -\tau^2 xy & h^2 + \tau^2 x^2 & -\tau x \\ \tau y & -\tau x & 1 \end{pmatrix}, \quad h = 1 - \kappa x. \quad (2.11)$$

Using the metric tensor and  $\chi = B_0 l$ , which implies  $B_v = B_0/h$ , (2.7b) can be written as

$$B_0^2 \frac{\partial}{\partial y} (\Psi + A) = -B_0^2 \left( \tau y + \frac{1}{\kappa} \frac{d\kappa}{dl} x \right). \quad (2.12)$$

The non-orthogonal Mercier coordinates have simplified the equation such that it can now be integrated:

$$\Psi + A = -\frac{\tau y^2}{2} - \frac{1}{\kappa} \frac{d\kappa}{dl} xy + a(x, l), \quad (2.13)$$

where  $a$  is an arbitrary function. This allows us to write (2.8) as an equation for  $\Psi$ :

$$\Delta_{\perp} \Psi + \tau - \frac{\partial^2 a}{\partial x^2} = \frac{2\mu_0}{B_0^2} \frac{dp}{d\Psi} \kappa x - H(\Psi), \quad (2.14)$$

where we have used  $B_v = B_0 |\nabla l| = B_0 / (1 - \kappa x)$  and  $\Delta^* = \Delta_{\perp} + O(\epsilon)$ , and dropped terms of order  $\epsilon^2$  and higher.

Although we now have an equation for  $\Psi$ , solving it is non-trivial, as  $\Psi$  is overconstrained, due to also having to satisfy the equation  $\mathbf{B} \cdot \nabla \Psi = 0$ , which, after inserting  $A$  from (2.13), can be written in Mercier coordinates as

$$\frac{\partial \Psi}{\partial l} - \frac{1}{\kappa} \frac{d\kappa}{dl} x \frac{\partial \Psi}{\partial x} + \left( \frac{1}{\kappa} \frac{d\kappa}{dl} y - \tau x - \frac{\partial a}{\partial x} \right) \frac{\partial \Psi}{\partial y} = 0, \quad (2.15)$$

where we have inserted the expression for  $A$  from (2.13). The characteristics of this equation are given by

$$\kappa x = C_1; \quad \frac{y}{\kappa} = -C_1 \int \frac{\tau}{\kappa^2} dl - \int \frac{1}{\kappa} \frac{\partial a}{\partial x} \Big|_{x=C_1/\kappa} dl + C_2. \quad (2.16a,b)$$

For  $\Psi$  to satisfy (2.15), it must be a function of only  $C_1$  and  $C_2$ , as defined in (2.16a,b). One must then ensure that the function also satisfies (2.14) at all values of  $l$ .

Having derived the model, we now show that adding an  $O(\epsilon)$  correction to  $\chi$  will not change it. Thus, if the assumption that the  $B_0 l$  is the dominant term in  $\chi$  holds, the model is fully general. Indeed, suppose that we add a correction  $\chi_1 = O(\epsilon)$  to  $\chi$ , then at order  $\epsilon$ , this term can be absorbed into the  $\nabla A \times \nabla \chi$  term by replacing  $A \mapsto A + \tilde{A}$ , where  $\nabla \chi_1 = B_0 \nabla \tilde{A} \times \nabla l$ . The fact that such an  $\tilde{A}$  exists can be seen by writing out the contravariant  $x$  and  $y$  components of this relation:

$$\frac{\partial \chi_1}{\partial x} = B_0 \frac{\partial \tilde{A}}{\partial y} + O(\epsilon^2), \quad \frac{\partial \chi_1}{\partial y} = -B_0 \frac{\partial \tilde{A}}{\partial x} + O(\epsilon^2). \quad (2.17a,b)$$

These are just the Cauchy–Riemann equations; since  $\chi_1$  must satisfy the Laplace equation at order  $\epsilon$ , a corresponding  $\tilde{A}$  is guaranteed to exist.

**3. Consistency with the first-order near-axis model**

The simplest ansatz for  $\Psi$  that will work in this model is a quadratic polynomial in  $C_1$  and  $C_2$ , which are defined in (2.16a,b):

$$\begin{aligned} \Psi &= s_1 C_1^2 + s_2 C_1 C_2 + s_3 C_2^2 = s_1 (\kappa x)^2 + s_2 (\kappa x) \left( \frac{y}{\kappa} + \kappa x \Lambda \right) + s_3 \left( \frac{y}{\kappa} + \kappa x \Lambda \right)^2 \\ &= (s_1 + s_2 \Lambda + s_3 \Lambda^2) (\kappa x)^2 + (s_2 + 2s_3 \Lambda) xy + s_3 \left( \frac{y}{\kappa} \right)^2, \end{aligned} \tag{3.1}$$

where  $\Lambda = \int \kappa^{-2} (\tau + a_2) dl$ ; and  $a$  was chosen as  $a = a_2(l)x^2/2$ . This represents a plasma with an elliptic cross-section that rotates about the magnetic axis as  $l$  varies. Inserting this expression into the Grad–Shafranov equation (2.14), we obtain an equation expressing  $a_2$  in terms of  $\Lambda$ . Note that, when inserting (3.1) into the Grad–Shafranov equation, all terms on the left-hand side of (2.14) will depend only on  $l$ , whereas terms on the right-hand side can depend on all three variables. Thus, we further assume  $p = O(\epsilon^2)$  and  $H(\Psi) = H_0 = \text{const}$ ; this is consistent with the near-axis model, where pressure enters only at the second order. Proceeding, we have

$$2(s_1 + s_2 \Lambda + s_3 \Lambda^2) \kappa^2 + 2 \frac{s_3}{\kappa^2} + \tau - a_2 + H_0 = 0. \tag{3.2}$$

Finally, we can combine  $d\Lambda/dl = (\tau + a_2)/\kappa^2$ , which follows from the definition of  $\Lambda$ , with the above equation, resulting in a Riccati equation for  $\Lambda$ :

$$\frac{d\Lambda}{dl} - 2(s_1 + s_2 \Lambda + s_3 \Lambda^2) - \frac{2s_3}{\kappa^4} - \frac{2\tau}{\kappa^2} - \frac{H_0}{\kappa^2} = 0. \tag{3.3}$$

The last step is to enforce periodicity on  $\Lambda$ . Averaging the above equation over  $l$  removes the derivative term, resulting in the following constraint:

$$-2(s_1 + s_2 \langle \Lambda \rangle + s_3 \langle \Lambda^2 \rangle) - 2s_3 \left\langle \frac{1}{\kappa^4} \right\rangle - 2 \left\langle \frac{\tau}{\kappa^2} \right\rangle - H_0 \left\langle \frac{1}{\kappa^2} \right\rangle = 0, \tag{3.4}$$

where  $\langle f(l) \rangle = L^{-1} \int_0^L f(l) dl$ , with  $L$  being the axis length. Thus, only two of the  $s_i$  constants are free, with the remaining one determined by the above constraint. In practice, since  $\Lambda$  is not known *a priori*, (3.3) and (3.4) must be solved iteratively: first, one makes an initial guess for the unknown  $s_i$ , then (3.3) is solved for  $\Lambda$ , which allows one to calculate the corrected value of the unknown  $s_i$  from (3.4); this cycle is repeated until convergence is achieved. This same method is used in the near-axis model to solve the  $\sigma$ -equation while simultaneously finding the correct rotational transform that is compatible with periodicity (Landreman, Sengupta & Plunk 2019).

Now consider the lowest order near-axis expression for  $\Psi$  (Jorge *et al.* 2020):

$$\begin{aligned} \Psi &\simeq \frac{\psi}{F_0} \simeq \frac{\pi B_0}{F_0} \rho^2 [\cosh \eta + \sinh \eta \cos 2(\theta + \delta)] \\ &= \frac{\pi B_0}{F_0} [(\cosh \eta + \sinh \eta \cos 2\delta)x^2 - 2 \sinh \eta \sin 2\delta xy + (\cosh \eta - \sinh \eta \cos 2\delta)y^2], \end{aligned} \tag{3.5}$$

where  $x = \rho \cos \theta$ ,  $y = \rho \sin \theta$ ,  $F_0$  is the value of  $F(\psi)$  on axis and  $\eta$  and  $\delta$  are functions of  $l$ . Note that  $F_0$ , as defined here, is the inverse of the  $F_0$  of Jorge *et al.* (2020).

In quasisymmetry, we have  $\cosh \eta - \sinh \eta \cos 2\delta = \bar{\eta}^2/\kappa^2$  (Jorge *et al.* 2020). Since  $\bar{\eta} = \text{const}^1$ , the  $y^2$  term of (3.1) agrees with that of (3.5) and  $s_3 = \pi B_0 \bar{\eta}^2/F_0$ . Further, taking  $4(\text{coef. of } x^2)(\text{coef. of } y^2) - (\text{coef. of } xy)^2$  from (3.5), we see that the  $l$ -dependent terms cancel and we are left with just a constant,  $4\pi^2 B_0^2/F_0^2$ . Likewise, when using the corresponding coefficients from (3.1), the terms with  $\Lambda$  cancel and we get  $4s_1 s_3 - s_2^2 = 4\pi^2 B_0^2/F_0^2$ .

Next, we show that (3.3) is equivalent to the  $\sigma$ -equation of Jorge *et al.* (2020). A relation between  $\Lambda$  and  $\sigma$  can be obtained by comparing the coefficients of  $xy$  in (3.1) and (3.5). Since  $\sigma = \sinh \eta \sin 2\delta$  (Jorge *et al.* 2020), we have

$$\Lambda = -\frac{\pi B_0}{F_0 s_3} \sigma - \frac{s_2}{2s_3}. \quad (3.6)$$

Inserting this into (3.3) and replacing the  $s_i$  variables with near-axis variables, we get

$$\frac{d\sigma}{dl} + \frac{2\pi B_0}{F_0} \left(1 + \sigma^2 + \frac{\bar{\eta}^4}{\kappa^4}\right) + 2\tau \frac{\bar{\eta}^2}{\kappa^2} + H_0 \frac{\bar{\eta}^2}{\kappa^2} = 0. \quad (3.7)$$

In the case where  $H_0 = 0$ , this equation reduces to (29) of Jorge *et al.* (2020).

To conclude this section, we show that when  $s_2 = 0$ , there is a simple relationship between the characteristics and Boozer angles. Choosing a value for  $s_2$  is similar to gauge fixing, since a non-zero  $s_2$  simply represents a deviation of the ellipse from the upright position in the  $(C_1, C_2)$  plane. The orientation of the ellipse in real space is determined by  $\sigma$ , which is unaffected by changes in  $s_2$ , as long as the initial condition for  $\Lambda$  is changed accordingly. The coordinates  $x$  and  $y$  can be written in terms of  $C_1, C_2$  as

$$x = \frac{C_1}{\kappa}, \quad y = \kappa C_2 - \kappa \Lambda C_1. \quad (3.8a,b)$$

Meanwhile, (1) and (54) from Jorge *et al.* (2020) give

$$\mathbf{r} = \mathbf{r}_0 + x\mathbf{n} + y\mathbf{b} = \mathbf{r}_0 + \sqrt{\frac{\psi}{\pi B_0}} \left[ \frac{\bar{\eta}}{\kappa} \cos \vartheta \mathbf{n} + \frac{\kappa}{\bar{\eta}} (\sin \vartheta + \sigma \cos \vartheta) \mathbf{b} \right]. \quad (3.9)$$

Equating the two and using the relations between  $\Lambda$  and  $\sigma$  as well as  $s_3$  and  $\bar{\eta}$ , we arrive at the following:

$$C_1 = \frac{F_0}{\pi B_0} \sqrt{s_3 \Psi} \cos \vartheta, \quad C_2 = \sqrt{\frac{\Psi}{s_3}} \sin \vartheta, \quad (3.10a,b)$$

where  $\vartheta = \theta_B - N\phi_B$ , with  $\theta_B$  and  $\phi_B$  being the Boozer angles.

#### 4. New classes of solutions

We now present three classes of exact and approximate solutions to (2.14) and (2.15) that fall outside the scope of the near-axis model. We will begin with a cubic polynomial solution with constant  $H$  and then proceed to a non-polynomial solution and a solution with variable  $H$ .

<sup>1</sup>In the near-axis literature,  $\bar{\eta}$  is a parameter that controls the extent of the flux surfaces in the normal direction for a given curvature. The maximum  $x$  for a flux surface  $\psi$  is  $x_{\max}(l) = \bar{\eta} \sqrt{\psi}/(\kappa(l) \sqrt{\pi B_0})$ ; see figure 2 of Rodríguez & Bhattacharjee (2021) for an illustration.

4.1. *Cubic solution with finite pressure*

Consider the following cubic polynomial in  $C_1$  and  $C_2$ :

$$\begin{aligned} \Psi &= s_1 C_1^2 + s_2 C_1 C_2 + s_3 C_2^2 + r C_1^3 \\ &= (s_1 + s_2 \Lambda + s_3 \Lambda^2)(\kappa x)^2 + (s_2 + 2s_3 \Lambda)xy + s_3 \left(\frac{y}{\kappa}\right)^2 + r(\kappa x)^3. \end{aligned} \tag{4.1}$$

Inserting this ansatz into (2.14), the following is obtained:

$$2(s_1 + s_2 \Lambda + s_3 \Lambda^2)\kappa^2 + \frac{2s_3}{\kappa^2} + 6r\kappa^3 x + \tau - a_2 = \frac{2\mu_0 p_1}{B_0^2} \kappa x - H_0, \tag{4.2}$$

where we have additionally assumed that  $dp/d\Psi = p_1 = \text{const}$ ,  $H(\Psi) = H_0 = \text{const}$  and  $a = a_2(l)x^2/2$ . Only two terms in the above equation depend on  $x$ ; equating those two terms, we get

$$r = \frac{\mu_0 p_1}{3B_0^2 \kappa^2}, \tag{4.3}$$

while the remaining terms match (3.2). Thus, a solution is obtained by first finding a quadratic solution, as discussed in § 3, and then adding a cubic term  $r(\kappa x)^3$  with  $r$  given by (4.3). Equation (4.3) also imposes the constraint that  $\kappa = \text{const}$ , but, since  $\tau$  is unconstrained, we can still get non-planar closed curves. At this point, we can see why  $rC_1^3$  is the only cubic term that we can include in (4.1). If we were to add cubic terms that involve  $C_2$ , then (4.2) would also have terms that depend on  $y$  and we would end up with an additional equation that would constrain  $\tau$ , meaning that we would not be able to ensure that the axis is closed.

Finally, note that while this cubic solution seems to be superficially similar to a second-order near-axis solution, there is a fundamental difference. Namely, the near-axis model assumes that the cubic term is a correction that is much smaller than the quadratic terms, whereas the present solution allows the cubic term to be of the same order as the quadratic ones. Also, unlike the second-order near-axis model, where the quasisymmetry error is  $O(\epsilon^3)$ , the present solution is still first order, so the quasisymmetry error will be  $O(\epsilon^2)$ .

4.2. *Non-polynomial approximate solutions*

In this subsection, we will show an approximate solution that consists of a rotating ellipse, which, as we have seen in the previous section, can be represented as a quadratic polynomial, and a non-polynomial perturbation. As we will perform a subsidiary expansion, it is convenient to rescale all quantities to be zeroth order in  $\epsilon$ . Thus, after performing the following replacement:  $\{\kappa, \bar{\eta}, \tau, a_2, H_0, \Psi, d/dl\} \mapsto \epsilon\{\kappa, \bar{\eta}, \tau, a_2, H_0, \Psi, d/dl\}$  and  $F_0 \mapsto F_0/\epsilon$ , we see that the  $\epsilon$  parameter is cancelled in (2.14) and the rotating ellipse solution (3.1).

Now consider the case when  $s_2 = 0$ ; then, using the expressions for the  $s_i$  variables and  $\Lambda$  that we found in § 3, we have  $s_3 = \pi B_0 \bar{\eta}^2 / F_0$ ,  $s_1 = \pi B_0 / (F_0 \bar{\eta}^2)$  and  $\Lambda = -\sigma / \bar{\eta}^2$ . The rotating ellipse solution can be represented as

$$\Psi_{re} = \frac{\pi B_0}{F_0} \left( \frac{1}{\bar{\eta}^2} C_1^2 + \bar{\eta}^2 C_2^2 \right) = \frac{\pi B_0}{F_0} (c_1^2 + c_2^2), \tag{4.4}$$



where the integration constants of the characteristics (2.16a,b) have been rescaled with respect to  $\bar{\eta}$ :

$$c_1 = \frac{C_1}{\bar{\eta}} = \frac{\kappa x}{\bar{\eta}}, \quad c_2 = \bar{\eta} C_2 = \frac{\bar{\eta} y}{\kappa} + \bar{\eta} \kappa x \Lambda = \frac{\bar{\eta} y}{\kappa} - \frac{\kappa x \sigma}{\bar{\eta}} \quad (4.5a,b)$$

yielding the normal form representation. We add to (4.4) a non-polynomial perturbation  $f$ , and use the relation  $F_0^{-1} = (\iota_0 - N)/(2\pi B_0 r_0)$ , which was found by Jorge *et al.* (2020), with  $\iota_0$  being the rotational transform on axis and  $N$  the helicity of the quasisymmetry. Here, the axis length  $L$  was replaced with the characteristic value of the minor radius  $r_0 = \epsilon L/(2\pi)$ , due to  $F_0$  having been rescaled with respect to  $\epsilon$ . The resulting ansatz is as follows:

$$\Psi = \frac{\iota_0 - N}{2r_0} \left[ (c_1^2 + c_2^2) + \left( \frac{\bar{\eta}}{\kappa_0} \right)^2 f(c_2 + \sigma_0 c_1) \right], \quad (4.6)$$

where  $\kappa_0$  is the minimum value of  $\kappa$ . Inserting this ansatz into the Grad–Shafranov equation (2.14), multiplying everything by  $(\bar{\eta}/\kappa)^2$  and using  $a_2 = \kappa^2 d\Lambda/dl - \tau = -(\kappa/\bar{\eta})^2 d\sigma/dl - \tau$ , the following is obtained:

$$\begin{aligned} & \frac{\iota_0 - N}{2r_0} \left[ 2(1 + \sigma^2) + 2 \left( \frac{\bar{\eta}}{\kappa} \right)^4 + \left( \frac{\bar{\eta}}{\kappa_0} \right)^2 \left( (\sigma_0 - \sigma)^2 + \left( \frac{\bar{\eta}}{\kappa} \right)^4 \right) f'' \right] \\ & + \frac{d\sigma}{dl} + \left( \frac{\bar{\eta}}{\kappa} \right)^2 (2\tau + H_0) = 0. \end{aligned} \quad (4.7)$$

We can now carry out a subsidiary expansion in the limit of small  $\bar{\eta}^2$ , i.e.  $(\bar{\eta}/\kappa_0)^2 \ll 1$ , but  $(\bar{\eta}/\kappa_0)^4 > \epsilon$  since we keep terms at that order in (4.6). This corresponds to the limit of a high aspect ratio stellarator with a highly elongated cross-section. Given that  $\sigma \sim \bar{\eta}^2$  and  $\iota_0 - N \sim \bar{\eta}^2$  (Rodríguez 2022), the two lowest orders of the above equation are as follows:

$$\left. \begin{aligned} O\left(\left(\frac{\bar{\eta}}{\kappa_0}\right)^2\right) : \quad & \frac{d\sigma_2}{dl} + \frac{\iota_{0,2}}{r_0} = -\left(\frac{\bar{\eta}}{\kappa}\right)^2 (2\tau + H_0), \\ O\left(\left(\frac{\bar{\eta}}{\kappa_0}\right)^6\right) : \quad & \frac{d\sigma_6}{dl} + \frac{\iota_{0,6}}{r_0} = -\frac{\iota_{0,2}}{r_0} \left[ \sigma_2^2 + \left(\frac{\bar{\eta}}{\kappa}\right)^4 \right], \end{aligned} \right\} \quad (4.8)$$

where  $\sigma = \sigma_2 + \sigma_6$  with  $\sigma_n \sim \bar{\eta}^n$ , and  $\iota_0 = N + \iota_{0,2} + \iota_{0,6}$  with  $\iota_{0,n} \sim \bar{\eta}^n$ . The values of  $\iota_{0,n}$  are determined by enforcing periodicity on  $\sigma_n$ : (4.8) is averaged over  $l$ , which removes  $d\sigma_n/dl$ ; the  $\iota_{0,n}$  term must then be equal to the average of the right-hand side. Since we only solve (4.7) up to  $O((\bar{\eta}/\kappa_0)^6)$  and  $f$  does not appear until  $O((\bar{\eta}/\kappa_0)^8)$ , we can treat  $f$  as a free function. If we were to attempt to solve this equation at  $O((\bar{\eta}/\kappa_0)^8)$ , then only the trivial solution  $f'' = \text{const}$  (i.e.  $f$  is a quadratic polynomial) would be permitted.

### 4.3. Approximate solution with variable $H$

We conclude this section by presenting a rotating ellipse approximate solution where  $H(\Psi) = H_0 + H_1\Psi$ . We also add a quartic term to  $a$ :  $a = a_2(l)x^2/2 + a_4(l)x^4/24$ , so the

rescaled characteristics become

$$\left. \begin{aligned} c_1 &= \frac{\kappa x}{\bar{\eta}}, & c_2 &= \frac{\bar{\eta}y}{\kappa} + \bar{\eta}\kappa x\Lambda + \frac{\bar{\eta}\kappa^3 x^3}{6}V, \\ \Lambda &= \int \frac{\tau + a_2}{\kappa^2} dl, & V &= \int \frac{a_4}{\kappa^4} dl. \end{aligned} \right\} \tag{4.9}$$

Just like in the previous subsection, we rescale all quantities to be zeroth order in  $\epsilon$  and let  $s_2 = 0$ . In addition, we order  $H_1 \sim \bar{\eta}^2$  and  $a_4 \sim \bar{\eta}^2$ . It can now be seen that in the rotating ellipse ansatz (4.4), terms with  $x$  or  $y$  raised to a power higher than two only appear at order  $(\bar{\eta}/\kappa_0)^6$  and higher. Thus, up to  $O((\bar{\eta}/\kappa_0)^4)$ , it is still purely a rotating ellipse.

Inserting the rotating ellipse solution (4.4) with the new characteristics (4.9) into the Grad–Shafranov equation (2.14) and multiplying everything by  $(\bar{\eta}/\kappa)^2$ , the following equation is obtained:

$$\begin{aligned} &\frac{\iota_0 - N}{r_0} \left[ (1 + \sigma^2) + \left(\frac{\bar{\eta}}{\kappa}\right)^4 \right] + \frac{d\sigma}{dl} + \left(\frac{\bar{\eta}}{\kappa}\right)^2 \left( 2\tau - a_4 \frac{x^2}{2} + H_0 \right) \\ &+ \frac{\iota_0 - N}{2r_0} H_1 x^2 + O\left(\left(\frac{\bar{\eta}}{\kappa_0}\right)^8\right) = 0, \end{aligned} \tag{4.10}$$

where we have again used  $a_2 = -(\kappa/\bar{\eta})^2 d\sigma/dl - \tau$ . Just as in the previous subsection, the above equation can be solved order by order. The three lowest orders are as follows:

$$\left. \begin{aligned} O\left(\left(\frac{\bar{\eta}}{\kappa_0}\right)^2\right): & \quad \frac{d\sigma_2}{dl} + \frac{\iota_{0,2}}{r_0} = -\left(\frac{\bar{\eta}}{\kappa}\right)^2 (2\tau + H_0), \\ O\left(\left(\frac{\bar{\eta}}{\kappa_0}\right)^4\right): & \quad \left(\frac{\bar{\eta}}{\kappa}\right)^2 a_4 = \frac{\iota_{0,2}}{r_0} H_1, \\ O\left(\left(\frac{\bar{\eta}}{\kappa_0}\right)^6\right): & \quad \frac{d\sigma_6}{dl} + \frac{\iota_{0,6}}{r_0} = -\frac{\iota_{0,2}}{r_0} \left[ \sigma_2^2 + \left(\frac{\bar{\eta}}{\kappa}\right)^4 \right]. \end{aligned} \right\} \tag{4.11}$$

Thus,  $\sigma$  is still determined by the same equations as (4.8); in addition to that, we also have an equation to determine  $a_4$ .

### 5. Numerical example

While the model presented in this paper was derived with high- $\beta$  equilibria in mind, the only high- $\beta$  equilibrium that we have been able to find analytically is the constant curvature one presented in § 4.1. In this section, we will instead illustrate the approximate solution discussed in § 4.2 with a numerical example of a quasisymmetric ( $N = 0$ ) device. This approximate solution has another important property that we wanted our model to include: non-zero shear. We were unable to get a numerical equilibrium with high  $\beta$  because VMEC failed to find a solution for all of the non-planar constant curvature axes that we could find, due to their high shaping. We leave the numerical verification of the analytical solution in § 4.1, as well as the numerical search for other high- $\beta$  solutions for future work.

The example is constructed by numerically solving (4.8), and is then compared with both VMEC (Hirshman & Whitson 1983) results and the closest near-axis approximation,

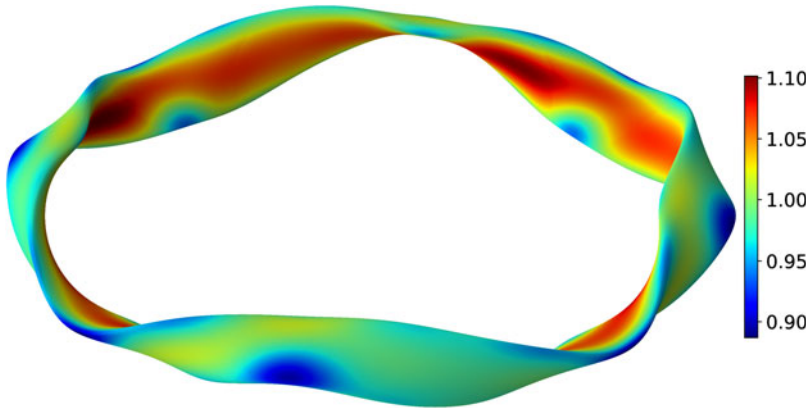


FIGURE 1. The outermost flux surface of the VMEC equilibrium based on the solution with a sixth-order polynomial perturbation. Colour represents  $|B|$ .

which is identical to taking  $f = 0$ . The pyQSC code (Landreman *et al.* 2020–2023) was used to solve the  $\sigma$  equations and generate VMEC input files, based on the method described by Landreman & Sengupta (2019).

While the solution in § 4.2 allows for arbitrary  $f$ , here we will consider a sixth-order polynomial,  $f(c) = (0.3 \text{ m}^{-2})c^4 + (0.3 \text{ m}^{-4})c^6$ , where  $c = c_2 + \sigma_0 c_1$ ; this is higher order than the cubic polynomials of the second-order near-axis model. We consider a four-field-period solution with  $\bar{\eta} = 2.01735426 \times 10^{-2} \text{ m}^{-1}$ ,  $B_0 = 1 \text{ T}$ ,  $H_0 = 0$ . Equations (4.8) result in rotational transforms  $\iota_{0,2} = 0.188923174$  and  $\iota_{0,6} = -0.031115615$ . The axis shape given by

$$\left. \begin{aligned} R(\phi) &= 29.7794783 - 3.63597602 \times 10^{-1} \cos 4\phi + 1.47477208 \times 10^{-1} \cos 8\phi \\ &\quad + 1.35576435 \times 10^{-2} \cos 12\phi, \\ z(\phi) &= 1.93173817 \sin 4\phi + 2.38762327 \times 10^{-2} \sin 8\phi - 7.72243217 \times 10^{-3} \sin 12\phi, \end{aligned} \right\} \quad (5.1)$$

where  $R$  and  $z$  are the cylindrical coordinates in metres. The value of  $(\bar{\eta}/\kappa_0)^2$  for this axis is approximately 0.3. The VMEC equilibrium, obtained by constructing the  $F_0\Psi/\pi = 4 \text{ T m}^2$  surface using (4.6) and passing it to VMEC as the boundary, is shown in figure 1. The aspect ratio of the resulting stellarator is 16.5. The assumption  $dp/d\Psi = 0$ , made when constructing the base solution in § 3, and the choice of  $H_0$  correspond to pressure and toroidal current profiles of zero; this was specified in the VMEC input. The total toroidal flux in the VMEC input was determined by multiplying  $B_0$  by the plasma cross-section area in the plane perpendicular to the axis.

Figure 2(a,b) compares the flux surfaces of the present Grad–Shafranov model, obtained from (4.6), to both the flux surfaces calculated by VMEC and the flux surfaces of the closest near-axis approximation, as given by (4.4). All flux surfaces are shown on the  $\phi = 0$  poloidal plane. As expected, near the axis, the near-axis flux surfaces closely match those of the Grad–Shafranov model, but further away from the axis, the Grad–Shafranov surfaces become non-elliptic as the contribution from  $f$  becomes non-negligible. Figure 2(c) compares the  $\iota$  profiles in the near-axis and Grad–Shafranov models. These are computed numerically by first calculating the toroidal flux  $\psi$  at each

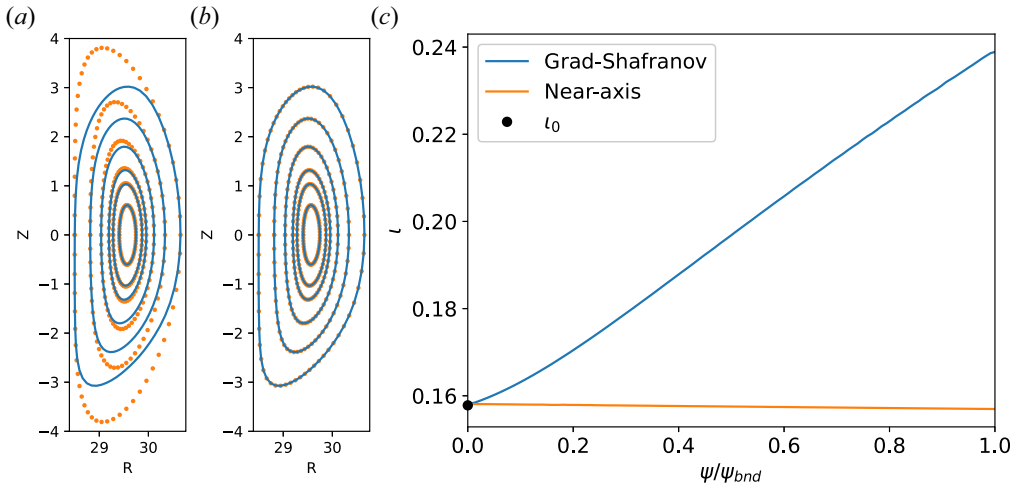


FIGURE 2. (a) A comparison of the flux surfaces computed from (4.6) (lines) to the closest near-axis approximation (dots). (b) Flux surfaces computed from (4.6) (lines) compared with the flux surfaces computed by VMEC (dots). The flux surfaces are shown for  $F_0\Psi/\pi = 0.1, 0.3, 0.5, 1, 2, 4 \text{ T m}^2$ . (c)  $\iota$  profiles computed in the Grad–Shafranov model, the near-axis model and the constant  $\iota_0$ .

value of  $\Psi$ , then finding  $F = d\Psi/d\psi$  and using the formula

$$F(\psi) = \frac{G(\psi) + NI(\psi)}{\iota(\psi) - N}, \tag{5.2}$$

which is given by Helander (2014) as an unnumbered equation. Again, these agree well with each other and the  $\iota_0$  constant near the axis, but the Grad–Shafranov solution has a noticeable shear. The  $\iota$  profiles calculated by VMEC (not shown here) do not match the Grad–Shafranov and near-axis profiles shown in figure 2(c), with the  $\iota$  on axis in the corresponding VMEC solutions being greater by approximately 0.06 and 0.024, respectively. A similar mismatch with VMEC has been observed by Sengupta *et al.* (2024b); it likely appears because the boundary (and not the axis) is fixed in VMEC, so for a finite aspect ratio, the VMEC axis does not match the axis given by (5.1). There are also additional complications arising from VMEC having a coordinate singularity on the axis (Panici *et al.* 2023a). Just as was reported by Sengupta *et al.* (2024b), the agreement between the  $\iota$  on axis in VMEC and that predicted by the Grad–Shafranov model improves as the aspect ratio is increased. Finally, figure 3 shows that the maximum quasisymmetry error, defined as

$$\max_{\psi} \sqrt{\frac{\sum_{n \neq 0} \hat{B}_{n,m}(\psi)^2}{\sum_{n,m} \hat{B}_{n,m}(\psi)^2}}, \tag{5.3}$$

and calculated in VMEC equilibria with varying aspect ratios, scales as  $\epsilon^2$ . Here,  $\hat{B}_{n,m}(\psi)$  are the Fourier modes of  $|\mathbf{B}|$  on flux surface  $\psi$ . This behaviour is to be expected, since we have only ensured quasisymmetry at order  $\epsilon$ .

### 6. Conclusion

We have derived a first-order asymptotic model for high- $\beta$  ( $\beta = O(\epsilon)$ ) quasisymmetric stellarators under the assumption that the vacuum magnetic field is dominant ( $\epsilon \ll 1$ ),

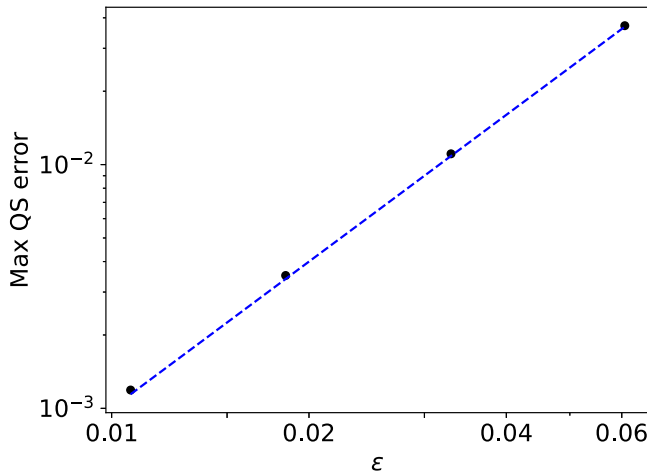


FIGURE 3. Maximum quasisymmetry error (black dots) scales as  $\epsilon^2$  (dashed blue line is  $10\epsilon^2$ ).

but non-vacuum effects can still contribute ( $\nabla\chi \cdot \nabla = O(\epsilon)$ ). We have shown that these assumptions require the aspect ratio to be large and a simple expression is obtained for the lowest-order vacuum field ( $\nabla\chi = B_0\nabla l$ ). The first-order correction, which involves both vacuum and current-driven fields, is then governed by a Grad–Shafranov equation and the requirement that flux surfaces exist ( $\mathbf{B} \cdot \nabla\Psi = 0$ ). Unlike the more simple near-axisymmetric case considered in our previous work (Sengupta *et al.* 2024a), which is a special case of the present model, the overconstraining problem cannot be resolved in general. Thus, one must look for special solutions that satisfy both equations simultaneously. One family of such solutions, and thus another special case of this model, is the first-order quasisymmetric near-axis solutions.

We also provide several new solutions which are outside the scope of both the near-axis model and our previous near-axisymmetric model, and show a numerical example of one of these new solutions. When comparing the numerical results from the present model to VMEC, the flux surfaces match well, but the  $\iota$  profiles show a mismatch which decreases as the aspect ratio is increased. The mismatch is most likely due to VMEC shifting the axis during minimization, so the VMEC axis does not exactly match the axis used when solving the Grad–Shafranov equation. We plan to use DESC in future work to re-evaluate the agreement between  $\iota$  profiles. DESC has the option of fixing the axis (Panici *et al.* 2023b) and avoids large errors on axis, which are typical for VMEC, by using Zernike polynomials (Panici *et al.* 2023a). Finally, DESC can handle highly shaped axes better than VMEC, and is a more appropriate tool for doing a numerical verification of the analytical solution in § 4.1, as all of the non-planar constant curvature axes that we have found so far are highly shaped.

We expect that the analytical solutions found in this paper are only a small subset of all possible solutions to the model we derived. In future work, we plan to implement a numerical solver that will look for solutions by constructing a function basis in  $C_1, C_2$  space and looking for functions and boundary shapes that minimize the residual in the Grad–Shafranov equation (2.14).

Another line of work that we intend to pursue is the construction of a similar model in the low- $\beta$  ( $\beta = O(\epsilon^2)$ ) regime. As discussed in § 2, in the low- $\beta$  regime, there are no constraints on  $\chi$ , aside from satisfying the Laplace equation. This allows for a richer set of solutions, including compact stellarators, at the expense of analytical progress being

more difficult. Nevertheless, as will be shown in a future publication, some analytical progress, such as finding the characteristics of the  $\mathbf{B} \cdot \nabla \Psi = 0$  equation, is still possible. Finally, the low- $\beta$  model is closely related to the reduced MHD model implemented in the JOEK code (Hoelzl *et al.* 2021; Nikulsin 2021; Nikulsin *et al.* 2021), with the main difference being that the JOEK models use an ansatz instead of an ordering. Implementing a numerical solver for this model will allow one to initialize JOEK simulations of quasisymmetric stellarators from the Grad–Shafranov solutions instead of importing GVEC equilibria.

### Acknowledgements

The authors thank P. Helander, S. Buller, E. Paul, E. Rodriguez, A. Brown, S. Hudson, M. Landreman, F.P. Diaz, H. Zhu and V. Duarte for fruitful discussions.

*Editor Per Helander thanks the referees for their advice in evaluating this article.*

### Funding

This research was supported by a grant from the Simons Foundation/SFARI (560651, AB) and the Department of Energy Award No. DE-SC0024548.

### Declaration of interests

The authors report no conflict of interest.

### Appendix A. Perpendicular force balance terms with general $B_v$

In this appendix, we will provide a rigorous proof that the two terms in (2.4) are linearly independent when  $B_v \neq B_0 + O(\epsilon)$ . First, suppose that the two terms are not linearly independent; then we can write

$$\nabla \left( \frac{B_1}{\mu_0 B_v} \right) + \frac{1}{B_v^2} \nabla p = f \nabla g, \quad (\text{A1})$$

where  $f$  and  $g$  are some functions to be determined. If the above equation holds, then the left-hand side should be orthogonal to its own curl:

$$0 = \left[ \nabla \left( \frac{B_1}{\mu_0 B_v} \right) + \frac{1}{B_v^2} \nabla p \right] \cdot \nabla \times \left[ \nabla \left( \frac{B_1}{\mu_0 B_v} \right) + \frac{1}{B_v^2} \nabla p \right] = \frac{2}{\mu_0 B_v^4} \nabla B_1 \cdot (\nabla p \times \nabla B_v). \quad (\text{A2})$$

Thus,  $B_1$  must be a function of only  $p$  and  $B_v$ . Using this fact and dotting (A1) with  $\nabla p \times \nabla B_v$ , we get  $\nabla g \cdot (\nabla p \times \nabla B_v) = 0$  if  $f \neq 0$ ; thus  $g$  must also be a function of only  $p$  and  $B_v$ .

For (2.4) to be satisfied, we must have either  $f = 0$  or  $g = \text{const}$  or  $g = g(\chi)$ . Consider the latter case first. Since  $g(\chi) = g(p, B_v)$ , this is an equation that can be solved for  $B_v$ , giving  $B_v = B_v(p, \chi)$ . If we insert this into the quasisymmetry condition (2.7a), we will get  $\mathbf{B} \cdot \nabla \chi \partial B_v / \partial \chi = 0$  since  $\mathbf{B} \cdot \nabla p = 0$ , meaning that  $B_v$  must be a flux function, which is a contradiction since  $|\mathbf{B}| = B_v + O(\epsilon^2)$  and  $|\mathbf{B}|$  can only be a flux function in axisymmetry (Schief 2003).

Alternatively, if either  $f = 0$  or  $g = \text{const}$ , (A1) will have the following two components in the  $\nabla p$  and  $\nabla B_v$  directions:

$$\frac{1}{\mu_0 B_v} \frac{\partial B_1}{\partial p} + \frac{1}{B_v^2} = 0, \quad \frac{1}{\mu_0 B_v} \frac{\partial B_1}{\partial B_v} - \frac{B_1}{\mu_0 B_v} = 0. \quad (\text{A3a,b})$$

These two equations for  $B_1$  are incompatible. Integrating the first one, we get  $B_1/\mu_0 = -p/B_v + u(B_v)$ , where  $u$  is an arbitrary function. Inserting this into the second equation, we get  $u' - u/B_v = -2p/B_v^2$ , which is a contradiction since  $u$  cannot depend on  $p$ . Finally, note that the contradiction is resolved if  $B_v = B_0 + O(\epsilon)$ , since  $\nabla B_v = O(\epsilon)$  in that case, which will remove the second equation, while the solution to the first equation will become the pressure-balance relation (2.5).

## Appendix B. Near-axisymmetric and near-helically-symmetric solutions

Previously (Sengupta *et al.* 2024a), we derived a condition for consistency between the Grad–Shafranov equation and the  $\mathbf{B} \cdot \nabla \Psi = 0$  equation, and then used it to obtain a condition on  $a$  under which the two equations are consistent for all solutions  $\Psi$ . We can attempt a similar approach for the present model, but it will exclude many solutions of interest, including all solutions discussed in §§ 3 and 4.

To obtain the consistency condition, we apply the  $\mathbf{B} \cdot \nabla$  operator to (2.14), and commute  $\mathbf{B} \cdot \nabla$  with  $\Delta_\perp$ :

$$\Delta_\perp \mathbf{B} \cdot \nabla \Psi + [\mathbf{B} \cdot \nabla, \Delta_\perp] \Psi + \mathbf{B} \cdot \nabla \left( \tau - \frac{\partial^2 a}{\partial x^2} - \frac{2\mu_0}{B_0^2} \frac{dp}{d\Psi} \kappa x \right) = 0. \quad (\text{B1})$$

The whole equation must be satisfied if (2.14) is satisfied; however, the first term must be individually zero since  $\mathbf{B} \cdot \nabla \Psi = 0$ . Removing the first term and working out the commutator, the following consistency condition is obtained:

$$\begin{aligned} 2 \frac{\partial^2 \Psi}{\partial x \partial y} \left( \tau + \frac{\partial^2 a}{\partial x^2} \right) + 2 \left( \frac{\partial^2 \Psi}{\partial x^2} - \frac{\partial^2 \Psi}{\partial y^2} \right) \frac{1}{\kappa} \frac{d\kappa}{dl} \\ + \frac{\partial^3 a}{\partial x^3} \left( \frac{\partial \Psi}{\partial y} + \frac{1}{\kappa} \frac{d\kappa}{dl} x \right) + \frac{\partial}{\partial \ell} \left( \tau - \frac{\partial^2 a}{\partial x^2} \right) = 0. \end{aligned} \quad (\text{B2})$$

The only case where it is satisfied independent of  $\Psi$  is if  $a(x, l) = a_0(l) + xa_1(l) - \tau x^2/2$  and  $d\tau/dl = d\kappa/dl = O(\epsilon^3)$ . When  $\tau = 0$  and  $\kappa = 1/R_0$ , this consistency condition reduces to that derived previously (Sengupta *et al.* 2024a), which limited the stellarator to being a perturbed tokamak. In addition to the vertical perturbation discussed previously (Sengupta *et al.* 2024a), the condition above explicitly allows the tokamak to be perturbed in the  $R$ -direction; such a perturbation, if it is of the order of the minor radius, will result in an  $O(\epsilon^2)$  correction to  $\kappa$ , which still satisfies  $d\kappa/dl = O(\epsilon^3)$ . Finally, when  $\tau \neq 0$ , the axis becomes an open helix, which corresponds to the straight stellarator limit. Similar to the tokamak, the straight stellarator can be perturbed in both the normal and binormal directions by manipulating  $\kappa$  and  $a_1$ .

## REFERENCES

- BOOZER, A.H. 1998 What is a stellarator? *Phys. Plasmas* **5** (5), 1647–1655.
- CONNOR, J.W., HASTIE, R.J. & TAYLOR, J.B. 1978 Shear, periodicity, and plasma ballooning modes. *Phys. Rev. Lett.* **40**, 396–399.
- FREIDBERG, J.P. 2014 *Ideal MHD*. Cambridge University Press.
- GARREN, D.A. & BOOZER, A.H. 1991a Existence of quasihelically symmetric stellarators. *Phys. Fluids* **B 3** (10), 2822–2834.
- GARREN, D.A. & BOOZER, A.H. 1991b Magnetic field strength of toroidal plasma equilibria. *Phys. Fluids* **B 3** (10), 2805–2821.
- HELANDER, P. 2014 Theory of plasma confinement in non-axisymmetric magnetic fields. *Rep. Prog. Phys.* **77** (8), 087001.

- HIRSHMAN, S.P. & WHITSON, J.C. 1983 Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria. *Phys. Fluids* **26** (12), 3553–3568.
- HOELZL, M., HUIJSMANS, G.T.A., PAMELA, S.J.P., BÉCOULET, M., NARDON, E., ARTOLA, F.J., NKONGA, B., ATANASIU, C.V., BANDARU, V., BHOLE, A., *et al.* 2021 The JOREK non-linear extended MHD code and applications to large-scale instabilities and their control in magnetically confined fusion plasmas. *Nucl. Fusion* **61** (6), 065001.
- JORGE, R., SENGUPTA, W. & LANDREMAN, M. 2020 Construction of quasisymmetric stellarators using a direct coordinate approach. *Nucl. Fusion* **60** (7), 076021.
- KRUGER, S.E., HEGNA, C.C. & CALLEN, J.D. 1998 Generalized reduced magnetohydrodynamic equations. *Phys. Plasmas* **5** (12), 4169–4182.
- LANDREMAN, M. 2022 Mapping the space of quasisymmetric stellarators using optimized near-axis expansion. *J. Plasma Phys.* **88** (6), 905880616.
- LANDREMAN, M., JORGE, R., RODRIGUEZ, E. & DUDT, D. 2020–2023 pyQSC. <https://github.com/landreman/pyQSC>. Software without accompanying journal publication.
- LANDREMAN, M. & SENGUPTA, W. 2018 Direct construction of optimized stellarator shapes. Part 1. Theory in cylindrical coordinates. *J. Plasma Phys.* **84** (6), 905840616.
- LANDREMAN, M. & SENGUPTA, W. 2019 Constructing stellarators with quasisymmetry to high order. *J. Plasma Phys.* **85** (6), 815850601.
- LANDREMAN, M., SENGUPTA, W. & PLUNK, G.G. 2019 Direct construction of optimized stellarator shapes. Part 2. Numerical quasisymmetric solutions. *J. Plasma Phys.* **85** (1), 905850103.
- NIKULSIN, N. 2021 Models and methods for nonlinear magnetohydrodynamic simulations of stellarators. PhD thesis, Technische Universität München.
- NIKULSIN, N., HOELZL, M., ZOCCO, A., LACKNER, K., GÜNTER, S. & THE JOREK TEAM 2021 Testing of the new JOREK stellarator-capable model in the tokamak limit. *J. Plasma Phys.* **87** (3), 855870301.
- PANICI, D., CONLIN, R., DUDT, D.W., UNALMIS, K. & KOLEMEN, E. 2023a The DESC stellarator code suite. Part 1. Quick and accurate equilibria computations. *J. Plasma Phys.* **89** (3), 955890303.
- PANICI, D., RODRIGUEZ, E., CONLIN, R., KIM, P., DUDT, D., UNALMIS, K. & KOLEMEN, E. 2023b Near-axis constrained stellarator equilibria with DESC. In *APS Division of Plasma Physics Meeting Abstracts*, Conference Presented at 2023 APS-DPP Denver, CO, pp. GP11–114.
- RODRÍGUEZ, E. & BHATTACHARJEE, A. 2021 Solving the problem of overdetermination of quasisymmetric equilibrium solutions by near-axis expansions. II. Circular axis stellarator solutions. *Phys. Plasmas* **28** (1), 012509.
- RODRIGUEZ, E. 2022 Quasisymmetry. PhD thesis, Princeton University.
- SCHIEF, W.K. 2003 Nested toroidal flux surfaces in magnetohydrostatics. Generation via soliton theory. *J. Plasma Phys.* **69** (6), 465–484.
- SENGUPTA, W., NIKULSIN, N., GAUR, R. & BHATTACHARJEE, A. 2024a Asymptotic quasisymmetric high-beta three-dimensional magnetohydrodynamic equilibria near axisymmetry. *J. Plasma Phys.* **90** (2), 905900209.
- SENGUPTA, W., RODRIGUEZ, E., JORGE, R., LANDREMAN, M., BHATTACHARJEE, A. 2024b Stellarator equilibrium axis-expansion to all orders in distance from the axis for arbitrary plasma beta. *J. Plasma Phys.* **90** (4), 905900407.
- SOLOV'EV, L.S. & SHAFRANOV, V.D. 1970 *Reviews of Plasma Physics*, vol. 5. Consultants Bureau.
- STRAUSS, H.R. 1997 Reduced MHD in nearly potential magnetic fields. *J. Plasma Phys.* **57** (1), 83–87.
- ZOCCO, A., HELANDER, P. & WEITZNER, H. 2020 Magnetic reconnection in 3D fusion devices: non-linear reduced equations and linear current-driven instabilities. *Plasma Phys. Control. Fusion* **63** (2), 025001.