J. Plasma Phys. (2025), *vol*. 91, E73 © The Author(s), 2025. Published by Cambridge University Press.
This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited. doi:10.1017/S0022377825000388

High-order magnetic near-axis expansion: ill-posedness and regularisation

Maximilian Ruth[1], **Rogerio Jorge**[2] **and David Bindel**[3]

[1]Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, TX 78712, USA
[2]Department of Physics, University of Wisconsin-Madison, WI 53706, USA
[3]Department of Computer Science, Cornell University, Ithaca, NY 14853, USA

Corresponding author: M. Ruth, maximilian.ruth@austin.utexas.edu

(Received 20 November 2024; revision received 17 March 2025; accepted 18 March 2025)

When analysing stellarator configurations, it is common to perform an asymptotic expansion about the magnetic axis. This so-called near-axis expansion is convenient for the same reason asymptotic expansions often are, namely, it reduces the dimension of the problem. This leads to convenient and quickly computed expressions of physical quantities, such as quasisymmetry and stability criteria, which can be used to gain further insight. However, it has been repeatedly found that the expansion diverges at high orders in the distance from axis, limiting the physics the expansion can describe. In this paper, we show that the near-axis expansion diverges in vacuum due to ill-posedness and that it can be regularised to improve its convergence. Then, using realistic stellarator coil sets, we demonstrate numerical convergence of the vacuum magnetic field and flux surfaces to the true values as the order increases. We numerically find that the regularisation improves the solutions of the near-axis expansion under perturbation, and we demonstrate that the radius of convergence of the vacuum near-axis expansion is correlated with the distance from the axis to the coils.

Key words: plasma confinement, plasma simulation

1. Introduction

The design of stellarators is a computationally intensive task. The most basic problem in stellarator design – that of computing the magnetic field – requires solving the steady-state magnetohydrostatics (MHS) equations. These equations are difficult to solve for reasons familiar to many problems in physics: they are nonlinear and three-dimensional. Popular MHS equilibrium solvers include VMEC (Hirshman 1983), DESC (Dudt & Kolemen 2020) and SPEC (Hudson *et al.* 2012), all of which take seconds to minutes to compute a single equilibrium. Beyond equilibrium solving, there are potentially many other stellarator objectives that are expensive to compute, with plasma stability metrics being a major example. When optimising for stellarators, the costs of equilibrium and objective solving can limit the speed

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

of the overall design process. This, in combination with the high dimensionality of specifying three-dimensional (3-D) fields, motivates a need for simpler alternatives.

Recently, near-axis expansion (Mercier 1964; Solov'ev & Shafranov 1970) has gained traction as an alternative to full 3-D MHS solvers. The near-axis expansion works by asymptotically expanding all of the relevant plasma variables (such as magnetic coordinates, pressure, rotational transform and plasma current) in the distance from the magnetic axis, which is assumed to be small relative to a characteristic magnetic scale length. The resulting equations are a hierarchy of one-dimensional ordinary differential equations (ODEs), which can be solved orders of magnitude faster than 3-D equilibria. This allows for one to quickly find large numbers of optimised stellarators (Landreman 2022; Giuliani 2024), something that was previously unavailable to the stellarator community.

In addition to the speed, the near-axis expansion has other benefits. For instance, Garren & Boozer (1991) showed that quasisymmetry imposes more constraints than free parameters in the expansion, leading to the conjecture that non-axisymmetric but perfectly quasisymmetric stellarators cannot exist. Many objectives have been defined and computed for the near-axis expansion, including quasisymmetry (Landreman & Sengupta 2019), quasi-isodynamicity (Mata *et al.* 2022), isoprominence (Burby *et al.* 2023), and Mercier and magnetic-well conditions for stability (Landreman & Jorge 2020; Kim *et al.* 2021). There is evidence that other higher-order effects such as ballooning and linear gyrokinetic stability could be investigated as well (Jorge & Landreman 2020). The near-axis expansion has also been combined with a type of quadratic flux minimising surfaces and coil optimisation to create free-field optimised quasi-axisymmetric (QA) equilibria (Giuliani 2024). In sum, the connection among easily expressed objectives, a relatively low-dimensional equilibrium description and fast computation has led to the increased use of the near-axis expansion.

However, the near-axis expansion is not without drawbacks. The primary drawback is fundamental: the expansion has limited accuracy far from the axis. For instance, in the 'far-axis' regime, there can be large errors in the magnetic shear and magnetic surfaces can self-intersect (Landreman 2022). The paper by Jorge & Landreman (2020) also indicates that higher-order terms may be needed for stability; such as magnetic curvature terms. Unfortunately, attempts to use higher-order terms have resulted in divergent asymptotic series, limiting the accuracy to small plasma volumes. Most series go to first, second or sometimes third order in the distance from the axis in the relevant quantities, with any more terms typically reducing accuracy rather than improving it. Therefore, if we want to include more physics objectives over larger volumes in the near-axis expansion, we must overcome the issue of series divergence.

Unfortunately, the issue of divergence is confounded by many of the assumptions that can be incorporated into the near-axis expansion. The most extreme case is that of QA stellarators, where it has been shown that the system of equations for quasi-axisymmetry is overdetermined beyond third order in the expansion. Obviously, unless one relaxes the problem (e.g. via anisotropic pressure; Rodríguez & Bhattacharjee 2021), one cannot generally ask for a convergent QA near-axis expansion in such a circumstance. In the simpler case of non-quasisymmetric stellarators with smooth pressure gradients and nested surfaces, it is still unknown whether there are non-axisymmetric solutions to MHS (Grad 1967; Constantin *et al.* 2021*a*).

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

Recent work has found that perturbing for small force (Constantin *et al.* 2021*b*) or non-flat metrics (Cardona *et al.* 2024) allow for integrable solutions, but currently, there is no guarantee of solutions of MHS, let alone convergent asymptotic expansions.

So, to begin the task of building convergent numerical methods for the near-axis expansion, we focus on a problem we know is solvable: Laplace's equation for vacuum magnetic potentials following Jorge *et al.* (2020). This can be solved in direct (Mercier) coordinates (Mercier 1964) with no assumption of nested surfaces. Additionally, because solutions of Laplace's equation are real analytic, there exist near-axis expansions of the equation that converge within a neighbourhood of the axis. Despite these guarantees, even the near-axis expansion of Laplace's equation diverges.

In this paper, we show that the vacuum near-axis expansion diverges for a reason: Laplace's equation as a near-axis expansion is ill-posed (§ 3, following background in § 2). To address this issue, we introduce a small regularisation term to Laplace's equation and expand to find a regularised near-axis expansion. We do this by including a viscosity term to Laplace's equation that damps the highly oscillatory unstable modes responsible for the ill-posedness. By appropriately bounding the input of the near-axis expansion in a Sobolev norm, we prove that this term results in a uniformly converging near-axis expansion within a neighbourhood of the axis.

Following the theory, we describe a pseudo-spectral method for finding solutions to the near-axis expansion to arbitrary order in § 4. In § 5, we use the numerical method to show two examples of high-order near-axis expansions: the rotating ellipse and Landreman–Paul (Landreman & Paul 2022). We find that the near-axis expansion magnetic field, rotational transform and magnetic surfaces can converge accurately near the axis for unperturbed initial data. The region of convergence is observed to be dictated by the distance from the magnetic axis to the coils. Then, by perturbing the on-axis inputs, we show that the regularised expansion obeys Laplace's more accurately farther from the axis. Finally, we conclude in § 6.

2. Background

In this section, we introduce the near-axis expansion for the vacuum field equilibrium problem. This presentation follows closely that of Jorge *et al.* (2020). We begin with a discussion of the geometry of the near-axis problem (§ 2.1), introduce the near-axis expansion (§ 2.2), define the magnetic field problem (§ 2.3) and finally discuss finding straight field-line magnetic coordinates (§ 2.4). For a fuller discussion of the near-axis expansion to all orders, including with pressure gradients, we recommend the papers by Jorge *et al.* (2020) and Sengupta *et al.* (2024). For ease of reference, we have summarised the equations in the background in box (2.30) for the magnetic field and box (2.60) for magnetic coordinates.

2.1. *Near-axis geometry*

We define a magnetic axis as a C^∞ closed curve $\boldsymbol{r}_0 : \mathbb{T} \to \mathbb{R}^3$ with $\boldsymbol{r}_0' \neq 0$ and a non-zero tangent magnetic field (see § 2.3 for details about the field). We define a near-axis domain about the axis with radius σ as

$$\Omega_\sigma = \left\{ \boldsymbol{r} \in \mathbb{R}^3 \mid \forall s \in \mathbb{T},\ \ \|\boldsymbol{r} - \boldsymbol{r}_0(s)\| < \sigma \right\}. \tag{2.1}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

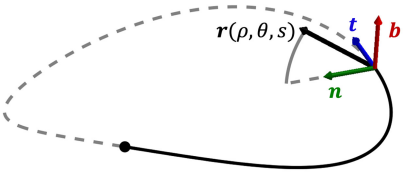


FIGURE 1. Schematic of the direct near-axis Frenet–Serret coordinate frame.

We note that the assumption that the axis is infinitely differentiable is necessary for the near-axis expansion to formally go to arbitrary order, and we will eventually reduce the required regularity for the inputs of the regularised expansion, summarised in box (3.30).

We define a direct coordinate system $\boldsymbol{r} : \Omega_\sigma^0 \to \Omega_\sigma$ where $(\rho, \theta, s) \in \Omega_\sigma^0 = [0, \sigma) \times \mathbb{T}^2$ is the solid torus as (see figure 1)

$$\boldsymbol{r}(\rho, \theta, s) = \boldsymbol{r}_0(s) + Q(s) \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \qquad (x,\ y) = (\rho \cos\theta,\ \rho \sin\theta), \tag{2.2}$$

where Q is an orthonormal basis for the local coordinates at the axis with the tangent vector in the third column, i.e. $\boldsymbol{t} = \boldsymbol{r}_0'/\ell' = Q\boldsymbol{e}_3$, where we assume $\ell' = \left\|\boldsymbol{r}_0'\right\| > 0$. Both the $\boldsymbol{x} = (x, y, s)$ and the $\boldsymbol{q} = (\rho, \theta, s)$ coordinate frames are useful, as $\boldsymbol{x}$ is non-singular with respect to the coordinate transformation, while $\boldsymbol{q}$ diagonalises the near-axis expansion operator. To perform calculus in the $\boldsymbol{q}$ basis, we require the induced metric from $\mathbb{R}^3$. For this, we first compute the coordinate derivative

$$F = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\boldsymbol{q}} = Q \begin{pmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & \ell' \end{pmatrix} + QK\boldsymbol{x}\boldsymbol{e}_3^T, \tag{2.3}$$

where K is an antisymmetric matrix determining the derivative of Q in the near-axis basis $Q'(s) = QK$. Using this matrix, the metric is computed as

$$g = F^T F. \tag{2.4}$$

For the numerical examples in this paper, we specifically consider the Frenet–Serret frame, meaning the local basis Q and its derivative are defined by

$$Q = \begin{pmatrix} | & | & | \\ \boldsymbol{n} & \boldsymbol{b} & \boldsymbol{t} \\ | & | & | \end{pmatrix}, \qquad K = \ell' \begin{pmatrix} 0 & -\tau & \kappa \\ \tau & 0 & 0 \\ -\kappa & 0 & 0 \end{pmatrix}, \tag{2.5}$$

where $\kappa = \|\boldsymbol{t}'\| / \ell'$ is the axis curvature, $\boldsymbol{n} = \|\boldsymbol{t}'\| / (\kappa\ell')$ is the normal vector, $\boldsymbol{b} = \boldsymbol{t} \times \boldsymbol{n}$ is the binormal vector and $\tau = -\left\|\boldsymbol{b}'\right\| / \ell'$ is the axis torsion. Alternative forms of the curvature and torsion are

$$\kappa = \frac{\left\|\boldsymbol{r}_0' \times \boldsymbol{r}_0''\right\|}{(\ell')^3}, \qquad \tau = \frac{(\boldsymbol{r}_0' \times \boldsymbol{r}_0'') \cdot \boldsymbol{r}_0'''}{\left\|\boldsymbol{r}_0' \times \boldsymbol{r}_0''\right\|^2}. \tag{2.6}$$

For the Frenet–Serret coordinate system to be well-defined and non-singular on Ω_σ^0, we require $\sigma^{-1} > \kappa > 0$. In particular, no straight segments are allowed in

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

the Frenet–Serret frame, disallowing quasi-isodynamic (QI) stellarators (Mata *et al.* 2022). An alternative choice for axis coordinates that allow for straight segments is Bishop's coordinates (Bishop 1975; Duignan & Meiss 2021).

Replacing the Frenet–Serret basis into (2.3), we obtain

$$F = Q\begin{pmatrix}\cos\theta & -\rho\sin\theta & -\ell'\tau\rho\sin\theta\\ \sin\theta & \rho\cos\theta & \ell'\tau\rho\cos\theta\\ 0 & 0 & h_s\end{pmatrix}, \qquad g = \begin{pmatrix}1 & 0 & 0\\ 0 & \rho^2 & \ell'\tau\rho^2\\ 0 & \ell'\tau\rho^2 & (\ell')^2\tau^2\rho^2 + h_s^2\end{pmatrix}, \tag{2.7}$$

where

$$h_s(\rho,\theta,s) = \ell'(1-\kappa\rho\cos\theta). \tag{2.8}$$

The local volume ratio is given by

$$\sqrt{g} = \det F = \rho h_s \tag{2.9}$$

and the inverse metric is

$$g^{-1} = \begin{pmatrix}1 & 0 & 0\\ 0 & (\ell')^2\tau^2 h_s^{-2} + \rho^{-2} & -\ell'\tau h_s^{-2}\\ 0 & -\ell'\tau h_s^{-2} & h_s^{-2}\end{pmatrix}. \tag{2.10}$$

To find metrics associated with the more general coordinate system (2.2), we consider transformations of the form $(\rho,\theta,s)\mapsto(\rho,\omega(\theta,s),s)$, where $\omega(\theta,s) = \theta+\lambda(s)$. This transformation rotates the orthonormal frame, changing the metric to

$$g = \begin{pmatrix}1 & 0 & 0\\ 0 & \rho^2 & \ell' T\rho^2\\ 0 & \ell' T\rho^2 & (\ell')^2T^2\rho^2 + h_s^2(\rho,\omega-\lambda,s),\end{pmatrix}, \qquad T = \tau - \frac{\lambda'}{\ell'}. \tag{2.11}$$

In this way, the general set of near-axis frames can be represented by a simple replacement of τ with T. A special case of this transformation is when $T=0$, yielding (Mercier 1964)

$$\lambda = \int_0^s \tau(s)\ell'(s)\,\mathrm{d}s. \tag{2.12}$$

In this coordinate frame, the metric becomes diagonal: $g = \mathrm{Diag}(1,\rho^2,h_s^2(\rho,\omega-\lambda,s))$. The fact that g is diagonalised is convenient for theoretical manipulations, but variables expressed in terms of ω are multivalued for curves with non-integer total torsion. This unfortunate consequence is important for numerical methods, as Fourier series cannot be used in s and additional consistency requirements are necessary. This is part of the motivation for using the Frenet–Serret frame for the numerical examples herein.

2.2. *Near-axis expansion*

Now, we consider expansions of functions $A\in C^\infty(\Omega_\sigma^0)$ about the magnetic axis. We formally expand in small distances from the axis $\rho \ll \min\kappa^{-1}$ as

$$A(\rho,\theta,s) = \sum_{m=0}^{\infty} A_m(\theta,s)\rho^m, \qquad A_m(\theta,s) = \sum_{n=0}^{m} A_{mn}(s)e^{(2n-m)\mathrm{i}\theta},$$
$$A_{mn}(s) = \sum_{k\in\mathbb{Z}} A_{mnk}e^{\mathrm{i}ks}. \tag{2.13}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

This expansion is not guaranteed to converge anywhere for $A \in C^\infty$, but it is asymptotic to A near the axis, i.e.

$$|A - A_{<m}| = \mathcal{O}(\rho^m), \tag{2.14}$$

where we define $A_{<m}$ as the partial sum

$$A_{<m} = \sum_{n=0}^{m-1} A_n \rho^n. \tag{2.15}$$

In defining magnetic coordinates, we also find it convenient to expand A in $\boldsymbol{x}$ as

$$A(x, y, s) = \sum_{\mu=0}^{\infty} \sum_{\nu=0}^{\mu} A_{\mu\nu}(s) x^{\mu-\nu} y^\nu, \tag{2.16}$$

where we use Latin indices for the $\boldsymbol{q}$ frame and Greek indices for the $\boldsymbol{x}$ frame. If we require A to be real-analytic on Ω^0_σ, there additionally exists a $\sigma' < \sigma$ so that the the asymptotic series converges uniformly on $\Omega^0_{\sigma'}$ (this does not necessarily extend to all of Ω^0_σ).

Throughout this paper, we attempt to minimise the number of complicated summation formulae resulting from the near-axis expansion (NAE). For instance, if we have two functions $A, B \in C^\infty$ and we want the mth component of the series of $C = AB$, we will write $C_m = (AB)_m$, rather than

$$C_m = \sum_{\ell=0}^{m} A_{m-\ell} B_\ell. \tag{2.17}$$

As expressions become increasingly complicated, this notation provides a concise description of the mathematics involved. In Appendix B, we define a number of relevant operations on series that the interested reader can use to expand the expressions within this paper. In § 4, we discuss how this is similarly convenient for the purposes of programming NAE operations. Rather than implementing residuals via complicated summation formulae, the operations in Appendix B are called, allowing for a simple framework for developing new code.

An important exception to the general rule of condensing notation is in defining any linear operators that must be inverted through the course of an asymptotic expansion. Detailed understanding of such operators are necessary for both numerical implementation and analysis on the series.

2.3. *Vacuum fields*

In the steady state, the vacuum magnetic field $\boldsymbol{B} \in C^\infty(\Omega^0_\sigma)$ satisfies

$$\nabla \cdot \boldsymbol{B} = 0, \qquad \boldsymbol{J} = \nabla \times \boldsymbol{B} = 0, \tag{2.18}$$

where $\boldsymbol{J}$ is the plasma current density. The fundamental near-axis assumption is that $\boldsymbol{B}$ evaluated on the axis is non-zero and parallel to the magnetic axis, i.e. for some $B_0 \in C^\infty(\mathbb{T})$,

$$\boldsymbol{B}(0, \theta, s) = B_0(s)\boldsymbol{t}(s). \tag{2.19}$$

Off the axis, we express the magnetic field on Ω^0_σ as

$$\boldsymbol{B} = \nabla\phi + \tilde{\boldsymbol{B}}, \qquad \tilde{\boldsymbol{B}} = \nabla\left[\int_0^s B_0(s')\ell'(s')\,\mathrm{d}s'\right] = B_0(s)\ell'(s)\nabla s, \tag{2.20}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

where $\phi \in C^\infty(\Omega_\sigma^0)$ is the magnetic potential satisfying $\nabla\phi|_{\rho=0} = 0$. Then, taking the divergence of (2.20), we find the magnetic scalar potential satisfies Poisson's equation,

$$\Delta\phi = -\nabla \cdot \tilde{\boldsymbol{B}}. \tag{2.21}$$

By construction, the field $\boldsymbol{B}$ in (2.20) is locally the gradient of some function, so it is curl-free. This means the contributions from $\tilde{\boldsymbol{B}}$ in (2.20) can locally be absorbed to recover Laplace's equation for the potential. However, because Ω_σ is not simply connected and $\boldsymbol{B} \cdot \boldsymbol{t}$ is single-valued on axis, the closed-loop axis integral $\oint \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{\ell}$ demonstrates that it is not possible for $\boldsymbol{B}$ to be globally the gradient of a single-valued function ϕ. In contrast, the Poisson's equation formulation contains only single-valued functions, which is convenient both numerically and analytically.

In coordinates, we can write the magnetic field in (2.20) as

$$B^i = g^{ij}\frac{\partial\phi}{\partial q^j} + \tilde{B}^i, \qquad \tilde{B}^i = B_0\ell' g^{ij}\frac{\partial s}{\partial q^j}, \tag{2.22}$$

where we assume summation over repeated indices and g_{ij} and g^{ij} are the components of the metric and inverse metric, respectively. Then, Poisson's equation in coordinates becomes

$$\sqrt{g}^{-1}\frac{\partial}{\partial q^i}\left(\sqrt{g}g^{ij}\frac{\partial\phi}{\partial q^j}\right) = -\sqrt{g}^{-1}\frac{\partial}{\partial q^i}\left(\sqrt{g}\tilde{B}^i\right). \tag{2.23}$$

Multiplying this by h_s and expanding this in the $\boldsymbol{q}$ coordinate system, we have

$$\begin{aligned}&\frac{1}{\rho^2}\left[\rho\frac{\partial}{\partial\rho}\left(\mathsf{A}\rho\frac{\partial\phi}{\partial\rho}\right) + \frac{\partial}{\partial\theta}\left(\mathsf{B}\frac{\partial\phi}{\partial\theta}\right)\right] = \\ &-\frac{\partial}{\partial\theta}\left(\mathsf{C}\left(\frac{\partial\phi}{\partial s} + B_0\ell'\right)\right) - \frac{\partial}{\partial s}\left(\mathsf{C}\frac{\partial\phi}{\partial\theta}\right) - \frac{\partial}{\partial s}\left(\mathsf{D}\left(\frac{\partial\phi}{\partial s} + B_0\ell'\right)\right),\end{aligned} \tag{2.24}$$

where

$$\mathsf{A} = h_s, \qquad \mathsf{B} = h_s + \rho^2 h_s^{-1}(\ell')^2\tau^2, \qquad \mathsf{C} = -h_s^{-1}\ell'\tau, \qquad \mathsf{D} = h_s^{-1}. \tag{2.25}$$

From here, we can substitute the asymptotic expansions of ϕ and the coefficients A, B, C and D into (2.24). At each order in ρ, Poisson's equation becomes

$$\begin{aligned}\ell'\left(m^2 + \frac{\partial^2}{\partial\theta^2}\right)\phi_m &= -(\nabla\cdot\tilde{B})_{m-2} - (\Delta\phi_{<m})_{m-2} \\ &= -m(m-1)\mathsf{A}_1\phi_{m-1} - \left[\frac{\partial}{\partial\theta}\left(\mathsf{B}\frac{\partial\phi_{<m}}{\partial\theta}\right)\right]_m \\ &\quad - \left[\frac{\partial}{\partial\theta}\left(\mathsf{C}\left(\frac{\partial\phi_{<m}}{\partial s} + B_0\ell'\right)\right) + \frac{\partial}{\partial s}\left(\mathsf{C}\frac{\partial\phi_{<m}}{\partial\theta}\right)\right]_{m-2} \\ &\quad - \left[\frac{\partial}{\partial s}\left(\mathsf{D}\left(\frac{\partial\phi_{<m}}{\partial s} + B_0\ell'\right)\right)\right]_{m-2},\end{aligned} \tag{2.26}$$

where $\mathsf{A}_1 = (h_s)_1 = -\ell'\kappa\cos\theta$. The right-hand side of (2.26) does not depend on orders of ϕ higher than ϕ_{m-1}, so inverting the left-hand-side operator gives an iteration for obtaining ϕ_m at each order.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

However, the operator $\ell'(m^2 + (\partial^2/\partial\theta^2))$ is singular, so we must confirm there are no secular terms. Specifically, we always have an unknown homogeneous solution at $\mathcal{O}(m)$ of the form $\phi_{m0}e^{-im\theta} + \phi_{mm}e^{im\theta}$. For this, we use Fredholm's alternative, which states that (2.26) is solvable if the right-hand side is orthogonal to the null space of the adjoint of the operator on the left-hand side. Because the operator is self-adjoint, the right-hand side must be orthogonal to $e^{\pm im\theta}$, i.e.

$$\left\langle m(m-1)\mathsf{A}_1\phi_{m-1} + \left[\frac{\partial}{\partial\theta}\left(\mathsf{B}_{>0} * \frac{\partial\phi_{<m}}{\partial\theta}\right)\right]_m + \left[\frac{\partial}{\partial\theta}\left(\mathsf{C}\left(\frac{\partial\phi_{<m}}{\partial s} + B_0\ell'\right)\right)\right]_{m-2} + \left[\frac{\partial}{\partial s}\left(\mathsf{C}\frac{\partial\phi}{\partial\theta}\right) + \frac{\partial}{\partial s}\left(\mathsf{D}\left(\frac{\partial\phi_{<m}}{\partial s} + B_0\ell'\right)\right)\right]_{m-2}, \; e^{\pm im\theta}\right\rangle = 0, \tag{2.27}$$

where the inner product is defined by

$$\langle f, \; g\rangle = \int_0^{2\pi} f(\theta)\overline{g}(\theta)\,\mathrm{d}\theta. \tag{2.28}$$

The $m-2$ coefficient of any analytic function is orthogonal to $e^{im\theta}$, so we can remove the C and D terms. The same argument allows us to remove the torsion terms in B. This means only contributions from $\mathsf{A}_1 = \mathsf{B}_1 = -\ell'\kappa\cos\theta$ and ϕ_{m-1} survive, so we only need to verify

$$\left\langle m(m-1)\phi_{m-1}\cos\theta + \left[\frac{\partial}{\partial\theta}\left(\frac{\partial\phi_{m-1}}{\partial\theta}\cos\theta\right)\right]_m, \; e^{\pm im\theta}\right\rangle = 0. \tag{2.29}$$

Using the identity $\cos\theta = (e^{i\theta} + e^{-i\theta})/2$, a quick calculation confirms the above identity holds.

Now, we consider the problem where ϕ is unknown. The fact that the Fredholm condition is automatically satisfied at each order implies that ϕ_{m0} and ϕ_{mm} are free parameters at each order in the near-axis expansion. So, these coefficients are an infinite-dimensional set of initial conditions for the near-axis expansion of Poisson's equation. Intuitively, one can think of the coefficients ϕ_{m0} and ϕ_{mm} as specifying the Fourier coefficients of an infinitely thin tube about the magnetic axis. In this way, the imposition of conditions at each order compensates for the fact that partial differential equations (PDEs) typically satisfy conditions on co-dimension-1 surfaces, whereas the near-axis expansion is specified on a co-dimension-2 curve.

In addition to ϕ_{m0} and ϕ_{mm} as free parameters, we also treat $\boldsymbol{r}_0$ and B_0 as inputs to the near-axis problem. The requirement that the magnetic field be tangent to the axis with magnitude B_0 results in a constraint that $\phi_{00} = \phi_{10} = \phi_{11} = 0$. In total, the direct vacuum near-axis problem can be written as

input:	axis $\boldsymbol{r}_0 \in C^\infty$, on-axis field $B_0 \in C^\infty$,
	higher moments ϕ_{m0}, $\phi_{mm} \in C^\infty$ for $m \geqslant 2$,
assuming:	$\ell' > 0, \; \kappa > 0, \; B_0 > 0, \; \phi_{00} = \phi_{10} = \phi_{11} = 0,$
solve:	$\ell'\left(m^2 + \dfrac{\partial^2}{\partial\theta^2}\right)\phi_m = -(\nabla\cdot\tilde{\boldsymbol{B}})_{m-2} - (\Delta\phi_{<m})_{m-2},$
output:	potential ϕ, magnetic field $\boldsymbol{B} = \nabla\phi + \tilde{\boldsymbol{B}}$.

(2.30)

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

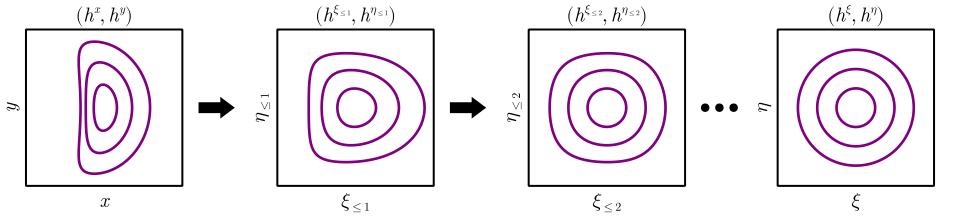


FIGURE 2. A schematic of the process of finding straight field-line coordinates. On the left, we plot the surfaces of the magnetic field (h^x, h^y) on a cross-section for fixed s. Moving one plot to the right, the leading correction transforms to a coordinate frame where the main elliptic component is eliminated. Going one further, the next correction accounts for the most prominent triangularity. This process continues until, in (ξ, η) coordinates, the magnetic surfaces are nested circles.

2.4. *Straight field-line coordinates: leading order*

Given a solution magnetic field from box (2.30), we consider the problem of finding straight field-line magnetic coordinates. We assume that the magnetic field is locally elliptic about the axis and the rotation number is irrational, so that the leading-order behaviour is rotation about the magnetic axis. This means that both hyperbolic orbits (x-points) and on-axis resonant perturbations are excluded from this work. In the language of Hamiltonian normal forms, the leading order field-line dynamics is conjugate to a non-resonant harmonic oscillator; see Burby *et al.* (2021) and Duignan & Meiss (2021) for a more rigorous derivation of magnetic coordinates in the near-axis expansion. We note that our process of finding coordinates is formal: we make no claims that this problem converges in the limit. However, in § 5, we find that this procedure appears to converge well numerically.

To find magnetic coordinates, we attempt to build a conjugacy between magnetic field-line dynamics $\dot{\boldsymbol{r}} = (B^s)^{-1}\boldsymbol{B}(\boldsymbol{r})$ and straight field-line dynamics $\dot{\boldsymbol{\xi}} = (-\iota(\psi)\eta,\ \iota(\psi)\xi,\ 1)$, where $\boldsymbol{\xi} = (\xi,\ \eta,\ s) \in \mathbb{R}^2 \times \mathbb{T}$ are Cartesian coordinates, $\psi = \xi^2 + \eta^2$ is a flux-like coordinate and ι is the rotational transform. To make the connection with straight field-line coordinates precise, consider the transformation to polar coordinates $(\xi, \eta) \to \sqrt{\psi}(\cos\gamma, \sin\gamma)$. Then, the field-line is traced by $(\dot{\psi},\ \dot{\gamma},\ \dot{s}) = (0,\ \iota(\psi),\ 1)$, i.e. magnetic field lines are straight with slope ι. However, we use the Cartesian version of magnetic coordinates because it removes the coordinate singularity associated with polar coordinates, simplifying the following steps.

There are two main steps to our process of finding magnetic coordinates: the leading-order problem and the higher-order problems (see figure 2 for a sketch of the process). If we use the notation $\tilde{\boldsymbol{x}} = (x,\ y)$ and $\tilde{\boldsymbol{\xi}} = (\xi,\ \eta)$ for the out-of-plane coordinates, the leading order transformation takes the form

$$\tilde{\boldsymbol{\xi}}_1 = \frac{\partial \tilde{\boldsymbol{\xi}}_1}{\partial \tilde{\boldsymbol{x}}}\tilde{\boldsymbol{x}} = G_0\tilde{\boldsymbol{x}} = \begin{pmatrix} \xi_{10}x + \xi_{11}y \\ \eta_{10}x + \eta_{11}y \end{pmatrix}. \tag{2.31}$$

We will find that the problem for G_0 is an eigenvalue problem for the on-axis rotation number, as is typical for the linearised dynamics about a fixed point. In the following section, we will discuss the inductive step to higher orders.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

To begin, consider the contravariant form of the Cartesian near-axis magnetic field

$$\frac{1}{B^s}\boldsymbol{B} = \frac{1}{B^s}\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\boldsymbol{x}}\begin{pmatrix} B^x \\ B^y \\ B^s \end{pmatrix} = \frac{B^x}{B^s}\frac{\partial \boldsymbol{r}}{\partial x} + \frac{B^y}{B^s}\frac{\partial \boldsymbol{r}}{\partial y} + \frac{\partial \boldsymbol{r}}{\partial s}. \tag{2.32}$$

We would like to equate this to the straight field-line dynamics as

$$\frac{1}{B^s}\boldsymbol{B} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\boldsymbol{\xi}}\begin{pmatrix} -\iota(\psi)\eta \\ \iota(\psi)\xi \\ 1 \end{pmatrix}, \tag{2.33}$$

where ι depends smoothly upon the radial label as

$$\iota = \iota_0 + \iota_2\psi + \iota_4\psi^2 + \ldots, \tag{2.34}$$

where we emphasise $\iota_\mu = 0$ for odd μ. Multiplying both sides by $\mathrm{d}\boldsymbol{\xi}/\mathrm{d}\boldsymbol{r}$, we find the Floquet conjugacy problem

$$\frac{\partial \tilde{\boldsymbol{\xi}}}{\partial s} = -G(\boldsymbol{\xi})\tilde{\boldsymbol{h}}^{\boldsymbol{x}} + \iota(\psi(\boldsymbol{\xi}))J\tilde{\boldsymbol{\xi}}, \tag{2.35}$$

where

$$G = \frac{\partial \tilde{\boldsymbol{\xi}}}{\partial \tilde{\boldsymbol{x}}}, \qquad \tilde{\boldsymbol{h}}^{\boldsymbol{x}} = \begin{pmatrix} h^x \\ h^y \end{pmatrix} = \frac{1}{B^s}\begin{pmatrix} B^x \\ B^y \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{2.36}$$

where J is known as the symplectic matrix. Our problem is to solve (2.35) for $\xi(x, y)$, $\eta(x, y)$ and $\iota(\psi)$.

To find the leading-order problem for (2.35), we note the magnetic field (h^x, h^y) is linear at leading order:

$$\tilde{\boldsymbol{h}}^{\boldsymbol{x}} = \begin{pmatrix} h^x_{10}x + h^x_{11}y \\ h^y_{10}x + h^y_{11}y \end{pmatrix} + \mathcal{O}(\psi) = H_0(s)\tilde{\boldsymbol{x}} + \mathcal{O}(\psi). \tag{2.37}$$

Substituting this, (2.31) and $\iota = \iota_0 + \mathcal{O}(\psi)$ into (2.35), we have

$$\frac{\partial G_0}{\partial s} + G_0 H_0 = \iota_0 J G_0. \tag{2.38}$$

The leading order problem (2.38) is a Floquet eigenvalue problem for the linearised field-line dynamics about the magnetic axis. Assuming that the near-axis expansion is elliptic at leading order, the value of ι_0 is real. Otherwise, ι_0 is not real, meaning ellipticity can be numerically verified for a given input.

There are many equivalent solutions to (2.38), owing to the symmetries that if (G_0, ι_0) satisfies (2.38), then the following are also solutions:

$$(R(ns)G_0, n + \iota_0), \qquad (JG_0, \iota_0), \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} G_0, -\iota_0, \tag{2.39}$$

where R is a rotation matrix

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{2.40}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

The question of which solution to choose is then a question of practicalities. For instance, one could choose the rotational transform corresponding to a non-twisting right-handed coordinate frame $\boldsymbol{\xi}_1$. Here, 'non-twisting' means that closed coordinate lines near the axis, implicitly defined as curves $\boldsymbol{r}(x, y, s)$ where $\tilde{\boldsymbol{\xi}}(x, y, s) \neq 0$ is held constant, can be continuously deformed to a point on $\mathbb{R}^3 \backslash \boldsymbol{r}_0$. In other words, the coordinates lines do not link with the axis. In this frame, ι_0 agrees with the intuitive definition of the rotational transform as the limiting ratio of poloidal turns divided by toroidal turns of fieldlines about the axis. Other choices may have other benefits, e.g. there may be an eigenfunction that G_0 behaves best numerically. In this paper, we opt for an option that is easy to implement: we take the real solution where ι_0 has the smallest magnitude and $\det G_0$ is positive. From here, other equivalent coordinates can easily be found by applying the transformations in (2.39).

For the scaling of G_0, we choose ψ to be the actual magnetic flux at leading order. The formula for the flux is

$$\psi = \int_{\boldsymbol{r}(D_\psi, 0)} \boldsymbol{B} \cdot \boldsymbol{t} \, \mathrm{d}A = \int_{\boldsymbol{r}(D_\psi, 0)} B_s \, \mathrm{d}A(\boldsymbol{r}), \tag{2.41}$$

where $D_\psi = \{(\xi, \eta) \mid \xi^2 + \eta^2 < \psi\}$ and

$$B_s = g_{sj} B^j = \frac{\partial \phi}{\partial s} + B_0 \ell'. \tag{2.42}$$

Pulling this back to the $\boldsymbol{\xi}$ frame, we have

$$\psi = \int_{D_\psi} B_s(\boldsymbol{x}(\boldsymbol{\xi}, s), 0)(\det G)^{-1} \, \mathrm{d}A(\boldsymbol{\xi}). \tag{2.43}$$

Both B_s and G_0 are constant in $\boldsymbol{\xi}$ to leading order, so (2.43) at leading order becomes

$$\int_{D_\psi} B_s(\boldsymbol{x}(\boldsymbol{\xi}, s), 0)(\det G)^{-1} \, \mathrm{d}A(\boldsymbol{\xi}) = \frac{(B_s)_0}{\det G_0} \pi \psi + \mathcal{O}(\psi^2). \tag{2.44}$$

Setting this equal to ψ, we find

$$\det G_0 = \pi (B_s)_0, \tag{2.45}$$

where we note that this is only possible when $(B_s)_0$ is chosen to be positive.

2.5. *Straight field-line coordinates: higher order*

Now that we have the leading-order behaviour, we iterate to go to higher order. To do so, first define the near-axis expansion of $\tilde{\boldsymbol{\xi}}$ near the axis as

$$\tilde{\boldsymbol{\xi}} = \sum_{\mu=1}^{\infty} \tilde{\boldsymbol{\xi}}_\mu(\tilde{\boldsymbol{x}}, s), \qquad \tilde{\boldsymbol{\xi}}_\mu = \sum_{\nu=0}^{\mu} \tilde{\boldsymbol{\xi}}_{\mu\nu}(s) x^{\mu-\nu} y^\nu, \tag{2.46}$$

and define the partial sums as

$$\tilde{\boldsymbol{\xi}}_{\leqslant \mu} = \sum_{\mu'=1}^{\mu} \tilde{\boldsymbol{\xi}}_{\mu'}, \tag{2.47}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

where at leading order, $\tilde{\boldsymbol{\xi}}_{\leqslant 1} = \tilde{\boldsymbol{\xi}}_1 = G_0 \tilde{\boldsymbol{x}}$.

At each order in the iteration, we consider the update to be a function of the previous coordinates, i.e.

$$\begin{aligned}\tilde{\boldsymbol{\xi}}_{\leqslant \mu+1}\left(\xi_{\leqslant \mu}, \eta_{\leqslant \mu}, s_{\leqslant \mu}\right) &= \tilde{\boldsymbol{\xi}}_{\leqslant \mu} + \sum_{\nu=0}^{\mu} \tilde{\boldsymbol{\xi}}_{\mu\nu}\left(s_{\leqslant \mu}\right) \xi_{\leqslant \mu}^{\mu-\nu} \eta_{\leqslant \mu}^{\nu}, \\ &= \tilde{\boldsymbol{\xi}}_{\leqslant \mu} + \sum_{\nu=0}^{\mu} \tilde{\boldsymbol{\xi}}_{\mu\nu} \xi_1^{\mu-\nu} \eta_1^{\nu} + \mathcal{O}\left(\psi^{(\mu+2)/2}\right). \end{aligned} \tag{2.48}$$

We explicitly write the transformed toroidal coordinate $s_{\leqslant \mu} = s$ so that it is clear that $(\partial/\partial s)$ and $(\partial/\partial s_{\leqslant \mu})$ are different operators. The purpose of performing the update in this way is primarily to make the update step as clear as possible.

To wit, the magnetic field in the new frame satisfies

$$\frac{1}{B^s}\boldsymbol{B} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\boldsymbol{\xi}_{\leqslant \mu}} \begin{pmatrix} h^{\xi_{\leqslant \mu}} \\ h^{\eta_{\leqslant \mu}} \\ 1 \end{pmatrix} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\boldsymbol{\xi}} \begin{pmatrix} -\iota\eta \\ \iota\xi \\ 1 \end{pmatrix}. \tag{2.49}$$

We can use the first equality to write

$$\begin{aligned}\tilde{\boldsymbol{h}}^{\xi_{\leqslant \mu}}(\boldsymbol{\xi}_{\leqslant \mu}) &= \begin{pmatrix} h^{\xi_{\leqslant \mu}}\left(\boldsymbol{\xi}_{\leqslant \mu}\right) \\ h^{\eta_{\leqslant \mu}}\left(\boldsymbol{\xi}_{\leqslant \mu}\right) \end{pmatrix} \\ &= \frac{\partial \tilde{\boldsymbol{\xi}}_{\leqslant \mu}}{\partial \tilde{\boldsymbol{x}}} \tilde{\boldsymbol{h}}^{x}\left(\boldsymbol{x}\left(\boldsymbol{\xi}_{\leqslant \mu}\right)\right) + \frac{\partial \tilde{\boldsymbol{\xi}}_{\leqslant \mu}}{\partial s}, \qquad (2.50) \\ &= [\iota(\psi) J \tilde{\boldsymbol{\xi}}]_{\leqslant \mu} + \left[\frac{\partial \tilde{\boldsymbol{\xi}}_{\leqslant \mu}}{\partial \tilde{\boldsymbol{x}}} \tilde{\boldsymbol{h}}^{x}(\boldsymbol{x}(\boldsymbol{\xi}_{\leqslant \mu}))\right]_{>\mu}, \qquad (2.51) \end{aligned}$$

where we use the notation $f_{>\mu} = \sum_{\nu>\mu} f_\nu$. To get from the second to the third line, we have used the inductive assumption that $\boldsymbol{\xi}_{\leqslant \mu}$ matches $\boldsymbol{\xi}$ up to order μ, meaning that $\boldsymbol{\xi}_{\leqslant \mu}$ is a straight field-line coordinate system up to order μ. In this way, we will find that the update residual depends neatly upon the transformed magnetic field.

It is worth noting that (2.51) has two new operations that have not been introduced so far. The first is that we are computing $\boldsymbol{x}(\boldsymbol{\xi}_{\leqslant \mu})$, i.e. we are inverting the coordinate transformation. The second is that we are composing functions with this inversion as $\tilde{\boldsymbol{h}}^{x}(\boldsymbol{x}(\boldsymbol{\xi}_{\leqslant \mu}))$. So, for any function $f(\boldsymbol{x})$ and coordinate transformation $\boldsymbol{\xi}(\boldsymbol{x})$, we can compute the equivalent function in indirect coordinates as $f(\boldsymbol{\xi})$ using these two steps. Moreover, the inverse transformation $\boldsymbol{x}(\boldsymbol{\xi})$ can be precomputed for all transformations one wishes to perform of this type. This gives a framework to move back and forth between direct and indirect near-axis formalisms to high order. For more details on how these transformations are computed, see Appendices (B.6) and (B.7).

Given the residual field (2.51), we can use the second equality in (2.49) to find the updated equation

$$\frac{\partial \tilde{\boldsymbol{\xi}}}{\partial s_{\leqslant \mu}} = -\frac{\partial \tilde{\boldsymbol{\xi}}}{\partial \tilde{\boldsymbol{\xi}}_{\leqslant \mu}} \tilde{\boldsymbol{h}}^{\xi_{\leqslant \mu}}\left(\boldsymbol{\xi}_{\leqslant \mu}\right) + \iota(\psi) J \tilde{\boldsymbol{\xi}}. \tag{2.52}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

Note that this has the exact same form as (2.35), except we have shifted the underlying coordinates. Substituting $\tilde{\boldsymbol{\xi}}_{\leqslant\mu+1} = \tilde{\boldsymbol{\xi}}_{\leqslant\mu} + \tilde{\boldsymbol{\xi}}_{\mu}$ into (2.52), we obtain

$$\frac{\partial \tilde{\boldsymbol{\xi}}_{\mu+1}}{\partial s_{\leqslant\mu}} + \iota_0 \frac{\partial \tilde{\boldsymbol{\xi}}_{\mu+1}}{\partial \tilde{\boldsymbol{\xi}}_{\leqslant\mu}} J\tilde{\boldsymbol{\xi}}_{\leqslant\mu} - \iota_0 J\tilde{\boldsymbol{\xi}}_{\mu+1} = -\tilde{\boldsymbol{h}}^{\xi_{\leqslant\mu}}_{\mu+1} + \iota_\mu \psi^{\mu/2} J\tilde{\boldsymbol{\xi}}_{\leqslant\mu} = \boldsymbol{F}_{\mu+1}. \tag{2.53}$$

Because the leading order problem (2.38) is an eigenvalue problem, (2.53) has the form of a higher-order correction to the eigenvalue and eigenfunction. To see what we might expect, consider the eigenvalue problem $K\boldsymbol{y} = \lambda M\boldsymbol{y}$, where each term is expanded in a small parameter, e.g. $K = K_0 + K_1\epsilon + K_2\epsilon^2 + \ldots$ for small ϵ. The analogous update equation would be

$$(K_0 - \lambda_0 B_0)\boldsymbol{y}_\mu = \lambda_\mu M_0 \boldsymbol{y}_0 + \boldsymbol{R}_\mu, \tag{2.54}$$

where $\boldsymbol{R}_\mu$ contains all of the residual terms. Assume for simplicity that λ_0 is an isolated eigenvalue. Then, there is a single secular term, which can be identified by taking the inner product of the above expression with the leading left eigenvector $\boldsymbol{z}_0$ to give $\lambda_\mu = -(\boldsymbol{z}_0^T \boldsymbol{R}_\mu)/(\boldsymbol{z}_0^T M_0 \boldsymbol{y}_0)$. Once this is satisfied, the equation can be solved for $\boldsymbol{y}_\mu$, where we typically choose the free component in $\boldsymbol{y}_0$ so that the norm is constant.

To perform the same steps on (2.53), we first diagonalise the left-hand-side operator by converting to polar coordinates $(\xi_{\leqslant\mu}, \eta_{\leqslant\mu}) = R_{\leqslant\mu}(\cos\Theta_{\leqslant\mu}, \sin\Theta_{\leqslant\mu})$ as

$$\frac{\partial \tilde{\boldsymbol{\xi}}_{\mu+1}}{\partial s_{\leqslant\mu}} + \iota_0 \frac{\partial \tilde{\boldsymbol{\xi}}_{\mu+1}}{\partial \Theta_{\leqslant\mu}} - \iota_0 J\tilde{\boldsymbol{\xi}}_{\mu+1} = \boldsymbol{F}_{\mu+1}, \tag{2.55}$$

where

$$\tilde{\boldsymbol{\xi}}_{\mu+1} = (R_{\leqslant\mu})^{\mu+1} \sum_{n=0}^{\mu} \sum_{\ell\in\mathbb{Z}} \tilde{\boldsymbol{\xi}}_{\mu+1,n\ell} e^{(2n-\mu)\mathrm{i}\theta + \mathrm{i}\ell s_{\leqslant\mu}}. \tag{2.56}$$

After substitution, we find that

$$((\iota_0(2n-\mu-1)+\ell)\mathrm{i}I - \iota_0 J)\,\tilde{\boldsymbol{\xi}}_{\mu+1,n\ell} = \boldsymbol{F}_{\mu+1,n\ell}, \tag{2.57}$$

where the 2×2 left-hand-side matrix has the eigenvalues $\lambda_{\mu+1,n\ell 0} = \iota_0\mathrm{i}(2n-\mu) + \mathrm{i}\ell$ and $\lambda_{\mu+1,n\ell 1} = \iota_0\mathrm{i}(2n-\mu-2) + \mathrm{i}\ell$ with corresponding right eigenvectors $(\boldsymbol{v}_0, \boldsymbol{v}_1) = ([-\mathrm{i}/\sqrt{2}, 1/\sqrt{2}], [\mathrm{i}/\sqrt{2}, 1/\sqrt{2}])$. So, the resulting updates in the coefficients are

$$\tilde{\boldsymbol{\xi}}_{\mu+1,n\ell} = \sum_{k\in\{0,1\}} \frac{1}{\mathrm{i}(\iota_0(2n-\mu-2k)+\ell)} \left\langle \overline{\boldsymbol{v}}_k,\ \boldsymbol{F}_{\mu+1,n\ell} \right\rangle \boldsymbol{v}_k, \tag{2.58}$$

where $\overline{\boldsymbol{v}}_k$ are the corresponding left eigenvectors due to $(\iota_0(2n-\mu-1)+\ell)\mathrm{i}I - \iota_0 J$ being skew-adjoint.

There are two cases where the update (2.58) fails. The first case is when $2n - \mu - 2k = 0$ and $\ell = 0$, occurring only when μ is even. This is the standard secularity that indicates that ι_μ must be updated, giving the condition that $\langle \overline{\boldsymbol{v}}_0,\ \boldsymbol{F}_{\mu+1,\mu/2,0}\rangle = \langle \overline{\boldsymbol{v}}_1,\ \boldsymbol{F}_{\mu+1,\mu/2+1,0}\rangle = 0$ for single-valued solutions, where we note that these formulae are equivalent for real magnetic fields. The resulting formula is

$$\iota_\mu = h^{\eta_{\leqslant\mu}}_{\mu+1,\mu/2,0} - \mathrm{i}h^{\xi_{\leqslant\mu}}_{\mu+1,\mu/2,0}. \tag{2.59}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

The second case of failure is when ι_0 is rational, as then there are other values of (μ, n, k, ℓ) such that (2.58) is singular. This is attributed to the expanding number of θ modes at each order, where higher-order poloidal perturbations resonate with the axis. To avoid this, extra resonant terms in the higher-order magnetic field must be introduced to avoid secularity (Duignan & Meiss 2021) (equivalently, this requires adding terms to the Hamiltonian normal form). Here, we assume that ι_0 is irrational so that the iteration is well defined.

We note that there is still one undetermined part of the problem: what value to choose for $\langle \overline{\boldsymbol{v}}_0, \tilde{\boldsymbol{\xi}}_{\mu+1,\mu/2,0} \rangle$ and its complex conjugate. Because this is arbitrary, we currently set this coefficient to 0. However, other options could be to choose this value to improve the radius of convergence or to match the flux (2.43). In summary, the straight field-line coordinate process is

input:	divergence-free magnetic field $\boldsymbol{B}$,
assuming:	elliptic on axis with irrational ι_0,
leading eigenvalue problem:	$\dfrac{\partial G_0}{\partial s} + G_0 H_0 = \iota_0 J G_0,$
	$\det G_0 = \pi (B_s)_0, \qquad \lvert\iota_0\rvert$ minimized,
where:	$\tilde{\boldsymbol{\xi}} = (\xi, \eta)^T = G_0(s)\tilde{\boldsymbol{x}} + \mathcal{O}(\psi),$
	$\tilde{\boldsymbol{h}}^{\boldsymbol{x}}(\boldsymbol{x}) = \dfrac{1}{B^s}\begin{pmatrix} B^x \\ B^y \end{pmatrix} = H_0(s)\tilde{\boldsymbol{x}} + \mathcal{O}(\psi),$
higher order linear solve:	$\dfrac{\partial \tilde{\boldsymbol{\xi}}_{\mu+1}}{\partial s_{\leqslant \mu}} + \iota_0 \dfrac{\partial \tilde{\boldsymbol{\xi}}_{\mu+1}}{\partial \tilde{\boldsymbol{\xi}}_{\leqslant \mu}} J\tilde{\boldsymbol{\xi}}_{\leqslant \mu} - \iota_0 J\tilde{\boldsymbol{\xi}}_{\mu+1} = \boldsymbol{F}_{\mu+1},$
	$\iota_\mu = h^{\eta_{\leqslant \mu}}_{\mu+1,\mu/2,0} - \mathrm{i} h^{\xi_{\leqslant \mu}}_{\mu+1,\mu/2,0}, \qquad (\mu \text{ even})$
where:	$\boldsymbol{F}_{\mu+1} = -\tilde{\boldsymbol{h}}^{\boldsymbol{\xi}_{\leqslant \mu}}_{\mu+1} + \iota_\mu \psi^{\mu/2} J\tilde{\boldsymbol{\xi}}_{\leqslant \mu},$
	$\tilde{\boldsymbol{h}}^{\boldsymbol{\xi}_{\leqslant \mu}}(\boldsymbol{\xi}_{\leqslant \mu}) = \dfrac{\partial \tilde{\boldsymbol{\xi}}_{\leqslant \mu}}{\partial \tilde{\boldsymbol{x}}} \tilde{\boldsymbol{h}}^{\boldsymbol{x}}(\boldsymbol{x}(\boldsymbol{\xi}_{\leqslant \mu})),$
output:	rotational transform ι,
	straight field-line coordinates $\boldsymbol{\xi}$.

(2.60)

3. Ill-posedness and regularisation

In this section, we describe how the near-axis problem is ill-posed (§ 3.1) and how we can regularise the problem (§ 3.2). In § 3.3, we state how the near-axis expansion of ϕ converges under suitable input assumptions (Theorem (3.8) and Corollary (3.9)). Most proofs can be found in Appendix A, where the individual sections are referred to after each statement.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

3.1. *Ill-posedness*

We define a problem as ill-posed if it is not well-posed, where the standard definition of a well-posed problem is that:

(i) the solution exists;

(ii) the solution is unique;

(iii) the solution is continuous in the initial data.

We note that the interpretations of these statements depend on what space we require the solution to belong to and over which topology continuity is described in. For instance, it is straightforward to show existence and uniqueness in the sense of a formal power series.

PROPOSITION 3.1. *Consider the near-axis problem in box (2.30), with all inputs in* C^∞. *Then, there exists a unique formal power series solution* $\phi_n(s)$ *at each order.*

Proof. Simply notice that the residual at each order is C^∞ if every previous order is. Then, because the inverse of $\Delta_\perp$ of C^∞ functions is C^∞, we satisfy the Fredholm alternative, and we specify the null space the operator at each order, we have a unique solution. □

Note that this proposition says nothing about the convergence of the power series to a solution off-axis; it only shows that we can find the coefficients of the power series. So, formal existence does not necessarily imply good or consistent computational results.

Another straightforward existence result for harmonic inputs is the following.

PROPOSITION 3.2. *Let* $\boldsymbol{B} = \nabla\phi + \tilde{\boldsymbol{B}}$ *be a valid vacuum magnetic field on* Ω_σ *with a real-analytic axis* $\boldsymbol{r}_0$ *and* $\sigma > 0$. *Then, the near-axis expansion using the coefficients corresponding to* ϕ *converges uniformly on a smaller domain* $\Omega_{\sigma'}$ *with* $0 < \sigma' < \sigma$.

Proof. See § A.1 □

Proposition (3.2) is a useful result because it says that vacuum fields can, in principle, be written as solutions to the infinite near-axis problem. However, this is a difficult theorem to use in practice, as it is difficult to verify *a priori* whether the input data to the near-axis expansion agree with a solution of Poisson's equation.

So, for a more computational approach, we must define normed spaces of inputs and outputs that agree with notions of convergence. To intuit what the correct space may be, we observe that the radial direction behaves as a 'time-like' variable, whereas the θ and s behave more like spatial variables. That is, the near-axis PDE can be thought of as propagating surface information off the axis. This motivates a decision to separate our treatment of these coordinates. Moreover, we desire convergence in a power series in the radial variable, so we choose to treat it in an analytic manner. In contrast, the coefficients $\phi_m(\theta, s)$ are obtained by solving a linear PDE at each order, so we treat them in a Sobolev sense as a function of θ and s.

To this end, let $\alpha = (\alpha_1, \alpha_2)$ be a multi-index of degree 2 and $|\alpha| = \sum_j \alpha_j$. We define the H^q Sobolev norm of functions $f : \mathbb{T}^2 \to \mathbb{R}$ as

$$\| f \|_{H^q}^2 = \sum_{|\alpha| \leqslant q} \| D_\alpha f \|_{L^2}^2 , \tag{3.1}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

where

$$D_\alpha f = \frac{\partial^{|\alpha|} f}{\partial \theta^{\alpha_1} \partial s^{\alpha_2}}, \qquad \|f\|_{L^2}^2 = \int_{\mathbb{T}^2} |f|^2 \, \mathrm{d}\mu, \tag{3.2}$$

and μ is the Lebesgue measure on $\mathbb{T}^2$. We additionally define the C^q norm of a q-times differentiable function $f : \mathbb{T}^2 \to \mathbb{R}$ as

$$\|f\|_{C^q} = \sum_{|\alpha| \leqslant q} \sup_{(\theta,s)\in\mathbb{T}^2} |D_\alpha f(\theta, s)| \,. \tag{3.3}$$

Then, we define the following convenient near-axis function space.

DEFINITION 3.3. *Let $\sigma > 0$ and W be a Banach space on functions on $\mathbb{T}^2$. We define a function $f : \Omega_\sigma^0 \to \mathbb{R}^d$ as (σ, W)-analytic if f has a convergent near-axis expansion of the form*

$$f(\rho, \theta, s) = \sum_{m=0}^{\infty} f_m(\theta, s)\rho^m, \qquad f_m(\theta, s) = \sum_{n=0}^{m} f_{mn}(s) e^{(2n-m)\mathrm{i}\theta}, \tag{3.4}$$

where the norm

$$\|f\|_{\sigma,W} = \sup_n \left(\sigma^n \|f_n\|_W\right) \tag{3.5}$$

is bounded. Here, convergence of the near-axis expansion is pointwise in ρ and in norm in (θ, s), i.e. for all $\rho < \sigma$,

$$\lim_{M\to\infty} \left\| f(\rho, \cdot, \cdot) - \sum_{m=0}^{M} f_m(\cdot, \cdot)\rho^m \right\|_W \to 0. \tag{3.6}$$

Paralleling standard linear regularity theory (Evans 2010), we will consider near-axis solutions ϕ to belong to a Sobolev (σ, H^q)-analytic space, while near-axis PDE coefficients belong to differentiable (σ, C^q)-analytic spaces. Because the coefficients of Poisson's equation (2.21) are functions of the metric, control on the (σ, C^q)-analytic norm of the coefficients in the Frenet–Serret coordinate system will be entirely determined by the differentiability class of the axis $\boldsymbol{r}_0 \in C^q(\mathbb{T})$.

To build some intuition for (σ, W)-analytic functions, we turn to some straightforward facts about their convergence. Let f be (σ, W)-analytic. Then, for any $F \geqslant \|f\|_{\sigma,W}$, the definition of the norm $\|\cdot\|_{\sigma,W}$ tells us that the coefficients are bounded as

$$\|f_m\|_W \leqslant F\sigma^{-m}. \tag{3.7}$$

This means that surfaces of f converge geometrically for $\rho^* < \sigma$ in W, i.e.

$$\left\| f|_{\rho=\rho^*} \right\|_W \leqslant F \sum_{m=0}^{\infty} \left(\frac{\rho^*}{\sigma}\right)^m = \frac{F}{1 - \rho^*/\sigma}. \tag{3.8}$$

For a more practical statement of pointwise convergence, we have the following proposition.

PROPOSITION 3.4. *Let f be (σ, C^q)-analytic for $q \geqslant 0$. Then, f is continuous and q-times continuously differentiable in Ω_σ^0.*

Proof. See § A.2 □

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

COROLLARY 3.5. *Let f be (σ, H^q)-analytic for $q \geqslant 2$. Then, f is (σ, C^{q-2})-analytic, continuous and $(q-2)$-times continuously differentiable in Ω^0_σ.*

Proof. Because $H^q(\mathbb{T}^2)$ is continuously embedded in $C^{q-2}(\mathbb{T}^2)$, f is also (σ, C^{q-2})-analytic. Apply Proposition (3.4). □

A direct consequence of the preceding statements is that both (σ, C^0)-analytic and (σ, H^2)-analytic functions are continuous in 3-D, but (σ, H^1)-analytic functions need not be.

Now, let us return to the question of ill-posedness. To define the norm of the input, let

$$\phi^{(\mathrm{IC})} = \sum_{m=0}^{\infty} \rho^m \left(\phi_{m0} e^{-\mathrm{i}m\theta} + \phi_{mm} e^{\mathrm{i}m\theta}\right). \tag{3.9}$$

This allows us to naturally define the norm of the input functions ϕ_{m0} and ϕ_{mm} via a single (σ, H^q)-analytic norm. Using this, we prove that the problem is ill-posed in the following sense.

THEOREM 3.6. *Let $\boldsymbol{r}_0, B_0, \phi_{m0}, \phi_{mm} \in C^\infty$ for $2 \leqslant m < \infty$ with $\ell', \kappa > 0$ and $q_0 \geqslant 0$. The near-axis solution ϕ of (2.30) is not continuous to perturbations $\delta\boldsymbol{r}_0, \delta B_0, \delta\phi_{m0}, \delta\phi_{mm} \in C^\infty$ under the $C^{4+q_0}(\mathbb{T})$ norm on $\boldsymbol{r}_0$, the $H^{1+q_0}(\mathbb{T})$ norm on B_0, the (σ_0, H^{2+q_0})-analytic norm on $\phi^{(\mathrm{IC})}$ with $\sigma_0 > 0$, and any (σ, H^q)-analytic norm on the output ϕ with $\sigma > 0$ and $q \geqslant 2$, i.e. the near-axis expansion is ill-posed.*

Proof. See § A.3 □

In other words, Theorem (3.6) tells us that smooth bounded perturbations in the input lead to unbounded deviations in the solution, even if the problem is initially prepared to be convergent. The characteristic form of the unbounded perturbations are high frequency in s, which grow exponentially off-axis according to their wavenumber. When the near-axis expansion is discretised, this appears to be a poor condition number for the truncated problem. This motivates us to introduce a term in the near-axis expansion that damps the behaviour of the high-frequency modes.

Before we continue, we note there is a strong connection between the ill-posedness of the near-axis expansion and the ill-posedness of coil design. It is typically the case that magnetic fields with large gradients are difficult to approximate using plasma coils far from the boundary (Kappel *et al.* 2024). This is because high frequencies in coil design decay quickly towards the surface of the plasma – the opposite view of the problem of high frequencies growing outward from the axis. The effect is that it is difficult to match high frequencies on the plasma boundary, and coil design codes also require some form of regularisation (Landreman 2017).

3.2. *Regularisation*

We have just seen in Theorem (3.6) that the near-axis expansion described in box (2.30) is ill-posed. Ill-posedness can potentially cause significant problems for numerical simulations, with the primary one being that discretisation refinement – both in toroidal resolution and in the order of expansion – will not lead to convergence to a solution. Instead, the deviation from the correct answer typically increases with refinement.

In the proof of Theorem (3.6) (§ A.3), the problematic perturbations to the input have small amplitude and high poloidal wavenumber, leading to exponential growth off the axis with a rate proportional to the wavenumber. Here, we would like to

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

damp the problematic high-wavenumber modes while maintaining the fidelity of the low-wavenumber modes. We do so by adding a high-order differential operator in the toroidal and poloidal directions, as high-order derivatives affect high wavenumbers significantly more than low wavenumbers. Specifically, we propose the following version of Poisson's equation (2.21) with regularisation:

$$\Delta\phi + \frac{\rho}{\sqrt{g}}\Delta_\perp P\phi = -\nabla\cdot\tilde{\boldsymbol{B}}. \tag{3.10}$$

The new term contains the regularising differential operator $\Delta_\perp P$ that is at least fourth order. The operator is composed of a product of the perpendicular Laplacian defined by

$$\Delta_\perp\phi = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\phi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\phi}{\partial\theta^2}, \tag{3.11}$$

and the regularising differential operator P satisfying the following hypotheses.

HYPOTHESES 3.7. *Let $q \geqslant 0$ and $D \geqslant 1$. We require P be an order $2D$ differential operator that satisfies the following:*

(i) *P takes the form*

$$P = \sum_{m=0}^{2D}\sum_{n=0}^{m} P^{(mn)}(s)\frac{\partial^{m+n}}{\partial s^m\partial\theta^n}, \tag{3.12}$$

where $P^{(mn)}(s) \in C^q$;

(ii) *P is strongly elliptic, i.e. for some $C > 0$ and for all $(x, y) \in \mathbb{R}^2\backslash\{0\}$,*

$$\sum_{n=0}^{2D} P^{(2D,n)}(s)x^{2D-n}y^n \geqslant C(x^2+y^2)^D, \tag{3.13}$$

(iii) *P is self-adjoint and semi positive-definite, i.e. for non-zero $f, g \in H^{2D}(\mathbb{T}^2)$,*

$$\langle f,\ Pg\rangle = \langle Pf,\ g\rangle, \qquad \langle f,\ Pf\rangle \geqslant 0, \qquad \langle f,\ g\rangle := \int_{\mathbb{T}^2} f(\theta,s)g(\theta,s)\,\mathrm{d}\theta\mathrm{d}s \tag{3.14}$$

We note that in hypothesis (i), the coefficients of P do not depend on θ. This means poloidal Fourier modes 'block diagonalise' P, i.e. for $f(s) \in H^{q+2D}$, we have $P(f(s)e^{im\theta}) = g(s)e^{im\theta}$ for some $g(s) \in H^q$. This is sufficient for P to map (σ, H^{q+2D})-analytic functions to (σ, H^q)-analytic functions, while θ dependence would make the operator non-analytic. It also means that P commutes with $\Delta_\perp$, i.e. $P\Delta_\perp = \Delta_\perp P$ on sufficiently regular functions. Moreover, the hypotheses (ii) and (iii) are sufficient for $1 + P$ to be invertible, which is necessary at each step of the iteration.

In practice, we use the operator

$$P = \left(-\frac{1}{K^2}\left(\frac{L}{2\pi\ell'}\frac{\partial}{\partial s}\right)^2\right)^D + \left(-\frac{1}{K^2}\frac{\partial^2}{\partial\theta^2}\right)^D, \tag{3.15}$$

where $K > 0$ is the characteristic wavenumber of regularisation, $2D \geqslant 2$ is the algebraic order of the frequency damping and $L = \oint \ell'\,\mathrm{d}s$ is the length of the magnetic

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

axis. We have chosen the specific form of P so that it is easy to implement numerically when the curve is in arclength coordinates (and therefore ℓ' is constant, satisfying the smoothness requirement). Indeed, any polynomial in the arclength derivative $(\ell')^{-1}(\partial/\partial s)$ and the poloidal derivative $(\partial/\partial\theta)$ is simply inverted in Fourier space. Additionally, the parameter K can be tuned from large (weak damping) to small (strong damping) to adjust the regularisation strength.

With regularisation, the new iterative form of the near-axis expansion (3.10) is

$$\left(m^2 + \frac{\partial^2}{\partial\theta^2}\right)(1+P)\phi_m = -(\nabla\cdot\tilde{\boldsymbol{B}} + \Delta\phi_{<m})_{m-2}, \tag{3.16}$$

where we discuss the relevant regularity in Corollary (3.9). If we use the regularisation (3.15) and s to be a scaled arclength coordinate so that $\ell' = L/2\pi$ is constant, the componentwise version of the iteration is

$$(m^2-(2n-m)^2)\left(1+\left(\frac{k}{K}\right)^{2D}+\left(\frac{2n-m}{K}\right)^{2D}\right)\phi_{mnk} = -(\nabla\cdot\tilde{\boldsymbol{B}} + \Delta\phi_{<m})_{m-2,n-1,k} \tag{3.17}$$

for $0 < n < m$. We see that in the near-axis iteration, the regularisation damps the high-order modes by dividing by high-order polynomials in the poloidal and toroidal wavenumbers.

By construction, the regularisation has another benefit: the full problem can be represented as the divergence of a perturbed magnetic field $\boldsymbol{B}_K$. If we let $G = Diag(1, \rho^{-2}, 0)$, we have

$$B_K^i = g^{ij}\frac{\partial\phi}{\partial q^j} + \tilde{B}^i + \frac{\rho}{\sqrt{g}}G^{ij}\frac{\partial(P\phi)}{\partial q^j}. \tag{3.18}$$

The fact that there is still an underlying divergence-free field is important for the relation between the near-axis expansion and Hamiltonian mechanics. Therefore, the regularised near-axis expansion of the field $\boldsymbol{B}_K$ can be expressed by the problem

$$\nabla\times\boldsymbol{B}_K = \boldsymbol{J}_K, \qquad \nabla\cdot\boldsymbol{B}_K = 0, \tag{3.19}$$

where $\boldsymbol{J}_K$ is a fictitious regularising current. In contrast, the regularisation means $\boldsymbol{B} = \nabla\phi + \tilde{\boldsymbol{B}}$ is not divergence-free. To find the flux surfaces, we note that the procedure in box (2.60) in no way depends on the magnetic field being curl-free. As such, we perform the same steps for finding flux coordinates for $\boldsymbol{B}_K$ as with $\boldsymbol{B}$.

3.3. *Convergence of the regularised expansion*

Consider a PDE of the form

$$\Delta_\perp(1+P)\phi = f + L^{(a)}\phi, \tag{3.20}$$

where P satisfies Hypotheses (3.7), f is (σ_0, H^q)-analytic for $q \geqslant 0$ and $\sigma_0 > 0$, and

$$L^{(a)} = a^{(1)}\phi + \frac{\partial a^{(2)}}{\partial\rho}\frac{\partial\phi}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial a^{(2)}}{\partial\theta}\frac{\partial\phi}{\partial\theta} + a^{(3)}\frac{\partial\phi}{\partial\theta} + a^{(4)}\frac{\partial\phi}{\partial s} + a^{(5)}\frac{\partial^2\phi}{\partial\theta^2} + a^{(6)}\frac{\partial^2\phi}{\partial\theta\partial s} + a^{(7)}\frac{\partial^2\phi}{\partial s^2}, \tag{3.21}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

and each $a^{(j)}$ is a (Σ, C^q)-analytic for $\Sigma > \sigma_0$. By a straightforward application of chain rule, we see that (3.10) satisfies this form for $\boldsymbol{r}_0 \in H^{4+q}$ and $B_0 \in H^{1+q}$, where $\Sigma \leqslant \min \kappa^{-1}$ (see proof of Corollary 3.9 in Appendix A.5). Equation (3.20) also encompasses other coordinate and regularisation assumptions of the near-axis expansion, including those not defined by Frenet–Serret coordinates.

The implicit near-axis iteration for (3.20) can be written as

$$(\Delta_\perp(1+P)\phi)_m = \left[f + L^{(a)}\right]_m. \tag{3.22}$$

The form of the iteration automatically respects the Fredholm condition, so it is solvable at each order (see the proof of Theorem (3.8)). So, to solve at each order of the iteration, we can invert to find that

$$\phi_m = \phi_m^{(\mathrm{IC})} + \left((1+P)^{-1}\Delta_\perp^+\left[f + L^{(a)}\phi\right]\right)_m, \tag{3.23}$$

where $\phi^{(\mathrm{IC})}$ is (σ_0, H^{q+2})-analytic of the form (3.9) and we define the pseudoinverse $\Delta_\perp^+$ as

$$\Delta_\perp^+(\rho^m e^{\mathrm{i}(2n-m)\theta}) = \frac{1}{(m+2)^2 - (2n-m)^2}\rho^{m+2}e^{\mathrm{i}(2n-m)\theta} \qquad \text{for } 0 \leqslant n \leqslant m. \tag{3.24}$$

Given this iteration, we find the following theorem for its convergence.

THEOREM 3.8. *Let $0 < \sigma_0 < \Sigma$, $q \geqslant 1$ and $D \geqslant 1$. Then, let f be (σ_0, H^q)-analytic, $a^{(j)}$ for $1 \leqslant j \leqslant 7$ be (Σ, C^q)-analytic, $\phi^{(\mathrm{IC})}$ as defined by (3.9) be (σ_0, H^{q+2D})-analytic and P satisfy Hypotheses (3.7). Then, there is a unique (σ, H^{q+2D})-analytic near-axis solution ϕ of (3.20) with initial data $\phi^{(\mathrm{IC})}$. Moreover, this solution operator is continuous in $a^{(j)}$ and satisfies the bound*

$$\|\phi\|_{\sigma, H^{q+2D}} \leqslant C\,\|f\|_{\sigma_0, H^q} + C\,\left\|\phi^{(\mathrm{IC})}\right\|_{\sigma_0, H^{q+2D}}. \tag{3.25}$$

Proof. See Appendix A.4. □

This theorem can be translated to the Frenet–Serret problem.

COROLLARY 3.9. *Let $\sigma > 0$, $q \geqslant 1$ and $D \geqslant 1$. Then, let $\boldsymbol{r}_0 \in C^{4+q}$ in scaled arclength coordinates with $\ell' = L/2\pi > 0$ and $\kappa > 0$, $B_0 \in H^{1+q}$, $\phi^{(\mathrm{IC})}$ be (σ_0, H^{q+2D})-analytic, and P be defined as in (3.15). Then, for some $\sigma > 0$ satisfying $\sigma \leqslant \sigma_0$ and $\sigma < \min \kappa^{-1}$, the near-axis solution ϕ of (3.10) is (σ, H^{q+2D})-analytic, continuous in $\boldsymbol{r}_0$ and satisfies*

$$\|\phi\|_{\sigma, H^{q+2D}} \leq C\,\|B_0\|_{H^{1+q}} + C\,\left\|\phi^{(\mathrm{IC})}\right\|_{\sigma_0, H^{q+2D}}. \tag{3.26}$$

Furthermore, the associated divergence-free field $\boldsymbol{B}_K$ is well-defined and is (σ', H^q)-analytic for all $\sigma' < \sigma$.

Proof. See Appendix A.5. □

There are two primary reasons why Theorem (3.8) is useful for computation. First, it gives a guarantee that the norm of the inputs controls the size of the output. For instance, in the context of stellarator optimisation, if an appropriate Sobolev penalty is put on $\boldsymbol{r}_0$, B_0 and $\phi^{(\mathrm{IC})}$, then one can expect the output to be appropriately bounded. Second, because the output is (σ, H^{q+2D})-analytic, it tells us that truncations of the near-axis expansion are approximations of the true solution. So, we can be more confident that finite asymptotic series are approximately correct, at least for the regularised problem.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

It is worth noting that while Theorem (3.8) tells us there is a solution to the regularised problem, it does not tell us that the solution solves the original problem. To address this, we can develop an *a posteriori* handle on the error.

PROPOSITION 3.10. *Consider the hypotheses of Theorem (3.8) with* $q \geqslant 2$. *Additionally, suppose* $0 < \sigma' < \sigma$, $L = \Delta_\perp - L^{(a)}$ *is a second-order negative-definite operator on* $H_0^2(\Omega^0_{\sigma'})$, *and* $\phi^{\sigma'} = \phi(\sigma', \theta, s)$. *Then, the solution* ϕ *of the near-axis expansion is the unique solution in* $H^{q+2D}(\Omega^0_{\sigma'})$ *of the boundary value problem*

$$P\Delta_\perp \phi + L\phi = f, \qquad \phi|_{\rho=\sigma'} = \phi^{\sigma'}. \tag{3.27}$$

Moreover, let $\tilde{\phi}$ *be the solution to*

$$L\tilde{\phi} = f, \qquad \tilde{\phi}\Big|_{\rho=\sigma'} = \phi^{\sigma'}. \tag{3.28}$$

Then,

$$\left\| \phi - \tilde{\phi} \right\|_{H^q\left(\Omega^0_{\sigma'}\right)} \leqslant C \left\| P\Delta_\perp \phi \right\|_{H^{q-2}\left(\Omega^0_{\sigma'}\right)}. \tag{3.29}$$

Proof. See Appendix A.6. □

In other words, this tells us that given a solution to the regularised problem, we can bound the distance to a non-regularised boundary value problem via the norm of $P\Delta_\perp\phi$. This applies directly to the Frenet–Serret case because the Laplacian is negative-definite. So, given a solution, this gives us an estimate of the error. As before, we can summarise the new problem (cf. box (2.30)):

input:	axis $\boldsymbol{r}_0 \in C^{4+q}$, on-axis field $B_0 \in H^{1+q}$,
	(σ_0, H^{q+2D})-analytic higher moments $\phi^{(\mathsf{IC})}$,
	P satisfying Hypotheses 3.7,
assuming:	$\ell' > 0,\ \kappa > 0,\ B_0 > 0,\ q, D \leqslant 1$
solve:	$\ell'\left(m^2 + \dfrac{\partial^2}{\partial\theta^2}\right)(1+P)\phi_m = -(\nabla\cdot\tilde{\boldsymbol{B}})_m - (\Delta\phi_{<m})_m,$
output:	(σ, H^q)-analytic potential ϕ,
	magnetic field $B^i_K = g^{ij}\dfrac{\partial\phi}{\partial q^j} + \tilde{B}^i + \dfrac{\rho}{\sqrt{g}}G^{ij}\dfrac{\partial(P\phi)}{\partial q^j}.$

(3.30)

4. Numerical method

To numerically solve the regularised near-axis expansion algorithm in (3.30), we use a pseudospectral method. Pseudospectral methods use spectral representations of the solution for derivatives, while scalar multiplication and other operations occur on a set of collocation points. The spectral form of the series is

$$X_{\leqslant N_\rho}(\rho,\theta,s) = \sum_{m=0}^{N_\rho} \rho^m X_m(\theta,s), \qquad X_m(\theta,s) = \sum_{n=0}^{m} X_{mn}(s)\mathcal{F}_{2n-m}(\theta),$$

$$X_{mn}(s) = \sum_{k=-N_s}^{N_s} X_{mnk}\mathcal{F}_k(s), \tag{4.1}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

where N_ρ and N_s are integers specifying the resolution of the series and

$$\mathcal{F}_k(s) = \begin{cases} \cos(ks), & k \geqslant 0, \\ \sin(-ks), & k < 0. \end{cases} \tag{4.2}$$

Derivatives of the series are numerically evaluated by

$$\begin{aligned} \frac{\partial X_{\leqslant N_\rho}}{\partial \rho} &= \sum_{m=0}^{N_\rho - 1} (m+1) X_m(\theta, s) \rho^m, \\ \frac{\partial X_{\leqslant N_\rho}}{\partial \theta} &= \sum_{m=1}^{N_\rho} \sum_{n=0}^{m} -(2n-m) X_{mn} \rho^m \mathcal{F}_{-(2n-m)}(\theta), \\ \frac{\partial X_{\leqslant N_\rho}}{\partial s} &= \sum_{m=0}^{N_\rho} \sum_{n=0}^{m} \sum_{k=-N_s}^{N_s} -k X_{mnk} \rho^m \mathcal{F}_{2n-m}(\theta) \mathcal{F}_{-k}(s). \end{aligned} \tag{4.3}$$

For algebraic operations such as series multiplication, composition and inversion, we discretise each X_m on a grid $X_m(\theta_{mj}, s_\ell)$, where

$$\theta_{mj} = \frac{\pi j}{m+1}, \qquad s_\ell = \frac{2\pi \ell}{M_s}, \tag{4.4}$$

where $0 \leqslant j \leqslant m$, $0 \leqslant \ell < M_s$ and $M_s \geqslant 2N_s + 1$ is the number of s-collocation points. Typically, we choose $M_s = 4N_s + 3$ to oversample in s by a factor of over 2. This choice anti-aliases the numerical method by removing high harmonics generated in the collocation space (Boyd 2001). We note that the θ-collocation points θ_{mj} are spaced around the half-circle instead of the full circle, owing to the fact that analytic coefficients satisfy the symmetry

$$X_m(\theta + \pi, s) = (-1)^m X_m(\theta, s). \tag{4.5}$$

For even m, (4.5) tells us that X_m is periodic in 2θ and the resulting transformation is the discrete Fourier transform (DFT) in that angular coordinate. However, X_m is anti-periodic in θ for odd m and the collocation on the half-circle can interpreted as a symmetry reduction of the DFT on the full circle.

To transform between Fourier and spatial representations at each order m, let $\mathsf{X}_m^{\mathrm{c}} = [X_m(\theta_{mj}, s_\ell)] \in \mathbb{R}^{(m+1)\times M_s}$ be the matrix of collocation values and $\mathsf{X}_m^s = [X_{mnk}] \in \mathbb{R}^{(m+1)\times(2N_s+1)}$ be the matrix of Fourier coefficients. Then, we define the transition matrices $[(\mathsf{F}_\theta^{m,m'})_{jn}] = [\mathcal{F}_{2n-m}(\theta_{m'j})] \in \mathbb{R}^{(m+1)\times(m'+1)}$ and $[(\mathsf{F}_s)_{k\ell}] = [\mathcal{F}_k(s_\ell)] \in \mathbb{R}^{(2N_s+1)\times M_s}$. The transformation from spectral coefficients to collocation nodes is expressed by

$$\mathsf{X}_m^c = (\mathsf{F}_\theta^{m,m})^T \mathsf{X}_m^s \mathsf{F}_s. \tag{4.6}$$

Similarly, the inverse transform happens via

$$\mathsf{X}_m^s = (\mathsf{F}_\theta^{m,m})^{-T} \mathsf{X}_m^c \mathsf{F}_s^+, \tag{4.7}$$

where the pseudo-inverse is $\mathsf{F}_s^+ = \mathsf{F}_s^T D_s^{-1}$ and $D_s = \mathsf{F}_s \mathsf{F}_s^T \in \mathbb{R}^{(2N_s+1)\times(2N_s+1)}$ is diagonal with $(D_s)_{jj} = (2N_s + 1)$ for $j = 0$ and $(D_s)_{jj} = (2N_s + 1)/2$ for $-N_s \leqslant j \leqslant N_s$, $j \neq 0$. This transformation is currently performed via full matrix–matrix multiplication, but it could be accelerated for large systems by the fast Fourier transform.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

The final basic operation we use is to raise the order of the θ-collocation. To see why this is necessary, consider the simple case of three monomial power series: $X = \rho^m X_m$, $Y = \rho^{m'} Y_{m'}$ and $Z = XY = \rho^{m+m'} Z_{m+m'} = \rho^{m+m'} X_m Y_{m'}$. Then, the multiplication on the collocation nodes as

$$Z(\theta_{m+m',j}, s_\ell) = \rho^{m+m'} X(\theta_{m+m',j}, s_\ell) Y(\theta_{m+m',j}, s_j). \tag{4.8}$$

So, to obtain the correct collocation on $Z_{m+m'}$, we need to change the θ-collocation on X_m from θ_{mj} to $\theta_{m+m',j}$ and similarly for $Y_{m'}$. To do this, we use the θ-collocation matrices to find

$$\mathsf{Z}^c_{m+m'} = \left[(F_\theta^{m,m+m'})^T (F_\theta^{m,m})^{-T} \mathsf{X}^c_m \right] \odot \left[(F_\theta^{m',m+m'})^T (F_\theta^{m',m'})^{-T} \mathsf{Y}^c_m \right], \tag{4.9}$$

where $\odot$ is the Hadamard (element-wise) product, and $\mathsf{Y}^c_{m'}$ and $\mathsf{Z}^c_{m+m'}$ are the collocation matrices of $Y_{m'}$ and $Z_{m+m'}$. With this operation, the operations outlined in Appendix B can be performed on the collocated nodes.

We note that choosing the correct amount of modes in s presents the most difficult numerical problem in computing the near-axis expansion. As the order increases, high-order residuals are increasingly nonlinear in the lower orders, causing a broadening of the spectrum in s. If the inputs to the expansion do not have a sufficiently narrow bandwidth, this will result in broad higher-order residuals, particularly when finding flux coordinates. Both the regularisation and the anti-aliasing effects of choosing F_s to be rectangular help alleviate the issue of broad bandwidth, but in practice, we have found that it remains important to choose smooth inputs, especially for finding flux surfaces.

5. Examples

We now investigate the numerical convergence of the near-axis expansion to high orders. Our focus is on characterising the convergence of the input (figure 4), the convergence of the output magnetic field (figure 5), the convergence of the magnetic surfaces (figures 7, 8) and the role of regularisation (figures 5, 6). Through our two examples – the rotating ellipse and the precise QA equilibrium of Landreman–Paul (Landreman & Paul 2022) – we find that the radius of convergence of every series is closely related to the distance from the magnetic axis to the coils σ_{coil}. This radius appears to limit the convergence of every other series of interest, including the magnetic surfaces that extend beyond this distance.

All computations in this section were performed on a personal laptop. The code used to perform the expansions can be found at the `StellaratorNearAxis.jl` package (Ruth 2024*a*).

5.1. *Equilibrium initialisation*

A major task in computing high-order near-axis expansions is choosing the input coefficients $\phi^{(\text{IC})}$ in box (3.30). While low-order expansions can often be expressed in physically intuitive variables parametrising the rotation and stretching of elliptical magnetic surfaces, it is not as intuitive how to determine the high-order coefficients of $\phi^{(\text{IC})}$. In practice, the best option for finding equilibria would likely be via optimisation. However, for the purposes of demonstration, we initialise our inputs by a more direct method: via magnetic coils (see figure 3).

The primary advantage of using coils for equilibrium initialisation is accuracy. The accuracy comes from the fact that the coil field can be expanded analytically

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

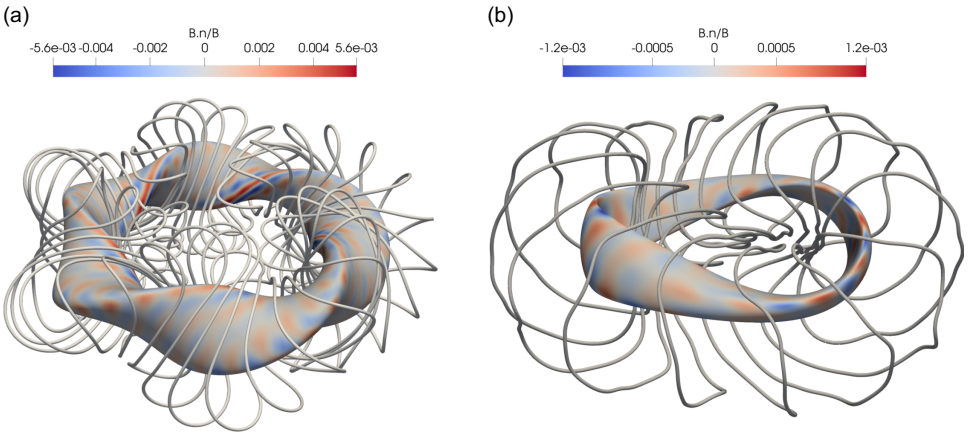


FIGURE 3. Coil sets for the rotating ellipse and Landreman–Paul examples. The colour indicates the normalised $\boldsymbol{B} \cdot \boldsymbol{N}$ error on the outer closed flux surface.

about the axis, giving a direct input to the near-axis expansion. This circumvents the potentially error-prone problem of interpolating stellarator equilibria. Coils also provide an accurate ground truth to which to compare our equilibrium. In the following subsections, we use this to assess the accuracy of the near-axis expansion, both close to the axis and farther away.

The coil optimisation method employed here follows the approach described by Wechsung *et al.* (2022) and Jorge *et al.* (2024) using the code SIMSOPT (Landreman *et al.* 2021). Coils are modelled as single closed 3-D filaments of current $\boldsymbol{\Gamma}^{(i)} : \mathbb{T} \to \mathbb{R}^3$. Each coil i is modelled as a periodic function in Cartesian coordinates where

$$\boldsymbol{\Gamma}^{(j)}(s) = \boldsymbol{c}_0^{(j)} + \sum_{\ell=1}^{N_F} \left[\boldsymbol{c}_\ell^{(j)} \cos(\ell\theta) + \boldsymbol{s}_\ell^{(j)} \sin(\ell\theta) \right], \tag{5.1}$$

where each $\boldsymbol{c}_\ell^{(j)}, \boldsymbol{s}_\ell^{(j)} \in \mathbb{R}^3$, yielding a total of $3 \times (2N_F + 1)$ degrees of freedom per coil. In this work, we used $N_F = 12$, with four coils per half-field period for Landreman–Paul case and eight coils per half-field period for the ellipse. The degrees of freedom for the coil shapes are then

$$\boldsymbol{x}_{\text{coils}} = [\boldsymbol{c}_l^{(j)}, \boldsymbol{s}_l^{(j)}, I_j], \tag{5.2}$$

with I_j the current that goes through each coil. We take advantage of stellarator and rotational symmetries to only optimise a set of N_c coils per half-field period. This leads to a total of $2 \times n_{\text{fp}} \times N_c$ modular coils, where n_{fp} is the number of toroidal field periods with $n_{\text{fp}} = 2$ for Landreman–Paul and $n_{\text{fp}} = 5$ for the rotating ellipse. The remaining coils are determined by symmetry. The magnetic field $\mathbf{B}_{\text{ext}}$ of each coil is evaluated using the Biot–Savart law

$$\boldsymbol{B}_{\text{coil}}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \sum_{j=1}^{2n_{\text{fp}}N_c} \int_0^{2\pi} \frac{I_j \frac{\partial \boldsymbol{\Gamma}_j}{\partial s'} \times (\boldsymbol{\Gamma}_j(s') - \boldsymbol{r})}{\left|\boldsymbol{\Gamma}_j(s') - \boldsymbol{r}\right|^3} \mathrm{d}s', \tag{5.3}$$

where s' parametrises the coil curve. Each coil is divided into 150 quadrature points, and the cost functions used to regularise the optimisation problem use the minimum

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

distance between two coils, the length of each coil, their curvature and mean-squared curvature (see Wechsung *et al.* 2022).

Using the coil magnetic field (5.3), we find the magnetic axis $\boldsymbol{r}_0$ via a shooting method. Then, using the near-axis coordinate representation of $\boldsymbol{r}$ in (2.2), we expand the quadrature rule of (5.3) using the operations in Appendix B to find a near-axis expansion for the magnetic field $\boldsymbol{B}(\rho, \theta, s)$. Given the near-axis field, it is straightforward to compute B_0 by

$$B_0(s) = \boldsymbol{t}(s) \cdot \boldsymbol{B}(0, 0, s), \tag{5.4}$$

and ϕ is found by a near-axis expansion of the path integral (note that $\phi = 0$ on the axis)

$$\phi(\rho, \theta, s) = \int_0^{\rho} (\boldsymbol{B}(\rho', \theta, s) - \tilde{\boldsymbol{B}}(\rho', \theta, s)) \cdot (\cos(\theta)\boldsymbol{n}(s) + \sin(\theta)\boldsymbol{b}(s))\, \mathrm{d}\rho'. \tag{5.5}$$

Finally, the coefficients ϕ_{m0} and ϕ_{mm} of this are used as input for $\phi^{(\mathrm{IC})}$. The input is computed to the $N_\rho = 9$ orders in ρ with $N_{s,\mathrm{RE}} = 100$ and $N_{,s\mathrm{LP}} = 50$ Fourier modes in s for the rotating ellipse and Landreman–Paul, respectively (see (4.1)), where we use the subscript 'RE' for the rotating ellipse and 'LP' for Landreman–Paul wherever necessary.

To begin our analysis of the examples, we consider the inputs to the near-axis expansion. Corollary (3.9) suggests that there are two length scales dictated by the inputs. First is the radius of curvature of the axis. Letting $\Sigma = \min_s \kappa^{-1}(s)$ be the radius of curvature, we find

$$\Sigma_{\mathrm{RE}} = 0.987, \qquad \Sigma_{\mathrm{LP}} = 0.681. \tag{5.6}$$

The other length scale of interest is the radius of convergence σ_0 of $\phi^{(\mathrm{IC})}$ in the (σ_0, H^{q+2D})-analytic norm. However, to the orders to which we compute, this radius appears to depend on the exponent $q + 2D$. To determine the most informative exponent, we consider the work by Kappel *et al.* (2024), where it was shown that the normalised gradient of the magnetic field is a strong predictor for plasma–coil separation. Because the $H^2(\mathbb{T}^2)$ norm measures the size of the second derivative of $\phi_m^{(\mathrm{IC})}$ (and therefore the gradient of the input magnetic field), we conjecture this corresponds to the most practical exponent.

To verify this, we first compute the minimum axis-to-coil distance σ_{coil} for both configurations to be

$$\sigma_{\mathrm{coil,RE}} = 0.241, \qquad \sigma_{\mathrm{coil,LP}} = 0.350. \tag{5.7}$$

We note that $\sigma_{\mathrm{coil}} < \Sigma$ for both configurations, indicating that the distance-to-coil is the limiting factor for convergence (cf. Corollary (3.9)). In figure 4, we plot the H^2 norms of $\phi_m^{(\mathrm{IC})}$ versus $A\sigma_{\mathrm{coil}}^{-m}$, where the coefficient A is found via a best fit for both configurations. In both cases, we find there is remarkable agreement, indicating that the H^2 radius of convergence of $\phi^{(\mathrm{IC})}$ could be used as a proxy for distance-to-coils.

5.2. *Magnetic field convergence*

Next, we compute the near-axis expansion via the procedure in box (3.30). We perform the expansion both without regularisation ($P = 0$) and with the regularisation operator in (3.15). For the regularised runs, we use $D = 2$ throughout and vary

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

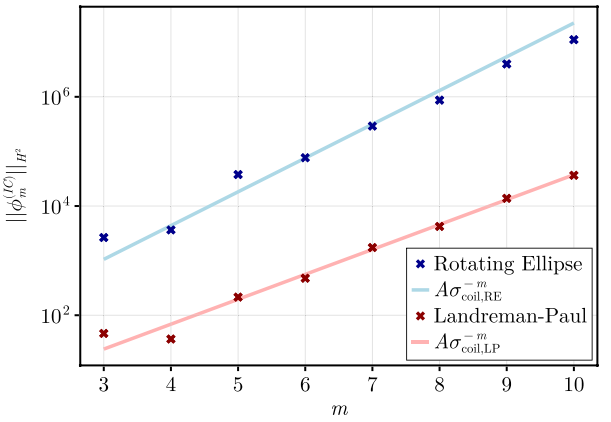


FIGURE 4. Plot of the coefficient norm $\|\phi_m^{(\mathrm{IC})}\|_{H^2}$ versus the order m (markers) and best-fit lines $A\sigma_{\mathrm{coil}}^{-m}$ (lines), where σ_{coil} is the axis-to-coil distance.

K between 10 and 200 to assess how the equilibrium changes between strong and weak regularisation, respectively.

As a first test of the output convergence, we compare the coil magnetic field against the unregularised expansion. Because the input is harmonic, we expect that the near-axis expansion will converge from Proposition (3.2) (unless floating-point errors overwhelm the solution, which is not observed to this order). We verify this by computing the L^2 magnetic field error on surfaces about the axis

$$\|\boldsymbol{B} - \boldsymbol{B}_{\mathrm{coil}}\|_{L^2(\partial\Omega_\sigma^0)} = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} |\boldsymbol{B}(\sigma,\theta,s) - \boldsymbol{B}_{\mathrm{coil}}(\sigma,\theta,s)|^2 \,\mathrm{d}\theta\mathrm{d}s, \tag{5.8}$$

where $\boldsymbol{B}$ is computed from the near-axis expansion and $\boldsymbol{B}_{\mathrm{coil}}$ is computed directly from Biot–Savart.

In figure 5(*a,b*), we plot the error (5.8) versus the normalised distance-from-axis $\sigma/\sigma_{\mathrm{coil}}$. This error is computed for the approximation

$$\phi_{\leqslant N_\rho} \approx \sum_{m=2}^{N_\rho} \phi_m \rho^m, \qquad \boldsymbol{B}_{\leqslant N_\rho - 1} = \nabla\phi_{\leqslant N_\rho} + \tilde{\boldsymbol{B}}_{\leqslant N_\rho - 1} \tag{5.9}$$

where N_ρ is varied from 1 to 8. For both configurations, we find that the error of the magnetic field obeys the expected power law

$$\left\|\boldsymbol{B}_{\leqslant N_\rho - 1} - \boldsymbol{B}_{\mathrm{coil}}\right\|_{L^2(\partial\Omega_\sigma^0)} = \mathcal{O}\left(\sigma^{-N_\rho}\right). \tag{5.10}$$

The error curves for varying N_ρ meet at $\sigma = \sigma_{\mathrm{coil}}$, indicating that the output radius of convergence is limited by the coils. This tells us that the limit of convergence $\sigma = \sigma_0$ is achievable in Corollary (3.9).

Turning to the effects of regularisation, we fix $N_\rho = 9$ and plot the error (5.8) versus $\sigma/\sigma_{\mathrm{coil}}$ for varying K between 10 and 200 in figure 5(*c,d*). We also include the unregularised solution, labelled with $K = \infty$. We find that as the regularisation becomes stronger (K decreases), the magnetic field loses fidelity near the core. We attribute to the increasing loss of accuracy of the high-wavenumber s modes,

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

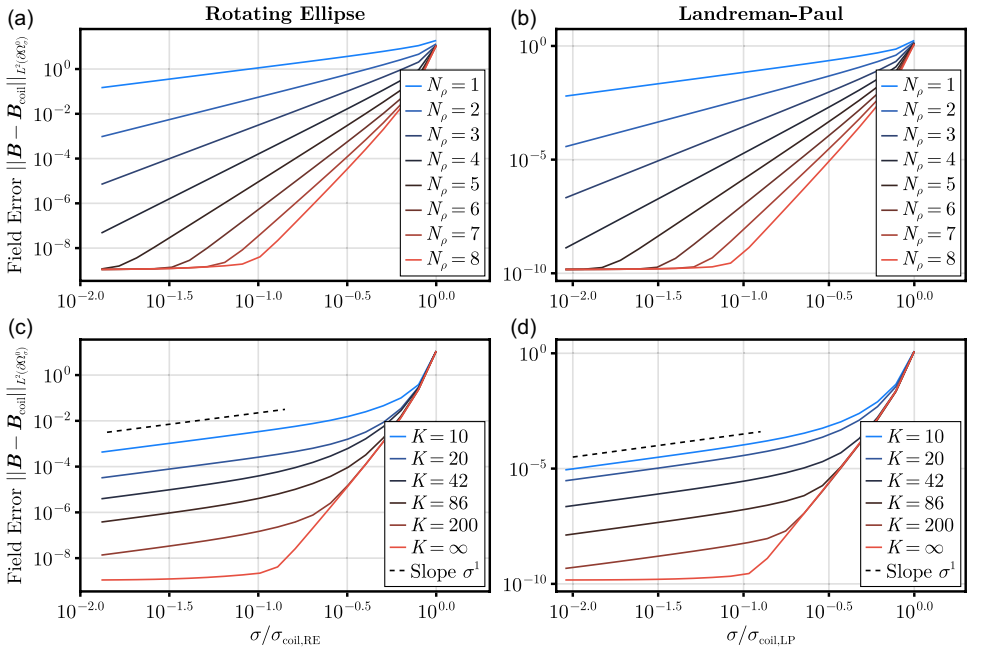


FIGURE 5. (*a,b*) The error (5.8) as a function of the normalised distance from axis $\sigma/\sigma_{\rm coil}$ for varying orders of approximation N_ρ. (*c,d*) The error (5.8) as a function of $\sigma/\sigma_{\rm coil}$ for varying values of the regularisation parameter K ($K = \infty$ is unregularised).

while the low-wavenumber modes maintain accuracy. Then, far from the axis, the regularised error inflects to begin to agree with the rate of convergence of the unregularised solution. So, while the solution loses a high-wavenumber fidelity, the low wavenumbers maintain a similar level of accuracy. Comparing figures 5(*a,b*) with figure 5(*c,d*), we see that a regularised high-order expansion can achieve an equivalent error to an unregularised lower-order expansion near the axis while maintaining that fidelity far from the axis.

To address the role of regularisation more fully, however, we need to consider the fidelity of the expansion on less tuned inputs. To do this, we perturb $\boldsymbol{r}_0$, B_0 and $\phi^{(\rm IC)}$ by the random functions as

$$\delta r_{0,i} = \epsilon \sum_{\ell=-N_s}^{N_s} \frac{X_{i\ell}}{1+(\ell/K_\epsilon)^{4+q}},$$
$$\delta B_0 = \epsilon \sum_{\ell=-N_s}^{N_s} \frac{Y_\ell}{1+(\ell/K_\epsilon)^{1+q}},$$
$$\delta \phi^{(\rm IC)}_{mn} = \epsilon \sum_{\ell=-N_s}^{N_s} \frac{Z_{mn\ell}}{1+(m/K_\epsilon)^{q+2D}+(\ell/K_\epsilon)^{q+2D}}, \qquad n \in \{0, m\}, \tag{5.11}$$

where $X_{i\ell}$, Y_ℓ and $Z_{mn\ell}$ are independent and identically distributed (i.i.d.) unit normal random variables, $q = 2$, and $\epsilon = 10^{-6}$ for the rotating ellipse and $\epsilon = 10^{-4}$ for Landreman–Paul. We have chosen the regularity of the perturbation to align with the inputs in box (3.30).

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

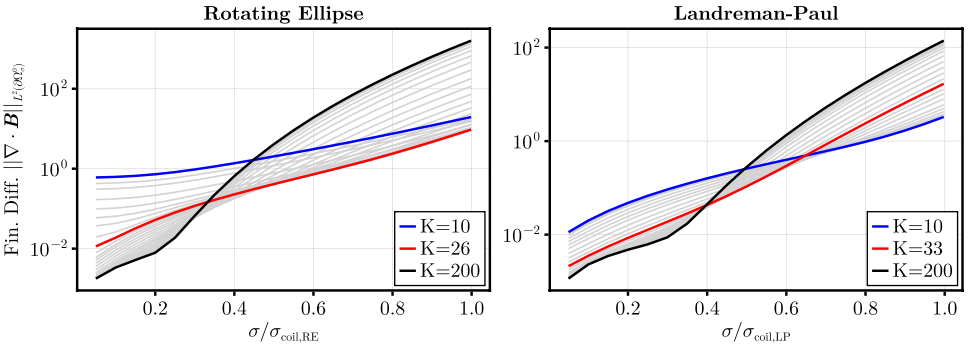


FIGURE 6. Finite difference residual (5.12) as a function of the normalised distance from axis $\sigma/\sigma_{\text{coil}}$ for the perturbed rotating ellipse and Landreman–Paul inputs (see (5.11)). For both plots, three lines are coloured and labelled, while the grey lines represent other values of K interpolating between $K = 10$ and $K = 200$.

In figure 6, we consider the accuracy of the solution to the perturbed problem for K varying between 10 and 200 for both examples with $N_\rho = 9$ fixed. To measure the accuracy, we no longer have a coil set with which to compare the solution directly. So, we instead measure the residual of Poisson's equation

$$\|\nabla \cdot \boldsymbol{B}\|_{L^2(\partial\Omega_\sigma^0)} = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \left|\nabla \cdot (\nabla\phi + \tilde{\boldsymbol{B}})\right|^2 \mathrm{d}\theta\,\mathrm{d}s, \tag{5.12}$$

where we evaluate every derivative (including in the metric) via finite differences. For both perturbed examples, the best solution near the axis is the lightly regularised $K = 200$ solution. However, beyond a certain radius between $0.3\sigma_{\text{coil}}$ and $0.4\sigma_{\text{coil}}$, more regularised solutions improve upon the less regularised ones in the finite difference metric. For our examples, we find that $K = 26$ for the rotating ellipse and $K = 33$ for Landreman–Paul are perhaps the best choices in practice. This figure potentially indicates a more general principle: the further from the axis one wants accuracy of the expansion, the more regularised the expansion likely has to be.

5.3. *Magnetic coordinate convergence*

To compute straight field-line coordinates, we return to the unregularised Landreman–Paul configuration. Then, using the straight field-line magnetic coordinate equation from box (2.60), we compute the approximate coordinates (ξ, η) for N_ρ varying between 2 and 9, where we note the $N_\rho = 2$ approximation of ϕ provides the leading-order field-line behaviour. To find flux surfaces, we then invert (ξ, η) to find the distance-to-axis coordinates $(x(\xi, \eta, s), y(\xi, \eta, s))$, where magnetic surfaces are parametrised by $\xi^2 + \eta^2 = \psi$.

In figure 7, we plot in black the computed surfaces on the $s = 0$ Poincaré section for varying values of N_ρ. For comparison, we plot the intersections of coil magnetic field lines in the background. At leading order, we see the surfaces are elliptical, while higher orders account for more shaping in the θ direction. Then, as the order increases beyond 5, the surfaces away from the core start diverging.

To investigate this divergence, we plot a red circle of constant radius $\sigma_{\text{coil,LP}}$ in the $N_\rho = 9$ panel. We see that the circle appears to separate the divergent surfaces from

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

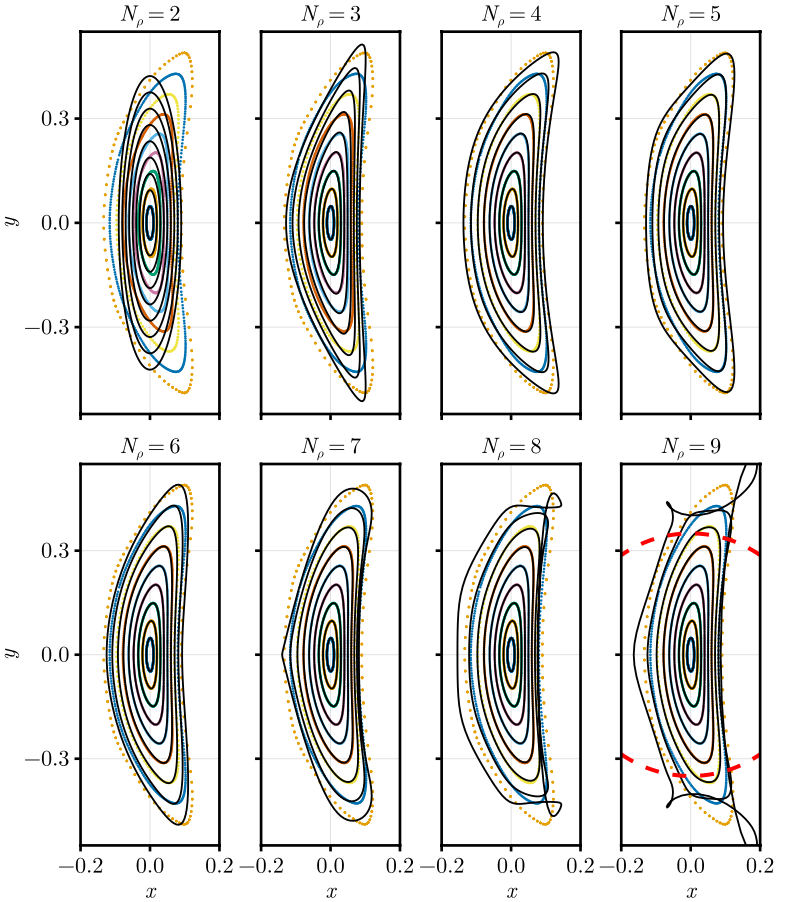


FIGURE 7. Near-axis approximations of flux surfaces for varying orders of approximation N_ρ (black); a Poincaré plot of the true coil magnetic field lines (colour). In the final $N_\rho = 9$ panel, we plot a circle with radius $\sigma_{\text{coil,LP}}$ in red.

the convergent ones. We believe this is the likely reason for the divergence; however, there are still other possibilities.

To assess the errors of surfaces closer to the axis, we turn to a more quantitative measure. To do this, we first use the method from Ruth & Bindel (2024) in the `SymplecticMapTools.jl` package (Ruth 2024*b*) to compute invariant circles $(x_{\text{coil}}(\theta), y_{\text{coil}}(\theta))$ and the rotational transform ι_{coil} on the cross-section from the Poincaré plot trajectories. Then, as a function of the inboard distance x from the axis (see figure 7), we compute rotational transform and parametrisation errors as

$$R_\iota = |\iota - \iota_{\text{coil}}|\,,$$
$$R_{\text{param}} = \left[\int_0^{2\pi} (x(\psi(x,0,0),\theta,0) - x_{\text{coil}}(\theta))^2 + (y(\psi(x,0,0),\theta,0) - y_{\text{coil}}(\theta))^2\,\mathrm{d}\theta\right]^{1/2}. \tag{5.13}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

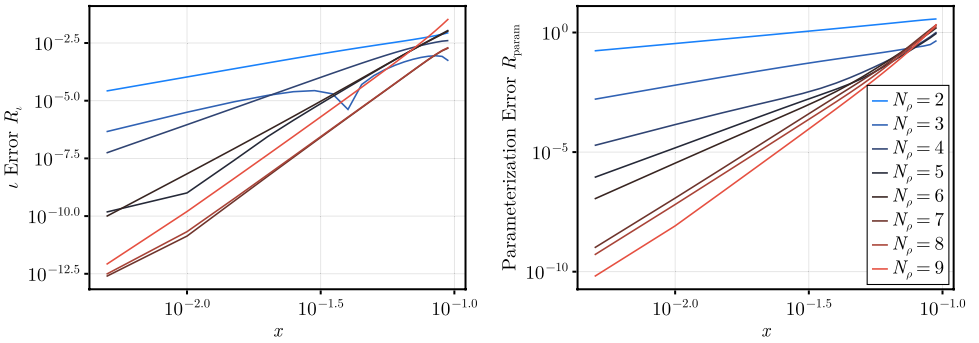


FIGURE 8. (*a*) The rotational transform and (*b*) the parametrisation error R_{param} defined in (5.13) as a function of the inboard x distance (cf. figure 7) for varying orders of approximation N_ρ.

In figure 8, we plot both errors with varying N_ρ. In both cases, the rotational transform and parametrisation converge to high accuracy near the core. However, they begin to diverge before the the outermost surface, agreeing with the visual divergence in figure 7.

6. Conclusion

In this paper, we have investigated the convergence of the near-axis expansion in vacuum, both theoretically and numerically. From the theoretical point of view, we showed in Theorem (3.6) that the near-axis expansion is ill-posed, even in the relatively simple case of vacuum fields. However, as shown in Theorem (3.8), we found the near-axis problem can be regularised giving a guarantee of convergence for appropriately smooth input data. In particular, this tells us that a truncated near-axis expansion is an approximation to the solution of the regularised problem. Combining this with Proposition (3.10), we find can estimate the error of the regularised expansion from a true solution.

From the numerical results, we have verified that the near-axis expansion can converge in vacuum. This includes convergence of surfaces, where we have shown that the rotational transform and surface parametrisations can be approximated near the axis to high accuracy. Moreover, we demonstrated that the radius of convergence of the expansion is directly tied to the minimum distance to coils. Under perturbation, we found that the regularisation reduces the residual of Poisson's equation far from the axis.

Our analysis suggests that the following four quantities should be kept in mind for future optimisation problems.

(i) The axis, on-axis field and higher moments should all be sufficiently regular for the expansion to converge (see box (3.30)). These can be enforced, e.g. by Sobolev norms on the inputs of the near-axis expansion.

(ii) In particular, the H^2-norm radius of convergence of $\phi^{(\text{IC})}$ appears to indicate the distance to coils from the axis (see figure 4). This gives a potential metric for plasma–coil distance.

(iii) The axis curvature also limits to the radius of convergence, so this should be small relative to the desired minimum distance to coils.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

(iv) In the case that the above terms are not sufficient, the error in Proposition (3.10) could be used to monitor the accuracy of the solutions.

Using these metrics, a moderate-order near-axis expansion (say, $4 \leqslant N \leqslant 6$) could be used to explore the space of stellarators more effectively. This could allow for the use of new near-axis optimisation problems.

Looking forward, these results indicate that regularisation is likely also required for the near-axis expansion to converge in pressure. The form of (3.19) gives a potential path forward, where the regularisation could be expressed as a fictitious current. Physically, a link between regularisation and extended MHD models that provide additional current contributions can be studied. However, the issue of small denominators for near-rational ι_0 (see (2.58)) will appear, which will combine with the regularised expansion in a non-trivial way in pressure. It remains to be seen whether regularisation can be used for improved convergence of these surfaces.

Acknowledgments

The authors would like to acknowledge the helpful conversations with Tony Xu, Sean Yang, Gokul Nair and Joshua Burby in the development of this work.

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

Funding

This work was supported by a grant from the Simons Foundation (No. 560651, D.B.) and by DOE DE-SC0024548. This material is based upon work supported by the National Science Foundation under Grant No.2409066. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Declaration of interests

The authors report no conflict of intererst.

Data availability statement

The data that support the finding of this study are openly available in the package StellaratorNearAxis.jl at the GitHub repository – `https://github.com/maxeruth/StellaratorNearAxis.jl`.

Appendix A. Proofs

A.1 *Proof of Proposition* (3.2)

Because $\boldsymbol{B}$ is a vacuum field, its Cartesian components (B^1, B^2, B^3) are also harmonic and therefore real-analytic, meaning at each point on the axis $\boldsymbol{r}_0(s)$, it has a uniformly convergent Taylor series in a ball of size $(\sqrt{2}-1)\sigma$ (Axler *et al.* 2001, Theorem 1.28). By choosing the coefficients ϕ_{mn} to match the Taylor series at each point, we find that the near-axis expansion is uniformly convergent near the axis. Because $\boldsymbol{B}$ is harmonic, the coefficients must satisfy the near-axis problem (2.30). Finally, because the solution to the near-axis expansion is unique (3.1), the proposition is proven.

A.2 *Proof of Proposition* (3.4)

We start with a lemma on derivatives on (σ, W)-analytic functions.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

LEMMA A.1. *Let* $(x, y) = (\rho\cos\theta, \rho\sin\theta)$ *and* W *be* $\mathcal{C}^q(\mathbb{T}^2)$ *or* $H^q(\mathbb{T}^2)$ *for* $q \geqslant 0$. *The derivatives* $(\partial/\partial x)$ *and* $(\partial/\partial y)$ *are bounded operators from* (σ, W)*-analytic functions to* (σ', W)*-analytic functions for all* $0 < \sigma' < \sigma$.

Proof. We will prove this for the x derivative, as the proof for the y derivative is identical. First, we observe that x derivatives preserve the analytic structure (3.4). Let f be (σ, W)-analytic, where W is H^q or $\mathcal{C}^q$ for $q \geqslant 1$. In polar coordinates, we have

$$\frac{\partial f}{\partial x} = \cos\theta \frac{\partial f}{\partial \rho} - \frac{\sin\theta}{\rho}\frac{\partial f}{\partial \theta}. \tag{A.1}$$

For both ρ and θ, we have

$$\left\| \frac{\partial}{\partial \rho}\left(\rho^m f_m\right)\right\|_W \leqslant m\rho^{m-1}\left\|f_m\right\|_W, \qquad \left\|\frac{1}{\rho}\frac{\partial}{\partial\theta}\left(\rho^m f_m\right)\right\|_W \leqslant m\rho^{m-1}\left\|f_m\right\|_W. \tag{A.2}$$

Multiplying by $\sin\theta$ and $\cos\theta$ is bounded on both H^q and $\mathcal{C}^q$, so

$$\begin{aligned}\left\|\left(\frac{\partial f}{\partial x}\right)_m\right\|_W &\leqslant C(m+1)\left\|f_{m+1}\right\|_W, \\ &\leqslant C(m+1)\left\|f\right\|_{\sigma,W}\sigma^{-m-1}, \\ &\leqslant \frac{C(m+1)(\sigma')^m}{\sigma^{m+1}}\left\|f\right\|_{\sigma,W}(\sigma')^{-m} \leqslant C\left\|f\right\|_{\sigma,W}(\sigma')^{-m}.\end{aligned} \tag{A.3}$$

□

LEMMA A.2. *Let* $q \geqslant 1$. *The derivatives* $(\partial/\partial s)$ *and* $(\partial/\partial\theta)$ *are bounded operators from:*

(i) $(\sigma, \mathcal{C}^q)$*-analytic functions to* $(\sigma, \mathcal{C}^{q-1})$*-analytic functions; and*

(ii) (σ, H^q)*-analytic functions to* (σ, H^{q-1})*-analytic functions.*

Proof. Simply notice $(\partial/\partial s)$ and $(\partial/\partial\theta)$ are bounded from $\mathcal{C}^q$ to $\mathcal{C}^{q-1}$ and from H^q to H^{q-1}. □

Combining the two above lemmas, if we choose $\sigma' < \sigma$, $q \geqslant 0$ and f to be a $(\sigma, \mathcal{C}^q)$-analytic function, then for all $m, n, \ell > 0$ such that $m + n + \ell \leqslant q$, the function

$$g = \frac{\partial f}{\partial x^n \partial y^m \partial s^\ell} \tag{A.4}$$

is $(\sigma', \mathcal{C}^0)$-analytic. So, it suffices to prove that g is continuous in $\Omega^0_{\sigma'}$.

Let $0 < \tilde\rho < \sigma'$ and $(\tilde\theta, \tilde s) \in \mathbb{T}^2$. Then, choose ρ^* such that $\tilde\rho < \rho^* < \sigma'$. We will show that g is continuous at the point $(\tilde\rho, \tilde\theta, \tilde s)$. Letting $\rho_1, \rho_2 \leqslant \rho^*$, we can establish a Lipschitz bound of g in ρ:

$$\begin{aligned}\left|g(\rho_1,\theta,s) - g(\rho_2,\theta,s)\right| &\leqslant \sum_{m=\beta}^{\infty}\left\|g_m\right\|_C\left|\rho_1^m - \rho_2^m\right| \\ &\leqslant G\frac{\left|\rho_1-\rho_2\right|}{\rho^*}\sum_{m=1}^{\infty} m\left(\frac{\rho^*}{\sigma}\right)^m \leqslant L\left|\rho_1-\rho_2\right|.\end{aligned} \tag{A.5}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

Now, let $(\rho_m, \theta_m, s_m) \to (\tilde{\rho}, \tilde{\theta}, \tilde{s}) \in \Omega^0_\sigma$ and $\sup(\rho_m) < \rho^* < \sigma$. We have

$$\left| g(\rho_m, \theta_m, s_m) - g(\tilde{\rho}, \tilde{\theta}, \tilde{s}) \right| \leqslant L \left| \rho_m - \rho \right| + \left| g(\tilde{\rho}, \theta_m, s_m) - g(\tilde{\rho}, \tilde{\theta}, \tilde{s}) \right|. \tag{A.6}$$

Because surfaces of g converge in C, both terms converge to zero giving our result.

A.3 *Proof of Theorem* (3.6)

To prove this theorem, we will begin with a few facts about operators on (σ, q)-analytic functions.

LEMMA A.3. *Let f be (σ, H^q)-analytic and g be $(\Sigma, \mathcal{C}^q)$-analytic where $0 < \sigma \leqslant \sigma_0 < \Sigma$ and $q \geqslant 0$. Then, fg is (σ, H^q)-analytic with*

$$\| fg \|_{\sigma,q} \leqslant C \, \| g \|_{\Sigma,\mathcal{C}^q} \, \| f \|_{\sigma,H^q}, \tag{A.7}$$

where C only depends on q, σ_0 and Σ.

Proof. The coefficients of fg are

$$(fg)_m = \sum_{n=0}^{m} f_n g_{m-n}. \tag{A.8}$$

We can bound the norm of $f_n g_{m-n}$ as

$$\| f_n g_{m-n} \|_{H^q} \leqslant C \, \| f_n \|_{H^q} \, \| g_{m-n} \|_{\mathcal{C}^q} \leqslant C \, \| f \|_{\sigma,H^q} \, \| g \|_{\Sigma,\mathcal{C}^q} \, \sigma^{-n} \Sigma^{-(m-n)}, \tag{A.9}$$

where the constant C only depends on q. So,

$$\begin{aligned} \| (fg)_m \|_{H^q} &\leqslant C \, \| f \|_{\sigma,H^q} \, \| g \|_{\Sigma,\mathcal{C}^q} \, \sigma^{-m} \sum_{n=0}^{m} \left(\frac{\sigma}{\Sigma} \right)^{-n} \\ &\leqslant C \, \| f \|_{\sigma,H^q} \, \| g \|_{\Sigma,\mathcal{C}^q} \, \sigma^{-m} \sum_{n=0}^{\infty} \left(\frac{\sigma_0}{\Sigma} \right)^{-n} \\ &\leqslant C \, \| f \|_{\sigma,H^q} \, \| g \|_{\Sigma,\mathcal{C}^q} \, \sigma^{-m} \frac{1}{1 - \sigma_0/\Sigma}, \end{aligned} \tag{A.10}$$

giving the result. □

LEMMA A.4. *Let a^α be $(\Sigma, \mathcal{C}^q)$-analytic for all degree 3 multi-indices $|\alpha| \leqslant m$, $0 < \sigma' < \sigma < \Sigma$ and $q \geqslant 0$. The operator*

$$L = \sum_{|\alpha| \leqslant m} a^\alpha \frac{\partial^{|\alpha|} f}{\partial x^{\alpha_1} \partial y^{\alpha_2} \partial s^{\alpha_3}} \tag{A.11}$$

is bounded from (σ, H^{q+m})-analytic functions to (σ', H^q)-analytic functions.

Proof. Combine Lemmas (A.1), (A.2) and (A.3). □

COROLLARY A.5. *Let $\boldsymbol{r}_0$ be $\mathcal{C}^{4+q}$ for $q \geqslant 0$ and $\ell', \kappa > 0$. There exists a $\Sigma > 0$ such that the Laplacian, defined in (x, y, s)-coordinates via the left-hand side of (2.23), is a bounded operator from (σ, H^{q+2}) to (σ', H^q). Similarly, the divergence operator,*

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

defined in (x, y, s) coordinates via the right-hand side of (2.23), is a bounded operator from (σ, H^{q+1}) to (σ', H^q).

Proof. In (x, y, s)-coordinates, we have $h_s = 1 - \kappa x$, $\sqrt{g} = \ell' h_s$ and

$$g^{-1} = \frac{1}{h_s^2} \begin{pmatrix} h_s^2 + \tau^2 y^2 & -\tau^2 xy & \frac{\tau y}{\ell'} \\ -\tau^2 xy & h_s^2 + \tau^2 x^2 & \frac{-\tau x}{\ell'} \\ \frac{\tau y}{\ell'} & -\frac{\tau x}{\ell'} & \frac{1}{(\ell')^2} \end{pmatrix}. \tag{A.12}$$

Because $\boldsymbol{r}_0 \in \mathcal{C}^{4+q}$, $\ell' \in \mathcal{C}^{3+q}$, $\kappa \in \mathcal{C}^{2+q}$ and $\tau \in \mathcal{C}^{1+q}$, so $g^{-1} \in \mathcal{C}^{1+q}$. This means the elements of $h_s^2 g^{-1}$ are $(\sigma, \mathcal{C}^{1+q})$-analytic with finite series. To include the factor of h_s^{-2}, we have

$$\frac{1}{h_s} = \sum_{m=0}^{\infty} (\kappa \cos\theta)^m \rho^m. \tag{A.13}$$

The function $\kappa \cos\theta$ is in $\mathcal{C}^{2+q}$, so $\|(\kappa \cos\theta)^m\|_{H^{2+q}} \leqslant C^m \|\kappa \cos\theta\|^m_{\mathcal{C}^{2+q}}$ for some $C \geqslant 1$ and h_s^{-1} is $(\Sigma, \mathcal{C}^{2+q})$ for some $0 < \Sigma \leqslant \|\kappa\|^{-1}_{C^0}$. Finally, after performing the chain rule to bring the operator to the form in Lemma (A.4), the coefficients are each in (Σ, H^q)-analytic, giving the result. The same argument applies to the divergence operator. □

The last ingredient needed for the proof of ill-posedness is Cauchy's estimates for harmonic functions.

THEOREM A.6 *(Cauchy's estimates (Axler et al. 2001)). Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ be a multi-index. Then, for some constant $C_\alpha > 0$, all harmonic functions ϕ bounded by M on the radius-R ball $B(x, R)$ satisfy the inequality*

$$|D^\alpha \phi(x)| \leqslant \frac{C_\alpha M}{R^{|\alpha|}}. \tag{A.14}$$

Now, we return to the proof of Theorem (3.6). It is sufficient to show that ϕ is not (σ, H^q)-analytic for $q = 2$. So, consider input data such that $\|\phi\|_{\sigma, H^2} < \infty$, say constructed via Proposition (3.2). We will focus on perturbations in B_0 of the form $\Delta B_0^{(j)} = c_j e^{\mathrm{i}js}$, where $c_j = j^{-N}$ and $N > q_0 + 1$ is a positive integer, while $\boldsymbol{r}_0$ and $\phi^{(\mathrm{IC})}$ remain constant. Clearly, $\|\Delta B_0\|_{H^{1+q_0}} \to 0$ as $j \to \infty$. Let $\phi^{(j)}$ be the formal power series solution of the perturbed problem.

We want to show that $\phi^{(j)}$ cannot converge to a (σ, H^q)-analytic solution. In the case where the perturbed near-axis expansion for fixed j does not converge in Ω_σ, then $\|\phi^{(j)}\|_{\sigma, H^2} = \infty$ for all σ and we are done, as the operator fails to be bounded for a specific function. Otherwise, suppose the near-axis expansion solution converges on Ω_σ and each $\phi^{(j)}$ is (σ, H^2)-analytic. Then, using Corollary (13), for $\sigma' < \sigma$, $\Delta\phi^{(j)} + \nabla \cdot \tilde{\boldsymbol{B}}$ is a (σ', L^2)-analytic function. Because ϕ satisfies the near-axis expansion, the solution must satisfy $(\Delta\phi^{(j)} + \nabla \cdot \tilde{\boldsymbol{B}})_m = 0$ in $L^2(\mathbb{T}^2)$, further implying that $\Delta\phi^{(j)} + \nabla \cdot \tilde{\boldsymbol{B}} = 0$ in $L^2(\Omega^0_{\sigma'})$. Pulling this back to $\Omega_{\sigma'}$, we are solving the standard Poisson's equation $\Delta\phi^{(j)} = -\nabla \cdot \tilde{\boldsymbol{B}}$ in $L^2(\Omega_{\sigma'})$. If we locally define

$$\psi^{(j)} = \phi^{(j)} + \int_0^s B_0(s')\ell'(s')\,\mathrm{d}s', \tag{A.15}$$

we find $\Delta\psi^{(j)} = 0$. That is, $\psi^{(j)}$ is analytic in simply connected subdomains of $\Omega_{\sigma'}$ and the magnetic field is locally the gradient of $\psi^{(j)}$.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

Then, let $B(x, R)$ be a ball around a point on the magnetic axis $\boldsymbol{r}_0(s_0)$. At $\boldsymbol{r}_0(s_0)$, the order $M > N$ derivative in the tangent direction of $\boldsymbol{r}_0$ of $\psi^{(j)}$ takes the polynomial form

$$\frac{\partial^M \psi^{(j)}}{\partial s^M} = \frac{\partial^M B_0}{\partial s^M} + c_j \sum_{\ell=0}^{M} a_\ell j^\ell, \tag{A.16}$$

where $a_M = (\mathrm{i}/\ell'(s_0))^M e^{\mathrm{i}js_0} \neq 0$ and a_ℓ for $\ell < M$ contain higher order derivatives of the axis. As such, there is a J such that $j > J$ implies that

$$\left| \frac{\partial^M \psi^{(j)}}{\partial s^M} \right| > \frac{|a_N|}{2} j^{M-N}. \tag{A.17}$$

Then, by Theorem (A.6), we have that

$$\max_{B(x,R)} \psi^{(j)} > \frac{R^N a_N}{2C_N} j^{M-N}. \tag{A.18}$$

So, as $j \to \infty$, $\psi^{(j)}$ cannot converge to a continuous function. However, because ϕ was assumed (σ, H^2)-analytic and by Corollary (3.5) it must be continuous, we have drawn a contradiction.

A.4 *Proof of Theorem* (3.8)

We will start with two lemmas. The first is on the boundedness of the right-hand-side operator of the PDE (3.20).

LEMMA A.7. *Let $q \geqslant 0$, $D \geqslant 1$, $0 < \sigma \leqslant \sigma_0 < \Sigma$, ϕ be (σ, H^{q+2D})-analytic, and $a^{(j)}$ be $(\Sigma, \mathcal{C}^q)$-analytic for $1 \leqslant j \leqslant 7$. Then, the operator $L^{(a)}$ defined in (3.21) preserves the analytic form (3.4) and satisfies the bound*

$$\left\| (L^{(a)}\phi)_m \right\|_{H^q} \leqslant C(m+1) \left(\sum_{j=1}^{7} \left\| a^{(j)} \right\|_{\Sigma, \mathcal{C}^q} \right) \|\phi\|_{\sigma, H^{q+2D}} \, \sigma^{-(m+1)}, \tag{A.19}$$

where C depends on σ_0 but not on σ.

Proof. For $j \neq 2$, Lemmas (A.2) and (A.3) tell us that

$$L^{(\neq 2)}\phi = a^{(1)}\phi + a^{(3)}\frac{\partial \phi}{\partial \theta} + a^{(4)}\frac{\partial \phi}{\partial s} + a^{(5)}\frac{\partial^2 \phi}{\partial \theta^2} + a^{(6)}\frac{\partial^2 \phi}{\partial \theta \partial s} + a^{(7)}\frac{\partial^2 \phi}{\partial s^2} \tag{A.20}$$

satisfies the analytic form and has the bound

$$\left\| L^{(\neq 2)}\phi \right\|_{\sigma, H^q} \leqslant C \left(\sum_{j \neq 2} \left\| a^{(j)} \right\|_{\Sigma, \mathcal{C}^q} \right) \|\phi\|_{\sigma, H^{q+2D}}, \tag{A.21}$$

where C does not depend on σ.

For the $j = 2$ term, define

$$L^{(2)}\phi = \frac{\partial a^{(2)}}{\partial \rho}\frac{\partial \phi}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial a^{(2)}}{\partial \theta}\frac{\partial \phi}{\partial \theta} = \frac{\partial a^{(2)}}{\partial x}\frac{\partial \phi}{\partial x} + \frac{\partial a^{(2)}}{\partial y}\frac{\partial \phi}{\partial y}, \tag{A.22}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

where $x = \rho\cos\theta$ and $y = \rho\sin\theta$. By Lemmas (A.1) and (A.3), this preserves the analytic form. To bound the operator, first note that

$$\left\| \left(\frac{\partial \phi}{\partial \rho} \right)_m \right\|_{H^q} = (m+1) \left\| \phi_{m+1} \right\|_{H^q} \leqslant (m+1)\sigma^{-(m+1)} \left\| \phi \right\|_{\sigma, H^{q+2D}}. \tag{A.23}$$

A similar bound is satisfied by the θ derivative:

$$\left\| \left(\frac{\partial \phi}{\partial \theta} \right)_m \right\|_{H^q} \leqslant m\sigma^{-m} \left\| \phi \right\|_{\sigma, H^{q+2D}} . \tag{A.24}$$

So, we focus on the ρ derivative term of $L^{(2)}$, where the same steps can be used to bound the θ derivative term. We have that

$$\begin{aligned}
\left\| \left(\frac{\partial a^{(2)}}{\partial \rho} \frac{\partial \phi}{\partial \rho} \right)_m \right\|_{H^q} &\leqslant \left\| \sum_{n=0}^{m} \left(\frac{\partial a^{(2)}}{\partial \rho} \right)_{n+1} \left(\frac{\partial \phi}{\partial \rho} \right)_{m-n+1} \right\|_{H^q} \\
&\leqslant C \left\| a^{(2)} \right\|_{\Sigma, \mathcal{C}^q} \left\| \phi \right\|_{\sigma, H^{q+2D}} \sum_{n=0}^{m} (n+1)(m-n+1)\Sigma^{-(n+1)}\sigma^{-(m-n+1)} \\
&\leqslant C(m+1) \left\| a^{(2)} \right\|_{\Sigma, \mathcal{C}^q} \left\| \phi \right\|_{\sigma, H^{q+2D}} \sigma^{-(m+1)} \sum_{n=0}^{m+1} (n+1) \left(\frac{\sigma_0}{\Sigma} \right)^n \\
&\leqslant C(m+1) \left\| a^{(2)} \right\|_{\Sigma, \mathcal{C}^q} \left\| \phi \right\|_{\sigma, H^{q+2D}} \sigma^{-(m+1)}.
\end{aligned} \tag{A.25}$$

Combining the estimates on $L^{(2)}$ and $L^{(\neq 2)}$, we have our theorem. □

Then, the main step in proving Theorem (3.8) is to show the inductive step in Lemma (A.9). For this, we depend on the following interior regularity theorem for the regularisation.

THEOREM A.8 ((Taylor (2011), Theorem 11.1)). *If P is elliptic of order $2D$ and $u \in \mathcal{D}'(M)$, $Pu = h \in H^q(M)$, then $u \in H^{q+2D}_{\text{loc}}(M)$, and, for each $U \subset\subset V \subset\subset M$, $\sigma < q + 2D$, there is an estimate*

$$\left\| u \right\|_{H^{q+2D}(U)} \leqslant C \left\| Pu \right\|_{H^q(V)} + C \left\| u \right\|_{H^\sigma(V)} . \tag{A.26}$$

Then, our inductive step is the following lemma.

LEMMA A.9. *Assume the hypotheses of Theorem (3.8), and let $\sigma \leqslant \sigma_0 \leqslant \sigma' < \Sigma$ and $m \geqslant 0$. Suppose we have computed the finite solution*

$$\phi_{<m} = \sum_{n=0}^{m-1} \rho^n \phi_n(\theta, s), \qquad \phi_n = \sum_{\ell=0}^{n} \phi_{n\ell}(s) e^{(2\ell - n)\mathrm{i}\theta}, \tag{A.27}$$

where

$$\left\| \phi_n \right\|_{H^{q+2D}} < \Phi\sigma^{-n} \qquad \text{for all } n < m. \tag{A.28}$$

Then,

$$\phi_m = \sum_{n=0}^{m} \phi_{mn}(s) e^{(2n-m)\mathrm{i}\theta} \tag{A.29}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

and there exists two constants $C_1, C_2 > 0$ independent of m, σ, σ_0, $\phi^{(\mathrm{IC})}$, f and Φ, where C_1 depends continuously on $a^{(j)}$ and σ', and C_2 is independent of $a^{(j)}$ such that

$$\begin{aligned}\|\phi_m\|_{H^{q+2D}} &< C_1\Phi\sigma^{-m+1} + \|\phi_m^{(\mathrm{IC})}\|_{\sigma,H^{q+2D}} + \frac{C_2}{m+1}\|f_m\|_{H^q}\,,\\ &< C_1\Phi\sigma^{-m+1} + (\|\phi^{(\mathrm{IC})}\|_{\sigma_0,H^{q+2D}} + C_2\|f\|_{\sigma_0,H^q})\sigma_0^{-m}. \end{aligned} \tag{A.30}$$

Proof. Let $L^{(a)}$ be as in (3.21). Then, the near-axis iteration is given by

$$\phi_m = \phi_m^{(\mathrm{IC})} + ((1+P)^{-1}\Delta_\perp^+(f + L^{(a)}\phi_{<m}))_m, \tag{A.31}$$

where $\Delta_\perp^+$ is defined in (3.24). As such, the triangle inequality gives

$$\|\phi_m\|_{H^{q+2D}} \leqslant \left\|\phi_m^{(\mathrm{IC})}\right\|_{H^{q+2D}} + \left\|(1+P)^{-1}(\Delta_\perp^+(f + L^{(a)}\phi_{<m}))_m\right\|_{H^{q+2D}}. \tag{A.32}$$

For the initial conditions, we have

$$\left\|\phi_m^{(\mathrm{IC})}\right\|_{H^{q+2D}} \leqslant \left\|\phi_m^{(\mathrm{IC})}\right\|_{\sigma_0,H^{q+2D}}\sigma_0^{-m}, \tag{A.33}$$

so we just need to focus on the second term.

Let $g = f + L^{(a)}\phi_{<m}$. By Lemma (A.7), we have that

$$\|g_m\|_{H^q} \leqslant \|f\|_{\sigma_0,H^q}\sigma_0^{-m} + C(m+1)\left(\sum_{j=1}^{7}\left\|a^{(j)}\right\|_{\Sigma,\mathcal{C}^q}\right)\Phi\sigma^{-(m+1)}, \tag{A.34}$$

where C depends on σ' and we have used $\|\phi_{<m}\|_{\sigma,H^q} \leqslant \Phi$. Next, we establish a bound for the inverse polar Laplacian. We have that

$$\left(\Delta_\perp^+ g\right)_m = \sum_{n=1}^{m-1}\frac{1}{m^2-(2n-m)^2}g_{m-2,n-1}e^{(2n-m)\mathrm{i}\theta}, \tag{A.35}$$

so for $m \geqslant 2$,

$$\begin{aligned}\left\|\left(\Delta_\perp^+ g\right)_m\right\|_{H^q}^2 &= \sum_{n=1}^{m-1}\frac{1}{m^2-(2n-m)^2}\left\|g_{m-2,n-1}e^{(2n-m)\mathrm{i}\theta}\right\|_{H^q}^2,\\ &\leqslant \frac{1}{4(m-1)}\sum_{n=1}^{m-1}\left\|g_{m-2,n-1}e^{(2n-m)\mathrm{i}\theta}\right\|_{H^q}^2,\\ &= \frac{1}{4(m-1)}\|g_{m-2}\|_{H^q}^2,\\ &\leqslant \frac{C^{(1)}}{m+1}\|f\|_{\sigma_0,H^q}\sigma_0^{-m} + C^{(2)}\left(\sum_{j=1}^{7}\left\|a^{(j)}\right\|_{\Sigma,\mathcal{C}^q}\right)\Phi\sigma^{-m+1}, \end{aligned} \tag{A.36}$$

where we used constants so that this is trivially extended to $m \in \{0,1\}$, where $(\Delta_\perp^+ g)_0 = (\Delta_\perp^+ g)_1 = 0$.

Now, we would like to bound the inverse operator $(1+P)^{-1}$. Specifically, we need it to be the case that

$$\|u\|_{H^{q+2D}} < C\,\|(1+P)u\|_{H^q}. \tag{A.37}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

We will prove this is true by the standard argument. For the sake of contradiction, suppose that there exists a sequence $u^{(n)} \in H^{q+2D}$ such that $\left\|u^{(n)}\right\|_{L^2} = 1$ and $\left\|(1+P)u^{(n)}\right\|_{H^q} \to 0$. By Theorem (A.8), this tells us that $\left\|u^{(n)}\right\|_{H^{q+2D}} \leqslant 1+\epsilon$ for some $\epsilon > 0$. By Rellich's theorem (Taylor 2011, Proposition 3.4), H^{q+2D} is compactly embedded in H^{q+2D-1}, so there exists a subsequence such that $u_n \to u$ in H^{q+2D-1}. This implies that $(1+P)u^{(n)} \to (1+P)u = 0$ in H^{q-1}, where we are using the assumption that $q \geqslant 1$. However, because $1+P$ is positive, it does not have a kernel, so the bound (A.37) must hold.

Equation (A.37) tells us that $1+P$ is one-to-one from H^{q+2D} to H^q. Moreover, because $1+P$ is positive and self-adjoint, it must be surjective, so it is invertible and we have the inequality

$$\left\|(1+P)^{-1}\left(\Delta_\perp^+ g\right)_m\right\|_{H^{q+2D}}^2 \leqslant \frac{C_1}{m+1}\left\|f\right\|_{\sigma_0,H^q}\sigma_0^{-m} + C_2\left(\sum_{j=1}^{7}\left\|a^{(j)}\right\|_{\Sigma,C^q}\right)\Phi\sigma^{-m+1}, \tag{A.38}$$

proving the lemma. □

We are now ready to prove Theorem (3.8). For continuity (boundedness) of the solution with respect to $\phi^{(\mathrm{IC})}$ and f, we choose a σ such that $\sigma C_1 < 1$ and $\sigma \leqslant \sigma_0$. Then, let $0 < \gamma < 1 - C_1\sigma$. We choose Φ such that

$$C_1\sigma\Phi + \left\|\phi^{(\mathrm{IC})}\right\|_{\sigma,H^{q+2D}} + C_2\left\|f\right\|_{\sigma_0,H^q} = (1-\gamma)\Phi. \tag{A.39}$$

Then, we perform induction. At $m = 0$, $\phi_{<m} = 0$, so we have trivially satisfied the initial case. For the inductive step, because $\sigma_0^{-m} \leqslant \sigma^{-m}$, we have

$$\left\|\phi_m\right\|_{H^{q+2D}} \leqslant (1-\gamma)\Phi\sigma^{-m}, \tag{A.40}$$

implying

$$\left\|\phi\right\|_{\sigma,H^{q+2D}} \leqslant \Phi = \frac{1-\gamma}{1-\gamma-C_1\sigma}\left(\left\|\phi^{(\mathrm{IC})}\right\|_{\sigma,H^{q+2D}} + C_2\left\|f\right\|_{\sigma_0,H^q}\right). \tag{A.41}$$

For continuity with respect to the coefficients $A = (a^{(1)}, \ldots, a^{(7)})$, consider fixing σ and γ as before. Then, because C_1 is continuous with respect to A, there is a small enough perturbation so that both σ and γ continue to satisfy $0 < \gamma < 1 - C_1\sigma$ and $C_1\sigma < 1$. So, there is a neighbourhood of U of A such that for all $A + \delta A \in U$ and some values of C, we have $\|\phi\|_{\sigma,H^{q+2D}} \leqslant C(\|\phi^{(\mathrm{IC})}\|_{\sigma_0,H^{q+2D}} + \|f\|_{\sigma_0,H^q})$. Because C_1 does not depend on σ_0, it is also the case that there is a neighbourhood of U_0 of A such that $\|\phi\|_{\sigma,H^{q+2D}} \leqslant C_0(\|\phi^{(\mathrm{IC})}\|_{\sigma,H^{q+2D}} + \|f\|_{\sigma_0,H^q})$ for all $A + \delta A \in U_0$.

Now, consider the full PDE operator

$$L = \Delta_\perp(1+P) - L^{(a)}, \tag{A.42}$$

and let $L + \delta L$ be the operator associated with substituting $a^{(j)}$ with $a^{(j)} + \delta a^{(j)} \in U \cap U_0$, i.e.

$$\begin{aligned}\delta L\phi = \delta L^{(a)}\phi = \delta a^{(1)}\phi + \frac{\partial(\delta a^{(2)})}{\partial\rho}\frac{\partial\phi}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial(\delta a^{(2)})}{\partial\theta}\frac{\partial\phi}{\partial\theta} \\ + \delta a^{(3)}\frac{\partial\phi}{\partial\theta} + \delta a^{(4)}\frac{\partial\phi}{\partial s} + \delta a^{(5)}\frac{\partial^2\phi}{\partial\theta^2} + \delta a^{(6)}\frac{\partial^2\phi}{\partial\theta\partial s} + \delta a^{(7)}\frac{\partial^2\phi}{\partial s^2}.\end{aligned} \tag{A.43}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

With fixed initial conditions $\phi^{(\mathrm{IC})}$ and f, let the solution to the original PDE be ϕ (i.e. $L\phi = f$) and the solution to the perturbed PDE be $\phi + \delta\phi$ (i.e. $(L + \delta L)(\phi + \delta\phi) = f$). Subtracting the two PDE formulae gives

$$(L + \delta L)\delta\phi = -\delta L\phi, \tag{A.44}$$

where $\delta\phi$ satisfies $\delta\phi^{(\mathrm{IC})} = 0$. Using Lemma (A.7), we have that

$$\|(\delta L\phi)_m\|_{H^q} \leqslant C(m+1)\left(\sum_{j=1}^{7} \left\|\delta a^{(j)}\right\|_{\Sigma,\mathcal{C}^q}\right) \|\phi\|_{\sigma,H^{q+2D}}\, \sigma^{-m}. \tag{A.45}$$

Then, we find that

$$C_1\Phi\sigma^{-m+1} + \frac{C_2}{m+1}\,\|f_m\|\,H^q \leqslant \left(C_1\sigma + C\left(\sum_{j=1}^{7} \left\|\delta a^{(j)}\right\|_{\Sigma,\mathcal{C}^q}\right)\right)\sigma^{-m}. \tag{A.46}$$

We can take δA small enough so the right term is less than 1, allowing us to proceed inductively as before, giving continuity in the coefficients.

A.5 *Proof of Corollary* (3.9)

For the coefficients of the PDE, we must only notice that

$$\frac{1}{1-\kappa\rho\cos\theta} = \sum_{n=1}^{\infty} \rho^n\,(\kappa\cos\theta)^n\,. \tag{A.47}$$

Because $\boldsymbol{r}_0 \in \mathcal{C}^{4+q}(\mathbb{T})$, $\kappa \in \mathcal{C}^{3+q}(\mathbb{T})$. This immediately tells us that

$$\|(\kappa\cos\theta)^n\|_{\mathcal{C}^{3+q}} \leqslant (C\,\|\kappa\cos\theta\|_{\mathcal{C}^{3+q}})^n \tag{A.48}$$

for some constant C, showing this converges. (In fact, the sum converges for all $\rho < \kappa$.)

We note that P satisfies the Hypotheses (3.7) by construction, so we can apply Theorem (3.8). The regularity of $\boldsymbol{B}_K$ is the obtained from (3.18), where the reduction in regularity comes from the order of P, combined with Lemmas (A.1) and (A.3).

A.6 *Proof of Propsition* (3.10)

For the statement about uniqueness, suppose $\phi' \in H^{q+2D}(\Omega^0_{\sigma'})$ with $\phi' - \phi = \delta\phi \neq 0$ satisfies the boundary value problem. Then, $\delta\phi$ satisfies

$$-(P\Delta_\perp + L)\delta\phi = 0, \qquad \delta\phi|_{\rho=\sigma'} = 0. \tag{A.49}$$

Then, after some algebra, we have

$$\left\langle \frac{\partial\delta\phi}{\partial\rho},\ P\frac{\partial\delta\phi}{\partial\rho}\right\rangle + \left\langle \frac{1}{\rho}\frac{\partial\delta\phi}{\partial\theta},\ P\frac{1}{\rho}\frac{\partial\delta\phi}{\partial\rho}\right\rangle - \langle\delta\phi,\ L\delta\phi\rangle = 0, \tag{A.50}$$

where the inner product is the L^2 inner product

$$\langle f,\ g\rangle = \int_0^{2\pi}\int_0^{2\pi}\int_0^{\sigma'} f(\rho,\theta,s)g(\rho,\theta,s)\,\rho\mathrm{d}\rho\mathrm{d}\theta\mathrm{d}s. \tag{A.51}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

By our assumptions on P and L, each term in (A.50) is positive. So, it then must be the case that $\langle \delta\phi,\ L\delta\phi\rangle = 0$. However, this is only possible when $\delta\phi = 0$, so the solution is unique.

For the error estimate, fix ϕ and subtract the two boundary value problems to find that

$$L(\phi - \tilde{\phi}) = -P\Delta_\perp \phi, \qquad (\phi - \tilde{\phi})\Big|_{\rho=\sigma'} = 0. \tag{A.52}$$

Because L is negative, there is a unique solution $\phi - \tilde{\phi}$ to this problem. Then, by standard regularity theory (Evans 2010, Theorem 6.3.5), we have the desired bound.

Appendix B. Asymptotic expansions of basic operations

To build the near-axis code, we need some facts from formal expansions. We let A, B and C be smooth formal power series of the generic form

$$A(\rho, \theta, s) = \sum_{n=0}^{\infty} A_n(\theta, s)\rho^n, \tag{B.1}$$

and let $\alpha \in \mathbb{R}$. Here, we explain how we numerically perform the following basic operations:

(i) multiplication (Appendix B.1), $C = AB$;

(ii) multiplicative inversion (Appendix B.2), $B = A^{-1}$;

(iii) differentiation with respect to ρ (Appendix B.3), $B = (\mathrm{d}A/\mathrm{d}\rho)$;

(iv) exponentiation (Appendix B.4), $B = e^{\alpha A}$;

(v) power (Appendix B.5), $B = A^{\alpha}$;

(vi) composition (Appendix B.6), $A(\tilde{\rho}, \tilde{\theta}, \phi)$ where $\begin{pmatrix} B & C \end{pmatrix}^T = \begin{pmatrix} \tilde{\rho}\cos\tilde{\theta} & \tilde{\rho}\sin\tilde{\theta} \end{pmatrix}^T$;

(vii) series inversion (Appendix B.6), find the inverse coordinate transformation $C^{(i)}$ of the transformation $B(i)$, i.e.

$$\begin{pmatrix} C^{(1)}(B^{(1)}(x, y, s), B^{(2)}(x, y, s), s) \\ C^{(2)}(B^{(1)}(x, y, s), B^{(2)}(x, y, s), s) \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}. \tag{B.2}$$

We find that these operations build upon each other, with multiplication, ρ-differentiation and series composition being the main building blocks of other, more complicated algorithms.

B.1 *Multiplication*

The most basic problem is that of (matrix) multiplication. Let C be the solution to

$$C = AB. \tag{B.3}$$

Via simple matching of orders, we find that

$$C_n(\theta, s) = \sum_{m=0}^{n} A_m(\theta, s) B_{n-m}(\theta, s). \tag{B.4}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

B.2 *Inversion*

Now, instead consider the problem of finding the (matrix) inverse

$$B = A^{-1}. \tag{B.5}$$

It is easier to write this in terms of the problem

$$AB = I. \tag{B.6}$$

So, using (B.4), we have

$$\sum_{m=0}^{n} A_m B_{n-m} = \begin{cases} I, & n = 0, \\ 0, & n > 0. \end{cases} \tag{B.7}$$

Assuming A_0 is invertible, an iterative method for finding B_n is

$$B_n = \begin{cases} A_0^{-1}, & n = 0, \\ -B_0 \left(\sum_{m=1}^{n} A_m B_{n-m} \right), & n > 0. \end{cases} \tag{B.8}$$

B.3 ρ *Derivatives*

Let

$$B = \frac{\mathrm{d}A}{\mathrm{d}\rho}. \tag{B.9}$$

Then, we have

$$B = \sum_{n=0}^{\infty} n A_n \rho^{n-1}, \tag{B.10}$$

or

$$B_n = (n+1) A_{n+1}. \tag{B.11}$$

B.4 *Exponentiation*

A more complicated series operation is scalar exponentiation (see Knuth 1997). We would like to find

$$B = e^{\alpha A}. \tag{B.12}$$

Taking a derivative with respect to ρ, we find

$$\begin{aligned} \frac{\mathrm{d}B}{\mathrm{d}\rho} &= \frac{\mathrm{d}e^{\alpha A}}{\mathrm{d}A} \frac{\mathrm{d}A}{\mathrm{d}\rho}, \\ &= \alpha e^{\alpha A} \frac{\mathrm{d}A}{\mathrm{d}\rho}, \\ &= \alpha B \frac{\mathrm{d}A}{\mathrm{d}\rho}. \end{aligned} \tag{B.13}$$

Now, we can write out the multiplication using (B.4) and (B.11), giving

$$(n+1) B_{n+1} = \alpha \sum_{m=0}^{n} (m+1) A_{m+1} B_{n-m}, \tag{B.14}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

where $B_0 = e^{\alpha A_0}$. From this, we have

$$B_n = \begin{cases} e^{\alpha A_0}, & n = 0, \\ \frac{\alpha}{n} \sum_{m=0}^{n-1} (m+1) A_{m+1} B_{n-1-m}, & n > 0. \end{cases} \tag{B.15}$$

Note that through exponentiation and inversion, we can obtain any trigonometric or hyperbolic trigonometric function. For instance, we have

$$\sin A = Im e^{\mathrm{i}A}, \tag{B.16}$$

and

$$\operatorname{sech} A = 2\left(e^{A} + e^{-A}\right)^{-1}. \tag{B.17}$$

As a more detailed example, consider the problem of simultaneously computing $\sin(A)$ and $\cos(A)$. Using $\alpha = i\omega$ in (B.15), we find

$$\cos(\omega A)_n + \mathrm{i}\sin(\omega A)_n = \begin{cases} \cos(\omega A_0) + \mathrm{i}\sin(\omega A_0), & n = 0, \\ -\sum_{m=0}^{n-1} \frac{\omega(m+1)}{n} A_{m+1} \sin(A)_{n-1-m} & \\ +\mathrm{i}\sum_{m=0}^{n-1} \frac{\omega(m+1)}{n} A_{m+1} \cos(A)_{n-1-m}, & n > 0. \end{cases} \tag{B.18}$$

From this, we can work with only real-valued series via the formulae

$$\cos(A)_n = \begin{cases} \cos(A_0), & n = 0, \\ -\sum_{m=0}^{n-1} \frac{\omega(m+1)}{n} A_{m+1} \sin(A)_{n-1-m}, & n > 0, \end{cases} \tag{B.19}$$

$$\sin(A)_n = \begin{cases} \sin(A_0), & n = 0, \\ \sum_{m=0}^{n-1} \frac{\omega(m+1)}{n} A_{m+1} \cos(A)_{n-1-m}, & n > 0. \end{cases} \tag{B.20}$$

B.5 *Power*

Now, we take a power $\alpha \neq 1$ of a series:

$$B = A^{\alpha}. \tag{B.21}$$

Assuming $A_0 \neq 0$, we have $B_0 = A_0^{\alpha}$ at leading order. At higher orders, we take the derivative with respect to ρ to find

$$\frac{\mathrm{d}B}{\mathrm{d}\rho} = \alpha A^{\alpha-1} \frac{\mathrm{d}A}{\mathrm{d}\rho}. \tag{B.22}$$

Multiplying both sides against A, we find

$$A\frac{\mathrm{d}B}{\mathrm{d}\rho} - \alpha B \frac{\mathrm{d}A}{\mathrm{d}\rho} = 0. \tag{B.23}$$

Term by term, we have

$$\sum_{j=0}^{n-1} (j+1) A_{n-1-j} B_{j+1} - \alpha (j+1) B_{n-1-j} A_{j+1} = 0. \tag{B.24}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

This gives

$$B_n = \frac{1}{nA_0}\left(n\alpha A_n B_0 + \sum_{j=0}^{n-2}(j+1)\big(-A_{n-1-j}B_{j+1} + \alpha B_{n-1-j}A_{j+1}\big)\right). \tag{B.25}$$

By substituting $j \to n-j-2$, we reorder the right sum to find

$$B_n = \frac{1}{nA_0}\left(n\alpha A_n B_0 + \sum_{j=0}^{n-2}\big(\alpha(n-j-1) - (j+1)\big)B_{j+1}A_{n-j-1}\right). \tag{B.26}$$

B.6 *Composition*

Series composition is an important operation for changing coordinates. Consider that $A = A(x, y, s)$, where $x = R\cos\Theta$ and $y = R\sin\Theta$. We would like to compose this with the functions $B(\rho, \theta)$ and $C(\rho, \theta)$ as $A(B, C, s)$, where B and C replace the Cartesian cross-section coordinates x and y. We note that this is preferable to composing with the polar coordinates R and Θ, as there are valid non-analytic transformations in these coordinates that keep A analytic. For this transformation, we assume that $B_0 = C_0 = 0$, i.e. there is no constant offset. We present the details for numerical Fourier series composition (4.1), but equivalent expressions could be used for other forms.

The main observation we make for series composition is that if we can compute the basis functions

$$\Phi_{mn}(\rho, \theta, s) = R(\rho, \theta, s)^m \mathcal{F}_{2n-m}(\Theta(\rho, \theta, s)), \tag{B.27}$$

then the composition is simply

$$A(B, C, s) = \sum_{m=0}^{\infty}\sum_{n=0}^{m} A_{mn}(s)\Phi_{mn}(\rho, \theta, s), \tag{B.28}$$

where the multiplication in s can be performed in spatial coordinates. So, the majority of the work is to find Φ_{mn}, with the added perk that once the basis is found, further series compositions are faster.

To find the basis, we first notice that

$$\Phi_{00} = 1, \qquad \Phi_{10} = R\cos\Theta = C, \qquad \Phi_{11} = R\sin\Theta = B. \tag{B.29}$$

Then, further functions can be found by using angle-sum identities in the cosine case $(m/2 \leqslant n \leqslant m+1)$

$$\begin{aligned}\Phi_{m+1,n} &= R^{m+1}\cos((2n-m-1)\Theta),\\ &= R^{m+1}\cos((2(n-1)-m)\Theta)\cos(\Theta) - R^{m+1}\sin((2(n-1)-m)\Theta)\sin(\Theta),\\ &= \Phi_{m,n-1}\Phi_{11} - \Phi_{m,m-n+1}\Phi_{10},\end{aligned} \tag{B.30}$$

and the sine case $(0 \leqslant n < m/2)$

$$\begin{aligned}\Phi_{m+1,n} &= R^{m+1}\sin((m+1-2n)\Theta),\\ &= R^{m+1}\sin((m-2n)\Theta)\cos(\Theta) + \cos((m-2n)\Theta)\sin(\Theta),\\ &= \Phi_{mn}\Phi_{11} + \Phi_{m,m-n}\Phi_{10}.\end{aligned} \tag{B.31}$$

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

B.7 *Series inversion*

Consider that we know the (Cartesian) flux coordinates $(B^{(1)}(x, y, s), B^{(2)}(x, y, s)) = (X, Y) = (R\cos\Theta, R\sin\Theta)$ and we want to know how to represent a function $A(\rho, \theta, s)$ in terms of R and Θ. For this, we would use the series composition step presented in the previous section, but we need to obtain ρ and θ in terms of R and Θ, i.e. we need $(C^{(1)}(X, Y, s), C^{(2)}(X, Y, s)) = (x, y) = (\rho\cos\theta, \rho\sin\theta)$, where

$$\begin{pmatrix} C^{(1)}(B^{(1)}, B^{(2)}, s) \\ C^{(2)}(B^{(1)}, B^{(2)}, s) \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}. \tag{B.32}$$

This is a step that is necessary if we have a series represented in direct coordinates, and we want to represent it in indirect (flux) coordinates.

We will solve this equation iteratively, assuming that we are using the real Fourier form (4.1). At leading order (assuming $C_0^{(i)} = B_0^{(i)} = 0$), we have

$$\begin{pmatrix} C_{11}^{(1)} & C_{10}^{(1)} \\ C_{11}^{(2)} & C_{10}^{(2)} \end{pmatrix} \begin{pmatrix} B_{11}^{(1)} & B_{10}^{(1)} \\ B_{11}^{(2)} & B_{10}^{(2)} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{B.33}$$

or

$$\begin{pmatrix} C_{11}^{(1)} & C_{10}^{(1)} \\ C_{11}^{(2)} & C_{10}^{(2)} \end{pmatrix} = \begin{pmatrix} B_{11}^{(1)} & B_{10}^{(1)} \\ B_{11}^{(2)} & B_{10}^{(2)} \end{pmatrix}^{-1}. \tag{B.34}$$

For the next orders, we note that we can compute $C_{<m+1}^{(i)}(B^{(1)}, B^{(2)}, s)$ using the composition formula, where we have chosen C_m so this is identity up to order $m+1$. Substituting this into (B.32), we have

$$\begin{pmatrix} C_{m+1}^{(1)}\left(\rho B_1^{(1)}, \rho B_1^{(2)}, s\right) \\ C_{m+1}^{(2)}\left(\rho B_1^{(1)}, \rho B_1^{(2)}, s\right) \end{pmatrix} = -\begin{pmatrix} \left(C_{<m+1}^{(1)}(B^{(1)}, B^{(2)}, s)\right)_{m+1} \\ \left(C_{<m+1}^{(2)}(B^{(1)}, B^{(2)}, s)\right)_{m+1} \end{pmatrix}. \tag{B.35}$$

If we build a basis Φ_{mn} from $B_1^{(i)}$ (see (B.27) and the following procedure) where

$$\Phi_{mn} = \sum_{k=0}^{\infty}\sum_{\ell=0}^{k} \Phi_{mnkl}\rho^k \mathcal{F}_{k-2\ell}(\theta), \tag{B.36}$$

the update can be expanded in index-notation as

$$\sum_{\ell=0}^{m+1} C_{m+1,\ell}^{(i)}\Phi_{m+1,\ell,m+1,n} = -\left(C_{<m+1}^{(1)}(B^{(1)}, B^{(2)}, s)\right)_{m+1,n}. \tag{B.37}$$

Then, by inverting the matrix $\Psi_{\ell n} = \Phi_{m+1,\ell,m+1,n}$ onto the right-hand side, we have the update for $C_{m+1}^{(i)}$.

REFERENCES

AXLER, S., BOURDON, P. & RAMEY, W. 2001 Harmonic function theory. In *Graduate Texts in Mathematics*, vol. **137**, Springer.

BISHOP, R.L. 1975 There is more than one way to frame a curve. *Am. Math. Mon.* **82** (3), 246–251.

BOYD, J.P. 2001 *Chebyshev and Fourier Spectral Methods*. 2nd edn. Dover Publications.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

BURBY, J.W., DUIGNAN, N. & MEISS, J.D. 2021 Integrability, normal forms, and magnetic axis coordinates. *J. Math. Phys.* **62** (12), 122901.

BURBY, J.W., DUIGNAN, N. & MEISS, J.D. 2023 Minimizing separatrix crossings through isoprominence, *Plasma Phys. Control. Fusion* **65** (4), 045004.

CARDONA, R., DUIGNAN, N. & PERRELLA, D. 2024 Asymmetry of MHD equilibria for generic adapted metrics. *Arch. Ration. Mech. Anal.* **249** (1), 1.

CONSTANTIN, P., DRIVAS, T.D. & GINSBERG, D. 2021*a* Flexibility and rigidity in steady fluid motion. *Commun. Math. Phys.* **385** (1), 521–563.

CONSTANTIN, P., DRIVAS, T.D. & GINSBERG, D. 2021*b* On quasisymmetric plasma equilibria sustained by small force. *J. Plasma Phys.* **87** (1), 905870111.

DUDT, D.W. & KOLEMEN, E. 2020 DESC: a stellarator equilibrium solver. *Phys. Plasmas* **27** (10), 102513.

DUIGNAN, N. & MEISS, J.D. 2021 Normal forms and near-axis expansions for Beltrami magnetic fields. *Phys. Plasmas* **28** (12), 122501.

EVANS, L.C. 2010 Partial differential equations. In *Graduate Studies in Mathematics*. 2nd edn, vol. **19**. American Mathematical Society.

GARREN, D.A. & BOOZER, A.H. 1991 Existence of quasihelically symmetric stellarators. *Phys. Fluids B: Plasma Phys.* **3** (10), 2822–2834.

GIULIANI, A. 2024 Direct stellarator coil design using global optimization: application to a comprehensive exploration of quasi-axisymmetric devices. *J. Plasma Phys.* **90** (3), 905900303.

GRAD, H. 1967 Toroidal containment of a plasma. *Phys. Fluids* **10** (1), 137–154.

HIRSHMAN, S.P. 1983 Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria. *Phys. Fluids* **26** (12), 3553–3568.

HUDSON, S.R., DEWAR, R.L., DENNIS, G., HOLE, M.J., MCGANN, M., VON NESSI, G. & LAZERSON, S. 2012 Computation of multi-region relaxed magnetohydrodynamic equilibria. *Phys. Plasmas* **19** (11), 112502.

JORGE, R., GIULIANI, A. & LOIZU, J. 2024 Simplified and flexible coils for stellarators using single-stage optimization. *Phys. Plasmas* **31** (11), 112501.

JORGE, R. & LANDREMAN, M. 2020 The use of near-axis magnetic fields for stellarator turbulence simulations, *Plasma Phys. Control. Fusion* **63** (1), 014001.

JORGE, R., SENGUPTA, W. & LANDREMAN, M. 2020 Near-axis expansion of stellarator equilibrium at arbitrary order in the distance to the axis, *J. Plasma Phys.* **86** (1), 905860106.

KAPPEL, J., LANDREMAN, M. & MALHOTRA, D. 2024 The magnetic gradient scale length explains why certain plasmas require close external magnetic coils, *Plasma Phys. Control. Fusion* **66** (2), 025018.

KIM, P., JORGE, R. & DORLAND, W. 2021 The on-axis magnetic well and Mercier's criterion for arbitrary stellarator geometries, *J. Plasma Phys.* **87** (2), 905870231.

KNUTH, D.E. 1997 *The Art of Computer Programming*. 3rd edn. Mass, Addison-Wesley.

LANDREMAN, M. 2017 An improved current potential method for fast computation of stellarator coil shapes, *Nucl. Fusion* **57** (4), 046003.

LANDREMAN, M. 2022 Mapping the space of quasisymmetric stellarators using optimized near-axis expansion. *J. Plasma Phys.* **88** (6), 905880616.

LANDREMAN, M. & JORGE, R. 2020 Magnetic well and Mercier stability of stellarators near the magnetic axis. *J. Plasma Phys.* **86** (5), 905860510.

LANDREMAN, M., MEDASANI, B., WECHSUNG, F., GIULIANI, A., JORGE, R. & ZHU, C. 2021 SIMSOPT: a flexible framework for stellarator optimization. *J. Open Source Softw.* **6** (65), 3525.

LANDREMAN, M. & PAUL, E. 2022 Magnetic fields with precise quasisymmetry for plasma confinement. *Phys. Rev. Lett.* **128** (3), 035001.

LANDREMAN, M. & SENGUPTA, W. 2019 Constructing stellarators with quasisymmetry to high order. *J. Plasma Phys.* **85** (6), 815850601.

MATA, K.C., PLUNK, G.G. & JORGE, R. 2022 Direct construction of stellarator-symmetric quasi-isodynamic magnetic configurations, *J. Plasma Phys.* **88**(5), 905880503.

MERCIER, C. 1964 Equilibrium and stability of a toroidal magnetohydrodynamic system in the neighbourhood of a magnetic axis. *Nucl. Fusion* **4** (3), 213–226.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press

RODRÍGUEZ, E. & BHATTACHARJEE, A. 2021 Solving the problem of overdetermination of quasisymmetric equilibrium solutions by near-axis expansions I. Generalized force balance.. *Phys. Plasmas* **28** (1), 012508.

RUTH, M. 2024*a* StellaratorNearAxis.jl. Available at https://github.com/maxeruth/StellaratorNearAxis.jl.

RUTH, M. 2024*b* SymplecticMapTools.jl. Available at: https://github.com/maxeruth/SymplecticMapTools.jl.

RUTH, M. & BINDEL, D. 2024 Finding Birkhoff averages via adaptive filtering. *Chaos: Interdisciplinary J. Nonlinear Sci.* **34** (12), 123109.

SENGUPTA, W., RODRIGUEZ, E., JORGE, R., LANDREMAN, M. & BHATTACHARJEE, A. 2024 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 Plasma confinement in closed magnetic systems. In *Reviews of Plasma Physics* (ed. LEONTOVICH M.A.), pp. 1–247, Springer US.

TAYLOR, M.E. 2011 Partial differential equations I: basic theory. In *Applied Mathematical Sciences*, vol. **115**. Springer.

WECHSUNG, F., LANDREMAN, M., GIULIANI, A., CERFON, A. & STADLER, G. 2022 Precise stellarator quasi-symmetry can be achieved with electromagnetic coils. *Proc. Natl. Acad. Sci. USA* **119** (13), e2202084119.

https://doi.org/10.1017/S0022377825000388 Published online by Cambridge University Press