Geophys. J. Int. (2004) **157,** 537–552 doi: 10.1111/j.1365-246X.2004.02216.x

Geomagnetic secular variation generated by a tangentially geostrophic flow under the frozen-flux assumption—II. Sufficient conditions

A. Chulliat*

Geophysical Fluid Dynamics Institute, Florida State University, Tallahassee 32306, FL, USA

Accepted 2003 November 24. Received 2003 November 17; in original form 2003 March 1

SUMMARY

Two assumptions are often made when computing Earth's core surface flows from geomagnetic data: the frozen-flux assumption and the tangentially geostrophic assumption. These assumptions impose some integral constraints on the geomagnetic field on secular timescales. It has been shown in the first paper of this series that all known integral constraints could actually be deduced from a single set of curvilinear integral constraints on the secular variation along level curves of $\zeta = B_r/\cos\theta$ at the core surface. In the present paper, these particular constraints are further proved to be sufficient conditions for the geomagnetic secular variation to be generated by a tangentially geostrophic flow under the frozen-flux assumption. This result means that no other independent constraint can be found and makes it possible to attempt building geomagnetic models consistent with both the frozen-flux and the tangentially geostrophic assumptions. Also, the complete set of tangentially geostrophic flow solutions of the induction equation under the frozen-flux assumption is exhibited. These solutions are amenable to direct numerical computation.

Key words: core flow, frozen-flux, geomagnetism, geostrophic, secular variation.

1 INTRODUCTION

This paper is the second of a series of papers (initiated by Chulliat & Hulot 2001, and hereafter referred to as Paper I) investigating the magnetohydrodynamic (MHD) properties of the Earth's core surface on secular timescales. The core surface is defined as the region just below the very thin viscous boundary layer under the core–mantle boundary (CMB) where the flow is a horizontal free-stream. It is of particular interest for understanding the geodynamo as it is the only region of the core where some information about the magnetic field can be directly inferred from observation. The field at the core surface is obtained by continuing the field measured at or near the Earth's surface downward through the weakly conducting mantle: the main limitation to this calculation is that the small-scale core field is masked by the crustal field at the Earth's surface as a result of geometrical attenuation (e.g. Gubbins & Roberts 1987). The secular variation (SV) of the core field is provided by historical data, which span approximately 400 yr, from early declination and inclination measurements made aboard ships to recent vector and scalar data systematically collected in observatories and by satellites (e.g. Jackson *et al.* 2000).

Two assumptions greatly simplify the equations relating the magnetic field and the flow at the core surface on secular timescales: the so-called frozen-flux (FF) and tangentially geostrophic (TG) assumptions. Both assumptions rely on order of magnitude estimates of the various terms of the equations. They have been widely used for computing core surface flows from geomagnetic data (e.g. Hulot *et al.* 2002, and more references in Paper I).

The FF assumption consists of neglecting the magnetic field diffusion on secular timescales (Roberts & Scott 1965). Under this assumption, the radial component of the induction equation at the core surface reads

$$\partial_t B_r = -\boldsymbol{\nabla}_H\cdot(\mathbf{u}B_r), \tag{1.1}$$

where $\mathbf{u}$ is the flow, B_r the radial component of the magnetic field and $\boldsymbol{\nabla}_H = \boldsymbol{\nabla} - \mathbf{n}\partial_r$, $\mathbf{n}$ the unit radial outward vector. Eq. (1.1) involves no radial derivative and relates the unknown flow $\mathbf{u}$ to the magnetic field and its secular variation. As only the radial component of the magnetic

*Now at: Laboratoire de Géomagnétisme, IPGP, 4 Place Jussieu, Paris 75252, Cedex 05, France.

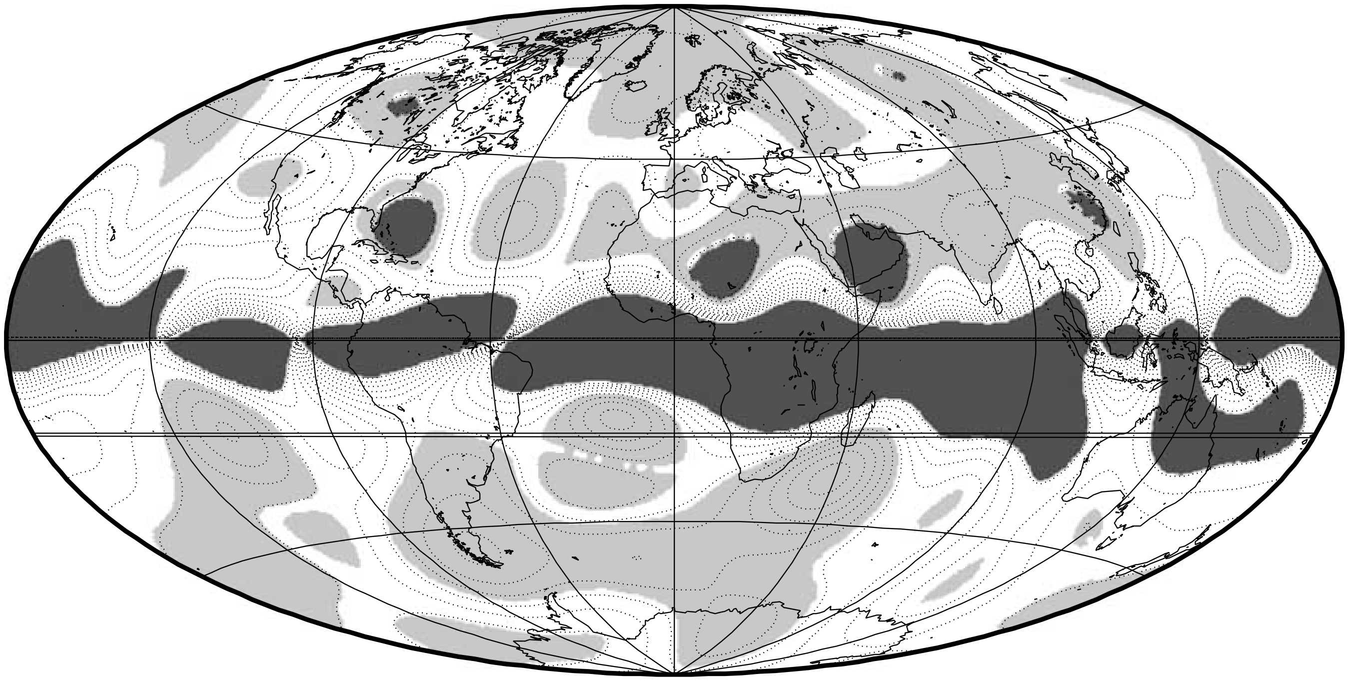

Figure 1. The iso-ζ curves and the geostrophic region in 1980, equal-area Hammer projection, model of Bloxham & Jackson (1992). Isovalues every 2×10^5 nT between -10^6 nT and 10^6 nT. The light grey areas are the ambiguous patches, defined as the set of iso-ζ curves that do not cross the geographic equator. The non-geostrophic region defined by $|\zeta| > |\zeta_{\max}|$ is the dark grey area. It is represented for $\zeta_{\max} = 10^{-3}$ T.

field can be assumed continuous through both viscous and magnetic boundary layers at the top of the core (Jault & Le Mouël 1991), no other component is generally considered.

The TG assumption consists of neglecting the Lorentz force in the upper layers of the core, where the toroidal field is weak (Le Mouël 1984). Under this assumption, the Coriolis force is dominant and the magnetogeostrophic balance generally assumed to hold throughout the core becomes

$$2\Omega\rho\mathbf{n} \times (\mathbf{u}\cos\theta) = -\nabla_H p \tag{1.2}$$

at the core surface, where Ω is the Earth's rotation rate, ρ the core density, θ the colatitude and p the dynamic pressure. Taking the curl of eq. (1.2) gives an equivalent formulation of the TG assumption:

$$\nabla_H\cdot(\mathbf{u}\cos\theta) = 0. \tag{1.3}$$

This assumption is useful for computing core surface flows because it reduces the otherwise very large number of solutions of the FF induction eq. (1.1). Yet, it cannot be used on the whole core surface because the Coriolis force vanishes on the geographic equator: it must be confined to a so-called geostrophic region (Backus & Le Mouël 1986). It has been argued in Paper I that the geostrophic region $\mathcal{G}$ could be defined consistently as the set of points where $|\zeta| \le \zeta_{\max}$, where $\zeta = B_r/\cos\theta$ and $\zeta_{\max} = 10^{-3}$ nT. As an illustration, Fig. 1 shows the geostrophic region in 1980, computed from the geomagnetic model of Bloxham & Jackson (1992).

Backus (1968) solved eq. (1.1) for continuous flows. He showed that: (a) the FF assumption imposes some constraints on the magnetic field at the CMB; and (b) if these constraints are satisfied, then eq. (1.1) has a non-empty set of continuous solutions that he exhibited and termed eligible flows. These constraints are thus necessary and sufficient conditions for the SV to be generated by a continuous flow under the FF assumption. They have been widely used for testing the FF assumption against geomagnetic data: most tests concluded with positive results (e.g. O'Brien *et al.* 1997, and more references in Paper I).

Unlike the FF assumption alone, the combined FF+TG assumption has never been properly tested against geomagnetic data. The reason is that theoretical tools for such tests were not available. As for the FF assumption alone, a consistent test of the FF+TG assumption would be to construct a field model constrained by the necessary and sufficient conditions imposed by this assumption and compare its misfit with the data error bars. However, while several sets of necessary conditions have been found, including the one presented in Paper I, none of these sets has been proven sufficient. [It is worth noting that the TG assumption alone cannot be tested, as eq. (1.2) does not involve the observable magnetic field.]

In the present paper, Backus's results in the lone FF case will be extended to the FF+TG case, providing the missing theoretical tools for testing the FF+TG assumption. A first step was made in Paper I, where a new set of necessary conditions was obtained and it was proven that several sets of necessary conditions published to date could actually be inferred from this new set. Here this set of necessary conditions will be proven to be sufficient for the SV to be generated by a TG flow under the FF assumption within a given geostrophic region $\mathcal{G}$. In addition, the complete set of eligible TG flows, i.e. continuously differentiable solutions of eq. (1.1) satisfying eq. (1.3) within $\mathcal{G}$, will be exhibited.

There are several motivations for testing the FF+TG assumption. Practically, tests would help us assessing the validity of core surface flow computations. Moreover, in the case of positive tests, constrained core field models would be as appropriate, if not more appropriate, than unconstrained ones for core studies. A new family of core field models with unknown properties could then be constructed. These models could be used, for example, to compute core surface flows by local methods such as the one developed by Chulliat & Hulot (2000). This first attempt was encouraging as the large-scale pattern of the locally computed pressure field was very similar to that obtained by the standard spectral method. However, the pressure field had spurious small-scale features that were attributed to the fact that the starting field model was unconstrained.

On the theoretical side, tests would provide us with some observational information on the dynamic regime at the top of the core, which could be compared to numerical geodynamo simulations. For example, Roberts & Glatzmaier (2000) found that one necessary condition imposed by the FF assumption, the constraint on the unsigned magnetic flux through patches delimited by $B_r = 0$ curves, was satisfied to within a few per cent in their simulation. Using a different geodynamo model, Rau *et al.* (2000) found that diffusion contributed to approximately 25 per cent of the energy of the secular variation and that less than 10 per cent of the flow at the core surface was ageostrophic.

The main limitation to such tests, and core surface flow computations in general, comes from the limit imposed by the crustal field on recovering core field scales smaller than the degree 13 of the spherical harmonics expansion. Yet, even this limited range of scales has not been fully exploited as existing tests have been performed with magnetic field models up to degree 13 and SV models up to degree 8 only. Moreover, the ØRSTED and CHAMP satellites have been mapping the magnetic field continuously since 2000 with an unprecedented precision. It is now possible to model the SV up to degree 13 , i.e. with the same spatial resolution as the magnetic field (Olsen 2002; Langlais *et al.* 2003). It seems reasonable to expect even greater spatial resolution for the SV in the near future as, unlike the stationary part of the field, it is not masked by the crustal field.

The present paper is organized as follows. In Section 2, basic assumptions, notations, definitions and useful formulae are presented. In Section 3, the set of constraints derived in Paper I is recalled and some useful consequences of these constraints are established. In Section 4, the existence of eligible TG flows is assumed and the general form of these flows is obtained. In Section 5, the set of necessary conditions of Paper I, assumed to be satisfied by the geomagnetic field, is proven to be sufficient by exhibiting one eligible TG flow; finally the complete set of eligible TG flows is exhibited.

2 PRELIMINARIES

The purpose of this section is to do some preliminary work. First, a few basic assumptions are made in order to ensure sufficient mathematical regularity of the magnetic field at the core surface. Next, several notations and definitions are introduced: most of them involve iso-ζ curves, which will be abundantly used in what follows. Finally, three lemmas, demonstrated in appendices, provide a general derivation formula and some asymptotic expansions that will simplify the mathematics in the remainder of the paper.

2.1 Basic assumptions

It is assumed that the core surface is a perfect sphere and that the geomagnetic field **B** meets the following conditions at all times at the core surface:

(i) the function B_r is four times continuously differentiable in θ and ϕ (longitude);
(ii) all iso-ζ curves are closed and of finite length;
(iii) there is at least one magnetic equator, i.e. a curve $B_r = 0$, crossing the geographic equator at some nodes;
(iv) the geomagnetic equator is nowhere tangent to the geographic equator;
(v) the critical points (where $\nabla_H \zeta = \mathbf{0}$) of the function ζ are not degenerate (i.e. the discriminant of the tensor of second derivatives of ζ is non-zero).

These are the same assumptions as in Paper I, except the first one, which is stronger (B_r was only assumed twice continuously differentiable). As argued in Paper I, they are not restrictive in a geophysical context.

2.2 Notations and definitions

As in Paper I: ∂S_{ζ_0} denotes an iso-ζ curve of equation $\zeta = \zeta_0$ encircling a surface ∂S_{ζ_0} (Fig. 2a); $\Gamma^{i,j}_{\zeta_0}$ denotes a portion of an iso-ζ curve of equation $\zeta = \zeta_0$ joining two nodes N_i and N_j (Fig. 2b); and, $\Sigma^{i,j}_{\infty,\zeta_0}$ denotes a surface delimited by a curve $\Gamma^{i,j}_{\zeta_0}$ and the shortest of the two segments $\Gamma^{i,j}_{\infty}$ of geographic equator (where $\zeta = \pm\infty$) joining N_i and N_j (Fig. 2b). Also, $\Gamma^{M_1,M_2}_{\zeta_0}$ denotes a portion of an iso-ζ curve of equation $\zeta = \zeta_0$ joining two points M_1 and M_2. [Note that Figs 2(a) and (b) are identical to figs 2(a) and (b) of Paper I.]

Because of the fifth basic assumption on **B**, there exists two kinds of critical points of the function ζ: saddle points, where the discriminant of the tensor of second derivatives of ζ is negative, and extrema, where the discriminant is positive. An iso-ζ curve containing one or more critical points is called critical.

The following definitions have been laid down by Backus & Le Mouël (1986): the visible belt is the set of all iso-ζ curves intersecting the geographic equator within the geostrophic region; an ambiguous patch is a maximal connected set of closed iso-ζ curves that never intersect

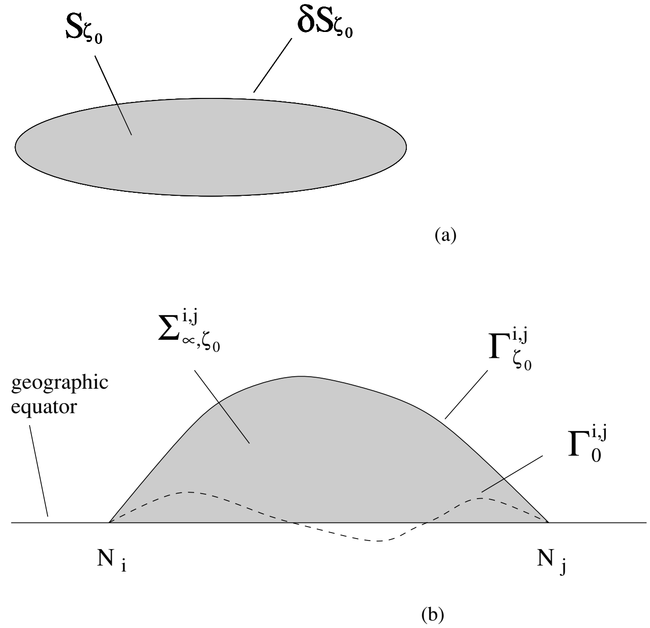

Figure 2. Generic iso-ζ curves and patches.

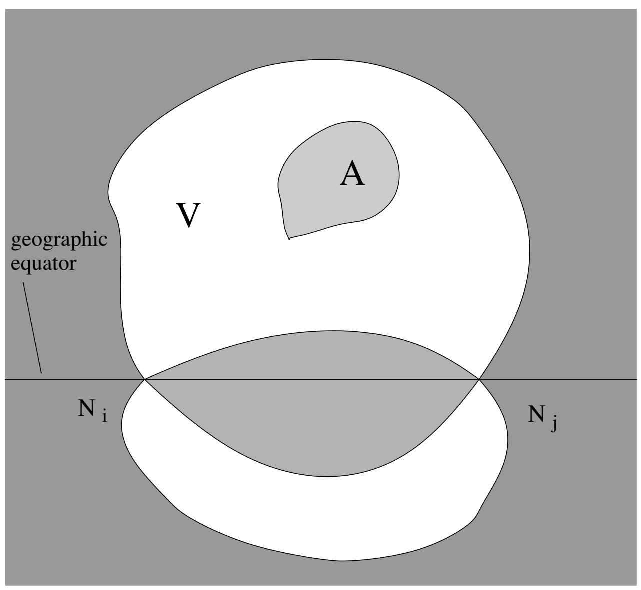

Figure 3. Generic visible domain. The visible domain $\mathcal{V}$ is depicted in white. The non-geostrophic belt (defined by $|\zeta| \geq \zeta_{\max}$) is the dark grey area containing the geographic equator. An ambiguous patch $\mathcal{A}$ is depicted in light grey.

the geographic equator. The meaning of the terms visible and ambiguous will be recalled in Section 5.2. Note that each ambiguous patch is delimited by a critical curve, contains at least one extremum of ζ and is simply connected.

For a given geostrophic region $\mathcal{G}$, a maximal connected set of iso-ζ curves intersecting the geographic equator within $\mathcal{G}$ is called a visible domain (Fig. 3). The visible belt within $\mathcal{G}$ is thus uniquely divided into one or more visible domains. For example, the visible belt in 1980 is made of only one visible domain within the geostrophic region defined by $|\zeta| \leq 10^{-3}$ nT (see Fig. 1).

The ambiguous patches may also be cut into smaller pieces. A maximal connected set of iso-ζ curves delimited by one or two critical curves and containing two, and only two, critical points (either two saddle points on its boundaries, or one saddle point on its boundary and one extremum) is called an ambiguous domain (Fig. 4). Thus, each ambiguous patch can be uniquely divided into one or more ambiguous domains. Note that an ambiguous domain is simply connected if, and only if, it contains an extremum. For example, the biggest ambiguous patch located in the south Atlantic region in 1980 is made of five ambiguous domains, three of them containing an extremum (see Fig. 1). For each ambiguous domain $\mathcal{A}$, there exists a curve $\Gamma_{\mathcal{A}}$ joining both critical points $C^{(1)}_{\mathcal{A}}$ and $C^{(2)}_{\mathcal{A}}$ of $\mathcal{A}$, and crossing each iso-ζ curve of $\mathcal{A}$ exactly once at a point $K_{\mathcal{A}}(\zeta)$ (Fig. 5). $\zeta^{(1)}_{\mathcal{A}}$ and $\zeta^{(2)}_{\mathcal{A}}$ denote, respectively, the values of ζ at $C^{(1)}_{\mathcal{A}}$ and $C^{(2)}_{\mathcal{A}}$.

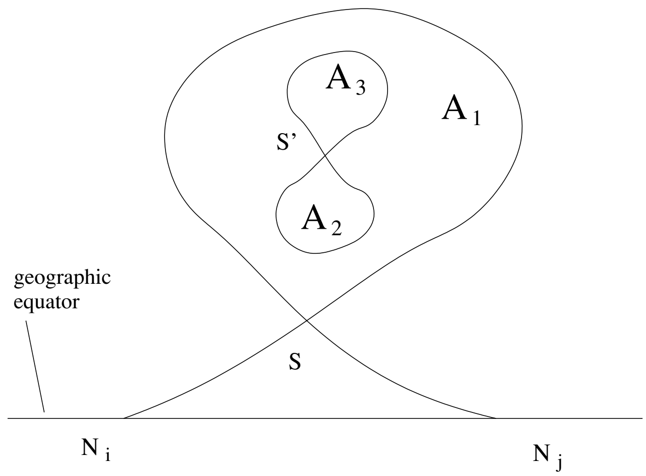

Figure 4. Generic ambiguous domains. The ambiguous domain $\mathcal{A}_1$ contains two saddles points (S and S'), while $\mathcal{A}_2$ and $\mathcal{A}_3$ both contain one saddle point (the same S') and one extremum (not represented).

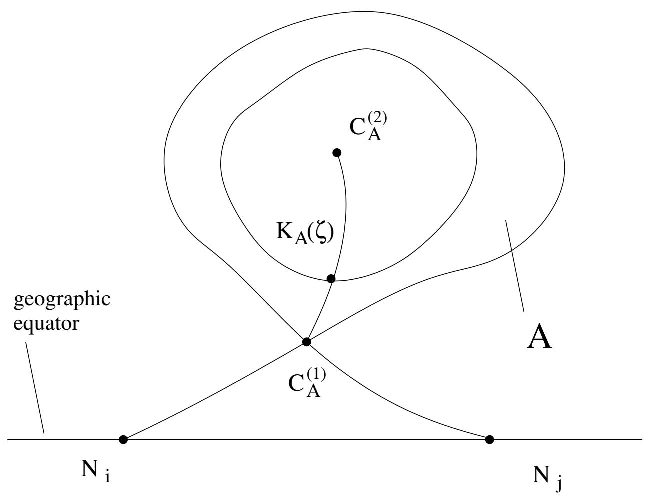

Figure 5. Notations relative to an ambiguous domain. The domain represented contains one saddle point and one extremum. Similar notations apply if the domain contains two saddle points.

The unit vectors normal and tangent to iso-ζ curves are, respectively,

$$\boldsymbol{\pi} \equiv \frac{\boldsymbol{\nabla}_H \zeta}{\|\boldsymbol{\nabla}_H \zeta\|} = \frac{1}{c\|\boldsymbol{\nabla}_H \zeta\|}\left(\partial_\theta \zeta\, \mathbf{e}_\theta + \frac{\partial_\phi \zeta}{\sin\theta}\mathbf{e}_\phi\right) \tag{2.1}$$

and

$$\boldsymbol{\tau} \equiv \mathbf{n} \times \frac{\boldsymbol{\nabla}_H \zeta}{\|\boldsymbol{\nabla}_H \zeta\|} = \frac{1}{c\|\boldsymbol{\nabla}_H \zeta\|}\left(-\frac{\partial_\phi \zeta}{\sin\theta}\mathbf{e}_\theta + \partial_\theta \zeta\, \mathbf{e}_\phi\right), \tag{2.2}$$

where $\mathbf{e}_\theta$ denotes the unit vector associated with the colatitude θ and $\mathbf{e}_\phi$ is the unit vector associated with the longitude ϕ. The vectors $\boldsymbol{\pi}$ and $\boldsymbol{\tau}$ are defined at any point except for poles and critical points. Also, the following 2×2 matrices defined everywhere except for poles introduce are introduced:

$$\mathsf{Z} \equiv \frac{1}{c^2}\begin{pmatrix} \partial_{\theta\theta}\zeta & \dfrac{\partial_{\theta\phi}\zeta}{\sin\theta} \\ \dfrac{\partial_{\theta\phi}\zeta}{\sin\theta} & \dfrac{\partial_{\phi\phi}\zeta}{\sin^2\theta} \end{pmatrix}, \tag{2.3}$$

$$\mathsf{X} \equiv \frac{1}{c^3}\begin{pmatrix} \partial_{\theta\theta\theta}\zeta & \dfrac{\partial_{\theta\theta\phi}\zeta}{\sin\theta} \\ \dfrac{\partial_{\theta\theta\phi}\zeta}{\sin\theta} & \dfrac{\partial_{\theta\phi\phi}\zeta}{\sin^2\theta} \end{pmatrix}, \tag{2.4}$$

$$\mathsf{Y} \equiv \frac{1}{c^3}\begin{pmatrix} \dfrac{\partial_{\theta\theta\phi}\zeta}{\sin\theta} & \dfrac{\partial_{\theta\phi\phi}\zeta}{\sin^2\theta} \\ \dfrac{\partial_{\theta\phi\phi}\zeta}{\sin^2\theta} & \dfrac{\partial_{\phi\phi\phi}\zeta}{\sin^3\theta} \end{pmatrix}, \tag{2.5}$$

$$\mathbf{W} \equiv \frac{1}{c^2}\begin{pmatrix} \partial_{\theta\theta} B_r & \dfrac{\partial_{\theta\phi} B_r}{\sin\theta} \\ \dfrac{\partial_{\theta\phi} B_r}{\sin\theta} & \dfrac{\partial_{\phi\phi} B_r}{\sin^2\theta} \end{pmatrix}, \tag{2.6}$$

$$\partial_t \mathbf{W} \equiv \frac{1}{c^2}\begin{pmatrix} \partial_{\theta\theta}\partial_t B_r & \dfrac{\partial_{\theta\phi}\partial_t B_r}{\sin\theta} \\ \dfrac{\partial_{\theta\phi}\partial_t B_r}{\sin\theta} & \dfrac{\partial_{\phi\phi}\partial_t B_r}{\sin^2\theta} \end{pmatrix}. \tag{2.7}$$

Finally, for all 2 × 2 matrices $\mathbf{A}$, vector fields $\mathbf{V}$ and scalar fields f: $\mathbf{I}$ denotes the 2 × 2 identity matrix, $Tr(\mathbf{A})$ the trace of $\mathbf{A}$, $\mathbf{V}^T$ the transpose of $\mathbf{V}$, V_θ the orthoradial coordinate of $\mathbf{V}$, V_ϕ the azimuthal coordinate of $\mathbf{V}$, $\mathbf{A}\,\mathbf{V}$ the product of $\mathbf{A}$ and the 2 × 1 matrix $\binom{V_\theta}{V_\phi}$, $\mathbf{V}_M$ the value of $\mathbf{V}$ at a point M and f_M the value of f at a point M.

2.3 Useful formulae

A proof of Lemma 1 is found in Appendix A. Proofs of Lemmas 2 and 3 are found in Appendix B.

2.3.1 Lemma 1

Consider a scalar field f, defined on the core surface and continuously differentiable at any non-critical point of ζ. Consider $\Gamma_\zeta^{M_1,M_2}$, a portion of an iso-ζ curve joining two points M_1 and M_2. The derivative with respect to ζ of the curvilinear integral of f along $\Gamma_\zeta^{M_1,M_2}$ may be expressed as

$$\frac{d}{d\zeta}\left(\int_{\Gamma_\zeta^{M_1,M_2}} f\boldsymbol{\tau}\cdot\mathbf{dl}\right) = \int_{\Gamma_\zeta^{M_1,M_2}} \nabla_H\cdot(f\boldsymbol{\pi})\frac{\boldsymbol{\tau}\cdot\mathbf{dl}}{\|\nabla_H\zeta\|}. \tag{2.8}$$

2.3.2 Lemma 2

Consider a node N. At any point $N + \boldsymbol{\delta}\,\mathbf{r}$ near N, the asymptotic expansions of the quantities $\partial_t B_r\|\nabla_H\zeta\|^{-1}$ and $\nabla_H\cdot(\partial_t B_r\|\nabla_H\zeta\|^{-1}\,\boldsymbol{\pi})$ are given by

$$\frac{\partial_t B_r}{\|\nabla_H\zeta\|} = -\boldsymbol{\pi}\cdot(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r})\cos\theta + O(\|\boldsymbol{\delta}\mathbf{r}\|^3) \tag{2.9}$$

and

$$\nabla_H\cdot\left(\frac{\partial_t B_r}{\|\nabla_H\zeta\|}\boldsymbol{\pi}\right) = \|\nabla_H\zeta\|(\boldsymbol{\tau}\cdot\nabla_H)\left[\frac{\boldsymbol{\tau}\cdot(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r})\cos\theta}{\|\nabla_H\zeta\|}\right] - \nabla_H\cdot[(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r})\cos\theta] + O(\|\boldsymbol{\delta}\mathbf{r}\|^2), \tag{2.10}$$

where

$$\mathbf{U}_N \equiv -c\left(\frac{\partial_t B_r}{\partial_\phi B_r}\right)_N \mathbf{e}_\phi \tag{2.11}$$

and $\mathbf{T}_N$ is a 2 × 2 matrix defined by

$$\mathbf{T}_N \equiv -\left\{\begin{matrix} 0 & 0 \\ \left[\partial_\theta\left(\dfrac{\partial_t B_r}{\partial_\phi B_r}\right)\right]_N & \left[\partial_\phi\left(\dfrac{\partial_t B_r}{\partial_\phi B_r}\right)\right]_N \end{matrix}\right\}. \tag{2.12}$$

2.3.3 Lemma 3

Consider a critical point C such that $(\partial_t B_r)_C = 0$. At any point $C + \boldsymbol{\delta}\mathbf{r}$ near C, the asymptotic expansions of the quantities $\partial_t B_r\|\nabla_H\zeta\|^{-1}$ and $\nabla_H\cdot(\partial_t B_r\|\nabla_H\zeta\|^{-1}\,\boldsymbol{\pi})$ are given by

$$\frac{\partial_t B_r}{\|\nabla_H\zeta\|} = -\boldsymbol{\pi}\cdot(\mathbf{U}_C + \mathbf{T}_C\boldsymbol{\delta}\mathbf{r})\cos\theta_C + O(\|\boldsymbol{\delta}\mathbf{r}\|^2) \tag{2.13}$$

and

$$\nabla_H\cdot\left[\frac{\partial_t B_r}{\|\nabla_H\zeta\|}\boldsymbol{\pi}\right] = \|\nabla_H\zeta\|(\boldsymbol{\tau}\cdot\nabla_H)\left[\frac{\boldsymbol{\tau}\cdot(\mathbf{U}_C + \mathbf{T}_C\boldsymbol{\delta}\mathbf{r})\cos\theta_C}{\|\nabla_H\zeta\|}\right] - \nabla_H\cdot[(\mathbf{U}_C + \mathbf{T}_C\boldsymbol{\delta}\mathbf{r})\cos\theta_C] + O(\|\boldsymbol{\delta}\mathbf{r}\|), \tag{2.14}$$

where

$$\mathbf{U}_C \equiv -\frac{1}{\cos\theta_C}\mathbf{Z}_C^{-1}(\nabla_H\partial_t B_r)_C \tag{2.15}$$

and $\mathbf{T}_C$ is a 2 × 2 matrix defined by

$$\mathbf{T}_C \equiv -\frac{1}{2}\mathbf{Z}_C^{-1}\left[(\mathbf{e}_\theta\cdot\mathbf{U}_C)\mathbf{X}_C + (\mathbf{e}_\phi\cdot\mathbf{U}_C)\mathbf{Y}_C + \partial_t\mathbf{W}_C\right] + \frac{1}{c}\cot\theta_C(\mathbf{e}_\phi\cdot\mathbf{U}_C)\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}. \tag{2.16}$$

3 NECESSARY CONDITIONS FOR THE EXISTENCE OF ELIGIBLE TG FLOWS

The constraints on the SV imposed by the combined FF+TG assumption are necessary conditions for the existence of eligible TG flows. It was shown in Paper I that several sets of such constraints could actually be inferred from one single set. The present section aims at recalling this particular set, which will subsequently be referred to as the basic set of constraints, and establishing some consequences (not mentioned in Paper I) of these constraints, which will prove useful in Sections 4 and 5.

3.1 Basic set of constraints

The basic set of constraints on the SV obtained in Paper I is:

$$\oint_{\partial S_{\zeta_0}} \frac{\partial_t B_r}{\|\nabla_H \zeta\|} \boldsymbol{\tau} \cdot \mathbf{dl} = 0, \tag{3.1}$$

$$\int_{\Gamma^{i,j}_{\zeta_0}} \frac{\partial_t B_r}{\|\nabla_H \zeta\|} \boldsymbol{\tau} \cdot \mathbf{dl}_\infty = \eta_{i,j} \big(q_{(j)} - q_{(i)}\big), \tag{3.2}$$

for all non-critical curves ∂S_{ζ_0} and $\Gamma^{i,j}_{\zeta_0}$ within $\mathcal{G}$. Here $\mathbf{dl}_\infty$ is oriented so that $\mathbf{dl}_\infty \times \mathbf{n}$ on $\partial \Sigma^{i,j}_{\infty,\zeta_0}$ points out of $\Sigma^{i,j}_{\infty,\zeta_0}$, $\eta_{i,j} = +1$ if $\mathbf{dl}_\infty$ on $\Gamma^{i,j}_{\zeta_0}$ is oriented from N_i to N_j, -1 otherwise, and $q_{(i)}$ is the renormalized pressure field $q = p/2\Omega\rho$ at the node N_i. The curvilinear integrals in eqs (3.1) and (3.2) were interpreted in Paper I as SV flux densities on the iso-ζ curves ∂S_{ζ_0} and $\Gamma^{i,j}_{\zeta_0}$.

3.2 Consequences

In the present subsection it is assumed that the SV satisfies the basic set of constraints eqs (3.1) and (3.2), with an arbitrary set $\{q_{(i)}\}$, for all non-critical curves ∂S_{ζ_0} and $\Gamma^{i,j}_{\zeta_0}$ within a geostrophic region $\mathcal{G}$ of arbitrary extent (defined by $\zeta_{\max}$). (No assumption is made regarding the existence of eligible TG flows.) This assumption has three useful consequences.

First, deriving eqs (3.1) and (3.2) with respect to ζ and using eq. (2.8) of Lemma 1 lead to another set of curvilinear integral constraints on the SV:

$$\oint_{\partial S_{\zeta_0}} \nabla_H \cdot \left(\frac{\partial_t B_r}{\|\nabla_H \zeta\|} \boldsymbol{\pi} \right) \frac{\boldsymbol{\tau} \cdot \mathbf{dl}}{\|\nabla_H \zeta\|} = 0, \tag{3.3}$$

$$\int_{\Gamma^{i,j}_{\zeta_0}} \nabla_H \cdot \left(\frac{\partial_t B_r}{\|\nabla_H \zeta\|} \boldsymbol{\pi} \right) \frac{\boldsymbol{\tau} \cdot \mathbf{dl}}{\|\nabla_H \zeta\|} = 0, \tag{3.4}$$

for all non-critical curves ∂S_{ζ_0} and $\Gamma^{i,j}_{\zeta_0}$ within $\mathcal{G}$. The convergence at nodes of the integral in eq. (3.4) is ensured by formula eq. (2.10) of Lemma 2.

Secondly, the SV satisfies the following two constraints for all critical points C:

$$(\partial_t B_r)_C = 0 \tag{3.5}$$

and, at any point $C + \boldsymbol{\delta}\mathbf{r}$ near C,

$$\nabla_H \cdot [(\mathbf{U}_C + \mathbf{T}_C \boldsymbol{\delta}\mathbf{r}) \cos\theta_C] = O(\|\boldsymbol{\delta}\mathbf{r}\|), \tag{3.6}$$

where $\mathbf{U}_C$ and $\mathbf{T}_C$ are given by eqs (2.15) and (2.16), respectively. Proofs of eqs (3.5) and (3.6) are found in Appendix C. [Note that, if an eligible TG flow exists, eq. (3.5) can also be straightforwardly deduced from the fact that the quantity $\partial_t B_r \|\nabla_H \zeta\|^{-1}$ is bounded everywhere at the core surface.]

Thirdly, as the SV vanishes at all critical points, the constraints of eqs (3.1) and (3.2) also hold for all critical curves ∂S_{ζ_0} and $\Gamma^{i,j}_{\zeta_0}$ within $\mathcal{G}$.

4 GENERAL FORM OF ELIGIBLE TG FLOWS

In the present section it is assumed that eligible TG flows exist within a geostrophic region $\mathcal{G}$ of arbitrary extent (defined by $\zeta_{\max}$). Let $\mathbf{u}$ be one eligible TG flow, p one of the pressure fields associated with it [i.e. satisfying eq. (1.2)] and $q = p/2\Omega\rho$ the corresponding renormalized pressure field (that will be considered instead of p hereafter). As $d\zeta/dt = 0$ [combine eqs (1.1) and (1.3)], the iso-ζ curves are material within $\mathcal{G}$; therefore, $\mathcal{G}$ is material and can be treated as a closed region. In addition, the basic set of constraints of eqs (3.1) and (3.2) is satisfied by the SV, as well as its consequences of eqs (3.3), (3.4), (3.5) and (3.6). The goal of the present section is to obtain the general form of $\mathbf{u}$ everywhere within $\mathcal{G}$. This form will be used in Section 5 to construct a solution.

4.1 Some useful expressions

Except for nodes (where $\cos\theta = 0$) and critical points (where $\|\nabla_H \zeta\| = 0$), the flow $\mathbf{u}$ may be expressed as

$$\mathbf{u} = \frac{1}{\cos\theta} \left(\|\nabla_H \zeta\| \frac{\partial q}{\partial \zeta} \boldsymbol{\tau} - \frac{\partial_t B_r}{\|\nabla_H \zeta\|} \boldsymbol{\pi} \right). \tag{4.1}$$

The component $\boldsymbol{\pi}\cdot\mathbf{u}$ is obtained by splitting $\mathbf{u}B_r$ into $(\mathbf{u}\cos\theta)\,\zeta$ in eq. (1.1) and using eq. (1.3); the component $\boldsymbol{\tau}\cdot\mathbf{u}$ is obtained by taking $\boldsymbol{\pi}\cdot$ (eq. 1.2).

Combining $\boldsymbol{\tau}\cdot$ (eq. 1.2) and $\boldsymbol{\pi}\cdot$ (eq. 4.1) gives the horizontal gradient of the (renormalized) pressure,

$$\boldsymbol{\tau}\cdot\boldsymbol{\nabla}_H q = \frac{\partial_t B_r}{\|\boldsymbol{\nabla}_H\zeta\|}. \tag{4.2}$$

Integrating eq. (4.2) along any segment iso-ζ curve $\Gamma_{\zeta_0}^{M_1,M_2}$ joining two points M_1 and M_2 yields the pressure difference between those two points,

$$q_{M_2} - q_{M_1} = \int_{\Gamma_{\zeta_0}^{M_1,M_2}} \frac{\partial_t B_r}{\|\boldsymbol{\nabla}_H\zeta\|}\boldsymbol{\tau}\cdot\mathbf{dl}, \tag{4.3}$$

where $\mathbf{dl}$ is oriented from M_1 to M_2. Note that the constraints of eqs (3.1) and (3.2) are expressions of eq. (4.3) applied to non-critical curves ∂S_{ζ_0} and $\Gamma_{\zeta_0}^{i,j}$ within $\mathcal{G}$.

Deriving eq. (4.3) with respect to ζ and applying eq. (2.8) of Lemma 1 with $f = \partial_t B_r\|\boldsymbol{\nabla}_H\zeta\|^{-1}$ further yields

$$\left(\frac{\partial q}{\partial\zeta}\right)_{M_2} - \left(\frac{\partial q}{\partial\zeta}\right)_{M_1} = \int_{\Gamma_{\zeta_0}^{M_1,M_2}} \boldsymbol{\nabla}_H\cdot\left(\frac{\partial_t B_r}{\|\boldsymbol{\nabla}_H\zeta\|}\boldsymbol{\pi}\right)\frac{\boldsymbol{\tau}\cdot\mathbf{dl}}{\|\boldsymbol{\nabla}_H\zeta\|}, \tag{4.4}$$

with the same convention for $\mathbf{dl}$, provided there is no node or critical point on $\Gamma_{\zeta_0}^{M_1,M_2}$.

4.2 Form of eligible TG flows at and near nodes

Consider a node N. Under the combined FF+TG assumption, N is material because it is the intersection of an infinite number of material curves. At any point $N + \boldsymbol{\delta}\mathbf{r}$ near N, taking eq. (2.9) of Lemma 2 into $\boldsymbol{\pi}\cdot$ (eq. 4.1) yields

$$\boldsymbol{\pi}\cdot\mathbf{u} = \boldsymbol{\pi}\cdot(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r}) + O(\|\boldsymbol{\delta}\mathbf{r}\|^2), \tag{4.5}$$

where $\mathbf{U}_N$ and $\mathbf{T}_N$ are defined by eq. (2.11) and eq. (2.12), respectively. Also, taking eq. (2.10) of Lemma 2 into eq. (4.4) (with $M_1 = N$ and $M_2 = N + \boldsymbol{\delta}\mathbf{r}$), next into $\boldsymbol{\tau}\cdot$ (eq. 4.1) and using

$$\boldsymbol{\nabla}_H\cdot[(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r})\cos\theta] = Tr(\mathbf{T}_N)\cos\theta + O(\|\boldsymbol{\delta}\mathbf{r}\|^2), \tag{4.6}$$

yields

$$\boldsymbol{\tau}\cdot\mathbf{u} = \boldsymbol{\tau}\cdot(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r}) - Tr(\mathbf{T}_N)\boldsymbol{\tau}\cdot\boldsymbol{\delta}\mathbf{r} + O(\|\boldsymbol{\delta}\mathbf{r}\|^2). \tag{4.7}$$

The first-order expansion of $\mathbf{u}$ near N is obtained by combining eqs (4.5) and (4.7) and noticing that $\boldsymbol{\pi}\cdot\boldsymbol{\delta}\mathbf{r} = O(\|\boldsymbol{\delta}\mathbf{r}\|^2)$:

$$\mathbf{u} = \mathbf{U}_N + [\mathbf{T}_N - Tr(\mathbf{T}_N)\mathbf{I}]\boldsymbol{\delta}\mathbf{r} + O(\|\boldsymbol{\delta}\mathbf{r}\|^2). \tag{4.8}$$

4.3 Form of eligible TG flows at and near critical points

Consider a critical point C. Under the combined FF+TG assumption, C is material because it is either the intersection of two material curves (if C is a saddle point) or an infinitely small material curve (if C is an extremum). As the constraint of eq. (3.5) is satisfied, Lemma 3 applies. At any point $C + \boldsymbol{\delta}\mathbf{r}$ near C, taking eq. (2.13) of Lemma 3 into $\boldsymbol{\pi}\cdot$ (eq. 4.1) yields

$$\boldsymbol{\pi}\cdot\mathbf{u} = \boldsymbol{\pi}\cdot(\mathbf{U}_C + \mathbf{T}_C\boldsymbol{\delta}\mathbf{r}) + O(\|\boldsymbol{\delta}\mathbf{r}\|^2), \tag{4.9}$$

where $\mathbf{U}_C$ and $\mathbf{T}_C$ are defined by eqs (2.15) and (2.16), respectively. Also, for any two points $M_1 = C + \boldsymbol{\delta}\mathbf{r}_1$ and $M_2 = C + \boldsymbol{\delta}\mathbf{r}_2$ near C on the same iso-ζ curve Γ_ζ, taking eq. (2.14) of Lemma 3 into eq. (4.4), next into $\boldsymbol{\tau}\cdot$ (eq. 4.1) and using the constraint of eq. (3.6) yields

$$\left[\frac{\boldsymbol{\tau}\cdot\mathbf{u}\cos\theta}{\|\boldsymbol{\nabla}_H\zeta\|}\right]_{M_1}^{M_2} = \left[\frac{\boldsymbol{\tau}\cdot(\mathbf{U}_C + \mathbf{T}_C\boldsymbol{\delta}\mathbf{r})\cos\theta_C}{\|\boldsymbol{\nabla}_H\zeta\|}\right]_{M_1}^{M_2} + O\big(\|\boldsymbol{\delta}\mathbf{r}_1\|\big) + O\big(\|\boldsymbol{\delta}\mathbf{r}_2\|\big). \tag{4.10}$$

Therefore, there exists a function $\alpha_C(\Gamma_\zeta)$ such that

$$\boldsymbol{\tau}\cdot\mathbf{u} = \boldsymbol{\tau}\cdot(\mathbf{U}_C + \mathbf{T}_C\boldsymbol{\delta}\mathbf{r}) + \alpha_C(\Gamma_\zeta)\|\boldsymbol{\nabla}_H\zeta\| + O(\|\boldsymbol{\delta}\mathbf{r}\|^2) \tag{4.11}$$

at any point $C + \boldsymbol{\delta}\mathbf{r}$ near C on an iso-ζ curve Γ_ζ. As the flow $\mathbf{u}$ is continuously differentiable, the function $\alpha_C(\Gamma_\zeta)$ is continuous at C; let $\alpha_C = \lim_{\zeta\to\zeta_C}\alpha_C(\Gamma_\zeta)$. Combining eqs (4.9) and (4.11) and expanding $\boldsymbol{\nabla}_H\zeta$ to first order, yields the first-order expansion of $\mathbf{u}$ near C:

$$\mathbf{u} = \mathbf{u}_C + \mathbf{T}_C\boldsymbol{\delta}\mathbf{r} + \alpha_C\|\mathbf{Z}_C\boldsymbol{\delta}\mathbf{r}\|\boldsymbol{\tau} + O(\|\boldsymbol{\delta}\mathbf{r}\|^2). \tag{4.12}$$

The dominant order in eq. (4.12) was first obtained by Hills (1979).

4.4 Form of eligible TG flows within visible domains

Consider a visible domain $\mathcal{V}$ and a point M within $\mathcal{V}$ connected to a node N_i on the geographic equator by an iso-ζ segment $\Gamma_{\zeta_M}^{N_i,M}$. Integrating eq. (4.4) along $\Gamma_{\zeta_M}^{N_i,M}$ and using eq. (4.1) yields

$$\mathbf{u}_M = \frac{\|\boldsymbol{\nabla}_H\zeta\|_M}{\cos\theta_M}\left[\int_{\Gamma_{\zeta_M}^{N_i,M}} \boldsymbol{\nabla}_H\cdot\left(\frac{\partial_t B_r}{\|\boldsymbol{\nabla}_H\zeta\|}\boldsymbol{\pi}\right)\frac{\boldsymbol{\tau}\cdot\mathbf{dl}}{\|\boldsymbol{\nabla}_H\zeta\|}\right]\boldsymbol{\tau}_M - \frac{1}{\cos\theta_M}\frac{(\partial_t B_r)_M}{\|\boldsymbol{\nabla}_H\zeta\|_M}\boldsymbol{\pi}_M, \tag{4.13}$$

where **dl** is oriented from N_i to M. Note, after eq. (4.1), $\partial q/\partial\zeta$ necessarily vanishes at N_i as $\cos\theta = 0$ and $\|\nabla_H\zeta\| \to +\infty$ on the geographic the equator. Therefore, the integral in eq. (4.13), which starts from a node, is well-defined. Although this integral diverges at critical points, its product with $\|\nabla_H\zeta\|_M$ (that goes to zero there) remains finite.

The pressure at M may be related to the pressure at the node N_i by integrating eq. (4.3) along $\Gamma_{\zeta_M}^{N_i,M}$:

$$q_M = q_{N_i} + \int_{\Gamma_{\zeta_M}^{N_i,M}} \frac{\partial_t B_r}{\|\nabla_H\zeta\|}\boldsymbol{\tau}\cdot\mathbf{dl}, \tag{4.14}$$

with the same convention for **dl**.

4.5 Form of eligible TG flows within ambiguous domains

Consider an ambiguous domain $\mathcal{A}$ and let

$$g_{\mathcal{A}}(\zeta) \equiv \boldsymbol{\tau}_{K_{\mathcal{A}}(\zeta)}\cdot\mathbf{u}_{K_{\mathcal{A}}(\zeta)} \tag{4.15}$$

for $\zeta_{\mathcal{A}}^{(1)} < \zeta < \zeta_{\mathcal{A}}^{(2)}$ (if, for example, $\zeta_{\mathcal{A}}^{(1)} < \zeta_{\mathcal{A}}^{(2)}$; see Fig. 5). Note that the definition of $\boldsymbol{\tau}_{K_{\mathcal{A}}(\zeta)}$ and $\boldsymbol{\pi}_{K_{\mathcal{A}}(\zeta)}$, hence of $g_{\mathcal{A}}(\zeta)$, may be extended by continuity along $\Gamma_{\mathcal{A}}$ at both critical points $C_{\mathcal{A}}^{(1)}$ and $C_{\mathcal{A}}^{(2)}$. Near $C = C_{\mathcal{A}}^{(1)}$ and $C_{\mathcal{A}}^{(2)}$, multiplying the first-order expansion of $\nabla_H\zeta$ by $\|\nabla_H\zeta\|^{-1}\mathbf{Z}_C^{-1}$ leads to

$$\frac{\delta\mathbf{r}}{\|\nabla_H\zeta\|} = \mathbf{Z}_C^{-1}\pi + O(\|\delta\mathbf{r}\|). \tag{4.16}$$

Applying eq. (4.12) to the point $K_{\mathcal{A}}(\zeta)$ and making use of eq. (4.16), the function $g_{\mathcal{A}}(\zeta)$ may then be expanded to first order in $\|\delta\mathbf{r}\|$ near $C = C_{\mathcal{A}}^{(1)}$ and $C_{\mathcal{A}}^{(2)}$:

$$g_{\mathcal{A}}(\zeta) = \boldsymbol{\tau}_{K_{\mathcal{A}}(\zeta_C)}\cdot\mathbf{U}_C + \delta\boldsymbol{\tau}_{K_{\mathcal{A}}(\zeta)}\cdot\mathbf{U}_C + \beta_C\|\mathbf{Z}_C\delta\mathbf{r}\| + O(\|\delta\mathbf{r}\|^2), \tag{4.17}$$

where $\delta\boldsymbol{\tau}_{K_{\mathcal{A}}(\zeta)}$ is the first-order term of the expansion of $\boldsymbol{\tau}_{K_{\mathcal{A}}(\zeta)} - \boldsymbol{\tau}_{K_{\mathcal{A}}(\zeta_C)}$ near C and

$$\beta_C \equiv \left[\boldsymbol{\tau}_{K_{\mathcal{A}}(\zeta_C)}\right]^T \mathbf{T}_C\mathbf{Z}_C^{-1}\boldsymbol{\pi}_{K_{\mathcal{A}}(\zeta_C)} + \alpha_C. \tag{4.18}$$

Now consider a point M within $\mathcal{A}$ on an iso-ζ curve ∂S_{ζ_M} crossing $\Gamma_{\mathcal{A}}$ at $K_{\mathcal{A}}(\zeta_M)$. Integrating eq. (4.4) from $K_{\mathcal{A}}(\zeta_M)$ to M along $\Gamma_{\zeta_M}^{K_{\mathcal{A}}(\zeta_M),M}$ and using eqs (4.1) and (4.15) yields

$$\mathbf{u}_M = \frac{\|\nabla_H\zeta\|_M}{\cos\theta_M}\left[\frac{\cos\theta_{K_{\mathcal{A}}(\zeta_M)}}{\|\nabla_H\zeta\|_{K_{\mathcal{A}}(\zeta_M)}}g_{\mathcal{A}}(\zeta_M) + \int_{\Gamma_{\zeta_M}^{K_{\mathcal{A}}(\zeta_M),M}} \nabla_H\cdot\left(\frac{\partial_t B_r}{\|\nabla_H\zeta\|}\pi\right)\frac{\boldsymbol{\tau}\cdot\mathbf{dl}}{\|\nabla_H\zeta\|}\right]\boldsymbol{\tau}_M - \frac{1}{\cos\theta_M}\frac{(\partial_t B_r)_M}{\|\nabla_H\zeta\|_M}\boldsymbol{\pi}_M, \tag{4.19}$$

where **dl** is oriented from $K_{\mathcal{A}}(\zeta_M)$ to M. Although the integral in eq. (4.19) diverges when $\zeta_M \to \zeta_{\mathcal{A}}^{(1)}$ or $\zeta_{\mathcal{A}}^{(2)}$ and $K_{\mathcal{A}}(\zeta_M) \to C_{\mathcal{A}}^{(1)}$ or $C_{\mathcal{A}}^{(2)}$, its sum with $g_{\mathcal{A}}(\zeta_M)/\|\nabla_H\zeta\|_{K_{\mathcal{A}}(\zeta_M)}$ remains finite.

The pressure at M may be related to the pressure at $C_{\mathcal{A}}^{(1)}$ by integrating eq. (4.3) successively from $C_{\mathcal{A}}^{(1)}$ to $K_{\mathcal{A}}(\zeta_M)$ along $\Gamma_{\mathcal{A}}$ and from $K_{\mathcal{A}}(\zeta_M)$ to M along $\Gamma_{\zeta_M}^{K_{\mathcal{A}}(\zeta_M),M}$, and by using $\boldsymbol{\tau}\cdot$ eqs (4.1) and (4.15):

$$q_M = q_{C_{\mathcal{A}}^{(1)}} + \int_{\zeta_{\mathcal{A}}^{(1)}}^{\zeta_M} \frac{\cos\theta_{K_{\mathcal{A}}(\zeta)}}{\|\nabla_H\zeta\|_{K_{\mathcal{A}}(\zeta)}}g_{\mathcal{A}}(\zeta)d\zeta + \int_{\Gamma_{\zeta_M}^{K_{\mathcal{A}}(\zeta_M),M}} \frac{\partial_t B_r}{\|\nabla_H\zeta\|}\boldsymbol{\tau}\cdot\mathbf{dl}, \tag{4.20}$$

with the same convention for **dl**. The first integral in eq. (4.20) is well-defined as $d\zeta\|\nabla_H\zeta\|_{K_{\mathcal{A}}(\zeta)}^{-1}$ remains bounded at critical points.

5 EXISTENCE AND COMPLETE SET OF ELIGIBLE TG FLOWS

In the present section, it is assumed that the SV satisfies the basic set of constraints of eqs (3.1) and (3.2), with an arbitrary set $\{q_{(i)}\}$, for all non-critical curves ∂S_{ζ_0} and $\Gamma_{\zeta_0}^{i,j}$ within a geostrophic region $\mathcal{G}$ of arbitrary extent (defined by $\zeta_{\max}$). Relying on the results of Sections 2, 3 and 4, an eligible TG flow $\mathbf{u}_0$ will be constructed step-by-step in Section 5.1. This will prove that the constraints of eqs (3.1) and (3.2) are sufficient conditions for the existence of eligible TG flows. The complete set of eligible TG flows will be exhibited in Section 5.2.

5.1 Construction of an eligible TG flow

To begin with, $\mathbf{u}_0$ is defined by eq. (4.13) everywhere within the visible belt except at nodes, critical points and on critical curves. The convergence of the integral in eq. (4.13) is ensured by eq. (2.10) of Lemma 2. This definition is unambiguous because of the consequence of eq. (3.4) of the basic set of constraints.

Next, for each critical point C belonging to an ambiguous patch $\mathcal{A}$, a scalar $(\beta_0)_C$ is chosen as follows: if C is a saddle point belonging to both $\mathcal{A}$ and the visible belt, $(\beta_0)_C$ is given by the behaviour of $\mathbf{u}_0$ near C in the visible belt [using eqs (4.12) and (4.18)]; in all other cases (saddle points not belonging to the visible belts, extrema), $(\beta_0)_C$ is arbitrarily chosen. The following two-step procedure is then repeated for each ambiguous domain $\mathcal{A}$.

(i) A continuously differentiable scalar function $(g_0)_{\mathcal{A}}$ is arbitrarily chosen on the interval $[\zeta_{\mathcal{A}}^{(1)}, \zeta_{\mathcal{A}}^{(2)}$; if, for example, $\zeta_{\mathcal{A}}^{(1)} \leq \zeta_{\mathcal{A}}^{(2)}]$, such that its behaviour near $C = C_{\mathcal{A}}^{(1)}$ and $C_{\mathcal{A}}^{(2)}$ is given by eq. (4.17) with $\beta_C = (\beta_0)_C$.

(ii) $\mathbf{u}_0$ is defined by eq. (4.19) with $g_{\mathcal{A}} = (g_0)_{\mathcal{A}}$ within $\mathcal{A}$, except at critical points and on critical curves. This definition is unambiguous because of the consequence of eq. (3.3) of the basic set of constraints.

Thus, the flow $\mathbf{u}_0$ is defined and continuously differentiable (because of the first basic assumption on B_r) everywhere within $\mathcal{G}$ except at nodes, critical points and on critical curves. Its definition will now be extended by continuity to the rest of the core surface.

First consider a node N and apply Lemma 2. Because of eqs (2.9) and (4.13), $\boldsymbol{\pi}\cdot\,\mathbf{u}_0$ satisfies eq. (4.5) near N. Because of eqs (2.10), (4.13) and (4.6), $\boldsymbol{\tau}\cdot\,\mathbf{u}_0$ satisfies eq. (4.7) near N. Therefore, eq. (4.8) applies to $\mathbf{u}_0$. Thus, the definition of $\mathbf{u}_0$ may be extended by continuity at nodes, $(\mathbf{u}_0)_N = \mathbf{U}_N$, and $\mathbf{u}_0$ is continuously differentiable at such points.

Next, consider a critical point C. Because of the consequence of eq. (3.5) of the basic set of constraints, Lemma 3 applies. Because of eqs (2.13), (4.13) and (4.19), $\boldsymbol{\pi}\cdot\,\mathbf{u}_0$ satisfies eq. (4.9) near C. Because of eqs (2.14), (4.13), (4.19) and the consequence eq. (3.6) of the basic set of constraints, $\boldsymbol{\tau}\cdot\,\mathbf{u}_0$ satisfies eq. (4.10) near C. Therefore, there exists a function $(\alpha_0)_C(\Gamma_\zeta)$ such that $\mathbf{u}_0$ satisfies eq. (4.11) near C. As $\mathbf{u}_0$ is continuously differentiable everywhere near C except for C, eq. (4.12) applies to $\mathbf{u}_0$ with $(\alpha_0)_C = \lim_{\zeta\to\zeta_C}(\alpha_0)_C(\Gamma_\zeta)$. [Note that if C belongs to an ambiguous patch $\mathcal{A}$, $(\alpha_0)_C$ is related to $(\beta_0)_C$ through eq. (4.18).] Thus, the definition of $\mathbf{u}_0$ may be extended by continuity at critical points, $(\mathbf{u}_0)_C = \mathbf{U}_C$, and $\mathbf{u}_0$ is continuously differentiable at such points.

At any saddle point C, eq. (4.12) means that $\mathbf{u}_0$ and its derivatives in θ and ϕ have the same finite limits on both sides of the critical curves crossing each other at C. Therefore, $\mathbf{u}_0$ and its derivatives have a finite limit everywhere on these critical curves because of eqs (4.13) and (4.19). The definition of $\mathbf{u}_0$ may thus be extended by continuity on all critical curves and $\mathbf{u}_0$ is continuously differentiable on such curves.

The constructed flow $\mathbf{u}_0$ is TG. To see this, take the divergence of the components of $\mathbf{u}_0\cos\theta$ along $\boldsymbol{\tau}$ and $\boldsymbol{\pi}$. Within the visible belt, except at nodes, critical points and on critical curves, eq. (4.13) leads to

$$\boldsymbol{\nabla}_H\cdot[(\boldsymbol{\tau}\cdot\mathbf{u}_0\cos\theta)\boldsymbol{\tau}] = \boldsymbol{\nabla}_H\cdot\left(\frac{\partial_t B_r}{\|\boldsymbol{\nabla}_H\zeta\|}\boldsymbol{\pi}\right) = -\boldsymbol{\nabla}_H\cdot[(\boldsymbol{\pi}\cdot\mathbf{u}_0\cos\theta)\boldsymbol{\pi}]. \tag{5.1}$$

Within ambiguous patches, except at nodes, critical points and on critical curves, eq. (4.19) and

$$\boldsymbol{\nabla}_H\cdot\left[\|\boldsymbol{\nabla}_H\zeta\|\frac{\cos\theta_{K_{\mathcal{A}(\zeta)}}}{\|\boldsymbol{\nabla}_H\zeta\|_{K_{\mathcal{A}(\zeta)}}}(g_0)_{\mathcal{A}}(\zeta)\boldsymbol{\tau}\right] = \|\boldsymbol{\nabla}_H\zeta\|(\boldsymbol{\tau}\cdot\boldsymbol{\nabla}_H)\left[\frac{\cos\theta_{K_{\mathcal{A}(\zeta)}}}{\|\boldsymbol{\nabla}_H\zeta\|_{K_{\mathcal{A}(\zeta)}}}(g_0)_{\mathcal{A}}(\zeta)\right] = 0 \tag{5.2}$$

[which comes from noticing that $(g_0)_{\mathcal{A}}(\zeta)\cos\theta_{K_{\mathcal{A}(\zeta)}}/\|\boldsymbol{\nabla}_H\zeta\|_{K_{\mathcal{A}(\zeta)}}$ does not vary with $\boldsymbol{\tau}$] also lead to eq. (5.1). Therefore, $\mathbf{u}_0$ satisfies eq. (1.3) everywhere within $\mathcal{G}$, except at nodes, critical points and on critical curves. As $\mathbf{u}_0$ is continuously differentiable everywhere within $\mathcal{G}$, it also satisfies eq. (1.3) at such points and on such curves.

Finally, it may be checked that $\mathbf{u}_0$ is the solution of eq. (1.1) by substituting $\boldsymbol{\pi}\cdot\,\mathbf{u}_0$, given by eqs (4.13) or (4.19), into $\boldsymbol{\nabla}_H\cdot(\mathbf{u}_0 B_r) = \boldsymbol{\pi}\cdot\,\mathbf{u}_0\cos\theta\|\boldsymbol{\nabla}_H\zeta\|$ everywhere within $\mathcal{G}$ except at nodes, critical points and on critical curves, and by making use of the property that $\mathbf{u}_0$ is continuously differentiable everywhere within $\mathcal{G}$. Thus, $\mathbf{u}_0$ is an eligible TG flow.

5.2 Complete set of eligible TG flows

Consider an eligible TG flow $\mathbf{u}$. Equation (4.13) implies that $\mathbf{u} = \mathbf{u}_0$ within the visible belt, i.e. there is only one eligible TG flow within this region of the core surface. The associated pressure field q is determined by eq. (4.14) to within an arbitrary constant within each visible domain. This result was obtained independently by Hills (1979) and Backus & Le Mouël (1986): the latter introduced the name visible belt. However, none of these authors provided an explicit expression of the solution such as eq. (4.13).

Equation (4.19) implies that for each ambiguous domain $\mathcal{A}$ there exists a function $g_{\mathcal{A}}(\zeta)$ such that $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_{\mathcal{A}}$, where $\mathbf{u}_{\mathcal{A}}$ is given by

$$\mathbf{u}_{\mathcal{A}} = \frac{\|\boldsymbol{\nabla}_H\zeta\|}{\cos\theta}\frac{\cos\theta_{K_{\mathcal{A}(\zeta)}}}{\|\boldsymbol{\nabla}_H\zeta\|_{K_{\mathcal{A}(\zeta)}}}[g_{\mathcal{A}}(\zeta) - (g_0)_{\mathcal{A}}(\zeta)]\boldsymbol{\tau}. \tag{5.3}$$

Similarly, eq. (4.20) implies that the pressure field q associated with $\mathbf{u}$ is obtained by adding

$$q_{K_{\mathcal{A}(\zeta)}} = \int_{\zeta_{\mathcal{A}}^{(1)}}^{\zeta}\frac{\cos\theta_{K_{\mathcal{A}(\zeta)}}}{\|\boldsymbol{\nabla}_H\zeta\|_{K_{\mathcal{A}(\zeta)}}}[g_{\mathcal{A}}(\zeta) - (g_0)_{\mathcal{A}}(\zeta)]\,d\zeta \tag{5.4}$$

to the pressure field q_0 associated with $\mathbf{u}_0$. There is more than one eligible TG flow within ambiguous patches, hence their name introduced by Backus & Le Mouël (1986). The function $g_{\mathcal{A}}(\zeta)$ is continuously differentiable and satisfies eq. (4.17) near the critical points $C_{\mathcal{A}}^{(1)}$ and $C_{\mathcal{A}}^{(2)}$, with $\beta_{C_{\mathcal{A}}^{(i)}} = (\beta_0)_{C_{\mathcal{A}}^{(i)}}$ if $C_{\mathcal{A}}^{(i)}$ belongs to both $\mathcal{A}$ and the visible belt, $\beta_{C_{\mathcal{A}}^{(i)}}$ not related to $(\beta_0)_{C_{\mathcal{A}}^{(i)}}$ otherwise.

Now for any ambiguous domain $\mathcal{A}$ it is possible to arbitrarily choose a continuously differentiable scalar function $g_{\mathcal{A}}(\zeta)$ satisfying eq. (4.17) near the critical points $C_{\mathcal{A}}^{(1)}$ and $C_{\mathcal{A}}^{(2)}$, with $\beta_{C_{\mathcal{A}}^{(i)}} = (\beta_0)_{C_{\mathcal{A}}^{(i)}}$ if $C_{\mathcal{A}}^{(i)}$ belongs to both $\mathcal{A}$ and the visible belt, whatever $\beta_{C_{\mathcal{A}}^{(i)}}$ otherwise. Consider the flow $\mathbf{u}$ defined by $\mathbf{u} = \mathbf{u}_0$ within the visible belt and by $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_{\mathcal{A}}$, where $\mathbf{u}_{\mathcal{A}}$ is given by eq. (5.3), within each ambiguous patch $\mathcal{A}$. Then in the whole geostrophic region: $\mathbf{u}$ is continuously differentiable because $\mathbf{u}_0$ is continuously differentiable and $g_{\mathcal{A}}(\zeta)$ satisfies eq. (4.17); $\mathbf{u}$ satisfies eq. (1.3) because of eq. (5.2) applied to $g_{\mathcal{A}}(\zeta)$; and $\mathbf{u}$ is solution of eq. (1.1) because $\boldsymbol{\pi}\cdot\,\mathbf{u} = \boldsymbol{\pi}\cdot\,\mathbf{u}_0$. Thus $\mathbf{u}$ is an eligible TG flow.

In conclusion, the eligible TG flows are the flows $\mathbf{u}$ defined by:

(i) $\mathbf{u} = \mathbf{u}_0$ within the visible belt;

(ii) $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_{\mathcal{A}}$ within each ambiguous patch $\mathcal{A}$, where $\mathbf{u}_{\mathcal{A}}$ is given by eq. (5.3) and $g_{\mathcal{A}}(\zeta)$ is any continuously differentiable function satisfying eq. (4.17), with $\beta_{C^{(i)}_{\mathcal{A}}} = (\beta_0)_{C^{(i)}_{\mathcal{A}}}$ if $C^{(i)}_{\mathcal{A}}$ belongs to both $\mathcal{A}$ and the visible belt, whatever $\beta_{C^{(i)}_{\mathcal{A}}}$ otherwise.

6 CONCLUSION

The first result of this paper is that the curvilinear integral constraints of eqs (3.1) and (3.2) found in Paper I are not only necessary conditions, but also sufficient conditions for the SV to be generated by a TG flow under the FF assumption within a given geostrophic region $\mathcal{G}$. The proof involved the construction of an explicit solution $\mathbf{u}_0$ within $\mathcal{G}$: this was made possible by a prior investigation of the general form of TG flows generating SV under the FF assumption. This result means that no other independent constraint can be found. It provides the theoretical tools for consistently testing the combined FF+TG assumption. Such tests would consist of constructing time-varying field models satisfying the constraints of eqs (3.1) and (3.2). This is all the more promising now that new satellite data are being made available. However, it should be noted that the methods used for implementing a finite number of constraints in the lone FF case (e.g. Bloxham & Gubbins 1986) cannot be straightforwardly applied to this case where the number of constraints is infinite. If tests do not dismiss the FF+TG assumption, geomagnetic models taking into account the MHD properties of the core surface within an estimated geostrophic region could then be constructed. Such models would be specifically designed for core studies and could be more relevant than existing ones when used as observational constraints for geodynamo modelling.

The second result of this study is the exhibition of the complete set of TG solutions of the induction equation under the FF assumption. Equations (4.13) and (4.19) (and their limits eqs (2.11) and (2.15) at nodes and critical points, respectively) are amenable to direct numerical computation (within an arbitrary flow tangent to the iso-ζ curves of the ambiguous patches). Also, eq. (B12) could be used to avoid finite difference computation of the quantity $\nabla_H \cdot (\partial_t B_r \|\nabla_H \zeta\|^{-1}\, \boldsymbol{\pi})$, as all quantities in this formula may be analytically evaluated from the spherical harmonics expansion of B_r and $\partial_t B_r$. In a previous study, the geostrophic pressure was computed using eq. (4.14) within the visible belt in an attempt to implement such a local method (Chulliat & Hulot 2000). Yet, the starting magnetic field model was unconstrained, generating spurious features in the pressure field. Provided constrained field models can be constructed from existing geomagnetic data, the present paper provides the tools for directly and locally computing the flow in a fully consistent way.

ACKNOWLEDGMENTS

The author thanks Gauthier Hulot for many discussions and for his comments on an early version of the manuscript. The author is grateful to Professor David Loper for encouraging this research during his stay at GFDI which was supported by NSF grant #0000400. A review by R. Holme helped improve the manuscript. This is GFDI contribution n^o 439.

REFERENCES

Backus, G.E., 1968. Kinematics of geomagnetic secular variation in a perfectly conducting core, *Phil. Trans. R. Soc. Lond.,* A,**263,** 239–266.

Backus, G.E., 1986. Poloidal and toroidal fields in geomagnetic field modeling, *Rev. Geophys.,* **24,** 75–109.

Backus, G.E. & Le Mouël, J.-L., 1986. The region on the core-mantle boundary where a geostrophic velocity field can be determined from frozen-flux magnetic data, *Geophys. J. R. astr. Soc.,* **85,** 617–628.

Bloxham, J. & Gubbins, D., 1986. Geomagnetic field analysis—IV. Testing the frozen—flux hypothesis, *Geophys. J. R. astr. Soc.,* **84,** 139–152.

Bloxham, J. & Jackson, A., 1992. Time-dependent mapping of the magnetic field at the core–mantle boundary, *J. geophys. Res.,* **97,** 19 537–19 563.

Chulliat, A. & Hulot, G., 2000. Local computation of the geostrophic pressure at the top of the core, *Phys. Earth planet. Int.,* **117,** 309–328.

Chulliat, A. & Hulot, G., 2001. Geomagnetic secular variation generated by a tangentially geostrophic flow under the frozen—flux assumption—I. Necessary conditions, *Geophys. J. Int.,* **147,** 237–246.

Gubbins, D. & Roberts, P.H., 1987. In: *Magnetohydrodynamics of the Earth's Core, Geomagnetism,* Press, New York Vol. 2, Academic, pp. 1–183, ed., Jacobs, J.

Hills, R.G., 1979. Convection in the Earth's mantle due to viscous shear at the core-mantle interface and due to large-scale buoyancy, *PhD thesis*, New Mexico State University, USA.

Hulot, G., Eymin, C., Langlais, B. & Mandea, M., 2002. Small-scale structure of the geodynamo inferred from Oersted and Magsat satellite data, *Nature,* **416,** 620–623.

Jackson, A., Jonkers, A.R.T. & Walker, M.R., 2000. Four centuries of geomagnetic secular variation from historical records, *Phil. Trans. R. Soc. Lond.,* A, **358,** 957–990.

Jault, D. & Le Mouël, J.-L., 1991. Physical properties at the top of the core and core surface motions, *Phys. Earth planet. Int.,* **68,** 76–84.

Langlais, B., Mandea, M. & Ultré-Guérard, P., 2003. High-resolution magnetic field modeling: application to MAGSAT and Orsted data, *Phys. Earth planet. Int.,* **135,** 77–91.

Le Mouël, J.-L., 1984. Outer core geostrophic flow and secular variation of Earth's geomagnetic field, *Nature,* **311,** 734–735.

O'Brien, M.S., Constable, C.G. & Parker, R.L., 1997. Frozen-flux modelling for epochs 1915 and 1980, *Geophys. J. Int.,* **128,** 434–450.

Olsen, N., 2002. A model of the geomagnetic field and its secular variation for epoch 2000 estimated from ørsted data, *Geophys. J. Int.,* **149,** 454–462.

Rau, S., Christensen, U., Jackson, A. & Wicht, J., 2000. Core flow inversion tested with numerical dynamo models, *Geophys. J. Int.,* **141,** 485–497.

Roberts, P.H. & Glatzmaier, G., 2000. A test of the frozen-flux approximation using a new geodynamo model, *Phil. Trans. R. Soc. Lond.,* A, **358,** 1109–1121.

Roberts, P.H. & Scott, S., 1965. On analysis of the secular variation, *J. Geomag. Geoelectr.,* **17,** 137–151.

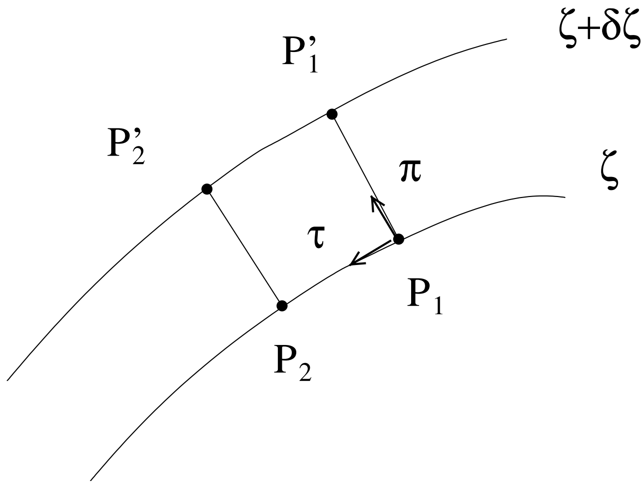

Figure A1. Illustration for Gauss theorem.

APPENDIX A: PROOF OF LEMMA 1

The derivative with respect to ζ of the curvilinear integral of f along $\Gamma_\zeta^{M_1,M_2}$ is defined by

$$\frac{d}{d\zeta}\left(\int_{\Gamma_\zeta^{M_1,M_2}} f\boldsymbol{\tau}\cdot\mathbf{dl}\right) \equiv \lim_{\delta\zeta\to 0}\frac{1}{\delta\zeta}\left[\int_{\Gamma_{\zeta+\delta\zeta}^{M'_1,M'_2}} f(\mathbf{r}')\boldsymbol{\tau}\cdot\mathbf{dl}' - \int_{\Gamma_\zeta^{M_1,M_2}} f(\mathbf{r})\boldsymbol{\tau}\cdot\mathbf{dl}\right], \tag{A1}$$

where

$$\mathbf{r}' = \mathbf{r} + \frac{\delta\zeta}{\|\nabla_H\zeta\|}\boldsymbol{\pi}. \tag{A2}$$

For $\delta\zeta$ small, $f(\mathbf{r}')$ and $f(\mathbf{r})$ are related by

$$f(\mathbf{r}') = f(\mathbf{r}) + \frac{\delta\zeta}{\|\nabla_H\zeta\|}\boldsymbol{\pi}\cdot\nabla_H f. \tag{A3}$$

The Gauss theorem for a continuously differentiable tangent vector field $\mathbf{V}$ on a sphere (Backus 1986) reads

$$(\nabla_H\cdot\mathbf{V})\frac{\delta\zeta\,\delta l}{\|\nabla_H\zeta\|} = (\boldsymbol{\pi}\cdot\mathbf{V})_{P'_1}\delta l' - (\boldsymbol{\pi}\cdot\mathbf{V})_{P_1}\delta l + (\boldsymbol{\tau}\cdot\mathbf{V})_{P_2}\frac{\delta\zeta}{\|\nabla_H\zeta\|_{P_2}} - (\boldsymbol{\tau}\cdot\mathbf{V})_{P_1}\frac{\delta\zeta}{\|\nabla_H\zeta\|_{P_1}}, \tag{A4}$$

where P_1 and P_2 are two points very close to each other on the same iso-ζ curve and, P'_1 and P'_2 are defined by

$$P'_{1,2} = P_{1,2} + \frac{\delta\zeta}{\|\nabla_H\zeta\|}\boldsymbol{\pi}, \tag{A5}$$

$\delta l = P_1P_2$ and $\delta l' = P'_1P'_2$ (Fig. A1). Taking $\mathbf{V} = \boldsymbol{\pi}$, $\delta l = \boldsymbol{\tau}\cdot\mathbf{dl}$ and $\delta l' = \boldsymbol{\tau}\cdot\mathbf{dl}'$ into eq. (A4) yields

$$\boldsymbol{\tau}\cdot\mathbf{dl}' = \left(1 + \frac{\delta\zeta}{\|\nabla_H\zeta\|}\nabla_H\cdot\boldsymbol{\pi}\right)\boldsymbol{\tau}\cdot\mathbf{dl}. \tag{A6}$$

Substituting eqs (A3) and (A6) into eq. (A1) leads to eq. (2.8).

APPENDIX B: PROOFS OF LEMMAS 2 AND 3

The purpose of this appendix is to state eqs (2.9), (2.10), (2.13) and (2.14) of the main text; that is, to obtain the asymptotic expansions of the quantities $\partial_t B_r\|\nabla_H\zeta\|^{-1}$ and $\nabla_H\cdot(\partial_t B_r\|\nabla_H\zeta\|^{-1}\boldsymbol{\pi})$ near nodes (where $\|\nabla_H\zeta\|$ goes to infinity) and critical points (where $\|\nabla_H\zeta\|$ vanishes).

B1 Asymptotic expansion of $\partial_t B_r\|\nabla_H\zeta\|^{-1}$ near nodes

Consider a node N. At any point $N + \boldsymbol{\delta}\mathbf{r}$ near N outside the geographic equator, the function $\nabla_H B_r$ may be expanded to first order in $\|\boldsymbol{\delta}\mathbf{r}\|$ as

$$\nabla_H B_r = (\nabla_H B_r)_N + \mathbf{W}_N\boldsymbol{\delta}\mathbf{r} + O(\|\boldsymbol{\delta}\mathbf{r}\|^2). \tag{B1}$$

Using this and $B_r = \zeta\cos\theta$, the first-order expansion of the function $\cos\theta\|\nabla_H\zeta\|$ near N may be expressed as

$$\cos\theta\,\nabla_H\zeta = (\nabla_H B_r)_N + \left[(\nabla_H B_r)_N\cdot\left(\frac{\boldsymbol{\delta}\mathbf{r}}{c\cos\theta}\right)\right]\mathbf{e}_\theta + \mathbf{W}_N\boldsymbol{\delta}\mathbf{r} + \frac{1}{2}\left[\left(\frac{\boldsymbol{\delta}\mathbf{r}}{c\cos\theta}\right)^T\mathbf{W}_N\boldsymbol{\delta}\mathbf{r}\right]\mathbf{e}_\theta + O(\|\boldsymbol{\delta}\mathbf{r}\|^2). \tag{B2}$$

Taking the scalar product of $-\cos\theta\ \nabla_H\zeta$ and $\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r}$ leads to the first-order expansion of $\partial_t B_r$ near N,

$$\partial_t B_r = (\partial_t B_r)_N + (\nabla_H \partial_t B_r)_N \cdot \boldsymbol{\delta}\mathbf{r} + O(\|\boldsymbol{\delta}\mathbf{r}\|^2), \tag{B3}$$

which proves eq. (2.9).

B2 Asymptotic expansion of $\partial_t B_r \|\nabla_H \zeta\|^{-1}$ near critical points where the SV vanishes

Consider a critical point C such that $(\partial_t B_r)_C = 0$. At any point $C + \boldsymbol{\delta}\mathbf{r}$ near C, the first basic assumption on $\mathbf{B}$ makes it possible to expand the function ζ to third order in $\|\boldsymbol{\delta}\mathbf{r}\|$ as

$$\zeta = \zeta_C + \frac{1}{2}\boldsymbol{\delta}\mathbf{r}^T \mathbf{Z}_C \boldsymbol{\delta}\mathbf{r} + \frac{1}{6}\left[(\partial_{\theta\theta\theta}\zeta)_C(\delta\theta)^3 + 3(\partial_{\theta\theta\phi}\zeta)_C(\delta\theta)^2\delta\phi + 3(\partial_{\theta\phi\phi}\zeta)_C\delta\theta(\delta\phi)^2 + (\partial_{\phi\phi\phi}\zeta)_C(\delta\phi)^3\right] + O(\|\boldsymbol{\delta}\mathbf{r}\|^4). \tag{B4}$$

Thus, $\nabla_H\zeta$ may be expanded to second order near C as

$$\nabla_H\zeta = \mathbf{Z}_C\boldsymbol{\delta}\mathbf{r} + \frac{1}{2}\left(\boldsymbol{\delta}\mathbf{r}^T\mathbf{X}_C\boldsymbol{\delta}\mathbf{r}\right)\mathbf{e}_\theta + \frac{1}{2}(\boldsymbol{\delta}\mathbf{r}^T\mathbf{Y}_C\boldsymbol{\delta}\mathbf{r})\mathbf{e}_\phi - \frac{1}{c}\cot\theta_C\left(\mathbf{e}_\theta\cdot\boldsymbol{\delta}\mathbf{r}\right)(\mathbf{e}_\phi^T\mathbf{Z}_C\boldsymbol{\delta}\mathbf{r})\mathbf{e}_\phi + O(\|\boldsymbol{\delta}\mathbf{r}\|^3). \tag{B5}$$

Taking the scalar product of $-\cos\theta_C\ \nabla_H\zeta$ and $\mathbf{U}_C + \mathbf{T}_C\boldsymbol{\delta}\mathbf{r}$ leads to the second-order expansion of $\partial_t B_r$ near C,

$$\partial_t B_r = (\nabla_H\partial_t B_r)_C\cdot\boldsymbol{\delta}\mathbf{r} + \frac{1}{2}\boldsymbol{\delta}\mathbf{r}^T\,\partial_\mathbf{t}\mathbf{W}_C\boldsymbol{\delta}\mathbf{r} + O(\|\boldsymbol{\delta}\mathbf{r}\|^3), \tag{B6}$$

which proves eq. (2.13).

B3 Expression of $\nabla_H\cdot(\partial_t B_r\|\nabla_H\zeta\|^{-1}\,\boldsymbol{\pi})$ in the general case and useful formulae

The first steps to obtain asymptotic expansions of $\nabla_H\cdot(\partial_t B_r\|\nabla_H\zeta\|^{-1}\,\boldsymbol{\pi})$ near nodes and critical points are the same. The divergence may be expanded as

$$\nabla_H\cdot\left(\frac{\partial_t B_r}{\|\nabla_H\zeta\|}\boldsymbol{\pi}\right) = \frac{\boldsymbol{\pi}\cdot\nabla_H\partial_t B_r}{\|\nabla_H\zeta\|} - 2\frac{\partial_t B_r}{\|\nabla_H\zeta\|^2}\boldsymbol{\pi}\cdot\nabla_H\|\nabla_H\zeta\| + \frac{\partial_t B_r}{\|\nabla_H\zeta\|^2}\nabla_H\cdot\nabla_H\zeta. \tag{B7}$$

The following intermediate formulae are obtained by expressing $\nabla_H\zeta$ in spherical coordinates and using eqs (2.1), (2.2) and (2.3):

$$\nabla_H(\nabla_H\zeta)_\theta = \frac{\partial_{\theta\theta}\zeta}{c^2}\mathbf{e}_\theta + \frac{\partial_{\theta\phi}\zeta}{c^2\sin\theta}\mathbf{e}_\phi, \tag{B8}$$

$$\nabla_H(\nabla_H\zeta)_\phi = \left(\frac{\partial_{\theta\phi}\zeta}{c^2\sin\theta} - \frac{1}{c}\|\nabla_H\zeta\|\cot\theta\pi_\phi\right)\mathbf{e}_\theta + \frac{\partial_{\phi\phi}\zeta}{c^2\sin^2\theta}\mathbf{e}_\phi, \tag{B9}$$

$$\nabla_H\|\nabla_H\zeta\| = \mathbf{Z}\boldsymbol{\pi} - \frac{1}{c}\|\nabla_H\zeta\|\cot\theta\pi_\phi^2\mathbf{e}_\theta, \tag{B10}$$

$$\nabla_H\cdot\nabla_H\zeta = \boldsymbol{\pi}^T\mathbf{Z}\boldsymbol{\pi} + \boldsymbol{\tau}^T\mathbf{Z}\boldsymbol{\tau} + \frac{1}{c}\|\nabla_H\zeta\|\cot\theta\pi_\theta. \tag{B11}$$

[Note that eq. (B10) can be straightforwardly deduced from eqs (B8) and (B9) by taking $\mathbf{V} = \nabla_H\zeta$ in the general relationship $\|\mathbf{V}\|\nabla_H\|\mathbf{V}\| = V_\theta\nabla_H V_\theta + V_\phi\nabla_H V_\phi$.] Substituting eqs (B10) and (B11) into eq. (B7) yields

$$\nabla_H\cdot\left(\frac{\partial_t B_r}{\|\nabla_H\zeta\|}\boldsymbol{\pi}\right) = \frac{\boldsymbol{\pi}\cdot\nabla_H\partial_t B_r}{\|\nabla_H\zeta\|} + \frac{\partial_t B_r}{\|\nabla_H\zeta\|^2}\left(\boldsymbol{\tau}^T\mathbf{Z}\boldsymbol{\tau} - \boldsymbol{\pi}^T\mathbf{Z}\boldsymbol{\pi}\right) + \frac{1}{c}\frac{\partial_t B_r}{\|\nabla_H\zeta\|}\cot\theta\pi_\theta(2\pi_\phi^2 + 1). \tag{B12}$$

This expression is valid at any point of the core surface where $\|\nabla_H\zeta\| \neq 0$ and ∞.

A few general formulae will prove useful in the remainder of the proof. For all horizontal vectors $\mathbf{V}$ at the core surface, let

$$(\boldsymbol{\tau}\cdot\nabla_H)^\star\mathbf{V} = \left(\boldsymbol{\tau}\cdot\nabla_H V_\theta\right)\mathbf{e}_\theta + \left(\boldsymbol{\tau}\cdot\nabla_H V_\phi\right)\mathbf{e}_\phi. \tag{B13}$$

This operator satisfies the following general properties for all horizontal vectors $\mathbf{V}_1$ and $\mathbf{V}_2$, for all 2×2 constant matrices $\mathbf{A}$ and for all small horizontal vectors $\boldsymbol{\delta}\mathbf{r}$:

$$(\boldsymbol{\tau}\cdot\nabla_H)(\mathbf{V}_1\cdot\mathbf{V}_2) = \mathbf{V}_1\cdot[(\boldsymbol{\tau}\cdot\nabla_H)^\star\mathbf{V}_2] + [(\boldsymbol{\tau}\cdot\nabla_H)^\star\mathbf{V}_1]\cdot\mathbf{V}_2, \tag{B14}$$

$$(\boldsymbol{\tau}\cdot\nabla_H)^\star(\mathbf{A}\boldsymbol{\delta}\mathbf{r}) = \mathbf{A}\boldsymbol{\tau} + O(\|\boldsymbol{\delta}\mathbf{r}\|). \tag{B15}$$

Also, eqs (B8), (B9) and (B10) yield

$$(\boldsymbol{\tau}\cdot\nabla_H)^\star\left(\frac{\boldsymbol{\tau}}{\|\nabla_H\zeta\|}\right) = -\frac{1}{\|\nabla_H\zeta\|^2}[(\boldsymbol{\tau}^T\mathbf{Z}\boldsymbol{\tau})\boldsymbol{\pi} + (\boldsymbol{\tau}^T\mathbf{Z}\boldsymbol{\pi})\boldsymbol{\tau}] - \frac{1}{c\|\nabla_H\zeta\|}\cot\theta\pi_\phi^2[(2\pi_\theta^2 - 1)\mathbf{e}_\theta + 2\pi_\theta\pi_\phi\mathbf{e}_\phi]. \tag{B16}$$

B4 Asymptotic expansion of $\nabla_H \cdot(\partial_t B_r \|\nabla_H\zeta\|^{-1}\, \boldsymbol{\pi})$ near nodes

Consider a node N. At any point $N + \boldsymbol{\delta}\mathbf{r}$ near N, the horizontal gradient of the SV may be expanded to first order in $\|\boldsymbol{\delta}\mathbf{r}\|$ as

$$\nabla_H \partial_t B_r = (\nabla_H \partial_t B_r)_N + \partial_t \mathbf{W}_N \boldsymbol{\delta}\mathbf{r} + O(\|\boldsymbol{\delta}\mathbf{r}\|^2), \tag{B17}$$

where $\partial_t\mathbf{W}$ is given by eq. (2.7). Using eqs (2.9), (B2) and (B3), $(\nabla_H \partial_t B_r)_N$ may be expressed as

$$(\nabla_H \partial_t B_r)_N = -\mathbf{W}_N \mathbf{U}_N - \cos\theta \|\nabla_H \zeta\| \mathbf{T}_N^T \boldsymbol{\pi} + O(\|\boldsymbol{\delta}\mathbf{r}\|). \tag{b18}$$

Also,

$$\frac{\boldsymbol{\delta}\mathbf{r}}{c\cos\theta} = -\frac{\boldsymbol{\tau}}{\tau_\theta} + O(\|\boldsymbol{\delta}\mathbf{r}\|). \tag{B19}$$

Substituting eqs (2.9), (B17), (B18) and (B19) into eq. (B12) yields

$$\nabla_H \cdot \left(\frac{\partial_t B_r}{\|\nabla_H \zeta\|}\boldsymbol{\pi}\right) = -\frac{\cos\theta}{\|\nabla_H\zeta\|}(\boldsymbol{\tau}^T \mathbf{Z}\boldsymbol{\tau} - \boldsymbol{\pi}^T\mathbf{Z}\boldsymbol{\pi})\boldsymbol{\pi}\cdot(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r}) - \frac{1}{\|\nabla_H\zeta\|}\boldsymbol{\pi}^T\mathbf{W}_N\mathbf{U}_N - \cos\theta\,\boldsymbol{\pi}^T\mathbf{T}_N\boldsymbol{\pi} + O(\|\boldsymbol{\delta}\mathbf{r}\|^2). \tag{B20}$$

(Note that $\boldsymbol{\pi}^T\mathbf{T}_N^T\boldsymbol{\pi} = \boldsymbol{\pi}^T\mathbf{T}_N\boldsymbol{\pi}$.) The part containing $\mathbf{U}_N$ in the first term of the right-hand side of eq. (B20) is $O(1)$; the other part and the other two terms behave like $\|\boldsymbol{\delta}\mathbf{r}\|$.

Now taking $\mathbf{V}_1 = \boldsymbol{\tau}\cos\theta/\|\nabla_H\zeta\|$ and $\mathbf{V}_2 = \mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r}$ into (B14), taking $\mathbf{A} = \mathbf{T}_N$ into eq. (B15) and using eq. (B16) yields

$$\begin{aligned}\|\nabla_H\zeta\|(\boldsymbol{\tau}\cdot\nabla_H)\left[\frac{\boldsymbol{\tau}\cdot(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r})\cos\theta}{\|\nabla_H\zeta\|}\right] &= -\frac{\cos\theta}{\|\nabla_H\zeta\|}(\boldsymbol{\tau}^T\mathbf{Z}\boldsymbol{\tau} - \boldsymbol{\pi}^T\mathbf{Z}\boldsymbol{\pi})\boldsymbol{\pi}\cdot(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r})\\ &\quad -\frac{\cos\theta}{\|\nabla_H\zeta\|}\boldsymbol{\pi}^T\mathbf{Z}\cdot(\mathbf{U}_N + \mathbf{T}_N\boldsymbol{\delta}\mathbf{r}) + 2\cos\theta\,\boldsymbol{\tau}^T\mathbf{T}_N\boldsymbol{\tau} - \frac{\tau_\theta}{c}\boldsymbol{\tau}\cdot\mathbf{U}_N + O(\|\boldsymbol{\delta}\mathbf{r}\|^2),\end{aligned} \tag{B21}$$

where eq. (2.11) and the identity

$$\boldsymbol{\pi}^T\mathbf{Z} = (\boldsymbol{\pi}^T\mathbf{Z}\boldsymbol{\pi})\boldsymbol{\pi}^T + (\boldsymbol{\pi}^T\mathbf{Z}\boldsymbol{\tau})\boldsymbol{\tau}^T \tag{B22}$$

have also been used.

The first term of the right-hand side of eq. (B20) is identical to that of eq. (B21). Combining eq. (2.11) and the following expansion of $\mathbf{Z}$ near N,

$$\mathbf{Z} = \frac{\|\nabla_H\zeta\|}{c\cos\theta}\begin{pmatrix}2\pi_\theta & \pi_\phi\\ \pi_\phi & 0\end{pmatrix} + \frac{1}{\cos\theta}\mathbf{W}_N + O(1), \tag{B23}$$

leads to a relationship between one term of eq. (B20) and two terms of eq. (B21):

$$-\frac{\cos\theta}{\|\nabla_H\zeta\|}\boldsymbol{\pi}^T\mathbf{Z}\mathbf{U}_N = \frac{\tau_\theta}{c}\boldsymbol{\tau}\cdot\mathbf{U}_N - \frac{1}{\|\nabla_H\zeta\|}\boldsymbol{\pi}^T\mathbf{W}_N\mathbf{U}_N + O(\|\boldsymbol{\delta}\mathbf{r}\|^2). \tag{B24}$$

Another relationship,

$$-\frac{\cos\theta}{\|\nabla_H\zeta\|}\boldsymbol{\pi}^T\mathbf{Z}\mathbf{T}_N\boldsymbol{\delta}\mathbf{r} = -\cos\theta\,\boldsymbol{\tau}^T\mathbf{T}_N\boldsymbol{\tau} + O(\|\boldsymbol{\delta}\mathbf{r}\|^2), \tag{B25}$$

may be obtained using eq. (B23) and the property $\mathbf{e}_\theta{}^T\ \mathbf{T}_N = \mathbf{0}$. Substituting eqs (B24) and (B25) into eq. (B21),

$$Tr(\mathbf{T}_N) = \boldsymbol{\tau}^T\mathbf{T}_N\boldsymbol{\tau} + \boldsymbol{\pi}^T\mathbf{T}_N\boldsymbol{\pi} \tag{B26}$$

into eq. (4.6) (a general formula valid near nodes) and combining both formulae with eq. (B20) leads to eq. (2.10).

B5 Asymptotic expansion of $\nabla_H \cdot(\partial_t B_r \|\nabla_H\zeta\|^{-1}\, \boldsymbol{\pi})$ near critical points where the SV vanishes

Consider a critical point C such that $(\partial_t B_r)_C = 0$. At any point $C + \boldsymbol{\delta}\mathbf{r}$ near C, the horizontal gradient of the SV may be expanded to first order in $\|\boldsymbol{\delta}\mathbf{r}\|$ as

$$\nabla_H\partial_t B_r = (\nabla_H\partial_t B_r)_C + \partial_t\mathbf{W}_C\boldsymbol{\delta}\mathbf{r} - \frac{1}{c}\cot\theta_C(\mathbf{e}_\theta\cdot\boldsymbol{\delta}\mathbf{r})[\mathbf{e}_\phi\cdot(\nabla_H\partial_t B_r)_C]\mathbf{e}_\phi + O(\|\boldsymbol{\delta}\mathbf{r}\|^2), \tag{B27}$$

where $\partial_t\mathbf{W}$ is given by eq. (2.7). Substituting eqs (2.13), (4.16) (obtained by multiplying eq. (B5) by $\|\nabla_H\zeta\|\mathbf{Z}_C^{-1}$) and (B27) into eq. (B12) yields

$$\begin{aligned}\nabla_H\cdot\left(\frac{\partial_t B_r}{\|\nabla_H\zeta\|}\boldsymbol{\pi}\right) = {}& \frac{1}{\|\nabla_H\zeta\|}\left[\boldsymbol{\pi}\cdot(\nabla_H\partial_t B_r)_C - (\boldsymbol{\pi}\cdot\mathbf{U}_C\cos\theta_C)\big(\boldsymbol{\tau}^T\mathbf{Z}_C\boldsymbol{\tau} - \boldsymbol{\pi}^T\mathbf{Z}_C\boldsymbol{\pi}\big)\right]\\ & -(\boldsymbol{\pi}\cdot\mathbf{U}_C\cos\theta_C)\left[\boldsymbol{\tau}^T\frac{(\mathbf{Z}-\mathbf{Z}_C)}{\|\nabla_H\zeta\|}\boldsymbol{\tau} - \boldsymbol{\pi}^T\frac{(\mathbf{Z}-\mathbf{Z}_C)}{\|\nabla_H\zeta\|}\boldsymbol{\pi}\right]\\ & -\boldsymbol{\pi}^T\mathbf{T}_C\mathbf{Z}_C^{-1}\boldsymbol{\pi}\big[\boldsymbol{\tau}^T\mathbf{Z}_C\boldsymbol{\tau} - \boldsymbol{\pi}^T\mathbf{Z}_C\boldsymbol{\pi}\big]\cos\theta_C + \boldsymbol{\pi}^T\partial_t\mathbf{W}_C\mathbf{Z}_C^{-1}\boldsymbol{\pi}\\ & +\frac{1}{c}\cot\theta_C\big[(\mathbf{e}_\theta^T\mathbf{Z}_C^{-1}\boldsymbol{\pi})\pi_\phi\mathbf{e}_\phi^T\mathbf{Z}_C - \pi_\theta(2\pi_\phi^2+1)\boldsymbol{\pi}\big]\cdot\mathbf{U}_C\cos\theta_C + O(\|\boldsymbol{\delta}\mathbf{r}\|).\end{aligned} \tag{B28}$$

The first term of the RHS of eq. (B28) behaves like $\|\boldsymbol{\delta}\mathbf{r}\|^{-1}$; the next four terms are $O(1)$.

Now taking $\mathbf{V}_1 = \boldsymbol{\tau}/\|\boldsymbol{\nabla}_H\zeta\|$, $\mathbf{V}_2 = \mathbf{U}_C + \mathsf{T}_C\boldsymbol{\delta}\mathbf{r}$ and $\mathbf{A} = \mathbf{T}_C$ into eqs (B14), (B15) and (B16) yields

$$\|\boldsymbol{\nabla}_H\zeta\|(\boldsymbol{\tau}\cdot\boldsymbol{\nabla}_H)\left[\frac{\boldsymbol{\tau}\cdot(\mathbf{U}_C+\mathsf{T}_C\boldsymbol{\delta}\mathbf{r})\cos\theta_C}{\|\boldsymbol{\nabla}_H\zeta\|}\right] = \frac{1}{\|\boldsymbol{\nabla}_H\zeta\|}\left[\boldsymbol{\pi}\cdot(\boldsymbol{\nabla}_H\partial_t B_r)_C - (\boldsymbol{\pi}\cdot\mathbf{U}_C\cos\theta_C)\left(\boldsymbol{\tau}^T\mathsf{Z}_C\boldsymbol{\tau} - \boldsymbol{\pi}^T\mathsf{Z}_C\boldsymbol{\pi}\right)\right]$$
$$-(\boldsymbol{\pi}\cdot\mathbf{U}_C\cos\theta_C)\left[\boldsymbol{\tau}^T\frac{(\mathsf{Z}-\mathsf{Z}_C)}{\|\boldsymbol{\nabla}_H\zeta\|}\boldsymbol{\tau} - \boldsymbol{\pi}^T\frac{(\mathsf{Z}-\mathsf{Z}_C)}{\|\boldsymbol{\nabla}_H\zeta\|}\boldsymbol{\pi}\right]$$
$$-\boldsymbol{\pi}^T\mathsf{T}_C\mathsf{Z}_C^{-1}\boldsymbol{\pi}\left[\boldsymbol{\tau}^T\mathsf{Z}_C\boldsymbol{\tau} - \boldsymbol{\pi}^T\mathsf{Z}_C\boldsymbol{\pi}\right]\cos\theta_C$$
$$-\boldsymbol{\pi}^T\frac{(\mathsf{Z}-\mathsf{Z}_C)}{\|\boldsymbol{\nabla}_H\zeta\|}\mathbf{U}_C\cos\theta_C - \boldsymbol{\pi}^T\mathsf{Z}_C\mathsf{T}_C\mathsf{Z}_C^{-1}\boldsymbol{\pi}\cos\theta_C + \boldsymbol{\tau}^T\mathsf{T}_C\boldsymbol{\tau}\cos\theta_C$$
$$+\frac{1}{c}\cot\theta_C\left[-\pi_\theta(2\pi_\phi^2+1)\boldsymbol{\pi} + \mathbf{e}_\theta + \pi_\theta\pi_\phi\mathbf{e}_\phi\right]\cdot\mathbf{U}_C\cos\theta_C + O\left(\|\boldsymbol{\delta}\mathbf{r}\|\right), \quad \text{(B29)}$$

where eqs (2.15) and (B22) have also been used.

The first three terms of the right-hand side of eq. (B29) are identical to the first three terms of the right-hand side of eq. (B28). The next two terms may be transformed as

$$-\boldsymbol{\pi}^T\frac{(\mathsf{Z}-\mathsf{Z}_C)}{\|\boldsymbol{\nabla}_H\zeta\|}\mathbf{U}_C\cos\theta_C - \boldsymbol{\pi}^T\mathsf{Z}_C\mathsf{T}_C\mathsf{Z}_C^{-1}\boldsymbol{\pi}\cos\theta_C = \boldsymbol{\pi}^T\mathsf{T}_C\boldsymbol{\pi}\cos\theta_C + \boldsymbol{\pi}^T\partial_{\mathbf{t}}\mathsf{W}_C\mathsf{Z}_C^{-1}\boldsymbol{\pi}$$
$$+\frac{1}{c}\cot\theta_C\left\{(\mathbf{e}_\theta^T\mathsf{Z}_C^{-1}\boldsymbol{\pi})\pi_\phi\left[(Z_C)_{\theta\phi}\mathbf{e}_\theta + (Z_C)_{\phi\phi}\mathbf{e}_\phi\right] - \pi_\theta\pi_\phi\mathbf{e}_\phi\right\}\cdot\mathbf{U}_C\cos\theta_C \quad \text{(B30)}$$

To prove eq. (B30), use the first-order expansion of Z near C,

$$\mathsf{Z} = \mathsf{Z}_C + \begin{pmatrix}\mathbf{e}_\theta^T\mathsf{X}_C\boldsymbol{\delta}\mathbf{r} & \mathbf{e}_\phi^T\mathsf{X}_C\boldsymbol{\delta}\mathbf{r}\\ \mathbf{e}_\theta^T\mathsf{Y}_C\boldsymbol{\delta}\mathbf{r} & \mathbf{e}_\phi^T\mathsf{Y}_C\boldsymbol{\delta}\mathbf{r}\end{pmatrix} - \frac{1}{c}\cot\theta_C(\mathbf{e}_\theta\cdot\boldsymbol{\delta}\mathbf{r})\begin{bmatrix}0 & (\mathsf{Z}_C)_{\theta\phi}\\ (\mathsf{Z}_C)_{\theta\phi} & 2(\mathsf{Z}_C)_{\phi\phi}\end{bmatrix} + O\left(\|\boldsymbol{\delta}\mathbf{r}\|^2\right), \quad \text{(B31)}$$

where X and Y are given by eqs (2.4) and (2.5), note that $\mathbf{e}_\phi{}^T\,\mathbf{X}_C = -\mathbf{e}_\theta{}^T\,\mathbf{Y}_C$ and use the property $\boldsymbol{\pi}^T\mathsf{A}_1\mathsf{A}_2\boldsymbol{\pi} = \boldsymbol{\pi}^T\mathsf{A}_2\mathsf{A}_1\boldsymbol{\pi}$ for any symmetrical 2×2 matrices A_1 and A_2. Substituting eq. (B30) into eq. (B29), subtracting the quantity

$$\boldsymbol{\nabla}_H\cdot\left[(\mathbf{U}_C+\mathsf{T}_C\boldsymbol{\delta}\mathbf{r})\cos\theta_C\right] = Tr(\mathsf{T}_C)\cos\theta_C + \frac{1}{c}\cot\theta_C(\mathbf{e}_\theta\cdot\mathbf{U}_C\cos\theta_C) + O\left(\|\boldsymbol{\delta}\mathbf{r}\|\right), \quad \text{(B32)}$$

and combining the result with eq. (B28) finally leads to eq. (2.14).

APPENDIX C: PROOFS OF EQUATIONS (3.5) AND (3.6)

C1 Proof of eq. (3.5)

Assume there exists a critical point C where eq. (3.5) does not hold. Then there exists a neighbourhood $\mathcal{N}$ of C where $\partial_t B_r \neq 0$, for example >0 and for any (segment or full) iso-ζ curve Γ_ζ within $\mathcal{N}$ we have

$$I(\Gamma_\zeta) \equiv \int_{\Gamma_\zeta}\partial_t B_r\frac{\boldsymbol{\tau}\cdot\mathbf{dl}}{\|\boldsymbol{\nabla}_H\zeta\|} > \min_{\mathcal{N}}(\partial_t B_r)\int_{\Gamma_\zeta}\frac{dl}{\|\boldsymbol{\nabla}_H\zeta\|}, \quad \text{(C1)}$$

where $\min_{\mathcal{N}}(f)$ denotes the minimum of a function f within $\mathcal{N}$. Introduce the eigenvalues λ_1 and λ_2 of Z (such that $\lambda_1 > \lambda_2$), the corresponding unit eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$ and the coordinates x_1 and x_2 of $C + \boldsymbol{\delta}\mathbf{r}$ in the reference frame (C, $\mathbf{e}_1$, $\mathbf{e}_2$). Recalling eq. (B4), the expansion of ζ to second order in $\boldsymbol{\delta}\mathbf{r}$ at any point $C + \boldsymbol{\delta}\mathbf{r}$ near C may be expressed as

$$\zeta = \zeta_C + \frac{1}{2}\lambda_1 x_1^2 + \frac{1}{2}\lambda_2 x_2^2 + O(\|\delta\mathbf{r}\|^3). \quad \text{(C2)}$$

If C is an extremum, $\lambda_1\lambda_2 > 0$ and the iso-ζ curves near C are ellipses to second order in $\|\boldsymbol{\delta}\mathbf{r}\|$. Let Γ_ζ be a curve ∂S_ζ within $\mathcal{N}$. Then inequality (eq. C1) implies $G(\partial S_\zeta) > 0$. However, eq. (3.1) implies $G(\partial S_\zeta) = 0$, hence the contradiction. If C is a saddle point, $\lambda_1\lambda_2 < 0$ ($\lambda_1 > 0$ and $\lambda_2 < 0$) and the iso-ζ curves near C are branches of hyperboles to second order in $\|\boldsymbol{\delta}\mathbf{r}\|$. Let Γ_ζ be a portion of a curve $\Gamma_\zeta^{i,j}$ going through $\mathcal{N}$ and choose the branch that crosses the x_2-axis and that runs through $[-x_0, x_0]$ for the x_1 coordinate; on that branch we have that $\zeta < \zeta_C$. Then the integral in the right-hand side of eq. (C1) may be lower bounded as

$$\int_{\Gamma_\zeta}\frac{dl}{\|\boldsymbol{\nabla}_H\zeta\|} > \int_{-x_0}^{x_0}\frac{dx_1}{\|\boldsymbol{\nabla}_H\zeta\|} = \frac{2}{b}\int_0^{u_0}\frac{du}{\sqrt{u^2+1}} + O(1), \quad \text{(C3)}$$

where $b = \sqrt{\lambda_1^2 - \lambda_1\lambda_2}$, $u_0 = x_0/a$ and $a^2 = 2\lambda_2(\zeta - \zeta_C)/(\lambda_1^2 - \lambda_1\lambda_2)$. As $\zeta \to \zeta_C$, $u_0 \to +\infty$ and the integral in u goes to infinity. Therefore, $I(\Gamma_\zeta) \to +\infty$ as $\zeta \to \zeta_C$. However, eq. (3.2) implies that $G(\partial S_\zeta)$ is bounded as $\zeta \to \zeta_C$, hence the contradiction.

C2 Proof of eq. (3.6)

Assume there exists a critical point C where eq. (3.5) holds while eq. (3.6) does not hold; i.e. where the dominant order term in eq. (B32) [which is otherwise valid provided eq. (3.5) is satisfied] does not vanish. Then there exists a neighbourhood $\mathcal{N}$ of C where $\boldsymbol{\nabla}_H \cdot [(\mathbf{U}_C + \mathsf{T}_C \boldsymbol{\delta}\mathbf{r}) \cos\theta_C] \neq 0$, for example > 0, and for any (segment or full) iso-ζ curve Γ_ζ within $\mathcal{N}$ we have

$$J(\Gamma_\zeta) \equiv \int_{\Gamma_\zeta} \boldsymbol{\nabla}_H \cdot \big[(\mathbf{U}_C + \mathsf{T}_C \boldsymbol{\delta}\mathbf{r}) \cos\theta_C\big] \frac{\boldsymbol{\tau} \cdot \mathbf{dl}}{\|\boldsymbol{\nabla}_H \zeta\|} > \min_{\mathcal{N}} \big\{\boldsymbol{\nabla}_H \cdot \big[(\mathbf{U}_C + \mathsf{T}_C \boldsymbol{\delta}\mathbf{r}) \cos\theta_C\big]\big\} \int_{\Gamma_\zeta} \frac{dl}{\|\boldsymbol{\nabla}_H \zeta\|}. \tag{C4}$$

Because of eq. (3.5), Lemma 3 applies. Using eq. (2.14), it may subsequently be proven along the same lines as the proof of eq. (3.5) that the inequality eq. (C4) is in contradiction with eqs (3.3) and (3.4) (which are direct consequences of eqs (3.1) and (3.2)).