CrossMark

J. Plasma Phys. (2022), *vol*. 88, 835880101 © The Author(s), 2022.
Published by Cambridge University Press
doi:10.1017/S0022377821001355

Relaxed magnetohydrodynamics with ideal Ohm's law constraint

R.L. Dewar [1],† **and Z.S. Qu** [1]

[1]Mathematical Sciences Institute, The Australian National University, Canberra, ACT 2601, Australia

(Received 5 July 2021; revised 15 December 2021; accepted 20 December 2021)

The gap between a recently developed dynamical version of relaxed magnetohydrodynamics (RxMHD) and ideal MHD (IMHD) is bridged by approximating the zero-resistivity 'ideal' Ohm's law (IOL) constraint using an augmented Lagrangian method borrowed from optimization theory. The augmentation combines a pointwise vector Lagrange multiplier method and global penalty function method and can be used either for iterative enforcement of the IOL to arbitrary accuracy, or for constructing a continuous sequence of magnetofluid dynamics models running between RxMHD (no IOL) and weak IMHD (IOL almost everywhere). This is illustrated by deriving dispersion relations for linear waves on an MHD equilibrium.

Key words: plasma confinement, plasma dynamics, plasma simulation

1. Introduction

1.1. *Basics*

In this paper choosing constraint equations is central to our approach to developing new fluid models. The concept of a constraint equation occurs in both the variational approach to classical mechanics (see e.g. Goldstein 1980) and optimization theory (see e.g. Nocedal & Wright 2006). While both traditionally treat finite-dimensional systems, the language and techniques of these fields can also help in understanding the infinite-dimensional dynamics of non-dissipative continuous media. In the following we shall distinguish between a hard constraint, i.e. one that is enforced exactly, a soft constraint, one that is enforced only approximately, and a weak version of a hard constraint, one that is enforced as the limiting case of a sequence of soft constraints (formulating such a method being the goal of this work, which it is hoped will lead to a physical regularization[1] of magnetohydrodynamics (MHD) that allows reconnection).

We also distinguish between microscopic, i.e. acting within each fluid element or infinitesimal parcel of fluid, and macroscopic constraints, i.e. global within a spatial domain Ω of the fluid (or subdomain if the system is partitioned into multiple regions).

† Email address for correspondence: robert.dewar@anu.edu.au

[1]We use regularization in the physics sense – adjusting for incipient singular behaviour in a way that is consistent with physics on scales outside the strict domain of applicability of a mathematical model. This goes somewhat beyond the mathematical sense of adjusting a problem to avoid ill-posedness.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

The mathematical model we seek to regularize is ideal MHD (IMHD), a special case in the general field of MHD. In the general, resistive case Ohm's law is $\boldsymbol{E}' = \eta \boldsymbol{j}$, where

$$\boldsymbol{E}'[\boldsymbol{u}] \stackrel{\text{def}}{=} \boldsymbol{E} + \boldsymbol{u}\times\boldsymbol{B}, \tag{1.1}$$

is the electric field observed in the local frame of each fluid element, $\boldsymbol{E}$ being the electric field in the lab frame. These elements are advected in the fluid velocity field $\boldsymbol{u}(\boldsymbol{x}, t)$ (i.e. $\dot{\boldsymbol{x}} = \boldsymbol{u}$ at each spatial point $\boldsymbol{x}$ and time t). Also $\boldsymbol{B}(\boldsymbol{x}, t)$ is the magnetic field, η is the resistivity and $\boldsymbol{j}(\boldsymbol{x}, t)$ is the electric current density (note that $\boldsymbol{j} = \nabla\times\boldsymbol{B}/\mu_0$ in standard non-relativistic MHD, where μ_0 is the vacuum permeability constant used in SI electromagnetic units). We have exhibited $\boldsymbol{u}$ as an explicit argument for use later in the paper, while leaving dependencies on $\boldsymbol{x}$, t, $\boldsymbol{E}$, and $\boldsymbol{B}$ implicit.

To get IMHD, set $\eta = 0$ so that $\boldsymbol{E}' = 0$, giving what is often called the ideal Ohm's law (IOL),

$$\boldsymbol{E} + \boldsymbol{u}\times\boldsymbol{B} = 0. \tag{1.2}$$

While $\boldsymbol{E}$ is not usually explicit in the IMHD equations, this is only because it is eliminated between (1.2), after taking the curl of both sides, and the Maxwell–Faraday induction equation

$$\nabla\times\boldsymbol{E} = -\partial_t\boldsymbol{B}, \tag{1.3}$$

to give the IMHD magnetic-field propagation equation

$$\partial_t\boldsymbol{B} + \nabla\times(\boldsymbol{u}\times\boldsymbol{B}) = 0. \tag{1.4}$$

With the 'pre-Maxwell' Ampére's law $\boldsymbol{j} = \nabla\times\boldsymbol{B}/\mu_0$ and $\nabla\cdot\boldsymbol{B} = 0$, the Maxwell–Faraday equation (1.3) plays the important role of preserving Galilean invariance (Hosking & Dewar 2015, § 5.4; Webb & Anco 2019), independent of whether or not the IOL equation is enforced. Thus we shall retain it in the following development of a dynamical relaxation theory.

Equation (1.3) can be viewed as a holonomic constraint on $\boldsymbol{E}$, and likewise $\nabla\cdot\boldsymbol{B} = 0$ is a holonomic constraint on $\boldsymbol{B}$, i.e. we can remove these constraints from consideration by expressing the constrained variables in terms of fewer unconstrained variables. Here these are the vector and scalar potentials $\boldsymbol{A}$ and Φ, respectively, in terms of which

$$\boldsymbol{B} = \nabla\times\boldsymbol{A}, \tag{1.5}$$

$$\boldsymbol{E} = -\partial_t\boldsymbol{A} - \nabla\Phi. \tag{1.6}$$

These imply $\nabla\cdot\boldsymbol{B} = 0$ and also (1.3), as is easily seen by calculating $\nabla\times\boldsymbol{E}$.

We restrict the choice of gauge to be such that Φ is a spatially single-valued potential and such that $\partial_t\boldsymbol{A} = 0$ in equilibrium cases in a frame (the laboratory frame) where $\partial_t\cdot = 0$, so $\boldsymbol{A}$ has no effect on $\boldsymbol{E}$ in that static case. Of course the vector potential still does play an explicit role in describing plasma equilibria because the magnetic flux threading a loop is $\oint\boldsymbol{A}\cdot\mathrm{d}\boldsymbol{l}$. Dynamically, only $\partial_t\boldsymbol{A}$ contributes to inductive electromagnetic forces $\oint\boldsymbol{E}\cdot\mathrm{d}\boldsymbol{l}$ around closed loops. In our case, we assume electromagnetic forces are zero around any loop on the boundary $\partial\Omega$ – the trapped-flux boundary condition of relaxed magnetohydrodynamics (RxMHD) (see Appendix B of Dewar *et al.* 2015). Aside from this restriction, there is still considerable gauge freedom in $\boldsymbol{A}$. If we choose Coulomb gauge, $\nabla\cdot\boldsymbol{A} = 0$, the potential representation is an example of the Helmholtz decomposition of an arbitrary vector field into the sum of curl-free and divergence-free vector fields, but we shall not make this gauge choice except in §§ 5.5 and 6 – we shall treat the magnetic helicity term carefully in our general derivation of the conservation form momentum equation in order to make it gauge invariant.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

1.2. *Methodology: variational principles and Euler–Lagrange equations*

In mechanics and optimization theory there are objective functions whose extrema – maxima, minima and saddle points – are given by Euler–Lagrange (EL) equations, which are found by setting first derivatives of these functions to zero. In mechanics such functions are Hamiltonians whose extrema give stable or unstable equilibria, or actions, time integrals of Lagrangians, whose extrema give physical time evolution equations (Hamilton's principle).

The main aim of this paper is to use an infinite-dimensional generalization of Hamilton's principle in which partial derivatives are replaced by functional derivatives (see e.g. Morrison 1998) of action integrals incorporating the IOL constraint, and also global entropy, magnetic helicity and cross-helicity constraints. These functional derivatives are with respect to the basic physical fields, e.g. $\Phi, \boldsymbol{A}, \boldsymbol{u}$, etc., describing the state of the system and are set to zero to find a set of EL equations which together are sufficient to describe the dynamics of the system. For brevity we shall refer, for example, to the equation found by setting the functional derivative with respect to Φ as the '$\delta\Phi$-EL equation'.

1.3. *Relaxation*

See Appendix A for a brief history of the variational approach to finding relaxed plasma equilibrium states by minimizing the IMHD energy functional using one or more IMHD invariants as global constraints. This construction implies immediately that such relaxed magnetostatic states are a special subset of all possible IMHD equilibria, most of which, being of higher energy, are likely to be more unstable than relaxed states.

In this paper we instead seek to find a time-dependent variational formulation for relaxed plasma systems going through a dynamical phase as they transition from one equilibrium state to another (e.g. due to boundary deformations). Thus, instead of minimizing energy, we use Hamilton's variational principle, widely regarded as the most fundamental principal in all mathematical physics, from general relativity through classical mechanics to quantum field theories (for instance connecting symmetries and conservation laws by Noether's theorem). As we are attempting to establish a new classical field theory related to, but different from, IMHD, it is appropriate to seek new magnetofluid models by modifying the IMHD Hamilton's principle.

Following this precept, Dewar *et al.* (2020) derived a new dynamical magnetofluid model, RxMHD, from Hamilton's action principle using a phase-space version of the magnetofluid Lagrangian with a non-canonical momentum field $\boldsymbol{u}$, physically identified as the laboratory frame mass flow velocity, and a kinematically constrained velocity field $\boldsymbol{v}$ (the fluid velocity relative to a magnetic-field-aligned flow). The resulting EL equations generalize from statics to dynamics the usual relaxation-by-energy-minimization concept developed by Taylor (1986) for flowless plasma equilibria, and its generalization to equilibria with steady flow by various authors (Finn & Antonsen 1983; Hameiri 1998; Vladimirov, Moffatt & Ilin 1999; Hameiri 2014; Dennis *et al.* 2014*b*). These generalized Taylor equilibria were shown by Dewar *et al.* (2020) to be consistent with RxMHD when time derivatives are set to zero. However, specific cases of equilibria with flows not aligned with the magnetic field have been limited to axisymmetric equilibria; whereas in this paper we aim to treat more general, non-axisymmetric (three-dimensional (3-D)) equilibria with flow, as well as time-dependent problems such as the calculation of the spectrum of normal modes of oscillation of 3-D relaxed equilibria.

The advection equation for $\boldsymbol{B}$, (1.4), implies the 'frozen-in flux constraint', which, as discussed by Newcomb (1958), preserves the topology of magnetic field lines. This prevents field-line breaking and reconnection from forming new structures, such

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

as magnetic islands, and this frustration of topological changes leads to singularities developing as time tends towards infinity (Grad 1967).

Though in this paper we proceed in a formal way by simply inserting constancy constraints of selected IMHD invariants as postulates, historically the heuristic assumption motivating relaxation theory is that, if it would be energetically favourable to do so, and on a long enough time scale, 'nature will find a way' for reconnection to occur, either due to the magnifying effect of large gradients on small but finite resistivity at singularities, or through 'anomalous' phenomena such as turbulence. Thus in the RxMHD of Dewar *et al.* (2020) the continuum of local frozen-in flux constraints is replaced by only two constraints involving $\boldsymbol{B}$, the two global IMHD invariants magnetic helicity and cross-helicity.

However, as will be argued in § 3.2, there is reason to believe that, for general 3-D equilibria with non-integrable magnetic field dynamics, imposing (1.2) as a hard constraint would lead to an ill-posed variational principle with no smooth extremum. In this case we regularize the problem by approaching an IOL-constrained state through a sequence of softly constrained states where the IOL constraint is not exactly satisfied.

For a dynamical RxMHD theory to be fully satisfactory we require it to be well-posed mathematically and desire it to agree with IMHD in two cases: (i) on the boundary $\partial\Omega(t)$, because multiregion RxMHD (MRxMHD) interfaces are regarded as arbitrarily thin sheets of IMHD fluid; and (ii) in an equilibrium state with steady flow, when one imagines any transient non-ideal behaviour to have died away, justifying the principle of IMHD-equilibrium consistency (Dewar *et al.* 2020).

This consistency principle was satisfied by the one flowing equilibrium test case Dewar *et al.* (2020) looked at using their RxMHD formulation: the rigidly rotating axisymmetric steady-flow equilibrium. However, RxMHD does not enforce the IOL constraint (1.2), so there is no reason to believe that ideal consistency would necessarily apply to more general relaxed equilibria. (Indeed, Dewar *et al.* (2020) showed that small dynamical perturbations about an equilibrium exhibited no tendency to preserve the IOL constraint.)

Specifically, we are interested in non-axisymmetric relaxed steady-flow toroidal equilibria such as may occur in stellarators. The elliptic nature of RxMHD (when flows are small) makes it reasonable to assume that smooth solutions of the RxMHD equations exist for such equilibria. We argue in § 3.2 that, generically, magnetic field and fluid flow lines on these smooth RxMHD solutions will be chaotic so their ergodic properties will exhibit complexity on all scales.

While RxMHD offers no impediment to the formation of such fractal structure (one of the principal motivations for the development of the stepped pressure equilibrium code (SPEC) code (Hudson *et al.* 2012)) the same is not true for IMHD where the topological constraints arising from its frozen-in-flux properties (see above) force the formation of singularities. The ability of SPEC to study difficult physical problems (Qu *et al.* 2020), and subtle fundamental problems involving chaos (Qu *et al.* 2021), motivates our current endeavour to extend the RxMHD formalism on which it is based to make it closer to IMHD but to retain sufficient topological relaxation to allow magnetic island formation and chaos, thus allowing further extension of SPEC to hande time-dependent problems in 3-D geometries.

1.4. *Background flow*

We define a fully relaxed RxMHD equilibrium as one where the electrostatic potential Φ has relaxed to a constant value throughout a volume Ω, so $\boldsymbol{E} = 0$. As Finn & Antonsen (1983) recognized, this would occur in the extreme case where magnetic field lines fill Ω ergodically, because dotting both sides of (1.2) with $\boldsymbol{B}$ gives the derivative along $\boldsymbol{B}$ as $\boldsymbol{B}\cdot\nabla\Phi = 0$. As we shall see, constant Φ implies purely parallel flow, $\boldsymbol{u} = \boldsymbol{u}_{\|}$, whose

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

magnitude is constrained by the steady-flow continuity equation $\nabla\cdot[(\rho/B)\boldsymbol{B}u_\parallel]=\boldsymbol{B}\cdot\nabla(\rho u_\parallel/B)=0$, where ρ is mass density. For consistency again with the (unachievable) fully ergodic limit, we define fully relaxed parallel flow as such that $\rho u_\parallel/B=\text{const}$. We denote this special parallel flow velocity as

$$\boldsymbol{u}^{\mathrm{Rx}}\stackrel{\mathrm{def}}{=}\frac{\nu_\Omega\boldsymbol{B}}{\mu_0\rho},\tag{1.7}$$

where ν_Ω is a constant throughout Ω – its significance in the RxMHD formalism is explained below.

In the variational $\boldsymbol{u},\boldsymbol{v}$ dynamical relaxation formalism of Dewar *et al.* (2020), EL equations for $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{B}$ and pressure p are derived variationally from Hamilton's principle, while the mass continuity equation is built-in as a holonomic constraint. The fully relaxed flow $\boldsymbol{u}^{\mathrm{Rx}}$ occurs in these EL equations, with ν_Ω arising as the Lagrange multiplier for the magnetic helicity constraint in the phase-space Lagrangian (PSL). Specifically, the EL equation arising from free variations of $\boldsymbol{u}$ is

$$\boldsymbol{u}=\boldsymbol{u}^{\mathrm{Rx}}+\boldsymbol{v},\tag{1.8}$$

so $\boldsymbol{v}$ is the relative flow, the fluid velocity relative to the fully relaxed flow velocity $\boldsymbol{u}^{\mathrm{Rx}}$.

Noting from (1.7) that $\nabla\cdot(\rho\boldsymbol{u}^{\mathrm{Rx}})=0$, we see that $\nabla\cdot(\rho\boldsymbol{u})=\nabla\cdot(\rho\boldsymbol{v})$. Thus the continuity equation holds for both $\boldsymbol{u}$ and $\boldsymbol{v}$, i.e. both flows are microscopically mass conserving. Also, $\boldsymbol{u}\times\boldsymbol{B}=\boldsymbol{v}\times\boldsymbol{B}$, so $\boldsymbol{E}'[\boldsymbol{u}]=\boldsymbol{E}'[\boldsymbol{v}]$. In order to preserve (1.8) in the variational formulation (see later), the version of the IOL constraint we shall be using in the body of this paper is $\boldsymbol{E}'[\boldsymbol{v}]=0$, which becomes equivalent to $\boldsymbol{E}'[\boldsymbol{u}]=0$ only after the EL equations are derived.

1.5. *Domains and boundaries*

For most purposes in this paper it is sufficient to restrict attention to plasma within a single domain $\Omega(t)$ that is closed, of genus at least 1, and whose boundary $\partial\Omega(t)$ is smooth, gapless, perfectly conducting and time dependent. However, we note this is part of a larger project, the development of MRxMHD (Dewar *et al.* 2015), in which Ω is but a subregion of a larger plasma region, partitioned into multiple relaxation domains physically separated by moving interfaces. As $\partial\Omega(t)$ is the union of the inward-facing sides of the interfaces $\Omega(t)$ shares with its neighbours, it transmits external forcing to the restricted subsystem within $\Omega(t)$ and imparts equal and opposite reaction forces on the neighbouring subdomains.

We take the interfaces to be perfectly flexible and impervious to mass and heat transport. We also take them to be impervious to magnetic flux like a superconductor, implying the tangentiality condition

$$\boldsymbol{n}\cdot\boldsymbol{B}=0\text{ on }\partial\Omega,\tag{1.9}$$

where $\boldsymbol{B}\stackrel{\mathrm{def}}{=}\nabla\times\boldsymbol{A}$ is the magnetic field and $\boldsymbol{n}$ is a unit normal at each point on $\partial\Omega$ (here and henceforth leaving the argument t implicit in Ω, $\boldsymbol{n}$, etc.). Also, to conserve magnetic fluxes trapped within Ω, loop integrals of the vector potential $\boldsymbol{A}$ within the interfaces must be conserved (see e.g. Dewar *et al.* 2015).

1.6. *Layout of this paper*

The PSL variational approach to deriving IMHD equations is briefly reviewed in § 2, then some general implications of the IOL when it is a hard constraint are discussed in § 3

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

including speculations in § 3.2 on the implications of chaos and ergodic theory on flows in 3-D systems, and in § 3.1 the $\boldsymbol{E}\times\boldsymbol{B}$ drift is derived.

In § 4 the adaptation of the augmented Lagrangian penalty function method from optimization theory to the physical purpose of approximating the IOL constraint is discussed as a softly constrained optimization problem in § 4.1.1, and the specific Lagrangian density constraint term for this method is given in § 4.1. The entropy, magnetic helicity and cross-helicity conservation constraints used in RxMHD theory are discussed in § 4.2, and the complete PSL to be used in this paper is constructed in § 4.3.

In § 5 the EL equations, including an equation of motion in momentum conservation form, are derived formally in § 5.1, and in specific forms in §§ 5.2–5.7 where the IOL constraint term provides new contributions that vanish only when the constraint is satisfied. In addition to the momentum equation form, an equation of motion in Bernoulli form is derived. A physical interpretation of the Lagrange multiplier for the IOL constraint in terms of a polarization field is also mentioned.

Section 6 illustrates the implementation of the augmented Lagrangian method for linear waves propagating on an IOL-compliant equilibrium in the Wentzel–Kramers–Brillouin (WKB) approximation. A continuous family of dispersion relations for wave residuals $\tilde{C}$ ranging from zero in the IMHD case to its value in the RxMHD case, where it is the perturbed Lagrange multiplier $\tilde{\boldsymbol{\lambda}}$ that is set to zero.

The conclusion, § 7, briefly summarizes what has been achieved in this paper and what more needs to be done. More detail on derivations of equations is available as online supplementary material which are available at https://doi.org/10.1017/S0022377821001355 in an unabridged version of this paper.

A brief historical overview of MHD relaxation theory is given in Appendix A and some useful vector and dyadic calculus identities are derived in Appendix B; in particular the little-known identity (B7), which is crucial for getting the general form of the momentum equation (5.8) into a general conservation form, (5.12).

2. Ideal MHD in phase space

The mathematical foundation on which our dynamical relaxation formalism is built is a non-canonical form (which we call the $\boldsymbol{u}$, $\boldsymbol{v}$ picture) of the canonical MHD Hamiltonian, and a PSL. Here we review how Hamilton's action principle leads to IMHD when microscopic constraints on entropy and magnetic flux are applied. Later we show how RxMHD arises when these are replaced by global constraints using the same PSL formalism.

Both IMHD and RxMHD start from the canonical MHD Hamiltonian

$$\left.\begin{aligned} H^{\mathrm{MHD}}[\boldsymbol{x},\boldsymbol{\pi},t] &= \int_{\Omega}\mathcal{H}^{\mathrm{MHD}}\,\mathrm{d}V,\\ \text{with } \mathcal{H}^{\mathrm{MHD}}(\boldsymbol{x},\boldsymbol{\pi},t) &\stackrel{\mathrm{def}}{=} \frac{\boldsymbol{\pi}^2}{2\rho}+\frac{p}{\gamma-1}+\frac{B^2}{2\mu_0}, \end{aligned}\right\}, \tag{2.1}$$

where $\boldsymbol{\pi}(\boldsymbol{x},t)$ is the canonical momentum density, the analogue of p in finite-dimensional classical dynamics.

The analogue of q is not $\boldsymbol{x}$ the Eulerian independent variable but $\boldsymbol{r}$, the Lagrangian position with respect to a given reference frame. We do not make this explicit as we shall always work in the Eulerian picture, but the Lagrangian picture in the background does manifest in interpreting variations. (This is discussed in more detail by Dewar *et al.* (2020).) For instance, the analogue of the variation δq at fixed t is $\Delta\boldsymbol{x}=\boldsymbol{\xi}(\boldsymbol{x},t)$, the Lagrangian fluid displacement in Eulerian representation, and the analogue of the

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

variation $\dot{q}\delta t$ is $\boldsymbol{v}(\boldsymbol{x},t)\delta t$, which we shall refer to as the Lagrangian velocity field (not always the same as the Eulerian velocity $\boldsymbol{u}$). Both IMHD and RxMHD also use the constrained kinematic variation Newcomb (1962),

$$\delta\boldsymbol{v} = \partial_t\boldsymbol{\xi} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{\xi} - \boldsymbol{\xi}\cdot\boldsymbol{\nabla}\boldsymbol{v}. \tag{2.2}$$

They also use the mass density variation

$$\delta\rho = -\boldsymbol{\nabla}\cdot(\rho\boldsymbol{\xi}) = -\rho\boldsymbol{\nabla}\cdot\boldsymbol{\xi} - \boldsymbol{\xi}\cdot\boldsymbol{\nabla}\rho, \tag{2.3}$$

which is an expression of the microscopic conservation of mass and can be found by integrating the perturbed continuity equation

$$\partial_t\rho + \boldsymbol{\nabla}\cdot(\rho\boldsymbol{v}) = 0 \Leftrightarrow \frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho\boldsymbol{\nabla}\cdot\boldsymbol{v}, \tag{2.4}$$

along varied Lagrangian trajectories $\boldsymbol{r}(t|\boldsymbol{x}_0)$ (Frieman & Rotenberg 1960) and expressing this Lagrangian variation in the Eulerian picture (Newcomb 1962).

Instead of seeking a Poisson bracket to get phase-space dynamics from H (see e.g. Morrison 1998), we instead work directly with the canonical PSL density $\mathcal{L}^{\mathrm{MHD}}$,

$$\begin{aligned} L^{\mathrm{MHD}}_{\mathrm{ph}}[\boldsymbol{x},\boldsymbol{v},\boldsymbol{\pi}] &= \int_\Omega [\boldsymbol{\pi}\cdot\boldsymbol{v} - \mathcal{H}^{\mathrm{MHD}}(\boldsymbol{x},\boldsymbol{\pi},t)]\,\mathrm{d}V, \\ &\stackrel{\mathrm{def}}{=} \int_\Omega \mathcal{L}^{\mathrm{MHD}}(\boldsymbol{x},\boldsymbol{\pi},t)\,\mathrm{d}V \\ &= \int_\Omega \left(\boldsymbol{\pi}\cdot\boldsymbol{v} - \frac{\boldsymbol{\pi}^2}{2\rho} - \frac{p}{\gamma-1} - \frac{B^2}{2\mu_0}\right)\mathrm{d}V \end{aligned} \tag{2.5}$$

and the corresponding canonical phase-space action

$$\mathcal{S}^{\mathrm{MHD}}_{\mathrm{ph}} \stackrel{\mathrm{def}}{=} \iint_\Omega \mathcal{L}^{\mathrm{MHD}}\,\mathrm{d}V\,\mathrm{d}t, \tag{2.6}$$

as the primary tools, deriving EL equations from Hamilton's principle of stationary action

$$\delta\mathcal{S}^{\mathrm{MHD}}_{\mathrm{ph}} = 0, \tag{2.7}$$

varying phase-space paths under appropriate constraints.

We have used the subscript notation $\cdot_{\mathrm{ph}}$ on the Lagrangian L_{ph} and the action $\mathcal{S}_{\mathrm{ph}}$ to make it clear the PSL defined in (2.5) is fundamentally different from the more usual configuration space Lagrangian and action. This is because, in (2.6), $\boldsymbol{\pi}$ is now regarded as freely variable, so the dimensionality of the space of allowed variations is doubled in the phase-space action principle.

For instance, varying $\boldsymbol{\pi}$ in (2.5) gives the $\delta\boldsymbol{\pi}$-EL equation $\delta\mathcal{S}_{\mathrm{ph}}/\delta\boldsymbol{\pi} = \boldsymbol{v} - \boldsymbol{\pi}/\rho = 0$, i.e. multiplying by ρ, the analogue of $p = m\dot{q}$ is seen to be $\boldsymbol{\pi} = \rho\boldsymbol{v}$, as expected. Likewise, using the microscopic holonomic constraints of entropy and flux, $\delta p = -\gamma p\boldsymbol{\nabla}\cdot\boldsymbol{\xi} - \boldsymbol{\xi}\cdot\boldsymbol{\nabla}p$ and $\delta\boldsymbol{B} = \boldsymbol{\nabla}\times(\boldsymbol{\xi}\times\boldsymbol{B})$, respectively, one can verify that the EL equation arising from Lagrange-varying $\boldsymbol{x}$ (i.e. varying $\boldsymbol{\xi}$) is just the IMHD equation of motion.

However, as it is not customary in fluid mechanics to work with canonical momenta, we follow Burby (2017) in exploiting the freedom afforded by the PSL to work with a velocity-like phase-space variable $\boldsymbol{u}$, obtained by the non-canonical change of variable

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

$\boldsymbol{\pi} = \rho \boldsymbol{u}$. Then the canonical Hamiltonian density (2.1) becomes the non-canonical Hamiltonian density

$$\mathcal{H}^{\rm MHD}_{\rm nc}(\boldsymbol{x}, \boldsymbol{u}, t) = \frac{\rho \boldsymbol{u}^2}{2} + \frac{p}{\gamma - 1} + \frac{B^2}{2\mu_0} \tag{2.8}$$

and the PSL density in non-canonical form becomes, from (2.5),

$$\begin{aligned}\mathcal{L}^{\rm MHD}_{\rm nc}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, t) &\stackrel{\rm def}{=} \rho \boldsymbol{u} \cdot \boldsymbol{v} - \mathcal{H}^{\rm MHD}_{\rm nc}(\boldsymbol{x}, \boldsymbol{u}, t) \\ &= \rho \boldsymbol{u} \cdot \boldsymbol{v} - \frac{\rho \boldsymbol{u}^2}{2} - \frac{p}{\gamma - 1} - \frac{B^2}{2\mu_0}.\end{aligned} \tag{2.9}$$

As neither p nor $\boldsymbol{B}$ depends on $\boldsymbol{u}$, the $\delta\boldsymbol{u}$-EL equation is $\delta \mathbb{S}^{\rm MHD}_{\rm ph}/\delta \boldsymbol{u} = \rho \boldsymbol{v} - \rho \boldsymbol{u} = 0$, i.e. in IMHD we have $\boldsymbol{v} = \boldsymbol{u}$. The IMHD equation of motion, which can be written in conservation form as

$$\partial_t(\rho \boldsymbol{u}) + \nabla \cdot \boldsymbol{T}_{\rm MHD} = 0, \tag{2.10}$$

where

$$\boldsymbol{T}_{\rm MHD} \stackrel{\rm def}{=} \rho \boldsymbol{u}\boldsymbol{u} + \left(p + \frac{B^2}{2\mu_0}\right)\boldsymbol{I} - \frac{\boldsymbol{B}\boldsymbol{B}}{\mu_0}, \tag{2.11}$$

follows as it does for the $\boldsymbol{\xi}$-EL equation in the canonical form.

It will be shown below that an isothermal version of IMHD can be derived by replacing the holonomic variational constraint $\delta p = -\gamma p \nabla \cdot \boldsymbol{\xi} - \boldsymbol{\xi} \cdot \nabla p$ with a global entropy conservation constraint, giving thermal relaxation, a more realistic model for hot plasmas than the microscopic entropy constraints implied by $\delta p = -\gamma p \nabla \cdot \boldsymbol{\xi} - \boldsymbol{\xi} \cdot \nabla p$.

3. Implications of the IOL constraint

In this section we examine consequences of applying the IOL constraint (1.2) in the form (see § 1.4) $\boldsymbol{E}'[\boldsymbol{v}] = 0$, which can be written $-\boldsymbol{E} = \boldsymbol{v} \times \boldsymbol{B}$ or $\partial_t \boldsymbol{A} + \nabla \Phi = \boldsymbol{v} \times \boldsymbol{B}$.

3.1. $\boldsymbol{E} \times \boldsymbol{B}$ *drift*

As $\boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}$ we have the two identities

$$\boldsymbol{B} \times \boldsymbol{E}' = -\boldsymbol{E} \times \boldsymbol{B} + (B^2 \boldsymbol{I} - \boldsymbol{B}\boldsymbol{B}) \cdot \boldsymbol{v} \tag{3.1}$$

$$\text{and} \quad \boldsymbol{v} \times \boldsymbol{E}' = -\boldsymbol{E} \times \boldsymbol{v} - (v^2 \boldsymbol{I} - \boldsymbol{v}\boldsymbol{v}) \cdot \boldsymbol{B}. \tag{3.2}$$

Equation (3.1) leads to a decomposition of the relative fluid flow into a component $v_\parallel$ tangential to $\boldsymbol{B}$ at $\boldsymbol{x}$, and a component $\boldsymbol{v}_\perp$, its projection onto the plane transverse to $\boldsymbol{B}$, the '$\boldsymbol{E} \times \boldsymbol{B}$ drift',

$$\boldsymbol{v}_\perp = \frac{\boldsymbol{E}_\perp \times \boldsymbol{B}}{B^2}. \tag{3.3}$$

It is usually safe to assume $B^2 \neq 0$ anywhere in toroidally confined plasmas, so the representation (3.3) generally applies everywhere, and to both equilibrium and dynamic MHD cases.[2]

[2]The case $u^2 = 0$ may well occur in plasma containment devices, so using (3.2) to make a decomposition of $\boldsymbol{B}$ in terms of $\boldsymbol{u}$ analogous to the reverse in (3.3) seems less useful.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

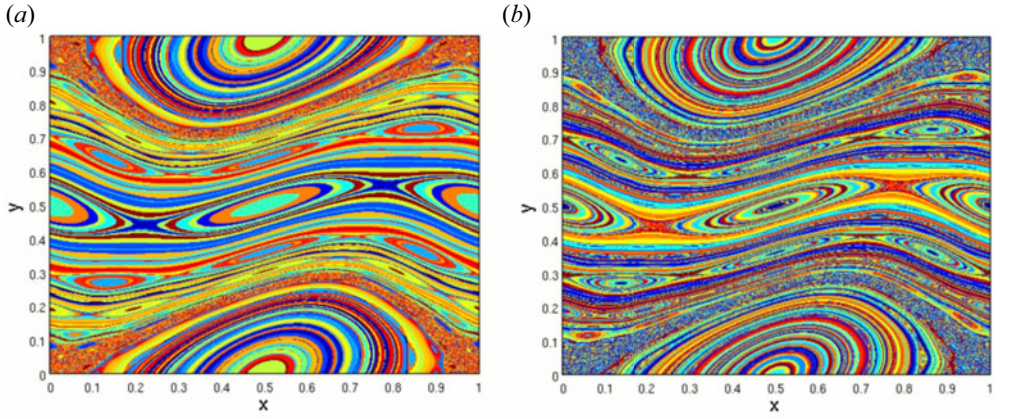


FIGURE 1. Ergodic partitions of iterates of the standard map as depicted in figure 11 of Levnajić & Mezić (2010). (Reprinted with permission from *Chaos*.)

3.2. *The equilibrium ergodicity problem*

Resolving the IOL onto the vectors $\boldsymbol{B}$ and $\boldsymbol{v}$ (or $\boldsymbol{u}$) eliminates the $\boldsymbol{v}\times\boldsymbol{B}$ term, so in equilibrium, when $\partial_t \boldsymbol{A} = 0$, these components of the IOL imply

$$\boldsymbol{B}\cdot\nabla\Phi = 0, \tag{3.4}$$

$$\boldsymbol{u}\cdot\nabla\Phi = 0. \tag{3.5}$$

This means $\Phi = $ const. on stream lines as well as magnetic field lines. As a consequence, level sets of Φ are invariant under magnetic and fluid flow. For instance, if Φ has smoothly nested level surfaces in a region then both $\boldsymbol{u}$ and $\boldsymbol{B}$ lie in the local tangent plane at each point on each isopotential surface – the magnetic and fluid flows are both locally integrable.

In the opposite extreme, Finn & Antonsen (1983) (after (29)) conclude from the constancy of Φ along a field line that 'if the turbulent relaxation has ergodic field lines throughout the plasma volume', then $\nabla\Phi = 0$, which implies that $\boldsymbol{u}\times\boldsymbol{B} = 0$ – the fluid flows along magnetic field lines. As already mentioned, we call such field-aligned steady flows fully relaxed equilibria (though the converse does not apply – field-aligned flows can be integrable).

However, field-aligned flow equilibria exclude many applications of physical interest – in particular tokamaks with strong toroidal flow. For such axisymmetric equilibria, Dewar *et al.* (2020) show RxMHD can give the same axisymmetric relaxed solutions with cross-field flow as found by Finn & Antonsen (1983) and Hameiri (1983), but without needing the angular momentum constraint used by these authors.

Unlike Finn & Antonsen (1983) we are not appealing to turbulence to justify relaxation, but, in fully 3-D plasmas, we may be able to appeal to the existence of chaotic magnetic field and stream lines. However, 'chaotic' is not the same as 'ergodic' – while chaotic flows do involve ergodicity, this is in an infinitely complicated way, visualized in figure 1 in terms of the fractal ergodic partition of Mezić & Wiggins (1999) and Levnajić & Mezić (2010). (This figure is generated for an iterated area-preserving map, but magnetic field-line flows being flux preserving, the magnetic field-line return map of a Poincaré section onto itself in a magnetic containment device is similar.)

A similar problem involving chaos and ergodicity arises in magnetohydrostatics (Hudson *et al.* 2012), where the equilibrium condition $\nabla p = \boldsymbol{j}\times\boldsymbol{B}$ implies $\boldsymbol{B}\cdot\nabla p = 0$,

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

analogously to (3.4) for Φ, so the fractal ergodic partition for field-line flow is as relevant to the pressure p as it is for the potential Φ. In their MRxMHD equilibrium code, Hudson *et al.* (2012) solved the puzzle posed by Grad (1967) (i.e. how to formulate the 3-D IMHD equilibrium problem so as to avoid a 'pathological' pressure profile) by using a much simpler ergodic partition obtained by aggregating contiguous elements of the fractal ergodic partition into a finite number of constant-pressure 'relaxation regions' Ω_i, with pressure changing (discontinuously) only across the interfaces between the Ω_is. The code was thus named the stepped pressure equilibrium code (SPEC).

3.2.1. *Continuity of electrostatic potential*

One might think that an analogous 'stepped potential equilibrium' could provide a solution to the problem of finding a non-trivial but tractable solution of $\nabla_\parallel \Phi \stackrel{\text{def}}{=} (\boldsymbol{B}/B)\cdot\nabla\Phi = 0$ in a chaotic magnetic-field-line flow. Unfortunately, however, we must restrict $\nabla_\perp \Phi \stackrel{\text{def}}{=} (\boldsymbol{I} - \boldsymbol{BB}/B^2)\cdot\nabla\Phi$ to square-integrable functions in order to keep the $\boldsymbol{E}\times\boldsymbol{B}$ drift (3.3) from acquiring a δ-function component.

This rules out having steps in Φ because δ-functions are not square integrable, so stepped potentials would make the kinetic energy integral infinite. However, this does not necessarily imply Φ is constant in weakly chaotic regions with a finite measure of Kolmogorov–Arnold–Moser (KAM) surfaces – perhaps weak KAM theory (see e.g. Fathi 2009) would allow fractal potential profiles having finite kinetic energy associated with them.

As a way to handle non-constant Φ computationally we propose using a penalty or augmented Lagrangian method (see e.g. Nocedal & Wright 2006). That is, we treat Hamilton's principle as a constrained saddle-point optimization problem and add a penalty functional to the Hamiltonian, which regularizes the variational problem by approaching the (perhaps fractal) IMHD 'feasible region' of configuration space from outside, in the less-constrained space on which RxMHD is defined (which is smoother, see figure 1 of Dewar *et al.* (2020)).

Another approach might be a time-evolution code with added dissipation such that the long-time solution is attracted to one having chaotic regions of constant pressure interspersed with integrable regions with changing pressure. This can be viewed as a steepest-descents solution of the same optimization problem.

4. Constraints and constrained optimization

In this section we first discuss the new aspect of variational relaxation theory introduced in this paper, namely the imposition of IOL as a constraint.

We then review use the Lagrange multiplier method in § 4.3 for imposing the conservation of entropy, magnetic helicity and cross-helicity as hard constraints, causing the EL equations, and hence the conserved quantities, to be parameterized by the triplet of multipliers τ_Ω, μ_Ω and ν_Ω (the subscripts Ω indicating they are constant throughout Ω, but may jump across $\partial\Omega$ if there are adjacent relaxation regions as in MRxMHD).

4.1. *Augmented Lagrangian constraint method*

In implementing the IOL constraint we propose to adapt the augmented Lagrangian method from finite-dimensional optimization theory, as described by Nocedal & Wright (2006, § 17.3), or for Banach spaces (see e.g. Kanzow, Steck & Wachsmuth (2018) and references therein). This is a hybrid numerical method that combines two constraint approaches: the Lagrange multiplier method and the penalty function method. We shall use the Lagrange multiplier method in § 4.3 for imposing the conservation of entropy,

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

magnetic helicity and cross-helicity as hard constraints, causing the EL equations, and hence the conserved quantities, to be parameterized by the triplet of multipliers τ_Ω, μ_Ω and ν_Ω (the subscripts Ω indicating they are constant throughout Ω, but may jump across $\partial\Omega$ if there are adjacent relaxation regions as in MRxMHD).

To impose the IOL as a hard constraint using the Lagrange multiplier method we would 'simply' add $\boldsymbol{\lambda}\cdot(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B})$ to the Lagrangian density, solve the resultant EL equations to give $\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B}$ as a function of the Lagrange multiplier $\boldsymbol{\lambda}$, and then solve for $\boldsymbol{\lambda}$ such that $\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B}=0$.

Apart from the unavoidable complication that $\boldsymbol{\lambda}$ is not just a 3-vector but also is a function of $\boldsymbol{x}$ and t, so infinite-dimensional, there is the more fundamental problem, flagged in § (3.2), that the limit $\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B}\to 0$ is likely singular in 3-D equilibria because $\boldsymbol{E}$, $\boldsymbol{v}$ and $\boldsymbol{B}$ presumably tend towards being fractal functions. Thus there is good reason to believe the hard IOL constraint problem is ill-posed in 3-D systems such as stellarators, which leads us to seek a soft IOL constraint approach in order to regularize the Hamilton's principle optimization problem.

We build in the Maxwell–Faraday induction constraint (1.3) as a hard constraint by using the potential representations (1.6), $\boldsymbol{E}=-\nabla\Phi-\partial_t\boldsymbol{A}$ and (1.5), $\boldsymbol{B}=\nabla\times\boldsymbol{A}$. Thus the set of primary variables subject to variation during an optimization is

$$X=\{\boldsymbol{r},\boldsymbol{u},p,\boldsymbol{A},\Phi\},\tag{4.1}$$

where $\boldsymbol{r}$ is the Lagrangian fluid-element position field discussed in § 2. (Note we have not included ρ and $\boldsymbol{v}$ as independent variables because they are functionals of $\boldsymbol{r}$, with variations given by (2.3) and (2.2).)

The simplest soft IOL constraint approach is to add $\frac{1}{2}\mu_\Omega^P(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B})^2$ to the Hamiltonian density (thus subtracting it from our Lagrangian density), where $\mu_\Omega^P\to+\infty$ is a penalty multiplier. In this limit the penalty term is supposed to dominate all other terms in the Hamiltonian or Lagrangian and enforce IOL feasibility through a sequence of infeasible solutions. However, this method is clearly ill-conditioned numerically, leading us to resort to the 'best of both worlds' augmented Lagrangian method described below.

4.1.1. *The IOL as a softly constrained optimization problem*

In implementing the parallel IOL constraint we propose to adapt the augmented Lagrangian method from finite-dimensional optimization theory, as described by Nocedal & Wright (2006, § 17.3), or for Banach spaces (see e.g. Kanzow *et al.* (2018) and references therein). This is a hybrid numerical method that combines two constraint approaches: the Lagrange multiplier method and the penalty function method sketched above.

As well as adapting notation and methods from optimization theory we have borrowed the term 'feasible region', meaning the range in which the vector X of variables to be solved for is such that a set of equality constraints $c_i[X]=0$ are satisfied (also inequality constraints $c_i[X]>0$, but we do not consider this case). The infeasible region, is its complement, where one or more constraints are violated. By hard constraint we mean one where X must be in the feasible region, and by soft constraint we mean one where X need only be in some neighbourhood of the feasible region, which is useful both practically and for regularizing when, as in MHD, defining the boundary between feasible and infeasible is complicated by the possibility of singular behaviour like current sheets and reconnection points.

We now formulate two related physical tasks, the simpler one being (task 1) the equilibrium problem. In toroidal plasma confinement theory the most physically desirable states are stable, time-independent equilibria, i.e. minima of a Hamiltonian functional

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

$\mathcal{H}[X]$, kinetic plus potential energy within a static boundary $\partial\Omega$. We seek a numerical algorithm that starts from an initial guess for the physical fields X and iterates to extremize (minimize if seeking a stable equilibrium) a Hamiltonian, under the IOL equality constraint, (1.2). Finding a stable equilibrium can be stated concisely as the optimization problem.

$$\textit{Equilibrium}: \quad \min_X \mathcal{H}[X] \quad \text{subject to } \boldsymbol{C}[X](\boldsymbol{x})] = 0, \\ \forall \boldsymbol{x} \in \Omega \text{ and boundary conditions} \quad \forall \boldsymbol{x} \in \partial\Omega, \tag{4.2}$$

where

$$\boldsymbol{C} \stackrel{\text{def}}{=} \boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B}, \tag{4.3}$$

is the generalization of the finite set of equality constraint functions $\{c_i\}$ of Nocedal & Wright (2006) (as a 3-vector it is finite-dimensional but as a function of $\boldsymbol{x}$ it is infinite-dimensional).

For the purposes of the present paper $\mathcal{H}$ is the non-canonical version, $H^{\text{MHD}}_{\text{nc}}$, of the Hamiltonian, H^{MHD} defined in (2.1) plus the global constraint terms described in the next subsection, § 4.2. The ideal boundary conditions are $\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B} = 0$, $\boldsymbol{n}\cdot\boldsymbol{v} = 0$, $\boldsymbol{n}\cdot\boldsymbol{B} = 0$ on $\partial\Omega$ and $\Phi = \text{const.}$ on each disjoint component of $\partial\Omega$ (think plates of a capacitor or electrodes of a vacuum tube).

To treat the implementation of the IOL in Hamilton's principle, a constrained saddle point optimization problem, we shall use the set of values of the components of $\boldsymbol{E}'(\boldsymbol{x} \in \Omega)$; note the identities

$$\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{E}} = \boldsymbol{I}, \quad \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{v}} = \boldsymbol{I}\times\boldsymbol{B} \quad \text{and} \quad \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{B}} = -\boldsymbol{I}\times\boldsymbol{v}. \tag{4.4a–c}$$

We seek a soft form of the equilibrium constraint, i.e. a formulation such that $\boldsymbol{C} \to 0$, $\forall \boldsymbol{x} \in \Omega$, where $\to$ denotes a limiting process whereby X moves from the infeasible class of states where $\boldsymbol{C} \neq 0$ towards the feasible class defined pointwise as $\boldsymbol{C} = 0$, $\forall \boldsymbol{x} \in \Omega(t)$, or, in a weak form, as $||\boldsymbol{C}|| = 0$. Such a soft constraint procedure is provided by the augmented Lagrangian (or, rather, Hamiltonian in the equilibrium problem) as prescribed by Nocedal & Wright (2006, § 17.3),

$$\mathcal{H}_A \stackrel{\text{def}}{=} \int_\Omega \left[\mathcal{H} - \boldsymbol{\lambda}\cdot\boldsymbol{C} + \tfrac{1}{2}\mu^P_\Omega \boldsymbol{C}^2\right] \mathrm{d}V, \tag{4.5}$$

where $\boldsymbol{\lambda}$ is a Lagrange multiplier and the spatial constant $\mu^P_\Omega \geq 0$ is a penalty multiplier of the non-negative quadratic penalty $\frac{1}{2}\int_\Omega \boldsymbol{C}^2\,\mathrm{d}V$.

Nocedal & Wright (2006, § 17.3) give an iterative algorithmic framework that combines the advantages of both the Lagrange multiplier and penalty function methods. In their algorithm the user provides an increasing sequence $\{\mu^P_\Omega|^n, : n = 0, 1, 2, \ldots\}$ of penalty multipliers and adjusts $\boldsymbol{\lambda}$ to solve for the minima of Hamiltonians with the Lagrange multiplier and penalty terms. The iteration update rule for the sequence of Lagrange multipliers and corresponding constraint residuals $\{\boldsymbol{\lambda}^n, \boldsymbol{C}^n : n = 0, 1, 2, \ldots\}$ is

$$\boldsymbol{\lambda}^{n+1} = \boldsymbol{\lambda}^n - \mu^P_\Omega|^n \boldsymbol{C}^n. \tag{4.6}$$

In the following sections we shall take the iteration index n as implicit unless needed for clarity, with the updated $\boldsymbol{\lambda}$ as given by the right-hand side of (4.6) denoted by

$$\boldsymbol{\lambda}_* \stackrel{\text{def}}{=} \frac{\partial}{\partial \boldsymbol{C}}\left(\boldsymbol{\lambda}\cdot\boldsymbol{C} - \frac{\mu^P_\Omega}{2}\boldsymbol{C}^2\right) = \boldsymbol{\lambda} - \mu^P_\Omega \boldsymbol{C}, \tag{4.7}$$

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

which is a 'best estimate' of the optimum Lagrange multiplier given the current estimate and penalty multiplier.

The more difficult second physical task (task 2) is the time evolution problem. This is similar to task 1 except we seek an evolution, a dynamical path in space-time $\Omega_t \times [t_1, t_2]$ given a time-dependent boundary $\partial\Omega_t$, the objective function for extremization now being an action integral. The stable minima of the Hamiltonian now becoming saddle points of the corresponding action functional $\mathcal{S}[X]$, kinetic minus potential energy. This task can be summarized as the pseudo optimization problem.

$$\textit{Dynamics}: \quad \underset{X}{\text{extr}}\, \mathcal{S}[X] \quad \text{subject to } \boldsymbol{C}[X](\boldsymbol{x}, t) = 0, \forall\, \boldsymbol{x} \in \Omega_t,\ t \in [t_1, t_2], \tag{4.8}$$

under the same boundary conditions as for equilibrium at each time t.

Although sometimes called the 'principle of least action', Hamilton's principle is often not an optimization problem but rather a saddle point problem, where the stationary point of $\mathcal{S}[X]$ cannot be found by a descent algorithm. (This is well known in nonlinear Hamiltonian dynamics (Meiss 1992) where periodic orbits are classified as (action) minimizing orbits, which are hyperbolic (unstable), or as minimax orbits, which are elliptic (stable).) Although 'extremum' or 'extremization' is not quite correct either, as extremum strictly means 'maximum or minimum', it is convenient to use the abbreviation 'extr' as an abbreviation for these words and add the rider 'depending on direction of traversal' (implying also the existence of neutral directions between max and min), so as to include saddle points.

To find saddle points requires some form of Newton method, needing at least estimates of the second variation (Hessian matrix) rather than a descent method. The augmented Lagrangian method still works if we solve (4.8) at each iteration, so here again we adopt it to solve for a stationary point of the augmented phase-space action functional (2.9),

$$\mathcal{S}^{\mathrm{A}}_{\mathrm{ph}} \overset{\text{def}}{=} \iint_{\Omega} (\rho \boldsymbol{u} \cdot \boldsymbol{v} - \mathcal{H} + \mathcal{L}^{C}_{\Omega}) \mathrm{d}V \mathrm{d}t, \tag{4.9}$$

where the augmented penalty constraint density $\mathcal{L}^{C}_{\Omega}$ is defined by

$$\mathcal{L}^{C}_{\Omega} \overset{\text{def}}{=} \boldsymbol{\lambda} \cdot \boldsymbol{C} - \frac{\mu^{P}_{\Omega} \boldsymbol{C}^2}{2}, \tag{4.10}$$

where $\boldsymbol{\lambda}$ and μ^{P}_{Ω} are taken as external parameters in the application of Hamilton's principle at each iteration, giving a sequence of regularized magnetofluid models. When $\mu^{P}_{\Omega} = 0$, the pure Lagrange multiplier method, feasible critical points of $\mathcal{H}_A$ might be saddle points with descending directions in the infeasible sector even if they are physically stable ideal equilibria where the IMHD Hamiltonian is minimized. When $\boldsymbol{\lambda} = 0$, the pure penalty function method, feasible stable equilibria could be approximated arbitrarily well in the limit as μ^{P}_{Ω} tends to infinity, but this becomes an increasingly ill-posed optimization problem. (It does, however, have the attractive feature of providing a continuous family of RxMHD models running from the RxMHD of Dewar *et al.* (2020) when $\mu^{P}_{\Omega} = 0$ to a subset of weak IMHD when $\mu^{P}_{\Omega} \to +\infty$.)

Remarks:

(i) task 1 can be treated as a subclass of task 2 in which time derivatives are set to zero and t is taken to be an irrelevant constant, but the Hamiltonian is more appropriate than the Lagrangian for treating it as an optimization problem;

(ii) the iteration method for implementing constraints is implicit, meaning that the state variables in the nth iteration need to be found by solving EL equations, taking it for granted the EL equations can be solved and any subiterations required have converged.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

We shall not discuss detailed implementation issues here, except to remark that time evolution over a large time interval can be implemented numerically in an outer time stepping loop in which a large time interval is split into multiple short time intervals (time steps) $[t_i, t_{i+1}]$, within each of which constraint iterations are repeated until converged to the required accuracy. Thus the evolutions required in implementing the constraint iterations are over short time intervals, with each initial guess being the converged evolution from the previous time step and the evolution representable to sufficient accuracy on a low-dimensional interpolation basis (e.g. dimension two for piecewise-linear representation of the full evolution) – the increase in difficulty in going from task 1 to task 2 may not be as great as at first it appears to be.

4.2. *Global constraints for isothermal RxMHD and IMHD*

We shall always retain the microscopic holonomic constraints (§ 2) on ρ and $\boldsymbol{v}$, but we relax the infinite number of microscopic dynamical constraints on p and $\boldsymbol{B}$ imposed in IMHD by replacing these constraints with only three macroscopic hard constraints. These three constraints, described below, are chosen to be quantities that are exact invariants under IMHD dynamics in order to ensure that relaxed equilibria are subsets of all ideal equilibria. Further, as we seek plasma relaxation formalisms applicable in arbitrary 3-D toroidal geometries, we invoke only the MHD invariants least dependent on integrability of the fluid and magnetic field line flows, the conservation of total mass $M_\Omega \stackrel{\text{def}}{=} \int_\Omega \rho \, \mathrm{d}V$ being the most fundamental (whose conservation is built in microscopically). While these global invariants are not as well conserved as mass under small resistive, viscous and 3-D chaos effects, in the spirit of Taylor (1986) we assume they are sufficiently robust that postulating their conservation produces a model that is useful in appropriate applications.

We can get IMHD by retaining all the microscopic holonomic constraints of § 2, but it seems more physically relevant to almost collisionless hot plasmas with high thermal conductivity along magnetic field lines to relax the plasma thermally by relaxing the microscopic dynamical constraint on p and replacing it with the first global constraint below (entropy) to give isothermal IMHD.

As just indicated, our first global constraint is the adiabatic-ideal-gas thermodynamic invariant, total entropy

$$S_\Omega[\rho, p] \stackrel{\text{def}}{=} \int_\Omega \frac{\rho}{\gamma - 1} \ln\left(\kappa \frac{p}{\rho^\gamma}\right) \mathrm{d}V, \tag{4.11}$$

where κ is, for our purposes, an arbitrary dimensionalizing constant, though it can be identified physically through a statistical mechanical derivation of (4.11) (see e.g. Dewar *et al.* 2015). Its functional derivatives are

$$\frac{\delta S_\Omega}{\delta \rho} = \frac{1}{\gamma - 1} \ln\left(\kappa \frac{p}{\rho^\gamma}\right) - \frac{\gamma}{\gamma - 1}, \tag{4.12}$$

$$\frac{\delta S_\Omega}{\delta p} = \frac{1}{\gamma - 1} \frac{\rho}{p}. \tag{4.13}$$

We also impose conservation of the magnetic helicity $2\mu_0 K_\Omega$, where, following Bhattacharjee & Dewar (1982), we define the invariant K_Ω as

$$K_\Omega[\boldsymbol{A}] \stackrel{\text{def}}{=} \frac{1}{2\mu_0} \int_\Omega \boldsymbol{A} \cdot \boldsymbol{B} \, \mathrm{d}V, \tag{4.14}$$

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

giving, with help of (B2), the functional derivative

$$\frac{\delta K_\Omega}{\delta \boldsymbol{A}} = \frac{\boldsymbol{B}}{\mu_0}. \tag{4.15}$$

As discussed by Hameiri (2014), in single-fluid IMHD we do not have a separate fluid helicity invariant, but do have the cross-helicity $\mu_0 K^X_\Omega$, which can be derived from a relabelling symmetry in the Lagrangian representation of the fields, see e.g. chapter seven of Webb (2018). Analogously to our other constraint parameters containing $\boldsymbol{B}$, we include μ_0^{-1} in the definition of the cross-helicity functional,

$$K^X_\Omega[\boldsymbol{u}, \boldsymbol{A}] \stackrel{\text{def}}{=} \frac{1}{\mu_0}\int_\Omega \boldsymbol{u}\cdot\boldsymbol{B}\,\mathrm{d}V, \tag{4.16}$$

which, like P_Ω and S_Ω, has two functional derivatives

$$\frac{\delta K^X_\Omega}{\delta \boldsymbol{u}} = \frac{\boldsymbol{B}}{\mu_0}, \quad \frac{\delta K^X_\Omega}{\delta \boldsymbol{B}} = \frac{\boldsymbol{u}}{\mu_0}. \tag{4.17a,b}$$

4.3. *IOL-constrained PSLs and actions*

As foreshadowed, our recipe for constructing a non-dissipative relaxed magnetofluid model is to start with the IMHD non-canonical Hamiltonian, (2.8), but to relax many, but not all, of the microscopic constraints to which it is subject when deriving the IMHD EL equations. Specifically, to retain the basic compressible Euler-fluid backbone of our RxMHD model (Dewar *et al.* 2015), we keep the microscopic kinematic and mass conservation constraints, (2.2) and (2.3).

However, we delete the microscopic ideal gas and flux-frozen magnetic field variational constraints, $\delta p = -\gamma p\boldsymbol{\nabla}\cdot\boldsymbol{\xi} - \boldsymbol{\xi}\cdot\boldsymbol{\nabla}p$ and $\delta\boldsymbol{B} = \boldsymbol{\nabla}\times(\boldsymbol{\xi}\times\boldsymbol{B})$, replacing these infinities of constraints with only the three robust IMHD global invariants (4.11)–(4.16). These global constraints are imposed by adding the global-invariants-constraint (GIC) Lagrange multiplier term

$$\mathcal{L}^{\text{GIC}}_\Omega \stackrel{\text{def}}{=} \tau_\Omega \frac{\rho}{\gamma-1}\ln\left(\kappa\frac{p}{\rho^\gamma}\right) + \mu_\Omega\frac{\boldsymbol{A}\cdot\boldsymbol{B}}{2\mu_0} + \nu_\Omega\frac{\boldsymbol{u}\cdot\boldsymbol{B}}{\mu_0}, \tag{4.18}$$

to $\mathcal{L}^{\text{MHD}}_{\text{nc}}$ to form the RxMHD PSL density (Dewar *et al.* 2020)

$$\begin{aligned}\mathcal{L}^{\text{Rx}}_\Omega \stackrel{\text{def}}{=} \mathcal{L}^{\text{MHD}}_{\text{nc}} + \mathcal{L}^{\text{GIC}}_\Omega &= \rho\boldsymbol{u}\cdot\boldsymbol{v} - \frac{\rho\boldsymbol{u}^2}{2} - \frac{p}{\gamma-1} - \frac{B^2}{2\mu_0}\\ &+ \tau_\Omega\frac{\rho}{\gamma-1}\ln\left(\kappa\frac{p}{\rho^\gamma}\right) + \mu_\Omega\frac{\boldsymbol{A}\cdot\boldsymbol{B}}{2\mu_0} + \nu_\Omega\frac{\boldsymbol{u}\cdot\boldsymbol{B}}{\mu_0}.\end{aligned} \tag{4.19}$$

In 4.18 the Lagrange multipliers τ_Ω, μ_Ω and ν_Ω are spatially constant throughout Ω, but can change in time to enforce constancy, respectively, of total entropy, magnetic helicity and cross-helicity in Ω. By removing the infinite numbers of microscopic constraints on p and $\boldsymbol{B}$ that are imposed in IMHD, in the RxMHD formalism (Dewar *et al.* 2020) we greatly increased the variationally feasible region of the state space, thus allowing the system to access a lower energy equilibrium. In fact, as the Eulerian fields $\delta p(\boldsymbol{x}, t)$ and $\delta\boldsymbol{A}(\boldsymbol{x}, t)$ are now locally free variations at each point $\boldsymbol{x}$, we have added two infinities of degrees of freedom, which turns out to be too many as the IOL constraint embedded in IMHD is entirely lost in RxMHD.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

Thus we reduce the degrees of freedom of RxMHD by imposing a soft penalty-function IOL constraint using the augmented Lagrangian constraint density $\mathcal{L}^C_\Omega$, (4.10). As the IOL constraint applies pointwise throughout Ω, it gives an infinite number of constraints, on Φ and $\boldsymbol{B}$. Adding the constraint term we get the full Lagrangian density with augmented constraint

$$\mathcal{L}^A_\Omega \stackrel{\text{def}}{=} \mathcal{L}^{\text{Rx}}_\Omega + \mathcal{L}^C_\Omega. \tag{4.20}$$

We shall also have need to define the gauge-invariant part of the Lagrangian density by subtracting the magnetic helicity term,

$$\mathcal{L}^{A_-}_\Omega \stackrel{\text{def}}{=} \mathcal{L}^A_\Omega - \mu_\Omega \frac{\boldsymbol{A}\cdot\boldsymbol{B}}{2\mu_0}. \tag{4.21}$$

(For derivatives of the Lagrangian density with respect to anything other than $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{x}$, or t, $\mathcal{L}^A_\Omega$ and $\mathcal{L}^{A_-}_\Omega$ can be used interchangeably.)

The augmented phase-space action integral is

$$\mathcal{S}^A_\Omega = \int \mathrm{d}t \int_\Omega \mathrm{d}V \mathcal{L}^A_\Omega. \tag{4.22}$$

As in (2.9), the fluid velocity $\boldsymbol{u}$ is treated as a non-canonical momentum variable that is freely variable in the phase-space version of Hamilton's principle, $\delta\mathcal{S}^A_\Omega = 0$, and $\boldsymbol{v}$ is a relative flow whose variation with respect to $\boldsymbol{\xi}$ obeys the kinematical constraint (2.2). It is also the flow appearing in the mass conservation constraint equations (2.3) and (2.4).

5. EL equations

5.1. *Formal view of EL equations*

The utility of Hamilton's action-principle approach is that a complete set of equations for our physical variables is provided by the EL equations following from the general variation of the generic augmented action $\mathcal{S}^A_\Omega$,

$$\begin{aligned}\delta\mathcal{S}^A_\Omega &= \int \mathrm{d}t \int_\Omega \mathrm{d}V \left[\delta\boldsymbol{u}\cdot\frac{\partial\mathcal{L}^A_\Omega}{\partial\boldsymbol{u}} + \delta\boldsymbol{A}\cdot\frac{\partial\mathcal{L}^A_\Omega}{\partial\boldsymbol{A}} + \delta\boldsymbol{B}\cdot\frac{\partial\mathcal{L}^A_\Omega}{\partial\boldsymbol{B}} + \delta\boldsymbol{E}\cdot\frac{\partial\mathcal{L}^A_\Omega}{\partial\boldsymbol{E}}\right.\\ &\quad\left. + \delta p\frac{\partial\mathcal{L}^A_\Omega}{\partial p} + \delta\rho\frac{\partial\mathcal{L}^A_\Omega}{\partial\rho} + \delta\boldsymbol{v}\cdot\frac{\partial\mathcal{L}^A_\Omega}{\partial\boldsymbol{v}}\right]\\ &\stackrel{\text{def}}{=} \int \mathrm{d}t \int_\Omega \mathrm{d}V \left[\delta\boldsymbol{u}\cdot\frac{\delta\mathcal{S}^A_\Omega}{\delta\boldsymbol{u}} + \delta\Phi\frac{\delta\mathcal{S}^A_\Omega}{\delta\Phi} + \delta\boldsymbol{A}\cdot\frac{\delta\mathcal{S}^A_\Omega}{\delta\boldsymbol{A}} + \delta p\frac{\delta\mathcal{S}^A_\Omega}{\delta p} + \boldsymbol{\xi}\cdot\frac{\delta\mathcal{S}^A_\Omega}{\delta\boldsymbol{x}}\right],\end{aligned} \tag{5.1}$$

where the top equation on the right-hand side is simply an integral over the first variation of $\mathcal{L}^A_\Omega$ and the second right-hand side equation defines the functional derivatives with respect to the independent variables by matching the corresponding terms in the top right-hand side equation after the variations of the explicit variables in $\mathcal{L}^A_\Omega$ are expanded and integrations by parts where necessary – ignoring boundary terms as we can assume the support of the variations does not include the boundary (note that there are no $\delta\lambda$ or $\delta\mu^P_\Omega$ terms as λ and μ^P_Ω are taken as given – see discussion around (4.10)). For instance, $\delta\mathcal{S}^A_\Omega/\delta\boldsymbol{x}$ is the sum of the terms linear in $\boldsymbol{\xi}$ obtained from $\delta\rho$ and $\delta\boldsymbol{v}$ given in (2.3) and (2.2), respectively. (Note that for notational convenience $\delta\boldsymbol{x}$ is used in the denominator of the functional derivative as an alternative to the Lagrangian variation of $\boldsymbol{x}$, denoted everywhere else as $\Delta\boldsymbol{x}$ or $\boldsymbol{\xi}$. It does not denote the Eulerian variation of $\boldsymbol{x}$, which is by definition zero.)

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

Inspecting (4.10) we see that $\mathcal{L}^C_\Omega$ contains $\boldsymbol{E}$ and $\boldsymbol{B}$ but does not contain $\boldsymbol{u}$, p or ρ, and no term in $\mathcal{L}^{\mathrm{Rx}}_\Omega$ contains $\boldsymbol{E}$, $\boldsymbol{\nabla u}$, $\boldsymbol{\nabla} p$ or $\boldsymbol{\nabla}\rho$, so the corresponding functional derivatives of $\mathcal{S}^A_\Omega$ are are simply partial derivatives of $\mathcal{L}^A_\Omega$, e.g. the $\delta\boldsymbol{u}$- and δp-EL equations are

$$\frac{\delta \mathcal{S}^A_\Omega}{\delta \boldsymbol{u}} = \frac{\partial \mathcal{L}^A_\Omega}{\partial \boldsymbol{u}} = 0, \tag{5.2}$$

$$\frac{\delta \mathcal{S}^A_\Omega}{\delta p} = \frac{\partial \mathcal{L}^A_\Omega}{\partial p} = 0. \tag{5.3}$$

The $\delta\boldsymbol{A}$-EL equation is best displayed by splitting $\mathcal{L}^A_\Omega$ into the gauge-invariant part $\mathcal{L}^{A_-}_\Omega$, (4.21), and the magnetic helicity constraint term $\mu_\Omega \boldsymbol{A}\cdot\boldsymbol{B}/2\mu_0$ in order to make manifest the explicit $\boldsymbol{A}$-dependence; thus

$$\frac{\partial \mathcal{L}^A_\Omega}{\partial \boldsymbol{A}} = \frac{\mu_\Omega \boldsymbol{B}}{2\mu_0}, \quad \frac{\partial \mathcal{L}^A_\Omega}{\partial \boldsymbol{B}} = \frac{\partial \mathcal{L}^{A_-}_\Omega}{\partial \boldsymbol{B}} + \frac{\mu_\Omega \boldsymbol{A}}{2\mu_0}. \tag{5.4a,b}$$

The $\delta\boldsymbol{A}$-EL equation is then found by using these results in the identity (B2) to give

$$\frac{\delta \mathcal{S}^A_\Omega}{\delta \boldsymbol{A}} = \frac{\mu_\Omega \boldsymbol{B}}{2\mu_0} + \boldsymbol{\nabla}\times\frac{\partial \mathcal{L}^{A_-}_\Omega}{\partial \boldsymbol{B}} + \boldsymbol{\nabla}\times\frac{\mu_\Omega \boldsymbol{A}}{2\mu_0} + \frac{\partial}{\partial t}\frac{\partial \mathcal{L}^{A_-}_\Omega}{\partial \boldsymbol{E}} = 0, \tag{5.5}$$

i.e.

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}^{A_-}_\Omega}{\partial \boldsymbol{E}} + \boldsymbol{\nabla}\times\frac{\partial \mathcal{L}^{A_-}_\Omega}{\partial \boldsymbol{B}} = -\frac{\mu_\Omega \boldsymbol{B}}{\mu_0}. \tag{5.6}$$

The $\delta\Phi$-EL equation is, using (B3),

$$\frac{\delta \mathcal{S}^A_\Omega}{\delta \Phi} = \boldsymbol{\nabla}\cdot\frac{\partial \mathcal{L}^A_\Omega}{\partial \boldsymbol{E}} = 0. \tag{5.7}$$

Using (2.2), $\delta\boldsymbol{v} = \partial_t\boldsymbol{\xi} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{\xi} - \boldsymbol{\xi}\cdot\boldsymbol{\nabla}\boldsymbol{v}$, the $\Delta\boldsymbol{x}$-EL (or $\boldsymbol{\xi}$-EL – see (5.1)) equation is

$$\begin{aligned}\frac{\delta \mathcal{S}^A_\Omega}{\delta \boldsymbol{x}} &= -\partial_t\boldsymbol{\Pi} - \boldsymbol{\nabla}\cdot(\boldsymbol{v}\boldsymbol{\Pi}) - (\boldsymbol{\nabla}\boldsymbol{v})\cdot\boldsymbol{\Pi} \\ &\quad + \boldsymbol{\nabla}\left(\rho\frac{\partial \mathcal{L}^A_\Omega}{\partial \rho}\right) - (\boldsymbol{\nabla}\rho)\frac{\partial \mathcal{L}^A_\Omega}{\partial \rho} = 0,\end{aligned} \tag{5.8}$$

where

$$\boldsymbol{\Pi} \overset{\text{def}}{=} \frac{\partial \mathcal{L}^A_\Omega}{\partial \boldsymbol{v}}, \tag{5.9}$$

is a new canonical momentum density (cf. $\boldsymbol{\pi}$ in § (2.5)).

5.2. *Formal conservation-form momentum equation*

A general form of the equation of motion is provided by the $\boldsymbol{\xi}$-EL equation (5.8), which agrees with (21) of Dewar *et al.* (2020) in the special case of their $\boldsymbol{\lambda} = [\rho]$, $\boldsymbol{V} = [0]$ and $\boldsymbol{\Lambda} = [1]$.

To get a more transparent version we now derive a canonical-momentum conservation form of the equation of motion, the existence of which is implied by Noether's theorem and translational invariance (within Ω, i.e. not including $\partial\Omega$). To do this we transform

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

(5.8) into the same form as (22) of Dewar *et al.* (2020)[3] by subtracting $\nabla\mathcal{L}^A_\Omega$ from both sides, giving, after a little rearrangement,

$$\partial_t\boldsymbol{\Pi}+\nabla\cdot\left[\boldsymbol{v}\boldsymbol{\Pi}+\boldsymbol{I}\left(\mathcal{L}^A_\Omega-\rho\frac{\partial\mathcal{L}^A_\Omega}{\partial\rho}\right)\right]$$
$$=\nabla\mathcal{L}^A_\Omega-(\nabla\boldsymbol{v})\cdot\boldsymbol{\Pi}-(\nabla\rho)\frac{\partial\mathcal{L}^A_\Omega}{\partial\rho}. \tag{5.10}$$

Local translational invariance implies the only $\boldsymbol{x}$ dependence of $\mathcal{L}^A_\Omega$ is through its component fields, so applying the chain rule, using (5.2), (5.3) and (5.4*a*,*b*) (here and elsewhere below see online supplement for intermediate steps) we get

$$\nabla\mathcal{L}^A_\Omega=\nabla\cdot\left[\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{E}}\boldsymbol{E}-\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{B}}\times\boldsymbol{I}\times\boldsymbol{B}+\boldsymbol{I}\frac{\mu_\Omega\boldsymbol{A}\cdot\boldsymbol{B}}{2\mu_0}\right]-\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{E}}\times\boldsymbol{B}\right)$$
$$+(\nabla\rho)\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\rho}+(\nabla\boldsymbol{v})\cdot\boldsymbol{\Pi}+(\nabla\boldsymbol{\lambda})\cdot\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{\lambda}}, \tag{5.11}$$

where we used the identity (B7), $(\nabla\boldsymbol{B})\cdot\boldsymbol{f}=(\nabla\times\boldsymbol{f})\times\boldsymbol{B}-\nabla\cdot[\boldsymbol{f}\times\boldsymbol{I}\times\boldsymbol{B}]$ of Appendix B, with $\boldsymbol{f}=\partial\mathcal{L}^{A_-}_\Omega/\partial\boldsymbol{B}$ and the $\delta\boldsymbol{A}$-EL equation (5.6), and also the identity (B5) $(\nabla\boldsymbol{E})\cdot\boldsymbol{f}=\nabla\cdot[\boldsymbol{f}\boldsymbol{E}]-\boldsymbol{E}\nabla\cdot\boldsymbol{f}-\boldsymbol{f}\times\partial_t\boldsymbol{B}$ with $\boldsymbol{f}=\partial\mathcal{L}^{A_-}_\Omega/\partial\boldsymbol{E}$, and the $\delta\Phi$-EL equation (5.7), to reduce all but the last three terms to divergence form. Eliminating these $\nabla\rho$ and $\nabla\boldsymbol{v}$ terms between those in (5.10) and $\nabla\mathcal{L}^A_\Omega$ above, and also cancelling the $\boldsymbol{A}\cdot\boldsymbol{B}$ terms, gives a general momentum equation in gauge-independent conservation form on the left-hand side, but with the $\nabla\boldsymbol{\lambda}$ term on the right-hand side acting as an external forcing term,

$$\partial_t\left(\boldsymbol{\Pi}+\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{E}}\times\boldsymbol{B}\right)+\nabla\cdot\boldsymbol{T}=(\nabla\boldsymbol{\lambda})\cdot\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{\lambda}}. \tag{5.12}$$

Here the tensor $\boldsymbol{T}$ is given by

$$\boldsymbol{T}=\boldsymbol{v}\boldsymbol{\Pi}+\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{B}}\times\boldsymbol{I}\times\boldsymbol{B}-\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{E}}\boldsymbol{E}+\boldsymbol{I}\left(\mathcal{L}^{A_-}_\Omega-\rho\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\rho}\right). \tag{5.13}$$

(See (B8) for a dyadic identity that is useful for interpreting the second term on the right-hand side.)

This construction illustrates that the momentum conservation form is a general property of any translation-invariant Lagrangian formulation (by Noether's theorem) and thus is preserved even with our augmented penalty function constraint (except for the forcing term from the symmetry-breaking Lagrange multiplier). It is not manifestly symmetric but we expect it to be symmetrizable from local rotational invariance (Dewar 1970, 1977).

We now examine the implications of our EL equations in more detail.

[3]Unfortunately the seemingly general stress tensor (27) derived by Dewar *et al.* (2020) was limited to scalar fields like Φ. Appendix B derives (B6) to handle vector fields like $\boldsymbol{A}$.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

5.3. *Explicit variation of Eulerian velocity*

From the $\delta\boldsymbol{u}$-EL equation (5.2),

$$\rho(\boldsymbol{v}-\boldsymbol{u})+\frac{\nu_\Omega \boldsymbol{B}}{\mu_0}=0, \tag{5.14}$$

which is equivalent to the relative flow formula (1.8) given in the introduction, thus both motivating and justifying substituting $\boldsymbol{v}$ for $\boldsymbol{u}$ (see text below (1.8)). We shall use this below in the form $\boldsymbol{v}=\boldsymbol{u}-\boldsymbol{u}^{\mathrm{Rx}}$ for eliminating $\boldsymbol{v}$ when required. (Recall $\boldsymbol{u}^{\mathrm{Rx}}\stackrel{\mathrm{def}}{=}\nu_\Omega \boldsymbol{B}/\mu_0\rho$.)

Note that taking the divergence of both sides of 1.8, the EL equation (5.14), we have $\boldsymbol{\nabla}\cdot(\rho\boldsymbol{v})=\boldsymbol{\nabla}\cdot(\rho\boldsymbol{u})$. Thus, as noted below (1.8), $\boldsymbol{u}$ obeys the same continuity equation as $\boldsymbol{v}$, (2.4), that is,

$$\partial_t\rho+\boldsymbol{\nabla}\cdot(\rho\boldsymbol{u})=0. \tag{5.15}$$

5.4. *Variation of pressure*

From the δp-EL equation (5.3)

$$\frac{1}{\gamma-1}\left(\frac{\tau_\Omega\rho}{p}-1\right)=0, \tag{5.16}$$

which leads to the isothermal equation of state

$$p=\tau_\Omega\rho. \tag{5.17}$$

A related result is sometimes useful: from (4.20) after a little algebra,

$$\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\rho}=\boldsymbol{u}\cdot\boldsymbol{v}-h_\Omega, \tag{5.18}$$

where the Bernoulli 'head' h_Ω is defined by

$$h_\Omega=\frac{u^2}{2}+\tau_\Omega\ln\frac{\rho}{\rho_\Omega}, \tag{5.19}$$

with ρ_Ω a spatially constant reference density that need not be given as it does not contribute to the $\Delta\boldsymbol{x}$-EL, (5.8). It has the property that $\rho\boldsymbol{\nabla}h_\Omega=\rho\boldsymbol{\nabla}\frac{1}{2}u^2+\boldsymbol{\nabla}p$.

5.5. *Explicit variation of scalar potential*

From the $\delta\Phi$-EL equation (5.7) and (5.29),

$$\boldsymbol{\nabla}\cdot\boldsymbol{\lambda}_*=0, \tag{5.20}$$

where $\boldsymbol{\lambda}_*=\boldsymbol{\lambda}-\mu^P_\Omega\boldsymbol{C}$ is as defined in (4.7).

Comparing the update rule (4.6), $\boldsymbol{\lambda}|^{n+1}=\boldsymbol{\lambda}|^n-\mu^P_\Omega|^n\boldsymbol{C}|^n$, with (4.7) we identify $\boldsymbol{\lambda}_*$ as the updated $\boldsymbol{\lambda}$ for initializing the next iteration, i.e. $\boldsymbol{\lambda}|^{n+1}=\boldsymbol{\lambda}_*|^n$. As $\boldsymbol{\nabla}\cdot\boldsymbol{\lambda}_*|^n=0$ we therefore have $\boldsymbol{\nabla}\cdot\boldsymbol{\lambda}^{n+1}=0$, and likewise for $\boldsymbol{\nabla}\cdot\boldsymbol{\lambda}|^{n+2}$ and all subsequent Lagrange multipliers in the iteration sequence. In fact, assuming integer n is a typical step in the iteration, we must also conclude that

$$\boldsymbol{\nabla}\cdot\boldsymbol{\lambda}|^n=0\quad\forall n,\ \text{including 0 and }\infty. \tag{5.21}$$

Thus we can eliminate both $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}_*$ from (4.7) by taking the divergence of both sides to give

$$\boldsymbol{\nabla}\cdot\boldsymbol{C}=0, \tag{5.22}$$

which, being a homogeneous equation, provides no driving term for $\boldsymbol{C}$.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

While, from (4.3), (5.22) implies an inhomogeneous equation for $\boldsymbol{E}$,

$$\nabla\cdot\boldsymbol{E}=-\nabla\cdot(\boldsymbol{u}\times\boldsymbol{B}), \tag{5.23}$$

this is also implied by the IOL, again showing that we cannot determine non-feasibility by taking divergences only.

However, we also have an expression for $\nabla\times\boldsymbol{E}$ from the Maxwell–Faraday induction equation (1.3), which, combined with (4.3), gives the inhomogeneous equation

$$\nabla\times\boldsymbol{C}=-\partial_t\boldsymbol{B}+\nabla\times(\boldsymbol{u}\times\boldsymbol{B}). \tag{5.24}$$

Thus the non-feasibility parameter $\boldsymbol{C}$ can be viewed as driven by the departure from the IMHD magnetic-field evolution equation.

This is seen better by rewriting (5.24) as an evolution equation for $\boldsymbol{B}$:

$$\partial_t\boldsymbol{B}=\nabla\times(\boldsymbol{u}\times\boldsymbol{B})-\nabla\times\boldsymbol{C}. \tag{5.25}$$

When $\boldsymbol{C}=0$ this is the IMHD evolution equation for $\boldsymbol{B}$, irrespective of our magnetic and cross-helicity constraints and confirms that allowing $\boldsymbol{C}\neq 0$ is sufficient to relax the flux-freezing topological constraints of IMHD.

However, to satisfy (1.3) automatically we use the potential representations (1.6), $\boldsymbol{E}=-\nabla\Phi-\partial_t\boldsymbol{A}$, and (1.5), $\boldsymbol{B}=\nabla\times\boldsymbol{A}$, so (4.3) becomes

$$\boldsymbol{C}=-\nabla\Phi-\partial_t\boldsymbol{A}+\boldsymbol{v}\times(\nabla\times\boldsymbol{A}), \tag{5.26}$$

showing $\boldsymbol{C}$ as the discrepancy between the potential representations of $-\boldsymbol{E}$ and of $\boldsymbol{v}\times\boldsymbol{B}$ (equivalently $\boldsymbol{u}\times\boldsymbol{B}$ after EL equations are derived). As (5.26) implies (5.25), the latter is now not an independent equation and is useful only for insight.

In potential representation (5.23) becomes the Poisson equation

$$\nabla^2\Phi=\nabla\cdot(\boldsymbol{u}\times\boldsymbol{B}), \tag{5.27}$$

where Coulomb gauge, $\nabla\cdot\boldsymbol{A}=0$, has been adopted to eliminate the explicit unknown $\boldsymbol{A}$, though it is still implicit through $\boldsymbol{B}$. The solution of this elliptic differential equation, using the Dirichlet boundary conditions discussed after (4.3), is such that Φ is a smooth function. However, as discussed in § 3.2, the electric and magnetic field lines it defines are not generically integrable and thus may represent chaotic flows in 3-D geometries.

Assuming $\boldsymbol{u}$ and $\boldsymbol{A}$ are determined using other EL equations, so (5.27) can be solved for Φ, (5.26) gives us $\boldsymbol{C}$ and then $\boldsymbol{\lambda}_*$ from (4.7). This will be illustrated in § 6.2.2.

5.6. *Explicit Variation of vector potential (3.3)*

From (4.21), (4.19) and (4.4*a–c*) we have

$$\frac{\partial\mathcal{L}_\Omega^{A_-}}{\partial\boldsymbol{B}}=-\frac{\boldsymbol{B}}{\mu_0}+\frac{\nu_\Omega\,\boldsymbol{u}}{\mu_0}-\boldsymbol{v}\times\boldsymbol{\lambda}_* \tag{5.28}$$

$$\text{and}\quad \frac{\partial\mathcal{L}_\Omega^{A_-}}{\partial\boldsymbol{E}}=\boldsymbol{\lambda}_*. \tag{5.29}$$

Inserting these identities in the $\delta\boldsymbol{A}$-EL equation (5.6) gives

$$\frac{\partial\boldsymbol{\lambda}_*}{\partial t}-\nabla\times(\boldsymbol{v}\times\boldsymbol{\lambda}_*)=\frac{1}{\mu_0}(\nabla\times\boldsymbol{B}-\mu_\Omega\boldsymbol{B}-\nu_\Omega\nabla\times\boldsymbol{u}), \tag{5.30}$$

displayed as an inhomogeneous hyperbolic equation for the Lagrange-multiplier field $\boldsymbol{\lambda}_*$. However, it can also be displayed in its dual form of an inhomogeneous elliptic equation

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

for $\boldsymbol{B}$ by multiplying both sides with $-\mu_0$ and rearranging to give

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \mu_{\Omega}\boldsymbol{B} + \nu_{\Omega}\boldsymbol{\nabla}\times\boldsymbol{u} + \mu_0[\partial_t\boldsymbol{\lambda}_* - \boldsymbol{\nabla}\times(\boldsymbol{v}\times\boldsymbol{\lambda}_*)]. \tag{5.31}$$

Apart from the terms in $\boldsymbol{\lambda}_*$ this is the RxMHD modified Beltrami equation found by Dewar *et al.* (2020).

The relation between Φ, $\boldsymbol{A}$, $\boldsymbol{\lambda}_*$ and $\boldsymbol{C}$ appears somewhat difficult to untangle in general so we shall defer detailed analysis of these equations to § 6, where the WKB approximation makes the task easier. Suffice it here simply to count equations to give confidence that the problem can be solved in principle – the four independent equations for these four unknowns are, in order of occurrence, (4.7), (5.26), (5.27) and (5.31). (Unless we set $\nu_{\Omega} = 0$, ρ occurs through the $\boldsymbol{u}^{\mathrm{Rx}}$ in $\boldsymbol{v}$, in which case we need to add (1.7), (1.8) and (5.15) to the list.) When solved, all variables should be known in terms of $\boldsymbol{u}$, whose evolution can then be determined from the $\boldsymbol{\xi}$-EL equation.

A final remark – taking the divergence of both sides the $\delta\boldsymbol{A}$ EL equation (5.31) verifies that it propagates the $\delta\Phi$ EL equation (5.20), $\boldsymbol{\nabla}\cdot\boldsymbol{\lambda}_* = 0$. That is, if $\boldsymbol{\nabla}\cdot\boldsymbol{\lambda}_* = 0$ initially, it will remain so even if $\boldsymbol{\lambda}_*$ changes as the plasma evolves in time, and at each step in the iteration to converge $\boldsymbol{C}\to 0$. So the two EL equations are consistent, though otherwise independent.

5.6.1. *Electric current*

We can also identify the electric current, $\boldsymbol{j} \overset{\text{def}}{=} \boldsymbol{\nabla}\times\boldsymbol{B}/\mu_0$, so (5.31) can be written as

$$\boldsymbol{j} = \frac{\mu_{\Omega}}{\mu_0}\boldsymbol{B} + \frac{\nu_{\Omega}}{\mu_0}\boldsymbol{\omega} - \boldsymbol{\nabla}\times(\boldsymbol{v}\times\boldsymbol{\lambda}_*) + \frac{\partial\boldsymbol{\lambda}_*}{\partial t}, \tag{5.32}$$

where $\boldsymbol{\omega} \overset{\text{def}}{=} \boldsymbol{\nabla}\times\boldsymbol{u}$ is the fluid vorticity. The first term on the right-hand side of (5.31) is the usual parallel electric current term of the linear-force-free (Beltrami) magnetic field model; the second term is a vorticity-driven current Yokoi (2013) term; while the last two terms form a new IOL constraint current which (taking into account the EL equation $\boldsymbol{\nabla}\cdot\boldsymbol{\lambda}_* = 0$) maintains the divergence-free nature of $\boldsymbol{j}$ as required to maintain quasineutrality.

5.6.2. *Physical interpretation of estimated Lagrange multiplier*

In the special case $\mu_{\Omega} = \nu_{\Omega} = 0$, $\boldsymbol{v} = \boldsymbol{u}$, if we make the identification $\boldsymbol{\lambda}_* = \boldsymbol{P}$ (5.32) becomes identical with the representation of $\boldsymbol{j}$ in terms of the electrostatic dipole moment per unit volume or polarization vector $\boldsymbol{P}$ (see e.g. §§ 1–10 of Panofsky & Phillips 1962). This representation is as given in (12) of Calkin (1963) and (1.2) of Webb & Anco (2017), specialized to the MHD case of a quasineutral moving medium, where $\boldsymbol{\nabla}\cdot\boldsymbol{P} = 0$ (consistently with (5.20)).

Calkin (1963) goes on to derive an IMHD action principle in terms of Clebsch potentials, but these are not globally defined in a 3-D plasma with non-integrable magnetic fields. Our derivation shows the Clebsch representation is not needed to apply this polarization representation for $\boldsymbol{j}$ in an action principle if we apply the Lagrangian variational approach of Newcomb (1962). (See also Webb & Anco (2019) for a discussion of the equivalence of Lagrangian and Eulerian variational approaches.)

5.7. *Explicit Lagrangian variation of fluid element position*

This final EL equation will in principle provide sufficient equations to solve for the unknowns.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

5.7.1. *Equations of motion*

From (4.4*a–c*) $\partial_{\boldsymbol{v}}\boldsymbol{C} = \boldsymbol{I}\times\boldsymbol{B}$, thus

$$\boldsymbol{\Pi} = \rho\boldsymbol{u} + \boldsymbol{B}\times\boldsymbol{\lambda}_*. \tag{5.33}$$

The EL equation obtained from setting $\delta\mathcal{S}^A_\Omega/\delta\boldsymbol{x} = 0$ in (5.8) thus becomes

$$\begin{aligned}&\partial_t(\rho\boldsymbol{u} + \boldsymbol{B}\times\boldsymbol{\lambda}_*) + \boldsymbol{\nabla}\cdot[\boldsymbol{v}(\rho\boldsymbol{u} + \boldsymbol{B}\times\boldsymbol{\lambda}_*)] + (\boldsymbol{\nabla}\boldsymbol{v})\cdot(\rho\boldsymbol{u} + \boldsymbol{B}\times\boldsymbol{\lambda}_*)\\ &\quad= \rho\boldsymbol{\nabla}\left(\frac{\partial\mathcal{L}^A_\Omega}{\partial\rho}\right) = \rho\boldsymbol{\nabla}(\boldsymbol{u}\cdot\boldsymbol{v} - h_\Omega),\end{aligned} \tag{5.34}$$

by (5.18). Cancelling the $\rho(\boldsymbol{\nabla}\boldsymbol{v})\cdot\boldsymbol{u}$ occurring on both sides and rearranging, we have

$$\begin{aligned}&\rho\partial_t\boldsymbol{u} + \rho\boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{u} - \rho(\boldsymbol{\nabla}\boldsymbol{u})\cdot\boldsymbol{v}\\ &\quad= -\rho\boldsymbol{\nabla}h_\Omega - \partial_t(\boldsymbol{B}\times\boldsymbol{\lambda}_*) - \boldsymbol{\nabla}\cdot[\boldsymbol{v}(\boldsymbol{B}\times\boldsymbol{\lambda}_*)] - (\boldsymbol{\nabla}\boldsymbol{v})\cdot(\boldsymbol{B}\times\boldsymbol{\lambda}_*),\end{aligned} \tag{5.35}$$

where we used (2.4), $\partial_t\rho + \boldsymbol{\nabla}\cdot(\rho\boldsymbol{v}) = 0$, to cancel all derivatives of ρ. Thus, dividing both sides by ρ we have the compact Bernoulli-like form

$$\partial_t\boldsymbol{u} + \boldsymbol{\omega}\times\boldsymbol{v} = -\boldsymbol{\nabla}h_\Omega - \boldsymbol{a}_\lambda, \tag{5.36}$$

the residual acceleration term containing $\boldsymbol{\lambda}_*$ being

$$\begin{aligned}\boldsymbol{a}_\lambda &\stackrel{\text{def}}{=} \rho^{-1}[\partial_t(\boldsymbol{B}\times\boldsymbol{\lambda}_*) + \boldsymbol{\nabla}\cdot[\boldsymbol{v}(\boldsymbol{B}\times\boldsymbol{\lambda}_*)] + (\boldsymbol{\nabla}\boldsymbol{v})\cdot(\boldsymbol{B}\times\boldsymbol{\lambda}_*)]\\ &= \partial_t\boldsymbol{w} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{w} + (\boldsymbol{\nabla}\boldsymbol{v})\cdot\boldsymbol{w},\end{aligned} \tag{5.37}$$

where

$$\boldsymbol{w} \stackrel{\text{def}}{=} \frac{\boldsymbol{B}\times\boldsymbol{\lambda}_*}{\rho}, \tag{5.38}$$

again using $\partial_t\rho + \boldsymbol{\nabla}\cdot(\rho\boldsymbol{v}) = 0$.

In the special case $\mu_\Omega = \nu_\Omega = 0$, $\boldsymbol{v} = \boldsymbol{u}$ we can use the identification in § 5.6.2 of $\boldsymbol{\lambda}_*$ as the polarization field $\boldsymbol{P}$ to write $\boldsymbol{w} = \boldsymbol{B}\times\boldsymbol{P}/\rho$. We can then recognize (5.36) as the Eulerian equation of motion, (23) of Calkin (1963), thus providing a physical interpretation of our equations of motion in terms of a Lagrange multiplier field.

5.7.2. *Conservation form*

Now consider the conservation form (5.12) where, from (4.4*a–c*), (5.9) and (4.10)

$$\boldsymbol{\Pi} + \frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\boldsymbol{E}}\times\boldsymbol{B} = \rho\boldsymbol{u}. \tag{5.39}$$

Starting with the coefficient of $\boldsymbol{I}$ in the tensor $\boldsymbol{T}$, (5.13), and referring to (4.19), (4.20) and (4.21) we find

$$\mathcal{L}^{A_-}_\Omega - \rho\frac{\partial\mathcal{L}^{A_-}_\Omega}{\partial\rho} = p - \frac{B^2}{2\mu_0} + \nu_\Omega\frac{\boldsymbol{u}\cdot\boldsymbol{B}}{\mu_0} + \boldsymbol{\lambda}\cdot\boldsymbol{C} - \frac{\mu^P_\Omega C^2}{2}. \tag{5.40}$$

The penultimate term in $\boldsymbol{T}$ is

$$-\frac{\partial\mathcal{L}^C_\Omega}{\partial\boldsymbol{E}}\boldsymbol{E} = -\boldsymbol{\lambda}_*\boldsymbol{E}, \tag{5.41}$$

which consists of a symmetric $\boldsymbol{E}'\boldsymbol{E}'$ term and a non-symmetric $\boldsymbol{E}'\boldsymbol{u}\times\boldsymbol{B}$ term.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

The preceding term of $\boldsymbol{T}$ is, using (5.28) and (B8),

$$\frac{\partial \mathcal{L}_\Omega^{A_-}}{\partial \boldsymbol{B}} \times \boldsymbol{I} \times \boldsymbol{B} - \frac{\boldsymbol{I}}{\mu_0}(-B^2 + \nu_\Omega \boldsymbol{u} \cdot \boldsymbol{B} - \mu_0 \boldsymbol{v} \times \boldsymbol{\lambda}_* \cdot \boldsymbol{B}), \tag{5.42}$$

and the remaining first term is

$$\boldsymbol{v\Pi} = \left(\boldsymbol{u} - \frac{\nu_\Omega \boldsymbol{B}}{\mu_0 \rho}\right)(\rho \boldsymbol{u} + \boldsymbol{B} \times \boldsymbol{\lambda}_*). \tag{5.43}$$

Thus, combining all terms, (5.12) becomes

$$\partial_t(\rho \boldsymbol{u}) + \nabla \cdot (\boldsymbol{T}_{\mathrm{MHD}} + \boldsymbol{T}_{\mathrm{Res}}) = (\nabla \boldsymbol{\lambda}) \cdot \boldsymbol{C}, \tag{5.44}$$

where $\boldsymbol{T}_{\mathrm{MHD}}$ is the momentum transport plus stress tensor for both IMHD and RxMHD (Dewar *et al.* 2020),

$$\boldsymbol{T}_{\mathrm{MHD}} = \rho \boldsymbol{u}\boldsymbol{u} + \left(p + \frac{B^2}{2\mu_0}\right)\boldsymbol{I} - \frac{\boldsymbol{B}\boldsymbol{B}}{\mu_0}, \tag{5.45}$$

the terms in ν_Ω that might have contributed to $\boldsymbol{T}_{\mathrm{MHD}}$ in the RxMHD case having cancelled.

The new term $\boldsymbol{T}_{\mathrm{Res}}$ is the 'internal' residual stress contribution arising when action-extremizing solutions are infeasible, i.e. when the IOL constraint is not satisfied exactly,

$$\begin{aligned}\boldsymbol{T}_{\mathrm{Res}} &\stackrel{\mathrm{def}}{=} \left(\boldsymbol{\lambda} \cdot \boldsymbol{C} - \frac{\mu_\Omega^P C^2}{2} - \boldsymbol{\lambda}_* \cdot \boldsymbol{v} \times \boldsymbol{B}\right)\boldsymbol{I} \\ &\quad - \boldsymbol{B}\boldsymbol{v} \times \boldsymbol{\lambda}_* + \boldsymbol{v}\boldsymbol{B} \times \boldsymbol{\lambda}_* - \boldsymbol{\lambda}_* \boldsymbol{E} \\ &= \left(\boldsymbol{\lambda}_* \cdot \boldsymbol{C} - \boldsymbol{\lambda}_* \cdot \boldsymbol{u} \times \boldsymbol{B} + \frac{\mu_\Omega^P C^2}{2}\right)\boldsymbol{I} \\ &\quad + \boldsymbol{B}\boldsymbol{\lambda}_* \times \boldsymbol{u} + \boldsymbol{u}\boldsymbol{B} \times \boldsymbol{\lambda}_* + \boldsymbol{\lambda}_* \boldsymbol{u} \times \boldsymbol{B} - \boldsymbol{\lambda}_* \boldsymbol{C}.\end{aligned} \tag{5.46}$$

(Interestingly, the ν_Ω cancellation also occurred in deriving $\boldsymbol{T}_{\mathrm{Res}}$ when $\boldsymbol{v}$ was replaced by $\boldsymbol{u} - \nu_\Omega \boldsymbol{B}/\mu_0 \rho$.)

The 'external' residual force on the right-hand side of (5.44) obviously vanishes for feasible solutions. It is not obvious that $\boldsymbol{T}_{\mathrm{Res}}$ vanishes when $\boldsymbol{C} = 0$ as it involves the unknown converged Lagrange multiplier $\boldsymbol{\lambda}|^\infty$ ($= \boldsymbol{\lambda}_*|^\infty$ as $\boldsymbol{C}|^\infty = 0$). However, it is easy to verify that both the diagonal and off-diagonal terms of $\boldsymbol{T}_{\mathrm{Res}}$ not involving $\boldsymbol{C}$ explicitly do vanish if $\boldsymbol{\lambda}_*$ is proportional to $\boldsymbol{B}$ pointwise, implying at least in this case $\boldsymbol{T}_{\mathrm{Res}} = 0$ if and only if $\boldsymbol{C} = 0$ (assuming $\mu_\Omega^P \neq 0$).

5.7.3. *Momentum and angular momentum conservation*

When a trial solution is IOL-infeasible, i.e. $\boldsymbol{C} \neq 0$, $\boldsymbol{T}_{\mathrm{Res}}$ is not a symmetric tensor, indicating it imparts both an isotropic pressure force and a torque on the plasma, presumably tending to change $\boldsymbol{u}$ in such a way as to 'bend' the flow towards conformity with the IOL. There is a cyclic symmetry in $\boldsymbol{T}_{\mathrm{Res}}$ in the sum of the three terms in $\boldsymbol{\lambda}_*, \boldsymbol{u}, \boldsymbol{B}$ that indicates that the magnetic field is coupled to $\boldsymbol{E}'$ in a similar fashion to $\boldsymbol{u}$, and indeed we see from (5.32) that that there is a 'dynamo' term depending on $\boldsymbol{E}'$ in $\boldsymbol{j}$ that modifies $\boldsymbol{B}$, by Ampére's law.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

6. Linearized dynamics in the WKB approximation

As indicated in the introduction, the present paper is a step towards an MRxMHD dynamics code in which the primary role of the relaxed fluid dynamics within an annular toroidal domain Ω is twofold: (i) to regularize IMHD by relaxing the topological constraint forbidding magnetic reconnection, so magnetic islands can form at resonances rather than singularities; and (ii) to transmit pressure disturbances across the thin layer of plasma between the two disjoint interfaces forming the boundary $\partial\Omega$, thereby coupling the interfaces and endowing them with the plasma's inertia. This section derives, in the WKB approximation, dispersion relations for the waves that transmit these disturbances.

6.1. *Linearization*

Thus, as a simple first step towards understanding the dynamical implications of the RxMHD equations we linearize around a steady, ($\partial_t \mapsto 0$), IOL-compliant ($C^{(0)} = 0$, $\boldsymbol{\lambda}_*^{(0)} = \boldsymbol{\lambda}^{(0)}$) solution of the EL equations in a domain Ω with either fixed boundaries or with only low-amplitude, short-wavelength perturbations on $\partial\Omega$. Thus, we insert in these equations the ansatz $\boldsymbol{u} = \boldsymbol{u}^{(0)} + \alpha\boldsymbol{u}^{(1)} + O(\alpha^2)$, where α is the amplitude expansion parameter, and similarly for other perturbations except we use their potential representations for $\boldsymbol{B}^{(1)}$ and $\boldsymbol{E}^{(1)}$ as this is important for enforcing 1.3. The entropy, helicity and cross-helicity integrals are conserved at $O(\alpha)$, with therefore no perturbation in the Lagrange multipliers. Thus here we take τ_Ω, μ_Ω and ν_Ω as time-independent constants. Also, from here onward we take the superscript (0) to be implicit, e.g. ρ means $\rho^{(0)}$, $\boldsymbol{u}$ means $\boldsymbol{u}^{(0)}$, $\boldsymbol{\lambda}$ means $\boldsymbol{\lambda}^{(0)}$, etc. While we assume the background equilibrium obeys the IOL, we do not assume the augmented-Lagrangian iteration for our perturbations is fully converged, so $\boldsymbol{C}^{(1)} \neq 0$ and our two successive Euler-Lagrange iterates are not equal, $\boldsymbol{\lambda}_*^{(1)} \neq \boldsymbol{\lambda}^{(1)}$.

6.1.1. *Linearization of Lagrange multiplier determination*

Focusing first on the novel part of the calculation we list the linearizations of immediate relevance to the augmented Lagrangian determination of the updated Lagrange multiplier field $\boldsymbol{\lambda}_*$.

From (1.5) and (1.6), $\boldsymbol{B}^{(1)} = \nabla\times\boldsymbol{A}^{(1)}$ and $\boldsymbol{E}^{(1)} = -\nabla\Phi^{(1)} - \partial_t\boldsymbol{A}^{(1)}$, so (4.3) becomes

$$\boldsymbol{C}^{(1)} = -\nabla\Phi^{(1)} - \partial_t\boldsymbol{A}^{(1)} + \boldsymbol{u}\times(\nabla\times\boldsymbol{A}^{(1)}) + \boldsymbol{u}^{(1)}\times\boldsymbol{B}, \tag{6.1}$$

with $\Phi^{(1)}$ to be determined from (5.27), which used the EL equation from the Φ variation in its derivation. This becomes

$$\nabla^2\Phi^{(1)} = \nabla\cdot[\boldsymbol{u}^{(1)}\times\boldsymbol{B} + \boldsymbol{u}\times(\nabla\times\boldsymbol{A}^{(1)})]. \tag{6.2}$$

When $\boldsymbol{C}^{(1)}$ is found, the updated Lagrange multiplier is determined from the linearization of (4.7), $\boldsymbol{\lambda}_*^{(1)} = \boldsymbol{\lambda}^{(1)} - \mu_\Omega^P\boldsymbol{C}^{(1)}$.

While one occurrence of $\boldsymbol{A}^{(1)}$ in (6.2) has been eliminated by assuming Coulomb gauge, $\nabla\cdot\boldsymbol{A}^{(1)} = 0$, it still arises in the $\nabla\times\boldsymbol{A}^{(1)}$ term arising from $\boldsymbol{B}^{(1)}$. Thus we also need the linearization of the $\delta\boldsymbol{A}$ EL equation to give us $\boldsymbol{A}^{(1)}$. We use the modified Beltrami form (5.31)

$$\begin{aligned}\nabla\times(\nabla\times\boldsymbol{A}^{(1)}) = \mu_\Omega\nabla\times\boldsymbol{A}^{(1)} + \nu_\Omega\nabla\times\boldsymbol{u}^{(1)} - \mu_0\nabla\times(\boldsymbol{v}^{(1)}\times\boldsymbol{\lambda})\\ + \mu_0[\partial_t\boldsymbol{\lambda}_*^{(1)} - \nabla\times(\boldsymbol{v}\times\boldsymbol{\lambda}_*^{(1)})],\end{aligned} \tag{6.3}$$

where

$$\boldsymbol{v}^{(1)} = \boldsymbol{u}^{(1)} - \frac{\nu_\Omega}{\mu_0}\left(\frac{\nabla\times\boldsymbol{A}^{(1)}}{\rho} - \frac{\rho^{(1)}}{\rho}\frac{\boldsymbol{B}}{\rho}\right), \tag{6.4}$$

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

with $\rho^{(1)}$ to be determined from

$$\partial_t \rho^{(1)} + \nabla \cdot (\rho \boldsymbol{u}^{(1)} + \rho^{(1)} \boldsymbol{u}) = 0, \tag{6.5}$$

and $\boldsymbol{u}^{(1)}$ to be treated as the one unknown in terms of which all other physical perturbations are to be expressed.

6.2. *Wave perturbations in WKB approximation*

6.2.1. *Eikonal ansatz and natural basis vectors*

For short-wavelength high-frequency velocity perturbations we use the eikonal ansatz

$$\boldsymbol{u}^{(1)} = \tilde{\boldsymbol{u}}(\boldsymbol{x}, t) \exp\left(\frac{\mathrm{i}\varphi(\boldsymbol{x}, t)}{\varepsilon}\right), \tag{6.6}$$

with similar notations for linear perturbations of other quantities, ε being the WKB (local plane wave) expansion parameter. The instantaneous local values of wavevector and frequency as seen in the laboratory frame are then defined as $\boldsymbol{k} \overset{\text{def}}{=} \nabla\varphi$ and $\omega(\boldsymbol{x}, t) \overset{\text{def}}{=} -\partial_t \varphi$.

In the following development we shall also encounter two 'shifted' frequencies:

$$(1) \quad \omega_{\boldsymbol{k}}'^{\boldsymbol{u}} \overset{\text{def}}{=} \omega - \boldsymbol{k} \cdot \boldsymbol{u}, \tag{6.7}$$

the Doppler-shifted frequency of the wave as seen in the local rest frame of a fluid element, velocity $\boldsymbol{u}(\boldsymbol{x}, t)$, and

$$(2) \quad \omega_{\boldsymbol{k}}'^{\boldsymbol{v}} \overset{\text{def}}{=} \omega - \boldsymbol{k} \cdot \boldsymbol{v}, \tag{6.8}$$

the same as frequency (1) except with $\boldsymbol{u}$ replaced by the relative velocity $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}^{\mathrm{Rx}} \equiv \boldsymbol{u} - \nu_\Omega \boldsymbol{B}/\mu_0 \rho$.

Taking φ and equilibrium quantities to vary $O(1)$ spatial and temporal scales, ω, $\boldsymbol{k}$, $\partial_t \boldsymbol{u}$, $\nabla \boldsymbol{u}$, μ_Ω, ν_Ω, etc. are $O(1)$, but $\partial_t \boldsymbol{u}^{(1)}$, $\nabla \boldsymbol{u}^{(1)}$, etc. are large, $O(\alpha\varepsilon^{-1})$, relative to $\boldsymbol{u}^{(1)} = O(\alpha)\varepsilon^0$, and similarly for spatio-temporal derivatives of $\rho^{(1)}$ and $\boldsymbol{B}^{(1)}$.

In order for $\boldsymbol{B}^{(1)}$ to be the same order as $\boldsymbol{u}^{(1)}$, the potentials $\Phi^{(1)}$ and $\boldsymbol{A}^{(1)}$ must be $O(\alpha\varepsilon)$, so we write

$$\Phi^{(1)} = \frac{\varepsilon}{\mathrm{i}} \tilde{\Phi}(\boldsymbol{x}, t) \exp\left(\frac{\mathrm{i}\varphi(\boldsymbol{x}, t)}{\varepsilon}\right) \quad \text{and} \quad \boldsymbol{A}^{(1)} = \frac{\varepsilon}{\mathrm{i}} \tilde{\boldsymbol{A}}(\boldsymbol{x}, t) \exp\left(\frac{\mathrm{i}\varphi(\boldsymbol{x}, t)}{\varepsilon}\right); \tag{6.9a,b}$$

(1.5) and (1.6) giving then $\tilde{\boldsymbol{B}} = \boldsymbol{k} \times \tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{E}} = -\boldsymbol{k}\tilde{\Phi} + \omega\tilde{\boldsymbol{A}}$.

As in Dewar *et al.* (2020) our strategy is to express all perturbations in terms of $\tilde{\boldsymbol{u}}$, in order to find a 3×3 matrix eigenvalue equation whose roots give the dispersion relations for the three propagating wave branches. We shall also include a forcing term of the form similar to (6.6) in the equation of motion for $\boldsymbol{u}^{(1)}$ so that this 3×3 matrix appears also as a response function, with the dispersion relations giving the location of its poles.

We shall find it useful to expand vectors and dyadics in the orthonormal MHD-wave basis,

$$\boldsymbol{e}_1 \overset{\text{def}}{=} \frac{\boldsymbol{k}_\perp}{k_\perp}, \quad \boldsymbol{e}_2 \overset{\text{def}}{=} \frac{\boldsymbol{B}}{B} \equiv \boldsymbol{e}_{\boldsymbol{B}} \quad \text{and} \quad \boldsymbol{e}_3 \overset{\text{def}}{=\!=} \boldsymbol{e}_1 \times \boldsymbol{e}_2 \equiv \frac{\boldsymbol{k} \times \boldsymbol{B}}{|\boldsymbol{k} \times \boldsymbol{B}|}, \tag{6.10a–c}$$

where $\boldsymbol{k}_\perp \overset{\text{def}}{=} \boldsymbol{P}_\perp \cdot \boldsymbol{k}$, so $\boldsymbol{k} = k_\perp \boldsymbol{e}_1 + k_\parallel \boldsymbol{e}_2$ (where $k_\perp \overset{\text{def}}{=} |\boldsymbol{k}_\perp|$, $k_\parallel \overset{\text{def}}{=} \boldsymbol{k} \cdot \boldsymbol{e}_{\boldsymbol{B}}$), $\boldsymbol{u}^{\mathrm{Rx}} = u^{\mathrm{Rx}} \boldsymbol{e}_2$ (where $u^{\mathrm{Rx}} \overset{\text{def}}{=} |\boldsymbol{u}^{\mathrm{Rx}}|$), $\boldsymbol{k} \times \boldsymbol{B}_A = k_\perp c_A \boldsymbol{e}_3$, $\boldsymbol{P}_\perp = \boldsymbol{e}_1\boldsymbol{e}_1 + \boldsymbol{e}_3\boldsymbol{e}_3$ and $\boldsymbol{I} = \boldsymbol{e}_1\boldsymbol{e}_1 + \boldsymbol{e}_2\boldsymbol{e}_2 + \boldsymbol{e}_3\boldsymbol{e}_3$. (There is of course a problem if $\boldsymbol{k}_\perp = 0$, but we are interested in low-frequency MHD waves around $k_\parallel = 0$ where $|k_\perp|$is maximal.)

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

6.2.2. *IOL constraint in WKB approximation*

We now use (6.6) and (6.9*a*,*b*) in the linearizations in § 6.1.1, working to leading order in ε (for instance $\mu_\Omega \boldsymbol{B}^{(1)}$ will be dropped as higher order in ε than other terms in (6.3)). Then (6.1) becomes

$$\tilde{\boldsymbol{C}} = -\boldsymbol{k}(\tilde{\Phi} - \boldsymbol{u}\cdot\tilde{\boldsymbol{A}}) + \omega_k'^{\boldsymbol{u}}\tilde{\boldsymbol{A}} + \tilde{\boldsymbol{u}}\times\boldsymbol{B}. \tag{6.11}$$

Also (6.2) becomes

$$\boldsymbol{k}^2\tilde{\Phi} = \boldsymbol{k}^2\boldsymbol{u}\cdot\tilde{\boldsymbol{A}} + \boldsymbol{k}\cdot\tilde{\boldsymbol{u}}\times\boldsymbol{B}, \tag{6.12}$$

assuming Coulomb gauge, $\boldsymbol{k}\cdot\tilde{\boldsymbol{A}} = 0$. Inserting (6.12) in (6.11) gives

$$\tilde{\boldsymbol{C}} = \left(\boldsymbol{I} - \frac{\boldsymbol{k}\boldsymbol{k}}{\boldsymbol{k}^2}\right)\cdot\tilde{\boldsymbol{u}}\times\boldsymbol{B} + \omega_k'^{\boldsymbol{u}}\tilde{\boldsymbol{A}}. \tag{6.13}$$

Next, (6.3) becomes

$$\boldsymbol{k}^2\tilde{\boldsymbol{A}} = -\nu_\Omega\boldsymbol{k}\times\tilde{\boldsymbol{u}} - \mu_0\boldsymbol{k}\cdot\tilde{\boldsymbol{v}}\boldsymbol{\lambda} + \mu_0\omega_k'^{\boldsymbol{v}}\tilde{\boldsymbol{\lambda}}_*, \tag{6.14}$$

where we used $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}^{\mathrm{Rx}} \equiv \boldsymbol{u} - \nu_\Omega\boldsymbol{B}/\mu_0\rho$.

Finally, (6.5) and (6.4) become

$$\frac{\tilde{\rho}}{\rho} = \frac{\boldsymbol{k}\cdot\tilde{\boldsymbol{u}}}{\omega_k'^{\boldsymbol{u}}}, \tag{6.15}$$

$$\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{u}} + \frac{\nu_\Omega}{\mu_0\rho}\left(\frac{\boldsymbol{k}\cdot\tilde{\boldsymbol{u}}}{\omega_k'^{\boldsymbol{u}}}\boldsymbol{B} - \boldsymbol{k}\times\tilde{\boldsymbol{A}}\right), \tag{6.16}$$

hence

$$\boldsymbol{k}\cdot\tilde{\boldsymbol{v}} = \left(\frac{\omega_k'^{\boldsymbol{v}}}{\omega_k'^{\boldsymbol{u}}}\right)\boldsymbol{k}\cdot\tilde{\boldsymbol{u}}. \tag{6.17}$$

Substituting (6.17) in (6.14) gives

$$\boldsymbol{k}^2\tilde{\boldsymbol{A}} = -\nu_\Omega\boldsymbol{k}\times\tilde{\boldsymbol{u}} - \mu_0\left(\frac{\omega_k'^{\boldsymbol{v}}}{\omega_k'^{\boldsymbol{u}}}\right)\boldsymbol{k}\cdot\tilde{\boldsymbol{u}}\,\boldsymbol{\lambda} + \mu_0\omega_k'^{\boldsymbol{v}}\tilde{\boldsymbol{\lambda}}_*, \tag{6.18}$$

which in (6.13) then gives

$$\tilde{\boldsymbol{C}} = \left(\boldsymbol{I} - \frac{\boldsymbol{k}\boldsymbol{k}}{\boldsymbol{k}^2}\right)\cdot\tilde{\boldsymbol{u}}\times\boldsymbol{B} - \nu_\Omega\omega_k'^{\boldsymbol{u}}\frac{\boldsymbol{k}\times\tilde{\boldsymbol{u}}}{\boldsymbol{k}^2} + \mu_0\frac{\omega_k'^{\boldsymbol{v}}}{\boldsymbol{k}^2}(\omega_k'^{\boldsymbol{u}}\tilde{\boldsymbol{\lambda}}_* - \boldsymbol{k}\cdot\tilde{\boldsymbol{u}}\boldsymbol{\lambda}). \tag{6.19}$$

Treating $\tilde{\boldsymbol{C}}$ for the moment as a known and solving for $\tilde{\boldsymbol{\lambda}}_*$ we have

$$\tilde{\boldsymbol{\lambda}}_* = \frac{\boldsymbol{k}\cdot\tilde{\boldsymbol{u}}}{\omega_k'^{\boldsymbol{u}}}\boldsymbol{\lambda} + \frac{\nu_\Omega}{\mu_0}\frac{\boldsymbol{k}\times\tilde{\boldsymbol{u}}}{\omega_k'^{\boldsymbol{v}}} + \frac{(\boldsymbol{k}^2\boldsymbol{I} - \boldsymbol{k}\boldsymbol{k})\times\boldsymbol{B}\cdot\tilde{\boldsymbol{u}} + \boldsymbol{k}^2\tilde{\boldsymbol{C}}}{\mu_0\omega_k'^{\boldsymbol{u}}\omega_k'^{\boldsymbol{v}}}. \tag{6.20}$$

Normally we do not need to know the residual IOL error term $\tilde{\boldsymbol{C}}$ exactly, but to convince ourselves it can be made arbitrarily small by iteration, replace $\tilde{\boldsymbol{\lambda}}_*$ with its explicit form from

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

the linearization of (4.7), $\tilde{\boldsymbol{\lambda}} - \mu_\Omega^P \tilde{\boldsymbol{C}}$, and collect both $\tilde{\boldsymbol{C}}$ terms on the left,

$$\left(1 + \mu_0 \omega_k'^u \omega_k'^v \frac{\mu_\Omega^P}{k^2}\right) \tilde{\boldsymbol{C}} = -\left(\boldsymbol{I} - \frac{\boldsymbol{k}\boldsymbol{k}}{k^2}\right) \times \boldsymbol{B} \cdot \tilde{\boldsymbol{u}} - \nu_\Omega \omega_k'^u \frac{\boldsymbol{k} \times \tilde{\boldsymbol{u}}}{k^2} + \mu_0 \frac{\omega_k'^v}{k^2} (\omega_k'^u \tilde{\boldsymbol{\lambda}} - \boldsymbol{k} \cdot \tilde{\boldsymbol{u}}\, \boldsymbol{\lambda}), \tag{6.21}$$

which confirms the implication in subsection (3.3) that we have enough equations to determine $\tilde{\boldsymbol{C}}$, and hence $\tilde{\Phi}$, $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{\lambda}}_*$, in terms of $\tilde{\boldsymbol{u}}$.

Dividing both sides of (6.21) by the large penalty multiplier μ_Ω^P we see that $\tilde{\boldsymbol{C}}$ is smaller than the other terms and the previous iterate of $\tilde{\boldsymbol{C}}$ by an $O(1/\mu_\Omega^P)$ factor. Thus the sequence $\{\ldots, \tilde{\boldsymbol{C}}^n, \tilde{\boldsymbol{C}}^{n+1}, \tilde{\boldsymbol{C}}^{n+2}, \ldots\}$ will converge exponentially towards 0, or superexponentially if μ_Ω^P is increased appropriately at each step. Also $\tilde{\boldsymbol{\lambda}}_*$ will converge to $\tilde{\boldsymbol{\lambda}}|^\infty$.

However, this linearized calculation is sufficiently simple that we do not actually need to carry out the iteration as we can find $\tilde{\boldsymbol{\lambda}}|^\infty$analytically from 6.21 by setting its left-hand side to zero and solving for $\tilde{\boldsymbol{\lambda}} = \tilde{\boldsymbol{\lambda}}|^\infty$. Or, if we want to investigate the hypothesis that terminating the iteration at finite n, so that $\tilde{\boldsymbol{C}} \neq 0$, will regularize MHD, we can prescribe $\tilde{\boldsymbol{C}}$ and use (6.20) to give $\tilde{\boldsymbol{\lambda}}_*$. To provide a continuous sequence of dynamical fluid models running between the unconstrained RxMHD perturbation dynamics of Dewar *et al.* (2020) to the converged, $\tilde{\boldsymbol{C}} = 0$, present model we choose

$$\tilde{\boldsymbol{C}} = \varepsilon_{\mathrm{Rx}} \tilde{\boldsymbol{C}}_0, \tag{6.22}$$

with the 'relaxedness' parameter $\varepsilon_{\mathrm{Rx}}$ running from 0 (IMHD, IOL-compliant) to 1 (RxMHD, may be IOL-infeasible). We shall later also have use of the complementary 'ideality' parameter $\varepsilon_I \stackrel{\text{def}}{=} 1 - \varepsilon_{\mathrm{Rx}}$.

Here $\tilde{\boldsymbol{C}}_0$ is the IOL error for unconstrained, $\varepsilon_I = 0$, $\varepsilon_{\mathrm{Rx}} = 1$, RxMHD perturbations, which we can find by setting $\tilde{\boldsymbol{\lambda}} = 0$, $\mu_\Omega^P = 0$, and thus $\tilde{\boldsymbol{\lambda}}_* = 0$, in (6.19) to give

$$\widetilde{C_0} = \left(\boldsymbol{I} - \frac{\boldsymbol{k}\boldsymbol{k}}{k^2}\right) \cdot \tilde{\boldsymbol{u}} \times \boldsymbol{B} - \nu_\Omega \omega_k'^u \frac{\boldsymbol{k} \times \tilde{\boldsymbol{u}}}{k^2} - \mu_0 \boldsymbol{\lambda} \omega_k'^v \frac{\boldsymbol{k} \cdot \tilde{\boldsymbol{u}}}{k^2}, \tag{6.23}$$

which is a linear tensor function of the form $\widetilde{C_0} = B\boldsymbol{C}_0(\boldsymbol{k}, \boldsymbol{B}) \cdot \tilde{\boldsymbol{u}}$, where the factor B is taken out to make $\boldsymbol{C}_0$ dimensionless. By inspection of (6.23),

$$\boldsymbol{C}_0 \stackrel{\text{def}}{=} -\frac{1}{k^2 B} [\mu_0 \omega_k'^v \boldsymbol{\lambda}\boldsymbol{k} + \nu_\Omega \omega_k'^u \boldsymbol{k} \times \boldsymbol{I} + (k^2 \boldsymbol{I} - \boldsymbol{k}\boldsymbol{k}) \times \boldsymbol{B}]. \tag{6.24}$$

The linear tensor form of $\widetilde{C_0}$ implies $\tilde{\boldsymbol{C}}$, $\tilde{\boldsymbol{\lambda}}$ and $\tilde{\boldsymbol{\lambda}}_*$ are of similar form,

$$\tilde{\boldsymbol{C}} = B\boldsymbol{C}(\boldsymbol{k}, \boldsymbol{B}) \cdot \tilde{\boldsymbol{u}}, \quad \tilde{\boldsymbol{\lambda}} = \boldsymbol{\Lambda}(\boldsymbol{k}, \boldsymbol{B}) \cdot \tilde{\boldsymbol{u}}, \quad \tilde{\boldsymbol{\lambda}}_* = \boldsymbol{\Lambda}_*(\boldsymbol{k}, \boldsymbol{B}) \cdot \tilde{\boldsymbol{u}}, \tag{6.25a–c}$$

where from $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_* + \mu_\Omega^P B\boldsymbol{C}$, and from (6.20) and (6.25*a–c*),

$$\boldsymbol{\Lambda}_* = \varepsilon_I \left[\frac{\boldsymbol{\lambda}\boldsymbol{k}}{\omega_k'^u} + \frac{\nu_\Omega}{\mu_0} \frac{\boldsymbol{k} \times \boldsymbol{I}}{\omega_k'^v} + \frac{(k^2 \boldsymbol{I} - \boldsymbol{k}\boldsymbol{k}) \times \boldsymbol{B}}{\mu_0 \omega_k'^u \omega_k'^v}\right], \tag{6.26}$$

using $\boldsymbol{C} = \varepsilon_{\mathrm{Rx}} \boldsymbol{C}_0$, 6.22.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

6.2.3. *Short-wavelength dynamical RxMHD equations*

We now consider the linearized equation of motion with the forcing term mentioned in § 6.2.1, a specific force (i.e. force/mass density) we denote as $\boldsymbol{f}^{(1)}$. Thus the linearized (5.36) with forcing term becomes

$$\partial_t \boldsymbol{u}^{(1)} + \boldsymbol{\omega}^{(1)}\times\boldsymbol{v} + \boldsymbol{\omega}\times\boldsymbol{v}^{(1)} = -\boldsymbol{a}_\lambda^{(1)} - \nabla h_\Omega^{(1)} + \boldsymbol{f}^{(1)}, \tag{6.27}$$

where, from (5.19),

$$h_\Omega^{(1)} = \boldsymbol{u}\cdot\boldsymbol{u}^{(1)} + \tau_\Omega \frac{\rho^{(1)}}{\rho}, \tag{6.28}$$

from (5.37),

$$\begin{aligned}\boldsymbol{a}_\lambda^{(1)} = {}& \partial_t \boldsymbol{w}^{(1)} + \boldsymbol{v}\cdot\nabla\boldsymbol{w}^{(1)} + (\nabla\boldsymbol{v})\cdot\boldsymbol{w}^{(1)} \\ & + \boldsymbol{v}^{(1)}\cdot\nabla\boldsymbol{w} + (\nabla\boldsymbol{v}^{(1)})\cdot\boldsymbol{w},\end{aligned} \tag{6.29}$$

and, from (5.38),

$$\boldsymbol{w}^{(1)} = \frac{\boldsymbol{B}^{(1)}\times\boldsymbol{\lambda}}{\rho} + \frac{\boldsymbol{B}\times\boldsymbol{\lambda}_*^{(1)}}{\rho} - \frac{\rho^{(1)}\boldsymbol{B}\times\boldsymbol{\lambda}}{\rho^2}. \tag{6.30}$$

Using the WKB representations in (6.9*a*,*b*) and in (6.6), and analogous representations for $\rho^{(1)}$, $\boldsymbol{B}^{(1)}$, $\boldsymbol{w}^{(1)}$, $\boldsymbol{a}_\lambda^{(1)}$ and $\boldsymbol{f}^{(1)}$ in the linearizations above, and working to leading order in ε as before we have

$$-\omega\tilde{\boldsymbol{u}} + (\boldsymbol{k}\times\tilde{\boldsymbol{u}})\times\boldsymbol{v} = -\boldsymbol{k}\left(\boldsymbol{u} + \tau_\Omega\frac{\boldsymbol{k}}{\omega_k'^u}\right)\cdot\tilde{\boldsymbol{u}} - \tilde{\boldsymbol{a}}_\lambda + \tilde{\boldsymbol{f}}, \tag{6.31}$$

which we shall show can be written as

$$\boldsymbol{D}(\omega,\boldsymbol{k})\cdot\tilde{\boldsymbol{u}} = -\tilde{\boldsymbol{f}}, \tag{6.32}$$

where, noting cancelling of $\boldsymbol{k}\boldsymbol{u}\cdot\tilde{\boldsymbol{u}}$ terms,

$$\boldsymbol{D}\cdot\tilde{\boldsymbol{u}} = \omega\tilde{\boldsymbol{u}} - \boldsymbol{k}\cdot\boldsymbol{u}\tilde{\boldsymbol{u}} - \boldsymbol{u}^{\mathrm{Rx}}\times(\boldsymbol{k}\times\tilde{\boldsymbol{u}}) - \tilde{\boldsymbol{a}}_\lambda - \tau_\Omega\frac{\boldsymbol{k}\boldsymbol{k}}{\omega_k'^u}\cdot\tilde{\boldsymbol{u}}. \tag{6.33}$$

From (6.29),

$$\tilde{\boldsymbol{a}}_\lambda = -\omega_k'^v\tilde{\boldsymbol{w}} + \boldsymbol{k}\tilde{\boldsymbol{v}}\cdot\boldsymbol{w}, \tag{6.34}$$

and, from (6.30), (6.15) and (6.18),

$$\begin{aligned}\rho\tilde{\boldsymbol{w}} = {}& \left[\frac{\boldsymbol{k}}{\boldsymbol{k}^2}\times\left(-\nu_\Omega\boldsymbol{k}\times\tilde{\boldsymbol{u}} - \mu_0\left(\frac{\omega_k'^v}{\omega_k'^u}\right)\boldsymbol{k}\cdot\tilde{\boldsymbol{u}}\boldsymbol{\lambda}\right) - \frac{\boldsymbol{k}\cdot\tilde{\boldsymbol{u}}}{\omega_k'^u}\boldsymbol{B}\right]\times\boldsymbol{\lambda} \\ & + \left[\frac{\mu_0\omega_k'^v}{\boldsymbol{k}^2}(\boldsymbol{\lambda}\cdot\boldsymbol{k}\,\boldsymbol{I} - \boldsymbol{k}\boldsymbol{\lambda}) + \boldsymbol{B}\times\boldsymbol{I}\right]\cdot\boldsymbol{\Lambda}_*\cdot\tilde{\boldsymbol{u}},\end{aligned} \tag{6.35}$$

where $\boldsymbol{\Lambda}_*$ was given in (6.26).

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

6.2.4. *WKB RxMHD response matrix*

To simplify the calculation of the response we expand around a relaxed equilibrium with $\boldsymbol{\lambda} = 0$, such as the axisymmetric tokamak equilibrium in Dewar *et al.* (2020), which has a steady flow field $\boldsymbol{u}$ that is the vector sum of an arbitrary rigid toroidal rotation carried by $\boldsymbol{v}$ and an axisymmetric magnetic-field-aligned flow $\boldsymbol{u}^{\mathrm{Rx}}$ proportional to ν_Ω, which equilibrium was shown to satisfy the IOL without needing a Lagrange multiplier.

In such a case $\boldsymbol{w} = 0$, and (6.35) becomes $\tilde{\boldsymbol{w}} = \boldsymbol{B}\times\boldsymbol{\Lambda}_* \cdot \tilde{\boldsymbol{u}}/\rho$. Then (6.34) becomes

$$\tilde{\boldsymbol{a}}_\lambda = -\varepsilon_I \boldsymbol{u}^{\mathrm{Rx}}\times(\boldsymbol{k}\times\tilde{\boldsymbol{u}}) - \frac{\varepsilon_I \boldsymbol{B}\times(k^2\boldsymbol{I} - \boldsymbol{k}\boldsymbol{k})\times\boldsymbol{B}}{\mu_0\rho\omega_k'^u} \cdot \tilde{\boldsymbol{u}}. \tag{6.36}$$

In (6.32) the $\boldsymbol{u}^{\mathrm{Rx}}$ terms not involving $\varepsilon_{\mathrm{Rx}}$ cancel, giving

$$\boldsymbol{D} = (\omega_k'^u + \varepsilon_{\mathrm{Rx}}\boldsymbol{k}\cdot\boldsymbol{u}^{\mathrm{Rx}})\boldsymbol{I} - \varepsilon_{\mathrm{Rx}}\boldsymbol{k}\boldsymbol{u}^{\mathrm{Rx}} + \varepsilon_I \frac{\boldsymbol{k}_\perp\times\boldsymbol{c}_A\boldsymbol{k}_\perp\times\boldsymbol{c}_A - k^2c_A^2\boldsymbol{P}_\perp}{\omega_k'^u} - c_s^2\frac{\boldsymbol{k}\boldsymbol{k}}{\omega_k'^u}, \tag{6.37}$$

where $\boldsymbol{u}^{\mathrm{Rx}} = \nu_\Omega \boldsymbol{B}/\mu_0\rho$ is defined in (1.7), $\boldsymbol{c}_A \overset{\mathrm{def}}{=} \boldsymbol{B}/(\mu_0\rho)^{1/2}$ is the Alfvén velocity, $c_s = \tau_\Omega^{1/2}$ is the isothermal sound speed, and we have used (B8) to write $\boldsymbol{c}_A\times\boldsymbol{I}\times\boldsymbol{c}_A = \boldsymbol{c}_A\boldsymbol{c}_A - c_A^2\boldsymbol{I} \equiv -c_A^2\boldsymbol{P}_\perp$.

To represent $\boldsymbol{D}$ as a matrix we project onto the orthonormal basis (6.10*a–c*), which can be written as $\boldsymbol{e}_1 = \boldsymbol{k}_\perp/k_\perp$, $\boldsymbol{e}_2 = \boldsymbol{c}_A/c_A$ and $\boldsymbol{e}_3 = \boldsymbol{k}_\perp\times\boldsymbol{c}_A/(k_\perp c_A)$. We thus have

$$\boldsymbol{D} = (\omega_k'^u + \varepsilon_{\mathrm{Rx}}k_\parallel u^{\mathrm{Rx}})\boldsymbol{I} - \varepsilon_{\mathrm{Rx}}u^{\mathrm{Rx}}(k_\perp\boldsymbol{e}_1 + k_\parallel\boldsymbol{e}_2)\boldsymbol{e}_2 - \varepsilon_I\frac{c_A^2}{\omega_k'^u}(k^2\boldsymbol{e}_1\boldsymbol{e}_1 + k_\parallel^2\boldsymbol{e}_3\boldsymbol{e}_3) - \frac{c_s^2}{\omega_k'^u}(k_\perp\boldsymbol{e}_1 + k_\parallel\boldsymbol{e}_2)(k_\perp\boldsymbol{e}_1 + k_\parallel\boldsymbol{e}_2), \tag{6.38}$$

which can be represented as the block-diagonal matrix

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{D}_{\mathrm{MS}} & 0 \\ 0 & \omega_k'^u + \varepsilon_{\mathrm{Rx}}k_\parallel u^{\mathrm{Rx}} - \varepsilon_I k_\parallel^2 c_A^2/\omega_k'^u \end{bmatrix}, \tag{6.39}$$

with the 1×1 Alfvén block on the lower right and the 2×2 magnetosonic (MS) block,

$$\boldsymbol{D}_{\mathrm{MS}} = \begin{bmatrix} \omega_k'^u + \varepsilon_{\mathrm{Rx}}k_\parallel u^{\mathrm{Rx}} - (\varepsilon_I k^2c_A^2 + k_\perp^2c_s^2)/\omega_k'^u & -k_\perp(\varepsilon_{\mathrm{Rx}}u^{\mathrm{Rx}} + k_\parallel c_s^2/\omega_k'^u) \\ -k_\parallel k_\perp c_s^2/\omega_k'^u & \omega_k'^u - k_\parallel^2 c_s^2/\omega_k'^u \end{bmatrix}, \tag{6.40}$$

upper left.

6.2.5. *Limiting cases*

Consider first the ideal, fully converged case $\boldsymbol{C} = 0$ ($\varepsilon_I = 1$, $\varepsilon_{\mathrm{Rx}} = 0$) and use (B8) to write $k^2\boldsymbol{I} - \boldsymbol{k}\boldsymbol{k} = -\boldsymbol{k}\times\boldsymbol{I}\times\boldsymbol{k}$ so

$$\begin{aligned}\boldsymbol{D} &= \omega_k'^u\boldsymbol{I} - \frac{\boldsymbol{B}\times(\boldsymbol{k}\times\boldsymbol{I}\times\boldsymbol{k})\times\boldsymbol{B}}{\mu_0\rho\omega_k'^u} - \tau_\Omega\frac{\boldsymbol{k}\boldsymbol{k}}{\omega_k'^u} \\ &= \omega_k'^u\boldsymbol{I} - \frac{(\boldsymbol{k}\boldsymbol{B} - \boldsymbol{k}\cdot\boldsymbol{B}\boldsymbol{I})\cdot(\boldsymbol{B}\boldsymbol{k} - \boldsymbol{k}\cdot\boldsymbol{B}\boldsymbol{I})}{\mu_0\rho\omega_k'^u} - \tau_\Omega\frac{\boldsymbol{k}\boldsymbol{k}}{\omega_k'^u},\end{aligned} \tag{6.41}$$

which, apart from the definitions of $\boldsymbol{D}$ differing by a factor of $\rho\omega_k'^u$, agrees with the IMHD form, (75), of Dewar *et al.* (2020).

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

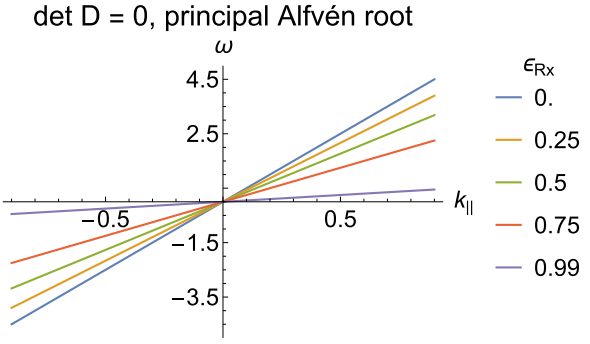


FIGURE 2. This figure shows the transition of the Alfvén-branch dispersion relation: ω versus $k_\parallel$; IMHD ($\varepsilon_{\rm Rx}=0$) to RxMHD ($\varepsilon_{\rm Rx}=1$; we have used $\varepsilon_{\rm Rx}=0.99$ for clarity as the $\varepsilon_{\rm Rx}=1$ line coincides with the $k_\parallel$ axis). Fixed parameters (in arbitrary units) are $k=1$, $\boldsymbol{u}=0$, $\boldsymbol{u}^{\rm Rx}=0$, $c_s=1$ and $c_A=5$. (Colour online – the vertical ordering of the lines in the $k_\parallel>0$ half-plane coincides with that of $\varepsilon_{\rm Rx}$ in the legend.)

In the pure RxMHD case $\varepsilon_I=0$, $\varepsilon_{\rm Rx}=1$,

$$\boldsymbol{D}=\omega_{\boldsymbol{k}}'^{\boldsymbol{u}}\boldsymbol{I}-(\boldsymbol{k}\boldsymbol{u}^{\rm Rx}-\boldsymbol{k}\cdot\boldsymbol{u}^{\rm Rx}\boldsymbol{I})-\tau_\Omega\frac{\boldsymbol{k}\boldsymbol{k}}{\omega_{\boldsymbol{k}}'^{\boldsymbol{u}}}. \tag{6.42}$$

Apart from the definitions of $\boldsymbol{D}$ again differing by a factor of $\rho\omega_{\boldsymbol{k}}'^{\boldsymbol{u}}$, this agrees with the RxMHD form, (88), of Dewar *et al.* (2020).

Thus $\varepsilon_{\rm Rx}$ parameterizes a continuous interpolation between RxMHD and IMHD.

6.2.6. *Dispersion relations – Alfvén branches*

Multiplying the first factor of the determinant

$$\det\boldsymbol{D}=(\omega_{\boldsymbol{k}}'^{\boldsymbol{u}}+\varepsilon_{\rm Rx}k_\parallel u^{\rm Rx}-\varepsilon_I k_\parallel^2c_A^2/\omega_{\boldsymbol{k}}'^{\boldsymbol{u}})\det\boldsymbol{D}_{\rm MS}, \tag{6.43}$$

by $\omega_{\boldsymbol{k}}'^{\boldsymbol{u}}$ gives the dispersion relation for the Alfvén-wave branch(es) as the quadratic equation

$$(\omega_{\boldsymbol{k}}'^{\boldsymbol{u}})^2+\varepsilon_{\rm Rx}k_\parallel u^{\rm Rx}\omega_{\boldsymbol{k}}'^{\boldsymbol{u}}-(1-\varepsilon_{\rm Rx})k_\parallel^2c_A^2=0. \tag{6.44}$$

The general solution of the quadratic equation is

$$\omega_{\boldsymbol{k}}'^{\boldsymbol{u}}=\frac{k_\parallel}{2}[-\varepsilon_{\rm Rx}u^{\rm Rx}\pm(4\varepsilon_Ic_A^2+(\varepsilon_{\rm Rx}u^{\rm Rx})^2)^{1/2}]. \tag{6.45}$$

Qualitative analysis is more informative: as $\varepsilon_{\rm Rx}\to0$ the dispersion relations for the two branches approach the Doppler-shifted Alfvén-wave dispersion relations $\omega-\boldsymbol{k}\cdot\boldsymbol{u}=\pm k_\parallel c_A$. Also, inspection shows that $\omega-\boldsymbol{k}\cdot\boldsymbol{u}\to0$ as $k_\parallel\to0$ quite generally, and when $|\varepsilon_{\rm Rx}|\ll1$ the modification of the dispersion departure from the standard Alfvén-wave dispersion relation is essentially determined by the product $\varepsilon_{\rm Rx}u^{\rm Rx}$. Thus, when $|\varepsilon_{\rm Rx}|\ll1$ and the parallel flow parameter is at most Alfvénic, $u^{\rm Rx}/c_A\le O(1)$, $\varepsilon_{\rm Rx}$ and $u^{\rm Rx}$ will have little effect on the Alfvén-wave branches.

The plots in figure 2 give a visualization of the dependence of the dispersion relation on $\varepsilon_{\rm Rx}$. The figure is for a case where $\boldsymbol{u}^{\rm Rx}=0$, when (6.45) simplifies to $\omega_{\boldsymbol{k}}'^{\boldsymbol{u}}=\pm\sqrt{\varepsilon_I}\,k_\parallel c_A$. (We call the + solution the principal branch.) The square root term $\sqrt{\varepsilon_I}=\sqrt{1-\varepsilon_{\rm Rx}}$ gives rise to a singular dependence on $\varepsilon_{\rm Rx}$ at $\varepsilon_{\rm Rx}=1$ but the vicinity of IMHD is regular.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

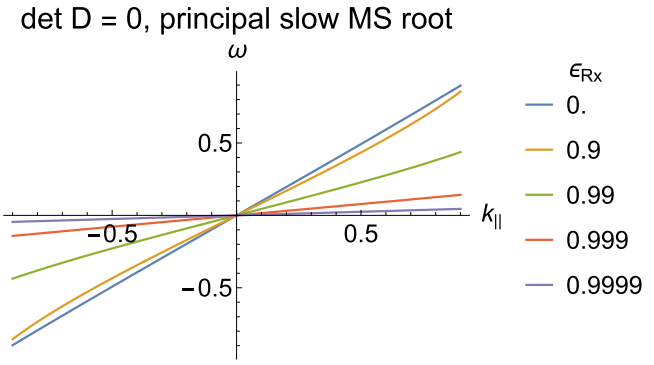


FIGURE 3. This figure shows the transition of the slow-magnetosonic-branch dispersion relation: ω versus $k_\parallel$; from IMHD ($\varepsilon_{\mathrm{Rx}} = 0$) to RxMHD ($\varepsilon_{\mathrm{Rx}} = 1$). Fixed parameters (in arbitrary units) are $k = 1$, $\boldsymbol{u} = 0$, $\boldsymbol{u}^{\mathrm{Rx}} = 0$, $c_s = 1$ and $c_A = 5$. (Colour online – the vertical ordering of the lines in the $k_\parallel > 0$ half-plane coincides with that of $\varepsilon_{\mathrm{Rx}}$ in the legend.)

6.2.7. *Dispersion relations – magnetosonic branches*

The magnetosonic dispersion relations are obtained by setting $\det \boldsymbol{D}_{\mathrm{MS}} = 0$, where

$$\det \boldsymbol{D}_{\mathrm{MS}} = \frac{c_A^2 \varepsilon_I k^2 (k_\parallel^2 c_s^2 - \omega^2) - \omega (c_s^2 k^2 - \omega^2)(\varepsilon_{\mathrm{Rx}} k_\parallel u^{\mathrm{Rx}} + \omega)}{\omega^2}. \tag{6.46}$$

The solution of the quartic equation $\omega^2 \det \boldsymbol{D}_{\mathrm{MS}} = 0$ is extremely complicated but figures 3 and 4 give an overview of the $\varepsilon_{\mathrm{Rx}}$ dependence. Again, the limit $\varepsilon_{\mathrm{Rx}} \to 1$ is clearly singular in the slow magnetosonic case but not $\varepsilon_{\mathrm{Rx}} \to 0$. This regularity around IMHD means our dispersion relation analysis is too crude to reveal the potential regularizing effect of softening the IOL constraint.

7. Conclusion

Invoking the augmented Lagrangian version of the penalty function method for constrained optimization, we have sketched out what we hope is a practical computational approach for iteratively solving the RxMHD EL equations of Dewar *et al.* (2020) with added IOL constraint terms.

This method depends crucially on the existence of a Lagrange multiplier field to be found using the augmented Lagrangian iteration algorithm borrowed from finite-dimensional optimization theory.

A formal proof of convergence may in general be difficult, but a practical approach will be to test the algorithm by perturbing away from IOL-feasible relaxed equilibria in simple geometries. A suitable such starting point is the rigidly rotating axisymmetric tokamak equilibrium discussed by Dewar *et al.* (2020). In this paper have illustrated the construction of the Lagrange multiplier field for linearized wave perturbations in the short-wavelength WKB approximation.

To find the constrained momentum equation we have used a little-known dyadic identity to derive a general conservation form. Substituting the constrained-RxMHD Lagrangian into this general form reveals residual terms in the stress tensor and a fictitious external force that should tend to zero uniformly in Ω if the constraint iteration converges so as to satisfy the IOL equality constraint.

However, in non-axisymmetric, 3-D plasma confinement systems such as stellarators and real tokamaks with field errors and intentionally resonant magnetic perturbations,

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

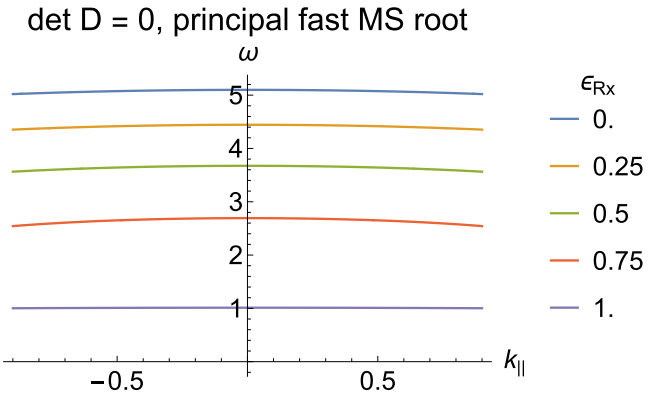


FIGURE 4. This figure shows the transition of the fast-magnetosonic-branch dispersion relation: ω versus $k_\parallel$; from IMHD ($\varepsilon_{\rm Rx}=0$) to RxMHD ($\varepsilon_{\rm Rx}=1$). Fixed parameters (in arbitrary units) are $k=1$, $\boldsymbol{u}=0$, $\boldsymbol{u}^{\rm Rx}=0$, $c_s=1$ and $c_A=5$. (Colour online – the vertical ordering of the lines in the $k_\parallel>0$ half-plane coincides with that of $\varepsilon_{\rm Rx}$ in the legend.)

there is good physical reason to believe uniform pointwise convergence is impossible. In such cases the best we can hope for is convergence in an L^2-norm, which will provide a weak-form regularization to cope with the singularities to which IMHD is prone in 3-D. This regularization should break the frozen-in flux condition of IMHD on small scales and allow interesting behaviour to be simulated without raising the order of the partial differential equations as adding resistivity does. Potential applications include reconnection events and the conjectured formation of equilibrium fractal magnetic and fluid flow patterns in 3-D systems. Other potential physical phenomena to investigate in 3-D systems include the linear normal mode spectrum, nonlinear saturation, bifurcations to oscillatory modes and the effect of quasisymmetry (Nührenberg & Zille 1988; Burby, Kallinikos & MacKay 2020; Rodriguez, Helander & Bhattacharjee 2020; Constantin, Drivas & Ginsberg 2021) on 3-D equilibria with flow (Vanneste & Wirosoetisno 2008).

Also, to improve the physical applicability of RxMHD it will be important to extend the handling of thermal relaxation beyond isotropic pressure. Relaxation parallel to the magnetic field is very reasonable physically but perpendicular relaxation has forced the use of discontinuous pressure profiles in the MRxMHD-based SPEC code described by Hudson *et al.* (2012). Thus it will be important to build on the work of Dennis *et al.* (2014*a*) to include an anisotropic pressure tensor in a weakly IOL-feasible model.

Note that an unabridged version of this paper with more detail on derivations of equations is available online as supplementary material at https://doi.org/10.1017/S0022377821001355.

Appendix A. A very brief history of RxMHD

The term relaxation in the physical sciences generally connotes a process by which a system tends towards an equilibrium state: thermodynamic, chemical, electrodynamic, mechanical or a combination of these. For example, in a closed, constant energy system initially out of thermodynamic equilibrium, relaxation occurs as the entropy increases towards a maximum. In an open system at a temperature above that of a surrounding heat bath, relaxation occurs as heat carries energy out of the system, so its thermal energy tends towards a minimum.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

In an open system with unbalanced mechanical forces, potential energy is converted into kinetic energy, which in turn is dissipated by friction into heat that is lost to the outside world, thus minimizing total energy, thermal and potential. This is the paradigm implicit in our use of the term 'relaxation', the assumption that a relaxed state is defined by the minimum of a Hamiltonian.

In plasma physics the first use of the term may have been in the paper by Chandrasekhar & Woltjer (1958), which proposes two variational principles other than maximizing entropy or minimizing energy: maximum energy for given mean-square current density and minimum dissipation for a given magnetic energy. The common element in these, and the minimum energy at constant magnetic helicity principle used by Woltjer (1958*a*) and Taylor (1974), is the derivation of a 'linear-force-free' magnetic field obeying the Beltrami equation $\nabla\times\boldsymbol{B}=\mu\boldsymbol{B}$, with μ constant, as the outcome. The Chandrasekhar and Woltjer work was in the context of plasma astrophysics, justifying the force-free assumption (where the force density in question is $\boldsymbol{j}\times\boldsymbol{B}$) basically on the assumption the plasma has low $\beta\stackrel{\text{def}}{=}p/(B^2/2\mu_0)$ and no confining forces that are strong compared with gradients of magnetic pressure. In contrast Taylor considered a toroidal terrestrial plasma confined in a metal shell and driven by a strong induced current, creating a turbulent state from which the plasma relaxes. Taylor regards the relaxation mechanism as the breaking of the microscopic IMHD topological invariants leaving only the global magnetic helicity as conserved.

Taylor (1974) is uncommitted as to the exact details of this breaking of microscopic invariants and is content to use successful comparison with experiment of the predictions flowing from his derivation of the Beltrami equation as sufficient validation of his elegantly simple model, a general philosophy we also adopt. However, in his later review, Taylor (1986) gives some more detail on the decay mechanism, citing some turbulence simulations and invokes turbulence scale length arguments to explain why it is energy that is minimized rather than magnetic helicity. Moffatt (2015) has recently critically reviewed the arguments of Taylor (1986) from a more modern perspective.

Woltjer (1958*b*) pointed out there were other global IMHD invariants beyond magnetic helicity, in particular his (2), the cross-helicity involving both flow and magnetic field. Bhattacharjee & Dewar (1982) pointed that in an axisymmetric system an infinity of additional global invariants could be generated by taking moments of $\boldsymbol{A}\cdot\boldsymbol{B}$ with powers of a flux function, and used lower moments to generate more physical pressure and current profiles for tokamak equilibria than the very restricted profiles given by Taylor's relaxation principle. Hudson *et al.* (2012) developed MRxMHD, a generalization of single-region Taylor relaxation by inserting thin IMHD barrier interface tori to frustrate global Taylor relaxation. This generalization is appropriate to non-axisymmetric equilibria in stellarators and in tokamaks with symmetry-breaking perturbations, where magnetic field-line flow can be chaotic even without turbulence. This MRxMHD formulation is implemented in the now well-established code SPEC.

A relaxation approach for finding equilibria with flow by adding a constraint additional to conservation of magnetic helicity, conservation of cross-helicity, was used by Finn & Antonsen (1983) using an entropy-maximization relaxation principle (see also the contemporaneous paper by Hameiri (1983)). However, they show this leads to the same equations as energy minimization. Thus we take, as in IMHD, the entropy in Ω to be conserved and follow Taylor in defining relaxed states as energy minima.

Pseudodynamical energy-descent relaxation processes that conserve topological invariants have been developed (Vallis, Carnevale & Young 1989; Vladimirov *et al.* 1999), but we stay within the framework of conservative classical mechanics by developing a

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

dynamical formalism, RxMHD, that includes relaxed equilibria as stationary points of a relaxation Hamiltonian, with Lagrange multipliers to constrain chosen macroscopic invariants, but which also allows non-equilibrium motions, most easily done using Hamilton's action principle. Stability can also be examined by taking the second variation of the Hamiltonian (Vladimirov *et al.* 1999), but in this paper, as in Dewar *et al.* (2015) and Dewar *et al.* (2020), we deal only with first variations.

However, the code SPEC of Hudson *et al.* (2012) implements a Newton method for finding energy minima and saddle points by calculating a Hessian matrix, which is the second variation of the MRxMHD energy; combined with a model kinetic energy obtained by loading all mass onto the interfaces between the relaxation regions. This has recently been used successfully by Kumar *et al.* (2021*a*,*b*) for calculating the spectrum of some linear eigenmodes in a tokamak, but comparison between the model kinetic energy and our new dynamical relaxation theory is desirable for determining the domain of applicability of the mass loading model.

Appendix B. Some vector and dyadic identities

In the body of this paper we have used the usual coordinate-free vector (and dyadic) calculus notations, but in this appendix we derive some identities that are more easily proved using elementary tensor notation. Detailed derivations have been deleted but are available in the online supplement.

Assuming an arbitrary fixed orthonormal basis $\{\boldsymbol{e}_i\}$, $i = 1, 2, 3 \mod 3$, a vector, $\boldsymbol{a}$ say, is represented as $\boldsymbol{a} = a_i\boldsymbol{e}_i$, the summation convention for contraction over repeated dummy indices being assumed throughout.

Thus dot and cross products are represented as $\boldsymbol{a}\cdot\boldsymbol{b} = a_ib_i$ and $\boldsymbol{a}\times\boldsymbol{b} = \boldsymbol{e}_i\varepsilon_{i,j,k}a_jb_k$, respectively, where the alternating Levi-Civita tensor ε_{ijk} is 1 or -1 according as $\{i, j, k\}$ is an even or odd permutation of $\{1, 2, 3\}$, or 0 if it is neither (e.g. if there are repeated integers). Also the operations of grad and curl acting on scalar and vector functions f and $\boldsymbol{f}$, respectively, are represented as $\nabla f = \partial f/\partial\boldsymbol{x} \stackrel{\text{def}}{=} \boldsymbol{e}_i\partial_i f$ and $\nabla\times\boldsymbol{a} \stackrel{\text{def}}{=} \boldsymbol{e}_i\varepsilon_{ijk}\partial_j a_k$, where $\partial_i \stackrel{\text{def}}{=} \partial/\partial x_i$. We use parentheses to limit the scope of the rightward differentiation of such operators. Note that left–right ordering is more important in vector notation; e.g. the dyadics $\boldsymbol{ab}$ and $\boldsymbol{ba}$ are distinct, but $a_ib_j = b_ja_i$.

First we derive three useful identities involving gradients with respect to $\boldsymbol{B} \equiv B_i\boldsymbol{e}_i$, and the unit vector parallel to $\boldsymbol{B}$, $\boldsymbol{e}_{\boldsymbol{B}}(\boldsymbol{x}) \equiv \boldsymbol{B}(\boldsymbol{x})/B(\boldsymbol{x})$. (By ï£¡ï£¡ï£¡parallel to $\boldsymbol{B}$' we mean locally tangent to the magnetic field line passing through any point $\boldsymbol{x}$. Henceforth the dependence on $\boldsymbol{x}$ is implicit as these identities concern functions purely of $\boldsymbol{B}$.)

Lemma B.1. *The gradients of $\boldsymbol{B}$, B and $\boldsymbol{e}_{\boldsymbol{B}}$ with respect to $\boldsymbol{B}$ are, in terms of the identity dyadic $\boldsymbol{I} \stackrel{def}{=} \sum_i \boldsymbol{e}_i\boldsymbol{e}_i$, the unit tangent vector $\boldsymbol{e}_{\boldsymbol{B}}$, and $\boldsymbol{P}_\perp \stackrel{def}{=} \boldsymbol{I} - \boldsymbol{e}_{\boldsymbol{B}}\boldsymbol{e}_{\boldsymbol{B}}$, the projector onto the plane perpendicular to $\boldsymbol{B}$,*

$$\frac{\partial\boldsymbol{B}}{\partial\boldsymbol{B}} = \boldsymbol{I}, \quad \frac{\partial B}{\partial\boldsymbol{B}} = \boldsymbol{e}_{\boldsymbol{B}} \quad \textit{and} \quad \frac{\partial\boldsymbol{e}_{\boldsymbol{B}}}{\partial\boldsymbol{B}} = \frac{\boldsymbol{P}_\perp}{B}. \tag{B 1a–c}$$

Lemma B.2. *Variational derivative of functional $F[\boldsymbol{A}, \Phi] = \int\!\!\int f(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{E})\,\mathrm{d}V\,\mathrm{d}t$ is*

$$\frac{\delta F}{\delta\boldsymbol{A}} = \frac{\partial f}{\partial\boldsymbol{A}} + \nabla\times\frac{\partial f}{\partial\boldsymbol{B}} + \frac{\partial}{\partial t}\frac{\partial f}{\partial\boldsymbol{E}}, \tag{B2}$$

where f is an arbitrary scalar-valued function of $\boldsymbol{A}$, $\boldsymbol{B} = \nabla\times\boldsymbol{A}$, and $\boldsymbol{E} = -\partial_t\boldsymbol{A} - \nabla\Phi$ from (1.6), $\boldsymbol{A}$ being an arbitrary vector field.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

Lemma B.3. *Variational derivative of functional F above is*

$$\frac{\delta F}{\delta \Phi} = \nabla \cdot \frac{\partial f}{\partial \boldsymbol{E}}. \tag{B3}$$

Two useful identities, closely related to integration by parts, for deriving conservation forms of EL equations for freely variable fields (members of the set denoted $\boldsymbol{\eta}$ in Dewar *et al.* (2020)) are as follows.

Lemma B.4. *For scalar fields, e.g.* Φ

$$(\nabla\nabla\Phi)\cdot\boldsymbol{f} = \nabla\cdot[\boldsymbol{f}\nabla\Phi] - (\nabla\Phi)\nabla\cdot\boldsymbol{f}, \tag{B4}$$

where $\boldsymbol{f}$ *is an arbitrary vector field, e.g.* $\partial\mathcal{L}/\partial\nabla\Phi$.

Corollary B.1. *For* $\boldsymbol{E} = -\nabla\Phi - \partial_t\boldsymbol{A}$,

$$(\nabla\boldsymbol{E})\cdot\boldsymbol{f} = \nabla\cdot[\boldsymbol{f}\boldsymbol{E}] - \boldsymbol{E}\nabla\cdot\boldsymbol{f} - \boldsymbol{f}\times\partial_t\boldsymbol{B}. \tag{B5}$$

Lemma B.5. *For* $\boldsymbol{A}$ *an arbitrary vector field and* $\boldsymbol{B} = \nabla\times\boldsymbol{A}$,

$$(\nabla\boldsymbol{B})\cdot\boldsymbol{f} = -\nabla\cdot[\boldsymbol{f}\times(\nabla\boldsymbol{A})^T] + (\nabla\boldsymbol{A})\cdot\nabla\times\boldsymbol{f}, \tag{B6}$$

where $\boldsymbol{f}$ *is an arbitrary vector field, e.g.* $\partial\mathcal{L}/\partial\boldsymbol{B}$. *(Note that for a more useful form see the corollary* (B7)*, below.)*

Corollary B.2.

$$(\nabla\boldsymbol{B})\cdot\boldsymbol{f} = (\nabla\times\boldsymbol{f})\times\boldsymbol{B} - \nabla\cdot[\boldsymbol{f}\times\boldsymbol{I}\times\boldsymbol{B}]. \tag{B7}$$

Lemma B.6.

$$\boldsymbol{f}\times\boldsymbol{I}\times\boldsymbol{g} = \boldsymbol{g}\boldsymbol{f} - \boldsymbol{f}\cdot\boldsymbol{g}\boldsymbol{I}. \tag{B8}$$

Supplementary material

Supplementary material are available at https://doi.org/10.1017/S0022377821001355.

Acknowledgements

We gratefully acknowledge useful discussions with N. Sato on constraint options, R. MacKay for suggesting the relevance of weak KAM theory to 3-D MHD equilibrium theory, J. Burby for discussions of an earlier version of this paper, M. Hegland for a discussion of regularization, L. Roberts for references on infinite-dimensional augmented Lagrangian optimization methods and M. Hole for reading the manuscript. We also thank Z. Levnajic and I. Mezic for consenting to use of their visualizations of chaos in our figure 1. Finally we thank an anonymous referee for pointing out the interpretation of the IOL Lagrange multiplier as an electrostatic polarization field.

Editor T. Passot thanks the referees for their advice in evaluating this article.

Declaration of interest

The authors report no conflict of interest.

Funding

The work of Z.Q. was supported by the Australian Research Council under grant DP170102606 and the Simons Foundation/SFARI (560651, AB). R.L.D. and Z.Q. also acknowledge travel support from the Simons Foundation/SFARI (560651, AB).

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

REFERENCES

BHATTACHARJEE, A. & DEWAR, R.L. 1982 Energy principle with global invariants. *Phys. Fluids* **25**, 887–897.

BURBY, J.W. 2017 Magnetohydrodynamic motion of a two-fluid plasma. *Phys. Plasmas* **24**, 082104–1–13.

BURBY, J.W., KALLINIKOS, N. & MACKAY, R.S. 2020 Some mathematics for quasi-symmetry. *J. Maths Phys*. **61**, 093503–1–22.

CALKIN, M.G. 1963 An action principle for magnetohydrodynamics. *Can. J. Phys*. **41**, 2241–2251.

CHANDRASEKHAR, S. & WOLTJER, L. 1958 On force-free magnetic fields. *Proc. Natl Acad. Sci. USA* **44**, 285.

CONSTANTIN, P., DRIVAS, T.D. & GINSBERG, D. 2021 On quasisymmetric plasma equilibria sustained by small force. *J. Plasma Phys*. **87**, 905870111–1–30.

DENNIS, G.R., HUDSON, S.R., DEWAR, R.L. & HOLE, M.J. 2014*a* Multi-region relaxed magnetohydrodynamics with anisotropy and flow. *Phys. Plasmas* **21**, 072512–1–10.

DENNIS, G.R., HUDSON, S.R., DEWAR, R.L. & HOLE, M.J. 2014*b* Multi-region relaxed magnetohydrodynamics with flow. *Phys. Plasmas* **21**, 042501–1–9.

DEWAR, R.L. 1970 Interaction between hydromagnetic waves and a time-dependent, inhomogeneous medium. *Phys. Fluids* **13**, 2710–2720.

DEWAR, R.L. 1977 Energy-momentum tensors for dispersive electromagnetic waves. *Austral. J. Phys*. **30**, 533.

DEWAR, R.L., BURBY, J.W., QU, Z.S., SATO, N. & HOLE, M.J. 2020 Time-dependent relaxed magnetohydrodynamics – inclusion of cross helicity constraint using phase-space action. *Phys. Plasmas* **27**, 062504–1–22.

DEWAR, R.L., YOSHIDA, Z., BHATTACHARJEE, A. & HUDSON, S.R. 2015 Variational formulation of relaxed and multi-region relaxed magnetohydrodynamics. *J. Plasma Phys*. **81**, 515810604.

FATHI, A. 2009 *Weak KAM Theorem in Lagrangian Dynamics*, Cambridge Studies in Advanced Mathematics, vol. 88. Cambridge University Press. 300 pp., ISBN: 0521822289.

FINN, J.M. & ANTONSEN, T.M. JR. 1983 Turbulent relaxation of compressible plasmas with flow. *Phys. Fluids* **26**, 3540.

FRIEMAN, E. & ROTENBERG, M. 1960 On hydromagnetic stability of stationary equilibria. *Rev. Mod. Phys*. **32**, 898–902.

GOLDSTEIN, H. 1980 *Classical Mechanics*, 2nd edn. Addison-Wesley.

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

HAMEIRI, E. 1983 The equilibrium and stability of rotating plasmas. *Phys. Fluids* **26**, 230–237.

HAMEIRI, E. 1998 Variational principles for equilibrium states with plasma flow. *Phys. Plasmas* **5**, 3270–3281.

HAMEIRI, E. 2014 Some improvements in the theory of plasma relaxation. *Phys. Plasmas* **21**, 044503–1–5.

HOSKING, R.J. & DEWAR, R.L. 2015 *Fundamental Fluid Mechanics and Magnetohydrodynamics*. Springer Singapore, published online 2015. Book copyright 2016.

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**, 112502–1–18.

KANZOW, C., STECK, D. & WACHSMUTH, D. 2018 An augmented Lagrangian method for optimization problems in Banach spaces. *SIAM J. Control Optim*. **56**, 272–291.

KUMAR, A., QU, Z., DOAK, J., DEWAR, R.L., HEZAVEH, H., NÜHRENBERG, C., ALEYNIKOVA, K., HOLE, M.J., HUDSON, S.R., LOIZU, J. & BAILLOD, A. 2021*a* Ideal MHD instabilities of multi-region relaxed MHD. *Plasma Phys. Control. Fusion* submitted.

KUMAR, A., QU, Z., HOLE, M.J., WRIGHT, A.M., LOIZU, J., HUDSON, S.R., BAILLOD, A., DEWAR, R.L. & FERRARO, N.M. 2021*b* Computation of linear MHD instabilities with the multi-region relaxed MHD energy principle. *Plasma Phys. Control. Fusion* **63**, 045006.

LEVNAJIĆ, Z. & MEZIĆ, I. 2010 Ergodic theory and visualization. I. Mesochronic plots for visualization of ergodic partition and invariant sets. *Chaos* **20**, 033114.

MEISS, J.D. 1992 Symplectic maps, variational principles, and transport. *Rev. Mod. Phys*. **64**, 795–848.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press

MEZIĆ, I. & WIGGINS, S. 1999 A method for visualization of invariant sets of dynamical systems based on the ergodic partition. *Chaos* **9**, 213.

MOFFATT, H.K. 2015 Magnetic relaxation and the Taylor conjecture. *J. Plasma Phys*. **81**, 905810608.

MORRISON, P.J. 1998 Hamiltonian description of the ideal fluid. *Rev. Mod. Phys*. **70**, 467–521.

NEWCOMB, W.A. 1958 Motion of magnetic lines of force. *Ann. Phys*. **3** (4), 347–385.

NEWCOMB, W.A. 1962 Lagrangian and Hamiltonian methods in magnetohydrodynamics. *Nucl. Fusion Suppl*. **Part 2**, 451–463.

NOCEDAL, J. & WRIGHT, S.J. 2006 *Numerical Optimization*, 2nd edn, Operation Research and Financial Engineering, vol. 25. Springer. 664 pp. ISBN-13: 978-0387-30303-1.

NÜHRENBERG, J. & ZILLE, R. 1988 Quasi-helically symmetric toroidal stellarators. *Phys. Lett.* A **129**, 113–117.

PANOFSKY, W.K.H. & PHILLIPS, M. 1962 *Classical Electricity and Magnetism*, 2nd edn. Addison-Wesley.

QU, Z.S., DEWAR, R.L., EBRAHIMI, F., ANDERSON, J.K., HUDSON, S.R. & HOLE, M.J. 2020 Stepped pressure equilibrium with relaxed flow and applications in reversed-field pinch plasmas. *Plasma Phys. Control. Fusion* **62**, 054002.

QU, Z.S., HUDSON, S.R., DEWAR, R.L., LOIZU, J. & HOLE, M.J. 2021 On the non-existence of stepped-pressure equilibria far from symmetry. *Plasma Phys. Control. Fusion* **63**, 125007.

RODRIGUEZ, E., HELANDER, P. & BHATTACHARJEE, A. 2020 Necessary and sufficient conditions for quasisymmetry. *Phys. Plasmas* **27**, 062501–1–5.

TAYLOR, J.B. 1974 Relaxation of toroidal plasma and generation of reverse magnetic fields. *Phys. Rev. Lett*. **33**, 1139–1141.

TAYLOR, J.B. 1986 Relaxation and magnetic reconnection in plasmas. *Rev. Mod. Phys*. **58**, 741–763.

VALLIS, G.K., CARNEVALE, G.F. & YOUNG, W.R. 1989 Extremal energy properties and construction of stable solutions of the Euler equations. *J. Fluid Mech*. **207**, 133–152.

VANNESTE, J. & WIROSOETISNO, D. 2008 Two-dimensional Euler flows in slowly deforming domains. *Physica* D **237**, 774–799.

VLADIMIROV, V.A., MOFFATT, H.K. & ILIN, K.I. 1999 On general transformations and variational principles for the magnetohydrodynamics of ideal fluids. Part 4. Generalized isovorticity principle for three-dimensional flows. *J. Fluid Mech*. **390**, 127–150.

WEBB, G.M. 2018 *Magnetohydrodynamics and Fluid Dynamics: Action Principles and Conservation Laws*, Lecture Notes in Physics, vol. 946. Springer International Publishing.

WEBB, G.M. & ANCO, S.C. 2017 On magnetohydrodynamic gauge field theory. *J. Phys.* A*: Math. Theor*. **50**, 255501–1–34.

WEBB, G.M. & ANCO, S.C. 2019 Conservation laws in magnetohydrodynamics and fluid dynamics: Lagrangian approach. *AIP Conf. Proc*. **2153**, 020024.

WOLTJER, L. 1958*a* A theorem on force-free magnetic fields. *Proc. Natl Acad. Sci. USA* **44**, 489–491.

WOLTJER, L. 1958*b* On hydromagnetic equilibrium. *Proc. Natl Acad. Sci. USA* **44**, 833.

YOKOI, N. 2013 Cross helicity and related dynamo. *Geophys. Astrophys. Fluid Dyn*. **107**, 114–184.

https://doi.org/10.1017/S0022377821001355 Published online by Cambridge University Press