

# A three-dimensional reduced MHD model consistent with full MHD

N. Nikulsin,<sup>1</sup> M. Hoelzl,<sup>1</sup> A. Zocco,<sup>2</sup> K. Lackner,<sup>1</sup> and S. Günter<sup>1</sup>

<sup>1</sup>*Max Planck Institute for Plasma Physics, Boltzmannstr. 2, 85748 Garching, Germany*

<sup>2</sup>*Max Planck Institute for Plasma Physics, Wendelsteinstr. 1, 17491 Greifswald, Germany*

(Dated: 23 September 2019)

Within the context of a viscoresistive magnetohydrodynamic (MHD) model with anisotropic heat transport and cross-field mass diffusion, we introduce novel three-term representations for the magnetic field (background vacuum field, field line bending and field compression) and velocity ( $\vec{E} \times \vec{B}$  flow, field-aligned flow and fluid compression), which are amenable to three-dimensional treatment. Once the representations are inserted into the MHD equations, appropriate projection operators are applied to Faraday's law and the Navier-Stokes equation to obtain a system of scalar equations that is closed by the continuity and energy equations. If the background vacuum field is sufficiently strong and the  $\beta$  is low, MHD waves are approximately separated by the three terms in the velocity representation, with each term containing a specific wave. Thus, by setting the appropriate term to zero, we eliminate fast magnetosonic waves, obtaining a reduced MHD model. We also show that the other two velocity terms do not compress the magnetic field, which allows us to set the field compression term to zero within the same reduced model. Dropping also the field-aligned flow, a further simplified model is obtained, leading to a fully consistent hierarchy of reduced and full MHD models for 3D plasma configurations. Finally, we discuss the conservation properties and derive the conditions under which the reduction approximation is valid. We also show that by using an ordering approach, reduced MHD equations similar to what we got from the ansatz approach can be obtained by means of a physics-based asymptotic expansion.

## I. INTRODUCTION

With the development of more powerful computers in the 1990s and 2000s, nonlinear numerical simulations began to play an increasingly important role in the interpretation of experimental results, planning of new experiments and the design of new machines. Nonlinear simulations allow one to simulate the operation of an entire machine on short time scales, typically thousands to hundreds of thousands of Alfvén times. In order to simulate such time scales with a reasonable spatial resolution and using a reasonable amount of computer time, one most often has to employ reduced MHD models, which eliminate fast magnetosonic waves while retaining the relevant physics<sup>1,2</sup>. The removal of fast magnetosonic waves, the fastest waves in the system, allows one to use larger time steps due to the Courant condition. Even when implicit time integration methods are used, and the Courant condition is no longer a hard limit, using time steps that are large compared to the shortest time scale can lead to poor accuracy<sup>2,3</sup>. In addition, reduced MHD has less unknowns compared to full MHD, which decreases the computational costs and memory requirements for simulations.

Reduced MHD, as first introduced by Greene and Johnson<sup>4</sup>, and later developed by Kadomtsev, Pogutse and Strauss<sup>5,6</sup>, relies on ordering in a small parameter, often taken to be the inverse aspect ratio. The ordering itself is a system of several approximations and assumptions involving the ordering parameter that allows one to determine the relative order (in terms of the ordering parameter) of any quantity with respect to any other

quantity of the same dimension. In this context, terms corresponding to fast magnetosonic waves have a higher order than the terms that one wants to keep, allowing the fast wave terms to be dropped. Naturally, there are many choices one can make in the ordering assumptions, depending on which physical effects one wants to keep, all of which result in different reduced equations<sup>1,3,6-8</sup>. The ideas of reduced MHD have also found use in astrophysics, where toroidal geometry cannot be assumed, and thus the inverse aspect ratio cannot be used as an ordering parameter<sup>9</sup>.

Starting in the 1980s, a new ansatz-based approach was introduced by Park *et al*<sup>10</sup>, where an ansatz form that eliminates fast magnetosonic waves is used for the velocity and terms of all orders are kept. Izzo *et al* used a similar ansatz in their study<sup>11</sup>. Later papers also adopt an ansatz for the magnetic field that eliminates field compression<sup>12,13</sup>. The ansatz approach allows one to make less assumptions (although, as we will show, an assumption on the relative magnitude of the induced magnetic field still needs to be made to eliminate fast waves) and keep more physical effects, while generally resulting in more complicated equations than the ordering approach. For example, internal kink modes cannot be modelled at the lowest order in inverse aspect ratio, but ansatz-based models overcome that by not neglecting higher order terms<sup>10</sup>. An ansatz-based approach can also guarantee exact conservation of energy, as in the models that we derive in this paper. Thus, while keeping more physics, the various terms in the equations of ansatz-based reduced MHD are harder to interpret due to their complexity. In addition, without an ordering parameter, error estimation becomes much more difficult.

Most of the previously published versions of reduced MHD focused on tokamaks, expanding the magnetic field around a tokamak-like vacuum field,  $\vec{B}_v = F_0 \nabla \phi$ . In addition, most of the reduced models for stellarators only approximately satisfy the condition  $\nabla \cdot \vec{B} = 0$ . A notable exception is the model by Strauss<sup>1</sup>, which expands the magnetic field around an arbitrary vacuum field, thus making no assumptions about the underlying geometry, while exactly satisfying the divergence-free condition. In this paper, we systematically follow the ansatz approach, proving our claims more rigorously than previous publications have. We arrive at a hierarchy of models that is mostly a generalization (with some modifications) of the results of Breslau *et al*<sup>12</sup>, while being compatible with an arbitrary vacuum field, just as Strauss's model<sup>1</sup>. The consistency with an ordering approach is also discussed.

The remainder of the paper is structured as follows. Section II introduces the specific viscoresistive MHD equations that serve as a starting point for the derivation. The same section also introduces Clebsch-type coordinates and our representations for magnetic field and velocity. In section III, we show that if the background vacuum field is sufficiently strong and the  $\beta$  is low, MHD waves are approximately separated by the three terms in the velocity representation, with each term containing a specific wave. In section IV, we derive a set of scalar full MHD equations by inserting the representations into the viscoresistive MHD equations and projecting the vector equations. In section V, we reduce our equations by dropping the term corresponding to fast magnetosonic waves from the velocity representation. Section VI considers the local conservation properties of the reduced MHD equations and derives validity conditions for the reduction. Finally, in appendix A we show how a similar, though not identical, system of equations can be

derived using an ordering approach.

## II. VISCORESISTIVE MHD AND MATHEMATICAL BACKGROUND

Ideal MHD, which assumes that the plasma is a perfectly conducting inviscid fluid and that there are no sources or sinks in any of the equations, remains the most well studied plasma fluid model. Due to its simplicity, it is often used in analytical calculations, and is commonly presented in introductory texts. Despite that, non-ideal effects, such as tearing modes and other resistive instabilities, become important on longer time scales, and an equilibrium that is ideally stable may not actually be stable. Thus, most modern fluid codes employ viscoresistive (and often extended) MHD, which includes non-ideal terms, making the model more realistic at the expense of increasing equation complexity.

### A. Viscoresistive MHD with heat conduction

In this subsection, the full MHD model is introduced in its usual formulation, and will be recast using potentials and stream functions in the following sections. The usual MHD notation is followed with  $\rho$ ,  $p$ ,  $\vec{v}$  and  $\vec{B}$  being density, pressure, velocity and magnetic field, respectively. In addition to that,  $\eta$  is the resistivity,  $\nu$  is the kinematic viscosity,  $D_{\perp}$  is the mass diffusion coefficient perpendicular to field lines,  $\kappa_{\perp}$  and  $\kappa_{\parallel}$  are the thermal conductivity across and along field lines, and  $S_{\rho}$  and  $S_e$  are source terms in the continuity and energy equations, respectively. The ideal gas law  $p = \rho RT$  is assumed to hold.

$$\begin{aligned}
 \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) &= P, \\
 \frac{\partial}{\partial t} (\rho \vec{v}) + \nabla \cdot (\rho \vec{v} \vec{v}) &= \vec{j} \times \vec{B} - \nabla p + \rho \nu \Delta \vec{v}, \\
 \frac{\partial}{\partial t} \left( \frac{\rho v^2}{2} + \frac{p}{\gamma - 1} + \frac{B^2}{2\mu_0} \right) + \nabla \cdot \left[ \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma - 1} \right) \vec{v} + \frac{p}{\gamma - 1} \frac{D_{\perp}}{\rho} \nabla_{\perp} \rho + \frac{\vec{E} \times \vec{B}}{\mu_0} - \kappa_{\perp} \nabla_{\perp} T - \kappa_{\parallel} \nabla_{\parallel} T \right] &= S_e - \frac{v^2}{2} P, \\
 \frac{\partial \vec{B}}{\partial t} &= -\nabla \times \vec{E}, \\
 \nabla \times \vec{B} = \mu_0 \vec{j}, \quad \nabla \cdot \vec{B} = 0, \quad \vec{E} = -\vec{v} \times \vec{B} + \eta \vec{j}, \quad P = \nabla \cdot (D_{\perp} \nabla_{\perp} \rho) + S_{\rho}. &
 \end{aligned} \tag{1}$$

The gradient operators parallel and perpendicular to the total magnetic field  $\vec{B}$  are defined as  $\nabla_{\parallel} = \frac{\vec{B}}{B^2} \vec{B} \cdot \nabla$  and  $\nabla_{\perp} = \nabla - \nabla_{\parallel}$ . The electric field can be eliminated from Faraday's law by inserting Ohm's law, the resulting equation will be referred to as the induction equation throughout the rest of this paper.

Note the form of the viscosity term in the Navier-Stokes equation. This is a rather simple approximation for the divergence of the viscous stress tensor in a plasma, made under the assumption that the viscosity is isotropic. While this may not be accurate for a magnetized plasma, it can nevertheless be sufficient to satisfactorily model plasma behavior<sup>14-16</sup>, and including

a viscosity term can sometimes help against numerical instabilities<sup>13</sup>. Due to the generic form of the viscosity term, it will not be treated in the derivations that follow, and instead a generic viscosity term will be added to the final equations.

## B. Clebsch-type coordinates

As shown in<sup>17,18</sup>, any magnetic field can be locally represented in the Clebsch form:

$$\vec{B} = \nabla\alpha \times \nabla\beta. \quad (2)$$

This means that the variables  $\alpha$  and  $\beta$  can be locally used to label field lines, and can parameterize any surface that the field lines intersect. A third coordinate that measures in the direction of the field is needed for a three-dimensional coordinate system. Let  $\chi$  be the magnetic scalar potential, so that  $\vec{B}_v = \nabla\chi$  is the vacuum component of the magnetic field, and  $\vec{B} - \nabla\chi$  is the induced magnetic field generated by currents flowing in the plasma. Assuming that such a background vacuum field exists, the magnetic scalar potential can be used as a third coordinate, forming the curvilinear coordinate system  $(\alpha, \beta, \chi)$ . Coordinate systems that rely on  $\alpha$  and  $\beta$  as the first two coordinates are called Clebsch-type coordinates, and are in general non-orthogonal<sup>17</sup>, with orthogonal coordinate systems being possible only in the special case of a shear-free magnetic field<sup>19</sup>.

Using the formalism of general curvilinear coordinates, one can define the contravariant and covariant basis vectors as:

$$\begin{aligned} \vec{e}^\alpha &\equiv \nabla\alpha, & \vec{e}^\beta &\equiv \nabla\beta, & \vec{e}^\chi &\equiv \nabla\chi, \\ \vec{e}_\alpha &\equiv J\nabla\beta \times \nabla\chi, & \vec{e}_\beta &\equiv J\nabla\chi \times \nabla\alpha, & \vec{e}_\chi &\equiv J\nabla\alpha \times \nabla\beta. \end{aligned} \quad (3)$$

Covariant and contravariant components of vectors will be represented by subscripts and superscripts, respectively. Note that the third covariant basis vector points in the direction of the magnetic field. The Jacobian for this coordinate system is  $J = [\nabla\chi \cdot (\nabla\alpha \times \nabla\beta)]^{-1} = (\nabla\chi \cdot \vec{B})^{-1} = 1/B^\chi$ . Also note that the vacuum field will be assumed to be static in time, whereas the total magnetic field can vary. Thus, while in many similarly derived coordinate systems (see for example<sup>17</sup>) it is assumed that  $B_\chi \equiv 1$ , this cannot hold in the most general case which we will consider.

We will make use of two different Clebsch-type coordinate systems in this paper. The first system is aligned to the total magnetic field and is defined as above with  $\vec{B}$  being the total magnetic field, while for the second system we use the vacuum field  $\vec{B}_v = \nabla\chi = \nabla\psi_v \times \nabla\beta_v$  in place of  $\vec{B}$ , aligning it to the vacuum field. To distinguish between the two systems, we label, in the vacuum field-aligned system, the first two coordinates as  $\psi_v$  and  $\beta_v$  and the basis vectors with the letter  $b$  instead of  $e$ . The third coordinate  $\chi$  is shared by both systems. When flux surfaces exist, it is convenient to choose

$\psi_v$  to be the flux surface label, however none of the results of this paper depend on the choice of  $\psi_v$ . For the vacuum field-aligned system, we have  $J = 1/B_v^2$  and the  $\chi$  basis vectors are related by  $\vec{b}_\chi = J\vec{b}^\chi$ . Thus,  $\nabla\chi$  is perpendicular to both  $\nabla\psi_v$  and  $\nabla\beta_v$ . This simplifies the coordinate system somewhat, as we have for some of the off-diagonal components of the metric tensor:  $g^{\psi_v\chi} = g^{\beta_v\chi} = g_{\psi_v\chi} = g_{\beta_v\chi} = 0$ , however the off-diagonal component  $g^{\psi_v\beta_v}$  is nonzero, and the coordinate system is still non-orthogonal. The off-diagonal components are related as  $g_{\psi_v\beta_v} = -Jg^{\psi_v\beta_v}$ . This becomes obvious when we recognize  $g^{ik}$  as the inverse matrix of  $g_{ik}$  and  $J$  as the determinant of  $g_{ik}$ . Given that the metric tensor is symmetric and  $g_{\psi_v\beta_v}$  is the only nonzero off-diagonal component, we can immediately see that  $g_{\psi_v\beta_v} = -Jg^{\psi_v\beta_v}$ .

We point out that satisfying the Laplace equation is the only requirement for  $\chi$ . While in this paper we refer to  $\nabla\chi$  as the vacuum field, nothing is stopping us from setting  $\chi = F_0\phi$ , where  $\phi$  is the toroidal angle, as is often done in reduced MHD for tokamaks. In that case,  $\nabla\chi$  would correspond to the dominant toroidal component of the vacuum field, while the field generated by poloidal coils will be grouped with the induced field. The magnetic scalar potential can in general be a multi-valued function. However, in the case of toroidal systems, such as tokamaks or stellarators,  $\chi$  is known analytically as a series of harmonics for arbitrary configurations, including stochastic field regions<sup>20,21</sup>, and the multi-valuedness is entirely contained in a tokamak-like  $F_0\phi$  term. To determine the coefficients of the toroidal and poloidal harmonics in the series, one needs to solve a system of linear algebraic equations, which fits the field  $\nabla\chi$  to a known curl-free field, usually obtained from the Biot-Savart law for the coils<sup>21</sup>.

It is important to note that in general the expression (2) is only valid locally within the vicinity of a given point. If we try to apply it globally, the functions  $\alpha$  and  $\beta$  can become multi-valued<sup>18</sup>. Consider, for example, the case of a tokamak with a stochastic magnetic field in a certain volume<sup>22</sup>, where the same field line circles the torus infinitely many times and covers all points within the volume. Since  $\alpha$  and  $\beta$  must be constant along a field line, if we require them to be single-valued, they will be constant within the volume, leading to zero gradients and forcing the field, as given by (2), to be zero, a contradiction. However,  $\alpha$  and  $\beta$  can be defined locally, starting in a given poloidal plane, parameterizing all points in the plane with coordinates  $\alpha$  and  $\beta$  (Note that once  $\alpha$  is selected,  $\beta$  has to be chosen appropriately to avoid having a scalar factor in front of the cross product in (2), see<sup>17,18</sup>.), labeling the field lines that pass through point  $(\alpha, \beta)$  with the same  $\alpha$  and  $\beta$ , and then following the field lines. As long as only a sector of the torus is considered, all coordinates can be single-valued, but once one transit around the torus is made, the field lines will encounter previously labeled points. This situation is similar to what happens in coordinate systems with angular coordinates, and the

solution is the same: we introduce a cut on which the  $\alpha$  and  $\beta$  undergo a discontinuity to avoid multi-valuedness, similarly to how most angular coordinates undergo a discontinuity and return to 0 after reaching  $2\pi$ .

As will be shown, only  $\chi$  and  $\psi_v$  appear in the equations that we derive. Thus, in practice, it is only necessary to implement  $\chi$  and  $\psi_v$  in a code. While  $\chi$  is known analytically as a series of harmonics and only the coefficients of the toroidal and poloidal harmonics need to be stored for each particular configuration<sup>20,21</sup>, implementing  $\psi_v$  is more difficult. Ideally,  $\psi_v$  should be chosen so that  $\nabla\psi_v$  is continuous. For example, in the case of a tokamak vacuum field  $\vec{B}_v = F_0\nabla\phi$ , one can set  $\psi_v$  to be the minor radius, or, if the vacuum field has flux surfaces, one can set  $\psi_v$  as the flux surface label. In both cases,  $\nabla\psi_v$  will be continuous everywhere, except for the  $\psi_v = 0$  axis. If magnetic islands are present, one can either group them with the induced field, keeping only the component of the vacuum field that has nested flux surfaces in  $\vec{B}_v$ , or  $\psi_v$  can be set to be the poloidal flux. In the latter case,  $\nabla\psi_v$  will be perpendicular to the total vacuum field even in the islands, but it will be discontinuous at the X and O points in addition to the magnetic axis. Finally, in stochastic field regions, both  $\psi_v$  and its gradient will inevitably be discontinuous across the cut. In such a case, we can define  $\psi_v$  as the flux surface label in the central region, where flux surfaces exist. Then, while on the cutting surface, we note the contours formed by the flux surfaces intersecting the cutting surface and extend those contours into the area where the stochastic field intersects the cutting surface. A  $\psi_v$  value will then be assigned to each stochastic field line on the basis of the contour that it intersects the cutting plane through. However, in a finite element code, it may not be possible to represent a discontinuous function in the basis used by that code, and so one may have to approximate  $\psi_v$  with a continuous function, and so  $\psi_v$  will no longer be constant along field lines. However, since the field line diffusion coefficient is small in typical devices<sup>23</sup>, the field line drift will be small compared to machine size, and the error will not be significant. Alternatively, one can forgo computing  $\psi_v$  and just find the components of its gradient by solving the system of linear ODEs derived by Xanthopoulos and Jenko<sup>24</sup>. However, in this case, since its components will only be known up to numerical accuracy,  $\nabla\psi_v$  will only be a gradient up to numerical accuracy, and so the form of the magnetic field introduced in subsection II C will only be divergence-free up to numerical accuracy, not machine precision.

### C. Magnetic field and velocity representations

For an arbitrary magnetic field, the vector potential can be represented in the vacuum field-aligned Clebsch-type coordinates as

$$\vec{A} = \Psi\nabla\chi + \Omega\nabla\psi_v, \quad (4)$$

where the  $\nabla\beta_v$  component was eliminated using a gauge transform. Expressing the vector potential of the induced magnetic field alone in the form (4), the total magnetic field can be expressed as

$$\vec{B} = \nabla\chi + \nabla\Psi \times \nabla\chi + \nabla\Omega \times \nabla\psi_v. \quad (5)$$

Since  $\chi$  satisfies the Laplace equation, this form guarantees that the magnetic field will be divergence free, even when the last term is dropped in the context of reduced MHD. Also note that we have partially fixed the gauge: the gradient of a scalar function  $F$  can only be added to the vector potential if  $\partial F/\partial\beta_v = 0$ .

Now consider the velocity field. To separate the three MHD waves, Izzo *et al*<sup>11</sup> and Breslau *et al*<sup>12</sup> used a three-term expansion of the velocity, with each wave contained in a specific term. We generalize their expression to the case of  $\chi \neq F_0\phi$  while keeping the first two terms consistent with<sup>1,13</sup>:

$$\vec{v} = \frac{\nabla\Phi \times \nabla\chi}{B_v^2} + v_{\parallel}\vec{B} + \nabla^{\perp}\zeta, \quad (6)$$

where  $\nabla^{\perp} = \nabla - \nabla^{\parallel}$  and  $\nabla^{\parallel} = \frac{\nabla\chi}{B_v^2}\nabla\chi \cdot \nabla$ . The superscripts are used to distinguish the parallel and perpendicular gradients with respect to the vacuum field from those defined with respect to the total field in the equations (1). Our expression matches that of Izzo *et al* and Breslau *et al* in the case of a tokamak vacuum field  $\chi = F_0\phi$ , except for the second term, which we have made to match the reduced expressions in<sup>1,13</sup>. Note that unlike the magnetic field in expression (5), which is in mixed form, the velocity in expression (6) is in contravariant form in the Clebsch-type coordinates aligned to the total magnetic field  $\vec{B}$ , as the cross product in the first term will produce covariant basis vectors, and the gradient in the last term can also be written in terms of covariant basis vectors due to being perpendicular to  $\vec{e}^{\chi}$ .

The terms in expressions (5) and (6) can be interpreted as, respectively, the background vacuum field, field line bending and field compression for the magnetic field, and  $\vec{E} \times \vec{B}$  velocity, field-aligned flow and fluid compression for the velocity. Note that these interpretations are not exact. In particular, the  $\nabla^{\parallel}\Omega \times \nabla\psi_v$  part of the third term in expression (5) is a correction to the field line bending. However, when  $\vec{B}_v$  is sufficiently strong, the perpendicular gradients of the hydromagnetic variables will dominate and  $\nabla^{\parallel}\Omega \times \nabla\psi_v$  will be small compared to  $\nabla^{\perp}\Omega \times \nabla\psi_v$ . Also, as will be shown in section V, the first term in expression (6) is exactly the  $\vec{E} \times \vec{B}$  flow only in ideal MHD. Finally, all terms in expression (6) have nonzero divergence, so the last term is not the exact fluid compression term.

We will now prove that any arbitrary vector field can be expressed in the form (6). Define three projection operators:

$$\begin{aligned} \nabla\chi \cdot \nabla \times [\nabla\chi \times (\vec{e}_{\chi} \times \\ \nabla\chi) \\ \nabla \cdot [B_v^2 \nabla\chi \times (\vec{e}_{\chi} \times \end{aligned} \quad (7)$$

Note that using the identity  $\nabla a \cdot \nabla \times \vec{A} = -\nabla \cdot (\nabla a \times \vec{A})$ , which follows directly from the cross product rule for the divergence, the first projection operator can alternatively be expressed as

$$-\nabla \cdot [\nabla \chi \times (\nabla \chi \times (\vec{e}_\chi \times \vec{v}))] \quad (8)$$

Also, notice that the effect of the  $\nabla \chi \times (\vec{e}_\chi \times \text{sub-operator})$  is to subtract out the contravariant  $\chi$  component of a vector:  $\nabla \chi \times (\vec{e}_\chi \times \vec{s}) = -\vec{s} + s^\chi \vec{e}_\chi$ . Applying each of the three operators to expression (6), we obtain equations for the scalar functions  $\Phi$ ,  $v_\parallel$  and  $\zeta$ :

$$\begin{aligned} \Delta^\perp \Phi &= \nabla \chi \cdot \nabla \times [\nabla \chi \times (\vec{e}_\chi \times \vec{v})], \\ v_\parallel &= \frac{v^\chi}{B\chi}, \end{aligned} \quad (9)$$

$$\nabla \cdot (B_v^2 \nabla^\perp \zeta) = -\nabla \cdot [B_v^2 \nabla \chi \times (\vec{e}_\chi \times \vec{v})],$$

where  $\Delta^\perp = \nabla \cdot \nabla^\perp$ . Thus, for any given  $\vec{v}$ , we have two uncoupled linear differential equations for  $\Phi$  and  $\zeta$  and a direct relation for  $v_\parallel$ . Both of the linear differential equations are generalized Poisson equations, and the boundary conditions can be obtained as follows. If  $\vec{v}$ ,  $\nabla \chi$  and  $\vec{B}$  are known everywhere, then a linear combination of Neumann boundary conditions for  $\Phi$  and  $\zeta$  can be determined by subtracting  $v_\parallel \vec{B}$  from (6), taking the cross product with  $\nabla \chi$  and then the dot product with the unit normal to the boundary:

$$-\vec{n} \cdot \nabla^\perp \Phi + \vec{n} \cdot (\nabla \zeta \times \nabla \chi) = \vec{n} \cdot [(\vec{v} - v_\parallel \vec{B}) \times \nabla \chi], \quad \vec{r} \in \partial V, \quad (10)$$

where  $\vec{n}$  is the unit normal vector to the boundary and  $\partial V$  is the boundary of the volume  $V$ . Having one boundary condition for two equations, we have the freedom to introduce a free function  $f(\vec{r})$  as follows:

$$\vec{n} \cdot \nabla^\perp \Phi = f(\vec{r}), \quad \vec{r} \in \partial V,$$

$$\vec{n} \cdot (\nabla \zeta \times \nabla \chi) = \vec{n} \cdot [(\vec{v} - v_\parallel \vec{B}) \times \nabla \chi] + f(\vec{r}), \quad \vec{r} \in \partial V. \quad (11)$$

The consistency condition for the first equation in (9)

$$\begin{aligned} \oint_{\partial V} \nabla^\perp \Phi \cdot d\vec{S} &= \int_V \nabla \chi \cdot \nabla \times [\nabla \chi \times (\vec{e}_\chi \times \vec{v})] dV \\ &= - \oint_{\partial V} [(\vec{v} - v_\parallel \vec{B}) \times \nabla \chi] \cdot d\vec{S} \end{aligned} \quad (12)$$

can be satisfied by requiring  $f(\vec{r})$  to satisfy  $\oint_{\partial V} f(\vec{r}) dS = - \oint_{\partial V} [(\vec{v} - v_\parallel \vec{B}) \times \nabla \chi] \cdot d\vec{S}$ . Thus,  $\Phi$  is guaranteed to exist, and, by the uniqueness theorem for Poisson's equation, is unique up to a constant when  $f(\vec{r})$  is specified.

For  $\zeta$ , another linear combination of Neumann boundary conditions can be obtained by subtracting  $v_\parallel \vec{B}$  from (6), multiplying by  $B_v^2$  and taking the dot product with the unit normal to the boundary:

$$\vec{n} \cdot (\nabla \Phi \times \nabla \chi) + B_v^2 \vec{n} \cdot \nabla^\perp \zeta = B_v^2 \vec{n} \cdot (\vec{v} - v_\parallel \vec{B}). \quad (13)$$

As before, a free function  $g(\vec{r})$  is introduced such that:

$$\begin{aligned} B_v^2 \vec{n} \cdot \nabla^\perp \zeta &= g(\vec{r}), \quad \vec{r} \in \partial V, \\ \vec{n} \cdot (\nabla \Phi \times \nabla \chi) &= \vec{n} \cdot (\vec{v} - v_\parallel \vec{B}) - g(\vec{r}), \quad \vec{r} \in \partial V, \end{aligned} \quad (14)$$

and the consistency condition for the third equation in (9)

$$\begin{aligned} \oint_{\partial V} B_v^2 \nabla^\perp \zeta \cdot d\vec{S} &= - \int_V \nabla \cdot [B_v^2 \nabla \chi \times (\vec{e}_\chi \times \vec{v})] dV \\ &= \oint_{\partial V} B_v^2 (\vec{v} - v_\parallel \vec{B}) \cdot d\vec{S} \end{aligned} \quad (15)$$

is satisfied by requiring  $\oint_{\partial V} g(\vec{r}) dS = \oint_{\partial V} B_v^2 (\vec{v} - v_\parallel \vec{B}) \cdot d\vec{S}$ . Thus,  $\zeta$  is also well-defined. In hindsight, after the existence of scalar functions  $\Phi$  and  $\zeta$  has been proven, the requirements imposed on  $f(\vec{r})$  and  $g(\vec{r})$  follow from the second conditions in (11) and (14) by integrating over  $\partial V$  and applying the divergence theorem to the LHS.

### III. MHD WAVES AND VELOCITY REPRESENTATION TERMS

In this section, we show in the framework of linearized MHD that the three MHD waves are approximately separated into three terms by the representation (6), with each term containing a wave and some instabilities. In addition to the usual assumptions of linearized MHD, we also assume that the induced equilibrium magnetic field is small compared to the vacuum field:

$$\frac{|\vec{B}_0 - \nabla \chi|}{|\nabla \chi|} \ll 1. \quad (16)$$

Thus, we can approximate  $\vec{B}_0$  by  $\nabla \chi$  in the following analysis. At the same time, we do not assume that  $\vec{j}_0 = \frac{1}{\mu_0} \nabla \times \vec{B}_0$  is negligible. However, the presence of equilibrium current will make no difference in the forthcoming analysis.

We begin with the linearized ideal MHD equation for velocity:

$$\begin{aligned} \rho_0 \frac{\partial^2 \vec{v}}{\partial t^2} &= \vec{j}_0 \times [\nabla \times (\vec{v} \times \nabla \chi)] \\ &+ \frac{1}{\mu_0} [\nabla \times [\nabla \times (\vec{v} \times \nabla \chi)]] \times \nabla \chi + \nabla (\vec{v} \cdot \nabla p_0) \\ &+ \gamma \nabla (p_0 \nabla \cdot \vec{v}). \end{aligned} \quad (17)$$

In general, both fluid-compressional and shear waves can propagate in a plasma, just like in an elastic solid. Comparing equation (17) with a typical elastic wave equation<sup>25</sup>, we see that the second term on the RHS has similar structure to the shear wave term in an elastic wave equation (this term can compress the magnetic field, but

not the fluid), whereas the last term is similar to the compressional wave term. The other two terms on the RHS of equation (17) do not have second order derivatives of  $\vec{v}$ .

We now evaluate the effect of each term in equation (6) by direct substitution into equation (17). Inserting just the first term from expression (6) into the above equation, one obtains:

$$\begin{aligned} \frac{\rho_0}{B_v^2} \frac{\partial^2 \nabla \Phi}{\partial t^2} \times \nabla \chi &= -\vec{j}_0 \times (\nabla \times \nabla^\perp \Phi) \\ &- \frac{1}{\mu_0} [\nabla \times (\nabla \times \nabla^\perp \Phi)] \times \nabla \chi \\ &+ \nabla \left( \frac{1}{B_v} [p_0, \Phi] \right) - 2\gamma \nabla \left( \frac{p_0}{B_v^2} [B_v, \Phi] \right), \end{aligned} \quad (18)$$

where  $[a, b] = B_v^{-1} \nabla \chi \cdot (\nabla a \times \nabla b)$  is the Poisson bracket for scalar fields  $a$  and  $b$ . The first term of expression (6) allows for shear Alfvén waves as well as various instabilities, for example, the first term on the RHS of equation (18) represents current-driven instabilities. We will only consider waves in this section, as a proper consideration of instabilities is better done with an energy principle. Several terms in equation (18) also account for distortions of waves due to inhomogeneity in the equilibrium, however we are only interested in identifying the MHD wave contained in each term of expression (6). Such identification is normally performed by determining the propagation speeds of wave-like perturbations, i.e. by determining the coefficient in the second order derivative (with respect to the unknown) in the wave equations<sup>26,27</sup>.

Let  $\nabla^\perp \Phi$ , which can be interpreted as the perpendicular component of the electric field in reduced ideal MHD (see section V), be the unknown in the vector wave equation. Then the second term on the RHS of (18), which is the only term that contains second order derivatives of  $\nabla^\perp \Phi$ , describes the propagation of waves. Note that what was the fluid-compressional wave term in equation (17) no longer contains second order derivatives of  $\nabla^\perp \Phi$  in equation (18). We focus attention on the second term, which can be rewritten as

$$-\frac{1}{\mu_0} [\nabla^\perp (\Delta^\perp \Phi) - \Delta^\perp (\nabla^\perp \Phi) - \Delta^\parallel (\nabla^\perp \Phi)] \times \nabla \chi. \quad (19)$$

We used the fact that  $\nabla \times (\nabla \times \vec{A}) = \nabla (\nabla \cdot \vec{A}) - \Delta \vec{A}$ ,  $\nabla = \nabla^\perp + \nabla^\parallel$  and  $\Delta = \Delta^\perp + \Delta^\parallel$ , where  $\Delta^\parallel = \nabla \cdot \nabla^\parallel$ . Now consider the two terms  $\nabla^\perp (\Delta^\perp \Phi)$  and  $-\Delta^\perp (\nabla^\perp \Phi)$ . The operators  $\nabla^\perp$  and  $\Delta^\perp$  do not commute when the background field is not uniform, however, as can be shown by lengthy algebraic manipulations, the highest order derivatives do indeed cancel, leaving only first and second order derivatives of  $\Phi$ . Thus, the only term containing third order derivatives of  $\Phi$  (i.e. second order derivatives of  $\nabla^\perp \Phi$ ) is  $-\Delta^\parallel (\nabla^\perp \Phi)$ . Cross multiplying equation (18)

by  $\nabla \chi$  from the the left and dividing by  $\rho_0$ , one obtains

$$\begin{aligned} \frac{\partial^2 \nabla^\perp \Phi}{\partial t^2} &= \frac{B_v^2}{\mu_0 \rho_0} [\Delta^\parallel (\nabla^\perp \Phi) + [\Delta^\perp, \nabla^\perp] \Phi] \\ &- \frac{\nabla \chi}{\mu_0 \rho_0} \cdot [\Delta^\parallel, \nabla^\perp] \Phi + [\Delta^\perp, \nabla^\perp] \Phi \\ &+ \frac{j_0^\chi}{\rho_0} \nabla \times \nabla^\perp \Phi + \frac{1}{\rho_0} \nabla \chi \times \nabla \left( \frac{1}{B_v} [p_0, \Phi] \right) \\ &- \frac{2\gamma}{\rho_0} \nabla \chi \times \nabla \left( \frac{p_0}{B_v^2} [B_v, \Phi] \right), \end{aligned} \quad (20)$$

where, by a slight abuse of notation, square brackets applied to two operators  $A, B$  that act on a function  $f$  is understood as the commutator of the operators:  $[A, B]f = A(Bf) - B(Af)$ . Similarly to  $[\Delta^\perp, \nabla^\perp]$ ,  $[\Delta^\parallel, \nabla^\perp]$  does not contain third order derivatives. Clearly, waves in the vector field  $\nabla^\perp \Phi$  will propagate along field lines with the Alfvén speed, while the velocity perturbation is perpendicular to the field lines. Thus, shear Alfvén waves are the only MHD waves allowed by the first term of expression (6). Note that, due to magnetic field nonuniformity, the first term in expression (6) has nonzero divergence. Although the Alfvén wave in a uniform background field is incompressible, this is not the case in a curved field<sup>26</sup>.

We now consider the second term in the expression (6). Note that in the context of linearized MHD with assumption (16) the second term can be approximated as:

$$v_\parallel \vec{B} = v_\parallel (\vec{B}_0 + \vec{B}_1) \approx v_\parallel \nabla \chi, \quad (21)$$

since both velocity and  $\vec{B}_1$  are first-order quantities. However, while using the full field instead of just the vacuum field in the second term makes no difference from the linear MHD wave point of view, the full field provides the advantage of allowing temperature and density profiles to flatten when the field becomes stochastic in a certain area, which would not be so simple in reduced MHD (see section V) if the parallel velocity was directed along the static background field.

Inserting the approximation for the second term (21) into equation (17), one obtains

$$\rho_0 \frac{\partial^2 v_\parallel}{\partial t^2} \nabla \chi = \gamma \nabla (B_v p_0 \partial^\parallel v_\parallel), \quad (22)$$

where  $\partial^\parallel = B_v^{-1} \nabla \chi \cdot \nabla$  is the spatial derivative along the vacuum field. The third term on the RHS of equation (17) was dropped due to the fact that  $\partial^\parallel p_0 \approx 0$  since  $\nabla p_0 = \vec{j}_0 \times \vec{B}_0$  and  $\vec{B}_0 \approx \nabla \chi$ . Note that only the fluid-compressional term from equation (17) survives. Expanding the RHS, multiplying by  $\nabla \chi$  and dividing by  $\rho_0 B_v^2$ , one obtains

$$\begin{aligned} \frac{\partial^2 v_\parallel}{\partial t^2} &= \frac{\gamma p_0}{\rho_0} (\partial^\parallel)^2 v_\parallel + \frac{\gamma p_0}{\rho_0 B_v} \partial^\parallel v_\parallel \partial^\parallel B_v \\ &= \frac{\gamma p_0}{\rho_0} \Delta^\parallel v_\parallel + \frac{2\gamma p_0}{\rho_0 B_v} \partial^\parallel v_\parallel \partial^\parallel B_v. \end{aligned} \quad (23)$$

We see that waves in the scalar field  $v_{\parallel}$  propagate with the sound speed along field lines while the velocity perturbation is parallel to the field lines. Thus, only slow magnetosonic waves are allowed by the second term of expression (6). The reason why these waves propagate with the sound speed instead of the slow magnetosonic speed is because we have constrained the velocity perturbation to be parallel to the background field, zeroing out the shear term and making it impossible for the wave to compress the magnetic field. Being able to only compress the fluid, the wave behaves as a sound wave. A true slow magnetosonic wave can exist when the third term of the expression (6) is also included, due to coupling between the second and third terms.

We now show that the first two terms in expression (6) do not compress the magnetic field even in the nonlinear regime. We start with the ideal MHD induction equation, insert expressions (5) and the first two terms of (6). Multiplying by  $\nabla\chi$  gives the component of  $\partial\vec{B}/\partial t$  along the vacuum magnetic field, which corresponds to field compression:

$$\begin{aligned} \frac{\partial B^{\chi}}{\partial t} &= B_v \left[ \frac{\partial\Omega}{\partial t}, \psi_v \right] = \nabla\chi \cdot \nabla \times (\vec{v} \times \vec{B}) \\ &= \nabla \cdot \left[ \left( -\nabla\Phi + \frac{\nabla\chi}{B_v^2} \nabla\chi \cdot \nabla\Phi - \frac{\nabla\chi}{B_v} [\Psi, \Phi] \right. \right. \\ &\quad \left. \left. + \frac{\nabla\Omega}{B_v} [\psi_v, \Phi] - \frac{\nabla\psi_v}{B_v} [\Omega, \Phi] \right) \times \nabla\chi \right] \\ &= B_v \left[ \frac{[\psi_v, \Phi]}{B_v}, \Omega \right] - B_v \left[ \frac{[\Omega, \Phi]}{B_v}, \psi_v \right]. \end{aligned} \quad (24)$$

As was mentioned in subsection II C,  $\nabla^{\perp}\Omega$  corresponds to field compression, and  $\nabla^{\parallel}\Omega$  will not contribute to the Poisson brackets above. From the above equation, if  $\nabla^{\perp}\Omega = 0$  initially, then  $\nabla^{\perp}\partial\Omega/\partial t = 0$ , and so  $\nabla^{\perp}\Omega$  will stay at zero. Thus, if there is no compression initially, the first two terms of expression (6) will not produce any compression. This is yet another advantage of using the full field in the second term of expression (6) instead of just the vacuum field. We note here that describing equilibria with a Shafranov shift is impossible without including vacuum field compression, and so setting  $\Omega = 0$  eliminates the possibility of consistently using tokamak equilibria obtained from the Grad-Shafranov equation as initial conditions. It also makes it impossible to directly use most stellarator equilibria, as those also include a Shafranov shift. One possibility is to obtain the initial conditions simply by neglecting the equilibrium field compression, as done in JOREK. In the case of a tokamak, this amounts to solving the Grad-Shafranov equation and using the results as initial conditions for  $\Psi$  and  $p$ , but keeping  $\Omega = 0$ , which is incompatible with the Grad-Shafranov expression for the magnetic field. Another possibility would be to allow field compression by evolving  $\Omega$  at the expense of more unknowns and more complicated equations, as done in the model by Izzo *et*

*al* and the M3D- $C^1$  four-field model<sup>11,12</sup>. However, it is usually advantageous to use the first method, as neglecting equilibrium field compression, which is anyway small due to (16), guarantees that the field will not be further compressed during the simulation, as shown above. Since field compression requires significant energy, there is a large restoring force associated with it, and unstable modes will generally not compress the field<sup>28</sup>. Thus, it makes sense to consistently neglect  $\Omega$ .

Finally, we consider the last term in the expression (6). Inserting it into equation (17), we obtain:

$$\begin{aligned} \rho_0 \frac{\partial^2 \nabla^{\perp} \zeta}{\partial t^2} &= \vec{j}_0 \times [\nabla \times (\nabla\zeta \times \nabla\chi)] \\ &\quad - \frac{1}{\mu_0} \Delta(\nabla\zeta \times \nabla\chi) \times \nabla\chi + \nabla(\zeta, p_0) \\ &\quad + \gamma \nabla(p_0 \Delta^{\perp} \zeta), \end{aligned} \quad (25)$$

where  $(a, b) = \nabla^{\perp}a \cdot \nabla^{\perp}b$  is the inner product of the perpendicular gradients of  $a$  and  $b$ . We take  $\nabla^{\perp}\zeta$  to be the unknown in the wave equation, the which is given by the perpendicular component of equation (25). Taking the perpendicular component of the equation, dividing by  $\rho_0$  and using  $\Delta(\vec{A} \times \vec{B}) = (\Delta\vec{A}) \times \vec{B} + 2\nabla\vec{A} \times \nabla\vec{B} + \vec{A} \times (\Delta\vec{B})$  to expand the RHS, we get:

$$\begin{aligned} \frac{\partial^2 \nabla^{\perp} \zeta}{\partial t^2} &= \frac{\vec{j}_0}{\rho_0} \times [\nabla \times (\nabla\zeta \times \nabla\chi)] \\ &\quad - \frac{\nabla\chi}{B_v \rho_0} \cdot [\vec{j}_0 \times [\nabla \times (\nabla\zeta \times \nabla\chi)]] \\ &\quad + \frac{B_v^2}{\mu_0 \rho_0} \Delta(\nabla^{\perp} \zeta) - \frac{\nabla\chi}{\mu_0 \rho_0} (\nabla\chi \cdot [\Delta, \nabla^{\perp}] \zeta) \\ &\quad + \frac{2}{\mu_0 \rho_0} \nabla\chi \times (\nabla \nabla^{\perp} \zeta \times \nabla \nabla\chi) + \frac{1}{\rho_0} \nabla^{\perp}(\zeta, p_0) \\ &\quad + \frac{\gamma}{\rho_0} \Delta^{\perp} \zeta \nabla^{\perp} p_0 + \frac{\gamma p_0}{\rho_0} \Delta^{\perp}(\nabla^{\perp} \zeta) \\ &\quad + \frac{\gamma p_0}{\rho_0} [\nabla^{\perp}, \Delta^{\perp}] \zeta, \end{aligned} \quad (26)$$

where  $\vec{T} \times \vec{U} = \epsilon_{ijk} J T^{lj} U_l^k \vec{e}^i$  is the dot-cross product of two tensors,  $\epsilon_{ijk}$  is the Levi-Civita symbol and  $J$  is the Jacobian. Just as before, the commutators do not contain third order derivatives. Only the third and second to last terms on the RHS of the above equation contain second order derivatives of  $\nabla^{\perp}\zeta$  (third order derivatives of  $\zeta$ ). The propagation of waves is described by these two terms, which can be rewritten as

$$\left( \frac{B_v^2}{\mu_0 \rho_0} + \frac{\gamma p_0}{\rho_0} \right) \Delta^{\perp}(\nabla^{\perp} \zeta) + \frac{B_v^2}{\mu_0 \rho_0} \Delta^{\parallel}(\nabla^{\perp} \zeta). \quad (27)$$

The wave can propagate both along and across field lines with different speeds. Now consider the fast magnetosonic wave speed in the form presented by Freidberg<sup>29</sup>:

$$c_f = \sqrt{\frac{1}{2}(c_A^2 + c_s^2) \left(1 + \sqrt{1 - 4 \cos^2 \theta \frac{c_A^2 c_s^2}{(c_A^2 + c_s^2)^2}}\right)}, \quad (28)$$

where  $\theta$  is the angle between the direction of wave propagation and the field,  $c_A = B_0/\sqrt{\mu_0 \rho_0}$  is the Alfvén speed and  $c_s = \sqrt{\gamma p_0/\rho_0}$  is the sound speed. To transform Freidberg's original expression, which was written in terms of the wave number  $k$ , the fact that  $k_{\parallel} = k \cos \theta$  was used. We see that the speed of the wave in equation (26) matches (28) both in the parallel ( $c_w = c_f = c_A$ ) and perpendicular ( $c_w = c_f = \sqrt{c_A^2 + c_s^2}$ ) directions, given that  $B_0 \approx B_v$ , as implied by assumption (16). In general, however, the speed of the wave in equation (26) is

$$\begin{aligned} c_w &= \sqrt{(c_A^2 + c_s^2) \sin^2 \theta + c_A^2 \cos^2 \theta} \\ &= \sqrt{c_A^2 + c_s^2 - c_s^2 \cos^2 \theta}. \end{aligned} \quad (29)$$

For  $\beta < 1$ , we have  $c_s < c_A$ , and as  $\beta \rightarrow 0$  both  $c_f \rightarrow c_A$  and  $c_w \rightarrow c_A$ . In the  $c_s \leq c_A$  regime, the discrepancy in the fast magnetosonic wave speed estimated by (29) will be maximized to about 9% when  $c_s = c_A$ , which corresponds to  $\beta = 2/\gamma$ , and when  $\theta = \arccos(\pm\sqrt{2\sqrt{2}-2})$ . The discrepancy arises because the third term of expres-

sion (6) was constrained to be orthogonal to magnetic field. Just like the slow magnetosonic wave, a true fast magnetosonic wave will only occur via coupling between the second and third terms. Nevertheless, the approximate separation of the slow and fast waves provided by second and third terms is fairly accurate for low  $\beta$ , and the third term does manage to separate out the fast  $\sqrt{c_A^2 + c_s^2}$  dynamics, the removal of which decreases the stiffness of the equations and is sufficient for the purposes of reduced MHD (see section V).

#### IV. DERIVATION OF THE EQUATIONS

In this section, we derive scalar equations for the potentials  $\Psi$ ,  $\Omega$ ,  $\Phi$  and  $\zeta$ , and the parallel component  $v_{\parallel}$  from the vector equations in (1) by inserting expressions (5) and (6) into them and applying projection operators. Since any arbitrary magnetic field and velocity can be represented in the forms (5) and (6), the scalar equations that we derive are still full MHD equations. The reduction procedure is applied separately from the derivation in section V.

The continuity and energy conservation equations (1), which we will use to evolve density and pressure, can be employed directly by inserting expressions (5) and (6) into them:

$$\frac{\partial \rho}{\partial t} = -B_v \left[ \frac{\rho}{B_v^2}, \Phi \right] - B_v \partial^{\parallel} (\rho v_{\parallel}) - B_v [\rho v_{\parallel}, \Psi] - F_v [\rho v_{\parallel}, \Omega]_{\psi_v} - (\rho, \zeta) - \rho \Delta^{\perp} \zeta + P, \quad (30)$$

and

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{\rho v^2}{2} + \frac{p}{\gamma-1} + \frac{B^2}{2\mu_0} \right) &= -B_v \left[ \frac{1}{B_v^2} \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0} \right), \Phi \right] - B_v \partial^{\parallel} \left[ v_{\parallel} \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0} \right) - \frac{\vec{v} \cdot \vec{B}}{\mu_0} \right] \\ &- B_v \left[ v_{\parallel} \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0} \right) - \frac{\vec{v} \cdot \vec{B}}{\mu_0}, \Psi \right] - F_v \left[ v_{\parallel} \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0} \right) - \frac{\vec{v} \cdot \vec{B}}{\mu_0}, \Omega \right]_{\psi_v} \\ &- \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0}, \zeta \right) - \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0} \right) \Delta^{\perp} \zeta - \frac{1}{\mu_0} \nabla \cdot (\eta \vec{j} \times \vec{B}) \\ &+ \nabla \cdot \left[ \frac{p}{\gamma-1} \frac{D_{\perp}}{\rho} \left( \nabla \rho - \frac{\vec{B}}{B^2} (B_v \partial^{\parallel} \rho + B_v [\rho, \Psi] + F_v [\rho, \Omega]_{\psi_v}) \right) \right] \\ &+ \frac{\kappa_{\perp}}{R} \nabla \left( \frac{p}{\rho} \right) + \frac{(\kappa_{\parallel} - \kappa_{\perp}) \vec{B}}{R B^2} \left( B_v \partial^{\parallel} \left( \frac{p}{\rho} \right) + B_v \left[ \frac{p}{\rho}, \Psi \right] + F_v \left[ \frac{p}{\rho}, \Omega \right]_{\psi_v} \right) \right] + S_e - \frac{v^2}{2} P, \end{aligned} \quad (31)$$

where  $F_v = |\nabla \psi_v|$  and  $[a, b]_{\psi_v} = F_v^{-1} \nabla \psi_v \cdot (\nabla a \times \nabla b)$  is the Poisson bracket for scalar fields  $a$  and  $b$  with respect to



$\nabla\psi_v$ . In addition,

$$\begin{aligned}
P &= \nabla \cdot \left[ D_{\perp} \nabla \rho - \frac{D_{\perp} \vec{B}}{B^2} (B_v \partial^{\parallel} \rho + B_v [\rho, \Psi] + F_v [\rho, \Omega]_{\psi_v}) \right] + S_{\rho}, \\
v^2 &= \frac{1}{B_v^2} (\Phi, \Phi) + 2v_{\parallel} \vec{v} \cdot \vec{B} + \frac{2}{B_v} [\zeta, \Phi] - v_{\parallel}^2 B^2 + (\zeta, \zeta), \\
\vec{v} \cdot \vec{B} &= (\Phi, \Psi) - \frac{F_v}{B_v} \partial_{\psi_v} \Phi \partial^{\parallel} \Omega + v_{\parallel} B^2 + B_v [\zeta, \Psi] + F_v [\zeta, \Omega]_{\psi_v} - \partial^{\parallel} \zeta [\Omega, \psi_v], \\
B^2 &= B_v^2 + 2B_v [\Omega, \psi_v] + B_v^2 (\Psi, \Psi) - 2B_v F_v \partial_{\psi_v} \Psi \partial^{\parallel} \Omega + F_v^2 (\Omega, \Omega)_{\psi_v}, \\
\vec{j} &= \frac{1}{\mu_0} [-\nabla \chi \Delta \Psi + B_v \partial^{\parallel} \nabla \Psi - (\nabla \Psi \cdot \nabla) \nabla \chi + \nabla \Omega \Delta \psi_v - \nabla \psi_v \Delta \Omega + F_v \partial_{\psi_v} \nabla \Omega - (\nabla \Omega \cdot \nabla) \nabla \psi_v].
\end{aligned} \tag{32}$$

where  $\partial_{\psi_v} = F_v^{-1} \nabla \psi_v \cdot \nabla$  is the spatial derivative in the direction of  $\nabla \psi_v$ , and  $(a, b)_{\psi_v} = \nabla a \cdot \nabla b - \partial_{\psi_v} a \partial_{\psi_v} b$  is the inner product of gradients of scalar functions  $a$  and  $b$  perpendicular to  $\nabla \psi_v$ .

We now proceed to derive the scalar equations for the magnetic potentials. Inserting expressions (5) and (6) into the induction equation (1), we obtain:

$$\begin{aligned}
\nabla \frac{\partial \Psi}{\partial t} \times \nabla \chi + \nabla \frac{\partial \Omega}{\partial t} \times \nabla \psi_v &= \nabla \times \left[ \frac{\nabla \chi}{B_v} (\partial^{\parallel} \Phi - [\Psi, \Phi]) + \frac{\nabla \Omega}{B_v} [\psi_v, \Phi] - \frac{\nabla \psi_v}{B_v} [\Omega, \Phi] + \nabla \zeta \times \nabla \chi \right. \\
&\quad \left. - \nabla \chi (\zeta, \Psi) + \nabla \Omega (\zeta, \psi_v) - \nabla \psi_v (\zeta, \Omega) - \eta \vec{j} \right].
\end{aligned} \tag{33}$$

Projecting this vector equation on the  $\nabla \psi_v$  and the  $\nabla \chi$  directions, we obtain scalar evolution equations for the  $\Psi$  and  $\Omega$  potentials:

$$\begin{aligned}
\left[ \psi_v, \frac{\partial \Psi}{\partial t} \right] &= \left[ \frac{[\Psi, \Phi] - \partial^{\parallel} \Phi}{B_v}, \psi_v \right] - \frac{F_v}{B_v} \left[ \Omega, \frac{[\psi_v, \Phi]}{B_v} \right]_{\psi_v} + \partial^{\parallel} (F_v \partial_{\psi_v} \zeta) + [(\zeta, \Psi), \psi_v] - \frac{F_v}{B_v} [\Omega, (\zeta, \psi_v)]_{\psi_v} \\
&\quad + \frac{1}{B_v} \nabla \cdot (\eta \nabla \psi_v \times \vec{j}), \\
\left[ \frac{\partial \Omega}{\partial t}, \psi_v \right] &= - \left[ \Omega, \frac{[\psi_v, \Phi]}{B_v} \right] + \left[ \psi_v, \frac{[\Omega, \Phi]}{B_v} \right] - 2(B_v, \zeta) - B_v \Delta^{\perp} \zeta - [\Omega, (\zeta, \psi_v)] + [\psi_v, (\zeta, \Omega)] \\
&\quad + \frac{1}{B_v} \nabla \cdot (\eta \nabla \chi \times \vec{j}).
\end{aligned} \tag{34}$$

where we have used the same identity as in (8) to simplify the projections on the right hand side. We could also have obtained evolution equations for  $\Psi$  and  $\Phi$  by using the potential form of Faraday's law, however in that case, due to our choice for the magnetic vector potential (4), we are no longer free in our choice of the electric potential  $V$ , which must be chosen so that  $\partial V / \partial \beta_v$  cancels with  $E_{\beta_v}$ , the covariant  $\beta_v$  component of the electric field. This would produce complicated integro-differential equations for  $\Psi$  and  $\Omega$ .

To obtain the scalar equations for the potentials  $\Phi$  and  $\zeta$  and the parallel component  $v_{\parallel}$ , we begin by inserting expressions (5) and (6) into the Navier-Stokes equation (1), dividing by  $\rho$  and then apply the projection operators (7). Expanding the time derivative, dividing by  $\rho$  and

inserting the appropriate expressions, we have:

$$\begin{aligned}
\nabla \frac{\partial \Phi}{\partial t} \times \frac{\nabla \chi}{B_v^2} + \frac{\partial v_{\parallel}}{\partial t} \vec{B} + v_{\parallel} \frac{\partial \vec{B}}{\partial t} + \nabla^{\perp} \frac{\partial \zeta}{\partial t} + \frac{\vec{v}}{\rho} P \\
+ \frac{1}{2} \nabla v^2 + \vec{\omega} \times \vec{v} = \frac{1}{\rho} \vec{j} \times \vec{B} - \frac{1}{\rho} \nabla p,
\end{aligned} \tag{35}$$

where we used the identity  $(\vec{v} \cdot \nabla) \vec{v} = \frac{1}{2} \nabla v^2 + \vec{\omega} \times \vec{v}$  and  $\vec{\omega}$  is the vorticity:

$$\begin{aligned}
\vec{\omega} &= \nabla \times \vec{v} = -\nabla \chi \nabla \cdot \left( \frac{\nabla \Phi}{B_v^2} \right) + B_v \partial^{\parallel} \frac{\nabla \Phi}{B_v^2} \\
&\quad - \frac{1}{B_v^2} (\nabla \Phi \cdot \nabla) \nabla \chi + \nabla v_{\parallel} \times \vec{B} + v_{\parallel} \vec{j} + \nabla \chi \times \nabla \frac{\partial^{\parallel} \zeta}{B_v}.
\end{aligned} \tag{36}$$

As discussed in subsection II A, the viscous term is not treated in this derivation. Applying the  $\nabla \chi \times$  sub-operator to equation (35), we obtain:

$$\begin{aligned}
& -\nabla \frac{\partial \Phi}{\partial t} \times \frac{\nabla \chi}{B_v^2} - v_{\parallel} \nabla \frac{\partial \Psi}{\partial t} \times \nabla \chi - v_{\parallel} \nabla \frac{\partial \Omega}{\partial t} \times \nabla \psi_v + \vec{e}_{\chi} v_{\parallel} B_v \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right] - \nabla^{\perp} \frac{\partial \zeta}{\partial t} = \frac{1}{2} \nabla v^2 - \vec{e}_{\chi} \frac{B_v}{2} \partial^{\parallel} v^2 \\
& + \left( \frac{\nabla \Phi \times \nabla \chi}{B_v^2} + \nabla^{\perp} \zeta \right) \frac{P}{\rho} - v_{\chi} \nabla \chi \times \vec{\omega} + \omega_{\chi} \nabla \chi \times \vec{v} + \frac{B^2}{\rho B^{\chi}} \nabla \chi \times \vec{j} - \frac{j_{\chi}}{\rho} \nabla \chi \times \vec{B} + \frac{\nabla p - \vec{e}_{\chi} B_v \partial^{\parallel} p}{\rho}.
\end{aligned} \tag{37}$$

Proceeding to obtain the equation for  $\Phi$ , we apply the remainder of the projection operator for  $\Phi$ , namely  $\nabla \chi \cdot (\nabla \times$ , or its equivalent  $-\nabla \cdot (\nabla \chi \times$ , when appropriate, to equation (37):

$$\begin{aligned}
& \Delta^{\perp} \frac{\partial \Phi}{\partial t} + \left( v_{\parallel} B_v^2, \frac{\partial \Psi}{\partial t} \right) + v_{\parallel} B_v^2 \Delta^{\perp} \frac{\partial \Psi}{\partial t} - F_v \partial_{\psi_v} \left( v_{\parallel} B_v \partial^{\parallel} \frac{\partial \Omega}{\partial t} \right) - v_{\parallel} B_v \partial^{\parallel} \frac{\partial \Omega}{\partial t} \Delta \psi_v - \left( \frac{v_{\parallel} B_v^2}{B_v + [\Omega, \psi_v]} \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right], \Psi \right) \\
& - \frac{v_{\parallel} B_v^2}{B_v + [\Omega, \psi_v]} \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right] \Delta^{\perp} \Psi + F_v \partial_{\psi_v} \left( \frac{v_{\parallel} B_v \partial^{\parallel} \Omega}{B_v + [\Omega, \psi_v]} \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right] \right) + \frac{v_{\parallel} B_v \partial^{\parallel} \Omega}{B_v + [\Omega, \psi_v]} \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right] \Delta \psi_v \\
& = \nabla \cdot \left[ \frac{B_v \nabla^{\perp} \Psi - \partial^{\parallel} \Omega \nabla \psi_v}{B_v + [\Omega, \psi_v]} \frac{B_v \partial^{\parallel} v^2}{2} + \omega_{\chi} (\nabla \Phi \times \nabla \chi + B_v^2 v_{\parallel} \vec{B}^{\perp} + B_v^2 \nabla^{\perp} \zeta) - v_{\chi} B_v^2 \vec{\omega}^{\perp} + \frac{B^2 B_v}{\rho (B_v + [\Omega, \psi_v])} \vec{j}^{\perp} \right. \\
& \left. - \frac{j_{\chi} B_v^2}{\rho} \vec{B}^{\perp} - (\nabla^{\perp} \Phi + \nabla \chi \times \nabla \zeta) \frac{P}{\rho} \right] + B_v \left[ \frac{1}{\rho}, p \right] + \left( \frac{B_v^2 \partial^{\parallel} p}{\rho (B_v + [\Omega, \psi_v])}, \Psi \right) + \frac{B_v^2 \partial^{\parallel} p}{\rho (B_v + [\Omega, \psi_v])} \Delta^{\perp} \Psi \\
& - F_v \partial_{\psi_v} \left( \frac{B_v \partial^{\parallel} \Omega \partial^{\parallel} p}{\rho (B_v + [\Omega, \psi_v])} \right) - \frac{B_v \partial^{\parallel} \Omega \partial^{\parallel} p}{\rho (B_v + [\Omega, \psi_v])} \Delta \psi_v + \nu \Delta \Delta^{\perp} \Phi,
\end{aligned} \tag{38}$$

where we have added a generic viscosity term  $\nu \Delta \Delta^{\perp} \Phi$ , as discussed in subsection II A. Following the same approach that was used by Franck *et al*<sup>13</sup>, we choose the form of the viscosity term by allowing the projection operator to act directly on  $\vec{v}$ , bypassing the  $\nu \Delta$  operator. Here, for any vector  $\vec{s}$ ,  $\vec{s}^{\perp} = \vec{s} - s^{\chi} \nabla \chi / B_v^2$  is the vector

component of  $\vec{s}$  perpendicular to the vacuum field. Subscripts and superscripts on  $v, \omega, j$  and  $B$  denote covariant and contravariant components, respectively.

To get the equation for the parallel component  $v_{\parallel}$ , we apply the projection operator  $\nabla \chi \cdot$  to equation (35) and divide by  $B^{\chi}$ :

$$\begin{aligned}
& \frac{\partial v_{\parallel}}{\partial t} + \frac{v_{\parallel}}{B_v + [\Omega, \psi_v]} \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right] = -\frac{v_{\parallel}}{\rho} P + \frac{1}{B_v + [\Omega, \psi_v]} \left[ \frac{\vec{\omega} \cdot \nabla^{\perp} \Phi}{B_v} + B_v v_{\parallel} \vec{\omega} \cdot \nabla^{\perp} \Psi - v_{\parallel} \omega^{\psi_v} \partial^{\parallel} \Omega - \frac{\vec{\omega} \cdot (\nabla \zeta \times \nabla \chi)}{B_v} \right. \\
& \left. - \frac{\partial^{\parallel} v^2}{2} - \frac{B_v \vec{j} \cdot \nabla^{\perp} \Psi + j^{\psi_v} \partial^{\parallel} \Omega - \partial^{\parallel} p}{\rho} \right] + \nu \Delta v_{\parallel},
\end{aligned} \tag{39}$$

where we have again allowed the projection operator  $(B^{\chi})^{-1} \nabla \chi \cdot$  to act directly on  $\vec{v}$ .

Finally, to get the equation for  $\zeta$ , we apply the remainder of the projection operator for  $\zeta$ , namely  $\nabla \cdot (B_v^2$  to equation (37):

$$\begin{aligned}
& B_v^2 \Delta^{\perp} \frac{\partial \zeta}{\partial t} + \left( B_v^2, \frac{\partial \zeta}{\partial t} \right) + B_v \left[ v_{\parallel} B_v^2, \frac{\partial \Psi}{\partial t} \right] + F_v \left[ v_{\parallel} B_v^2, \frac{\partial \Omega}{\partial t} \right]_{\psi_v} - B_v \partial^{\parallel} \left( \frac{v_{\parallel} B_v^2}{B_v + [\Omega, \psi_v]} \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right] \right) \\
& - B_v \left[ \frac{v_{\parallel} B_v^2}{B_v + [\Omega, \psi_v]} \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right], \Psi \right] - F_v \left[ \frac{v_{\parallel} B_v^2}{B_v + [\Omega, \psi_v]} \left[ \frac{\partial \Omega}{\partial t}, \psi_v \right], \Omega \right]_{\psi_v} = \nabla \cdot \left[ B_v^2 v_{\chi} \nabla \chi \times \vec{\omega} - \frac{B_v^2}{2} \nabla v^2 \right. \\
& \left. - B_v^2 \omega_{\chi} (\nabla^{\perp} \Phi + v_{\parallel} B_v^2 \nabla^{\perp} \Psi - v_{\parallel} B_v \partial^{\parallel} \Omega \nabla \psi_v + \nabla \chi \times \nabla \zeta) - (\nabla \Phi \times \nabla \chi + B_v^2 \nabla^{\perp} \zeta) \frac{P}{\rho} - \frac{B_v B^2}{\rho (B_v + [\Omega, \psi_v])} \nabla \chi \times \vec{j} \right] \\
& + \frac{B_v^3 j_{\chi}}{\rho} (B_v \nabla^{\perp} \Psi - \partial^{\parallel} \Omega \nabla \psi_v) + \frac{B_v^2 (\nabla \chi + \nabla \Psi \times \nabla \chi + \nabla \Omega \times \nabla \psi_v) \partial^{\parallel} v^2}{2 (B_v + [\Omega, \psi_v])} - \frac{B_v^2}{\rho} \Delta p - \nabla \left( \frac{B_v^2}{\rho} \right) \cdot \nabla p \\
& + B_v \partial^{\parallel} \left( \frac{B_v^2 \partial^{\parallel} p}{\rho (B_v + [\Omega, \psi_v])} \right) + B_v \left[ \frac{B_v^2 \partial^{\parallel} p}{\rho (B_v + [\Omega, \psi_v])}, \Psi \right] + F_v \left[ \frac{B_v^2 \partial^{\parallel} p}{\rho (B_v + [\Omega, \psi_v])}, \Omega \right]_{\psi_v} + \nu B_v^2 \Delta \Delta^{\perp} \zeta.
\end{aligned} \tag{40}$$

In this equation, in addition to allowing the projection operator to act directly on  $\vec{v}$ , we also allow  $B_v^2$  to pass

through the divergence and Laplacian operators after the projection operator has acted on  $\vec{v}$ . This does not introduce any new error since we already allowed  $B_v^2$  to pass through the Laplacian as part of the projection operator. Strictly speaking, for non-negligible viscosity, the approximations applied to the viscosity term in equations (38), (39) and (40) are only valid when both the vacuum field and the full magnetic field are approximately uniform, however the viscosity term in the Navier-Stokes equation (1) does not accurately model viscous effects in a plasma anyway, and the approximations should still give the correct order of magnitude even when the magnetic field cannot be approximated as uniform<sup>13</sup>.

We point out that third- and fourth-order spatial derivatives arise in equations (34), (38) and (40). This can be problematic when the unknown functions are interpolated with third-order polynomials, an approach used by the JOREK code<sup>30</sup>. To mitigate this problem, we express terms with third order derivatives as divergences and terms with fourth order derivatives as Laplacians, which allows one to reduce the order of the derivatives by applying integration by parts in the weak form of the equations. Finally, we note that, except for the viscosity term, no other approximations were made in this section, and the equations we derived still correspond to full MHD, albeit in a potential form.

Finally, from the form of equation (31), it may seem that at  $\beta \ll 1$  the magnetic energy to be the dominant term in the time derivative. Since equation (31) is used to evolve the pressure, which requires subtracting the changes in kinetic and magnetic energy from the change in total local energy over one time step, this would lead to a large numerical error in the pressure change due to being mixed with the error in the magnetic energy change. Fortunately, this is not the case. If we apply an ordering, then the time derivatives of the kinetic, internal and magnetic energies will all have exactly the same order in  $\lambda$ , as shown in appendix A. Without a strict ordering, we can still argue that the time derivatives of internal and magnetic energy will be comparable. Using the expression for  $B^2$  given in (32) we have:

$$\frac{\partial}{\partial t} \left( \frac{B^2}{2\mu_0} \right) \sim \frac{B_v^2}{2\mu_0} \frac{\partial}{\partial t} (\Psi, \Psi),$$

where we have assumed that  $\partial\Omega/\partial t$  is negligible. This is because field compression is associated with a large restoring force. We have also assumed that the magnitude of the  $\nabla\Omega \times \nabla\psi_v$  term in the magnetic field is of the same order or smaller than the magnitude of the  $\nabla\Psi \times \nabla\chi$  term. We can estimate the relative magnitudes of  $p$  and  $\Psi$  from the Grad-Shafranov equation (the orders of magnitude should be about the same for a stellarator):

$$\begin{aligned} \Delta^* \Psi &= -\frac{\mu_0 R^2}{F_0^2} \frac{dp}{d\Psi} - \frac{F}{F_0} \frac{dF}{d\Psi} \\ \Rightarrow p &\sim \frac{F_0^2}{\mu_0 R^2} \Psi \Delta^* \Psi \sim \frac{F_0^2}{\mu_0 R^2} |\nabla^\perp \Psi|^2, \end{aligned}$$

where we have assumed that the transient values of  $p$  and  $\Psi$  persist at the same order of magnitude as the equilibrium values  $p_0$  and  $\Psi_0$ . As is usual in the tokamak limit, we have  $F_0 = RB_v = \text{const.}$  Comparing the time derivatives, we see that  $\partial_t[B^2/(2\mu_0)] \div \partial_t p \sim 1$ .

## V. REDUCED MHD

Although there are many approaches to reduced MHD, which often involve an expansion with respect to the inverse aspect ratio, the common goal of all these models is the elimination of fast magnetosonic waves<sup>1,3,6-8,13</sup>. By eliminating the fastest propagating waves, we decrease the maximum velocity in the system, thus increasing the maximum time step allowed by the Courant condition in numerical simulations with explicit time integration. When implicit methods are used, the Courant condition is no longer a hard limit, however using time steps that are large compared to the shortest time scale can lead to particularly stiff matrix systems and poor accuracy<sup>2,3</sup>.

In this paper, we adopt an ansatz approach to reduced MHD, which does not rely on a large aspect ratio, does not assume an ordering and, in fact, does not require a toroidal geometry at all. While the assumption (16) is identical to the first assumption of the ordering in<sup>1</sup>, we do not *a priori* introduce further assumptions on the magnitudes of the other hydromagnetic variables. We will, however, derive further conditions from the equations themselves, see section VI B. In appendix A, we will show an alternative ordering-based approach which also does not assume a toroidal geometry. As long as assumption (16) is met, i.e. in the presence of a strong guiding field, we can eliminate fast magnetosonic waves by setting  $\zeta = 0$ . In addition, since the first two terms of the velocity expression (6) do not compress the magnetic field, we can also set  $\vec{\Omega} = 0$ , further simplifying the equations. However, since  $\vec{B}_0 \approx \nabla\chi$  is only an approximation, the wave equations in section III will only hold approximately. Thus, some residual effects of the fast magnetosonic wave can remain even after setting  $\zeta = 0$ . Care must be taken when using an explicit time stepping method, as these residual effects can prevent the usage of time steps larger than the fast magnetosonic time scale, for this reason it is preferable to use implicit schemes as is done in modern codes. When implicit time stepping is used, the residual effects amount to large eigenvalues in the resulting matrices, which can lead to convergence issues with iterative solvers. These problems can usually be mitigated by applying a preconditioner<sup>2,13</sup>.

Having eliminated two of the dependent variables, we can also drop the corresponding equations. In such a manner, we obtain the following set of reduced MHD equations:

$$\frac{\partial \rho}{\partial t} = -B_v \left[ \frac{\rho}{B_v^2}, \Phi \right] - B_v \partial^{\parallel} (\rho v_{\parallel}) - B_v [\rho v_{\parallel}, \Psi] + P, \quad (41)$$

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{\rho v^2}{2} + \frac{p}{\gamma-1} + \frac{B^2}{2\mu_0} \right) &= -B_v \left[ \frac{1}{B_v^2} \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0} \right), \Phi \right] - B_v \partial^{\parallel} \left[ v_{\parallel} \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0} \right) - \frac{\vec{v} \cdot \vec{B}}{\mu_0} \right] \\ &- B_v \left[ v_{\parallel} \left( \frac{\rho v^2}{2} + \frac{\gamma p}{\gamma-1} + \frac{B^2}{\mu_0} \right) - \frac{\vec{v} \cdot \vec{B}}{\mu_0}, \Psi \right] - \frac{1}{\mu_0} \nabla \cdot (\eta \vec{j} \times \vec{B}) \\ &+ \nabla \cdot \left[ \frac{p}{\gamma-1} \frac{D_{\perp}}{\rho} \left( \nabla \rho - \frac{\vec{B}}{B^2} (B_v \partial^{\parallel} \rho + B_v [\rho, \Psi]) \right) \right] \\ &+ \frac{\kappa_{\perp}}{R} \nabla \left( \frac{p}{\rho} \right) + \frac{(\kappa_{\parallel} - \kappa_{\perp}) \vec{B}}{R B^2} \left( B_v \partial^{\parallel} \left( \frac{p}{\rho} \right) + B_v \left[ \frac{p}{\rho}, \Psi \right] \right) \right] + S_e - \frac{v^2}{2} P \end{aligned} \quad (42)$$

where we now have

$$\begin{aligned} P &= \nabla \cdot \left[ D_{\perp} \nabla \rho - \frac{D_{\perp} \vec{B}}{B^2} (B_v \partial^{\parallel} \rho + B_v [\rho, \Psi]) \right] + S_{\rho}, \\ \vec{B} &= \nabla \chi + \nabla \Psi \times \nabla \chi, \\ \vec{v} &= \frac{\nabla \Phi \times \nabla \chi}{B_v^2} + v_{\parallel} \vec{B}, \\ v^2 &= \frac{1}{B_v^2} (\Phi, \Phi) + 2v_{\parallel} (\Phi, \Psi) + v_{\parallel}^2 B_v^2 + v_{\parallel}^2 B_v^2 (\Psi, \Psi), \\ \vec{v} \cdot \vec{B} &= (\Phi, \Psi) + v_{\parallel} B_v^2 + v_{\parallel} B_v^2 (\Psi, \Psi), \\ B^2 &= B_v^2 + B_v^2 (\Psi, \Psi), \\ \vec{j} &= \frac{1}{\mu_0} [-\nabla \chi \Delta \Psi + B_v \partial^{\parallel} \nabla \Psi - (\nabla \Psi \cdot \nabla) \nabla \chi], \\ \vec{\omega} &= \nabla \times \vec{v} = -\nabla \chi \nabla \cdot \left( \frac{\nabla \Phi}{B_v^2} \right) + B_v \partial^{\parallel} \frac{\nabla \Phi}{B_v^2} - \frac{1}{B_v^2} (\nabla \Phi \cdot \nabla) \nabla \chi + \nabla v_{\parallel} \times \vec{B} + v_{\parallel} \vec{j}. \end{aligned} \quad (43)$$

Reduction of the flux and momentum equations results in the following three equations:

$$\left[ \psi_v, \frac{\partial \Psi}{\partial t} \right] = \left[ \frac{[\Psi, \Phi] - \partial^{\parallel} \Phi}{B_v}, \psi_v \right] + \frac{1}{B_v} \nabla \cdot (\eta \nabla \psi_v \times \vec{j}) \quad (44)$$

$$\begin{aligned} \Delta^{\perp} \frac{\partial \Phi}{\partial t} + \left( v_{\parallel} B_v^2, \frac{\partial \Psi}{\partial t} \right) + v_{\parallel} B_v^2 \Delta^{\perp} \frac{\partial \Psi}{\partial t} &= \nabla \cdot \left[ \frac{B_v \partial^{\parallel} v^2}{2} \nabla^{\perp} \Psi + \omega_{\chi} (\nabla \Phi \times \nabla \chi + B_v^2 v_{\parallel} \nabla \Psi \times \nabla \chi) - v_{\chi} B_v^2 \vec{\omega}^{\perp} \right] \\ &+ \frac{B^2}{\rho} \vec{j}^{\perp} - \frac{j_{\chi} B_v^2}{\rho} \nabla \Psi \times \nabla \chi - \frac{P}{\rho} \nabla^{\perp} \Phi \left] + B_v \left[ \frac{1}{\rho}, p \right] + \left( \frac{B_v \partial^{\parallel} p}{\rho}, \Psi \right) + \frac{B_v \partial^{\parallel} p}{\rho} \Delta^{\perp} \Psi + \nu \Delta \Delta^{\perp} \Phi \end{aligned} \quad (45)$$

$$\frac{\partial v_{\parallel}}{\partial t} = -\frac{v_{\parallel}}{\rho} P + \frac{1}{B_v} \left[ \frac{\vec{\omega} \cdot \nabla^{\perp} \Phi}{B_v} + B_v v_{\parallel} \vec{\omega} \cdot \nabla^{\perp} \Psi - \frac{\partial^{\parallel} v^2}{2} - \frac{B_v \vec{j} \cdot \nabla^{\perp} \Psi - \partial^{\parallel} p}{\rho} \right] + \nu \Delta v_{\parallel}, \quad (46)$$

A further simplification would be to set  $v_{\parallel} = 0$ , reducing the number of unknowns to four at the expense of field-aligned flows. Similar reduced models for tokamaks that include field-aligned flows, as well as variants which

only allow flow perpendicular to the background field, are used in the JOREK and M3D- $C^1$  codes<sup>12,14</sup>.

As a final remark, we show that in the reduced ideal case the scalar function  $\Phi$  introduced in the velocity ex-

pression (6) corresponds, up to an additive constant, to the electric potential taken with the opposite sign. Faraday's law in potential form states

$$\frac{\partial \vec{A}}{\partial t} = -\vec{E} - \nabla V, \quad (47)$$

where  $V$  is the electric potential and  $\vec{A}$  is the magnetic vector potential. Using the ideal Ohm's law  $\vec{E} = -\vec{v} \times \vec{B}$  with the reduced expressions for  $\vec{v}$  and  $\vec{B}$  from (43) and the expression (4) for  $\vec{A}$  with  $\Omega = 0$  (where we have dropped the vacuum field vector potential due to its static nature), we obtain

$$\frac{\partial \Psi}{\partial t} \nabla \chi = -\nabla^\perp \Phi - \frac{\nabla \chi}{B_v} [\Psi, \Phi] - \nabla V. \quad (48)$$

Taking just the components perpendicular to the vacuum field, we get  $\nabla^\perp \Phi = -\nabla^\perp V$ , which requires  $\Phi = -V + c(\chi)$ , where  $c(\chi)$  is an arbitrary function. The component along the vacuum field is essentially an evolution equation for  $\Psi$ :

$$\frac{\partial \Psi}{\partial t} \nabla \chi = (\partial^\parallel \Phi - c' - [\Psi, \Phi]) \frac{\nabla \chi}{B_v}, \quad (49)$$

where we have replaced  $V$  with  $-\Phi + c(\chi)$ . To show that this equation is consistent with equation (44), we take the curl, project it on  $\nabla \psi_v$  and divide by  $B_v$ , obtaining

$$\left[ \psi_v, \frac{\partial \Psi}{\partial t} \right] = \left[ \frac{[\Psi, \Phi] - c' - \partial^\parallel \Phi}{B_v}, \psi_v \right], \quad (50)$$

which, if  $c' = 0$ , is exactly the ideal version of equation (44). Thus,  $c$  is an additive constant.

## VI. CONSERVATION PROPERTIES

In this section we consider sources of error and validity conditions for the reduced MHD approximation by looking at the components of the MHD equations (1) that are dropped in the reduction. Since all of the MHD equations (1), except for the induction equation, are local conservation laws, any error introduced by the reduction amounts to a non-conservation of the corresponding quantity. For the induction equation, which allows for non-conservation of flux even when it is satisfied exactly, we will consider both the physical non-conservation of flux due to resistivity as well as the reduction error.

The conservation of mass and energy is exact, due to the fact that the continuity and energy equations are used directly to evolve density and pressure, with non-conservation being only due to the terms  $S_\rho$  and  $S_e$ , which correspond to physically meaningful sources. On the other hand, momentum is not conserved due to equation (40) being discarded after the reduction. Indeed, if one were to set  $\zeta = 0$  and attempt to retain equation (40), one would be left with an overconstrained system, with

more equations than unknowns. For the same reason the second equation in (34) is discarded. Unlike momentum, flux is physically not conserved due to resistivity, and the reduction leads to errors in the rate of change of flux.

### A. Non-conservation of flux

We follow the same general procedure to show non-conservation of flux due to finite resistivity as Freidberg<sup>29</sup> does to show conservation in the ideal case. Magnetic flux through an arbitrary surface  $S(t)$  is defined as

$$\psi = \int_{S(t)} \vec{B} \cdot d\vec{S}, \quad (51)$$

where the surface  $S(t)$  is advected with the plasma, hence its dependence on time. Taking the time derivative, and then applying the induction equation (1) and Stokes' theorem, we obtain

$$\frac{d\psi}{dt} = \int_{S(t)} \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S} + \oint_{\partial S(t)} \vec{B} \cdot (\vec{v} \times d\vec{l}) = - \oint_{\partial S(t)} \eta \vec{j} \cdot d\vec{l}, \quad (52)$$

where  $\partial S(t)$  is the loop enclosing  $S(t)$ . This is the non-conservation of flux due to resistivity, which is present in the full MHD model, and, as expected, is locally proportional to the resistivity. Clearly, the resistivity term in the induction equation (1) is responsible for this non-conservation. Interestingly enough, when we apply the reduction, only the resistivity term is left in the second equation in (34). Indeed, setting  $\Omega = \zeta = 0$  and multiplying by  $B_v$ , the second equation in (34) becomes:

$$0 = \nabla \cdot (\eta \nabla \chi \times \vec{j}),$$

which is satisfied when  $\eta = 0$ . Thus, in the ideal case, both equations (34) can be satisfied even after a reduction, which means that the induction equation (1) will be satisfied in reduced ideal MHD. As such, when  $\eta = 0$ , flux will be conserved and magnetic field lines will be frozen into the plasma even in the reduced MHD model. For nonzero resistivity, the second equation in (34) cannot be satisfied and must be dropped. This amounts to neglecting a component of the resistivity term in the induction equation and, depending on the relative orientations of  $\vec{j}^\perp$  and the loop  $\partial S(t)$ , underestimating or overestimating  $\partial \psi / \partial t$ . Since the term in the above equation is also locally proportional to the perpendicular components of the current, we need  $|j^\parallel| \gg |\vec{j}^\perp|$  in order for  $\Omega = \zeta = 0$  to be a valid approximation. In other words,  $|\vec{j}^\perp|/j$  can serve as an order of magnitude estimate of the relative reduction error in  $\partial \psi / \partial t$ . As we will see below, the perpendicular components of the current, which arise due to nonzero parallel derivatives of  $\Psi$  and components of the metric tensor of the vacuum field-aligned coordinate system, also contribute to momentum conservation error.

## B. Non-conservation of momentum

The action of the first projection operator (7) on a vector  $\vec{s}$  can be written as  $\nabla\chi \cdot \nabla \times [\nabla\chi \times (\vec{e}_\chi \times \vec{s})] = \nabla\chi \cdot \nabla \times (\vec{e}_\chi s^\chi - \vec{s})$ . Thus, equation (45) is the contravariant  $\chi$  component of vorticity-type equation, which we will refer to as the reduced vorticity equation. If all three components of this reduced vorticity equation were

$$\begin{aligned} & \frac{\partial \vec{\omega}}{\partial t} + v_{\parallel} \frac{\partial \vec{j}}{\partial t} + \nabla v_{\parallel} \times \frac{\partial \vec{B}}{\partial t} - \frac{\vec{j}}{B_v} \partial^{\parallel} v^2 - \nabla \left( \frac{\partial^{\parallel} v^2}{B_v} \right) \times \vec{B} + \nabla \times (\vec{\omega} \times \vec{v}) - \frac{\vec{j}}{B_v^2} \nabla\chi \cdot (\vec{\omega} \times \vec{v}) \\ & - \nabla \left[ \frac{\nabla\chi \cdot (\vec{\omega} \times \vec{v})}{B_v^2} \right] \times \vec{B} + \vec{\omega} \frac{P}{\rho} + \nabla \left( \frac{P}{\rho} \right) \times \vec{\omega} = \frac{\vec{B}}{\rho^2} (\vec{j} \cdot \nabla \rho) - \frac{\vec{j}}{\rho^2} (\vec{B} \cdot \nabla \rho) + \frac{1}{\rho} (\vec{B} \cdot \nabla) \vec{j} - \frac{1}{\rho} (\vec{j} \cdot \nabla) \vec{B} \\ & - \frac{\vec{j}}{\rho B_v^2} \nabla\chi \cdot (\vec{j} \times \vec{B}) - \nabla \left[ \frac{\nabla\chi \cdot (\vec{j} \times \vec{B})}{\rho B_v^2} \right] \times \vec{B} + \frac{1}{\rho^2} \nabla \rho \times \nabla p + \frac{\vec{j}}{\rho B_v} \partial^{\parallel} p + \nabla \left( \frac{\partial^{\parallel} p}{\rho B_v} \right) \times \vec{B}, \end{aligned} \quad (53)$$

where we have introduced the reduced velocity  $\vec{\omega} = -\nabla\chi \times (\vec{e}_\chi \times \vec{v}) = \nabla\Phi \times \nabla\chi / B_v^2$  and the reduced vorticity  $\vec{\omega} = \nabla \times \vec{\omega}$ . The viscosity term is not considered here since we have not done a proper derivation of it but simply added a generic term after the fact. If the components of this equation perpendicular to  $\nabla\chi$  are identically zero, then there is no approximation in the velocity reduction as nothing is being neglected, and momentum is still conserved. The most general case in which the perpendicular components are zero is the following:

$$\partial^{\parallel} u = 0, \quad u \in \{g^{ik}, \Phi, \Psi, v_{\parallel}, p, \rho, P\} \quad (54)$$

where  $g^{ik}$  are the components of the metric tensor of the vacuum field-aligned coordinate system. As can be shown by a simple calculation, in this case both  $\vec{\omega}$  and  $\vec{j}$  will be directed strictly along  $\nabla\chi$ . If we allow either the metric tensor,  $\Phi$  or  $\Psi$  to vary along  $\nabla\chi$ , the same calculation will show that  $\vec{\omega}$  has nonzero perpendicular components. This will cause  $\partial\vec{\omega}/\partial t$  to have nonzero perpendicular components, which cannot be canceled by any other terms since there are no more time derivatives involving  $\Phi$  in the equation. Similarly, if we let any of the other quantities vary along  $\nabla\chi$ , the last term and the seventh term on the RHS (pressure), the first term and the seventh term on the RHS (density), the last term on the LHS ( $P$ ) and third term on the LHS ( $v_{\parallel}$ ) will be nonzero and will not be canceled by any other terms. If the conditions (54) are met, then only the fourth and sixth terms on the LHS are nonzero. As can be shown by a simple expansion of the sixth term:

$$[\nabla \times (\vec{\omega} \times \vec{v})]^\perp - \left[ \nabla \left[ \frac{\nabla\chi \cdot (\vec{\omega} \times \vec{v})}{B_v^2} \right] \times \vec{B} \right]^\perp \equiv 0.$$

A major simplification comes from the fact that, due to  $\partial^{\parallel} g^{\psi\beta\nu} = 0$ , the Clebsch-type coordinate system aligned

satisfied (which, in general, is not possible as the system of equations would be overconstrained), then the original Navier-Stokes equation would also be satisfied and momentum would be conserved exactly. We can therefore estimate the magnitude of momentum conservation error by considering the components of the vorticity-type equation perpendicular to  $\nabla\chi$ . This vorticity-type equation can be written as

to the vacuum field can be made orthogonal<sup>19</sup>. In addition, the conditions (54) also require that the lengths of the basis vectors do not vary along  $\nabla\chi$ , which forces all of the co- and contravariant components of  $\vec{v}$  and  $\vec{\omega}$  to be constant along vacuum field lines.

In the general case, when the conditions (54) are no longer satisfied, the reduced velocity is no longer an exact solution to the Navier-Stokes equation, and momentum is not conserved exactly. Nevertheless, as long as we have  $|\partial^{\parallel} u| \ll |\nabla^\perp u|$ , the approximation  $\nabla u \approx \nabla^\perp u$  is valid and errors introduced by the reduction should be small. The smallness of the parallel derivative is, in most cases, a reasonable assumption, and is included in most orderings<sup>1,3,6-8</sup>. A notable example when this assumption is not valid is pellet injection<sup>31</sup>, when, at the tip of simulated pellet, the local gradients of density, pressure and  $P$  can be comparable in the parallel and perpendicular directions. However, during pellet ablation, the density and pressure perturbations will quickly equilibrate along the total magnetic field, and if the direction of the total field is not too different from the direction of the vacuum field, as implied by (16), the parallel derivatives will return to being small even in such a scenario.

The metric tensor is the only quantity in (54) that is determined solely by the vacuum field and is not affected by the dynamics of the system. The assertions (54) imply that the vacuum field has zero local shear everywhere<sup>19</sup>, and that its strength does not change along field lines. As mentioned previously in section II C, we have a degree of control over what we choose to be the "vacuum field", as long as the assumption (16) is valid, otherwise the MHD waves will not separate in the velocity representation (6). Of course,  $\chi$  must always satisfy the Laplace equation in order for  $\nabla\chi$  to be a valid magnetic field. In the case of a tokamak, choosing  $\chi = F_0\phi$ , where  $F_0$  is a constant and  $\phi$  is the toroidal angle, will satisfy  $\partial^{\parallel} g^{ik} = 0$  exactly. The part of the vacuum field created by the poloidal

coils will then be grouped with the induced field. If a two-dimensional, completely axisymmetric tokamak simulation is run, then the conditions (54) will be satisfied exactly. In the more general case of an arbitrary three-dimensional magnetic configuration, as in a stellarator, it may not be possible to choose  $\chi$  so that  $\partial^{\parallel} g^{ik} = 0$  is satisfied exactly while fulfilling the assumption (16) at the same time. The choice of  $\chi$  would then be a compromise between  $|\partial^{\parallel} g^{ik}| \ll |\nabla^{\perp} g^{ik}|$  and assumption (16), with the parallel derivative of the metric tensor contributing to the error, which should still be small if the perpendicular derivatives are sufficiently large.

## VII. CONCLUSION

In the present article, a hierarchy of models suitable for stellarator geometries with excellent conservation properties was derived. We introduced representations that consist of a background vacuum field, a field line bending term and a field compression term for the magnetic field, and an  $\vec{E} \times \vec{B}$  term, a field-aligned flow term and a fluid compression term for the velocity. We also showed that any arbitrary magnetic and velocity fields can be expressed in this form. Thus, when we insert the representations into the viscoresistive MHD equations and apply appropriate projection operators to Faraday's law and the Navier-Stokes equation, obtaining a system of scalar equations that is closed by the continuity and energy equations, the scalar equations are identical to the original full MHD equations in the inviscid case.

Importantly, we showed that, if the background vacuum field is stronger than the bending and compression terms, and if the  $\beta$  is sufficiently low, MHD waves are approximately separated in the velocity representation, with Alfvén waves contained in the  $\vec{E} \times \vec{B}$  term, slow magnetosonic waves in the field-aligned flow term and fast magnetosonic waves in the fluid compression term. Thus, by setting the fluid compression term to zero, we eliminated fast magnetosonic waves, obtaining a reduced MHD model. We also showed that the  $\vec{E} \times \vec{B}$  and field-aligned flow terms do not compress the magnetic field, which allows us to set the field compression term in the magnetic field representation to zero within the same reduced model. As an optional further reduction, we also considered a model where the field-aligned flow term is set to zero. This is similar to the approach followed by Breslau *et al*<sup>12</sup> and Izzo *et al*<sup>11</sup> for tokamaks.

Finally, by considering the terms that were neglected in the reduction, we showed that there is no approximation associated with the reduction if the background vacuum field is shear-free, all the unknown scalar fields do not vary in the direction of the background field, and the mass diffusivity  $D_{\perp}$  and density source  $S_{\rho}$  also do not vary in the direction of the background field. When this is not the case, the reduction leads to violations of the conservation of momentum and errors in the evolution of magnetic flux. Therefore, the reduction approximation

is valid as long as the shear in the background vacuum field is low and parallel derivatives are small compared to perpendicular derivatives, which is often the case since parallel dynamics are much faster than perpendicular dynamics.

Although we derived these equations with the intention of eventually implementing them in the JOREK code, we deliberately made no assumptions about the underlying geometry of the problem. It should therefore be possible to apply these equations not only to toroidal fusion machines, such as tokamaks and stellarators, but also to non-toroidal and open field line configurations, which are often encountered in astrophysics, provided a strong guide field.

It remains an open question of how well various instabilities are represented in the two reduced MHD models presented here. We are currently working on implementing the reduced and full models in JOREK, after which we will benchmark the simulation results of the reduced models against the full model. We also plan to study the (1,1) internal kink mode in a tokamak analytically within the context of our reduced models.

## ACKNOWLEDGMENTS

The authors thank Omar Maj, Florian Hindenlang, Per Helander, Boniface Nkonga, Guido Huijsmans, Erika Strumberger, Alessandro Biancalani and Rohan Ramasamy for fruitful discussions.

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training program 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

## Appendix A: Ordering

Reduced MHD equations similar, though not identical, to those considered in section V can be obtained from the full MHD equations for  $\rho$ ,  $p$ ,  $\Phi$ ,  $\Psi$  and  $v_{\parallel}$  in section IV using an ordering approach. We introduce a small parameter  $\lambda \equiv L_{\perp}/L_{\parallel}$ , where  $L_{\perp}$  and  $L_{\parallel}$  are the perpendicular and parallel length scales, respectively. Then clearly the spatial derivatives must be ordered as follows:

$$|\partial^{\parallel}| \sim \lambda |\nabla^{\perp}|. \quad (\text{A1})$$

Since we intend to eliminate fast magnetosonic waves, the fastest remaining waves will be Alfvén waves, which travel along field lines. The shortest time scale will then be the Alfvén time  $\tau_A \equiv L_{\parallel}/c_A$ , and the partial time derivative will be ordered as:

$$\left| \frac{\partial}{\partial t} \right| \sim \frac{1}{\tau_A} \equiv \frac{c_A}{L_{\parallel}}, \quad (\text{A2})$$

where  $c_A^2 \equiv B_s^2/(\mu_0\rho_s)$ , and  $B_s$  and  $\rho_s$  are typical reference values for the magnetic field and density in the system. In addition, we make the following assumptions on the relative orders of the other quantities. The terms in the magnetic field representation (5) are ordered as:

$$\frac{|\nabla\Psi \times \nabla\chi|}{|\nabla\chi|} \sim |\nabla^\perp\Psi| \sim \lambda, \quad (\text{A3})$$

and

$$\frac{|\nabla\Omega \times \nabla\psi_v|}{|\nabla\chi|} \sim \frac{F_v}{B_v} |\nabla\Omega| \sim \lambda^3. \quad (\text{A4})$$

For the terms in the velocity representation (6) we require:

$$|v_{\parallel}\vec{B}| \div \frac{|\nabla\Phi \times \nabla\chi|}{B_v^2} \sim B_v^2 \frac{|v_{\parallel}|}{|\nabla^\perp\Phi|} \sim 1, \quad (\text{A5})$$

and

$$|\nabla^\perp\zeta| \div \frac{|\nabla\Phi \times \nabla\chi|}{B_v^2} \sim B_v \frac{|\nabla^\perp\zeta|}{|\nabla^\perp\Phi|} \sim \lambda. \quad (\text{A6})$$

Finally, we assume that the partial and convective terms in the material derivative are of the same order:

$$\left| \frac{\partial}{\partial t} \right| \div |\vec{v} \cdot \nabla| \sim \left| \frac{\partial}{\partial t} \right| \div \frac{|\nabla^\perp\Phi||\nabla^\perp|}{B_v} \sim 1. \quad (\text{A7})$$

We also assume that

$$\begin{aligned} \beta &\sim \lambda^2, & \kappa_\perp &\sim \lambda\kappa_\parallel, & \text{Re} &\sim \text{Re}_m \sim 1, \\ P &\sim \partial\rho/\partial t, & S_e &\sim \partial\mathcal{E}/\partial t \end{aligned} \quad (\text{A8})$$

Here,  $\mathcal{E} = \rho v^2/2 + p/(\gamma-1) + B^2/(2\mu_0)$  is the total energy.

If we normalize all lengths by  $L_\perp$ , velocities by  $c_A$ , densities by  $\rho_s$ , thermal conductivities by  $\kappa_s$ ,  $\nabla\psi_v$  by  $F_s$ , and magnetic fields by  $B_s \equiv c_A\sqrt{\mu_0\rho_s}$ , deriving all other normalization factors from simple combinations of these, we obtain the following ordering for the normalized quantities:

- $O(1)$ :  $\tilde{\rho}$ ,  $\tilde{B}_v$ ,  $\tilde{F}_v$ ,  $\tilde{\kappa}_\parallel$ ,  $\tilde{\nabla}^\perp$
- $O(\lambda)$ :  $\tilde{\Psi}$ ,  $\tilde{\Phi}$ ,  $\tilde{v}_\parallel$ ,  $\tilde{\eta}$ ,  $\tilde{\nu}$ ,  $\tilde{D}_\perp$ ,  $\tilde{\kappa}_\perp$ ,  $\tilde{S}_\rho$ ,  $\tilde{\partial}_\parallel$ ,  $\partial/\partial\tilde{t}$
- $O(\lambda^2)$ :  $\tilde{\zeta}$ ,  $\tilde{p}$
- $O(\lambda^3)$ :  $\tilde{\Omega}$ ,  $\tilde{S}_e$

where, for each quantity  $u$ ,  $\tilde{u} = u/u_s$  and  $u_s$  is the normalization factor. The scales  $L_\perp$ ,  $L_\parallel$ ,  $\rho_s$ ,  $\kappa_s$ ,  $B_s$  and  $F_s$  were chosen so that  $L_\perp|\nabla^\perp| \sim L_\parallel|\partial^\parallel| \sim \tilde{\rho} \sim \tilde{\kappa}_\parallel \sim \tilde{B} \sim \tilde{F}_v \sim 1$ .

We can now apply our ordering to the equations in section IV either directly using relations (A1–A8) or by first normalizing the equations and then using the ordering for the normalized quantities above. In any case, we get the same result, and if we drop the tildes from

the normalized quantities, the resulting equations will be visually identical to the non-normalized equations.

Applying the ordering to equation (30), and keeping only the lowest order terms ( $O(\lambda)$  in this case), we obtain:

$$\frac{\partial\rho}{\partial t} = -B_v \left[ \frac{\rho}{B_v^2}, \Phi \right] + P, \quad (\text{A9})$$

where  $P = \nabla \cdot (D_\perp \nabla \rho) + S_\rho$ . There are two differences this equation has with equation (41). First, the  $v_\parallel$  terms are eliminated by the ordering, which means that advection is dominated by the  $\vec{E} \times \vec{B}$  flow. This is a consequence of constraining the field-aligned flow to be of the same order of magnitude as the  $\vec{E} \times \vec{B}$  flow, as, in the lowest order, field-aligned flow is in the  $\nabla\chi$  direction, which produces a parallel derivative. The second difference is that the diffusion in the  $P$  term is no longer anisotropic. This can be justified by considering that diffusive transport is negligible compared to the transport due to  $v_\parallel$ , so failing to subtract out diffusion along  $\vec{B}$  makes no difference.

When applying the ordering to equation (31), we note that while  $\mathcal{E}$  itself has order  $O(1)$  due to the  $B_v^2/(2\mu_0)$  term,  $\partial\mathcal{E}/\partial t$  will be  $O(\lambda^3)$  and  $\nabla^\perp(\mathcal{E}/B_v^2)$  will be  $O(\lambda^2)$ . In addition, in the expression  $v_\parallel\mathcal{E} - \vec{v} \cdot \vec{B}/\mu_0$ , the  $O(\lambda)$  term  $v_\parallel B_v^2/\mu_0$  is cancelled by a similar term in  $\vec{v} \cdot \vec{B}$ . Keeping just the lowest order terms ( $O(\lambda^2)$  this time), the following is obtained after some simplifications:

$$2(B_v, \zeta) + B_v \Delta^\perp \zeta = 0. \quad (\text{A10})$$

We will treat this as the governing equation for  $\zeta$ . We do not need to actually solve it since we are only interested in the first two terms of  $\vec{v}$  and  $\zeta$  does not appear in the governing equations of any quantities that we are interested in. It should be pointed out that the perpendicular components of  $\vec{j}$  are all  $O(\lambda^2)$  or higher. In particular, if we expand the  $(\nabla\Psi \cdot \nabla)\nabla\chi$  term and evaluate the Christoffel symbols, we see that its perpendicular components are also  $O(\lambda^2)$ . Taking the next lowest order ( $O(\lambda^3)$ ) terms in equation (31), we have

$$\begin{aligned} \frac{\partial\mathcal{E}}{\partial t} &= -B_v \left[ \frac{\mathcal{E} + p}{B_v^2}, \Phi \right] + \frac{B_v}{\mu_0} \partial^\parallel(\Phi, \Psi) + \frac{B_v}{\mu_0} [(\Phi, \Psi), \Psi] \\ &\quad - \frac{1}{\mu_0} \nabla \cdot [\eta \vec{j} \times (\nabla\chi + \nabla\Psi \times \nabla\chi)] + \nabla \cdot \left[ \frac{\kappa_\perp}{R} \nabla \left( \frac{p}{\rho} \right) \right. \\ &\quad \left. + \frac{p}{\gamma-1} \frac{D_\perp}{\rho} \nabla\rho + \frac{\kappa_\parallel \nabla\chi}{RB_v} \left( \partial^\parallel \left( \frac{p}{\rho} \right) + \left[ \frac{p}{\rho}, \Psi \right] \right) \right] \\ &\quad + S_e - \frac{v^2}{2} P, \end{aligned} \quad (\text{A11})$$

where the expressions for  $v^2$ ,  $B^2$  and  $\vec{j}$  match those given in (43). Comparing this to equation (42), we see that, again, the  $\vec{E} \times \vec{B}$  flow dominates advection and field-aligned advection is not present at this order. The subtracting out of the component of  $\kappa_\perp$  heat transport in the



$\vec{B}$  direction is neglected since  $\kappa_{\parallel} \gg \kappa_{\perp}$ , similarly to our treatment of mass diffusion. Finally, in the parallel heat flow term,  $\vec{B}$  and  $B^2$  have been approximated by  $\nabla\chi$  and  $B_v^2$ , respectively, neglecting the higher order differences.

We proceed to equation (33). The lowest order ( $O(\lambda^2)$ ) terms are

$$\nabla \frac{\partial \Psi}{\partial t} \times \nabla \chi = \nabla \times \left[ \frac{\nabla \chi}{B_v} (\partial^{\parallel} \Phi - [\Psi, \Phi]) + \nabla \zeta \times \nabla \chi - \eta \vec{j} \right]. \quad (\text{A12})$$

Projecting this equation on  $\nabla \chi$ , we get exactly the same equation as (A10). Projecting on  $\nabla \psi_v$ , we get

$$\left[ \psi_v, \frac{\partial \Psi}{\partial t} \right] = \left[ \frac{[\Psi, \Phi] - \partial^{\parallel} \Phi}{B_v}, \psi_v \right] + \frac{\nabla \cdot (\eta \nabla \psi_v \times \vec{j})}{B_v}, \quad (\text{A13})$$

which exactly matches equation (44). We do not need an equation for  $\Omega$  since  $\Omega$  does not appear in the governing equations for any of quantities that we are interested in.

Now consider equation (38). The lowest order terms are  $O(\lambda^2)$ :

$$\begin{aligned} \Delta^{\perp} \frac{\partial \Phi}{\partial t} = & \nabla \cdot \left[ \frac{\omega \chi}{B_v^2} (\nabla \Phi \times \nabla \chi + B_v^2 v_{\parallel} \nabla \Psi \times \nabla \chi) \right. \\ & \left. - v_{\parallel} B_v^2 \vec{\omega}^{\perp} + \frac{B_v^2}{\rho} \vec{j}^{\perp} - \frac{j^{\chi}}{\rho} \nabla \Psi \times \nabla \chi - \frac{P}{\rho} \nabla^{\perp} \Phi \right] \\ & + B_v \left[ \frac{1}{\rho}, p \right] + \nu \Delta \Delta^{\perp} \Phi. \end{aligned} \quad (\text{A14})$$

This result is equivalent to approximating  $\vec{B}$  by  $\nabla \chi$  in equation (35), which will cause the third term on the LHS to disappear, and then applying the projection operator, which is now  $\nabla \chi \cdot \nabla \times [B_v^{-2} \nabla \chi \times (\nabla \chi \times$  due to the approximation we made. Note that we have also approximated  $\vec{B}$  by  $\nabla \chi$  in the projections ( $s_{\chi}$  becomes  $s^{\chi}/B_v^2$  for any vector  $\vec{s}$ ) and  $B^2$  by  $B_v^2$  in the coefficient in front of  $\vec{j}^{\perp}$ .

The last equation that we consider is equation (39). The lowest order terms are again  $O(\lambda^2)$ :

$$\frac{\partial v_{\parallel}}{\partial t} = -\frac{v_{\parallel}}{\rho} P + \frac{\vec{\omega} \cdot \nabla^{\perp} \Phi}{B_v^2} + \nu \Delta v_{\parallel}. \quad (\text{A15})$$

Just as equation (A14), this equation corresponds to approximating  $\vec{B}$  by  $\nabla \chi$  in equation (35) and applying the projection operator  $\nabla \chi \cdot$ . In addition, both the hydrodynamic pressure and ram pressure are neglected.

Finally, we must simply drop equation (40) to avoid having an overconstrained system. If we attempt to keep it,  $\zeta$  and  $\Omega$  will drop from the equation in the lowest order, leaving us with an equation involving just  $\Psi$ ,  $\Phi$ ,  $v_{\parallel}$  and  $p$ , and overconstraining these variables. Our decision is in line with what is generally done in ordering approaches<sup>1,3,6-8</sup>, where one considers the parallel projection of the vorticity equation and the parallel projection of the Navier-Stokes equation, but not the divergence of the Navier-Stokes equation.

We have thus shown that a similar, though much simpler set of reduced MHD equations can be derived using an ordering. However, more assumptions are needed for ordering than for a consistent ansatz approach, and these extra assumptions are what allows us to obtain simpler equations.

- <sup>1</sup>H. Strauss, *Journal of Plasma Physics* **57**, 83 (1997).
- <sup>2</sup>S. Jardin, N. Ferraro, J. Breslau, and J. Chen, *Computational Science & Discovery* **5**, 014002 (2012).
- <sup>3</sup>S. E. Kruger, C. C. Hegna, and J. D. Callen, *Physics of Plasmas* **5**, 4169 (1998).
- <sup>4</sup>J. M. Greene and J. L. Johnson, *The Physics of Fluids* **4**, 875 (1961).
- <sup>5</sup>B. Kadomtsev and O. Pogutse, *Zh. Eksp. Teor. Fiz* **65**, 575 (1973).
- <sup>6</sup>H. Strauss, *The Physics of Fluids* **19**, 134 (1976).
- <sup>7</sup>H. Strauss, *The Physics of Fluids* **20**, 1354 (1977).
- <sup>8</sup>H. Strauss, *Plasma Physics* **22**, 733 (1980).
- <sup>9</sup>S. Oughton, W. H. Matthaeus, and P. Dmitruk, *The Astrophysical Journal* **839**, 2 (2017).
- <sup>10</sup>W. Park, D. Monticello, R. White, and S. Jardin, *Nuclear fusion* **20**, 1181 (1980).
- <sup>11</sup>R. Izzo, D. Monticello, J. DeLucia, W. Park, and C. Ryu, *The Physics of fluids* **28**, 903 (1985).
- <sup>12</sup>J. Breslau, N. Ferraro, and S. Jardin, *Physics of Plasmas* **16**, 092503 (2009).
- <sup>13</sup>E. Franck, M. Hölzl, A. Lessig, and E. Sonnendrücker, *ESAIM: Mathematical Modelling and Numerical Analysis* **49**, 1331 (2015).
- <sup>14</sup>G. Huysmans and O. Czarny, *Nuclear Fusion* **47**, 659 (2007).
- <sup>15</sup>G. Huysmans, S. Pamela, E. Van Der Plas, and P. Ramet, *Plasma Physics and Controlled Fusion* **51**, 124012 (2009).
- <sup>16</sup>S. Pamela, G. Huijsmans, T. Eich, S. Saarelma, I. Lupelli, C. Maggi, C. Giroud, I. Chapman, S. Smith, L. Frassinetti, M. Becoulet, M. Hoelzl, F. Orain, S. Futatani, and JET Contributors, *Nuclear Fusion* **57**, 076006 (2017).
- <sup>17</sup>W. D. D'haeseleer, W. N. Hitchon, J. D. Callen, and J. L. Shohet, *Flux coordinates and magnetic field structure: a guide to a fundamental tool of plasma theory* (Springer Science & Business Media, 2012).
- <sup>18</sup>D. P. Stern, *American Journal of Physics* **38**, 494 (1970).
- <sup>19</sup>A. Salat and J. Tataronis, *Journal of Geophysical Research: Space Physics* **105**, 13055 (2000).
- <sup>20</sup>W. Dommashch, *Computer Physics Communications* **40**, 203 (1986).
- <sup>21</sup>E. Thomas Jr, G. Sasser, S. Knowlton, J. Hanson, and R. Gandy, *Computer Physics Communications* **100**, 31 (1997).
- <sup>22</sup>R. Wolf, W. Biel, M. De Bock, K. Finken, S. Günter, G. Hogewij, S. Jachmich, M. Jakubowski, R. Jaspers, A. Krämer-Flecken, *et al.*, *Nuclear fusion* **45**, 1700 (2005).
- <sup>23</sup>S. Abdullaev, T. Van Rompuy, K. Finken, M. Lehnen, and M. Jakubowski, arXiv preprint arXiv:0805.1477 (2008).
- <sup>24</sup>P. Xanthopoulos and F. Jenko, *Physics of Plasmas* **13**, 092301 (2006).
- <sup>25</sup>J. A. Hudson, *The excitation and propagation of elastic waves* (CUP Archive, 1980).
- <sup>26</sup>D. Southwood and M. Saunders, *Planetary and space science* **33**, 127 (1985).
- <sup>27</sup>D. Y. Klimushkin, P. Mager, and K.-H. Glassmeier, *Annales Geophysicae* **22**, 267 (2004).
- <sup>28</sup>N. Ferraro, S. Jardin, M. Shephard, A. Bauer, J. Breslau, J. Chen, F. Delalondre, X. Luo, and F. Zhang, in *SciDAC Conference* (2011).
- <sup>29</sup>J. P. Freidberg, *Ideal MHD* (Cambridge University Press, 2014).
- <sup>30</sup>O. Czarny and G. Huysmans, *Journal of Computational Physics* **227**, 7423 (2008).
- <sup>31</sup>S. Futatani, G. Huijsmans, A. Loarte, L. Baylor, N. Commaux,

T. Jernigan, M. Fenstermacher, C. Lasnier, T. Osborne, and B. Pegourié, *Nuclear Fusion* **54**, 073008 (2014).