MHD Simulations for Fusion Applications*

Stephen C. Jardin†
Princeton Plasma Physics Laboratory
P.O. Box 451, Princeton, NJ 08543

July 19, 2010

Abstract

Nuclear fusion holds forth the promise of being a clean and safe solution to meet the world's energy demand in the foreseeable future without producing long-lived radioactive waste or weapons-grade material. The most mature configuration for magnetically confining a fusion plasma is the tokamak; a current carrying toroidal plasma characterized by strong externally produced magnetic fields. The temperatures, densities, and current will diffuse across the magnetic field lines at some rate, determining the confinement properties of the tokamak. The tokamak can also develop global instabilities if the current and/or pressure exceed certain instability thresholds. This set of lectures is aimed at describing analytical formulations and associated numerical methods for quantitatively describing both the slow (diffusive) motion associated with transport and the faster (wavelike) motion associated with instabilities. The former uses slow time scale ordering to remove the wavelike motion, and a time-dependent field-aligned coordinate transformation to isolate the cross-field transport from the faster transport along the magnetic field lines. The latter uses a combination of high-order finite elements, a particular representation of the magnetic and velocity vector fields, and an implicit time advance algorithm with desirable properties.

*This work was supported by USDOE Contract No. DE-AC02-76CH03073
†email: jardin@princeton.edu

1 Tokamak Fusion Basics and the MHD Equations

1.1 Tokamak Fusion Basics

A strong case can be made for the development of fusion energy [1]. Worldwide demand for energy continues to increase due to both population increases and economic development. Most of this population growth and new energy demand is in urban areas, which implies the need for large centralized power generation. By many estimates, worldwide oil and gas production is near or past its peak, which implies the need for an alternative source. The only possibilities are: coal, nuclear fission, or nuclear fusion.

There is increasing evidence that release of greenhouse gasses is causing global climate change. This comes both from historical data and from detailed climate projections. This makes nuclear energy (both fission and fusion) preferable to fossil (coal). Nuclear fusion has three advantages over nuclear fission that could become critical: (i) It has inherent safety, so there is no possibility of a meltdown accident; (ii) There are no weapons proliferation considerations; and (iii) The waste disposal problems are greatly alleviated.

Controlled fusion uses isotopes of Hydrogen. A Deuterium (D) nuclei that collides with a Tritium (T) nuclei can fuse to produce a Helium nuclei, and neutron, and 17.5 MeV of energy. The neutron can be captured by a Lithium nuclei to produce another Helium nuclei and a Tritium nuclei. Because the process breeds its own Tritium and both Deuterium and Lithium are naturally abundant, there is essentially an unlimited supply of fuel.

In order to create the conditions for D and T to fuse, you must create a mixture of DT nuclei (or ions) and their associated electrons, called a plasma, and heat it to high temperature and high pressure. You need about 5 atmosphere pressure at a temperature of 10 keV ($100,000,000^0$ degrees K). A plasma at that temperature can never come into contact with material walls so we must confine it with magnetic fields.

To a first approximation, both the electrons and DT ions exhibit spiral orbits around magnetic field lines as indicated in Fig. 1. They describe circular orbits perpendicular to the field and free-stream in the direction of the field lines. In a *tokamak* [2], the magnetic field is bent into a torus so that the free-streaming motion does not lead to loss of particles. The tokamak has large electromagnets to produce the confining magnetic fields and to induce electrical current into the plasma. This current both heats the plasma and provides an essential "twist" to the magnetic field in the plasma so that the electrons and ions do not drift out of the confinement region. The toroidal-field coils produce the strongest magnetic field. These fields are directed in the toroidal direction (the long way around the torus). The poloidal-field coils produce weaker fields in the same plane as the fields produced by the plasma currents. These are orthogonal to the toroidal field, and serve to shape the plasma cross section. The magnetic field from the central field coil changes in time such as to induce current into

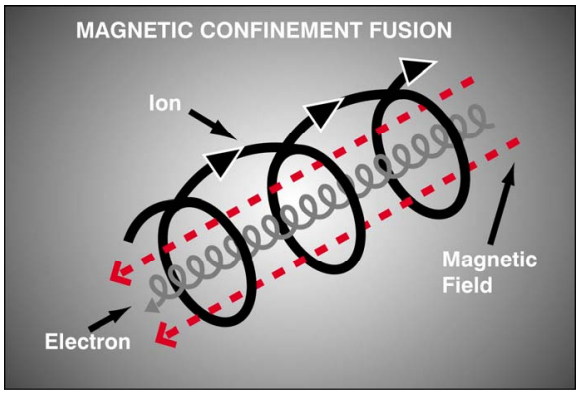

Figure 1: Charged particles have helical orbits in a magnetic field; they describe circular orbits perpendicular to the field and free-stream in the direction of the field.

the plasma through transformer action. The coil arrangement in a standard tokamak is shown in Fig. 2.

The doughnut-shaped tokamak plasma can develop global instabilities if the current it carries is too large or poorly distributed, or if the pressure is too high. This is the motivation for developing a set of partial differential equations that describe the dynamics of the plasma in the tokamak. The mathematical description of a plasma we consider here is called *extended magnetohydrodynamics*, or just simply *magnetohydrodynamics* (MHD). This treats the plasma as a conducting fluid that interacts with the magnetic fields produced by the electrical currents flowing in the plasma and from external sources. A tokamak is characterized by having a very strong externally imposed toroidal magnetic field (going the long way around the torus) so that the plasma pressure, p is much less than the magnetic pressure, $B^2/2\mu_0$ (SI Units). This is referred to as *low-*β, where $\beta \equiv 2\mu_0 p/B^2$. Special numerical methods are required to obtain accurate numerical solutions of low-β plasmas [3].

1.2 Magnetohydrodynamic (MHD) Equations

For simplicity, we consider here only a single species of ions with unit charge. The field equations in MHD are then [4]:

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \tag{1.1}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.2}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \tag{1.3}$$

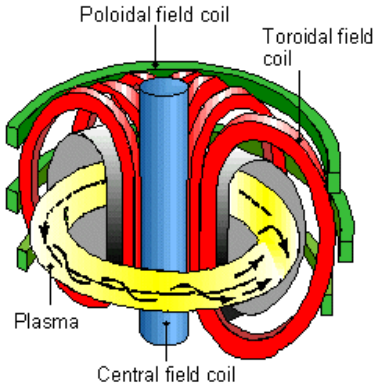

Figure 2: The tokamak creates toroidal magnetic field to confine the plasma and induces a toroidal plasma current to heat and confine the plasma.

$$n_i = n_e = n. \tag{1.4}$$

Equations (1.1) and (1.2) for the magnetic field $\mathbf{B}$ and the electric field $\mathbf{E}$ are exact and always valid. Note that Eq. (1.2) can be regarded as an initial condition for Eq. (1.1). Equation (1.3) can be taken as a defining equation for the electrical current density $\mathbf{J}$. Equation (1.4), stating the equivalence of the electron and ion number density, is referred to as *quasineutrality* .

The fluid equations are the continuity equation for the number density

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{u}) = S_m, \tag{1.5}$$

the internal energy equation for the plasma pressure

$$\begin{aligned} \frac{3}{2}\frac{\partial p}{\partial t} \quad &+ \quad \nabla \cdot \left(\mathbf{q} + \frac{3}{2}p\mathbf{u}\right) = -p\nabla \cdot \mathbf{u} + \left(\frac{3}{2}\nabla p_e - \frac{5}{2}\frac{p_e}{n}\nabla n\right) \cdot \frac{\mathbf{J}}{ne} \\ &+ \quad \mathbf{R}_e \cdot \frac{\mathbf{J}}{ne} - \boldsymbol{\pi} : \nabla\mathbf{u} + \boldsymbol{\pi}_e : \nabla\frac{\mathbf{J}}{ne} + Q, \end{aligned} \tag{1.6}$$

the electron internal energy equation for the electron pressure

$$\begin{aligned} \frac{3}{2}\frac{\partial p_e}{\partial t} \quad &+ \quad \nabla \cdot \left(\mathbf{q}_e + \frac{3}{2}p_e\mathbf{u}\right) = -p_e\nabla \cdot \mathbf{u} + \left(\frac{3}{2}\nabla p_e - \frac{5}{2}\frac{p_e}{n}\nabla n\right) \cdot \frac{\mathbf{J}}{ne} \\ &+ \quad \mathbf{R}_e \cdot \frac{\mathbf{J}}{ne} - \boldsymbol{\pi}_e : \nabla\left(\mathbf{u} - \frac{\mathbf{J}}{ne}\right) + Q_{\Delta ei} + Q_e, \end{aligned} \tag{1.7}$$

and the force balance equation for the fluid velocity

$$nm_i\left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla\mathbf{u}\right) + \nabla \cdot (\boldsymbol{\pi}_e + \boldsymbol{\pi}_i) = -\nabla p + \mathbf{J} \times \mathbf{B}. \tag{1.8}$$

In addition, from the momentum equation for the electrons, we have the *generalized Ohm's law* equation, which in the limit of vanishing electron mass ($m_e = 0$) becomes:

$$\mathbf{E} + \mathbf{u} \times \mathbf{B} = \frac{1}{ne} \left[\mathbf{R}_e + \mathbf{J} \times \mathbf{B} - \nabla p_e - \nabla \cdot \boldsymbol{\pi}_e \right]. \tag{1.9}$$

The variables $\mathbf{B}$, $\mathbf{u}$, n, p, p_e are the fundamental variables in that they obey time advancement equations. The variables $\mathbf{E}$, $\mathbf{J}$, and the ion pressure $p_i = p - p_e$ are auxiliary variables which we define only for convenience. We have also introduced the total heat flux $\mathbf{q} = \mathbf{q}_e + \mathbf{q}_i$ and the total non-isotropic stress tensor $\boldsymbol{\pi} = \boldsymbol{\pi}_e + \boldsymbol{\pi}_i$.

To proceed with the solution, one needs *closure* relations for the remaining terms; the collisional friction term $\mathbf{R}_e$, the random heat flux vectors $\mathbf{q}_j$, the anisotropic part of the stress tensor $\boldsymbol{\pi}_j$, and the equipartition term $Q_{\Delta ei}$. It is the objective of transport theory to obtain closure expressions for these in terms of the fundamental variables and their derivatives. There is extensive and evolving literature on deriving closures that are valid in different parameter regimes. We will discuss some of the more standard closures in the following sections. We note here that the external sources of particles, total energy, and electron energy S_m, Q, and Q_e must also be supplied, as must initial and boundary conditions.

1.2.1 Two-Fluid MHD

The set of MHD equations as written in Section 1.2 is not complete because of the closure issue. The most general closures that have been proposed are presently too difficult to solve numerically. However, there are some approximations that are nearly always valid and are thus commonly applied. Other approximations are valid over limited time scales and are useful for isolating specific phenomena.

The two-fluid magnetohydrodynamic equations are obtained by taking an asymptotic limit of the extended MHD equations in which first order terms in the ratio of the ion Larmor radius to the system size are retained. This is sometimes called the *finite Larmor radius* or FLR approximation. The principal effects are to include expressions for the parts of the ion stress tensor $\boldsymbol{\pi}_i$ and ion and electron heat fluxes $\mathbf{q}_i$ and $\mathbf{q}_e$ that do not depend on collisions. This contribution to the ion stress tensor is known as the *gyroviscous stress* or *gyroviscosity* . Let $\mathbf{b}$ be a unit vector in the direction of the magnetic field. The general form of gyroviscosity, assuming isotropic pressure and negligible heat flux, can be shown to be [4, 5, 6]:

$$\boldsymbol{\pi}_i^{gyr} = \frac{m_i p_i}{4eB} \left\{ \mathbf{b} \times \mathbf{W} \cdot (\mathbf{I} + 3\mathbf{bb}) + [\mathbf{b} \times \mathbf{W} \cdot (\mathbf{I} + 3\mathbf{bb})]^\dagger \right\}. \tag{1.10}$$

Here, the rate of strain tensor is

$$\mathbf{W} = \nabla \mathbf{u} + (\nabla \mathbf{u})^\dagger - \frac{2}{3} \mathbf{I} \nabla \cdot \mathbf{u}. \tag{1.11}$$

The expressions for the gyroviscosity in Eqs. (1.10) and (1.11) are quite complex and difficult to implement. A common approximation is to replace Eq. (1.10) with what has become known as the *gyroviscous cancellation* approximation [7, 8],

$$\nabla \cdot \boldsymbol{\pi}_i^{gyr} \approx -m_i n \mathbf{u}_* \cdot \nabla \mathbf{u}, \tag{1.12}$$

where the ion magnetization velocity is defined by

$$\mathbf{u}_* = -\frac{1}{ne} \nabla \times \left(\frac{p_i}{B^2} \mathbf{B} \right). \tag{1.13}$$

At the same order in this expansion, it can be shown that for each species there exists a heat flux $\mathbf{q}$ that is independent of the collision frequency:

$$\mathbf{q}_{\wedge j} = \frac{5}{2} \frac{p_j k_B}{q_j B} \mathbf{b} \times \nabla T_j. \tag{1.14}$$

The remaining contributions to the ion stress tensor are dependent on the collisionality regime and magnetic geometry in a complex way. These are often approximated by an isotropic part and a parallel part as follows:

$$\begin{aligned} \boldsymbol{\pi}_i^{iso} &= -\mu \left[\nabla \mathbf{u} + \nabla \mathbf{u}^\dagger \right] - 2(\mu_c - \mu) (\nabla \cdot \mathbf{u}) \mathbf{I}, & (1.15) \\ \nabla \cdot \boldsymbol{\pi}_i^{iso} &= -\mu \nabla^2 \mathbf{u} - (2\mu_c - \mu) \nabla (\nabla \cdot \mathbf{u}), & (1.16) \\ \boldsymbol{\pi}_i^{\parallel} &= \mu_\parallel (\mathbf{b} \cdot \mathbf{W} \cdot \mathbf{b})(\mathbf{I} - 3\mathbf{bb}). & (1.17) \end{aligned}$$

We note the positivity constraints $\mu \geq 0$, $\mu_c \geq \frac{2}{3}\mu$, $\mu_\parallel \geq 0$.

The electron gyroviscosity can normally be neglected due to the small electron mass. However, for many applications, an electron parallel viscosity can be important. This is similar in form to the ion one in Eq. (1.17) but with a coefficient $\mu_\parallel^e$. In addition, it has been shown by several authors that in certain problems involving two-fluid magnetic reconnection, an electron viscosity term known as *hyper-resistivity* is required to avoid singularities from developing in the solution [9, 10]. It is also useful for modeling the effect of fundamentally three-dimensional reconnection physics in a two-dimensional simulation [11, 12]. Introducing the coefficient λ_H, we can take the hyper-resistivity to be of the form:

$$\boldsymbol{\pi}_e^{hr} = \lambda_H \eta_\parallel n e \nabla \mathbf{J}. \tag{1.18}$$

If h is the smallest dimension that can be resolved on a numerical grid, then it has been proposed that $\lambda_H \sim h^2$ is required for the current singularity to be resolvable, while at the same time vanishing in the continuum limit $h \to 0$.

The plasma friction force is generally taken to be of the form:

$$\mathbf{R}_e = ne \left(\eta_\parallel \mathbf{J}_\parallel + \eta_\perp \mathbf{J}_\perp \right), \tag{1.19}$$

where $\|$ and $\perp$ are relative to the magnetic field direction. The classical values for these resistivity coefficients are [4]:

$$\eta_\perp = 1.03 \times 10^{-4} Z \ln\lambda [T(ev)]^{-3/2} \ \ \Omega\ m, \tag{1.20}$$

$$\eta_\| = 0.51\ \eta_\perp. \tag{1.21}$$

Here Z is the effective charge of the ion species and $\ln\lambda \sim 20$ is the Coulomb logarithm [13]. The electron-ion temperature equilibration term is related to the perpendicular resistivity by:

$$Q_{\Delta ei} = \frac{3e^2 n}{m_i} \eta_\perp n k_B (T_i - T_e). \tag{1.22}$$

The forms used for the viscosity coefficients μ, μ_c, $\mu_\|$ and the heat conduction coefficients $\kappa^e_\|$, $\kappa^e_\perp$, $\kappa^i_\|$, $\kappa^i_\perp$ can be the classical value for the collisional regime [4], but are normally chosen according to other *anomalous* (or empirical) models or so as to perform parametric studies. However, from basic physical considerations (that the particles essentially free-stream with very long collisional mean-free-paths parallel to the magnetic field but not perpendicular to it) we normally have:

$$\kappa^{e,i}_\| \gg \kappa^{e,i}_\perp. \tag{1.23}$$

The standard two-fluid MHD model can now be summarized as follows. Using the equations of Section 1.2, we use the following closure model:

$$\mathbf{R}_e = ne\left(\eta_\| \mathbf{J}_\| + \eta_\perp \mathbf{J}_\perp\right), \tag{1.24}$$

$$\mathbf{q}_e = -\kappa^e_\| \nabla_\| T_e - \kappa^e_\perp \nabla_\perp T_e + \mathbf{q}_{\wedge e}, \tag{1.25}$$

$$\mathbf{q}_i = -\kappa^i_\| \nabla_\| T_i - \kappa^i_\perp \nabla_\perp T_i + \mathbf{q}_{\wedge i}, \tag{1.26}$$

$$\boldsymbol{\pi}_i = \boldsymbol{\pi}^{gyr}_i + \boldsymbol{\pi}^{iso}_i + \boldsymbol{\pi}^\|_i, \tag{1.27}$$

$$\boldsymbol{\pi}_e = \boldsymbol{\pi}^\|_e + \boldsymbol{\pi}^{hr}_e, \tag{1.28}$$

$$Q_{\Delta ei} = \frac{3e^2 n}{m_i} \eta_\perp n k_B (T_i - T_e). \tag{1.29}$$

To complete this model, one must specify the 11 scalar functions $\eta_\|$, $\eta_\perp$, $\kappa^e_\|$, $\kappa^e_\perp$, $\kappa^i_\|$, $\kappa^i_\perp$, μ, μ_c, $\mu_\|$, $\mu^e_\|$, and λ_H. These can be either constants or functions of the macroscopic quantities being evolved in time according to the transport model being utilized.

1.2.2 Resistive MHD

The resistive MHD model treats the electrons and ions as a single fluid with pressure $p = p_e + p_i$. This model can formally be derived by taking a limiting case of the two-fluid equations. The first limit is that of *collision dominance*, which allows one to neglect the parallel viscosities compared to ∇p. The second limit is that of *zero Larmor radius*. This implies the neglect of the gyroviscous

stress and the collision-independent perpendicular heat fluxes $\mathbf{q}_{\wedge j}$ as well as most terms involving $\mathbf{J}/ne$ compared to those involving $\mathbf{u}$. It follows that the electron hyperviscosity is also not required. With these simplifications, neglecting external sources, and introducing the mass density $\rho \equiv nm_i$, the equations of Section 1.2 become:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{1.30}$$

$$\rho(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u}) = -\nabla p + \mathbf{J} \times \mathbf{B} - \nabla \cdot \boldsymbol{\pi}_i^{iso}, \tag{1.31}$$

$$\frac{3}{2}\frac{\partial p}{\partial t} + \nabla \cdot (\mathbf{q} + \frac{3}{2} p\mathbf{u}) = -p\nabla \cdot \mathbf{u} - \boldsymbol{\pi}_i^{iso} : \nabla \mathbf{u} + \eta J^2, \tag{1.32}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B} - \eta \mathbf{J}), \tag{1.33}$$

$$\mathbf{J} = \frac{1}{\mu_0} \nabla \times \mathbf{B}, \tag{1.34}$$

$$\mathbf{q}_i = -\kappa_\parallel \nabla_\parallel T - \kappa_\perp \nabla_\perp T. \tag{1.35}$$

Here the fluid temperature is defined as $T = p/2nk_B$ and the viscosity term in Eq. (1.31) is evaluated using Eq. (1.16). This model requires only the five transport coefficients: η, $\kappa_\parallel^e$, $\kappa_\perp$, μ, and μ_c.

1.2.3 Ideal MHD

A further approximation has to do with the smallness of the plasma resistivity and the time scales over which resistive effects are important, $\tau_R = \mu_0 a^2/\eta$, where a is the minor radius or other typical global dimension. If we non-dimensionalize the MHD equations, using the Alfvén velocity $V_A = B/\sqrt{\mu_0 n M_i}$, a, and the Alfvén time $\tau_A = a/V_A$, we find that the plasma resistivity becomes multiplied by the inverse *magnetic Lundquist number* S^{-1}, where

$$S \equiv \frac{\tau_R}{\tau_A}. \tag{1.36}$$

In modern fusion experiments, this number is typically in the range $S \sim 10^6 - 10^{12}$.

The other dissipative quantities $\mathbf{q}_j$ and $\boldsymbol{\pi}^{iso}$ are related to the resistivity and thus also become multiplied by S^{-1}. Although these terms are very important for the longer time dynamics of the plasma or to describe *resistive instabilities* that involve internal boundary layers [14], if we are only interested in the fastest

time scales present in the equations, we can neglect all these dissipative terms which scale as S^{-1}. Doing so leaves the *ideal MHD equations* [17].

These equations have been extensively studied by both physicists and mathematicians. They have a seemingly simple symmetrical structure with well-defined mathematical properties, but at the same time can be exceedingly rich in the solutions they admit. Adding back the dissipative and dispersive terms will enlarge the class of possible solutions by making the equations higher order, but will not fundamentally change the subset of non-dissipative solutions found here for macroscopic motions. It is therefore important to understand the types of solutions possible for these equations before studying more complex equation sets.

The ideal MHD equations can be written (in SI units), as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0, \tag{1.37}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}), \tag{1.38}$$

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) + \nabla p = \mathbf{J} \times \mathbf{B}, \tag{1.39}$$

$$\frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p + \gamma p \nabla \cdot \mathbf{u} = 0, \tag{1.40}$$

where the current density is given by $\mathbf{J} \equiv \mu_o^{-1} \nabla \times \mathbf{B}$. Here we have introduced $\gamma = \frac{5}{3}$, which is the ratio of specific heats, sometimes called the adiabatic index. It is sometimes useful to also define another variable

$$s \equiv p/\rho^{\gamma},$$

the entropy per unit mass. It then follows from Eqs. (1.37) and (1.40) that s obeys the equation

$$\frac{\partial s}{\partial t} + \mathbf{u} \cdot \nabla s = 0. \tag{1.41}$$

Eq. (1.41) can be used to replace Eq. (1.40).

1.3 Characteristics

The ideal MHD equations are a quasilinear symmetric hyperbolic system of equations and thus have real characteristics. Because they are real, the characteristic directions or characteristic manifolds have important physical meaning since all information is propagated along them. Here we determine these characteristics for the ideal MHD equations [18, 19].

We start with the ideal MHD equations, in the following form:

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t}+\mathbf{v}\cdot\nabla\mathbf{v}\right)=-\nabla p-\frac{1}{\mu_o}\nabla\mathbf{B}\cdot\mathbf{B}+\frac{1}{\mu_0}(\mathbf{B}\cdot\nabla)\mathbf{B}, \tag{1.42}$$

$$\frac{\partial \mathbf{B}}{\partial t}+\mathbf{v}\cdot\nabla\mathbf{B}=\mathbf{B}\cdot\nabla\mathbf{v}-\mathbf{B}\nabla\cdot\mathbf{v}, \tag{1.43}$$

$$\frac{\partial p}{\partial t}+\mathbf{v}\cdot\nabla p=-\gamma p\nabla\cdot\mathbf{v}, \tag{1.44}$$

$$\frac{\partial s}{\partial t}+\mathbf{v}\cdot\nabla s=0. \tag{1.45}$$

We use the fact that if the boundary data are given only along the characteristic curves, then the solution cannot be determined away from those curves. Let us assume that the boundary data for variables $\mathbf{v}$, $\mathbf{B}$, p, s are given on the three-dimensional surface $\phi(\mathbf{r},t)=\phi_0$, and ask under what conditions is this insufficient to determine the solution away from this surface.

We perform a coordinate transformation to align the boundary data surface with one of the coordinates. We consider ϕ as a coordinate and introduce additional coordinates χ, σ, τ within the three-dimensional boundary data manifold. Thus we transform

$$(\mathbf{r},t)\rightarrow(\phi,\chi,\sigma,\tau). \tag{1.46}$$

On the boundary data manifold $\phi=\phi_0$, we specify $\mathbf{v}(\mathbf{r},t)=\mathbf{v}_0(\chi,\sigma,\tau)$, $\mathbf{B}(\mathbf{r},t)=\mathbf{B}_0(\chi,\sigma,\tau)$, $p(\mathbf{r},t)=p_0(\chi,\sigma,\tau)$, and $s(\mathbf{r},t)=s_0(\chi,\sigma,\tau)$. Since $\mathbf{v}_0$, $\mathbf{B}_0$, p_0, and s_0 are known functions, the derivatives with respect to χ, σ, and τ are also known. We ask under what conditions can the solutions $\mathbf{v}(\phi,\chi,\sigma,\tau)$, $\mathbf{B}(\phi,\chi,\sigma,\tau)$, $p(\phi,\chi,\sigma,\tau)$, and $s(\phi,\chi,\sigma)$ be obtained away from the boundary $\phi=\phi_0$?

We look for a power series solution of the form

$$\begin{aligned}\mathbf{v}\,(\phi,\chi,\sigma,\tau) &= \mathbf{v}_0(\chi,\sigma,\tau)+(\phi-\phi_0)\left.\frac{\partial\mathbf{v}}{\partial\phi}\right|_{\phi_0}\\ &+ (\chi-\chi_0)\left.\frac{\partial\mathbf{v}}{\partial\chi}\right|_{\phi_0}+(\sigma-\sigma_0)\left.\frac{\partial\mathbf{v}}{\partial\sigma}\right|_{\phi_0}+(\tau-\tau_0)\left.\frac{\partial\mathbf{v}}{\partial\tau}\right|_{\phi_0}+\cdots,\end{aligned}$$

and similarly for $\mathbf{B}$, p, and s. The problem is solvable if the normal derivatives $\partial\mathbf{v}/\partial\phi|_{\phi_0}$, $\partial\mathbf{B}/\partial\phi|_{\phi_0}$, $\partial p/\partial\phi|_{\phi_0}$, and $\partial s/\partial\phi|_{\phi_0}$ can be constructed since all surface derivatives are known and higher-order derivatives can be constructed by differentiating the original PDEs.

Using the chain rule and using subscripts to denote partial derivatives with respect to the time t, $\phi_t = \partial\phi/\partial t$, etc., we can calculate

$$\begin{aligned}
\frac{\partial \mathbf{v}}{\partial t} &= \frac{\partial \mathbf{v}}{\partial \phi}\phi_t + \frac{\partial \mathbf{v}}{\partial \chi}\chi_t \cdots , \\
\mathbf{v}\cdot\nabla\mathbf{v} &= \mathbf{v}\cdot\nabla\phi\frac{\partial \mathbf{v}}{\partial \phi} + \mathbf{v}\cdot\nabla\chi\frac{\partial \mathbf{v}}{\partial \chi} + \cdots , \\
\nabla\cdot\mathbf{v} &= \nabla\phi\cdot\frac{\partial \mathbf{v}}{\partial \phi} + \nabla\chi\cdot\frac{\partial \mathbf{v}}{\partial \chi} + \cdots ,
\end{aligned}$$

etc. It is convenient to define some new notation. Define the *spatial normal*,

$$\hat{\mathbf{n}} = \nabla\phi/\mid\nabla\phi\mid,$$

the *characteristic speed*,

$$u \equiv -\left(\phi_t + \mathbf{v}\cdot\nabla\phi\right)/|\nabla\phi|,$$

which is the normal velocity of the characteristic measured with respect to the fluid moving with velocity $\mathbf{v}$, and let a prime denote the normal derivative,

$$(\ )' \equiv \frac{\partial}{\partial\phi}(\ ).$$

Using this notation, the ideal MHD equations take the form

$$-\rho u\mathbf{v}' + \hat{\mathbf{n}}p' + \frac{1}{\mu_0}\hat{\mathbf{n}}\mathbf{B}\cdot\mathbf{B}' - \frac{1}{\mu_0}\hat{\mathbf{n}}\cdot\mathbf{B}\mathbf{B}' = \cdots , \tag{1.47}$$

$$-u\mathbf{B}' - \hat{\mathbf{n}}\cdot\mathbf{B}\mathbf{v}' + \mathbf{B}\hat{\mathbf{n}}\cdot\mathbf{v}' = \cdots , \tag{1.48}$$

$$-up' + \gamma p\hat{\mathbf{n}}\cdot\mathbf{v}' = \cdots , \tag{1.49}$$

$$-us' = \cdots , \tag{1.50}$$

where the right side contains only known derivative terms. For definiteness, now choose $\mathbf{B}$ along the z axis and $\hat{\mathbf{n}}$ in the ($\hat{\mathbf{x}}$, $\hat{\mathbf{z}}$) plane so that in Cartesian coordinates,

$$\begin{aligned}
\mathbf{B} &= (0, 0, B)\,, \\
\hat{\mathbf{n}} &= (n_x, 0, n_z)\,.
\end{aligned}$$

We define the quantities $V_A \equiv B/\sqrt{\mu_0\rho}$ and $c_S \equiv \sqrt{\gamma p/\rho}$, and premultiply Eq. (1.48) by $\sqrt{\rho/\mu_0}$ and Eq. (1.49) by c_S^{-1}. This allows the system of equations Eqs. (1.47)–(1.50) to be written as

$$\mathbf{A}\cdot\mathbf{X} = \cdots , \tag{1.51}$$

where

$$\mathbf{A} = \begin{bmatrix} -u & 0 & 0 & -n_z V_A & 0 & n_x V_A & n_x c_S & 0 \\ 0 & -u & 0 & 0 & -n_z V_A & 0 & 0 & 0 \\ 0 & 0 & -u & 0 & 0 & 0 & n_z c_S & 0 \\ -n_z V_A & 0 & 0 & -u & 0 & 0 & 0 & 0 \\ 0 & -n_z V_A & 0 & 0 & -u & 0 & 0 & 0 \\ n_x V_A & 0 & 0 & 0 & 0 & -u & 0 & 0 \\ n_x c_S & 0 & n_z c_S & 0 & 0 & 0 & -u & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -u \end{bmatrix}$$

and

$$\mathbf{X} = \begin{bmatrix} \rho v_x' \\ \rho v_y' \\ \rho v_z' \\ \sqrt{\rho/\mu_0} B_x' \\ \sqrt{\rho/\mu_0} B_y' \\ \sqrt{\rho/\mu_0} B_z' \\ \frac{1}{c_S} p' \\ s' \end{bmatrix}.$$

Note that the matrix in Eq. (1.51) is symmetric, which guarantees that the eigenvalues will be real and that the system is hyperbolic. The characteristics are obtained when the determinant vanishes so that solutions cannot be propagated away from the boundary data manifold $\phi = \phi_0$. The determinant is given by

$$D = u^2 \left(u^2 - V_{An}^2\right) \left[u^4 - \left(V_A^2 + c_S^2\right) u^2 + V_{An}^2 c_S^2\right] = 0, \tag{1.52}$$

where we have used the relation $n_x^2 + n_z^2 = 1$ and let $V_{An}^2 = n_z^2 V_A^2$. The eight roots are given by:

$$\begin{aligned} u &= u_0 = \pm 0 \quad \text{entropy disturbances,} \\ u &= u_A = \pm V_{An} \quad \text{Alfvén waves,} \\ u &= u_s = \pm \left\{ \frac{1}{2}\left(V_A^2 + c_S^2\right) - \frac{1}{2}\left[\left(V_A^2 + c_S^2\right)^2 - 4V_{An}^2 c_S^2\right]^{1/2} \right\}^{1/2} \quad \text{slow wave,} \\ u &= u_f = \pm \left\{ \frac{1}{2}\left(V_A^2 + c_S^2\right) + \frac{1}{2}\left[\left(V_A^2 + c_S^2\right)^2 - 4V_{An}^2 c_S^2\right]^{1/2} \right\}^{1/2} \quad \text{fast wave.} \end{aligned}$$

The latter two roots are also known as the slow magnetoacoustic and fast magnetoacoustic waves.

In normal magnetically confined fusion plasmas, we can take the *low-β limit*, $c_S^2 \ll V_A^2$, which implies

$$\begin{aligned} u_s^2 &\cong n_z^2 c_S^2, \\ u_f^2 &\cong V_A^2 + n_x^2 c_S^2. \end{aligned}$$

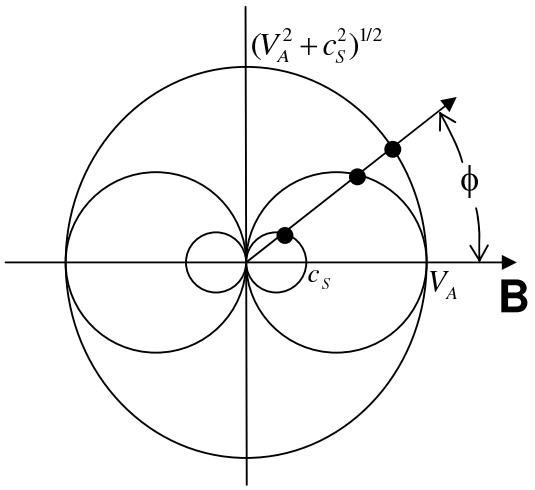

Figure 3: Reciprocal normal surface diagram in low-β limit.

The solutions following Eq. (1.52) are represented pictorially by the *reciprocal normal surface diagram* of Figure 3. The intersection points u_j give the speed of a plane wavefront whose normal $\hat{\mathbf{n}}$ is at an angle ϕ with the magnetic field $\mathbf{B}$.

We list here some properties of the characteristic speeds of Eq. (1.52) and Figure 3:

1. $|u_0| \leq |u_s| \leq |u_A| \leq |u_f| < \infty$.
2. For propagation along $\mathbf{B}$, $V_{An} = V_A$, and
$$\begin{aligned} |u_s| &= \min(V_A, c_S), \\ |u_A| &= V_A, \\ |u_f| &= \max(V_A, c_S). \end{aligned}$$
3. For propagation perpendicular to $\mathbf{B}, V_{An} = 0$, and
$$\begin{aligned} |u_s| &= |u_A| = 0, \\ |u_f| &= \left(V_A^2 + c_S^2\right)^{1/2}. \end{aligned}$$
4. If $V_A = 0$ (no magnetic field) the gas dynamics equations are obtained:
$$|u_s|, |u_A| \to 0,$$
$$|u_f| \to c_S \quad \text{(ordinary sound wave } c_S\text{)}.$$
5. Incompressible plasma limit is $\gamma \to \infty$, *or* $c_S \to \infty$:
$$|u_s| \to |u_A| \cdots \quad \text{these coincide,}$$
$$|u_f| \to \infty \cdots \quad \text{instantaneous propagation.}$$

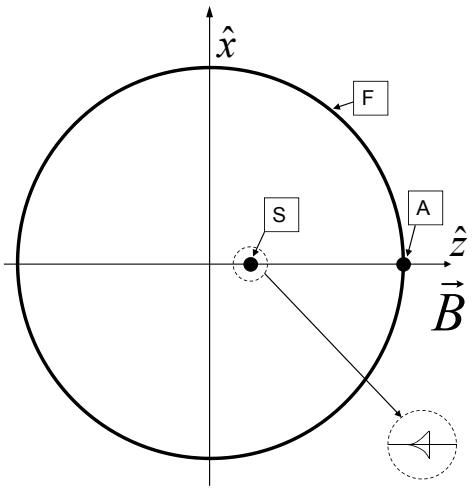

Figure 4: Ray surface diagram in low-β limit.

It is also instructive to compute a *ray surface diagram*. Suppose the initial disturbance is on a circle of radius R_0. After a time t, the equation for the wave surface is

$$z\cos\phi + x\sin\phi = R_0 \pm u_j(\phi)t. \tag{1.53}$$

We can take the ϕ derivative of Eq. (1.53) to obtain

$$-z\sin\phi + x\cos\phi = \pm\frac{d}{d\phi}u_j(\phi)t. \tag{1.54}$$

Now, invert Eqs. (1.53) and (1.54) for (x, z) to obtain

$$\begin{aligned} x &= R_0\sin\phi + \left[\sin\phi u_j + \cos\phi\frac{d}{d\phi}u_j(\phi)\right]t, \\ z &= R_0\cos\phi + \left[\cos\phi u_j - \sin\phi\frac{d}{d\phi}u_j(\phi)\right]t. \end{aligned}$$

Next, we let $R_0 \to 0$ represent a point disturbance and plot $[x(\phi), z(\phi)]$ for $0 < \phi < 2\pi$ to obtain the ray surface diagram in Figure 4. This shows clearly the extreme anisotropy of the Alfvén wave and also the slow waves in the low-β limit $c_S^2 \ll V_A^2$. Point disturbances just travel along the magnetic field.

Finally, let us consider the eigenvectors of the matrix in Eq. (1.51). The eigenvectors corresponding to each of the 4-pair of eigenvalues have physical significance. On a characteristic manifold, relations exist between $\mathbf{v}'$, $\mathbf{B}'$, p', s'. Discontinuities satisfying these relations are propagated along with the characteristics. Substitution of the roots of Eq. (1.52) into the matrix in Eq. (1.51)

yields the following eigenvectors:

$$\begin{matrix} & & \underline{\text{Entropy}} & \underline{\text{Alfvén}} & \underline{\text{Magnetoacoustic}} \end{matrix}$$

$$\begin{bmatrix} \rho v_x' \\ \rho v_y' \\ \rho v_z' \\ \sqrt{\rho/\mu_0} B_x' \\ \sqrt{\rho/\mu_0} B_y' \\ \sqrt{\rho/\mu_0} B_z' \\ \frac{1}{c_S} p' \\ s' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \pm 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \pm 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} n_x u^2/(u^2 - V_A^2) \\ 0 \\ n_z \\ -n_x u V_A n_z/(u^2 - V_A^2) \\ 0 \\ n_x^2 u V_A/(u^2 - V_A^2) \\ u/c_S \\ 0 \end{bmatrix}. \tag{1.55}$$

Note that the Alfvén wave is purely transverse, only perturbing $\mathbf{v}$ and $\mathbf{B}$ perpendicular both to the propagation direction and to the equilibrium magnetic field, while the magnetoacoustic waves involve perturbations in the other two directions. Note also that the different eigenvectors are orthogonal.

It is instructive to again take the low-β limit $c_S^2 \ll V_A^2$, and also to examine propagation parallel to $(n_z = 1,\ n_x = 0)$, and perpendicular to $(n_z = 0,\ n_x = 1)$ the equilibrium magnetic field. We find

$$\begin{matrix} & & \underline{\begin{matrix} Fast \\ n_z = 0 \\ n_x = 1 \end{matrix}} & \underline{\begin{matrix} Fast \\ n_z = 1 \\ n_x = 0 \end{matrix}} & \underline{\begin{matrix} Slow \\ n_z = 1 \\ n_x = 0 \end{matrix}} \end{matrix}$$

$$\begin{bmatrix} \rho v_x' \\ \rho v_y' \\ \rho v_z' \\ \sqrt{\rho/\mu_0} B_x' \\ \sqrt{\rho/\mu_0} B_y' \\ \sqrt{\rho/\mu_0} B_z' \\ \frac{1}{c} p' \\ s' \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ c_S/V_A \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ \pm 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ \pm 1 \\ 0 \end{bmatrix}. \tag{1.56}$$

It is seen that in a low-β ($\mu_0 p \ll B^2$) magnetized plasma, there exists a dramatic difference in the nature of the three non-trivial families of characteristics. All three waves can propagate parallel to the magnetic field, with the fast wave and the Alfvén wave having approximately the same velocity, but different polarities with respect to the perturbed velocities and fields, and the slow wave being much slower. The slow wave is the only one that involves a velocity component parallel to the background magnetic field.

The fast wave alone can propagate perpendicular to the background magnetic field and does so by compressing and expanding the field. Note that since the energy in the magnetic field is $B^2/2\mu_0$, perturbing the background field by an amount δB_z will require an energy $B_z \delta B_z/\mu_0 \gg \delta B_z^2/2\mu_0$. Thus, from

energetic considerations we expect these perpendicularly propagating disturbances to be of very small amplitude compared to other disturbances that do not compress the background field. However, because the length scales perpendicular to the field are normally much smaller than those parallel to the field, the time scales associated with the fast wave can be much shorter than those associated with either the Alfvén or slow waves. This is the fundamental reason for the stiffness of the ideal MHD equations when applied to a low-β magnetic confinement device.

1.3.1 Wave Dispersion Relation in Two-Fluid MHD

The additional terms present in the two-fluid MHD model of Section 1.2.1 introduce many new effects that are not present in the ideal MHD equations. The resistivities and viscosities generally damp the wave motion found in the last section, but resistivity can also allow plasma instabilities by relaxing the flux constrains of ideal MHD [14].

Here we discuss the effect of the non-dissipative terms in Ohm's law, Eq. (1.9), that are present in the two-fluid model but not in the ideal MHD description,

$$\mathbf{E} + \mathbf{u} \times \mathbf{B} = \frac{1}{ne} \left[\mathbf{J} \times \mathbf{B} - \nabla p_e + \cdots \right]. \tag{1.57}$$

When we non-dimensionalize the equations, these new terms on the right bring in a new dimensionless parameter,

$$d_i \equiv \frac{1}{\Omega_{ci} \tau_A}, \tag{1.58}$$

where $\Omega_{ci} = eB_0/M_i$ is the ion cyclotron frequency. The parameter d_i is called the *ion skin depth*. With these terms included, commonly called the *Hall terms*, the equations are no longer purely hyperbolic and become dispersive, i.e., different wavelength disturbances will propagate at different velocities.

If we linearize the equations about an equilibrium state with no flow and assume a periodic time and space dependence $\sim \exp i(\omega t - \mathbf{k} \cdot \mathbf{x})$, and now take the velocity u to be the phase velocity, $u \equiv \omega/k$, the analogue of the dispersion relation, Eq. (1.52) (after removing the entropy roots), becomes [20]

$$\begin{aligned} D \quad = \quad & \left(u^2 - V_{An}^2\right) \left[u^4 - \left(V_A^2 + c_S^2\right) u^2 + V_{An}^2 c_S^2\right] \\ & - V_{An}^2 u^2 d_i^2 k^2 \left(u^2 - c_S^2\right) = 0. \end{aligned} \tag{1.59}$$

Equation (1.59) is a cubic equation in u^2. In Figure 5 we show the roots as a function of the parameter $d_i k$ corresponding to the low-β parameters of Figure 3 with $c_S = 0.3 V_A$ for three angles of propagation relative to the magnetic field, $n_z = V_{An}/V_A$. The curves are labeled according to which ideal MHD characteristic they match onto in the limit $k \to 0$. It is seen that the Hall terms increase the propagation speed of the fast wave, either increase or decrease the propagation velocity of the Alfvén wave, and decrease the velocity of the slow wave. The modification is most pronounced for propagation parallel to the background magnetic field. Only wavelengths such that $kd_i > 1$ are substantially affected.

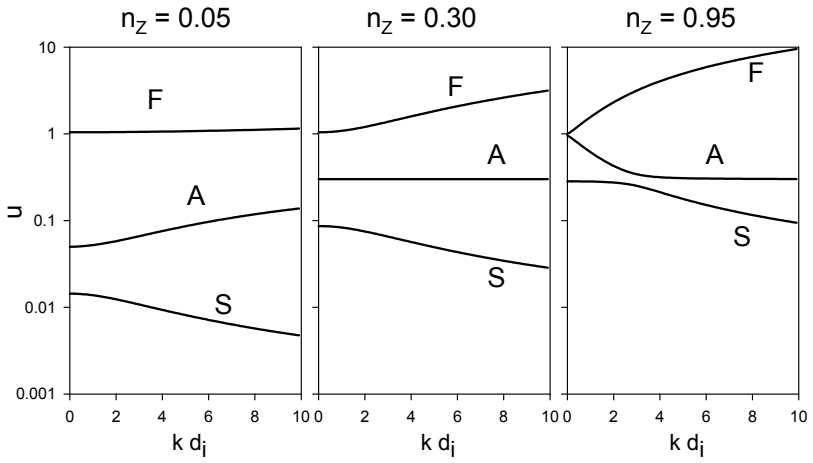

Figure 5: Typical dispersion relation for low-β two-fluid MHD for different angles of propagation relative to the background magnetic field.

2 Diffusion and Transport in Axisymmetric Geometry

In this lecture we consider the time evolution of magnetically confined tokamak plasmas over time scales that are very long compared to the Alfvén transit time, and are thus characterized by resistive diffusion and particle and heat transport. Because the electron and ion heat fluxes $\mathbf{q}_e$ and $\mathbf{q}_i$ are extremely anisotropic in a highly magnetized plasma, it becomes essential to work in a coordinate system that is aligned with the magnetic field. The derivation of an appropriate set of transport equations to deal with this is presented in Section 2.1 and its subsections. These transport equations need to be supplemented by an equilibrium constraint obtained by solving a particular form of the equilibrium equation as described in Section 2.2. Together, this system of equations provides an accurate description of the long time scale evolution of a MHD stable toroidal plasma.

2.1 Basic Equations and Orderings

Here we consider the scalar-pressure two-fluid MHD equations of Section 1.2.1, which for our purposes can be written:

$$nm_i\left(\frac{\partial \mathbf{u}}{\partial t}+\mathbf{u}\cdot\nabla\mathbf{u}\right)+\nabla p=\mathbf{J}\times\mathbf{B}, \tag{2.1}$$

$$\frac{\partial n}{\partial t}+\nabla\cdot(n\mathbf{u})=S_n, \tag{2.2}$$

$$\frac{\partial \mathbf{B}}{\partial t}=-\nabla\times\mathbf{E}, \tag{2.3}$$

$$\mathbf{E}+\mathbf{u}\times\mathbf{B}=\mathbf{R}, \tag{2.4}$$

$$\frac{3}{2}\frac{\partial p}{\partial t}+\nabla\cdot\left[\mathbf{q}+\frac{3}{2}p\mathbf{u}\right]=-p\nabla\cdot\mathbf{u}+\mathbf{J}\cdot\mathbf{R}+S_e, \tag{2.5}$$

$$\frac{3}{2}\frac{\partial p_e}{\partial t}+\nabla\cdot\left[\mathbf{q}_e+\frac{3}{2}p_e\mathbf{u}\right]=-p_e\nabla\cdot\mathbf{u}+\mathbf{J}\cdot\mathbf{R}+Q_{\Delta ei}+S_{ee}. \tag{2.6}$$

We have denoted by $\mathbf{R}$ the inhomogeneous term in the generalized Ohm's law, Eq. (2.4), by $\mathbf{q}=\mathbf{q}_i+\mathbf{q}_e$ the random heat flux vector due to ions and electrons, by $Q_{\Delta ei}$ the electron-ion equipartition term, and by S_n and $S_e=S_{ei}+S_{ee}$ sources of particles and energy. Note that in comparing Eq. (2.4) with Eq. (1.9), we have the relation

$$\mathbf{R}=\frac{1}{ne}[\mathbf{R}_e+\mathbf{J}\times\mathbf{B}-\nabla p_e-\nabla\cdot\boldsymbol{\pi}_e]. \tag{2.7}$$

We are also neglecting (assumed small) terms involving the ion and electron stress tensors π_e, π_i, and the heating due to $\nabla T_e \cdot \mathbf{J}$.

We now apply a *resistive time scale ordering* [21, 22, 24] to these equations to isolate the long time scale behavior. This consists of ordering all the source and transport terms to be the order of the inverse magnetic Lundquist number, Eq. (1.36), $S^{-1} = \epsilon \ll 1$, where ϵ is now some small dimensionless measure of the dissipation. Thus

$$\eta \sim \mathbf{R} \sim S_n \sim S_e \sim \mathbf{q} \sim \epsilon \ll 1. \tag{2.8}$$

We look for solutions in which all time derivatives and velocities are also small, of order ϵ ,

$$\frac{\partial}{\partial t} \sim \mathbf{u} \sim \epsilon \ll 1, \tag{2.9}$$

as is the electric field, $\mathbf{E} \sim \epsilon$.

Applying this ordering to Eqs. (2.1)–(2.6), we find that the last five equations remain unchanged, merely picking up the factor ϵ in every term, which can then be canceled. However, the momentum equation, Eq. (2.1), does change, with a factor of ϵ^2 multiplying only the inertial terms,

$$\epsilon^2 n m_i \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) + \nabla p = \mathbf{J} \times \mathbf{B}. \tag{2.10}$$

Thus in the limit $\epsilon \to 0$ we can neglect the inertial terms, replacing the momentum equation with the equilibrium condition

$$\nabla p = \mathbf{J} \times \mathbf{B}. \tag{2.11}$$

Equation (2.11) is correct to second order in ϵ, and provides significant simplifications since replacing Eq. (2.10) by this removes all the wave propagation characteristics from the system.

We note here that the system of equations given by Eqs. (2.11) and (2.2) through (2.6) involve the plasma velocity $\mathbf{u}$, but there is no longer a time advancement equation for $\mathbf{u}$. We will derive a method to solve this system asymptotically in spite of this apparent difficulty. We restrict consideration here to axisymmetric geometry. The most general form for an axisymmetric magnetic field consistent with Eq. (2.11) is given by

$$\mathbf{B} = \nabla\phi \times \nabla\Psi + g(\Psi)\nabla\phi. \tag{2.12}$$

Let us first consider the poloidal part of the magnetic field evolution equation, Eq. (2.3). Insertion of Eq. (2.12) into Eq. (2.3) and taking the $(\hat{R}, \hat{Z})$ projections gives an equation to evolve the poloidal flux function

$$\frac{\partial \Psi}{\partial t} = R^2 \mathbf{E} \cdot \nabla\phi + C(t). \tag{2.13}$$

The integration constant $C(t)$ can be set to zero by adopting the convention that Ψ be proportional to the actual poloidal flux that must vanish at $R = 0$. Setting $C(t) = 0$ and using Eq. (2.4) to eliminate the electric field, we have

$$\frac{\partial \Psi}{\partial t} + \mathbf{u} \cdot \nabla \Psi = R^2 \nabla \phi \cdot \mathbf{R}. \tag{2.14}$$

(Note that the symbol R is being used to represent the cylindrical coordinate, while $\mathbf{R}$, defined in Eq. (2.7), is the vector inhomogeneous term in the generalized Ohm's law.) The $\nabla\phi$ projection of Eq. (2.3) gives

$$\frac{\partial g}{\partial t} = R^2 \nabla \cdot [\nabla \phi \times \mathbf{E}],$$

or, upon substituting from Eq. (2.4),

$$\frac{\partial g}{\partial t} + R^2 \nabla \cdot \left[\frac{g}{R^2} \mathbf{u} - (\nabla \phi \cdot \mathbf{u}) \nabla \phi \times \nabla \Psi - \nabla \phi \times \mathbf{R} \right] = 0. \tag{2.15}$$

The system of time evolution equations that we are solving is thus reduced to the five scalar equations (2.2), (2.5), (2.6), (2.14), and (2.15) as well as the equilibrium equation (2.11).

2.1.1 Time-Dependent Coordinate Transformation

We adopt here an axisymmetric magnetic flux coordinate system (ψ, θ, ϕ) where ψ is a magnetic flux coordinate, θ is a poloidal angle (that goes the short way around the torus), and ϕ is the standard toroidal angle in a cylindrial coordinate system. [3]. The infinitesimal volume element associated with the differentials $d\psi, d\theta, d\phi$ is given by the *Jacobian* J defined by

$$d\tau = J d\psi d\theta d\phi = \frac{d\psi d\theta d\phi}{[\nabla \psi \times \nabla \theta \cdot \nabla \phi]} \quad . \tag{2.16}$$

Since the magnetic field and flux surfaces evolve and change in time, the coordinate transformation being considered will be a time-dependent one. At any given time we have the flux coordinates (ψ, θ, ϕ) and the inverse representation $\mathbf{x}(\psi, \theta, \phi)$, where $\mathbf{x}$ are Cartesian coordinates. We define the *coordinate velocity* at a particular (ψ, θ, ϕ) location as the time rate of change of the Cartesian coordinate at a fixed value of the flux coordinate as shown in Figure 6,

$$\mathbf{u}_C = \left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\psi,\theta,\phi}. \tag{2.17}$$

Consider now a scalar function, α, that may be thought of either as a function of Cartesian coordinates and time, $\alpha(\mathbf{x}, t)$, or of flux coordinates and time, $\alpha(\psi, \theta, \phi, t)$. Time derivatives at a fixed spatial location $\mathbf{x}$ and at fixed coordinates (ψ, θ, ϕ) are related by the chain rule of partial differentiation,

$$\left. \frac{\partial \alpha}{\partial t} \right|_{\psi,\theta,\phi} = \left. \frac{\partial \alpha}{\partial t} \right|_{\mathbf{x}} + \frac{\partial \alpha}{\partial \mathbf{x}} \cdot \left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\psi,\theta,\phi}.$$

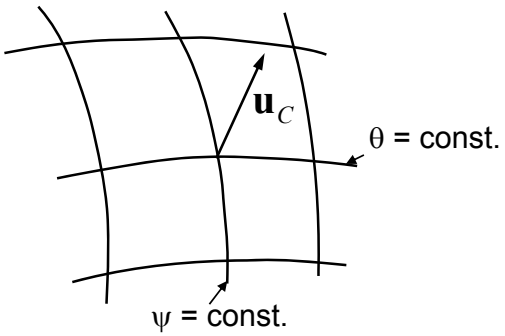

Figure 6: $\mathbf{u}_C$ is the velocity of a fluid element with a given ψ,θ value relative to a fixed Cartesian frame.

Using Eq. (2.17), we therefore have the relation

$$\left.\frac{\partial}{\partial t}\right|_{\mathbf{x}} = \left.\frac{\partial}{\partial t}\right|_{\psi,\theta,\phi} - \mathbf{u}_C \cdot \nabla. \tag{2.18}$$

We will also make use of the relation for the time derivative of the Jacobian,

$$\left.\frac{\partial J}{\partial t}\right|_{\psi,\theta,\phi} = J\nabla \cdot \mathbf{u}_C. \tag{2.19}$$

This may be verified directly.

With the introduction of the coordinate velocity in Eq. (2.17), it follows that the fluid velocity appearing in the MHD equations can be thought of as consisting of two parts,

$$\mathbf{u} = \mathbf{u}_C + \mathbf{u}_R, \tag{2.20}$$

where $\mathbf{u}_C$ is the coordinate velocity already discussed, and $\mathbf{u}_R$ is the velocity of the fluid relative to the coordinates. Since the total velocity $\mathbf{u}$ is a physical quantity which must be determined by the MHD equations, constraining either $\mathbf{u}_C$ or $\mathbf{u}_R$ to be of a particular form will determine the other.

The two parts of the velocity field can be thought of as a *Lagrangian* part and an *Eulerian* part. If the total velocity were represented with $\mathbf{u}_C$ only and there were no dissipation, the coordinates would be frozen into the fluid as it moves and distorts. If the total velocity were represented with $\mathbf{u}_R$, the coordinates would be fixed in space, and the fluid would move through them. We will see that an attractive choice is to split the velocity between these two parts so that the coordinates can move just enough to stay flux coordinates, but we allow the fluid to diffuse relative to them.

2.1.2 Evolution Equations in a Moving Frame

We now transform each of the scalar evolution equations, Eqs. (2.14), (2.15), (2.2), (2.5), and (2.6), into the moving flux coordinate frame by using the identities in Eqs. (2.18), (2.19), and (2.20). After some manipulations, we obtain the scalar equations

$$\frac{\partial \Psi}{\partial t} + \mathbf{u}_R \cdot \nabla \Psi = R^2 \nabla \phi \cdot \mathbf{R}, \tag{2.21}$$

$$\frac{\partial}{\partial t}\left(g\frac{J}{R^2}\right) + J\nabla \cdot \left[\frac{g}{R^2}\mathbf{u}_R - (\nabla\phi \cdot \mathbf{u}_R)\nabla\phi \times \nabla\Psi - \nabla\phi \times \mathbf{R}\right] = 0, \tag{2.22}$$

$$\frac{\partial}{\partial t}(nJ) + J\nabla \cdot (n\mathbf{u}_R) = JS_n, \tag{2.23}$$

$$\frac{\partial}{\partial t}\left(p^{3/5}J\right) + J\nabla \cdot \left[p^{3/5}\mathbf{u}_R\right] + \frac{2}{5}Jp^{-2/5}[\nabla \cdot q - \mathbf{J} \cdot \mathbf{R} - S_e] = 0, \tag{2.24}$$

$$\frac{\partial}{\partial t}\left(p_e^{3/5}J\right) + J\nabla \cdot \left[p_e^{3/5}\mathbf{u}_R\right] \quad + \quad \frac{2}{5}Jp_e^{-2/5}[\nabla \cdot q_e - \mathbf{J} \cdot \mathbf{R} - Q_{\Delta ei} - S_{ee}] = 0. \tag{2.25}$$

Here, and in what follows, the time derivatives are with ψ and θ held fixed, and are thus in a moving flux coordinate frame. We note that the coordinate velocity $\mathbf{u}_C$ does not appear in Eqs. (2.21)–(2.25), but only the velocity of the fluid relative to the moving coordinate, $\mathbf{u}_R = \mathbf{u} - \mathbf{u}_C$, is present. This is because the equations are of conservation form, and thus valid in a moving frame. We will use this fact to define a coordinate transformation in which the velocity vanishes altogether.

We first derive a reduced set of equations by integrating Eqs. (2.22)–(2.25) over the angle θ at fixed values of the coordinate ψ. This *flux surface averaging* leads to a set of one-dimensional evolution equations that depend only on the coordinate ψ and time t. Furthermore, by using the fact that $g = g(\psi, t)$ and $p = p(\psi, t)$ from the equilibrium constraint, and that $n \simeq n(\psi, t)$ and $p_e \simeq p_e(\psi, t)$ from the fact that the temperatures are nearly constant on flux surfaces because the parallel conductivities are large compared to their perpendicular values, we can obtain a closed set of equations that depend only on these surface averages.

We introduce the *differential volume* as

$$V'(\psi) \equiv \frac{dV}{d\psi} = 2\pi \int_0^{2\pi} d\theta J = 2\pi \oint \frac{R d\ell}{|\nabla\psi|}, \tag{2.26}$$

where $V(\psi)$ is the volume enclosed by the flux surface ψ=constant. For any scalar quantity $a(\psi, \theta)$ we define the *surface average* of a as the following integral over a constant ψ contour:

$$\langle a \rangle = \frac{2\pi}{V'} \int_0^{2\pi} d\theta J a. \tag{2.27}$$

Using this notation, it follows that for any vector $\mathbf{A}$, we have the identity

$$2\pi \int_0^{2\pi} J\nabla \cdot \mathbf{A} d\theta = \frac{\partial}{\partial \psi} \left[V' \left\langle \mathbf{A} \cdot \nabla \psi \right\rangle \right].$$

Using this, the following set of equations are obtained by integrating Eqs. (2.22)–(2.25) over the angle θ at fixed value of the flux coordinate ψ:

$$\frac{\partial}{\partial t} \left[gV' \left\langle R^{-2} \right\rangle \right] + \frac{\partial}{\partial \psi} \left[gV' \left\langle R^{-2} \mathbf{u}_R \cdot \nabla \psi \right\rangle - V' \left\langle \nabla \phi \times \mathbf{R} \cdot \nabla \psi \right\rangle \right] = 0, \quad (2.28)$$

$$\frac{\partial}{\partial t} \left[nV' \right] + \frac{\partial}{\partial \psi} \left[nV' \left\langle \nabla \psi \cdot \mathbf{u}_R \right\rangle \right] = V' \left\langle S_n \right\rangle, \quad (2.29)$$

$$\begin{aligned} &\frac{\partial}{\partial t} \left[p^{\frac{3}{5}} V' \right] + \frac{\partial}{\partial \psi} \left[p^{\frac{3}{5}} V' \left\langle \mathbf{u}_R \cdot \nabla \psi \right\rangle \right] \\ &\quad + \frac{2}{5} p^{-\frac{2}{5}} \left[\frac{\partial}{\partial \psi} \left(V' \left\langle \mathbf{q} \cdot \nabla \psi \right\rangle \right) - V' \left\langle \mathbf{J} \cdot \mathbf{R} \right\rangle - V' \left\langle S_e \right\rangle \right] = 0. \end{aligned} \quad (2.30)$$

$$\begin{aligned} &\frac{\partial}{\partial t} \left[p_e^{\frac{3}{5}} V' \right] + \frac{\partial}{\partial \psi} \left[p_e^{\frac{3}{5}} V' \left\langle \mathbf{u}_R \cdot \nabla \psi \right\rangle \right] \\ &\quad + \frac{2}{5} p_e^{-\frac{2}{5}} \left[\frac{\partial}{\partial \psi} \left(V' \left\langle \mathbf{q}_e \cdot \nabla \psi \right\rangle \right) - V' \left(\left\langle \mathbf{J} \cdot \mathbf{R} \right\rangle + \left\langle Q_{\Delta ei} \right\rangle + \left\langle S_{ee} \right\rangle \right) \right] = 0. \end{aligned} \quad (2.31)$$

As discussed above, an important constraint that must be incorporated is that the coordinates (ψ, θ, ϕ) remain *flux coordinates* as they evolve in time. To this end, we require that in the moving frame, the flux function Ψ evolve in time in such a way that the coordinate ψ remain a flux coordinate, i.e.,

$$\nabla \phi \times \nabla \psi \cdot \frac{\partial \Psi}{\partial t} = 0. \quad (2.32)$$

From Eq. (2.21), this implies

$$\nabla \Psi \cdot \mathbf{u}_R - R^2 \mathbf{R} \cdot \nabla \phi = f(\psi). \quad (2.33)$$

Here $f(\psi)$ is a presently undetermined function only of ψ. Equation (2.33) puts an important constraint on the relative velocity $\mathbf{u}_R$ and hence on the coordinate velocity $\mathbf{u}_C$ through Eq. (2.20), but it also leaves some freedom in that we are free to prescribe the function $f(\psi)$. This freedom will be used to identify the flux coordinate ψ with a particular surface function.

The system of equations given by Eqs. (2.21), (2.28)–(2.31), and (2.33) still depend upon the relative velocity $\mathbf{u}_R \cdot \nabla \psi$. This is determined up to a function only of ψ by Eq. (2.33), which follows from the constraint that constant Ψ surfaces align with constant ψ surfaces as they both evolve. We thus have a freedom in the velocity decomposition that we can use to simplify the problem.

The remaining function of ψ is determined by specifying which flux function ψ is. Three common choices are the following.

(i) Constant poloidal flux:

$$\mathbf{u}_R \cdot \nabla \Psi = R^2 \nabla \phi \cdot \mathbf{R}. \tag{2.34}$$

(ii) Constant toroidal flux:

$$-g \left\langle R^{-2} \mathbf{u}_R \cdot \nabla \psi \right\rangle = \left\langle \nabla \phi \times \nabla \psi \cdot \mathbf{R} \right\rangle. \tag{2.35}$$

(iii) Constant mass:

$$\langle \nabla \psi \cdot \mathbf{u}_R \rangle = 0. \tag{2.36}$$

Here, we choose number (ii), the toroidal magnetic flux, as it is most appropriate for most magnetic fusion applications, particularly for describing tokamaks. The toroidal field in the tokamak is primarily produced by the external field magnets, and is generally much stronger than the poloidal field. This makes it the most immobile, and thus most suitable for use as a coordinate. Also, unlike the poloidal magnetic flux, the toroidal flux at the magnetic axis does not change in time, always remaining zero.

2.1.3 Evolution in Toroidal Flux Coordinates

By combining Eq. (2.35) and the constraint Eq. (2.33), we can eliminate the free function $f(\psi)$ and solve explicitly for the normal relative velocity. This gives

$$f(\psi) = -\frac{\langle \mathbf{B} \cdot \mathbf{R} \rangle}{\langle \mathbf{B} \cdot \nabla \phi \rangle},$$

which when inserted into Eq. (2.33) yields

$$\mathbf{u}_R \cdot \nabla \psi = \frac{1}{\Psi'} \left[R^2 \mathbf{R} \cdot \nabla \phi - \frac{\langle \mathbf{B} \cdot \mathbf{R} \rangle}{\langle \mathbf{B} \cdot \nabla \phi \rangle} \right]. \tag{2.37}$$

Using Eq. (2.37) to eliminate the relative velocity $\mathbf{u}_R$ from the surface averaged transport equations allows us to identify the flux coordinate ψ with the toroidal magnetic flux inside a constant flux surface, Φ.

Following Kruskal and Kulsrud [23], we compute the *toroidal flux* inside a magnetic surface explicitly as

$$\Phi(\psi) = \frac{1}{2\pi} \int \mathbf{B} \cdot \nabla \phi d\tau = \frac{1}{2\pi} \int_{\psi_0}^{\psi} d\psi g(\psi) V' \left\langle R^{-2} \right\rangle. \tag{2.38}$$

The time derivative of this can be verified to vanish by calculating directly from the definition of Φ,

$$\left. \frac{\partial \Phi}{\partial t} \right|_{\psi} = \frac{1}{2\pi} \int_{\psi_0}^{\psi} \left[g V' \left\langle R^{-2} \right\rangle \right]_t d\psi = 0,$$

where we used Eqs. (2.28) and (2.37). It is seen that the use of Eq. (2.37) causes Φ and ψ to be stationary with respect to each other, and we may therefore adopt Φ as the flux surface label. Equation (2.28) need no longer be solved as it is intrinsically satisfied by our adoption of Φ as the flux coordinate. Since ψ and Φ are the same, we obtain a useful identity, valid for toroidal flux coordinates, by differentiating each side of Eq. (2.38) by $\psi \equiv \Phi$,

$$V' = \frac{2\pi}{g\left\langle R^{-2}\right\rangle}. \tag{2.39}$$

Using Eq. (2.37) to eliminate the relative velocity from Eqs. (2.21), (2.29), (2.30), and (2.31) then yields the surface-averaged transport equations relative to surfaces of constant toroidal flux. In deriving Eqs. (2.50) and (2.51) that follow, we make use of the equilibrium conditions to express the equilibrium current density as

$$\mathbf{J} = -R^2\frac{dp}{d\Psi}\nabla\phi - \frac{1}{\mu_0}\frac{dg}{d\Psi}\mathbf{B}.$$

We also use Eq. (2.37) and the surface average of the inverse equilibrium equation to obtain the intermediate result

$$V'\left\langle \mathbf{J}\cdot\mathbf{R}\right\rangle = -p'V'\left\langle \mathbf{u}_R\cdot\nabla\Phi\right\rangle + \frac{\left\langle \mathbf{B}\cdot\mathbf{R}\right\rangle}{\left\langle \mathbf{B}\cdot\nabla\phi\right\rangle}\frac{d}{d\Phi}\left[\frac{V'}{2\pi\mu_0 q}\left\langle\frac{|\nabla\Phi|^2}{R^2}\right\rangle\right]. \tag{2.40}$$

To express the final form of the surface-averaged transport equations, we first define the rotational transform

$$\iota \equiv \frac{1}{q} = 2\pi\frac{d\Psi}{d\Phi}, \tag{2.41}$$

the loop voltage

$$V_L \equiv 2\pi\frac{\left\langle \mathbf{B}\cdot\mathbf{R}\right\rangle}{\left\langle \mathbf{B}\cdot\nabla\phi\right\rangle}, \tag{2.42}$$

the differential particle number, or number of particles in the differential volume between surfaces Φ and $\Phi + d\Phi$,

$$N' \equiv nV', \tag{2.43}$$

the particle flux

$$\Gamma \equiv 2\pi qn\left[\left\langle R^2\mathbf{R}\cdot\nabla\phi\right\rangle - \frac{\left\langle \mathbf{B}\cdot\mathbf{R}\right\rangle}{\left\langle \mathbf{B}\cdot\nabla\phi\right\rangle}\right], \tag{2.44}$$

the differential total and electron entropy densities

$$\sigma = pV'^{\frac{5}{3}}, \qquad \sigma_e = p_eV'^{\frac{5}{3}}, \tag{2.45}$$

the surface integrated current density

$$K \equiv \frac{V'}{(2\pi)^2 \mu_0 q} \left\langle \frac{|\nabla\Phi|^2}{R^2} \right\rangle, \tag{2.46}$$

and the electron and ion heat fluxes

$$\begin{aligned} Q_e &\equiv V' \left[\langle \mathbf{q}_e \cdot \nabla\Phi \rangle + \frac{5}{2} \frac{p_e}{n} \Gamma \right], \\ Q_i &\equiv V' \left[\langle \mathbf{q}_i \cdot \nabla\Phi \rangle + \frac{5}{2} \frac{p_i}{n} \Gamma \right]. \end{aligned} \tag{2.47}$$

With these definitions, the basic transport equations take on the compact form:

$$\frac{\partial \Psi}{\partial t} = \frac{1}{2\pi} V_L, \tag{2.48}$$

$$\frac{\partial N'}{\partial t} + \frac{\partial}{\partial \Phi} V' \Gamma = V' \langle S_n \rangle, \tag{2.49}$$

$$\frac{3}{2} (V')^{-\frac{2}{3}} \frac{\partial \sigma}{\partial t} + \frac{\partial}{\partial \Phi} (Q_e + Q_i) = V_L \frac{\partial K}{\partial \Phi} + V' \langle S_e \rangle, \tag{2.50}$$

$$\frac{3}{2} (V')^{-\frac{2}{3}} \frac{\partial \sigma_e}{\partial t} + \frac{\partial Q_e}{\partial \Phi} + V' \left(\frac{\Gamma}{n} \frac{\partial p_i}{\partial \Phi} - Q_{\Delta ei} \right) = V_L \frac{\partial K}{\partial \Phi} + V' \langle S_{ee} \rangle. \tag{2.51}$$

By differentiating Eq. (2.48) with respect to the toroidal flux Φ, we obtain an evolution equation for the rotational transform, defined in Eq. (2.41),

$$\frac{\partial \iota}{\partial t} = \frac{\partial}{\partial \Phi} V_L. \tag{2.52}$$

In a similar manner, we can obtain an evolution equation for the *toroidal angular momentum density*, $\Omega(\Phi) = m_i N' \left\langle R^2 \right\rangle \omega(\Phi)$, where ω is the toroidal angular velocity:

$$\frac{\partial \Omega}{\partial t} + \frac{\partial}{\partial \Phi} V' \Gamma_\Omega = V' \langle S_\Omega \rangle. \tag{2.53}$$

The physical reason we were able to eliminate the fluid velocity entirely from the system of transport equations derived here is that transport coefficients only determine the *relative transport* of the magnetic field and plasma densities and energies with respect to one another. We cast these equations in a frame moving with the toroidal magnetic flux, and so only the relative motion remains. The absolute motion of the toroidal magnetic surfaces is determined by the equilibrium constraint and the interaction with externally applied magnetic fields. This is addressed in Section 2.2.

2.1.4 Specifying a Transport Model

Specifying a transport model consists of providing the transport fluxes, Γ, $\langle \mathbf{q}_i \cdot \nabla\Phi \rangle$, $\langle \mathbf{q}_e \cdot \nabla\Phi \rangle$, V_L, and Γ_Ω and the equipartition term $Q_{\Delta ei}$ as functions of the thermodynamic and magnetic field variables and the metric quantities. Also required are the source functions for mass, energy, electron energy, and angular momentum, $\langle S_n \rangle$, $\langle S_e \rangle$,$\langle S_{ee} \rangle$,$\langle S_\Omega \rangle$. The radiative loss function $\langle S_{RAD} \rangle$ must be subtracted from the energy and electron energy source functions.

At a given time, when the geometry is fixed, the plasma pressures, densities, and toroidal angular velocities are obtained from the adiabatic variables through the relations, valid for a two-component plasma with charge $Z = 1$ and with $n_e = n_i = n$,

$$\begin{aligned} p_e(\Phi) &= \sigma_e / V'^{5/3}, \\ p_i(\Phi) &= p - p_e = (\sigma - \sigma_e)/V'^{5/3}, \\ n(\Phi) &= N'/V', \\ k_B T_e(\Phi) &= p_e/n, \\ k_B T_i(\Phi) &= p_i/n, \\ \omega(\Phi) &= \Omega / \left(m_i N' \left\langle R^2 \right\rangle \right). \end{aligned}$$

The electron-ion equipartition term, $Q_{\Delta ei}$, is normally taken to be the classical value given by Eq. (1.22).

It is convenient to define a five-component force vector given by Φ derivatives of the density, total and electron pressures, current density, and toroidal angular velocity relative to the toroidal magnetic flux,

$$\mathbf{F} = \left[n'(\Phi), p'(\Phi), p_e'(\Phi), \left(A(\Phi)\iota(\Phi)\right)', \omega'(\Phi) \right]. \tag{2.54}$$

Here we have defined the geometrical quantity

$$A(\Phi) \equiv \frac{V'}{g} \left\langle \frac{|\nabla\Phi|^2}{R^2} \right\rangle.$$

This is related to the surface-averaged parallel current density by

$$\frac{\langle \mathbf{J} \cdot \mathbf{B} \rangle}{\langle \mathbf{B} \cdot \nabla\phi \rangle} = \frac{g^2}{(2\pi)^2 \mu_0} \left[\frac{V'}{g} \left\langle \frac{|\nabla\Phi|^2}{R^2} \right\rangle \iota \right]' = \frac{g^2}{(2\pi)^2 \mu_0} \left(A(\Phi)\iota(\Phi)\right)'.$$

The transport fluxes can now be expressed as a matrix of functions multiplying the force vector:

$$\begin{aligned} \Gamma &= \sum_{j=1}^{5} \Gamma^j F_j, \qquad \langle \mathbf{q}_i \cdot \nabla\Phi \rangle = \sum_{j=1}^{5} q_i^j F_j, \qquad \langle \mathbf{q}_e \cdot \nabla\Phi \rangle = \sum_{j=1}^{5} q_e^j F_j, \\ V_L &= \sum_{j=1}^{5} V_L^j F_j + V_L^6, \qquad \Gamma_\Omega = \sum_{j=1}^{5} \Gamma_\Omega^j F_j. \end{aligned}$$

With this convention, specification of the 25 scalar functions $\Gamma^1, \Gamma^2, \cdots \Gamma^5_\Omega$ will specify the transport model. The additional term V_L^6 will be used to incorporate a source term for current drive.

A. Pfirsch–Schlüter regime plasma

The non-zero scalar coefficients needed to evaluate the transport fluxes for an electron-ion plasma in the collision-dominated regime are as follows [25, 26]:

$$\begin{aligned}
\Gamma^1 &= -L_{12}p_e, \qquad \Gamma^2 = -L_{11}n, \qquad \Gamma^3 = L_{12}n, \\
\Gamma^4 &= -\frac{\eta_\parallel q g^2 n}{2\pi\mu_0}\left\langle \frac{|\nabla\psi|^2}{R^2}\right\rangle \left\langle B^2\right\rangle^{-1}, \quad q_i^1 = L_i p_i^2/n, \\
q_i^2 &= -L_i p_i, \qquad q_i^3 = L_i p_i, \qquad q_e^1 = L_{22}p_e^2/n, \\
q_e^2 &= L_{12}p_e, \qquad q_e^3 = -L_{22}p_e, \qquad V_L^4 = \frac{\eta_\parallel g^2}{2\pi\mu_0}.
\end{aligned}$$

The transport coefficients are:

$$\begin{aligned}
L_{11} &= L_0\left[1 + 2.65(\eta_\parallel/\eta_\perp)q_*^2\right], \\
L_{12} &= (3/2)L_0\left[1 + 1.47(\eta_\parallel/\eta_\perp)q_*^2\right], \\
L_{22} &= 4.66 L_0\left[1 + 1.67(\eta_\parallel/\eta_\perp)q_*^2\right], \\
L_i &= \sqrt{2}L_0(m_i/m_e)^{1/2}(T_e/T_i)^{3/2}\left[1 + 1.60 q_*^2\right], \\
L_0 &= \frac{\eta_\perp}{\mu_0}\left\langle |\nabla\Phi|^2/B^2\right\rangle.
\end{aligned}$$

Here, we have introduced the function

$$q_*^2 = \frac{1}{2}\left[\left\langle B^{-2}\right\rangle - \left\langle B^2\right\rangle^{-1}\right]\frac{g^2}{\langle|\nabla\Psi|^2/B^2\rangle}, \tag{2.55}$$

which reduces to the square of the safety factor, q^2, in the low beta, large aspect ratio limit. The resistivity functions $\eta_\parallel$ and $\eta_\perp$ are given by Eqs. (1.20) and (1.21).

If an external source of current drive is present, we can incorporate it into this model with the V_L^6 coefficient by defining

$$V_L^6 = -2\pi\eta_\parallel J_\parallel^{CD}, \tag{2.56}$$

where

$$J_\parallel^{CD} \equiv \frac{\left\langle \mathbf{J}^{CD}\cdot\mathbf{B}\right\rangle}{\langle\mathbf{B}\cdot\nabla\phi\rangle}, \tag{2.57}$$

with $\mathbf{J}^{CD}$ being the external current drive vector.

B. Banana Regime Plasma

The transport coefficients for a low-collisionality plasma including trapped and circulating particles have been computed for an arbitrary aspect ratio and cross section shape [26] with the restriction that $|\mathbf{B}|$ have only a single maximum on each flux surface [24, 26, 27, 28]. The fraction of trapped particles on a flux surface is given by [24]

$$f_t = 1 - (3/4)\left\langle B^2 \right\rangle \int_0^{B_c^{-1}} \lambda d\lambda / \left\langle (1-\lambda B)^{1/2} \right\rangle, \tag{2.58}$$

where B_c is the maximum value of $|B|$ on a flux surface. Using this, we can express the banana regime transport model as follows:

$$\begin{aligned}
\Gamma^1 &= p_i L_{11}^{bp} y - p_e \left(L_{12}^{bp} + \tilde{L}_{12} + L_E L_{13} L_{23} \right), \\
\Gamma^2 &= n(1+y)\left[-L_{11}^{bp} - L_E (L_{13})^2 \right] - n\tilde{L}_{11}, \\
\Gamma^3 &= n\left[L_{11}^{bp} y + L_{12}^{bp} + \tilde{L}_{12} + L_E L_{13}(L_{23} + y L_{13}) \right], \\
\Gamma^4 &= -L_E L_{13} (2\pi)^{-2} g^3 n \left\langle R^{-2} \right\rangle,
\end{aligned}$$

$$\begin{aligned}
q_i^1 &= L_i^{nc} p_i^2 / n^2, \\
q_i^2 &= -L_i^{nc} p_i / n, \\
q_i^3 &= L_i^{nc} p_i / n,
\end{aligned}$$

$$\begin{aligned}
q_e^1 &= \left[-L_{12}^{bp} y p_i + L_{22}^{nc} p_e + L_E L_{23} \left(-L_{13} y p_i + L_{23} p_e \right) \right] p_e / n, \\
q_e^2 &= \left[\left(L_{12}^{bp} + L_E L_{23} L_{13} \right)(1+y) + \tilde{L}_{12} \right] p_e, \\
q_e^3 &= \left[-L_{22}^{nc} - L_{12}^{bp} y - L_E L_{23} (L_{13} y + L_{23}) \right] p_e, \\
q_e^4 &= L_E L_{23} (2\pi)^{-2} g^3 \left\langle R^{-2} \right\rangle,
\end{aligned}$$

$$\begin{aligned}
V_L^1 &= L_E^* L_{13} p / n, \\
V_L^2 &= L_E^* L_{13} n (y+1), \\
V_L^3 &= L_E^* \left(L_{13} - L_{23} \right) n, \\
V_L^4 &= L_E g^2 \left\langle B^2 \right\rangle / 2\pi.
\end{aligned}$$

Here, the transport coefficients are:

$$\begin{aligned}
L_{11}^{bp} &= L_*(1.53 - 0.53 f_t), \\
L_{12}^{bp} &= L_*(2.13 - 0.63 f_t), \\
L_{13} &= (2\pi q) g f_t (1.68 - 0.68 f_t), \\
L_{23} &= 1.25 (2\pi q) g f_t (1 - f_t),
\end{aligned}$$

$$
\begin{aligned}
\tilde{L}_{11} &= L_0(1+2q_*^2), \\
\tilde{L}_{12} &= (3/2)L_0(1+2q_*^2),
\end{aligned}
$$

$$
\begin{aligned}
L_{22}^{nc} &= 4.66\left[L_0(1+2q_*^2)+L_*\right], \\
L_i^{nc} &= \sqrt{2}(m_i/m_e)^{1/2}(T_e/T_i)^{3/2}\left[L_0(1+2q_*^2)+0.46L_*(1-0.54f_t)^{-1}\right], \\
y &= -1.17(1-f_t)(1-0.54f_t)^{-1},
\end{aligned}
$$

$$
\begin{aligned}
L_{33} &= -1.26f_t(1-0.18f_t), \\
L_* &= f_t(2\pi qg)^2\frac{\eta_\perp}{\mu_0}/\left\langle B^2\right\rangle, \\
L_E &= \frac{\eta_\parallel(1+L_{33})^{-1}}{\mu_0\left\langle B^2\right\rangle}, \\
L_E^* &= L_E\frac{2\pi\left\langle B^2\right\rangle}{g\left\langle R^{-2}\right\rangle}
\end{aligned}
$$

Note that in steady state, when the loop voltage V_L is a spatial constant, the banana regime model implies a current driven by the temperature and density gradients,

$$
\left\langle\frac{1}{R^2}\right\rangle\frac{\langle\mathbf{J}\cdot\mathbf{B}\rangle}{\langle\mathbf{B}\cdot\nabla\phi\rangle} = -f_t\,p\left[\frac{L_{13}^*}{n}\frac{dn}{d\Psi}+\frac{L_{13}^*(1+y)}{T_e+T_i}\frac{dT_i}{d\Psi}+\frac{L_{13}^*-L_{23}^*}{T_e+T_i}\frac{dT_e}{d\Psi}\right]. \tag{2.59}
$$

Here $L_{13}^* \equiv L_{13}/(2\pi qgf_t)$, $L_{23}^* \equiv L_{23}/(2\pi qgf_t)$, and y are all dimensionless and of order unity. In deriving Eq. (2.59) we have used the relation $2\pi qd/d\Phi = d/d\Psi$, where Ψ is the poloidal flux function. This current is called the *bootstrap current*. These relations have been extended to multi-charged ions [29] and to arbitrary collisionality [30]. An excellent overview of the results for the transport coefficients in all of the collisionality regimes is given in Helander and Sigmar [31].

As in the collisional regime case, an external source of current drive can be included by defining the equivalent of Eq. (2.56). In the presence of external current drive, we would add the coefficient

$$
V_L^6 = -2\pi\eta_\parallel\left(1+L_{33}\right)^{-1}J_\parallel^{CD},
$$

where $J_\parallel^{CD}$ is defined in Eq. (2.57).

C. Anomalous Transport Model

Although this is still an area of active research, there are a number of anomalous transport models available that purport to calculate local values of the surface averaged transport coefficients based on local values of the surface averaged profiles of density, the temperatures, angular velocity, and current profile, and

their gradients. These profiles are often obtained from a fit to a subsidiary micro-instability calculation [32, 33, 34, 35].

A typical model would return the particle diffusivity D, the electron thermal diffusivity χ_e, the ion thermal diffusivity χ_i, and the toroidal angular velocity diffusivity, χ_ω, all with dimension m^2/s, and a parallel resistivity function $\eta_\parallel$ with units $\Omega - m$. A relatively simple diagonal model would fit into the above formalism as follows [36]:

$$\begin{aligned}
\Gamma^1 &= -D\left\langle|\nabla\Phi|^2\right\rangle, \\
q_i^1 &= \chi_i \frac{p_i}{n}\left\langle|\nabla\Phi|^2\right\rangle, \qquad q_i^2 = -\chi_i\left\langle|\nabla\Phi|^2\right\rangle, \qquad q_i^3 = \chi_i\left\langle|\nabla\Phi|^2\right\rangle, \\
q_e^1 &= \chi_e \frac{p_e}{n}\left\langle|\nabla\Phi|^2\right\rangle, \qquad q_e^3 = -\chi_e\left\langle|\nabla\Phi|^2\right\rangle, \\
V_L^4 &= \frac{\eta_\parallel g^2}{2\pi\mu_0}, \\
\Gamma_\Omega^5 &= -\chi_\omega\left\langle|\nabla\Phi|^2\right\rangle m_i n\left\langle R^2\right\rangle.
\end{aligned}$$

Bootstrap and current drive terms are included in these models by defining the additional V_L^i coefficients as discussed above.

The micro-instability-based models tend to return transport coefficients that have strong dependences on the gradients of the corresponding surface averaged profiles. Special computational techniques have been developed for dealing with these[3].

2.2 Equilibrium Constraint

The variables N', σ, ι, σ_e, and Ω introduced in Section 2.1 are called *adiabatic variables*. If there is no dissipation and no explicit sources of mass or energy, then the time derivatives of these quantities are zero in the toroidal flux coordinate system being used here.

In the presence of dissipation, the surface-averaged transport equations of the last section describe how these adiabatic variables evolve relative to equilibrium magnetic surfaces with fixed values of toroidal magnetic flux. To complete the description, we need to solve a global equation to describe how these surfaces evolve relative to a fixed laboratory frame in which the toroidal field and poloidal field magnetic coils are located. This is the associated equilibrium problem. In the next subsection, we describe the circuit equations that describe how the nearby coil currents evolve in time due to applied and induced voltages. Then, we describe two approaches for incorporating the equilibrium constraint, the Grad–Hogan method and an accelerated form of the Taylor method.

2.2.1 Circuit Equations

The toroidal plasma is coupled electromagnetically to its surroundings; both passive structures and poloidal field coils that are connected to power supplies. We assume here that the external conductors are all axisymmetric, and that

their currents and applied voltages are in the toroidal (ϕ) direction. Although the passive structures are continuous, it is normally adequate to subdivide them into discrete elements, each of which obeys a discrete circuit equation. Thus, each of the poloidal field coils and passive structure elements obeys a circuit equation of the form

$$\frac{d}{dt}\Psi_{Pi} + R_i I_i = V_i, \tag{2.60}$$

where the poloidal flux at each coil i is defined by

$$\Psi_{Pi} = L_i I_i + \sum_{i \neq j} M_{ij} I_j + 2\pi \int_P J_\phi(\mathbf{R}') G(\mathbf{R}_i, \mathbf{R}') d\mathbf{R}'. \tag{2.61}$$

Here, R_i and V_i are the resistance and applied voltage at coil i, L_i is the self-inductance of coil i, and M_{ij} is the mutual inductance of coil i with coil j. For a passive conductor, the corresponding $V_i = 0$.

The last term in Eq. (2.61) is an integral over the plasma volume. This represents the mutual inductance between the distributed plasma current and the conductor with index i.

2.2.2 Grad–Hogan Method

The Grad–Hogan method [21, 37] splits every time step into two parts. In the first part, the adiabatic variables, including the poloidal flux at the conductors, are advanced from time t to time $t + \delta t$ by solving Eqs. (2.49)–(2.53) and (2.60) using standard techniques for parabolic equations [3]. In the second part, these adiabatic variables are held fixed while we solve the appropriate form of the equilibrium equation, where the "free functions" $p'(\Psi)$ and $gg'(\Psi)$ have been expressed in terms of the adiabatic variables $\sigma(\Phi)$ and $\iota(\Phi)$. The individual PF coil and conductor currents will change during this part of the time step in order to keep the poloidal flux fixed at each coil location and at the plasma magnetic axis. This part of the time step effectively determines the absolute motion of the toroidal flux surfaces relative to a fixed frame.

The adiabatic variables may also be expressed in terms of the poloidal flux function Ψ. This is most convenient when solving the Grad–Shafranov equation in the Grad–Hogan method as the poloidal flux function is what is being solved for. Thus, if we define $V_\Psi \equiv dV/d\Psi$ and $\sigma_\Psi \equiv pV_\Psi^{5/3}$, the form of the equilibrium equation that needs to be solved is:

$$\Delta^*\Psi + \mu_0 R^2 \frac{d}{d\Psi}\left[\frac{\sigma_\Psi(\Psi)}{V_\Psi^{5/3}}\right] + \frac{(2\pi)^4 q(\Psi)}{V_\Psi \langle R^{-2}\rangle}\frac{d}{d\Psi}\left[\frac{q(\Psi)}{V_\Psi \langle R^{-2}\rangle}\right] = 0. \tag{2.62}$$

The functions $\sigma_\Psi(\Psi)$ and $q(\Psi)$ must be held fixed while finding the equilibrium solution.

Equation (2.62) for Ψ can be solved using standard techniques [3]. Note that the poloidal flux at the magnetic axis, Ψ_{MA}, must be held fixed during the

equilibrium solution as well. Since $\nabla\Psi = 0$ at the magnetic axis, the evolution equation for Ψ at the axis is simply

$$\frac{\partial \Psi_{MA}}{\partial t} = R^2 \nabla\phi \cdot \mathbf{R}. \tag{2.63}$$

(Note that this is equivalent to applying Eq. (2.48) to Ψ at the magnetic axis.) The value of Ψ_{MA} is evolved during the transport part of the time step using Eq. (2.63) and Ψ_{MA} and the adiabatic variables are then held fixed during the equilibrium solution part of the time step.

This completes the formalism needed to describe the transport in an axisymmetric system. The one-dimensional evolution equations, Eqs. (2.49)–(2.53), advance the adiabatic variables in time on the resistive time scale. Equation (2.63) is used to advance the value of Ψ at the magnetic axis, and Eq. (2.60) is used to advance the value of Ψ at the nearby conductors. The equilibrium equation, Eq. (2.62), defines the flux surface geometry consistent with these adiabatic variables and the boundary conditions. It does not introduce any new time scales into the equations.

2.2.3 Taylor Method (Accelerated)

J. B. Taylor [38] suggested an alternative to the Grad–Hogan method that does not require solving the equilibrium equation with the adiabatic constraints, Eq. (2.62). His approach involves solving for the velocity field $\mathbf{u}$, which when inserted into the field and pressure evolution equations, Eqs. (2.3), (2.4), and (2.5), will result in the equilibrium equation, Eq. (2.11), continuing to be satisfied as time evolves. It was shown by several authors [22, 24] that an elliptic equation determining this velocity field could be obtained by time differentiating the equilibrium equation and substituting in from the time evolution equations. Taking the time derivative of Eq. (2.11) gives

$$\nabla\dot{p} = \dot{\mathbf{J}} \times \mathbf{B} + \mathbf{J} \times \dot{\mathbf{B}}, \tag{2.64}$$

or, by substituting in from Eqs. (2.3), (2.4), and (2.5),

$$\begin{aligned} &\frac{2}{3}\nabla\left[-\nabla\cdot\left(\mathbf{q}+\frac{5}{2}p\mathbf{u}\right)+\mathbf{J}\cdot(-\mathbf{u}\times\mathbf{B}+\mathbf{R})+S_e\right] \\ &+\mu_0^{-1}\nabla\times[\nabla\times(\mathbf{u}\times\mathbf{B}-\mathbf{R})]\times\mathbf{B}+\mathbf{J}\times[\nabla\times(\mathbf{u}\times\mathbf{B}-\mathbf{R})]\,. \end{aligned} \tag{2.65}$$

This equation can, in principle, be solved for $\mathbf{u}$, and that velocity field $\mathbf{u}$ can be used to keep the system in equilibrium without repeatedly solving the equilibrium equation. While this approach has been shown to be viable [39], it suffers from "drifting" away from an exact solution of the equilibrium equation. A preferred approach [40, 41] to obtaining this velocity is to use the accelerated steepest descent algorithm which involves obtaining the velocity from the residual equation,

$$\dot{\mathbf{u}} + \frac{1}{\tau}\mathbf{u} = D\left[\mathbf{J}\times\mathbf{B} - \nabla p\right]. \tag{2.66}$$

By choosing the proportionality and damping factors, D and τ, appropriately, the system can be kept arbitrarily close to an equilibrium state as it evolves. This is equivalent to applying the dynamic relaxation method to the plasma equilibrium problem.

In using this approach, the magnetic field variables Ψ and g must be evolved in time as two-dimensional functions from Eqs. (2.14) and (2.15) in order to preserve the magnetic flux constraints. Equation (2.52) for ι is then redundant but is useful as a check. The total entropy constraint is preserved by representing the pressure as $p(\psi) = \sigma(\psi)/V'^{5/3}$. The other adiabatic variables, including the poloidal flux at the conductors, are still evolved by solving Eqs. (2.49), (2.50), (2.51), (2.53), and (2.60).

3 The Galerkin Finite Element Method

The finite element method has emerged as an extremely powerful technique for solving systems of partial differential equations. As described in the now-classic textbook by Strang and Fix [42], the method was originally developed on intuitive grounds by structural engineers. It was later put on a sound theoretical basis by numerical analysts and mathematicians once it was realized that it was, in fact, an instance of the Rayleigh–Ritz technique. It is similar to the finite difference method in that it is basically a technique for generating sparse matrix equations that describe a system. However, it provides a prescription for doing this that is in many ways more straightforward than for finite differences, especially when going to high-order approximations.

3.1 Galerkin Method in One Dimension

The original Rayleigh-Ritz (or just Ritz) technique applies only to problems of the classical variational type, in which a convex functional is being minimized. The corresponding Euler differential equation is self-adjoint and elliptic. The Galerkin method is a means for extending the finite element technique to differential equations (or systems of differential equations) which are not necessarily the Euler equation for corresponding variational statements.

Consider the ordinary differential equation in one dimension

$$\mathcal{L}\{U\} = f, \tag{3.1}$$

where $\mathcal{L}$ is a linear differential operator, $f(x)$ is a known function, and $U(x)$ is the function to be solved for. We multiply by a test function $V(x)$ and integrate over the domain,

$$(\mathcal{L}\{U\}, V) = (f, V). \tag{3.2}$$

Here we have introduced the notation that for two functions f and g, and a one-dimensional domain $a \leq x \leq b$, we define the inner product to be

$$(f, g) \equiv \int_a^b f(x)g(x)dx. \tag{3.3}$$

If we introduce a solution space to which U belongs, and require that Eq. (3.2) be satisfied for every test function $V(x)$ in that function space, it is called the *weak form* of Eq. (3.1). Mathematically, the solution space might be the infinite dimensional Hilbert space with finite energy, H^1. In applying the finite element method, we introduce a finite dimensional subspace of H^1 we call S. The Galerkin method is just the discretization of the weak form within the subspace S.

There is a close connection between the Galerkin method and the Ritz method if $\mathcal{L}$ is a *self-adjoint operator*. Let us take for an example the operator

$$\mathcal{L} = -\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x),$$

and as our interval $[a, b] = [0, 1]$. After an integration by parts, we obtain

$$a(U, V) = (f, V), \tag{3.4}$$

where $a(U, V)$ is the energy integral that appears in the Ritz variational statement of the problem:

$$a(U, V) \equiv \int_0^1 [p(x)U'(x)V'(x) + q(x)U(x)V(x)]\, dx. \tag{3.5}$$

Here, we assumed that the boundary conditions are such that the boundary terms in the integration by parts vanish. In general, if the order of $\mathcal{L}$ is $2m$, we can shift m derivatives to v by repeated integration by parts.

To discretize Eq. (3.2), let us take $\phi_1, \cdots, \phi_N$ as the basis for both the solution space and for the test function space. We represent the solution as a linear combination of the ϕ_j,

$$U = \sum_{j=1}^{N} q_j \phi_j,$$

where the amplitudes q_j are to be solved for. We then have the system

$$\sum_{j=1}^{N} q_j \left(\mathcal{L}\{\phi_j\}, \phi_i\right) = (f, \phi_i)\,; \qquad\qquad i = 1, \cdots, N,$$

or, in matrix form

$$\mathbf{G} \cdot \mathbf{Q} = \mathbf{F}, \tag{3.6}$$

where

$$\begin{aligned} G_{ij} &= (\mathcal{L}\{\phi_j\}, \phi_i), \\ F_i &= (f, \phi_i). \end{aligned}$$

Assuming that the solution and test space are polynomials with finite energy for the mth derivative, the elements in the matrix in Eq. (3.6) can be integrated by parts m times so that

$$(\mathcal{L}\{\phi_j\}, \phi_i) \to a(\phi_j, \phi_i),$$

in which case it becomes identical with the discretization of the Ritz case.

The expected rate of convergence for the Galerkin method is [42]

$$\|U - u\|_0 = \mathcal{O}\left(h^k, h^{2(k-m)}\right),$$

where the finite element solution space and test space S is of degree $k - 1$, and the differential equation is order $2m$.

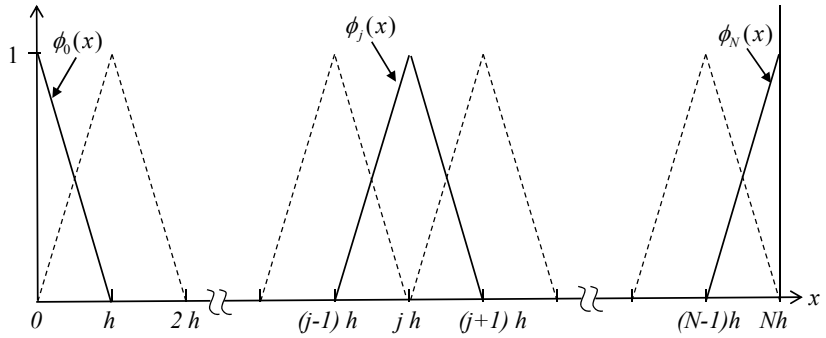

Figure 7: Linear finite elements ϕ_j equals 1 at node j, 0 at all others, and are linear in between. A half element is associated with the boundary nodes at $j = 0$ and $j = N$.

3.1.1 Linear Elements

We divide the interval $[0, 1]$ into N equally spaced intervals of spacing $h = 1/N$. The simplest choice for the subspace S is the space of functions that are linear over each interval $[(j-1)h, jh]$ and continuous at the nodes $x = jh$. For $j = 0, \cdots, N$ let ϕ_j be the function in S which equals one at the particular node $x = jh$, and vanishes at all the others (Figure 7). Every member of S can be written as a linear combination of the ϕ_j,

$$U(x) = \sum_{j=0}^{N} q_j \phi_j. \tag{3.7}$$

The "tent functions" ϕ_j are the simplest members of the more general set of piecewise polynomials. Note that the amplitudes q_j are the nodal values of the function and thus have direct physical significance. Also, note that the functions ϕ_j form a local basis in that they are orthogonal to all other elements except ϕ_{j+1} and ϕ_{j-1}. Only adjacent elements are directly coupled.

Let us now return to the problem at hand using the linear elements so that we approximate $U(x)$ as in Eq. (3.7). For simplicity, we will now take the functions p and q to be constants and the boundary conditions to be $U(0) = U(1) = 0$. To compute the terms on row i of the matrix equation, we compute separately each subinterval in which ϕ_i is non-zero. Evaluating the elements as shown in Figure 8a, we have for $i = 1, N-1$

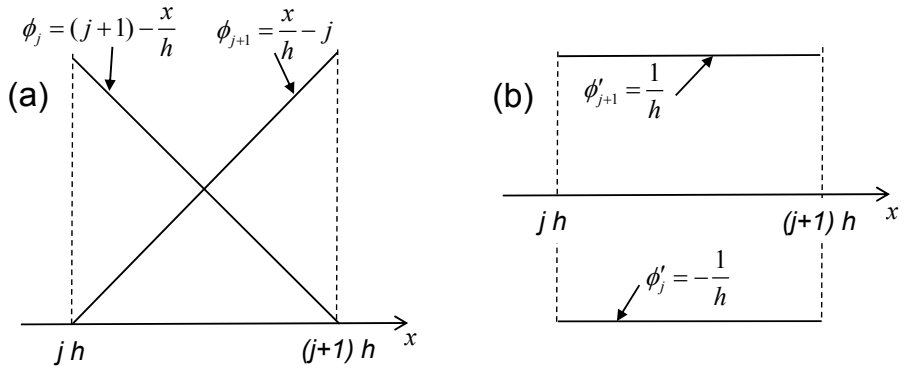

Figure 8: Expressions needed for overlap integrals for functions (a) and derivatives (b) for linear elements.

$$\begin{aligned}
\int_0^1 \phi_i \sum_{j=0}^{N} q_j \phi_j dx &= \int_{(i-1)h}^{ih} \left[\frac{x}{h} - i + 1\right] \sum_{j=i-1}^{i} q_j \phi_j dx \\
&+ \int_{ih}^{(i+1)h} \left[i + 1 - \frac{x}{h}\right] \sum_{j=i}^{i+1} q_j \phi_j dx \\
&= \int_{(i-1)h}^{ih} \left[\frac{x}{h} - i + 1\right] \left\{q_{i-1} \left[i - \frac{x}{h}\right] + q_i \left[\frac{x}{h} - i + 1\right]\right\} dx \\
&+ \int_{ih}^{(i+1)h} \left[i + 1 - \frac{x}{h}\right] \left\{q_i \left[i + 1 - \frac{x}{h}\right] + q_{i+1} \left[\frac{x}{h} - i\right]\right\} dx \\
&= \frac{h}{6} \left[q_{i-1} + 4q_i + q_{i+1}\right] \qquad (3.8)
\end{aligned}$$

Similarly, using the relations as shown in Figure 8b,

$$\int_0^1 \phi'_i \sum_{j=0}^{N} q_j \phi'_j dx = \frac{1}{h} \left[-q_{i-1} + 2q_i - q_{i+1}\right].$$

We also need to compute the vector corresponding to the linear term, which corresponds to a load. This is given by

$$\int_0^1 \phi_i f(x) dx. \qquad (3.9)$$

The usual method to evaluate Eq. (3.9) is to approximate f by linear interpolation at the nodes, i.e.,

$$f(x) = \sum_{j=0}^{N} f_j \phi_j(x), \qquad (3.10)$$

where f_j is the value of f at the node points $x = jh$.

We next take note of the boundary condition to set $q_0 = 0$ and $q_N = 0$ and thus eliminate them from the system so that we only need to solve for the amplitudes $q_1, \cdots, q_{N-1}$. We then have the matrix equation

$$\mathbf{K} \cdot \mathbf{q} = \mathbf{F}, \tag{3.11}$$

where $\mathbf{q}$ is the vector of unknown amplitudes

$$\mathbf{q} = [q_1, q_2, \cdots, q_{N-1}], \tag{3.12}$$

and the matrix $\mathbf{K}$ is given by the sum of two matrices, commonly called the "stiffness" matrix and the "mass" matrix.

$$\mathbf{K} = \frac{p}{h}\begin{bmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & \cdot & & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{bmatrix} + \frac{qh}{6}\begin{bmatrix} 4 & 1 & & & & & \\ 1 & 4 & 1 & & & & \\ & 1 & 4 & \cdot & & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & 1 & 4 & 1 \\ & & & & & 1 & 4 \end{bmatrix} \tag{3.13}$$

The load vector is given by

$$\mathbf{F} = \frac{h}{6}\begin{bmatrix} f_0 + 4f_1 + f_2 \\ f_1 + 4f_2 + f_3 \\ \cdots \\ f_{j-1} + 4f_j + f_{j+1} \\ \cdots \\ f_{N-2} + 4f_{N-1} + f_N \end{bmatrix}. \tag{3.14}$$

Finally, the solution is obtained by solving the matrix equation in Eq. (3.11). We see that in this example $\mathbf{K}$ is a tridiagonal matrix (that is also diagonally dominant and symmetric) so that this matrix equation can be solved efficiently with standard methods [3].

3.1.2 Some Definitions

The finite element subspace is said to be of degree $k - 1$ if it contains in each element a complete polynomial of this degree. For linear elements, as considered in Section 3.1.1, the degree is one and $k = 2$.

The subspace S is said to possess continuity C^{q-1} if the q^{th} derivative exists everywhere but is discontinuous at element boundaries. The linear elements have $q = 1$ and are therefore C^0. The Galerkin method can be applied to a differential equation of order $2m$ if $q \geq m$. This is because the associated variational statement is only of order m, but the integrations by parts can increase the order of the differential equation over that of the variational integral by a factor of 2.

For the subsequent error analysis, we introduce norms of the square roots of the energies. For an interval $[a, b]$ and a function $u(x)$, we define the norms

$$\begin{aligned} \|u\|_0 &\equiv \left[\int_a^b \left(u(x)\right)^2 dx\right]^{\frac{1}{2}}, \\ \|u\|_2 &\equiv \left[\int_a^b \left[\left(u''(x)\right)^2 + \left(u'(x)\right)^2 + \left(u(x)\right)^2\right] dx\right]^{\frac{1}{2}}, \end{aligned}$$

etc.

3.1.3 Hermite Cubic Elements

The linear elements discussed in Section 3.1.1 are continuous, but their first derivatives are not. (However, the first derivatives have finite energy, and hence they belong to H^1.) These elements are therefore classified as C^0. It is also possible to define a class of piecewise quadratic or piecewise cubic elements that are C^0 [42]. However, there is a special class of C^1 cubic elements, called *Hermite cubics*, that are constructed such that not only is the solution guaranteed to be continuous across element boundaries, but so is the first derivative. These can therefore be used on variational problems up to second order, or on differential equations up to fourth order by performing two integrations by parts.

There are two elements, or basis functions, associated with each node j that we will refer to as $\Phi_{1,j}(x)$ and $\Phi_{2,j}(x)$ (or sometimes just as Φ_1 and Φ_2 in context). The first, $\Phi_{1,j}(x)$, has value unity at node j but its derivative $\Phi'_{1,j}(x)$ vanishes there. The second basis function, $\Phi_{2,j}(x)$, is zero at node j but its derivative, $\Phi'_{2,j}(x)$, takes on the value of unity there. Both $\Phi_{1,j}(x)$ and $\Phi_{2,j}(x)$ and their derivatives $\Phi'_{1,j}(x)$ and $\Phi'_{2,j}(x)$ vanish at nodes $j+1$ and $j-1$. These elements are defined as follows and illustrated in Figure 9:

$$\begin{aligned} \Phi_1(y) &= \left(|y| - 1\right)^2 \left(2|y| + 1\right), \\ \Phi_2(y) &= y\left(|y| - 1\right)^2, \\ \Phi_{1,j}(x) &= \Phi_1\left(\frac{x - x_j}{h}\right), \\ \Phi_{2,j}(x) &= h\,\Phi_2\left(\frac{x - x_j}{h}\right). \end{aligned}$$

Since a cubic polynomial has four coefficients, it is completely determined in an interval by the value of the function and its derivative at the endpoints. The cubic polynomial in the interval between grid point j and $j + 1$ can therefore

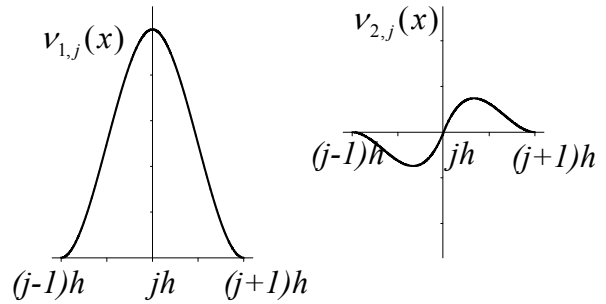

Figure 9: C^1 Hermite cubic functions enforce continuity of $v(x)$ and $v'(x)$. A linear combination of these two functions is associated with node j.

be written

$$\begin{aligned} v(x) &= v_j\,\Phi_{1,j}(x) + v'_j\,\Phi_{2,j}(x) + v_{j+1}\,\Phi_{1,j+1}(x) + v'_{j+1}\,\Phi_{2,j+1}(x), \\ &= v_j + v'_j x + \left(-3v_j - 2hv'_j + 3v_{j+1} - hv'_{j+1}\right)\left(\frac{x}{h}\right)^2 \\ &\quad + \left(2v_j + hv'_j - 2v_{j+1} + hv'_{j+1}\right)\left(\frac{x}{h}\right)^3, \\ &= a_0 + a_1 x + a_2 x^2 + a_3 x^3. \end{aligned}$$

In matrix form, the four coefficients of the cubic polynomial in the interval $[x_j, x_{j+1}]$ can thereby be related to the values of the function and its derivative at the endpoints by

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{3}{h^2} & -\frac{2}{h} & \frac{3}{h^2} & -\frac{1}{h} \\ \frac{2}{h^3} & \frac{1}{h^2} & -\frac{2}{h^3} & \frac{1}{h^2} \end{bmatrix} \cdot \begin{bmatrix} v_j \\ v'_j \\ v_{j+1} \\ v'_{j+1} \end{bmatrix}. \tag{3.15}$$

The error $\|u - U\|_0$ in these elements when applied to a second-order differential equation such as Eq. (3.1) to be $\mathcal{O}(h^4)$. This is also what we would infer from the fact that a local piecewise cubic expansion can match the terms in a Taylor series expansion through x^3 and so we would expect the local error would be of order $x^4 \approx h^4$. Boundary conditions at node J are readily incorporated into these matrices by replacing the row corresponding to either v_J or v'_J by a row with a 1 on the diagonal and the other elements zero, and with the boundary value on the right side.

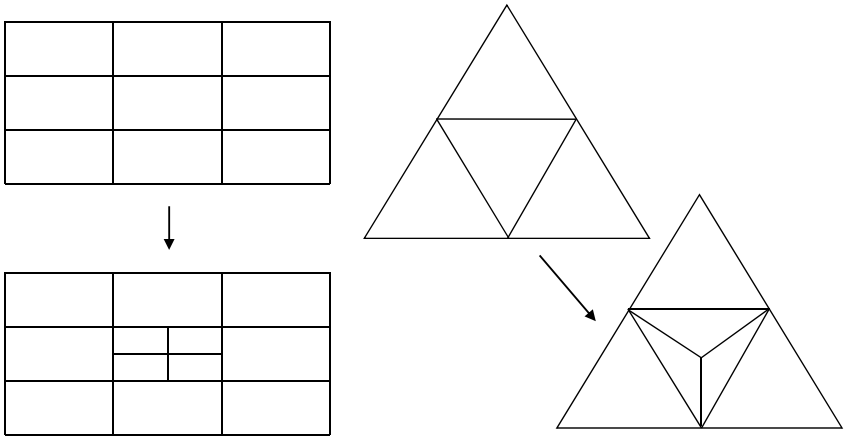

Figure 10: Rectangular elements cannot be locally refined without introducing hanging nodes. Triangular elements can be locally refined, but require unstructured mesh data structures in programming.

3.2 Finite Elements in Two Dimensions

In two dimensions, finite elements occupy a finite area rather than a finite interval. This area can be a rectangle or a triangle as shown in Figure 10. Rectangular elements have the advantage that they have a regular structure, which will simplify programming requirements, but they generally cannot be locally refined. Triangular elements have the advantage that they can more easily fit complex shapes and can be locally refined, but they generally require unstructured mesh logic which can make the programming more complex.

Elements are also characterized by the order of the polynomial that is defined within them. If an element with typical size h contains a complete polynomial of order M, then the error will be at most of order h^{M+1}. This follows directly from a local Taylor series expansion:

$$\phi(x,y) = \sum_{k=0}^{M} \sum_{l=0}^{k} \frac{1}{l!(k-l)!} \left[\frac{\partial^k \phi}{\partial x^k \partial y^{k-1}} \right]_{x_0,y_0} (x-x_0)^l (y-y_0)^{k-l} + \mathcal{O}(h^{M+1}).$$

Linear elements will therefore have an error $\mathcal{O}(h^2)$, quadratic elements will have an error $\mathcal{O}(h^3)$, etc.

Another property that characterizes elements is their interelement continuity. As discussed in Section 3.1.2, a finite element with continuity C^{q-1} belongs to Hilbert space H^q, and hence can be used for differential operators with order up to $2q$. This applicability is made possible using the corresponding variational statement of the problem (Ritz method) or by performing integration by parts in the Galerkin method, and thereby shifting derivatives from the unknown to the trial function.

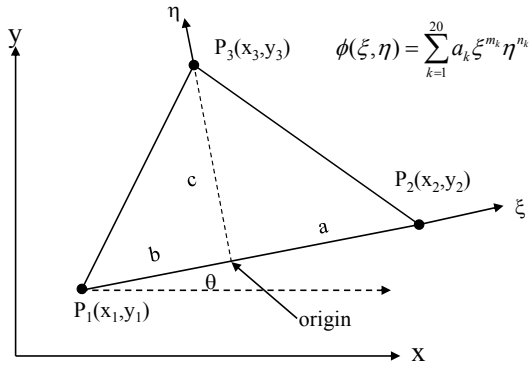

Figure 11: Reduced quintic finite element Q_{18} is defined by the four geometric parameters a, b, c, θ. A local (ξ, η) Cartesian system is used. The function and first two derivatives are constrained at the three vertex points, and C^1 continuity is imposed at the edges. Exponents m_i and n_i are given in Table 1.

3.2.1 Triangular Elements with C^1 Continuity

A particularly useful triangular finite element in two dimensions with C^1 continuity is known as the *reduced quintic* [10, 43, 44, 45], or Q_{18}, and is depicted in Figure 11. In each triangular element, the unknown function $\phi(x, y)$ is written as a general polynomial of 5th degree in the local Cartesian coordinates ξ and η: $\phi(\xi, \eta) = \sum_{i=1}^{20} a_i \xi^{m_i} \eta^{n_i}$ (where the exponents m_i and n_i are given in Table 1), which would have 21 coefficients were not there additional constraints. Eighteen of the coefficients are determined from specifying the values of the function and its first five derivatives: $\phi, \phi_x, \phi_y, \phi_{xx}, \phi_{xy}, \phi_{yy}$ at each of the three vertices, thus guaranteeing that globally all first and second derivatives will be continuous at each vertex. Since the one-dimensional quintic polynomial along each edge is completely determined by these values specified at the endpoints, it is guaranteed that the expansion is continuous between elements.

The remaining three constraints come from the requirement that the normal derivative of ϕ at each edge, ϕ_n, reduce to a one-dimensional cubic polynomial along that edge. This implies that the two sets of nodal values completely determine ϕ_n everywhere on each edge, guaranteeing its continuity from one triangle to the next so that the element is C^1. One of these three constraints is trivial and has been used to reduce the number of terms from 21 to 20 in the sum.

In imposing these continuity constraints, the expansion is no longer a complete quintic, but it does contain a complete quartic with additional constrained quintic coefficients to enforce C^1 continuity between elements. Thus, the name "reduced quintic." If the characteristic size of the element is h, then it follows from a local Taylor's series analysis that the approximation error in the

Table 1: Exponents of ξ and η for the reduced quintic expansion $\phi(\xi,\eta) = \sum_{i=1}^{20} a_i \xi^{m_i} \eta^{n_i}$.

k	m_k	n_k	k	m_k	n_k	k	m_k	n_k	k	m_k	n_k
1	0	0	6	0	2	11	4	0	16	5	0
2	1	0	7	3	0	12	3	1	17	3	2
3	0	1	8	2	1	13	2	2	18	2	3
4	2	0	9	1	2	14	1	3	19	1	4
5	1	1	10	0	3	15	0	4	20	0	5

unknown function, $\phi - \phi^h$, will be of order h^5.

Another advantage of this element is that all of the unknowns (or degrees of freedom, DOF) appear at the vertices, and are thus shared with all the triangular elements that connect to this vertex. This leads to a very compact representation with relatively small matrices compared to other representations. Suppose that we are approximating a square domain by partitioning it into n^2 squares or $2n^2$ triangles. The reduced quintic will asymptotically have $N = 6n^2$ unknowns, or three unknowns for each triangle. This scaling can be verified by the fact that if we introduce a new point into any triangle and connect it to the three nearby points, we will have generated two new triangles and introduced six new unknowns. We contrast this with a Lagrange quartic element which has the same formal order of accuracy but asymptotically has eight unknowns per triangle [42].

For a given triangle, if we locally number the unknowns ϕ, ϕ_x, ϕ_y, ϕ_{xx}, ϕ_{xy}, ϕ_{yy} at vertex P_1 as $\Phi_1 - \Phi_6$, at P_2 as $\Phi_7 - \Phi_{12}$, and at P_3 as $\Phi_{13} - \Phi_{18}$, then we seek a relation between the polynomial coefficients a_i and the Φ_j. This is done in two parts. We first define the 20×20 matrix in each triangle that relates the 18 local derivatives and two constraints to the 20 polynomial coefficients. We call this matrix $\mathbf{T}$. The 20 rows of $\mathbf{T}$ are given by:

$$
\begin{aligned}
\phi^1 &= a_1 - b a_2 + b^2 a_4 - b^3 a_7 + b^4 a_{11} - b^5 a_{16} \\
\phi^1_\xi &= a_2 - 2b a_4 + 3b^2 a_7 - 4b^3 a_{11} + 5b^4 a_{16} \\
\phi^1_\eta &= a_3 - b a_5 + b^2 a_8 - b^3 a_{12} \\
\phi^1_{\xi\xi} &= 2a_4 - 6b a_7 + 12 b^2 a_{11} - 20 b^3 a_{16} \\
\phi^1_{\xi\eta} &= a_5 - 2b a_8 + 3b^2 a_{12} \\
\phi^1_{\eta\eta} &= 2a_6 - 2b a_9 + 2b^2 a_{13} - 2b^3 a_{17}
\end{aligned}
$$

$$
\begin{aligned}
\phi^2 &= a_1 + aa_2 + a^2a_4 + a^3a_7 + a^4a_{11} + a^5a_{16} \\
\phi^2_\xi &= a_2 + 2aa_4 + 3a^2a_7 + 4a^3a_{11} + 5a^4a_{16} \\
\phi^2_\eta &= a_3 + aa_5 + a^2a_8 + a^3a_{12} \\
\phi^2_{\xi\xi} &= 2a_4 + 6aa_7 + 12a^2a_{11} + 20a^3a_{16} \\
\phi^2_{\xi\eta} &= a_5 + 2aa_8 + 3a^2a_{12} \\
\phi^2_{\eta\eta} &= 2a_6 + 2aa_9 + 2a^2a_{13} + 2a^3a_{17}
\end{aligned}
$$

$$
\begin{aligned}
\phi^3 &= a_1 + ca_3 + c^2a_6 + c^3a_{10} + c^4a_{15} + c^5a_{20} \\
\phi^3_\xi &= a_2 + ca_5 + c^2a_9 + c^3_{14} + c^4a_{19} \\
\phi^3_\eta &= a_3 + 2ca_6 + 3c^2a_{10} + 4c^3a_{15} + 5c^4a_{20} \\
\phi^3_{\xi\xi} &= 2a_4 + 2ca_8 + 2c^2a_{13} + 2c^3a_{18} \\
\phi^3_{\xi\eta} &= a_5 + 2ca_9 + 3c^2a_{14} + 4c^3a_{19} \\
\phi^3_{\eta\eta} &= 2c_6 + 6ca_{10} + 12c^2a_{15} + 20c^3a_{20}
\end{aligned}
$$

$$
\begin{aligned}
0 &= 5b^4ca_{16} + (3b^2c^3 - 2b^4c)a_{17} + (2bc^4 - 3b^3c^2)a_{18} \\
&\quad +(c^5 - 4b^2c^3)a_{19} - 5bc^4a_{20} \\
0 &= 5a^4ca_{16} + (3a^2c^3 - 2a^4c)a_{17} + (-2ac^4 - 3a^3c^2)a_{18} \\
&\quad +(c^5 - 4a^2c^3)a_{19} - 5ac^4a_{20}.
\end{aligned}
$$

This satisfies $\mathbf{\Phi'} = \mathbf{T}\,\mathbf{A}$, where $\mathbf{\Phi'}$ denotes the vector of length 20 produced by stringing together the 18 values of the function and derivatives with respect to the local Cartesian coordinates ξ and η at the three vertices, and with the final two elements zero, and $\mathbf{A}$ is the vector produced by the 20 polynomial coefficients. This can be solved for the coefficient matrix by inverting $\mathbf{T}$, thus $\mathbf{A} = \mathbf{T}^{-1}\mathbf{\Phi'}$. A useful check is to verify that the numerically evaluated determinant of $\mathbf{T}$ has the value $-64(a+b)^{17}c^{20}(a^2+c^2)(b^2+c^2)$. Since the final two elements of $\mathbf{\Phi'}$ are zero, we can replace $\mathbf{T}^{-1}$ by the 20×18 matrix $\mathbf{T}_2$ which consists of the first 18 columns of $\mathbf{T}^{-1}$.

To get the coefficient matrix $\mathbf{A}$ in terms of the vector containing the actual derivatives with respect to (x, y), we have to apply the rotation matrix $\mathbf{R}$. This is compactly defined in terms of the angle θ appearing in Figure 11 by

$$
\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & & \\ & \mathbf{R}_1 & \\ & & \mathbf{R}_1 \end{bmatrix}, \tag{3.16}
$$

where

$$
\mathbf{R}_1 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & \cos\theta & \sin\theta & 0 & 0 & 0 \\
0 & -\sin\theta & \cos\theta & 0 & 0 & 0 \\
0 & 0 & 0 & \cos^2\theta & 2\sin\theta\cos\theta & \sin^2\theta \\
0 & 0 & 0 & -\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta & \sin\theta\cos\theta \\
0 & 0 & 0 & \sin^2\theta & -2\sin\theta\cos\theta & \cos^2\theta
\end{bmatrix}.
$$

If we then define the matrix $\mathbf{G} = \mathbf{T}_2\mathbf{R}$, this relates the coefficient matrix directly to the unknown vector consisting of the function and derivatives with respect to (x, y), i.e., $\mathbf{A} = \mathbf{G}\Phi$, or in component notation: $a_i = \sum_{j=1}^{18} g_{i,j}\Phi_j$ for $i = 1, 20$. The 20×18 matrix $g_{i,j}$ depends only on the shape and orientation of the individual triangle, and in general will be different for each triangle. The general expression for the unknown function ϕ in a given triangle is

$$\begin{aligned} \phi(\xi,\eta) &= \sum_{i=1}^{20} a_i \xi^{m_i}\eta^{n_i} = \sum_{i=1}^{20}\sum_{j=1}^{20} g_{i,j}\phi_j \xi^{m_i}\eta^{n_i}, \\ &= \sum_{j=1}^{18} \nu_j \phi_j, \end{aligned} \tag{3.17}$$

where we have defined the basis functions as

$$\nu_j \equiv \sum_{i=1}^{20} g_{i,j}\xi^{m_i}\eta^{n_i}; \qquad\qquad j = 1, 18. \tag{3.18}$$

The 18 basis functions for each triangle, as defined by Eq. (3.18), have the property that they have a unit value for either the function or one of its first or second derivatives at one vertex and and zero for the other quantities at this and the other nodes. They also have the C^1 property embedded.

All of the integrals that need to be done to define the matrices that occur in the Galerkin method are of the form of 2D integrals of polynomials in ξ or η over the triangles. These can either be evaluated by numerical integration [46], or by making use of the analytic formula:

$$F(m,n) \equiv \int\int_{triangle} \xi^m \eta^n d\xi d\eta = c^{n+1} \frac{\left[a^{m+1} - (-b)^{m+1}\right] m!n!}{(m+n+2)!}. \tag{3.19}$$

For example, to evaluate the mass matrix, we would have

$$\begin{aligned} \int\int \nu_i(\xi,\eta)\phi(\xi,\eta)d\xi d\eta &= \sum_{k=1}^{18}\left[\sum_{i=1}^{20}\sum_{l=1}^{20} g_{i,j}g_{l,k}F(m_i+m_l, n_i+n_l)\right]\Phi_k \\ &\equiv \sum_{k=1}^{18} M_{jk}\Phi_k. \end{aligned} \tag{3.20}$$

The implementation of boundary conditions requires some discussion [47]. Recall that for each scalar variable being solved for, there are six unknowns at each node corresponding to the function and all its first and second derivatives. If Dirichlet or Neumann boundary conditions are being applied at a boundary that is aligned with the $\hat{x}$ or $\hat{y}$ axis, the imposition of boundary conditions is straightforward. Consider homogeneous Dirichlet boundary conditions being applied at a boundary node where the boundary lies on the x axis. We then replace the rows of the matrix that correspond to the trial functions ν_i having

non-zero values of ϕ, ϕ_x and ϕ_{xx} with a row from the identity matrix of all zeros but a one on the diagonal, and with zero on the right side of the equation. This is equivalent to removing these basis functions from the system. If Neumann conditions were being applied, we would similarly replace the rows corresponding to ϕ_y and ϕ_{xy}. If we define the operator

$$\mathcal{L}\{\phi\} = [\phi, \phi_x, \phi_y, \phi_{xx}, \phi_{xy}, \phi_{yy}], \tag{3.21}$$

then these conditions are all of the form $\mathcal{L}_i\{\phi\} = C$. This is easily implemented because the basis functions ν_j at the boundary point satisfy the orthogonality conditions

$$\mathcal{L}_i\{\nu_j\} = \delta_{ij}. \tag{3.22}$$

But what if the boundaries are not aligned with the x or y axis? In this case, if we define the vector corresponding to normal and tangential derivatives of the solution as

$$\mathcal{L}'\{\phi\} = [\phi, \phi_n, \phi_t, \phi_{nn}, \phi_{nt}, \phi_{tt}], \tag{3.23}$$

then the boundary conditions we need to impose are of the form $\mathcal{L}'_i\{\phi\} = C$. Here n is the coordinate locally normal to the boundary, and t is locally tangent to the boundary. We also define the local curvature as $\kappa = \hat{n} \cdot d\hat{t}/ds$ where ds is the arc length along the boundary curve. The derivatives of ϕ with respect to the normal and tangential coordinates (n, t) are related to those with respect to the global coordinates (x, y) at a given boundary point by the transformation

$$\mathcal{L}'_i\{\phi\} = M_{ij}\mathcal{L}_j\{\phi\}, \tag{3.24}$$

where M is the unimodular matrix

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & n_x & n_y & 0 & 0 & 0 \\ 0 & -n_y & n_x & 0 & 0 & 0 \\ 0 & 0 & 0 & n_x^2 & 2n_xn_y & n_y^2 \\ 0 & -\kappa n_y & \kappa n_x & -n_xn_y & n_x^2 - n_y^2 & n_xn_y \\ 0 & -\kappa n_x & -\kappa n_y & n_y^2 & -2n_xn_y & n_x^2 \end{bmatrix}. \tag{3.25}$$

In order to impose a boundary condition on a curved boundary (or even a straight boundary not aligned with the $\hat{x}$ or $\hat{y}$ axis) and for the residual to be constrained as much as possible, we want to define a set of trial functions μ_i at each boundary point that satisfy a similar orthogonality condition to that in the axis-aligned boundary case, i.e.,

$$\mathcal{L}'_i\{\mu_j\} = \delta_{ij}. \tag{3.26}$$

The new trial functions must be a linear combination of the old basis functions since the latter span the solution space, and thus we can write for some matrix $\mathbf{N}$,

$$\mu_i = N_{ij}\nu_j. \tag{3.27}$$

It follows from Eqs. (3.22), (3.24), (3.26), and (3.27) that $N_{ij} = (M_{ji})^{-1}$,

$$\mathbf{N} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & n_x & n_y & \kappa n_y^2 & -\kappa n_x n_y & \kappa n_x^2 \\ 0 & -n_y & n_x & 2\kappa n_x n_y & -\kappa(n_x^2 - n_y^2) & -2\kappa n_x n_y \\ 0 & 0 & 0 & n_x^2 & n_x n_y & n_y^2 \\ 0 & 0 & 0 & -2n_x n_y & n_x^2 - n_y^2 & 2n_x n_y \\ 0 & 0 & 0 & n_y^2 & -n_x n_y & n_x^2 \end{bmatrix}. \tag{3.28}$$

For each point that lies on the boundary, we therefore substitute the six trial functions μ_i for the original six trial functions ν_i. These boundary vertex trial functions obey an orthogonality condition on their derivatives with respect to a local coordinate system (n, t) defined relative to the local boundary orientation. We then substitute the appropriate rows for rows imposing the boundary conditions as in the case of the coordinate aligned boundaries. Using these linearly transformed trial functions to impose the boundary conditions is optimal in the sense that the components of the residual associated with the boundary vertex are maximally constrained.

4 Implicit Methods and the M3D-C^1 Approach

In this lecture we discuss implicit methods for solving predominantly hyperbolic systems of equations, and a particular method that is well suited for the MHD equations being applied to tokamak plasmas. If we denote by $\mathbf{U}$ the vector of unknowns, then a conservative system of equations can be written (in one dimension) either as

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0, \tag{4.1}$$

or, in the non-conservative form

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A} \cdot \frac{\partial \mathbf{U}}{\partial x} = 0. \tag{4.2}$$

The matrix $\mathbf{A} = \partial \mathbf{F}/\partial \mathbf{U}$ has real eigenvalues for a pure hyperbolic system. Since the stability analysis we perform is linear, the stability properties of the same method as applied to either Eq. (4.1) or Eq. (4.2) will be identical. Systems of linear equations will be of the form Eq. (4.2) from the outset.

If all the eigenvalues of the matrix $\mathbf{A}, \lambda_A$, are close together in magnitude, it is generally not worthwhile to use an implicit method for a purely hyperbolic problem. The reasons for this are as follows: It requires more computational work per time step using an implicit method compared to that required for an explicit method. This difference becomes greater as the rank of the matrix $\mathbf{A}$ increases or especially when working in two or higher dimensions. Also, to take advantage of the unconditional stability of the implicit method, we would need to significantly violate the CFL condition so that $\delta t \gg \delta x/|\lambda_A|$. Since the truncation error is normally of the same order in both δt and δx, this would imply that the $\mathcal{O}(\delta t^2)$ term would dominate in the truncation error and thus solutions using larger time steps would be less accurate than those with smaller time steps. This is in contrast to a Crank–Nicolson implicit solution of a diffusion equation where the time step can be increased from $\delta t \sim \delta x^2$ to $\delta t \sim \delta x$ without substantially increasing the overall truncation error.

Another possible disadvantage of implicit methods applied to hyperbolic equations is that they result in an infinite signal propagation speed. Physically, all information is propagated along the characteristics, which corresponds to a finite velocity of propagation. An explicit finite difference method normally has a domain of dependence that very closely resembles the domain of dependence of the physical system, while the domain of dependence of the implicit method is totally different.

Explicit finite difference methods are also ideally suited to pure hyperbolic problems with with a wide range of characteristic velocities but where interest lies in phenomena associated with the fastest wave speed (or equivalently, the largest eigenvalue λ_A). However, in a strongly magnetized plasma we normally have the situation where there are multiple characteristic velocities that differ

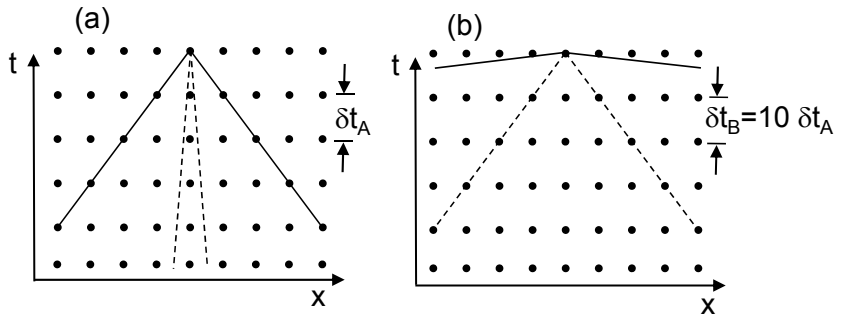

Figure 12: Characteristics for the Alfvén wave (dotted) and the fast magnetoacoustic wave (solid) are shown in this schematic diagram. (a) Space-time plot for fully explicit method. (b) Same for partially implicit method where the time step is at CFL limit for the Alfvén wave.

widely, and where primary interest lies not with the fastest wave speed (which is the fast magnetoacoustic wave which compresses the strong background field), but rather with the slower Alfvén wave, which is nearly incompressible [49]. The situation is then as depicted in Figure 12, where we draw the characteristics associated with the fast wave as a solid line, and those associated with the slower shear Alfvén wave with a dotted line.

Figure 12a illustrates the space-time points in an explicit calculation where the time step is determined by the CFL condition for the fast wave. It is seen that it takes many time steps for the shear Alfvén wave to traverse one zone spacing. In contrast, Figure 12b shows the same characteristics but with a time step chosen so that the CFL condition is just satisfied for the shear Alfvén wave. It is seen that the CFL condition for the fast wave is strongly violated and so an implicit treatment is needed, at least for the eigencomponents associated with the fast wave.

There are also many situations where there are additional terms in the equations, such as resistivity, that lead to even longer time scales that need to be followed, for example the slow growth of a resistive instability in a fusion device [14, 15, 16]. In these cases, the CFL time step restriction based on the fast wave would be much too restrictive and some form of implicit solution is required. The next two sections describe some different approaches to solving the implicit system in multiple dimensions. These methods are aimed at producing efficient algorithms for studying plasma motion that is associated with other than the fastest characteristic.

4.0.2 θ-Implicit Method

We first consider an implicit method for the hyperbolic system of Eq. (4.2) that is similar to the θ-implicit and Crank–Nicolson methods for parabolic equations,

$$\frac{\mathbf{U}_j^{n+1} - \mathbf{U}_j^n}{\delta t} + \mathbf{A} \cdot \left[\theta \left(\frac{\mathbf{U}_{j+1}^{n+1} - \mathbf{U}_{j-1}^{n+1}}{2\delta x} \right) + (1-\theta) \left(\frac{\mathbf{U}_{j+1}^{n} - \mathbf{U}_{j-1}^{n}}{2\delta x} \right) \right], \quad (4.3)$$

where $0 \leq \theta \leq 1$ is the implicit parameter. Von Neumann analysis yields for the amplification factor

$$r = \frac{1 + is(\theta - 1)}{1 + is\theta}, \quad (4.4)$$

where $s = (\delta t \lambda_A / \delta x) \sin\theta_k$, with λ_A being each of the eigenvalues of the matrix $\mathbf{A}$. We see that $|r| \leq 1$, implying unconditional stability, for $0.5 \leq \theta \leq 1.0$. For $\theta = 0.5$ we have $|r| = 1$, implying no dissipation, and the method described by Eq. (4.3) is centered in both space and time and therefore has a truncation error $T_\Delta = \mathcal{O}(\delta x^2, \delta t^2)$. In this sense, it is analogous to the Crank–Nicolson method for parabolic equations.

The θ-implicit method as described by Eq. (4.3) uses the quasi-linear form of conservation equations, and does not deal with any non-linearity present in $\mathbf{A}$. If $\mathbf{A}$ is a strongly non-linear function of $\mathbf{U}$, or if we wish to solve the problem in conservation form, the method can be modified in several ways. One obvious way is to update the solution to a provisional time, $(n+1)^*$, to evaluate the matrix $\mathbf{A}$ at this time, and then repeat the update of the solution from time n to $(n+1)$ but using some linear combination of $\mathbf{A}$ at time n and $(n+1)^*$.

A straightforward linearization technique can also be used to apply the θ-implicit method to the conservative formulation of a hyperbolic system as given by Eq. (4.1) in one dimension [50]. We first Taylor expand $\mathbf{F}^{n+1}$ in time about $\mathbf{F}^n$

$$\mathbf{F}^{n+1} = \mathbf{F}^n + \mathbf{A}^n \cdot \left(\mathbf{U}^{n+1} - \mathbf{U}^n \right) + \mathcal{O}(\delta t^2). \quad (4.5)$$

Applying the θ-implicit method to the system of conservation equations, Eq. (4.1), and rearranging terms gives

$$\left[\mathbf{I} + \theta \delta t \frac{\partial}{\partial x} \mathbf{A}^n \right] \cdot \mathbf{U}^{n+1} = \left[\mathbf{I} + \theta \delta t \frac{\partial}{\partial x} \mathbf{A}^n \right] \cdot \mathbf{U}^n - \delta t \left. \frac{\partial \mathbf{F}}{\partial x} \right|^n. \quad (4.6)$$

Note that the derivative is acting on $\mathbf{A} \cdot \mathbf{U}$ for both bracketed terms. If $\mathbf{A}$ is a constant matrix, Eq. (4.6) is equivalent to Eq. (4.3) if centered spatial difference operators are used. For $\theta = 1/2$, the method is seen to be time and space centered and thus have a truncation error $\mathcal{O}(\delta x^2, \delta t^2)$.

4.0.3 Method of Differential Approximation

Here we describe an approach for solving stiff hyperbolic systems that is based on the method of differential approximations [51] and is now used in several

contemporary multidimensional MHD codes [10, 52]. We start by considering the model second-order hyperbolic system in one dimension given by:

$$\frac{\partial v}{\partial t} = c\frac{\partial \omega}{\partial x}, \tag{4.7}$$

$$\frac{\partial \omega}{\partial t} = c\frac{\partial v}{\partial x}. \tag{4.8}$$

Here, the wave speed, c, is assumed to be a constant for simplicity. These can be combined to give the single second-order equation,

$$\frac{\partial^2 v}{\partial t^2} = \mathcal{L}(v), \tag{4.9}$$

where we have introduced the linear second-order spatial operator,

$$\mathcal{L}(v) \equiv c^2 \frac{\partial^2}{\partial x^2} v. \tag{4.10}$$

Consider now the following algorithm for advancing the system in time:

$$\left(1 - \theta^2 \delta t^2 \mathcal{L}\right) v^{n+1} = \left(1 - \alpha \delta t^2 \mathcal{L}\right) v^n + \delta t c \left(\frac{\partial \omega}{\partial x}\right)^m \tag{4.11}$$

$$\omega^{m+1} = \omega^m + \delta t c \phi \left(\frac{\partial v}{\partial x}\right)^{n+1} + \delta t c (1-\phi) \left(\frac{\partial v}{\partial x}\right)^n. \tag{4.12}$$

We have introduced the implicit parameters θ, α, and ϕ and the time step δt. The derivatives are to be replaced by centered finite difference (or spectral or finite element) operators. For $\theta = \alpha = 0$ and $\phi = 1$, the algorithm described by Eqs. (4.11) and (4.12) is just the explicit leapfrog method.

To examine the numerical stability of the general case, we make the substitutions

$$\frac{\partial}{\partial x} \to -ik_e,$$
$$\mathcal{L} \to -c^2 k_e^2,$$

where $k_e = -(1/\delta x)\sin\theta_k$ is the effective wavenumber when the centered finite difference operator is substituted for the spatial derivative. Letting $v^{n+1} = rv^n$ and $\omega^{m+1} = r\omega^m$, the amplification factor r can then be determined from the quadratic equation

$$(1 + \theta^2 D)(r-1)^2 + D(\theta^2 + \phi - \alpha)(r-1) + D = 0, \tag{4.13}$$

where we have defined $D \equiv \delta t^2 c^2 k_e^2$. Equation (4.13) has roots

$$r = \frac{1 + \frac{1}{2}D\left(\theta^2 - \phi + \alpha\right) \pm i\sqrt{D + \left[\theta^2 - \frac{1}{4}\left(\theta^2 + \phi - \alpha\right)^2\right] D^2}}{1 + \theta^2 D}, \tag{4.14}$$

for which

$$|r|^2 = \frac{1 + (1 + \alpha - \phi)D}{1 + \theta^2 D} \tag{4.15}$$

when the quantity within the square root in Eq. (4.16) is non-negative. There are two cases that are of interest.

For the *Caramana method* [51], we set $\phi = 1$ and $\alpha = \theta^2$. From Eq. (4.15), this yields $|r|^2 = 1$ for any D, as long as $\theta \geq \frac{1}{2}$. Thus the method is linearly stable and non-dissipative for $\theta \geq \frac{1}{2}$. Truncation error analysis shows the time discretization error to be second-order in δt for any stable value of θ. This method has the additional feature that the multiplier of the operator $\mathcal{L}$ is the same on both sides of Eq. (4.12), so that in steady state, when $v^{n+1} = v^n$, the operator will have no affect on the solution.

For the *split θ-implicit method*, we set $\phi = \theta$ and $\alpha = \theta(\theta - 1)$. This is what one would obtain if both Eqs. (4.7) and (4.8) were time differenced using the θ-implicit method, and then the differenced form of Eq. (4.8) was used to algebraically eliminate ω^{n+1} from the differenced form of Eq. (4.7). The amplification factor is $|r|^2 = 1 + (1 - 2\theta)D/(1 + \theta^2 D)$, which is less than or equal to 1 (and hence stable) when $\theta \geq \frac{1}{2}$. The quantity within the square root in Eq. (4.14) is exactly D for this method, which is assumed to be positive. Note that $|r|^2 = 1$ only when $\theta = \frac{1}{2}$ for this method. The truncation error analysis for this method is exactly the same as it would be for the θ-implicit method, yielding second-order accuracy only for $\theta = \frac{1}{2}$.

In each of these algorithms, the operator $\left(1 - \theta^2 \delta t^2 \mathcal{L}\right)$ needs to be inverted each time step. However, this is a well-conditioned symmetric diagonally dominant operator, and so iterative methods should perform well on this. Also, we see that the two equations, Eqs. (4.11) and (4.12), can be solved sequentially, and so the associated sparse matrix equation is only for a single scalar variable. Both of these features offer clear advantages over the unsplit θ-implicit method discussed in Section 4.0.2.

To obtain the operator $\mathcal{L}$ for the ideal MHD equations, we start with the momentum equation, again ignoring for simplicity the density and convective derivative terms, assumed small. Denoting time derivatives with a dot, i.e., $\dot{\mathbf{u}} \equiv \partial \mathbf{u}/\partial t$, we have

$$\rho_0 \dot{\mathbf{u}} + \nabla p = \frac{1}{\mu_0} \left[(\nabla \times \mathbf{B}) \times \mathbf{B} \right]. \tag{4.16}$$

Time differentiating gives

$$\rho_0 \ddot{\mathbf{u}} + \nabla \dot{p} = \frac{1}{\mu_0} \left[\left(\nabla \times \dot{\mathbf{B}} \right) \times \mathbf{B} + (\nabla \times \mathbf{B}) \times \dot{\mathbf{B}} \right]. \tag{4.17}$$

Next, we take the ideal MHD components of the magnetic field and pressure equations:

$$\dot{\mathbf{B}} = \nabla \times [\mathbf{u} \times \mathbf{B}], \tag{4.18}$$

$$\dot{p} = -\mathbf{u} \cdot \nabla p - \gamma p \nabla \cdot \mathbf{u}. \tag{4.19}$$

We substitute for $\dot{\mathbf{B}}$ and $\dot{p}$ from Eqs. (4.18) and (4.19) into Eq. (4.16), whereby in analogy with Eq. (4.9), we can identify

$$\begin{aligned}\mathcal{L}(\mathbf{u}) &\equiv \frac{1}{\mu_0}\{\nabla\times[\nabla\times(\mathbf{u}\times\mathbf{B})]\}\times\mathbf{B} \\ &+ \frac{1}{\mu_0}\{(\nabla\times\mathbf{B})\times\nabla\times(\mathbf{u}\times\mathbf{B})\} \\ &+ \nabla(\mathbf{u}\cdot\nabla p+\gamma p\nabla\cdot\mathbf{u}). \end{aligned} \tag{4.20}$$

We note that this is equivalent to the ideal MHD operator introduced in Bernstein, et. al. [53].

Applying the Caramana method to this equation gives the following equation to advance the velocity to the new time level

$$\begin{aligned}\left\{\rho_0-\theta^2\delta t^2\mathcal{L}\right\}\mathbf{u}^{n+1} &= \left\{\rho_0-\theta^2\delta t^2\ \mathcal{L}\right\}\mathbf{u}^n \\ &+ \delta t\left\{-\nabla p+\frac{1}{\mu_0}[(\nabla\times\mathbf{B})\times\mathbf{B}]\right\}^{n+\frac{1}{2}}. \end{aligned} \tag{4.21}$$

When finite difference or finite element methods are used to discretize the spatial operators in Eq. (4.21), it becomes a sparse matrix equation for the advanced time velocity $\mathbf{u}^{n+1}$. When this sparse matrix equation is solved, the new velocity can be used to advance the magnetic field and pressure according to Eqs. (4.18) and (4.19), which now take the form:

$$\dot{\mathbf{B}} = \nabla\times\left[\left(\theta\mathbf{u}^{n+1}+(1-\theta)\mathbf{u}^n\right)\times\mathbf{B}\right], \tag{4.22}$$

$$\begin{aligned}\dot{p} &= -\left(\theta\mathbf{u}^{n+1}+(1-\theta)\mathbf{u}^n\right)\cdot\nabla p \\ &\quad -\gamma p\nabla\cdot\left(\theta\mathbf{u}^{n+1}+(1-\theta)\mathbf{u}^n\right). \end{aligned} \tag{4.23}$$

A von Neumann stability analysis shows this method to be linearly stable for all δt as long as $\theta\geq 1/2$. We also note that it gives accurate solutions in steady state when the terms in Eq. (4.21) on the two sides of the equation which involve the operator $\mathcal{L}$ cancel.

Non-ideal terms such as resistivity and viscosity can be added to Eqs. (4.21)–(4.23) by treating them implicitly and this will not affect the form of the operator in Eq. (4.20) or the numerical stability. The challenge in this method is to find an efficient method for inverting the matrix corresponding to the operator on the left in Eq. (4.21). This is presently an area of active research.

4.0.4 The Vector Fields

In describing unstable motions of a low $\beta\equiv 2\mu_0 p/B^2$ plasma with a strong background magnetic field, such as a tokamak, it is essential to employ high-accuracy spatial representations. One approach [52] is to represent the velocity and magnetic fields by their cylindrical coordinate projections and use high-order finite elements to accurately compute the product terms. Another approach [48] is to represent the velocity and magnetic fields in terms of stream functions and

potentials that exhibit analytic cancelations in the asymptotic large-guide-field limit.

Consider the following forms for the velocity field and magnetic vector potential:

$$\mathbf{u} = R^2 \nabla U \times \nabla\phi + \omega R^2 \nabla\phi + R^{-2} \nabla_\perp \chi, \tag{4.24}$$

$$\mathbf{A} = R^2 \nabla\phi \times \nabla f + \psi \nabla\phi - F_0 \ln R\hat{Z}. \tag{4.25}$$

Here (R, ϕ, Z) form a cylindrical coordinate system, the subscript $\perp$ indicates "perpendicular to $\nabla\phi$", i.e., in the (R, Z) plane, and F_0 is a constant which represents an externally imposed toroidal magnetic field. The velocity is thus determined by the three scalar functions (U, ω, χ). The magnetic vector potential is defined by the two scalar functions (f, ψ), the constant F_0, and the gauge condition $\nabla_\perp \cdot R^{-2}\mathbf{A} = 0$.

The magnetic field is given by $\mathbf{B} = \nabla \times \mathbf{A}$, or

$$\mathbf{B} = \nabla\psi \times \nabla\phi - \nabla_\perp \frac{\partial f}{\partial \phi} + F\nabla\phi. \tag{4.26}$$

$$= \nabla\psi \times \nabla\phi - \nabla \frac{\partial f}{\partial \phi} + F^*\nabla\phi. \tag{4.27}$$

Here,we have defined $F \equiv F_0 + R^2\nabla_\perp^2 f$ and $F^* \equiv F_0 + R^2\nabla^2 f$. The first term in the velocity representation, Eq. (4.24), represents motion in a 2D toroidal plane that does not compress the background toroidal field, the second term represents the motion in the toroidal direction, and the third term, which is normally very small, represents compressible motion in the toroidal plane. Note that it is orthogonal to the first term in that the inner product of those two terms vanishes when integrated over the torus.

To see that the first velocity term does not compress the background toroidal field, we can substitute $\mathbf{u} = R^2\nabla U \times \nabla\phi$ and $\mathbf{B} = F_0\nabla\phi$ into the ideal-MHD field evolution equation, Eq. (1.38), and compute the toroidal component:

$$\begin{aligned} \nabla\phi \cdot \frac{\partial \mathbf{B}}{\partial t} &= \nabla\phi \cdot \nabla \times [\mathbf{u} \times \mathbf{B}] \\ &= \nabla\phi \cdot \nabla \times \left[\left(R^2\nabla U \times \nabla\phi\right) \times (F_0\nabla\phi)\right] \\ &= F_0 \nabla \cdot [\nabla\phi \times \nabla U] \\ &= 0 \end{aligned} \tag{4.28}$$

We are thus able to eliminate this possibly large error term associated with anomalous compression of the externally imposed toroidal field analytically from the equations, resulting in a large increase in accuracy. The velocity field associated with this first term always dominates over that associated with the third term in Eq. (4.24).

4.0.5 Projections

When using a velocity representation such as that in Eq. (4.24), it is also important what projections one takes of the momentum equation. When applying

the Galerkin finite element method, one operates on the momentum equation with a differential operator, multiplies by the finite element ν_i, and integrates over all space to obtain the weak form of the equation. Consider the following annihilation operators applied to the modified momentum equation, Eq. (4.21), and then integrated by parts:

$$\int d\tau \nu_i \nabla\phi \cdot \nabla_\perp \times R^2 (4.21) \quad \rightarrow \quad \int d\tau R^2 \nabla_\perp \nu_i \times \nabla\phi \cdot (4.21) \tag{4.29}$$

$$\int d\tau \nu_i R^2 \nabla\phi \cdot (4.21) \quad \rightarrow \quad \int d\tau \nu_i R^2 \nabla\phi \cdot (4.21) \tag{4.30}$$

$$-\int d\tau \nu_i \nabla_\perp \cdot R^{-2} (4.21) \quad \rightarrow \quad \int d\tau R^{-2} \nabla_\perp \nu_i \cdot (4.21) \tag{4.31}$$

The boundary terms from the integration by parts are assumed to vanish here. By comparing the integrands on the right in Eq. (4.29)-(4.31) with the form of the velocity in Eq. (4.24), we see that after the integration by parts, these projection operators are equivalent to taking the inner product of the momentum equation, Eq. (4.21), separately with each of the three terms in the velocity field, but with the trial function ν_i replacing each of the three scalar functions (U, ω, χ). This property leads to an energy-conserving set of discrete equations, to two energy-conserving subsets of reduced equations, and to self-adjoint energy terms, called *partial energy terms*, being introduced into the implicit time advance [48].

The partial energy terms come from the inner product of the velocity-like terms, on the right in Eqs. (4.29)-(4.31), with the ideal MHD operator $\mathcal{L}$ in Eq. (4.21), which is defined in Eq. (4.20). These add to the mass matrix, improving its condition number.

Similar projections and integrations are performed with the magnetic field advance equation, with the projection operators being given by:

$$\int d\tau \nu_i \nabla\phi \cdot \nabla_\perp \times (4.22) \quad \rightarrow \quad \int d\tau \nabla_\perp \nu_i \times \nabla\phi \cdot (4.22) \tag{4.32}$$

$$\int d\tau \nu_i \nabla\phi \cdot (4.22) \quad \rightarrow \quad \int d\tau \nu_i \nabla\phi \cdot (4.22) \tag{4.33}$$

As in the discussion following Eq. (4.31), if we compare the integrands on the right in Eqs. (4.32) and (4.33) with the form of the magnetic field in Eq. (4.27), we see that these projection operators are equivalent to taking the inner product of the magnetic field evolution equation, Eq. (4.22), with the first and third terms in the magnetic field, but with the trial function ν_i replacing the scalar quantities ψ and F^*. In this case, there is no need to take the third projection, which would be

$$-\int d\tau \nu_i \nabla_\perp \cdot (4.22) \quad \rightarrow \quad \int d\tau \nabla_\perp \nu_i \cdot (4.22), \tag{4.34}$$

since the divergence constraint on the magnetic field assures that this is satisfied.

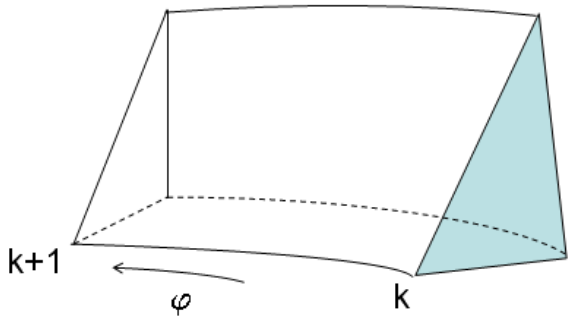

Figure 13: Integration volume is the triangular wedge between two toroidal planes.

4.0.6 An Appropriate Finite Element Basis

When the projection operators defined in Section 4.0.5 are applied to the modified momentum equation, Eq. (4.21), and the forms of the vector fields given by Eq. (4.24)-(4.27) are used, the resulting equations contain 4^{th} order spacial derivatives of the scalar variables. As discussed in Section 3.1.2, this requires the use of C^1 finite elements.

An appropriate set of 3D C^1 basis functions can be formed by taking tensor products of the $Q_{18}(R,Z)$ 2D basis functions defined on a triangle in Section 3.2.1 with the set of 1D Hermite cubic polynomial basis functions $\Phi_i(\phi)$ defined in Section 3.1.3. Thus, if we denote by $\nu_j(R,Z)$ the 2D basis functions defined at the nodes of the 2D triangles, we form 3D basis functions as follows:

$$E_{i,j}(R,\phi,Z) = \nu_j(R,Z)\Phi_i(\phi). \tag{4.35}$$

There are two Hermite basis functions associated with each toroidal plane $\phi = \phi_k$. They each extend over a single mesh spacing in ϕ in each direction, $[\phi_k - h, \phi_k + h]$. Since $\Phi_1(x)$ has zero derivative at local coordinate $x = 0$ and $\Phi_2(x)$ has zero value at $x = 0$, the basis function $E_{i,j}$ has the simple physical interpertation that $E_{1,j}$ is the value of the 3D function at that plane and $E_{2,j}$ is the value of the toroidal derivative of the 3D function at that plane. The 3D integration volume is the triangular wedge between two toroidal planes as depicted in Figure 13. In the volume bounded by planes k and $k+1$ and a given triangle, shown in the figure, we can express each scalar variable such as U in terms of 72 time-varying scalars:

$$\begin{aligned} U(R,\phi,Z) &= \sum_{j=1}^{18} \nu_j(R,Z)\left[U^1_{j,k}\Phi_1(\phi/h) + U^2_{j,k}\Phi_2(\phi/h)\right. \\ &+ \left. U^1_{j,k+1}\Phi_1(\phi/h-1) + U^2_{j,k+1}\Phi_2(\phi/h-1)\right] \end{aligned} \tag{4.36}$$

With these conventions, the standard Galarkin method is applied, turning the

differential equations into discrete matrix equations which are then solved using standard iterative methods [3].

A computer code based on these methods has recently been developed [48] and early applications show excellent accuracy and efficiency properties [54] for linear calculations. Extension to fully nonlinear calculations is in progress.

References

[1] Freidberg, J. P. 2008. *Plasma Physics and Fusion Energy* Cambridge: Cambridge University Press.

[2] Wesson, J. 2004. *Tokamaks.* USA: Oxford University Press.

[3] Jardin, S. C. 2010. *Computational Methods in Plasma Physics* Boca Raton, FL: Taylor and Francis.

[4] Braginski, S. I. 1966. Transport processes in a plasma. In *Review of Plasma Physics, Vol. I*, ed. M.A. Leontovich, 205-311. New York: Consultants Bureau.

[5] Ramos, J. J. 2005. General expression of the gyroviscous force. *Phys. Plasmas* **12**:112301.

[6] Simakov, A. N. and P. J. Catto. 2005. Evaluation of the neoclassical radial electric field in a collisional tokamak. *Phys. Plasmas* **12**:012105.

[7] Chang, Z. and J. D. Callen. 1992. Generalized gyroviscous force and its effect on the momentum balance equation. *Phys. Fluids B* **4**:1766-71.

[8] Ferraro, N. and S. C. Jardin. 2006. Finite element implementation of Braginskii's gyroviscous stress with application to the gravitational instability. *Phys. Plasmas* **13**:092101.

[9] Breslau, J. A. and S. C. Jardin. 2003. Global extended magnetohydrodynamic studies of fast magnetic reconnection. *Phys. Plasmas* **10**:1291-1298.

[10] Jardin, S. C., J. Breslau, and N. Ferraro. 2007. A high-order implicit finite element method for integrating the two-fluid magnetohydrodynamic equations in two dimensions. *J. Comput. Phys.* **226**:2146-2174.

[11] Jardin, S. C., A. Janos, and M. Yamada. 1986. The effect of a column inductive transformer on the S-1 spheromak. *Nuclear Fusion* **26**:647-55.

[12] Ward, D. J. and S. C. Jardin. 1989. Modeling the effects of the sawtooth instability in tokamaks using a current viscosity term. *Nuclear Fusion* **29**:905-14.

[13] Hubba, J. D. 2007. NRL Plasma Formulary, NRL/PU/6790-07-500.

[14] Furth, H., J. Killeen, and M. Rosenbluth. 1960. Finite-resistivity instabilities of a sheet pinch. *Phys. Fluids* **6**:459-84.

[15] Coppi, B., J. M. Greene, and J. L. Johnson. 1966. Resistive instabilities in a diffuse linear pinch. *Nucl. Fusion* **6**:101-17.

[16] Glasser, A. H., S. C. Jardin, and G. Tesauro. 1984. Numerical solution of the resistive magnetohydrodynamic boundary-layer equations. *Phys. Fluids* **27**:1225-1242.

[17] Freidberg, J. P. 1987. *Ideal Magnetohydrodynamics*. New York: Plenum Press.

[18] Friedrichs, K. O. 1955. Nonlinear wave motion in magnetohydrodynamics. LAMS–2105, Los Alamos Scientific Laboratory.

[19] Friedrichs, K. O. and H. Kranzer. 1958. Notes on Magnetohydrodynamics VIII – Nonlinear Wave Motion. New York University NYO–6486.

[20] Ramos, J. J. private communication. 2009.

[21] Grad H. and J. Hogan. 1970. Classical diffusion in a tokamak. *Phys. Rev. Lett.* **24**:1337.

[22] Pao, Y. P. 1976. Classical diffusion in toroidal plasmas. *Phys. Fluids* **19**:1177-82.

[23] Kruskal M. and R. Kulsrud. 1958. Equilibrium of a magnetic confined plasma in a toroid. *Phys. Fluids* **1**:265.

[24] Hirshman, S. P. and S. C. Jardin. 1979. Two-dimensional transport of tokamak plasmas. *Phys. Fluids* **22**:731.

[25] Hazeltine, R. D. and F. L. Hinton. 1973. Collision-dominated plasma transport in toroidal confinement systems. *Phys. Fluids.* **16**:1883-9.

[26] Hirshman, S. P. 1977. Transport of a multiple-ion species plasma in Pfirsch–Schluter regime. *Phys. Fluids* **20**:589.

[27] Hirshman, S. P. 1978. Moment equation approach to neoclassical transport-theory. *Phys. Fluids* **21**:224.

[28] Hirshman, S. P. and D. J. Sigmar. 1981. Neoclassical transport of impurities in tokamak plasmas. *Nuc. Fusion* **21**:1079-1201.

[29] Hirshman, S. P. 1988. Finite-aspect-ratio effects on the bootstrap current in tokamaks. *Phys. Fluids* **31**:3150-52.

[30] Sauter, O., C. Angioni, and Y. R. Lin-Liu. 1999. Neoclassical conductivity and bootstrap current formulas for general axisymmetric equilibria and arbitrary collisionality regime. *Phys. Plasmas* **6**:2834.

[31] Helander, P. and D. J. Sigmar. 2002. *Collisional Transport in Magnetized Plasmas*. Cambridge: Cambridge University Press.

[32] Kinsey, J. E., G. M. Staebler, and R. E. Waltz. 2005. Predicting core and edge transport barriers in tokamaks using the GLF23 drift-wave transport model. *Phys Plasmas* **12**:052503.

[33] Bateman, G., A. H. Kritz, and J. E. Kinsey, et. al. 1998. Predicting temperature and density profiles in tokamaks. *Phys. Plasmas* **5**:1793.

[34] Weiland, J. 1999. *Collective Modes in Inhomogeneous Plasmas and Advanced Fluid Theory*. Bristol: Institute of Physics.

[35] Erba, M., T. Aniel, and V. Basiuk. 1998. Validation of a new mixed Bohm/gyro-Bohm model for electron and ion heat transport against the ITER, Tore Supra and START database discharges. *Nucl. Fusion* **38**:1013.

[36] Jardin, S. C., M. G. Bell, and N. Pomphrey. 1993. TSC simulation of Ohmic discharges in TFTR. *Nuclear Fusion* **33**:371-82.

[37] Blum J. and J. LeFoll. 1984. Plasma equilibrium evolution at the resistive diffusion timescale. *J. Comp. Phys. Reports* **1**:465-494.

[38] Taylor, J. B., Private communication. 1975.

[39] Jardin, S. C. 1981. Self-consistent solutions of the plasma transport equations in an axisymmetric toroidal system. *J. Comput. Phys.* **43**:31-60.

[40] Jardin, S. C. and W. Park. 1981. Two-dimensional modeling of the formation of spheromak configurations. *Phys. Fluids* **24**:679-88.

[41] Jardin, S. C., N. Pomphrey, and J. DeLucia. 1986. Dynamic modeling of transport and positional control of tokamaks. *J. Comput. Phys.* **66**:481-507.

[42] Strang C. and G. J. Fix. 1973. *An Analysis of the Finite Element Method*. Englewood Cliffs, NJ: Prentice-Hall.

[43] Cowper, G. R., E. Kosko, G. Lindberg, et al. 1969. Static and dynamic applications of a high-precision triangular plate bending element. *AIAA J.* **7**:1957.

[44] Jardin, S. C. 2004. A triangular finite element with first-derivative continuity applied to fusion MHD applications. *J. Comput. Phys.* **200**:133-152.

[45] Jardin, S. C. and J. A. Breslau. 2005. Implicit solution of the four-field extended-magnetohydrodynamic equations using high-order high-continuity finite elements. *Phys. Plasmas* **12**:056101.

[46] Dunavant, D. A. 1985. High degree efficient symmetrical Gaussian quadrature rules for the triangle. *Int. J. Numer. Methods Eng.* **21**:1129-1148.

[47] Ferraro, N., S. C. Jardin, X. Luo. 2009. Boundary conditions with reduced quintic finite elements. *PPPL-4497* Princeton, NJ: Princeton University Plasma Physics Laboratory.

[48] Breslau, J., N. Ferraro, S. C. Jardin. 2009. Some properties of the M3D-C^1 form of the 3D magnetohydrodynamics equations. *Phys. Plasmas* **16**:092503.

[49] Jardin, S. C. 1985. Multiple time-scale methods in tokamak magnetohydrodynamics. In *Multiple Timescales*, ed. J. Brackbill and B. Cohen. New York: Academic Press.

[50] Beam, R. M. and R. F. Warming. 1976. An implicit finite-difference algorithm for hyperbolic systems in conservation-law form. *J. Comput. Phys.* **22**:87.

[51] Caramana, E. 1991. Derivation of implicit difference schemes by the method of differential approximation. *J. Comput. Phys.* **96**:484-493.

[52] Sovinec, C. R., A. H. Glasser, T. A. Gianakon, et al. 2004. Nonlinear magnetohydrodynamics simulation using high-order finite elements. *J. Comput. Phys.* **195**:355-386.

[53] Bernstein, I. B., E. A. Frieman, M. D. Kruskal, and R. M. Kulsrud. 1958. An energy principle for hydromagnetic stability problems. *Proc. Royal Soc. London Ser. A* **244**:17

[54] Ferraro, N.M., S. C. Jardin, and P. Snyder. (2010) "Ideal and resistive edge stability calculations with M3D-C^1", submitted to Phys. Plasmas. [available at http://w3.pppl.gov/cemm/Project/ELMbenchmark.pdf]