Mathematics and Maxwell's Equations

Allen H. Boozer
Department of Applied Physics and Applied Mathematics
Columbia University, New York, NY 10027
ahb17@columbia.edu
(Dated: August 12, 2010)

The universality of mathematics and Maxwell's equations is not shared by specific plasma models. Computations become more reliable, efficient, and transparent if specific plasma models are used to obtain only the information that would otherwise be missing. Constraints of high universality, such as those from mathematics and Maxwell's equations, can be obscured or lost by integrated computations. Recognition of subtle constraints of high universality is important for (1) focusing the design of control systems for magnetic field errors in tokamks from perturbations that have little effect on the plasma to those that do. (2) clarifying the limits of applicability to astrophysics of computations of magnetic reconnection in fields that have a double periodicity or have $\vec{B} = 0$ on a surface, as in a Harris sheet. Both assume a symmetry not expected in natural systems. Mathematics and Maxwell's equations imply that neighboring magnetic field lines characteristically separate exponentially with distance along a line. This remarkably universal phenomenon has been largely ignored, though it defines a trigger for reconnection through a critical magnitude of exponentiation. These and other examples of the importance of making distinctions and understanding constraints of high universality are explained.

I. INTRODUCTION

An integrated plasma computation is generally assumed to be preferable to a separation of the information that depends on the plasma model from that determined by mathematics and Maxwell's equations. This assumption is false, because mathematics and Maxwell's equations have a universality not shared by specific plasma models. Computations become more reliable, efficient, and transparent if specific plasma models are used to obtain only the information that would otherwise be missing.

Many phenomena of importance to plasma physics are highly constrained or essentially determined by mathematics and Maxwell's equations. This can be obscured or lost by integrated computations.

The laws of mathematics and Maxwell's equations may be universal, but their constraints can be trumped by the laws of sociology. Familiar models that violate constraints are assumed validated by conventional wisdom. Conclusions derived without the expected detailed plasma model are assumed ignorable. The views of the first reviewer of a recent paper [1] are not atypical: *The derivations are very imprecise, verging on hand-waving.*

Recognition of subtle constraints of high universality is important for clarifying the applicability of theoretical mechanisms for magnetic reconnection in astrophysics, Section II. Two common assumptions on the initial magnetic field used in reconnection simulations are not generic: double periodicity and a Harris sheet, $\vec{B} = B_0 \tanh(x/L_0)\hat{z}$, where B_0 and L_0 are constants. Generic means the assumptions remain valid in the presence of small perturbations. Perturbations are endemic to natural systems, so the applicability of a non-generic model can only be assessed when the effect of breaking its assumptions is known. Another reconnection concept, which is generic and provides a trigger for reconnection, is essentially ignored. Mathematics and Maxwell's equations imply that neighboring magnetic field lines characteristically separate exponentially with distance ℓ along a line, $\delta(\ell) = \delta_0 e^{\sigma(\ell)}$. The exponentiation σ changes even when the magnetic field evolution is ideal. Reconnection becomes generic for evolving magnetic fields when $\sigma \gtrsim 20$, which defines a reconnection trigger.

Recognition of subtle constraints of high universality is also important for focusing the design of control systems for magnetic field errors in tokamks from perturbations that have little effect on the plasma to those that do, Section III.

The importance of making distinctions based on universality extends beyond just mathematics and Maxwell's equations to areas such as classical mechanics and kinetic theory. Not everything can be calculated, but much can be constrained, and what is not constrained may be possible. Constraints of high universality may neither be obvious, nor well known, but of great importance. Examples from kinetic theory are given in Section IV.

Even more general principles provide guidance for research directions in an applied program, such as the effort to demonstrate the feasibility of magnetic fusion energy, Section V.

If the paper can motivate a few individuals in a way of thinking, it will have served its purpose, Section VI.

II. MAGNETIC RECONNECTION

Fundamental differences of opinion on the theory of the reconnection of magnetic fields in three dimensions are apparent in the major reviews [2], [3], [4], [5]. The importance of X-points to reconnection, Figure (1), has been clear in the astrophysical literature since Dungey's

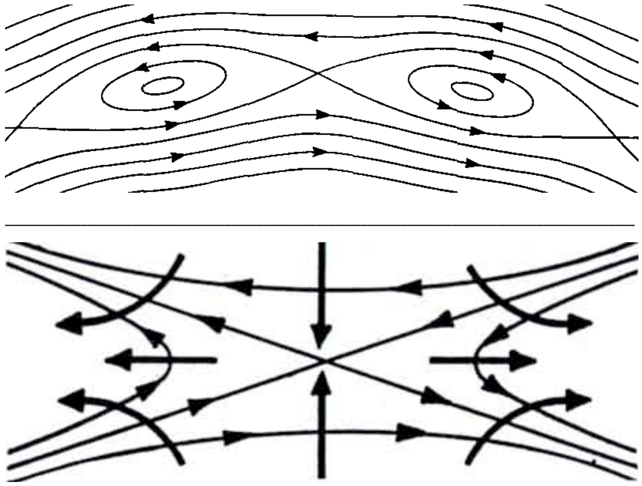

FIG. 1: The top figure shows the way magnetic field lines in a torus appear to advance at a fixed toroidal position near a rational surface, $q = m/n$, that has been split by a chain of magnetic islands. The X-point in the center of the figure is formed by a field line that closes on itself after m toroidal circuits, which is also the number of islands in the chain, and n poloidal circuits. Field lines at the top of the figure have $q > m/n$, so they fail to close on themselves by moving to the left. The field lines at the bottom have $q < m/n$, so they fail to close on themselves by moving to the right. The bottom figure is Dungey's expectation [6] of the plasma flows as well as the field-line trajectories during reconnection, which for the toroidal case is the opening of an island. In three dimensions, an X-point is is the projection of a curve onto a plane, but as discussed in the text, the meaning in an astrophysical context is obscure.

work [6] in the 1950's, and his insights were extended to reconnection in laboratory plasmas by Furth, Killeen, and Rosenbluth [7]. None of the reviews on reconnection deal with the constraints of mathematics and Maxwell's equations, which imply (1) that X-point reconnection has a clear meaning in toroidal but not in astrophysical plasmas and (2) the exponentially increasing separation of neighboring field lines provides a ubiquitous trigger for reconnection.

Magnetic reconnection is an evolution of the magnetic field that cannot be represented by the magnetic field lines moving through space with a well-behaved velocity $\vec{u}(\vec{x},t)$ while preserving their identity. Faraday's law, $\partial\vec{B}/\partial t = -\vec{\nabla}\times\vec{E}$ implies that a magnetic field evolution is completely characterized by an electric field $\vec{E}(\vec{x},t)$. Consequently, mathematical constraints on the form of an arbitrary vector in three dimensions place constraints on reconnection.

Reconnection does not take place in a bounded region of space if the electric field has the mathematical representation

$$\vec{E} + \vec{u}(\vec{x},t)\times\vec{B} = -\vec{\nabla}\Phi(\vec{x},t), \tag{1}$$

where $\vec{u}(\vec{x},t)$ and $\Phi(\vec{x},t)$ are well-behaved functions of position and time [9]. The reason is that the vector $\vec{u}(\vec{x},t)$ gives the evolution of the magnetic field lines and can be interpreted as their velocity. Section II B will discuss this interpretation and its subtleties. When Equation (1) holds the magnetic evolution will be described as ideal.

Equation (1) is not an Ohm's law, but an abstract mathematical statement of the representation of a vector in three space. For example, Equation (1) holds in a vacuum region that is sufficiently small that no field lines close on themselves and in which the magnetic field has no nulls. If the magnetic field is embedded in a perfectly conducting fluid, which has an Ohm's law, $\vec{E}+\vec{v}\times\vec{B}=0$, then an identification of the field line velocity $\vec{u}$ and the fluid velocity $\vec{v}$ shows reconnection cannot take place.

Three scalar functions of position and time are required to represent an arbitrarily evolving vector in three dimensional space. In Equation (1) these three functions are Φ and the two components of $\vec{u}$ that are orthogonal to $\vec{B}$.

The functions $\vec{u}$ and Φ of Equation (1) are well behaved if the differential equation

$$\vec{B}\cdot\vec{\nabla}\Phi = -\vec{E}\cdot\vec{B} \tag{2}$$

has a solution for which $\vec{\nabla}\Phi$ is well behaved. The equation for Φ can be written as $d\Phi/d\ell = -\hat{b}\cdot\vec{E}$, where $\hat{b} = \vec{B}/|\vec{B}|$ and ℓ is the distance along a magnetic field line. The expression,

$$\vec{u} = \frac{(\vec{E}+\vec{\nabla}\Phi)\times\vec{B}}{B^2}, \tag{3}$$

gives a $\vec{u}$ consistent an ideal evolution, Equation (1). Such solutions for Φ and $\vec{u}$ always exist in a sufficiently small bounded region that contains no nulls of the magnetic field, $\vec{B}(\vec{x},t)=0$ and in which the current density has no singularities.

Sometimes reconnection is identified with a large power transfer from the magnetic field to a plasma, but this can occur even when the magnetic evolution is ideal. Poynting's theorem says the power per unit volume transfered from the magnetic field to a conducting medium with a current density $\vec{j}$ is $p_p = \vec{j}\cdot\vec{E}$, so when Equation (1) holds $p_p = \vec{u}\cdot(\vec{j}\times\vec{B}) - \vec{\nabla}\cdot(\Phi\vec{j})$. The last term vanishes when integrated over the region in which $\vec{j}$ is non-zero, but the first term, which is the dot product of field line velocity with the magnetic force on the conducting medium, is not in general zero.

Except for mathematical subtleties, Equation (2) can always be solved consistent with a well behaved field line velocity $\vec{u}$, Equation (3), so it is these subtleties that define when reconnection is possible, Section II A. The details of the plasma model have no relevance to the analysis of Section II A, so the results hold for any plasma model.

The discussion of mathematical subtleties requires a common understanding of the adjective generic. A mathematical statement is generic if it continues to be true in the presence of arbitrarily small perturbations. For example, the generic number of distinct solutions to the equation $x^2 - 2x + c = 0$ is two, $x = 1 \pm \sqrt{1-c}$. The equation has only one solution if $c = 1$, but an arbitrarily small perturbation to c will split the single solution at $c = 1$ into two distinct solutions. When the coefficient c is obtained from complicated physical phenomena, the case that $c = 1$ with absolute precision will arise only in the sense that a stopped watch is absolutely correct twice a day. Perturbations are endemic to natural systems, so the applicability of a non-generic model can only be assessed when the effect of breaking its assumptions is known.

A. Locations at which reconnection can occur

A well behaved velocity of the magnetic field lines, $\vec{u}$ of Equation (3), may not exist at spatial locations of four distinct types:

1. Where $\vec{B}(\vec{x}, t) = 0$.

Singularities arise at places at which $\vec{B}(\vec{x}, t) = 0$ [10], but these singularities are weaker than might be expected because only nulls at isolated points are generic and for an isolated point null the most singular features cancel out. A characteristic singularity is logarithmic. Because of the importance of magnetic nulls in the reconnection literature, they will be explored further in Section II C.

Magnetic field nulls do not arise in the toroidal plasmas that are the focus of laboratory research. Nulls of the magnetic field can arise at isolated points in an astrophysical plasma, though generically neither nulls along a curve nor over a surface can occur Section II C. Let $\langle B^2 \rangle$ be the average over a bounded region of space, then generically $B^2(\vec{x})/\langle B^2 \rangle$ can vanish only at isolated points. Along a curve or on a surface, the smallness of $B^2(\vec{x})/\langle B^2 \rangle$ is a measure of the symmetry of the system—an absolute zero of $B^2(\vec{x})/\langle B^2 \rangle$ along a curve or on a surface requires an absolute symmetry. To be significant, reconnection in astrophysical systems requires enhanced plasma dissipation. If that enhancement depends upon the magnitude of $\langle B^2 \rangle / B^2$ along a curve, then the required symmetry is astronomical in another sense of that word.

The Harris sheet, $\vec{B} = B_0 \tanh(x/L_0)\hat{z}$, where B_0 and L_0 are constants, has become such a standard model of a magnetic field prone to reconnection that Harris' paper [8] has had over 500 citations. A Harris sheet is a surface null of the magnetic field, which is not generic. Any magnetic perturbation with $\hat{x}$ or $\hat{y}$ components removes the null. The applicability of models based on the Harris sheet to astrophysical reconnection cannot be assessed without knowledge of how small $B^2(\vec{x})/\langle B^2 \rangle$ must for $B^2(\vec{x})/\langle B^2 \rangle = 0$ over a surface to be a reasonable approximation.

2. Where magnetic field lines are closed.

Reconnection occurs where magnetic field lines are closed, as on a rational surface of a toroidal plasma, and $d\Phi/d\ell = -\hat{b}\cdot\vec{E}$ is not consistent with $\oint \hat{b}\cdot\vec{E} d\ell = 0$. When this occurs, a solution for Φ does not exist. In toroidal laboratory plasmas, this situation can arise at rational magnetic surfaces, which is the only place reconnection is believed to occur in such plasmas. Reconnection at rational surfaces is the opening of magnetic islands, which splits the rational surface, Figure (1).

Reconnection forced by field line closure arises not only in toroidal plasmas but also in numerical simulations that have periodicity in two directions. Double periodicity provides clear and efficient boundary conditions and is, therefore, used in many simulations. The resulting reconnection depends sensitively on the behavior of the plasma near X-points on the separatrices of islands. This theory is very important in toroidal laboratory plasmas, but its relevance to astrophysics is difficult to comprehend.

The definition of an X-pont in an astrophysical plasma is obscure, Figure (1). An X-point in a toroidal plasma is has two characteristics: (1) The X-point is formed by a field line that closes on itself. (2) The magnetic field lines in the neighborhood of that line are of two types: those that exponentially separate from and those exponentially approach the X-point with distance ℓ along each line. Since the concept of a closed magnetic field line in an astrophysical plasma is not taken seriously, an X-point is sometimes taken to be a null of the magnetic field along a line, but a line null is destroyed by small perturbations—only point nulls are generic, Section II C.

3. Where a magnetic field line goes from one boundary to another.

Reconnection arises when a magnetic field line goes from one boundary to another on which Φ is specified—for example $\Phi = 0$ on a perfectly conducting plate—and $d\Phi/d\ell = -\hat{b} \cdot \vec{E}$ does not give a solution that obeys this condition, which means a solution for Φ does not exist.

Reconnection that is forced by the boundary conditions has two subtleties associated with the distance to the boundaries. (1) In a highly conducting plasma, the plasma is isolated from boundary effects for a time less than the transit time of a shear Alfvén wave, which limits the rate of reconnection when the distance along the field lines from one boundary to another is great. (2) Numerical studies are expensive with distant boundaries unless careful approximations are made. A comprehensive code that does not make such approximations is difficult to apply in a way that is relevant to common astrophysical situations.

4. Where adjacent magnetic field lines separate exponentially.

Magnetic field lines that have an infinitesimal separation $\vec{\delta}_0$ a one point in space generically have a exponentially increasing separation, $\vec{\delta}(\ell) \sim \delta_0 e^{\sigma(\ell)}$, with distance along the lines ℓ. When σ is large, the electric potential Φ of Equation (2), $d\Phi/d\ell = -\hat{b}\cdot\vec{E}$, naturally develops exponentially large gradients across the magnetic field lines since the potential at adjacent points comes from integrations along widely separated field lines. For example, the radius of the sun, $1.4 \times 10^6 km$, divided by the ion gyroradius in the photoshere, $\rho_i \approx 70cm$, is approximately $e^{21.4}$.

Reconnection characteristically occurs [11] for $\sigma \gtrsim 20$. Even in an ideal evolution σ changes, so σ reaching a critical value can be a trigger for reconnection.

A clear demonstration that in a smooth magnetic field with no nulls neighboring magnetic field lines generically separate exponentially has not be given for open magnetic field lines. However when the field lines lie within a region bounded by a torus, it is well known that trajectories characteristically separate with $k_L \equiv \lim_{\ell\rightarrow\infty} \sigma/\ell$ called the Lyapunov exponent. When k_L is non-zero, the magnetic field is said to be chaotic. Toroidal laboratory experiments can have non-chaotic magnetic fields, $k_L = 0$, but even small perturbations, $\delta B/B < 1\%$ can make an non-chaotic field chaotic.

This paragraph is a demonstration that generically $\delta \sim \delta_0 e^{\sigma(\ell)}$ and can be skipped on a first reading. At an arbitrary point along a magnetic field line, define Cartesian coordinates (x_0, y_0) in a plane perpendicular to that line. Magnetic field lines are the solutions to $d\vec{x}/d\tau = \vec{B}(\vec{x})$. The trajectory of a field line that passes through the plane at (x_0, y_0) is $\vec{x}(x_0, y_0, \tau)$, where $\tau = 0$ on the perpendicular plane. A trajectory started at $|x_0| \rightarrow 0$ and $y_0 = 0$ is separated from the trajectory that passes through $x_0 = 0, y_0 = 0$ by $\vec{\delta}_x = (\partial\vec{x}/\partial x_0)x_0$. Initial conditions do not change along a trajectory, so $(\partial\vec{x}/\partial\tau)_{x_0 y_0} = \vec{B}(\vec{x})$, and

$$\frac{\partial}{\partial\tau}\left(\frac{\partial\vec{x}}{\partial x_0}\right) = \overleftrightarrow{\mathcal{B}}\cdot\left(\frac{\partial\vec{x}}{\partial x_0}\right), \tag{4}$$

where $\overleftrightarrow{\mathcal{B}}\cdot(\partial\vec{x}/\partial x_0) \equiv (\partial\vec{x}/\partial x_0)\cdot\vec{\nabla}\vec{B}$. In Cartesian coordinates $\mathcal{B}_{xy} = \partial B_x/\partial y$, *etc.* The derivatives of the magnetic field are to be evaluated along the trajectory of the field line that passes through $x_0 = 0, y_0 = 0$, so $\overleftrightarrow{\mathcal{B}}$ is a function of τ alone. The magnitude of the separation $\vec{\delta}_x$ is determined by the evolution of $(\partial\vec{x}/\partial x_0)^2$, which is given by

$$\frac{\partial}{\partial\tau}\left(\frac{\partial\vec{x}}{\partial x_0}\right)^2 = \frac{\partial\vec{x}}{\partial x_0}\cdot\left(\overleftrightarrow{\mathcal{B}}^\dagger + \overleftrightarrow{\mathcal{B}}\right)\cdot\frac{\partial\vec{x}}{\partial x_0}, \tag{5}$$

where $\mathcal{B}^\dagger_{ij} \equiv \mathcal{B}_{ji}$. Equation (5) has the solution

$$\left(\frac{\partial\vec{x}}{\partial x_0}\right)^2 = e^{2\sigma_x(\tau)}, \text{ where } \sigma_x(\tau) \equiv \int_0^\tau \hat{e}_x\cdot\overleftrightarrow{\mathcal{B}}\cdot\hat{e}_x d\tau', \tag{6}$$

$\hat{e}_x(\tau) \equiv (\partial\vec{x}/\partial x_0)/|\partial\vec{x}/\partial x_0|$ is a unit vector, and $|\partial\vec{x}/\partial x_0| = 1$ on the $\tau = 0$ plane. The absolute value of the function σ_x tends to become large when the distance along the arbitrary field line, $\ell = \int B d\tau$, is large. If σ_x is large and negative, the separation $|\vec{\delta}_x|$ becomes exponentially small, but a field line started at $x_0 = 0$ and $|y_0| \rightarrow 0$ must then have an exponentially large separation $\vec{\delta}_y = (\partial\vec{x}/\partial y_0)y_0$. To show this, note that the Jacobian, $\mathcal{J}$, of (x_0, y_0, τ) coordinates must be constant along the field line trajectories. The magnetic flux $\int(\vec{B}\cdot\vec{\nabla}\tau)\mathcal{J}dx_0 dy_0$ must be independent of τ since $\vec{\nabla}\cdot\vec{B} = 0$. The orthogonality condition of general coordinates says $\vec{\nabla}\tau\cdot\partial\vec{x}/\partial\tau = 1$, which implies $\vec{B}\cdot\vec{\nabla}\tau = 1$. In the limit as $|x_0| \rightarrow 0$ and $|y_0| \rightarrow 0$, the flux is just $\mathcal{J}x_0 y_0$, which implies $\mathcal{J}$ is a constant $\mathcal{J}_0$ along the field line. The Jacobian $\mathcal{J} \equiv \{(\partial\vec{x}/\partial x_0)\times(\partial\vec{x}/\partial y_0)\}\cdot(\partial\vec{x}/\partial\tau) \leq |\partial\vec{x}/\partial x_0||\partial\vec{x}/\partial y_0||\partial\vec{x}/\partial\tau|$. Now $|\partial\vec{x}/\partial\tau| = B$, so the inequality $|\partial\vec{x}/\partial x_0||\partial\vec{x}/\partial y_0|B \geq \mathcal{J}_0$ must hold. Since $|\partial\vec{x}/\partial x_0| = \exp(\sigma_x)$, one finds $|\partial\vec{x}/\partial y_0| \geq (\mathcal{J}_0/B)e^{-\sigma_x(\tau)}$.

B. Field line velocity

If a well behaved $\vec{u}(\vec{x}, t)$ of Equation (1) exists in a bounded region of space, then magnetic field lines preserve their identity if they are assumed to move with the velocity $\vec{u}$.

The concept of a field line velocity without reference to the medium in which it is embedded has been subject to criticism because of non-uniqueness, most notably by William Newcomb [12]. However, the important point is whether a velocity $\vec{u}$ of the magnetic file lines exists and not whether more than one mathematically valid expression for $\vec{u}$ can be found.

The non-uniqueness in the magnetic field line velocity $\vec{u}$ comes from the arbitrariness in the value of potential Φ_0 in the $\ell = 0$ surface when $d\Phi/d\ell = -\hat{b}\cdot\vec{E}$ is integrated to obtain Φ through out a bounded region of space. An arbitrary velocity $\vec{u}_0 = (\vec{B}\times\vec{\nabla}\Phi_0)/B^2$ can be added to the field line motion driven by the electric field. Since magnetic field lines are the trajectories of a Hamiltonian systems, the freedom in $\vec{u}$ can also be taken to be the freedom of canonical transformations as shown below.

1. Field line velocity and the Clebsch representation

The interpretation of $\vec{u}$ as the velocity of the magnetic field lines can be obtained from the Clebsch representation

$$\vec{B} = \vec{\nabla}\psi\times\vec{\nabla}\vartheta, \tag{7}$$

which holds locally for an arbitrary divergence-free field without nulls. The two functions $\psi(\vec{x}, t)$, which has units of magnetic flux, and $\vartheta(\vec{x}, t)$, which is dimensionless, are known as Clebsch potentials to physicists and Euler potentials to mathematicians [13]. Neither ψ nor ϑ change along a magnetic field line, $\vec{B}\cdot\vec{\nabla}\psi=0$ and $\vec{B}\cdot\vec{\nabla}\vartheta=0$, so the locations of the magnetic field lines at time t are given by $\vec{x}(\psi,\vartheta,\ell,t)$, where ℓ can be interpreted as the distance along a magnetic field line. That is,

$$\frac{\partial\vec{x}(\psi,\vartheta,\ell,t)}{\partial\ell}=\frac{\vec{B}}{B}, \tag{8}$$

and the velocity of a magnetic field line through space is

$$\vec{u}_f\equiv\frac{\partial\vec{x}(\psi,\vartheta,\ell,t)}{\partial t}, \tag{9}$$

which is the velocity of a (ψ,ϑ,ℓ) point through space.

The mathematics of general coordinates implies that

$$\frac{\partial\vec{B}}{\partial t}=\vec{\nabla}\times(\vec{u}_f\times\vec{B}). \tag{10}$$

Faraday's law then gives $\vec{E}+\vec{u}_f\times\vec{B}=-\vec{\nabla}\Phi$, which has the same form as Equation (1), when the field line velocity $\vec{u}_f$ and $\vec{u}$ are identified. The derivation Equation (10), which is involved and not required to understand the remainder of the paper, is given below.

Equation (10) has a remarkable corollary. If a Clebsch representation exists, $\vec{B}(\vec{x},t)=\vec{\nabla}\psi\times\vec{\nabla}\vartheta$ with a well-defined $\psi(\vec{x},t)$ and $\vartheta(\vec{x},t)$, then the evolution of the magnetic field appears ideal.

Magnetic nulls are important, for Graham and Henyey have shown [15] that the Clebsch representation, Equation (7), does not generically exist near a null. The word generic is important, for otherwise a counter example can be given. The generic curl-free magnetic field near a point null is $\vec{B}=ax\hat{x}+by\hat{y}-(a+b)z\hat{z}$, where a and b are constants. If a and b are equal, then the field has a Clebsch representation, $\psi=a(x^2+y^2)z$ and $\vartheta=\arctan(x/y)$. Even in first order in $\epsilon\equiv(a-b)/a$, the required change to the Clebsch potentials has singularities, $\delta\psi=\epsilon a(x^2-y^2)z\ln z$ and $\delta\vartheta=\epsilon\frac{xy}{x^2+y^2}\ln z$, which is consistent with the failure of a Clebsch representation to exist near a generic null.

The relation between the freedom in $\vec{u}$ and the freedom of canonical transformations to the Clebsch representation, $\vec{B}(\vec{x})=\vec{\nabla}\psi\times\vec{\nabla}\vartheta$, can be derived in a few lines. Given an arbitrary function $S(\Psi,\vartheta)$, where $\psi=\partial S/\partial\vartheta$ and $\Theta=\partial S/\partial\Psi$, the magnetic field also has the representation $\vec{B}(\vec{x})=\vec{\nabla}\Psi\times\vec{\nabla}\Theta$. This result is proven by substituting $\psi=\partial S/\partial\Psi$ into $\vec{B}=\vec{\nabla}\times(\psi\vec{\nabla}\vartheta)$, which gives $\vec{B}=\vec{\nabla}\times\{\vec{\nabla}S-(\partial S/\partial\Psi)\vec{\nabla}\Psi\}$.

The derivation of Equation (10) will be given in this paragraph but can be skipped. Let (ξ^1,ξ^2,ξ^3) denote (ψ,ϑ,ℓ), then

$$0=\left(\frac{\partial\xi^i}{\partial t}\right)_{\xi^i}=\frac{\partial\xi^i}{\partial\vec{x}}\cdot\left(\frac{\partial\vec{x}}{\partial t}\right)_{\xi^i}+\left(\frac{\partial\xi^i}{\partial t}\right)_{\vec{x}}. \tag{11}$$

The implication is that $\partial\psi/\partial t=-\vec{u}_f\cdot\vec{\nabla}\psi$, $\partial\vartheta/\partial t=-\vec{u}_f\cdot\vec{\nabla}\vartheta$, and $\partial\ell/\partial t=-\vec{u}_f\cdot\vec{\nabla}\ell$. The orthogonality relations of general coordinates, which are derived in the appendix of [14], then imply

$$\vec{u}_f=-\left(\frac{\partial\psi}{\partial t}\right)_{\vec{x}}\frac{\partial\vec{x}}{\partial\psi}-\left(\frac{\partial\vartheta}{\partial t}\right)_{\vec{x}}\frac{\partial\vec{x}}{\partial\vartheta}-\left(\frac{\partial\ell}{\partial t}\right)_{\vec{x}}\frac{\partial\vec{x}}{\partial\ell}. \tag{12}$$

Using the dual relations of general coordinates, which are derived in the appendix of [14], Equation (7) can be rewritten as $\vec{B}=(\partial\vec{x}/\partial\ell)/\mathcal{J}$ and

$$\vec{u}_f\times\vec{B}=\left(\frac{\partial\psi}{\partial t}\right)_{\vec{x}}\vec{\nabla}\vartheta-\left(\frac{\partial\vartheta}{\partial t}\right)_{\vec{x}}\vec{\nabla}\psi, \tag{13}$$

where $\mathcal{J}$ is the Jacobian of (ψ,ϑ,ℓ) coordinates. The time derivative of $\vec{B}=\vec{\nabla}\psi\times\vec{\nabla}\vartheta$ is $\partial\vec{B}/\partial t=\vec{\nabla}(\partial\psi/\partial t)_{\vec{x}}\times\vec{\nabla}\vartheta+\vec{\nabla}\psi\times\vec{\nabla}(\partial\vartheta/\partial t)_{\vec{x}}$, so

$$\frac{\partial\vec{B}}{\partial t}=\vec{\nabla}\times\left(\frac{\partial\psi}{\partial t}\vec{\nabla}\vartheta-\frac{\partial\vartheta}{\partial t}\vec{\nabla}\psi\right)=\vec{\nabla}\times(\vec{u}_f\times\vec{B}) \tag{14}$$

using Equation (13).

C. Magnetic nulls

Magnetic nulls have a special importance to both the existence of the Clebsch representation, Section II B, and the representation of the electric field given in Equation (1). A generic magnetic field can only have nulls at isolated points. To prove this, it will be shown that an arbitrarily small magnetic perturbation $\delta\vec{B}$ can (1) eliminate all nulls if the unperturbed field were zero on a surface as in a Harris sheet, (2) either eliminate all nulls or convert them into isolated nulls if the unperturbed field were zero on a curve, but (3) only move the location of a null at a point by a distance proportional to the perturbation, not eliminate the null.

The time evolution of a magnetic field can be viewed as the response of a magnetic field to a series perturbations. This will be shown to imply that nulls can move and that a single null cannot appear in a field that had no nulls. Nevertheless, space-time points can exist in which two nulls separate in a region in which no nulls previously existed, and an example will be given. An example will also be given to show that an evolving magnetic field with a point null can force $\vec{\nabla}_\perp\Phi$ to have a logarithmic singularity.

1. *Only point nulls generic*

In Cartesian coordinates, the nulls of a magnetic field are given by the solving three simultaneous equations for three unknowns:

$$B_x(x,y,z)=0;\quad B_y(x,y,z)=0;\quad B_z(x,y,z)=0. \tag{15}$$

The generic existence of only isolated magnetic nulls is a statement about the nature of solutions of three equations for three unknowns.

If a magnetic field is zero at a point, the null point is moved by a small perturbation but not eliminated. Near a point null at $\vec{x}=\vec{x}_0$, the magnetic field has the form

$$\vec{B}=\overleftrightarrow{\mathcal{B}}\cdot(\vec{x}-\vec{x}_0). \tag{16}$$

In the presence of a perturbation, $\vec{B}=\delta B+\overleftrightarrow{\mathcal{B}}\cdot(\vec{x}-\vec{x}_0)$. The perturbed field has a zero at $\vec{x}=\vec{x}_0-\overleftrightarrow{\mathcal{B}}^{-1}\cdot\delta\vec{B}$, which is nearby unless the inverse matrix $\overleftrightarrow{\mathcal{B}}^{-1}$ does not exist. Since $\overleftrightarrow{\mathcal{B}}$ is a square matrix, a nearby null of the magnetic field exists for a sufficiently small perturbation unless the determinant of the matrix $\overleftrightarrow{\mathcal{B}}$ vanishes. The vanishing of the determinant would be a fourth equation in addition to the three of Equation (15). Four simultaneous equations for three unknowns is said to be overdetermined, and generically overdetermined systems of equations have no solution.

If a magnetic field is zero along a curve, then a magnetic perturbation that is non-zero along that curve will either remove the field null altogether or turn it into point nulls. In Cartesian coordinates, a null along the x axis, which means at $y=0$, $z=0$, implies that near the x axis a divergence-free field obeys

$$B_x = y\beta_{xy}(x)+z\beta_{xz}(x) \tag{17}$$

$$B_y = y\beta_{yy}(x)-\frac{y^2}{2}\frac{d\beta_{xy}}{dx}+z\beta_{yz}(x) \tag{18}$$

$$B_z = y\beta_{zy}(x)+z\beta_{zz}(x)-\frac{z^2}{2}\frac{d\beta_{xz}}{dx}. \tag{19}$$

In the presence of a perturbation, the null is moved to $y=y_0(x)$ and $z=z_0(x)$, which are given by the solution to the equations $\delta B_y+y\beta_{yy}-y^2\beta'_{xy}/2+z\beta_{yz}=0$ and $\delta B_z+y\beta_{zy}+z\beta_{zz}-z^2\beta'_{xz}/2=0$, where primes denote x differentiation. The magnetic field then has a null only if $B_x(x)=\delta B_x(x,y_0,z_0)+y_0\beta_{xy}(x)+z_0\beta_{xz}(x)=0$. Unless at least one quantity, δB_x, y_0, β_{xy}, z_0, and β_{xz}, is x dependent, the magnetic field has no nulls except for a special perturbation direction. If at least one quantity has x dependence, then $B_x(x)$ can pass through zero, which reduces the line null to one or more point nulls. The distance between the point nulls is proportional to the perturbation.

If a magnetic field has a null that extends over a surface of non-zero area, as in a Harris sheet, then a magnetic perturbation $\delta\vec{B}$ generically removes the magnetic field null. For a Harris sheet, $\vec{B}=B_0\tanh(x/L_0)\hat{z}$, the addition of a perturbation that has $\hat{x}$ or $\hat{y}$ components completely removes the null. A magnetic field can have a more general null over the $x=0$ plane than the Harris sheet. Near that plane the y and z components of the unperturbed magnetic field have the forms $B_y=xB'_y$ and $B_z=xB'_z$. Since $\vec{\nabla}\cdot\vec{B}=0$, the unperturbed x component has the form $B_x=x^2B''_x/2$. If a perturbation is applied that has an x component, δB_x, of the same sign as B''_x, then no solution for a null exists in the range of the validity of the approximation $B_x=x^2B''_x/2$. Even if δB_x is applied with the opposite sign to B''_x, then the null in B_x moves to $x_\pm=\pm\sqrt{2|\delta B_x/B''_x|}$ so the B_y and B_z components become non-zero and the null disappears unless both $x_\pm B'_y+\delta B_y$ and $x_\pm B'_z+\delta B_z$ vanish at a common point, which is not a generic condition. The minimum of the field strength in the vicinity of null surface of the unperturbed field in the presence of the perturbation is $|\vec{B}|\approx|\delta\vec{B}|$.

The proof that that line and surface nulls are not generic can be extended to arbitrary curves and surfaces using the methods of general coordinates, appendix of [14].

2. Evolution of nulls

The evolution of magnetic nulls is closely related to the response of a null to a perturbation for $\vec{B}(\vec{x},t+\delta t)=\vec{B}(\vec{x},t)+\delta\vec{B}$, where the perturbation is $\delta\vec{B}\equiv(\partial\vec{B}/\partial t)\delta t$. The implication is that a null can neither be created nor destroyed unless $\overleftrightarrow{\mathcal{B}}$ of Equation (16) has a vanishing determinant at the location of a null. The vanishing of the determinant at a null is four conditions on (B_x,B_y,B_z) as functions of (x,y,z,t). A set of four simultaneous equations with four unknowns is not an overdetermined system, so one expects the four conditions to be satisfied at isolated space-time points.

The formation of a pair of nulls at space-time points at which four constraints are obeyed, which are $\vec{B}=0$ and the determinant of $\overleftrightarrow{\mathcal{B}}$ equals zero, is illustrated by the curl-free magnetic field $B_x=ax$, $B_y=(b-cz)y$, $B_z=-(a+b)z-c(y^2-z^2)/2$, where a and c are positive constants, but $b(t)$ is a function of time. The matrix

$$\overleftrightarrow{\mathcal{B}}=\begin{pmatrix} a & 0 & 0\\ 0 & b-cz & -cy\\ 0 & -cy & -(a+b-cz)\end{pmatrix}. \tag{20}$$

A magnetic field null always exists at $x=0$, $y=0$, $z=0$. A more interesting pair of nulls occurs at $x=0$, $y^2=-(b/c^2)(b+2a)$, $z=b/c$ when $y^2>0$. The determinant at this pair of nulls is $ab(b+2a)$. If $b(t)=\dot{b}t-2a$, then for $t<0$ the only null is at $x=0$, $y=0$, $z=0$. However, for t small but positive an additional pair of nulls exists at $x=0$, $y=\pm\sqrt{(2a/c^2)\dot{b}t}$, and $z=-2a/c$.

3. Non-ideal behavior due to a point null

A point null of the magnetic field can produce a logarithmic singularity in the field line velocity $\vec{u}$, which violates the validity of an ideal evolution, Equation (1).

A logarithmic singularity in $\vec{u}$ is illustrated by the evolution of curl-free magnetic field $\vec{B}(\vec{x},t)=ax\hat{x}+by\hat{y}-(a+b)z\hat{z}$, where $\partial\vec{B}/\partial t=\dot{a}(x\hat{x}-z\hat{z})$ and $\dot{a}\equiv da/dt$. An electric field that gives this evolution is $\vec{E}=\dot{a}xz\hat{y}$. The potential is given by the equation

$$ax\frac{\partial\Phi}{\partial x}+by\frac{\partial\Phi}{\partial y}-(a+b)z\frac{\partial\Phi}{\partial z}=-\dot{a}bxyz. \tag{21}$$

A solution is $\Phi=-(d\ln a/dt)bxyz\ln x$. Solutions that are proportional to $xyz\ln y$ and $xyz\ln z$ also exist. The gradient of the potential that goes as $xyz\ln x$ is

$$\vec{\nabla}\Phi=-\frac{da}{dt}\frac{b}{a}\left\{(yz+yz\ln x)\hat{x}+(xz\ln x)\hat{y}+(xy\ln x)\hat{z}\right\}. \tag{22}$$

Although $\vec{E}\times\vec{B}/B^2$ is well behaved in the vicinity of the null, $\vec{\nabla}\Phi$ is not, and the velocity $\vec{u}$ has a logarithmic singularity as the x axis is approached. Nevertheless, the parallel electric field, which drives the singularity in $\vec{u}$, is not only non-singular but actually zero along the x axis,

$$\frac{\vec{E}\cdot\vec{B}}{B}=\frac{da}{dt}b\frac{xyz}{\sqrt{a^2x^2+b^2y^2+(a+b)^2z^2}}. \tag{23}$$

III. CONTROL OF EXTERNAL MAGNETIC PERTURBATIONS TO TOKAMAKS

The constraints of mathematics and Maxwell's equations also have important implications for the control of magnetic field errors in tokamaks.

Tokamak plasmas can be thrown into a disruptive state by non-axisymmetric magnetic perturbations [16] as small as $\delta B/B\sim10^{-4}$. The maintenance of sufficiently tight construction tolerances to eliminate such perturbations is prohibitively expensive in cost and schedule. Correction coils are used to control the external magnetic field errors, but the results may seem paradoxical [17]:

1. Successful error field control does not mean error field reduction.

   The error field control system on DIII-D increases the toroidal asymmetry of the magnetic field when it optimally mitigates the effect of the error field.

2. The optimal location for error field control coils may be rotated from the poloidal location of the source of the error.

   On NSTX an inboard field error is controlled by an outboard magnetic field an order of magnitude smaller.

3. The drive for islands at the $q=2$ surface has little relation to the resonant $m=2,n=1$ part of the external perturbation.

   The Fourier component of the external magnetic field with the largest drive for islands at the $q=2$ surface is $m\sim10,n=1$.

Mathematics and Maxwell's equations provide clear guidance on the design of an error field control system though this guidance is just being understood and has not yet been fully implemented. Studies have not been made of the engineering tradeoff between a more complete error field control system and lowered construction tolerances.

Traditional analyses have studied the effect on the plasma of various Monte Carlo realizations of coil displacements [18], which presuppose that the plasma effect of an arbitrary external perturbation can be assessed. If true, the required construction accuracy and the effectiveness of the control system can be better determined by finding the external field distributions of high plasma sensitivity. If false, error field analyses should be restricted to what is known, which is the form and the critical amplitude of the error with the highest plasma sensitivity, but then the construction tolerances are very tight.

A. Representation of magnetic fields using fluxes

Magnetic fluxes on an arbitrary closed surface, called a control surface, provide an efficient and intuitive description of magnetic fields. Given a control surface, the magnetic field $\vec{B}$ throughout space can be uniquely separated into two parts: the field $\vec{B}_i$ produced by currents within the region enclosed by the control surface and the field $\vec{B}_x$ produced by currents external to the control surface, $\vec{B}=\vec{B}_i+\vec{B}_x$.

Within the region enclosed by a control surface the external field satisfies $\vec{B}_x=\vec{\nabla}\phi$ with $\nabla^2\phi=0$ and is completely determined by either its normal component, $\vec{B}_x\cdot\hat{n}$ on that surface.

Magnetic fluxes are defined on a control surface using a set of orthonormal functions $f_i(\theta,\varphi)$, such as the Fourier functions. A set of functions is orthonormal if $\oint f_if_jwda=\delta_{ij}$ when integrated over the control surface, where $w>0$ is a weight function with $\oint wda=1$.

The magnetic fluxes that define the external field within the region enclosed by a control surface are

$$\Phi_i^{(x)}=\oint f_i\vec{B}_x\cdot\hat{n}da,\ \text{so}\ \vec{B}_x\cdot\hat{n}=w\sum_i\Phi_i^{(x)}f_i. \tag{24}$$

In the region outside of the control surface, the magnetic field $\vec{B}_i$, which is due to currents within the region enclosed by the control surface, can also be represented by a magnetic flux on the surface.

B. Plasma sensitivity

For error fields the important issue is the sensitivity of the plasma to external perturbations. The distributions of external magnetic field can be ordered by the plasma sensitivity. The first is that external magnetic perturbation $\left(\delta\vec{B}\cdot\hat{n}\right)_i$ that with the smallest flux,

$$\sigma_i \equiv \sqrt{\oint \left(\delta\vec{B}\cdot\hat{n}\right)_i^2 \frac{da}{w}}, \tag{25}$$

has a significant effect on the plasma. The second is the orthogonal distribution which has the second smallest flux σ_i for significance, and so forth. Two external magnetic distributions are orthogonal if

$$\oint \left(\delta\vec{B}\cdot\hat{n}\right)_i \left(\delta\vec{B}\cdot\hat{n}\right)_j \frac{da}{w} = 0. \tag{26}$$

The sensitive distributions define an orthonormal set of expansion functions, where $\left(\delta\vec{B}\cdot\hat{n}\right)_i \propto w(\theta)f_i(\theta,\varphi)$, where the weight function $w>0$ satisfies $\oint w da = 1$. Using this set of orthonormal functions the plasma sensitivity is given by a diagonal matrix $S_{ij}\equiv\delta_{ij}/\sigma_i^2$. The quantity $\vec{\Phi}_x^\dagger\cdot\overleftrightarrow{S}\cdot\vec{\Phi}_x$ is the dimensionless measure of the strength of the external perturbation compared to the level at which significant plasma effects occur.

The external perturbation to which the plasma is most sensitive drives the least stable kink mode of ideal MHD, which at a sufficiently high plasma pressure becomes the resistive wall mode [17]. However, this perturbation was not the basis of the error field control system on ITER, [16]. The sensitivity to the least stable mode is not surprising, since a system at marginal stability can be displaced an arbitrarily large amount by an infinitesimal force. The $\vec{B}_x\cdot\hat{n}$ of this perturbation has very narrow lobes of width Δ near the outboard midplane, Figure (2). For ITER, the expected width of the lobes [19] compared to the minor radius is $1/2 \gtrsim \Delta/a \gtrsim 1/3$, and perturbations with $\delta B/B\sim 10^{-4}$ are expected to cause disruptions, so error field control is needed. Studies have been carried out for ITER on the effect of external magnetic field distribution on the drive for islands at a number of the rational surfaces [19]. These studies define external magnetic perturbations of secondary plasma sensitivity, but whether these are the external distributions of the greatest sensitivity is not clear.

C. Magnetic field errors

The magnetic field errors produced by dislocations in positions of the main equilibrium coils can be represented by a flux vector $\delta\vec{\Phi}_m$, which is defined on a control surface just on the plasma side of the coils, which is called the coil surface. The external normal field on the unperturbed plasma boundary is

$$\vec{\Phi}_x = \overleftrightarrow{T}\cdot\delta\vec{\Phi}_m + \overleftrightarrow{M}\cdot\vec{J}, \tag{27}$$

where $\overleftrightarrow{T}$ is called a transfer matrix, which relates fluxes of external magnetic field on the plasma boudary to those on the coil surface, $\delta\vec{\Phi}_m$, and $\overleftrightarrow{M}$ is a mutual inductance

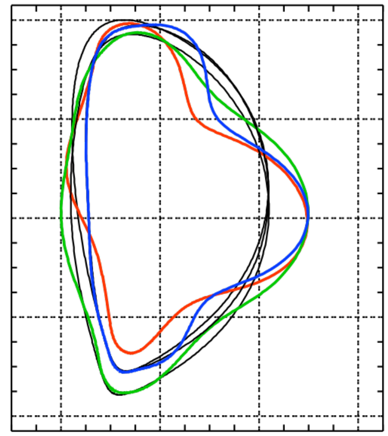

FIG. 2: The $n=1$ external magnetic perturbation, $\vec{B}_x\cdot\hat{n}$, is illustrated to which ITER plasmas are most sensitive in three scenarios [19]. The back contours give the plasma boundaries. The deviations of the colored from the black contours are proportional $\vec{B}_x\cdot\hat{n}$. The lobe of $\vec{B}_x\cdot\hat{n}$ on the outboard side of the plasma has a width Δ, where $1/2\gtrsim\Delta/a\gtrsim 1/3$. This lobe is too narrow for distant coils to efficiently cancel this perturbation. Even a coil that is sufficiently small to be represented by a dipole can effective push back against this lobe only if its distance from the plasma is $\sim a/8$, where a is the plasma half-width or minor radius.

between the fluxes on the plasma surface and the currents $\vec{J}$ in error field control coils. Neil Pomphrey has written a code that can determine $\overleftrightarrow{T}$, and the matrix $\overleftrightarrow{M}$ can be determined with standard Biot-Savart codes.

The error field control coils can null some of the external fluxes on the plasma surface. The practical issue is how well can this be done. Naively it would appear that currents in the control coils $\vec{J} = -\overleftrightarrow{M}^{-1}\cdot\overleftrightarrow{T}\cdot\delta\vec{\Phi}_m$ would null the entire external perturbation $\vec{\Phi}_x=0$. The subtlety, which makes this impossible, is contained in the matrix inverse $\overleftrightarrow{M}^{-1}$.

The inverse of a matrix is best interpreted using Singular Value Decomposition (SVD). SVD represents any matrix M_{ij} in the form $\overleftrightarrow{M}=\overleftrightarrow{U}_m\cdot\overleftrightarrow{m}\cdot\overleftrightarrow{V}_m^\dagger$, where $\overleftrightarrow{m}$ is diagonal, with real diagonal components, or singular values, $m_i\geq 0$ arranged with the largest first. The matrices $\overleftrightarrow{U}_m$ and $\overleftrightarrow{V}_m$ are orthogonal, which means $\overleftrightarrow{U}_m^\dagger\cdot\overleftrightarrow{U}_m=\overleftrightarrow{1}$. If $\vec{f}$ is a matrix vector with the expansion functions $f_i(\theta,\varphi)$ as its elements and if $\vec{F}\equiv\overleftrightarrow{U}^\dagger\cdot\vec{f}$, then $\oint F_i\vec{B}_x\cdot\hat{n}da$ are the external fluxes on the plasma in the order of ease with which they can be driven by the control coils.

It is easily checked that the inverse of $\overleftrightarrow{M}$ is $\overleftrightarrow{M}^{-1} = \overleftrightarrow{V}_m\cdot\overleftrightarrow{m}^{-1}\cdot\overleftrightarrow{U}_m^\dagger$, where $\overleftrightarrow{m}^{-1}$ is diagonal with elements $1/m_i$. The singular values can span a large range of values or even be zero, so the $1/m_i$ can become large, or even infinite. The pseudoinverse, $\overleftrightarrow{M}_k^{-1}$ is the matrix in which only the k largest singular values are retained in forming the inverse with the other diagonal elements chosen to be zero rather that $1/m_i$.

As discussed in Section III B the plasma sensitivity is measured by $\vec{\Phi}_x^\dagger \cdot \overleftrightarrow{S} \cdot \vec{\Phi}_x$. The currents in the control coils $\vec{J}$ of Equation (27) should be chosen to minimize $\vec{\Phi}_x^\dagger \cdot \overleftrightarrow{S} \cdot \vec{\Phi}_x$, which gives $\overleftrightarrow{M}^\dagger \cdot \overleftrightarrow{S} \cdot \overleftrightarrow{M} \cdot \vec{J} = -\overleftrightarrow{M}^\dagger \cdot \overleftrightarrow{S} \cdot \overleftrightarrow{T} \cdot \delta\vec{\Phi}_m$. The optimal control currents are then $\vec{J} = -\overleftrightarrow{\mathcal{C}}_k \cdot \delta\vec{\Phi}_m$, where

$$\overleftrightarrow{\mathcal{C}}_k \equiv \left(\overleftrightarrow{M}^\dagger \cdot \overleftrightarrow{S} \cdot \overleftrightarrow{M}\right)_k^{-1} \cdot \overleftrightarrow{M}^\dagger \cdot \overleftrightarrow{S} \cdot \overleftrightarrow{T}, \tag{28}$$

and the subscript k means k singular values are retained in the pseudoinverse of $\overleftrightarrow{M}^\dagger \cdot \overleftrightarrow{S} \cdot \overleftrightarrow{M}$.

Mathematics says the maximum number of non-zero singular values of the matrix $\overleftrightarrow{M}^\dagger \cdot \overleftrightarrow{S} \cdot \overleftrightarrow{M}$ is the number of control coils, so k has to be less than or equal to that number. As k is increased two problems can occur: (1) The magnitude of the currents in the control coils can reach an unacceptable value given a set of expected field errors. (2) The matrix $\overleftrightarrow{\mathcal{C}}_k$, which generally has k non-zero singular values $\mathcal{C}_i$, will have too large a ratio between its largest and smallest non-zero singular values. This ratio, called the condition number, defines the accuracy with which the currents in the control coils must be specified to prevent a field error associated with a large $\mathcal{C}_i$ from preventing control of a field error associated with a small $\mathcal{C}_i$. These two problems determine the optimal k, which is the effective number of control coils, and help determine the optimal locations for the error field control coils.

The effective transfer matrix, $\vec{\Phi}_x = \overleftrightarrow{\mathcal{T}}_k \cdot \delta\vec{\Phi}_m$, is defined assuming the control currents have their optimal values, $\overleftrightarrow{\mathcal{T}}_k \equiv \overleftrightarrow{T} - \overleftrightarrow{M}.\overleftrightarrow{\mathcal{C}}_k$. With optimal error field correction, the remaining effect on the plasma $\vec{\Phi}_x^\dagger \cdot \overleftrightarrow{S} \cdot \vec{\Phi}_x = \delta\vec{\Phi}_m^\dagger \cdot \overleftrightarrow{\mathcal{R}}_k \cdot \delta\vec{\Phi}_m$, where

$$\overleftrightarrow{\mathcal{R}}_k \equiv \overleftrightarrow{\mathcal{T}}_k^\dagger \cdot \overleftrightarrow{S} \cdot \overleftrightarrow{\mathcal{T}}_k. \tag{29}$$

Knowledge of $\overleftrightarrow{\mathcal{R}}_k$ allows one to determine what are the worst error fields, study the adequacy of a given set of control coils, and determine how the engineering trade-off should be made between extra control coils and reduced construction tolerances. Since $\overleftrightarrow{\mathcal{R}}_k$ is Hermitian it can be diagonalized with real positive eigenvalues $\mathcal{R}_i$. For a given set of control coils, the worst machine error is associated with the largest $\mathcal{R}_i$. The control coils should be optimized to make this $\mathcal{R}_i$ as small as possible.

The i^{th} singular value of the matrix $\overleftrightarrow{T}$ becomes exponentially small as i is increased due to the decay of curl-free magnetic fields through space. A curl-free magnetic field decays through space as $\exp(-Kx)$, where $K^2 \sim (m/a)^2 + (n/R)^2$, the poloidal and toroidal mode numbers are m and n, and the major and minor radius are R and a. When there are a large number of control coils, the singular values of $\overleftrightarrow{M}$ also decrease exponentially though at lower rate if the control coils are closer to the plasma.

If the control coils are closer to the plasma than the source of the errors, an arbitrarily large number of error field distributions can in principle be controlled if there are more control coils than error field distributions. However, if the source of the errors is closer to the plasma than the control coils, the difficulty of controlling error field distributions increases exponentially with their number. Using control coils to drive particular magnetic fields on the plasma boundary $\vec{B}_x \cdot \hat{n}$ for beneficial effects, as in the control of Edge Localized Modes (ELM's) [20], while not driving perturbations that degrade the plasma performance, becomes exponentially more difficult the further back the coils. The exponential drop in the singular values of $\overleftrightarrow{M}$ implies the control currents must be both exponentially larger and specified with exponentially greater accuracy the further back the coils.

IV. LIMITS FROM KINETIC THEORY

Kinetic theory is the basis of most plasma models, but definite results can be both complicated and involve questionable approximations. However, certain limits on plasma behavior can be simply obtained.

One simple but important limit, is the minimum power required to drive a current [21]. The current in steady-state current drive is most efficiently driven by waves interacting with mildly relativistic electrons of kinetic energy $(\gamma - 1)m_e c^2$. The number of relativistic electrons required to carry the driven current is $2\pi R_0 I_d/ec$, where R_0 is the major radius. These electrons slow down on the background electrons, so a power

$$P_d > 2\pi R_0(\gamma - 1)m_e c^2 (I_d/ec)\nu_{ee}(\gamma) \tag{30}$$

is required to offset the slowing down. The electron-electron collision frequency $\nu_{ee}(\gamma)$ is proportional to the background electron density n_e. The quantity $(\gamma - 1)\nu_{ee}(\gamma)$ has a minimum as a function of γ, which defines the minimum power required to maintain a driven current. The power per unit volume p_d required to maintain a current is proportional to the current of density j that is driven. The quantity $\mathcal{E}_d \equiv p_d/j$ has units of volts per meter. For electrons at $\gamma \approx 2$, which is the most efficient energy for current drive, $\mathcal{E}_d \gtrsim \mathcal{E}_r$ where

$$\mathcal{E}_r \equiv \frac{e\ln(\Lambda)}{4\pi\epsilon_0(c/\omega_{pe})^2} \simeq \left(0.087\frac{Volts}{meter}\right)\left(\frac{n}{10^{20}\frac{1}{m^3}}\right), \tag{31}$$

the background electron density is n, the electron plasma frequency is $\omega_{pe} \equiv \sqrt{ne^2/\epsilon_0 m_e}$, and the Coulomb logarithm is $\ln(\Lambda) \approx 17$.

Boltzmann and Gibbs, [22] determined the kinetic expression for entropy per unit volume,

$$s(\vec{x}, t) = -\int f\ln(f)d^3v, \tag{32}$$

and a number of important constraints can be obtained from the rate entropy is produced by collisions,

$$\dot{s}_c \equiv -\int (1 + \ln f)C(f)d^3v, \tag{33}$$

which must be positive to be consistent with the second law of thermodynamics. The rate of collisional entropy production per unit volume is [14]

$$\dot{s}_c = \vec{\mathcal{F}}_\epsilon \cdot \vec{\nabla}\frac{1}{T} - \vec{\Gamma}\cdot\vec{\nabla}\frac{\mu}{T} \geq 0, \tag{34}$$

where $\vec{\mathcal{F}}_\epsilon$ is the energy flux $\partial\epsilon/\partial t + \vec{\nabla}\cdot\vec{\mathcal{F}}_\epsilon$ with $\epsilon = \int (mv^2/2) f d^3v$ the energy density, $\vec{\Gamma}$ the particle flux with $\partial n/\partial t + \vec{\nabla}\cdot\vec{\Gamma} = 0$,

$$\frac{\mu}{T} = c_0 + \ln\left(\frac{n}{T^{3/2}}\right) \tag{35}$$

the chemical potential μ divided by the temperature, and c_0 a constant,.

Equation (34) essentially follows from the fundamental thermodynamic relation $dU = TdS - pdV + \mu dN$, where U is the total energy in a system, T the temperature, S the entropy, p the pressure, V the volume, μ the chemical potential and N the number of particles. In a plasma that is large compared to the Debye length, the thermodynamic properties become independent of the overall volume of the plasma; $\epsilon = U/V$, $s = S/V$, and $n = N/V$ are independent of the volume of the plasma. The chain rule of calculus implies $(d\epsilon - Tds - \mu dn)V = -(\epsilon - Ts + p - \mu n)dV$. Volume independence implies both sides must be zero. That is

$$d\epsilon = Tds + \mu dn; \tag{36}$$

$$\frac{\mu}{T} = \frac{(\epsilon + p) - Ts}{nT} \tag{37}$$

Equation (35) for μ/T is obtained from Equation (37) using the expression for the entropy density of a local Maxwellian given by Equation (32). Equation (36) implies $T\partial s/\partial t = \partial\epsilon/\partial t - \mu\partial n/\partial t$. Energy and particle conservation, $\partial\epsilon/\partial t + \vec{\nabla}\cdot\vec{\mathcal{F}}_\epsilon$ and $\partial n/\partial t + \vec{\nabla}\cdot\vec{\Gamma} = 0$, then give

$$\frac{\partial s}{\partial t} + \vec{\nabla}\cdot\vec{\mathcal{F}}_s = \vec{\mathcal{F}}_\epsilon\cdot\vec{\nabla}\frac{1}{T} - \vec{\Gamma}\cdot\vec{\nabla}\frac{\mu}{T} \tag{38}$$

The entropy flux is $\vec{\mathcal{F}}_s = \vec{\mathcal{F}}_\epsilon/T - (\mu/T)\Gamma$. The rate of entropy creation per unit volume is $\vec{\mathcal{F}}_\epsilon\cdot\vec{\nabla}(1/T) - \vec{\Gamma}\cdot\vec{\nabla}(\mu/T)$, which is $\dot{s}_c$ in kinetic theory, Eq. (33).

Equation (34) implies the entropy production due to collisions $\dot{S}_c(\psi_t) \equiv \int \dot{s}_c d^3x$ in a region enclosed by a magnetic surface containing toroidal flux ψ_t obeys [23], [24]

$$\frac{d\dot{S}_c}{d\psi_t} = -\left\{\frac{\mathcal{F}_\epsilon}{T} - \left(\frac{3}{2} - \frac{1}{\eta}\right)\Gamma\right\}\frac{d\ln T}{d\psi_t} \geq 0, \tag{39}$$

where $\eta \equiv d\ln T/d\ln n$, the flow of energy (Joules per second) across a ψ_t surface is $\mathcal{F}_\epsilon \equiv \oint \vec{\mathcal{F}}_\epsilon\cdot d\vec{a}$, and the flow of particles is $\Gamma \equiv \oint \vec{\Gamma}\cdot d\vec{a}$.

Equation (39) gives a limit on an inward particle pinch $\Gamma dn/d\psi_t > 0$. However, a pinch Γ that is comparable to $\mathcal{F}_\epsilon/T$ should not be surprising due to thermodynamic cross terms [24]. When $\eta = 2/3$, the particle flux Γ has no limit from entropy production. A change at fixed entropy is called adiabatic and obeys $p \propto n^\gamma$, where $\gamma = 5/3$, so adiabatic changes have $T \propto n^{\gamma-1}$.

Equation (39) also gives a limit on the deviation of the distribution functions from a local Maxwellian [23]. If the distribution function is written as $f = f_M \exp(\hat{f})$, where f_M is a local Maxwellian, then the approximate form for the limit on the deviation is $\hat{f}^2 \lesssim 1/\nu\tau_E$, where ν is the collision frequency and τ_E is the energy confinement time of the plasma.

V. IMPLICATIONS FOR MAGNETIC FUSION

The perception of realistic scientific options guides research strategy in an applied program such as the quest for magnetic fusion energy. Specific plasma models can be used in statements of the *if-then* form to define these options. As examples, this section identifies two major missing elements in the world magnetic fusion program.

The first example of an *if-then* statement providing guidance on a missing research element comes from the theory of plasma force balance. If $\vec{\nabla}p = \vec{j}\times\vec{B}$, then a solution is specified by the shape of the plasma boundary and the pressure and safety factor profiles [25], [14]. In fusion plasmas, pressure and the current profile are expected to be largely self determined, so shape becomes the primary determinant of the plasma equilibrium.

Engineering and physics must be consistent in a successful fusion reactor. If the physics produced by a line of experiments does not appear to give consistency, then causality says the outcome can only be modified by changing the input parameters. For magnetic fusion systems, shape is the primary freedom to control the plasma equilibrium, and most of that freedom is in non-axisymmetric shaping. Non-axisymmetric shaping can be applied to tokamaks if the constraint of quasi-axisymmetry is maintained on the magnetic field strength, $B(\ell) = B(\ell + L)$, where ℓ is the distance along field lines and L is a field-line constant. The standard view is that the plasma in a tokamak reactor will be in a self-organized microturbulent state with little effective external control. This is a high risk design choice and not a requirement [26], [27]. Non-axisymmetric fields may be used to: (a) form a cage around a plasma, which makes it robust against disruptions and eliminates the need for technically demanding feedback systems, (b) allow the plasma edge to have a high density and a low temperature plasma, which would greatly ease the interface between plasma and the wall, (c) maintain the magnetic configuration. The allowable ratio of the driven to the bootstrap current in a fusion reactor is several times smaller than that expected [28] in ITER, so the determination of the sustainability of the magnetic configuration of an axisymmetric tokamak must await an experiment beyond ITER. Quasi-axisymmetry appears central to reducing the risks and time scale of tokamak development and to its optimization as a power plant, but no program of theory and experiments on quasi-axisymmetry is being pursued

The second, but less mathematical, example of an *if-then* statement identifying a missing research element uses a principle, which also follows from causality: a large modification in the input parameters of an experiment may change the outcome. Tokamak confinement is largely determined by a temperature pedestal at the plasma edge [29]. If this is true, then the properties of the pedestal are probably affected by edge recycling, which is exponentially dependent on wall temperature [30]. No experimental program exists to study tokamak confinement with a fusion relevant wall temperature, which thermodynamic efficiency and material issues imply is in the range of about 600^oC to 1000^oC.

VI. DISCUSSION

This paper will have served its purpose if it motivates a few individuals to think more broadly on the theory of plasmas that just highly integrated computations.

Even within the area of integrated computations, a different way of thinking has implications. Integrated computations are circumscribed by two concepts: validation, which means consistency with experiments, and verification, which means consistency with the equations [31]. However, the philosopher of science, Karl Popper (1902-1994), has noted experiments can only invalidate a theory, which makes a validated code an oxymoron. Of the two concepts, verification seems the more important. When a code, consistent with a set of equations, is compared to a data set: Agreement implies those equations are adequate for that data set. Disagreement implies equations are missing or inappropriate. Either advances science. Only when a code is consistent with a set of known equations should our confidence grow as it is found consistent with a wider and wider data set.

When dealing with plasma equations, which are not universal in application, knowning the minimum set of equations to obtain consistency with data implies what is not constrained. What is not constrained is available for invention. How are constraints and possibilities obtained? (1) Focus on what makes a difference. (2) Look for paradoxes. (3) Remember that science advances through the posing of questions that (a) make a difference, (b) have not been answered, but (c) can be answered. (4) When a critical question can not be answered, ask whether anything can be said.

Whether in the laboratory or in space, mathematics and Maxwell's equations delineate directions for plasma research that can outweigh decades of accepted wisdom and thousands of hours of computer time.

Acknowledgments

My thanks for the Alfvén prize. Over my career my European colleagues have been very kind to me. I owe thanks to physicists who are no longer with us. Edwin Salpeter, my graduate advisor, imparted the importance of physics constraints, Morton Levine introduced me to plasma physics and imparted the importance of physical reasoning, as did another physicist who benefited from Mort's tutelage, Harold Furth. Other physicists, Arnulf Schlüter and Karl Lackner did much to strengthen my ties with Europe. Arnulf Schlüter, Jürgen Nührenberg, and Peter Merkel gave me important lessons on optimization. My thanks to my family. My career in plasma physics would not have been possible without the long-term support of U.S. Department of Energy; for this work grant ER5433 to Columbia University.

[1] A. H. Boozer, Phys. Plasmas **17**, 072503 (2010).
[2] M. Yamada, R. Kulsrud, and H. T. Ji, Rev. Mod. Phys. **82**, 603 (2010).
[3] *Reconnection of Magnetic Fields: Magnetohydrodynamics and Collisionless Theory and Observations* (J. Birn and E. R. Priest editors, Cambridge University Press, 2007).
[4] E. Priest and T. Forbes, *Magnetic Reconnection: MHD Theory and Applications*, (Cambridge University Press, 2007).
[5] D. Biskamp, *Magnetic Reconnection in Plasmas* (Cambridge University Press, 2005).
[6] J. W. Dungey in the *Cosmic Electrodynamics*, (Cambridge University Press, New York 1958) p. 99.
[7] H. P. Furth, J. Killeen, and M. N. Rosenbluth, Phys. Fluids **6**, 459 (1963).
[8] E. G. Harris, Nuovo Cimento, **23**, 115 (1962).
[9] A. H. Boozer, Phys. Rev. Lett. **88**, 215005 (2002).
[10] J. M. Greene, Phys. Fluids B **5**, 2355 (1993).
[11] A. H. Boozer, Phys. Plasmas **12**, 070706 (2005).
[12] W. A. Newcomb, Ann. Phys. (NY) **3**, 347 (1958).
[13] D. P. Stern, Am. J. Phys. **38**, 494 (1970).
[14] A. H. Boozer, Rev. Mod. Phys. **76**, 1071 (2004).
[15] C. R. Graham and F. S. Henyey, Phys. Fluids **12**, 744 (2000).
[16] T.C. Hender, J.C Wesley, J. Bialek, A. Bondeson, A.H. Boozer, R.J. Buttery, A. Garofalo, T.P Goodman, R.S. Granetz, Y. Gribov, O. Gruber, M. Gryaznevich, G. Giruzzi, S. Gnter, N. Hayashi, P. Helander, C.C. Hegna, D.F. Howell, D.A. Humphreys, G.T.A. Huysmans, A.W. Hyatt, A. Isayama, S.C. Jardin, Y. Kawano, A. Kellman, C. Kessel, H.R. Koslowski, R.J. La Haye, E. Lazzaro, Y.Q. Liu, V. Lukash, J. Manickam, S. Medvedev, V. Mertens, S.V. Mirnov, Y. Nakamura, G. Navratil, M. Okabayashi, T. Ozeki, R. Paccagnella, G. Pautasso, F. Porcelli, V.D. Pustovitov, V. Riccardo, M. Sato, O. Sauter, M.J. Schaffer, M. Shimada, P. Sonato, E.J. Strait, M. Sugihara, M. Takechi, A.D. Turnbull, E. Westerhof, D.G. Whyte, R. Yoshino, H. Zohm and the ITPA MHD, Disruption and Magnetic Control Topical Group, Nucl. Fu-

sion **47**, S128 (2007).
[17] J-K Park, M. J. Schaffer, J. E. Menard, and A. H. Boozer, Phys. Rev. Lett. **99**, 195003 (2007).
[18] V. Amoskov, A. Belov, V. Belyakov, O. Filatov, Yu. Gribov, E. Lamzin, N. Maximenkova, B. Mingalev, and S. Sytchevsky, Plasma Devices and Operations **12**, 285 (2004).
[19] J.-K. Park, A. H. Boozer, J. E. Menard, and M. J. Schaffer, Nucl. Fusion **48**, 045006 (2008).
[20] O. Schmitz, T. E. Evans, M. E. Fenstermacher, E. A. Unterberg, M. E. Austin, B. D. Bray, N. H. Brooks, H. Frerichs, M. Groth, M. W. Jakubowski, C. J. Lasnier, M. Lehnen, A. W. Leonard, S. Mordijck, R. A. Moyer, T. H. Osborne, D. Reiter, U. Samm, M. J. Schaffer, B. Unterberg, W. P. West, and the DIII-D Research Team, and the TEXTOR Research Team, Phys. Rev. Lett. **103**, 165005 (2009).
[21] A. H. Boozer, Phys. Fluids **31**, 591 (1988).
[22] E. T. Jaynes, American Journal of Physics, **33**, 391 (1965).
[23] A. H. Boozer, Phys. Fluids **24**, 1382 (1988).
[24] A. H. Boozer, Phys. Fluids B **4**, 2845 (1992).
[25] F. Bauer, O. Betancourt, and P. Garabedian, *Magnetohydrodynamic Equilibrium and Stability of Stellarators*, (Springer, New York, 1984).
[26] A. H. Boozer, Plasma Phys. Control. Fusion **50**, 124005 (2008).
[27] A. H. Boozer, Phys. Plasmas **16**, 058102 (2009).
[28] V. Mukhovatov, M. Shimada, K. Lackner, D.J. Campbell, N.A. Uckan, J.C. Wesley, T.C. Hender, B. Lipschultz, A. Loarte, R.D. Stambaugh, R.J. Goldston, Y. Shimomura, M. Fujiwara, M. Nagami, V.D. Pustovitov, H. Zohm, ITPA CC Members, ITPA Topical Group Chairs and Co-Chairs and the ITER International Team, Nuclear Fusion **47**, S404 (2007).
[29] E.J. Doyle, W.A. Houlberg, Y. Kamada, V. Mukhovatov, T.H. Osborne, A. Polevoi, G. Bateman, J.W. Connor, J.G. Cordey, T. Fujita, X. Garbet, T.S. Hahm, L.D. Horton, A.E. Hubbard, F. Imbeaux, F. Jenko, J.E. Kinsey, Y. Kishimoto, J. Li, T.C. Luce, Y. Martin, M. Ossipenko, V. Parail, A. Peeters, T.L. Rhodes, J.E. Rice, C.M. Roach, V. Rozhansky, F. Ryter, G. Saibene, R. Sartori, A.C.C. Sips, J.A. Snipes, M. Sugihara, E.J. Synakowski, H. Takenaga, T. Takizuka, K. Thomsen, M.R. Wade, H.R. Wilson, ITPA Transport Physics Topical Group, ITPA Confinement Database and Modelling Topical Group, and ITPA Pedestal and Edge Topical Group, Nucl. Fusion **47**, S18 (2007).
[30] D. G. Whyte, Journal of Nuclear Materials, **390-391**, 911 (2009).
[31] M. Greenwald, Phys. Plasmas **17**, 058101 (2010).