NAMT
94-032

A Statistical Equilibrium Model of Coherent Structures in Magnetohydrodynamics

Richard Jordan
Carnegie Mellon University

Research Report No.94-NA-032

October 1994

Sponsors

U.S. Army Research Office
Research Triangle Park
NC 27709

National Science Foundation
1800 G Street, N.W.
Washington, DC 20550

A Statistical Equilibrium Model of Coherent Structures in Magnetohydrodynamics

by
Richard Jordan*
Department of Mathematics and Center for Nonlinear Analysis,
Carnegie Mellon University,
Pittsburgh, PA 15213-3890.

Abstract

A statistical equilibrium theory is developed which characterizes the large-scale coherent structures that emerge during the course of the evolution of an ideal two-dimensional magnetofluid. Macrostates are defined to be local joint probability distributions, or Young measures, on the values of the fluctuating magnetic field and velocity field at each point in the spatial domain. The most probable macrostate is found by maximizing a Kullback-Liebler entropy functional subject to constraints dictated by the conserved integrals of the ideal dynamics. This maximum entropy macrostate is, for each point in the spatial domain, a Gaussian probability distribution, whose local mean is an exact stationary solution of the evolution equations of the magnetohydrodynamic system. The predictions of the statistical equilibrium model are found to be in excellent qualitative and quantitative agreement with recent high resolution numerical simulations of turbulence in slightly dissipative two-dimensional magnetofluids.

* Research supported by the ARO and the NSF through grants to the Center for Nonlinear Analysis.

1. INTRODUCTION

1.1 Features of Two-Dimensional Magnetodydrodynamic Turbulence

The turbulent behavior of a two-dimensional magnetofluid is a fascinating phenomenon that has attracted the interests of theoreticians and experimentalists for many decades. Magnetohydrodynamic (MHD) turbulence in plasmas occurs in astrophysical systems, such as the solar wind [9, 37, 48], and in devices developed for controlled nuclear fusion research [17]. It is believed that certain large-scale features of solar dynamics can be modeled by the equations of two-dimensional MHD [36, 37, 48]. It has also been argued that, in strongly magnetized situations, MHD turbulence becomes essentially two-dimensional with increasing field strength [18, 43].

Apart from the relevance of MHD to problems of plasma physics and astrophysics, two-dimensional MHD turbulence is a particularly nice prototype for the general turbulence problem. High-resolution numerical simulations are available for slightly dissipative two-dimensional magnetofluids [7, 9, 10, 11]. These simulations clearly display the dominant features of the turbulence - namely, the generation of small-scale fluctuations, and the emergence and persistence of large-scale coherent structures amidst the turbulent fluctuations. It is expected that a turbulent three-dimensional magnetofluid will also display these generic features. Unlike nonmagnetic fluid dynamics, a two-dimensional magnetofluid produces energetic small scales, which makes the difference between two-dimensional and three-dimensional MHD turbulence less severe than that for ordinary hydrodynamic turbulence [24]. In addition, a two-dimensional magnetofluid exhibits a rich set of phenomena not found in ordinary hydrodynamics, which challenge any proposed model of two-dimensional MHD turbulence. For example, such a model must account for correlations between the magnetic field and the velocity field, and must predict how energy is divided among its magnetic and kinetic parts.

1.2 An Overview of Our Approach

In this work, we develop a model that predicts the properties of the long-lived organized macroscopic structures which emerge during the course of the evolution of an ideal (or a slightly dissipative) two-dimensional magnetofluid. These coherent structures are naturally modeled by appealing to the methods of equilibrium statistical mechanics. Our approach is information-theoretic in spirit [3, 26, 30]. In characterizing the relaxation of the magnetofluid into a coherent state, we appeal to the general principle that entropy is to be maximized subject to constraints imposed by the underlying dynamics. These constraints are dictated by the global conservation of energy, cross-helicity, and flux under the evolution of an ideal two-dimensional magnetofluid. By solving this constrained maximum entropy problem, we obtain a most probable macrostate, which quantifies the macroscopic mean field-flow, as well as the small-scale fluctuations that are present in the turbulent relaxed state.

The essence of our method is to introduce a macroscopic description of the MHD system, which only partially captures the highly complicated small-scale behavior of the microscopic field-flow state. A macrostate is taken to be a local joint probability distribution, or Young measure [14, 41], on the values of the magnetic field and velocity field at each point x in the spatial domain D. Such a description measures the fluctuations of the field and the flow in an infinitesimal neighborhood of any point x in D.

There is a natural description of the system by such a macrostate for any finite period of time. Indeed, if $Y(x,t)=(B(x,t),V(x,t))$ denotes the field-flow state of the system (B is the magnetic field and V is the velocity field), then the Dirac mass $p_x^t(dy)=\delta_{Y(x,t)}(dy)$ provides such a description. The final turbulent relaxed state is then conceptualized as a possible weak limit, as $t\to\infty$, of the trivial macrostates p^t. Constraints on the admissible macrostates are derived from the global conserved quantities of the ideal MHD system in a manner that is consistent with the weak convergence. The most probable admissible macrostate is determined by maximizing a Kullback-Liebler type entropy functional subject to these constraints.

The most probable macrostate is, for each $x \epsilon D$, a Gaussian probability distribution on the values of $B(x)$ and $V(x)$, with local mean field-flow $\overline{Y}(x)=(\overline{B}(x),\overline{V}(x))$, which is an exact stationary solution of the evolution equations of the ideal MHD system. This leads us to conclude, therefore, that an initial state $Y^{\circ}(x)$ will eventually evolve, under the ideal dynamics, into a stationary coherent structure (the mean field-flow) with Gaussian local fluctuations about this coherent structure. We find that both the fluctuations and the mean field-flow contribute to the total energy and to the cross-helicity, while only the mean field contributes to the flux integrals.

Our model for coherent structures in two-dimensional MHD turbulence is largely motivated by the recent statistical equilibrium theories of Robert *et al.* [33, 41, 42] and Miller *et al.* [34, 35] for two-dimensional hydrodynamic turbulence. These authors introduced the use of local probability distributions to encode the small-scale fluctuations of the turbulent scalar vorticity. In their theories, a macrostate is taken to be a local probability distribution, or Young measure, on the values of the vorticity field at each point in the spatial domain. Robert and Miller also showed how the invariance of the energy and of the generalized enstrophy integrals under the Euler dynamics translate into corresponding constraints on the macrostates.

Robert has characterized most probable macrostates as maximizers of an appropriate Kullback-Liebler entropy functional, and has made great strides toward a rigorous mathematical justification of the model within the framework of the theory of large deviations [33, 40, 41, 42]. In addition, many predictions of the Robert-Miller theory have been confirmed numerically [46, 52, 55] and experimentally [45].

We have attempted throughout to justify, both mathematically and heuristically, our model of two-dimensional MHD turbulence, although we recognize that there is considerable work remaining along those lines. Perhaps the most convincing argument in favor of the model, however, is its remarkable agreement with the recent direct high-resolution numerical simulations of Biskamp *et al.* [7, 9, 10, 11] of slightly dissipative two-dimensional magnetofluids. The comparison of our predictions with theirs is made in Section 4.5.

1.3 The Gibbs Ensemble Theory for MHD Turbulence.

The classical approach to a statistical equilibrium theory for hydrodynamics or magnetohydrodynamics is based upon the canonical Gibbs ensemble for a truncated spectral representation of the full system of equations [18, 19, 28, 29, 44]. Such theories have led to some very interesting and useful qualitative predictions about the nature of turbulence in fluids and magnetofluids. For example, Fyfe and Montgomery [19] have used the canonical ensemble theory for two-dimensional MHD turbulence to predict an approximate equipartition of magnetic energy and kinetic energy contributions from the shortest wavelengths. This prediction will be seen to have a counterpart in our theory. Fyfe and Montgomery [19] were also able to use this theory to demonstrate that, as the number of modes in the spectral representation is increased at fixed values of energy, (quadratic) cross-helicity, and (quadratic) flux, the flux spectral density becomes more and more dominated by the lowest wavenumbers. This observation led them to conclude that, in turbulent two-dimensional MHD, flux is cascaded to large scales, resulting in the formation of macroscopic coherent magnetic structures.

While the qualitative predictions of the classical canonical ensemble approach to two-dimensional MHD turbulence appear to be quite sound, the theory suffers from at least two major limitations. First, the truncation to a finite number of spectral modes destroys all but three of the infinitely many conserved integrals. Only the energy, quadratic cross-helicity, and quadratic flux integral are known to survive such a truncation [19]. Furthermore, the theory yields ensemble-averaged quantities that diverge as the number of modes is taken to infinity. This is the so-called ultraviolet catastrophe [21, 24, 35]. As a result, the classical approach provides very little information about the precise form of the coherent structures.

Very recently, Gruzinov and Isichenko [20, 21, 24] have proposed a statistical equilibrium model for two-dimensional MHD turbulence based upon a Gibbs ensemble theory, which does not suffer from the defects mentioned above for the classical approach. Their theory rests upon the assumption that there exists a meaningful N-dimensional approximation of the two-dimensional ideal MHD system having $N'(N)$ conserved integrals, where both N' and $N-N'$ go to infinity as $N\to\infty$. It is this assumption that enables them to incorporate the complete family of conserved integrals into their statistical equilibrium model. By appropriately rescaling the inverse temperature parameters that arise as multipliers for the conserved integrals, Gruzinov and Isichenko are able to formally obtain a continuum limit in which the values of the ensemble-averaged conserved quantities are equal to their finite prescribed initial values. The key observation is that the inverse temperatures must be allowed to diverge with N in order to carry out the limit.

While the Gruzinov-Isichenko approach and our approach are conceptually different, the predictions of the two models are quite similar. Indeed, the Gruzinov-Isichenko model also predicts the emergence of a coherent stationary mean field-flow amidst Gaussian fluctuations. In their theory, as in ours, the energy and cross-helicity are shared by the mean field-flow and the fluctuations, whereas the flux integrals are determined entirely by the mean field. We feel that this remarkable agreement of the predictions of the two models lends additional credibility to them both.

2. A PRIMER ON INCOMPRESSIBLE MAGNETOHYDRODYNAMICS

2.1 The Equations of Motion

The equations of ideal incompressible magnetohydrodynamics (MHD) in nondimensional variables are [17, 25, 37]

$$\frac{\partial B}{\partial t}=\nabla\times(V\times B), \tag{2.1}$$

$$\frac{\partial V}{\partial t}+(V.\nabla)V=(\nabla\times B)\times B-\nabla p, \tag{2.2}$$

$$\nabla.B=0,\ \nabla.V=0. \tag{2.3}$$

B is the magnetic field, V is the fluid velocity, and p is the pressure. These quantities are appropriately normalized to eliminate physical constants. The incompressible fluid medium is ideal in the sense that the fluid viscosity and the electrical resistivity are taken to be zero.

We are concerned with the two-dimensional form of these equations in a spatial domain $D \subset \mathbb{R}^2$, which is simply connected and bounded, with smooth boundary ∂D. We assume, for simplicity, that ∂D is perfectly conducting, so that the appropriate boundary conditions are [17]

$$B.n=0,\ V.n=0 \ on\ \partial D, \tag{2.4}$$

where n is the outwardly directed normal to the boundary ∂D. The statistical equilibrium theory developed below also applies, with minor modifications, to the case where D is a fundamental period domain corresponding to the periodicity of B and V in x_1 and x_2. Equation (2.3) implies the existence of a magnetic flux function ψ and a velocity stream function φ such that

$$B = curl\,\psi = \left(\frac{\partial\psi}{\partial x_2}, -\frac{\partial\psi}{\partial x_1}\right), \tag{2.5}$$

$$V = curl\,\varphi = \left(\frac{\partial\varphi}{\partial x_2}, -\frac{\partial\varphi}{\partial x_1}\right). \tag{2.6}$$

It is also useful to introduce the scalar current density j and the scalar vorticity ω, which are given by

$$j = Curl\,B, \tag{2.7}$$

$$\omega = Curl\,V. \tag{2.8}$$

Note that the operator *curl* acts on a scalar to produce a vector as in (2.5) or (2.6), whereas the operator *Curl* acts on a vector valued function $B=(B_1, B_2)$ to produce the scalar function

$$Curl\,B = \frac{\partial B_2}{\partial x_1} - \frac{\partial B_1}{\partial x_2}.$$

In terms of the four scalars ψ, φ, j, and ω, the ideal MHD equations can be expressed as

$$\frac{\partial\psi}{\partial t} = \partial(\varphi, \psi), \tag{2.9}$$

$$\frac{\partial\omega}{\partial t} = \partial(\varphi, \omega) + \partial(j, \psi), \tag{2.10}$$

in which

$$\partial(f, g) =: \frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial g}{\partial x_1},$$

is the canonical Poisson bracket in $\mathbb{R}^2$ [22]. The system (2.9)

-(2.10) is completed by the boundary value problems:

$$-\Delta\psi=j \text{ in } D; \quad \psi=0 \text{ on } \partial D,$$
$$-\Delta\varphi=\omega \text{ in } D; \quad \varphi=0 \text{ on } \partial D.$$

This form of the ideal MHD equations is particularly important because of its relation to the Hamiltonian structure of the system [22]. It is therefore useful for deriving properties of solutions, especially those properties which are related to the conserved quantities of the ideal dynamics.

2.2 Invariants of the Motion

A classical solution of the ideal MHD equations conserves energy, flux, and cross-helicity [22, 56]. These conserved functionals are given by, respectively,

$$E=\frac{1}{2}\int_D (B^2+V^2)\,dx, \tag{2.11}$$

$$F_f=\int_D f(\psi)\,dx, \tag{2.12}$$

$$H_f=\int_D \omega f(\psi)\,dx=\int_D V.Bf'(\psi)\,dx. \tag{2.13}$$

The function $f(s)$, for s in the invariant range of the flux function ψ, must satisfy certain regularity conditions, but is otherwise arbitrary. Hence, there is an infinite family of conserved flux integrals, and an infinite family of conserved cross-helicity integrals. That these integrals are constants of the motion follows easily from equations (2.9)-(2.10), the boundary conditions (2.4), properties of the Poisson bracket, and equations (2.5)-(2.6). It is generally accepted in the literature that these are the only conserved quantities, aside from those which may arise from spatial symmetries corresponding to certain domains and boundary conditions.

The conservation of flux places important constraints on the structure of the magnetic field. It is a consequence of equation (2.9) that each magnetic surface $\{\psi=\sigma\}$ moves with the flow. Furthermore, the area within a flux tube, $\int_{\{\psi>\sigma\}}dx$, is a constant of the motion. This follows easily from equation (2.12) with the particular choice $f(s)=I_{\{s>\sigma\}}$, where I_A is the indicator function of the set A. While a given flux tube must preserve its connectivity, it may become highly distorted and convoluted under the evolution of the magnetofluid [17].

The invariance of the generalized cross-helicities, in addition to imposing certain topological relations between the magnetic field and velocity field, implies that the vorticity within a given flux tube is conserved. That is, for any real number σ, the quantity $\int_{\{\psi>\sigma\}}\omega dx$ is conserved by the dynamics [24].

3. THE STATISTICAL EQUILIBRIUM MODEL

3.1 Motivation for the Model

The direct numerical simulations of large Reynolds number two-dimensional MHD recently performed by Biskamp *et al.* [7, 9, 10, 11] clearly demonstrate that the evolution of a slightly dissipative two-dimensional magnetofluid is turbulent. As time proceeds, the magnetic field and velocity field develop fluctuations on finer and finer scales, and the field and flow lines are stretched and folded in a highly convoluted manner. The onset of turbulence is accompanied by the emergence of large-scale coherent structures which persist amidst the sea of fluctuations.

Coherent magnetic structures, for example, survive in regions where the magnetic field is strong, and turbulent current sheets are concentrated in regions of weak magnetic field. The velocity field also exhibits persistent macroscopic organized structures as well as turbulent vorticity sheets localized in regions where the velocity field is weak. Due to the effects of finite dissipation, these organized structures, as well as the fluctuations surrounding them, are observed in the numerical studies to gradually decay. However, it is expected that if the dynamics were truly ideal (i.e., if there were no dissipation), then the mixing would continue indefinitely, and would excite infinitesimally small scales about each point in the spatial domain. Furthermore, the coherent structures which emerge during the evolution would not be dissipated. We expect, therefore, that under the ideal dynamics, the magnetofluid will approach some final turbulent relaxed state consisting of large-scale coherent structures and finite amplitude fluctuations on arbitrarily small scales about each point in the domain. The statistical equilibrium model that is developed in this chapter attempts to predict this final state of turbulent relaxation, given some initial state $(B^\circ(x),V^\circ(x))$.

The existence and regularity of solutions of the evolution equations of ideal MHD do not appear to be well understood. The best results known to us along these lines concern existence and regularity for only a finite time [47]. These theoretical difficulties, as well as the intricate behavior demonstrated by the numerical simulations for the slightly dissipative dynamics, strongly suggest that the evolution equations should be viewed as governing the microscopic state of the ideal MHD system. The microstate $Y(x,t):=(B(x,t),V(x,t))$ then provides a fine-grained description of the system. The pressure is ignored in the state variable Y because it is determined instantaneously in response to the incompressibility condition, $\nabla\cdot V=0$. Unfortunately, this microstate is not very useful for describing the long-time behavior of the system due to its highly complicated behavior and its tendency to develop fluctuations on increasingly fine scales as time proceeds.

In order to gain an understanding of the long-lived large-scale organized structures which emerge during the course of the evolution of the MHD system, we introduce a coarse-grained, or macroscopic, description of the system, which only partially encodes the intricate small-scale behavior of the microscopic state. A macrostate is taken to be a local probability distribution on the values of the microstate $Y(x)$ at each point x in the domain D. We now proceed to define microstates and macrostates more precisely.

3.2 Microscopic and Macroscopic Descriptions of the MHD System

As the ideal MHD system evolves from some initial state, the magnetic field and velocity field may very well develop singularities. However, since the total energy is conserved, we know that $Y(x,t)$ must remain bounded in $L^2(D:\mathbb{R}^6)$ for all t. Consequently, a microstate is defined to be any function $Y(x)\in L^2(D:\mathbb{R}^6)$. At a given point x in the domain D, such a microstate could theoretically take on any value in $\mathbb{R}^6$. We are, therefore, led to define a macrostate as any local probability measure $p_x(dy)$ on $\mathbb{R}^6$ for almost every $x\in D$. We interpret $p_x(dy)$ as the probability that (or frequency with which) Y takes values in dy when sampled at points infinitesimally close to x. Intuitively, for any Borel subset $T \subset \mathbb{R}^6$

$$p_x(T) = \lim_{\ell\to 0} \frac{|x'\in N_\ell(x) : Y(x')\in T|}{|N_\ell(x)|}, \tag{3.1}$$

where $N_\ell(x)$ is a neighborhood of x satisfying diam $N_\ell(x)\leq\ell$.

Equation (3.1) establishes a many-to-one correspondence between microstates and macrostates. The macrostate p_x varies slowly in x, while the microstate Y fluctuates rapidly in x. For any cell dx, over which p is effectively constant, $Y(x)$ behaves like a random variable having distribution p_x. The macrostate has the virtue that it encodes only partially the infinitesimal-scale fluctuations of the microstates, as it ignores the extremely intricate local spatial configurations realized by these fluctuations.

The local probability distributions $p_x(dy)$ are referred to in the literature as Young measures [14, 41], and, as mentioned above, have been employed recently in the statistical equilibrium theories of Robert *et al.* [33, 40, 42] and Miller *et al.* [34, 35] for two-dimensional ideal hydrodynamics. In the context of hydrodynamics, the Young measure $p_x(dy)$ represents a local probability distribution on the values of the fluctuating scalar vorticity $\omega(x)$. In the Robert theory of two-dimensional hydrodynamic turbulence, the most probable macrostate is found by maximizing an appropriate entropy functional subject to constraints on macrostates which are derived from the conserved quantities of the two-dimensional Euler dynamics. In the next section, we demonstrate how the energy, cross-helicity, and flux constraints on the microstates Y translate into corresponding constraints on the macrostates p, and in Section (3.4), we propose a maximum entropy principle for determining the most probable macrostate which satisfies these constraints. First, we give a more mathematically precise definition of the macroscopic states.

We have defined macrostates to be local probability measures or Young measures $p_x(dy)$ on $\mathbf{R}^4$ for almost every $x \in D$. At a given point $x \in D$ we no longer have a well determined value of $Y(x)$, but only a probability distribution on the values ($y \in \mathbf{R}^4$) of $Y(x)$. Technically, a macrostate is a measurable mapping $p: x \to p_x$ from D to the space $M_1(\mathbf{R}^4)$ of Borel probability measures on $\mathbf{R}^4$ endowed with the topology of weak convergence associated with bounded continuous functions [4, 6]. Such a measurable map defines a positive Radon measure on $D \times \mathbf{R}^4$ (which we also denote by p) by the relation

$$\langle p, f \rangle =: \int_{D \times \mathbf{R}^4} f dp = \int_D \left(\int_{\mathbf{R}^4} f(x, y) \, p_x(dy) \right) dx, \qquad (3.2)$$

for all $f \in C_c(D \times \mathbf{R}^4)$, the space of compactly supported continuous functions. In equation (3.2), $dp = dx \otimes p_x(dy)$. For $f \in C_c(D)$, there holds

$$\langle p, f \rangle = \int_D f(x) \, dx.$$

That is, the projection of p on D is Lebesque measure dx. Conversely, given any bounded Radon measure p on $D\times\mathbb{R}^4$ whose projection on D is dx, there is a measurable map $x\to p_x$, unique dx - almost everywhere, such that (3.2) holds [33, 41]. In what follows, we refer to both p, and the corresponding map $x\to p_x$ as the macrostate or the Young measure, and we denote by M the space of Young measures on $D\times\mathbb{R}^4$.

3.3 Constraints on the Macrostates

As discussed in Chapter 2, the energy, cross-helicity, and flux are conserved by the dynamics of ideal two-dimensional MHD, and the conservation of these quantities places important constraints on the microstate of the system. Indeed, since the actual evolution of the field-flow state Y is extremely complicated and impossible to predict after a certain period of time, these constraints furnish, for all practical purposes, the only useful source of information about the state of the system in the long-time limit. We suspect that there are corresponding constraints on the macrostate. Here we concern ourselves only with the following conserved functionals:

The total energy:

$$E=\frac{1}{2}\int_D Y^2dx=\frac{1}{2}\int_D B^2+V^2dx, \tag{3.3}$$

The quadratic cross helicity:

$$H=\int_D B.\,Vdx, \tag{3.4}$$

and an arbitrarily large but finite collection of flux integrals:

$$F_i=\int_D f_i(\psi)\,dx=F_i^{\circ},\, i=1,\ldots,n, \tag{3.5}$$

where the f_i are chosen from some appropriate basis. We may assume as in [15], for example, that $|f_i(s)| \le C|s|^{\ell}$ for some $\ell>0$, and that the f_i' are linearly independent. We point out that it is quite natural from an analytical point of view to interpolate the infinite family of flux constraints on a finite basis in order to surmount some of the mathematical difficulties associated with the continuously infinite family. Due to the smoothness properties of ψ, as opposed to B itself, this interpolation approximates the exact constraints quite accurately [51].

In considering only the quadratic cross-helicity, (3.4), we are simplifying considerably the actual statistical equilibrium problem which incorporates an arbitrary finite collection of cross-helicity integrals (see equation (2.13)). This full problem, which is rather lengthy and highly technical, will be addressed in a future investigation. The simplified problem that we consider here has the virtue that it captures the essence of the effects of magnetic-kinetic correlations, while it yields nicely to a rigorous mathematical treatment.

In order to carry out our program, it is desirable to replace the energy and quadratic cross-helicity functionals by an equivalent pair of conserved quantities, namely $M=E-H$ and $N=E+H$. From (3.3)-(3.4) it is easy to see that these functionals take the forms

$$M=\frac{1}{2}\int_D (B-V)^2 dx, \tag{3.6}$$

$$N=\frac{1}{2}\int_D (B+V)^2 dx. \tag{3.7}$$

It will become clear later why we have made this technical modification. The variables $U=B-V$, $W=B+V$ are the well known Elsasser variables of magnetohydrodynamics [7].

We suppose that some initial state $Y^\circ(x)=(B^\circ(x),V^\circ(x))$ is given, which fixes the values $M^\circ=E^\circ-H^\circ$, $N^\circ=E^\circ+H^\circ$, and $F_i^\circ, i=1,\ldots,n$

of the functionals M, N, and $F_i, i=1,\ldots,n$, respectively. Here, $E^\circ$and H° are the initial values of energy and cross-helicity. In order to determine corresponding constraints on the macrostates, we examine more closely the connections between the microscopic and macroscopic descriptions of the ideal MHD system.

There is an obvious connection between these two levels of description, at least for a finite period of time. For each time t, one may define the trivial macrostate p^t, which, for each $x\epsilon D$, is a Dirac mass at $Y(x,t)$. That is, for any Borel subset $T \subset \mathbb{R}^4$, we have

$$p_x^t(T)=\delta_{Y(x,t)}(T) .$$

The macrostate p^t, in particular, contains all of the information supplied by the invariance of the flux integrals, the energy, and the cross-helicity. In fact, it is easy to see that

$$M^\circ=\frac{1}{2}\int_D(B(x,t)-V(x,t))^2dx=M(p^t), \tag{3.8}$$

$$N^\circ=\frac{1}{2}\int_D(B(x,t)+V(x,t))^2dx=N(p^t), \tag{3.9}$$

$$F_i^\circ=\int_D f_i(\psi(x,t))\,dx=F_i(p^t),\, i=1,\ldots,n \quad , \tag{3.10}$$

where for a Young measure p, we have defined the functionals

$$M(p)=\frac{1}{2}\int_D(\int_{\mathbb{R}^4}(b-v)^2p_x(dy))\,dx=\frac{1}{2}\int_{D\times\mathbb{R}^4}(b-v)^2dp, \tag{3.11}$$

$$N(p)=\frac{1}{2}\int_D(\int_{\mathbb{R}^4}(b+v)^2p_x(dy))\,dx=\frac{1}{2}\int_{D\times\mathbb{R}^4}(b+v)^2dp, \tag{3.12}$$

$$F_i(p)=\int_D f_i(\overline{\psi}(x))\,dx. \tag{3.13}$$

Here, $y=(b,v)$, where $b\in\mathbb{R}^2$ and $v\in\mathbb{R}^2$ run over the ranges of the magnetic field and the velocity field, respectively. In equation (3.10), $\overline{\psi}$ is the mean flux function corresponding to the (local) mean magnetic field $\overline{B}(x)$, defined by the relation

$$\overline{B}(x)=\int_{\mathbb{R}^4} bp_x(dy)\,, \tag{3.14}$$

Some care must be taken in defining $\overline{\psi}$, since for an arbitrary $p\in M$, it need not be the case that $\overline{B}$ is divergence free, nor is it necessarily true that $B\cdot n$ vanish on ∂D. However, we can associate with $\overline{B}$ a unique divergence free field $P_H\overline{B}$ which satisfies $P_H\overline{B}.n|_{\partial D}=0$. P_H is the projection of $L^2(D:\mathbb{R}^2)$ onto the subspace H of $L^2(D:\mathbb{R}^2)$ defined below, which consists of divergence free fields whose normal components vanish on ∂D [49]. The mean flux function $\overline{\psi}$ is then defined by the relation

$$\overline{\psi}=curl^{-1}\, P_H\overline{B}. \tag{3.15}$$

The operator $curl^{-1}$ is a compact operator from the space

$$H=\{U\in L^2(D:\mathbb{R}^2):\nabla\cdot U=0, U\cdot n|_{\partial D}=0\}$$

to the Sobolev space $H_0^1(D)$ [1, 49]. The flux $\overline{\psi}$ defined by (3.15) necessarily satisfies $\overline{\psi}=0$ on ∂D. Of course, for the trivial macrostates P^t, we have $\overline{B}^t(x)=B(x,t)\in H$, and $\overline{\psi}(x)=\psi(x,t)$.

The macroscopic description is really intended to capture the long-time behavior of the MHD system, and is therefore conceptualized as a possible long-time limit of the Dirac masses p^t. The natural mode of convergence for the p^t is that of weak convergence. Indeed, from the conditions $M(p^t)=M^\circ$, $N(P^t)=N^\circ$, it follows easily (see [27]) that the family $\{p^t\}$ of bounded Radon measures on $D\times\mathbb{R}^4$ is tight, so that each sequence contains a subsequence that converges weakly to some bounded Radon measure p on $D\times\mathbb{R}^4$ [5, 6]. As the space of Young measures on $D\times\mathbb{R}^4$ is closed in the space of bounded Radon measures on $D\times\mathbb{R}^4$ with respect to the topology of weak convergence [33, 40], p must itself be a Young measure on $D\times\mathbf{R}^4$. Thus, it is justifiable to model the long-time behavior of the MHD system by a local probability distribution on the values of the fluctuating field-flow variable $Y(x)$ for each $x\epsilon D$.

We can now pose constraints on the macrostates in a manner that is consistent with the weak convergence of subsequences of the p^j. It follows from the analysis of Section 5.1 that if a subsequence of the p^j converges weakly to a macrostate p then p must satisfy

$$M(p)\leq M^0, \tag{3.16}$$

$$N(p)\leq N^0, \tag{3.17}$$

$$F_i(p)=F_i^\circ,\, i=1,\dots,n. \tag{3.18}$$

We remark that if p is a weak limit of some subsequence of $\{p^j\}$, then the mean magnetic field $\overline{B}$ defined by (3.14) corresponding to p necessarily is divergence free, and satisfies $\overline{B}\cdot n|_{\partial D}=0$. However, we will find it mathematicaly convenient to allow arbitrary Young measures that satisfy (3.16)-(3.18) to compete in the variational problem (MEP) (posed in the next section), which characterizes the most probable macrostates. Therefore, we define the admissible set of macrostates,

$$W=\{p\in\mathbf{M}: M(p)\leq M^0,\, N(p)\leq N^0,\, F_i(p)=F_i^\circ,\, i=1,\dots,n\}.$$

The constraints (3.18), which can also be written as $\int_D f_i(\overline{\psi})\,dx=\int_D f_i(\psi^\circ(x))\,dx$, where ψ° is the inital flux function, have the important interpretation that the mean field satisfies the same flux constraints as the initial field. In other words, $\overline{\psi}$ is a rearrangement of the initial flux function ψ° [12]. Since only the mean field, and not fluctuations, contributes to the flux integrals, we might say that mean field theory is exact [33]. This conclusion might also have been reached by noticing that the compact operator $curl^{-1}:B\to\psi$ tends to damp fluctuations, so that there are no fluctuations in the flux function, even though the magnetic field itself exhibits fluctuating behavior.

It is also worth commenting on the constraints (3.16)-(3.17) on macrostates. It may seem a bit troublesome that these constraints are in the form of inequalities rather than equalities. The interpretation is that, under the weak convergence of (subsequences of) the p^t to a possible long-time limit macrostate p, important information about the system could be lost. At any finite time t, the complete information about the conservation of energy and cross-helicity is easily extracted from p^t. Indeed, if we define the energy $E(p)$ and cross-helicity $H(p)$ of a macrostate by the formulas

$$E(p)=\frac{1}{2}\int_D\left(\int_{\mathbf{R}^4}(b^2+v^2)\,p_x(dy)\right)dx=\frac{1}{2}\left(M(p)+N(p)\right), \tag{3.19}$$

$$H(p)=\int_D\left(\int_{\mathbf{R}^4}b\cdot v\,p_x(dy)\right)dx=\frac{1}{2}\left(N(p)-M(p)\right), \tag{3.20}$$

it then follows readily that $E(p^t)=E^\circ$ and $H(p^t)=H^\circ$ for all finite t. However, if all we know about a macrostate p is that it is a weak limit of some sequence of the p^t, then the only information concerning the energy of the system that we can obtain from p is that $E(p)\le E^\circ$. Furthermore, p may tell us nothing at all about the cross-helicity of the system, since $H(p)$ is, a priori, undetermined (except for the necessary inequality $|H(p)|\le E(p)$). The problem, of course, is that the functionals $H(p)$ and $E(p)$ (or $M(p)$ and $N(p)$) are not continuous under the weak convergence, since the functions $g(b,v)=b^2+v^2$ and $h(b,v)=b\cdot v$ are not bounded on $\mathbf{R}^4$. We will see shortly that these potential difficulties are not as severe as they may seem. The most probable macrostate p^* that satisfies the constraints (3.16)-(3.18) actually satisfies $M(p^*)=M^\circ$ and $N(p^*)=N^\circ$, and consequently it satisfies $E(p^*)=E^\circ$ and $H(p^*)=H^\circ$. Therefore, the information provided by the conservation of energy and (quadratic) cross-helicity under the ideal dynamics, as well as the conservation of the flux integrals, can be extracted from p^*. Of course, we must explain the notion of most probable macrostate. That is the topic of the next section.

3.4 Most Probable Macrostates: The Maximum Entropy Principle

Now that we have defined an admissible class of macrostates, namely those which satisfy the constraints (3.16)-(3.18), we seek to determine the macrostate from this class which is most probable, or most likely to be observed. For this purpose, we introduce the Kullback-Liebler entropy of a Young measure p [13, 33, 40, 41, 53]

$$K(p:\pi)=-\int_D\left(\int_{\mathbb{R}^4}\log\frac{dp_x}{d\pi^\circ}(y)\,p_x(dy)\right)dx, \tag{3.21}$$

where $\pi^\circ(dy)$ is a probability measure on $\mathbb{R}^4$, π is the Young measure on $D\times\mathbb{R}^4$ defined by $\pi=dx\otimes\pi^\circ$, and $dp_x/d\pi^\circ$ is the Radon-Nikodym derivative of p_x with respect to π°. If p_x is not absolutely continuous with respect to π° for almost every $x\in D$, or if $|\log dp/d\pi|$ is not in $L^1(p)$, then $K(p:\pi)$ is set equal to $-\infty$. We remark that $p<<\pi$ if and only if $p_x<<\pi^\circ$ for almost every x, and when such is the case, $dp/d\pi=dp_x/d\pi^\circ$ π - almost everywhere [38].

As an integral over y, $K(p:\pi)$ admits either of the standard interpretations as a measure of (the logarithm of) the number of microscopic realizations of p, or as a measure of the uncertainty of p [3]. The functional $I(p:\pi)=-K(p:\pi)$ is the Kullback-Liebler information functional, and can be interpreted as the statistical distance from p to the spatially homogenous macrostate π [13, 30, 40]. The form of $K(p:\pi)$ as an integral over x implies that the local fluctuations at two separated points in D are treated as independent. This implicit assumption reflects the ergodicity of the local mixing of the microscopic field-flow system, and is adopted as a hypothesis of the model. The most probable macrostate (actually, the set of most probable macrostates) is found by maximizing the Kullback-Liebler entropy functional subject to the constraints (3.16)-(3.18), once an appropriate reference measure π has been chosen.

The assumption of ergodicity in the dynamic evolution of the field-flow state $Y(x,t)$ implies that, in the absence of information that constrains spatial variations of the macrostates, the most probable macrostate should be spatially homogeneous. It is easy to see that the constraints $M(p)\leq M^{\circ}$ and $N(p)\leq N^{\circ}$ impose no particular spatial structure on the admissible macrostates. The flux constraints, on the other hand, do impose spatial structure on the macrostates. For these reasons we choose π to be the macrostate whose local distribution is for all $x\epsilon D$

$$\pi^{\circ}(dy)=\frac{|D|^2}{\pi^2 M^{\circ} N^{\circ}}\exp(-\frac{|D|}{2M^{\circ}}(b-v)^2-\frac{|D|}{2N^{\circ}}(b+v)^2)dy. \qquad (3.22)$$

π is the most probable spatially homogenous macrostate in the sense that it maximizes the Boltzmann-Gibbs-Shannon entropy functional [3],

$$K(p)=-\int_D(\int_{\mathbf{R}^4}\log\frac{dp_x}{dy}p_x(dy))\,dx,$$

subject to the constraints $M(p)\leq M^{\circ}$ and $N(p)\leq N^{\circ}$, (see [27]). We point out that π actually satisfies $M(\pi)=M^{\circ}$ and $N(\pi)=N^{\circ}$, and, therefore, $E(\pi)=E^{\circ}$ and $H(\pi)=H^{\circ}$. Thus, in the absence of flux constraints, the most probable macrostate is a spatially homogeneous Gaussian distribution with mean zero, whose energy and cross-helicity are equal to the initial values of these quantities.

Of course, π does not satisfy the flux constraints (3.18). To determine the most probable macrostate which satisfies all of the constraints (3.16)-(3.18), we solve the following constrained entropy maximization problem:

$$K(p:\pi)\rightarrow max, \text{ subject to } p\in W. \qquad \text{(MEP)}$$

We may sometimes indicate the dependence of W and π on the constraint values M° and N° by writing $W=W(M^\circ,N^\circ)$, $\pi=\pi_{M^\circ,N^\circ}$, and $\pi^\circ=\pi^\circ_{M^\circ,N^\circ}$. The maximum entropy principle (MEP) selects those admissible macrostates that minimize the distance to the most probable spatially homogenous macrostate π. Unlike π, however, a macrostate which solves (MEP) exhibits spatial variations characteristic of a coherent structure.

By appealing to Robert's Concentration Theorem for Young measures [33, 40, 41], we can give a more exact meaning to the statement that solutions of (MEP) are most probable. That is done in Chapter 5. There, we also examine the existence and uniqueness of solutions to (MEP). In Chapter 4, the predictions based upon this model of MHD turbulence are discussed and compared with the results of the numerical simulations of Biskamp *et al.* [7, 9, 10, 11].

4. PREDICTIONS OF THE STATISTICAL EQUILIBRIUM MODEL

4.1 Some Existence and Uniqueness Results for Solutions to (MEP)

Our primary goals in this chapter are to present and discuss the predictions of the statistical equilibrium theory which was developed in the previous chapter, and to compare these results with the known behavior of a slightly dissipative two-dimensional magnetofluid as simulated numerically. In this section, we state some results concerning the existence and uniqueness of solutions of the maximum entropy principle (MEP) proposed in Chapter 3. We are mostly concerned with conditions under which solutions with finite entropy exist, since such solutions can be explicitly caluclated, as in Section 4.2, by appealing to the Kuhn-Tucker theory.

Concerning the existence of finite entropy solutions of (MEP), we have the following result.

Theorem 4.1

Assume that:

(a) The functions f_i, and the constraint values F_i°, $i=1,\ldots,n$, are given so that there is at least one critical point $B_c \epsilon H$ of the magnetic energy functional

$$E(B)=\frac{1}{2}\int_D B^2 dx,$$

subject to the flux constraints

$$F_i(B)=\int_D f_i(\psi)\,dx=F_i^\circ, i=1,\ldots,n.$$

In addition, assume that E° and H° satisfy either

(b) $|H^\circ|>2E_m,\quad E^\circ>|H^\circ|,$

or

(c) $|H^\circ|\leq 2E_m,\quad E^\circ>E_m+(H^\circ)^2/4E_m,$

where E_m is the minimum value of $E(B)$ subject to the flux constraints; then there is a solution of the constrained optimization problem,

(MEP) $\qquad K(p:\pi)\rightarrow\max$ over W.

Furthermore, any such solution $p*$ satisfies

$$K(p*:\pi)>-\infty.$$

Some comments are in order concerning conditions (a)-(c) in the theorem. The significance of the hypothesis (a) will become clear in the next section when we appeal to the Lagrange multiplier rule to calculate the solutions of (MEP). It should be noted that any critical point of $E(B)$ subject to the constraints $F_i(B)=F_i^\circ$, $i=1,\ldots,n$ takes the form [22]

$$j=\Sigma\lambda_i f_i'(\psi),$$

or equivalently,

$$B=\Sigma\lambda_i \,\mathrm{curl}\, G f_i'(\psi),$$

where G is the Green operator on D, corresponding to Dirichlet boundary conditions. Thus, condition (a) is equivalent to the condition that there exists a solution $(\lambda,\ldots,\lambda_n,\psi)$, with $\lambda_i \epsilon \mathbf{R}$, and $\psi \epsilon H_0^1(D)$, of either of these equations.

Conditions (b) and (c) both imply that $E^\circ > E_m$, and that $E^\circ > |H^\circ|$, which are quite natural from a physical standpoint. In fact, the initial energy E° must be larger than the minimum energy E_m consistent with the constraints on flux if the initial velocity field is to be nonzero. Also, unless the initial magnetic field B° and initial velocity field V° satisfy $B^\circ = \pm V^\circ$, it follows that $2|B^\circ \cdot V^\circ| < (B^\circ)^2 + (V^\circ)^2$, and, consequently, that $|H^\circ| < E^\circ$. When $|H^\circ|$ is small, $E^\circ > |H^\circ|$ does not necessarily imply $E^\circ > E_m$. This partially explains (c). The limits $E^\circ \to E_m$ *and* $E^\circ \to |H^\circ|$ are interesting to examine, and we will do so later in this chapter.

We should point out that when $|H^\circ| \leq 2E_m$, the condition $E^\circ > E_m + (H^\circ)^2/4E_m$ is not, in general, necessary for the existence of finite entropy solutions of (MEP). It appears to be a difficult problem to determine necessary conditions. This difficulty is closely related to the open problem of classifying all the critical points of the magnetic energy functional $E(B)$ subject to the flux constraints $F_i(B)=F_i^\circ$ for a given domain D. See also the discussion in Section (4.4) for more comments on condition (c). Under certain circumstances, there is a unique critical point $B_c \epsilon H$. When this is the case, we have the following uniqueness result for solutions of (MEP).

Corollary 4.2

If there is a unique critical point of the functional $E(B)$ subject to the constraints $F_i(B)=F_i^\circ, i=1,\dots,n,$ and if either of the conditions (b) or (c) of Theorem 4.1 are met, then there is a unique finite entropy solution of (MEP).

The uniqueness part of Corollary 4.2 will follow from the analysis in Section 4.4. The existence follows, of course, from Theorem 4.1, which is proved in Chapter 5.

4.2 Calculation of the Equilibrium States

Any solution $p*$ of (MEP) under the conditions of Theorem 4.1 has finite entropy, so that $p*=\rho*(x,y)\pi^\circ$, where the density $\rho*(x,y)$ satisfies $\rho*\geq 0$, and

$$\int_{\mathbf{R}^4}\rho*(x,y)\,\pi^\circ(dy)=1,$$

for almost every $x\epsilon D$. For densities $\rho\epsilon L^1(\pi)$, we introduce the notation

$$S(\rho)=K(\rho\pi:\pi)=-\int_D\left(\int_{\mathbf{R}^4}\rho(x,y)\log\rho(x,y)\,\pi^\circ(dy)\right)dx.$$

To avoid introducing any further notation, we will simply write $M(\rho)$, $N(\rho)$, and $F_i(\rho)$ for $M(\rho\pi)$, $N(\rho\pi)$ and $F_i(\rho\pi)$, respectively.

It follows that $\rho*$ maximizes $S(\rho)$ on the subspace of $L^1(\pi)$ defined by the constraints

$$\rho(x,y)\geq 0,\ \int_{\mathbb{R}^4}\rho(x,y)\,\pi^\circ(dy)=1 \quad dx\ a.e., \tag{4.1}$$

$$M(\rho)=\frac{1}{2}\int_D\left(\int_{\mathbb{R}^4}(b-v)^2\rho(x,y)\,\pi^\circ(dy)\right)dx\leq M^\circ, \tag{4.2}$$

$$N(\rho)=\frac{1}{2}\int_D\left(\int_{\mathbb{R}^4}(b+v)^2\rho(x,y)\,\pi^\circ(dy)\right)dx\leq N^\circ, \tag{4.3}$$

$$F_i(\rho)=\int_D f_i(\overline{\Psi})\,dx=F_i^\circ,\ i=1,\dots,n. \tag{4.4}$$

The equilibrium equations for (MEP) then follow from a formal application of the Lagrange Multiplier Rule, or, more precisely, the Kuhn-Tucker Rule [23, 57] (see [27] for a rigorous treatment). These equations are

$$S'(\rho*)=\Sigma\alpha_i F_i'(\rho*)+\sigma M'(\rho*)+\eta N'(\rho*), \tag{4.5}$$

on the subspace of variations $\delta\rho$ which satisfy $\int\delta\rho(x,y)dy=0$. The multipliers α_i, σ, and η are real numbers, and are analogous to the inverse temperatures of usual statistical mechanics. The multipliers σ and η satisfy the side conditions [23, 57]

$$\sigma\geq 0,\ \eta\geq 0, \tag{4.6}$$

$$\sigma(M(\rho*)-M^\circ)=0,\ \eta(N(\rho*)-N^\circ)=0. \tag{4.7}$$

In particular, if $\sigma>0$ and $\eta>0$, then $M(\rho*)=M^\circ$ and $N(\rho*)=N^\circ$, and, as a consequence, $E(\rho*)=E^\circ$ and $H(\rho*)=H^\circ$. We will see shortly that this is indeed the case.

The functional derivatives appearing in (4.5) are easily found to be

$$F_i'(\rho*) = b.\ \mathrm{curl}\,Gf_i'(\overline{\Psi}),$$

$$M'(\rho*) = \frac{1}{2}(b-v)^2,$$

$$N'(\rho*) = \frac{1}{2}(b+v)^2,$$

$$S'(\rho*) = -(1+\log\rho*),$$

where $\overline{\Psi} = \mathrm{curl}^{-1}P_H\overline{B}$ is the mean flux function corresponding to the mean magnetic field

$$\overline{B}(x) = \int_{\mathbb{R}^4} b\rho*(x,y)\,\pi^\circ(dy).$$

These equations, together with (4.5), yield the following form for the equilibrium state:

$$\rho*(x,y) = Z^{-1}\exp\{-\frac{1}{2}\sigma(b-v)^2 - \frac{1}{2}\eta(b+v)^2 - b\cdot\Sigma\alpha_i\,\mathrm{curl}\,Gf_i'(\overline{\Psi})\},$$

where the partition function,

$$Z(x) = \int_{\mathbb{R}^4}\exp\{-\frac{1}{2}\sigma(b-v)^2 - \frac{1}{2}\eta(b+v)^2 - b\cdot\Sigma\alpha_i\,\mathrm{curl}\,Gf_i'(\overline{\Psi})\}\pi^\circ(dy),$$

enforces the normalization constraint (4.1).

After a straightforward calculation, we arrive at the following expression for the local distribution $p_x*(dy) = \rho*(x,y)\pi^\circ(dy)$:

$$p_x*(dy) = \frac{\beta^2-\gamma^2}{4\pi^2}\exp\{-\frac{\beta}{2}(b-\overline{B}(x))^2 - \frac{\beta}{2}(v-\overline{V}(x))^2 - \gamma(b-\overline{B}(x))\cdot(v-\overline{V}(x))\}dy, \tag{4.8}$$

where

$$\beta=\eta+\sigma+|D|(N^{o})^{-1}+|D|(M^{o})^{-1},\ \gamma=\eta-\sigma+|D|(N^{o})^{-1}-|D|(M^{o})^{-1}, \quad (4.9)$$

and the mean field and flow satisfy

$$\overline{B}(x)=\Sigma\lambda_i \, curl\, Gf_i'(\overline{\Psi}(x)),\ \overline{V}(x)=\mu\overline{B}(x), \quad (4.10)$$

with the constants λ_i and μ given by the formulas

$$\mu=-\gamma/\beta,\ \lambda_i=-\alpha_i/\beta(1-\mu^2),\ i=1,\ldots,n. \quad (4.11)$$

Notice that $|\mu|<1$ by definition. Also, notice that $\overline{B}$ and $\overline{V}$ are divergence free, and that the mean field-flow is a stationary solution of the ideal MHD equations. It is a further consequence of (4.10) that $\overline{B}$ is a critical point of the magnetic energy functional $E(B)$ subject to the constraints $F_i(B)=F_i^{\circ}, i=1,...,n,$ which explains the necessity of condition (a) of Theorem 4.1. As a result, there holds

$$E(\overline{B})=:\frac{1}{2}\int_D(\overline{B})^2dx\geq E_m. \quad (5.12)$$

We are now in a position to prove the following essential result concerning solutions of (MEP).

<u>Theorem 4.3</u>

Under the hypotheses (a) and either (b) or (c) of Theorem 4.1, any solution $p*$ of (MEP) satisfies

$$M(p*)=M^{o},\ N(p*)=N^{o},$$

and as a result

$$E(p*)=E^{o},\ H(p*)=H^{o}.$$

Proof:

By the conditions (4.6) and (4.7), it is sufficient to show that $\sigma>0$ and $\eta>0$. Using (4.8) and (4.9) we calculate

$$M(p*)=\frac{1}{2}\int_D(\overline{B}-\overline{V})^2dx+(M^\circ\sigma+|D|)^{-1}|D|M^\circ, \tag{4.13}$$

$$N(p*)=\frac{1}{2}\int_B(\overline{B}+\overline{V})^2dx+(N^\circ\eta+|D|)^{-1}|D|N^\circ. \tag{4.14}$$

If it is assumed that $\sigma=0$, then (4.13), together with the fact that $p*$ is an admissible macrostate, yields

$$M^\circ\geq\frac{1}{2}\int_D(\overline{B}-\overline{V})^2dx+M^\circ,$$

which can hold only if $\overline{B}(x)=\overline{V}(x)$ for almost every $x\epsilon D$. Now, (4.10) and the fact that $|\mu|<1$ imply that this is possible only if $\overline{B}=0$ almost everywhere. However, it is impossible that $\overline{B}=0$ almost everywhere since (4.12) must hold. We have, therefore, arrived at a contradiction; it must be that $\sigma>0$. The proof that $\eta>0$ follows along similar lines using equation (4.14).

■

Recall that the possible macroscopic descriptions p of the system have been conceptualized as long-time weak limits of (subsequences) of the Dirac masses $p^t=\delta_{Y(x,t)}$ and that the constraints $M(p)\leq M^\circ$, $N(p)\leq N^\circ$ and $F_i(p)=F_i^\circ$ on admissible macrostates have been formulated so as to be consistent with the information afforded by the conservation of energy, cross-helicity, and flux under the weak convergence. Theorem 4.3 has the important consequence that the most probable admissible macrostate $p*$ contains all of the information furnished by the conservation of energy, quadratic cross-helicity, and flux, since $E(p*)=E^\circ$, $H(p*)=H^\circ$, and $F_i(p*)=F_i^\circ$. By considering constraints on $M(p)$ and $N(p)$ in our maximum entropy principle (rather than on $E(p)$ and $H(p)$ themselves), we have been able to recapture the full information provided by the invariance of energy and quadratic cross-helicity under the ideal dynamics. As was discussed in Chapter 3, if p^j converges weakly to p, then all we know, offhand, about $E(p)$ is that it cannot exceed E°, and nothing at all can be said about $H(p)$. The constraints $M(p)\leq M^\circ$ and $N(p)\leq N^\circ$, which are valid under weak convergence of p^j to p, contain more information about admissible macrostates than the single constraint $E(p)\leq E^\circ$. The difference turns out to be crucial. Using the complimentary conditions (4.7) in the Kuhn-Tucker Rule, we have been able to conclude that $M(p^*)=M^\circ$, $N(p^*)=N^\circ$ for the entropy maximizer p^*.

4.3 A Further Analysis of Equilibrium States

It is evident from equation (4.8) for the most probable states p^* that, for each $x\epsilon D$, the magnetic field $B(x)$ and the velocity field $V(x)$ have Gaussian distributions with means $\overline{B}(x)$ and $\overline{V}(x)$, respectively, given by (4.10), and with variances and correlations satisfying

$$Var\,B_j(x)=Var\,V_j(x)=\beta^{-1}(1-\mu^2)^{-1},\, j=1,2, \tag{4.15}$$

$$corr\,(B_j(x),V_j(x))=\mu,\;\; j=1,2. \tag{4.16}$$

The other components of $B(x)$ and $V(x)$ are independent (since these components are Gaussian and have correlations equal to 0). We emphasize that the correlations and variances are independent of $x\epsilon D$ only for the simplified problem we consider here, which accounts for only the quadratic cross-helicity constraint. When generalized cross-helicities are incorporated into the theory, the variances and correlation depend on x in a nontrivial fashion. This more general problem will be undertaken fully in a future investigation. See also [50].

While the flux F_i° is contained entirely in the mean field, the energy E° and the cross-helicity H° are shared by the mean field-flow and the fluctuations. Indeed, direct calculations yield the energy and cross-helicity of a macrostate:

$$E^0 = E(p*) = \frac{1}{2}\int_D (\overline{B}^2+\overline{V}^2)\,dx + 2|D|\beta^{-1}(1-\mu^2)^{-1}, \tag{4.17}$$

$$H^0 = H(p*) = \int_D \overline{B}\cdot\overline{V}dx + 2|D|\mu\beta^{-1}(1-\mu^2)^{-1}. \tag{4.18}$$

The integral of the right hand side of (4.17) is the energy of the mean field-flow $\overline{Y}=(\overline{B},\overline{V})$, and the second term is the contribution of the fluctuations to the total energy. In fact, it follows from (4.15) that

$$2|D|\beta^{-1}(1-\mu^2)^{-1} = \frac{1}{2}\int_D \sum_{j=1}^{2} (Var\,B_j(x) + Var\,V_j(x))\,dx.$$

We remind the reader that the variance of a random variable is a natural measure of the expected magnitude of its fluctuations about its mean. In equation (4.18), the first term on the right hand side is evidently the cross-helicity of the mean field-flow. Recalling the formula $cov(U,W)=corr(U,W)\,(VarW)^{1/2}(VarU)^{1/2}$ for the covariance of random variables U and W, we see that the contribution of the fluctuations to the cross-helicity can be expressed as

$$2\mu|D|\beta^{-1}(1-\mu^2)^{-1} = \int_D \sum_{j=1}^{2} cov(B_j(x), V_j(x))\,dx.$$

Using the relation $\overline{V}=\mu\overline{B}$, equations (4.17)-(4.18) can be rewritten as

$$E^0 = (1+\mu^2)E(\overline{B}) + 2|D|\beta^{-1}(1-\mu^2)^{-1}, \tag{4.19}$$

$$H^0 = 2\mu E(\overline{B}) + 2|D|\mu\beta^{-1}(1-\mu^2)^{-1}. \tag{4.20}$$

The contribution of the magnetic field to the total energy,

$$E_{mag} = \frac{1}{2}\int_D \left(\int_{\mathbb{R}^4} b^2 p_x*(dy)\right)dx,$$

satisfies

$$E_{mag}=E(\overline{B})+|D|\beta^{-1}(1-\mu^2)^{-1}, \tag{4.21}$$

and the kinetic energy,

$$E_{kin}=E^{\circ}-E_{mag}=\frac{1}{2}\int_D\left(\int_{\mathbf{R}^4}v^2p_x*(dy)\right)dx,$$

is given by

$$E_{kin}=\mu^2E(\overline{B})+|D|\beta^{-1}(1-\mu^2)^{-1}. \tag{4.22}$$

By virtue of the fact that $|\mu|<1$, we see that, in equilibrium, the ratio of kinetic to magnetic energy is less than one. Since the initial value of this ratio can be arbitrarily large, this prediction about the turbulent relaxed state is an especially useful test of the model. We will have more to say about this in Section 4.5 when we compare the predictions of the model to those of numerical simulations.

It is also interesting to observe from equations (4.21)-(4.22) that there is an equipartition of fluctuating energy between its magnetic and kinetic components. This is reminiscent of the prediction of the classical statistical equilibrium theory for MHD that the smallest scale energy contributions are divided equally into magnetic and kinetic parts [19].

4.4 Complete Determination of the Equilibrium State for a Simple Case

In order to determine the most probable equilibrium states completely, it is necessary to solve for the Lagrange multipliers σ, η, and $\alpha_i, i=1,\ldots,n$. In view of equations (4.9) and (4.11), it is equivalent to solve for the constants β, μ, and $\lambda_i, i=1,\ldots,n$. In the case when there is a unique critical point of the (deterministic) magnetic energy functional $E(B)$ subject to the flux constraints $F_i(B)=F_i^{\circ}, i=1,\ldots,n$, the λ_i are uniquely determined. We shall assume for the remainder of this chapter that this is the case. Accordingly, there holds $E(\overline{B})=E_m$, with E_m defined as in Theorem 4.1 to be the minimum value of the functional $E(B)$ consistent with the constraints on flux. Equations (4.19) and (4.20) become

$$E^{\circ}=(1+\mu^2)E_m+2|D|\beta^{-1}(1-\mu^2)^{-1}, \tag{4.23}$$

$$H^{\circ}=2\mu E_m+2|D|\mu\beta^{-1}(1-\mu^2)^{-1}, \tag{4.24}$$

which provide two simultaneous equations for β and μ. We recall that β must be positive by definition (see equation (4.9)), and also that the correlation μ must satisfy $|\mu|<1$ (see equation (4.11)). However, equations (4.23)-(4.24) furnish us with additional information about the correlation μ. It has the same sign as H°, and must also satisfy

$$|\mu|<min\left(1,\frac{|H^{\circ}|}{2E_m},\left(\frac{E^{\circ}-E_m}{E_m}\right)^{\frac{1}{2}}\right).$$

(since $\beta>0$). In particular, if $H^{\circ}=0$, then $\mu=0$, and β^{-1} is determined uniquely by (4.23). In the next section, we will discuss the zero cross-helicity case in detail.

Assume now that $H^{\circ}\neq 0$. Upon multiplying equation (4.23) by μ and comparing the resulting equation (4.24), we discover the following interesting dependence of the correlation μ on E_m, and on the given values of total energy E° and cross-helicity H°:

$$E_m\mu^3-(E_m+E^{\circ})\mu+H^{\circ}=0. \tag{4.25}$$

The following result, whose proof is elementary, shows that either of the conditions (b), (c) of Theorem 4.1 is sufficient to guarantee a unique solution μ of (4.25) which satisfies all of the properties mentioned following equation (4.24).

Lemma 4.4

(1) If $E^{\circ}>H^{\circ}>0$, then there exists a unique $\mu\in(0,1)$ which satisfies (4.25). If, in addition, E° and H° satisfy $E^{\circ}>E_m+(H^{\circ})^2/4E_m$, then $\mu<min(1,H^{\circ}/2E_m)$.

(2) If $E^\circ > -H^\circ > 0$, then there is a unique $\mu \in (-1,0)$ satisfying (4.25). If, in addition, $E^\circ > E_m + (H^\circ)^2/4E_m$, then $\max(-1, H^\circ/2E_m) < \mu < 0$.

(3) If $H^\circ = 0$, then $\mu = 0$ is the unique solution of (4.25) over the interval [-1,1].

We remark that the condition $E^\circ > E_m + (H^\circ)^2/4E_m$ is equivalent to the condition

$$\frac{|H^\circ|}{2E_m} < \left(\frac{E^\circ - E_m}{E_m}\right)^{\frac{1}{2}}.$$

Also, note that if μ exists which satisfies (4.25), then necessarily

$$(H^\circ - 2\mu E_m) = \mu (E^\circ - 1 + \mu^2) E_m).$$

The lemma guarantees that when either of the conditions (b), (c) of Theorem 4.1 holds, a unique μ exists which satisfies (4.25), $sgn\ \mu = sgn\ H^\circ$, and

$$\mu < min\left(1, \frac{|H^\circ|}{2E_m}, \left(\frac{E^\circ - E_m}{E_m}\right)^{\frac{1}{2}}\right).$$

Therefore $\beta^{-1} > 0$ is given consistently, when $H^\circ \neq 0$, by

$$\beta^{-1} = \frac{(E^\circ - (1+\mu^2) E_m)(1-\mu^2)}{2|D|} = \frac{(H^\circ - 2\mu E_m)(1-\mu^2)}{2\mu |D|}.$$

We thereby have a unique solution pair (μ, β^{-1}) of (4.23)-(4.24), which satisfies all of the necessary conditions as mentioned above.

We now know that if there is a unique critical point of the functional $E(B)$ subject to the constraints $F_i(B) = F_i^\circ, i=1,\ldots,n$, and if either condition (b), (c) of Theorem 4.1 holds, then the statistical equilibrium macrostate p^*, which solves (MEP), is uniquely determined. Consequently, the uniqueness assertion of Corollary 4.2 has been established by the analysis of this section.

4.5 Equilibrium States Without Magnetic-Kinetic Correlations: Comparison with Numerical Simulations

Most of what is known of the behavior of a slightly dissipative two-dimensional magnetofluid comes from direct high-resolution numerical simulations. The best computations of this kind are reported by Biskamp *et al.* [7, 9, 10, 11]. Those authors focus much of their attention on cases where the initial ratio of quadratic cross-helicity to total energy is small. In some of their simulations, this ratio is taken to be as small as 0.1. They observe local Gaussian distributions on the magnetic and velocity fields, and they verify the direct cascade of energy to small scales, and an inverse cascade of flux (with $f(s)=\frac{1}{2}s^2$) to large scales, resulting in the emergence of large-scale coherent magnetic structures amongst the turbulent fluctuations. Most remarkably, however, they find that E_{kin}/E_{mag} relaxes to a value less than one, even when the initial ratio is as large as 25 [10].

As we have already seen, our model also predicts local Gaussian distributions on the field and flow, as well as an equilibrium value of E_{kin}/E_{mag} which is less than one, whatever the initial ratio, and regardless of the initial values of energy and cross-helicity (as long as they satisfy either condition (b) or condition (c) of Theorem 4.1). To gain a more complete understanding of the predictions of the statistical equilibrium model in the regime of small H°/E°, we examine the limit $H^\circ \to 0$. For the sake of clarity, we shall assume that the hypotheses of Corollary 4.2 are met. Then the analysis and the formulas of Section 4.4 are valid.

Fix E° and $f_i, F_i^\circ, i=1,\dots,n$. Given $H^\circ \neq 0$, denote by p^{H° the unique solution of (MEP), and by $\rho^{H^\circ}(x,y)$ its density with respect to Lebesque measure $dx \otimes dy$ on $D \times \mathbb{R}^4$ (p^{H° is given by equation 4.8). From Lemma 4.4 and from equation (4.25) it follows that

$$\frac{|H^\circ|}{E^\circ + E_m} \leq |\mu| \leq \frac{|H^\circ|}{2E_m}.$$

Thus, as $H^\circ \to 0$, the correlation μ converges to 0. Equation (4.23) gives for β the expression

$$\beta = \frac{2|D|}{(1-\mu^2)\,(E^0 - (1+\mu^2)\,E_m)}.$$

We also have $\gamma = -\beta\mu$. Hence, when $H^0 \to 0$, β converges to β_0, where

$$\beta_0 = 2|D|\,(E^0 - E_m)^{-1}, \tag{4.26}$$

and γ converges to zero. As a result, the density ρ^{H^0} converges pointwise on $D\times\mathbb{R}^4$ to the density ρ, where

$$\rho(x,y) = \frac{\beta_0^2}{4\pi^2}\exp\{-\frac{\beta_0}{2}(b-\overline{B}(x))^2 - \frac{\beta_0}{2}v^2\}. \tag{4.27}$$

Scheffe's Theorem [5] implies that p^{H^0} converges weakly to the measure p on $D\times\mathbb{R}^4$ defined by

$$p(A) = \int_A \rho(x,y)\,dxdy,$$

for Borel sets $A\subset D\times\mathbb{R}^4$. The measure p is a Young measure on $D\times\mathbb{R}^4$, whose local distribution p_x has density $\rho(x,y)$ with respect to Lebesque measure dy on $\mathbb{R}^4$.

It is readily verified by the analysis in Sections 4.2 and 4.3, that p is the unique solution of the constrained entropy maximization problem,

$$K(p:\pi_{E^0,E^0}) \to max$$
$$\textit{subject to} \quad p\in W(E^0,E^0).$$

(See the discussion immediately following the statement of (MEP) in Chapter 3 for an explanation of the notation $W(E^0,E^0)$, π_{E^0,E^0}.) In fact, p maximizes $K(p:\pi_{E^0,E^0})$ over the larger set

$$\{p\in M: E(p)\leq E^{\circ}, F_i(p)=F_i^{\circ}, i=1,\ldots,n\}.$$

In this limit of zero cross-helicity, then, we see from (4.27) that $B(x)$ and $V(x)$ have Gaussian distributions for all x, each component having variance

$$\frac{1}{\beta_0}=\frac{E^{\circ}-E_m}{2|D|}. \tag{4.28}$$

However, now $B(x)$ and $V(x)$ are independent for all x, the local mean velocity field $\overline{V}(x)$ is identically zero, and the mean magnetic field $\overline{B}(x)$ is given by (4.10). We also have $E(p)=E^{\circ}$ and, of course, $H(p)=0$.

As before, the flux is contained entirely in the mean field, and the energy splits into mean and fluctuating parts. The formulas for the magnetic and kinetic energy are

$$E_{mag}=E_m+|D|\beta_0^{-1},\ E_{kin}=|D|\beta_0^{-1}. \tag{4.29}$$

It is evident, therefore, that the difference between $E^{\circ}=E_m+2|D|\beta_0^{-1}$, the given initial energy, and E_m, the minimum energy consistent with the given constraints on flux, resides in the local fluctuations, where it is divided equally into magnetic and kinetic parts. Using equations (4.28) and (4.29) we arrive at the particularly simple expression for the ratio of kinetic to magnetic energy

$$\frac{E_{kin}}{E_{mag}}=\frac{E^{\circ}-E_m}{E^{\circ}+E_m}.$$

As before, it follows that this ratio is less than one, which is in accord with the numerical simulations.

Biskamp *et al*. [9, 10, 11] also study the difference between the low energy regime ($E^{\circ}/E_m\approx 1$) and the high energy regime ($E^{\circ}/E_m>>1$), for the case of small cross-helicity. In the low energy regime, the fluctuations are small, while a distinct coherent magnetic structure emerges in the mean field through the process of quasi-static coalescence of flux tubes. The high energy regime is characterized by large fluctuations, extending uniformly over the domain, which obscure the spatial structure of the mean field-flow. In this sense, the high energy regime resembles homogeneous turbulence.

The statistical equilibrium model captures these qualitative properties of the low and high energy regimes. When $E^{\circ} \to E_m$, it follows from (4.28) that the variance β_0^{-1} tends to zero, so that the fluctuations in B and V disappear, and there results a deterministic nonturbulent magnetostatic equilibrium with $B=\overline{B}$ and $V=\overline{V}=0$. As a matter of fact, it is an easy exercise to show that the macrostate p given by (4.27) converges weakly in this limit to a Dirac mass at $(\overline{B}(x),0)$ for all $x \in D$. On the other hand, for large values of E°, equation (4.28) shows that the variance is large, so that there are large fluctuations about the mean field-flow. Also, when $E^{\circ} \to \infty$, the ratio E_{kin}/E_{mag} tends to one, as the turbulent energy is equipartitioned globally between its magnetic and kinetic constituents. Thus, as E° is taken larger and larger, the equilibrium state resembles more and more a state of homogeneous turbulence, as is predicted by the numerical simulations of Biskamp *et al.* [9, 10, 11].

4.6 Equilibrium States With Aligned Field and Flow

Having already discussed the zero cross-helicity limit, we now look at what happens as $|H^{\circ}|/E^{\circ} \to 1$, its largest possible value. It is the regime of large $|H^{\circ}|/E^{\circ}$ in which the effects of magnetic-kinetic correlation are most vivid. To illustrate this limit, we assume that $E^{\circ} > 2E_m$ is fixed, and allow H° to approach the fixed value E°. It is a result of equation (4.25) that the correlation μ must converge to one as $H^{\circ} \to E^{\circ}$. By virtue of (4.23) and (4.24), the variance $\beta^{-1}(1-\mu^2)^{-1}$ tends to the limiting value $(E^{\circ}-2E_m)/2|D|$. The marginal distributions of $B(x)$ and $V(x)$ converge weakly to the same Gaussian measure, which has mean $\overline{B}(x)$ (the unique solution of the variational problem $E(B) \to min$ subject to $F_i(B)=F_i^{\circ}, i=1,\ldots,n$), and finite variance $(E^{\circ}-2E_m)/2|D|$. However, the limiting macrostate is degenerate since $corr(B_i(x),V_i(x)) \to 1$. In other words, the field and flow are statistically indistinguishable in the limit $H^\circ/E^\circ \to 1$. This remarkable effect is confirmed by numerical simulations in which the magnetic field and velocity field are observed to align when the initial ratio of quadratic cross-helicity to energy is taken above a certain threshold value [31]. Similarly, when $H^\circ/E^\circ \to -1$, it follows that $\mu \to -1$, and we might say that B and V become antialigned.

5. A MATHEMATICAL ANALYSIS OF THE MAXIMUM ENTROPY PRINCIPLE

5.1 Proof of the Existence Theorem

In this section, it is assumed for convenience that $|D|=1$. Then, the macrostates are probability measures on $D\times\mathbb{R}^4$. (We could define, for each macrostate p, the normalized macrostate $\bar{p}=|D|^{-1}p$, and work with the $\bar{p}$ instead.) We begin by showing that the class W of admissible macrostates is closed in the weak topology. Since the family $\{p^t\}$ is contained in W, it is an immediate consequence of this result that, if p is a weak limit of a sequence of elements in $\{p^t\}$, then p satisfies the constraints (3.16)-(3.18), which justifies the form of these constraints on the admissible macrostates.

Lemma 5.1:

The constraint set,

$$W=\{p\in\mathcal{M}: M(p)\le M^\circ, N(p)\le N^\circ, F_i(p)=F_i^\circ, i=1,\dots,n\},$$

is closed in the space of Radon probability measures on $D\times\mathbb{R}^4$ in the weak topology.

Proof:

Let p^k, $k=1,2,\dots$ be elements of W, and assume that p^k converges weakly to p. We must show that $p\in W$. As mentioned above, we know that $p\in\mathcal{M}$. Now, by the Skorokhod Representation Theorem [4, 5, 6], there is a probability space $(\Omega,\mathcal{F},q)$, and random elements Z^k, Z on Ω taking values in $D\times\mathbb{R}^4$, with distributions p^k,p respectively, such that $\lim_{k\to\infty}Z^k(\omega)=Z(\omega)$ for all $\omega\in\Omega$. Let us write

$$Z=(X,\underline{Y}),Z^k=(X^k,\underline{Y}^k),$$

$$\underline{Y}=(\underline{B},\underline{V}),\underline{Y}^k=(\underline{B}^k,\underline{V}^k),$$

with X, $X^k:\Omega\to D$, $\underline{B},\underline{B}^k:\Omega\to\mathbb{R}^2$, and $\underline{V}$, $\underline{V}^k:\Omega\to\mathbb{R}^2$. Then, there holds,

$$\begin{aligned}M(P)&=\frac{1}{2}\int_{D\times\mathbb{R}^4}(b-v)^2dp\\&=\frac{1}{2}\int_{\Omega}(\underline{B}(\omega)-\underline{V}(\omega))^2q(d\omega)\\&\le liminf_k\ \frac{1}{2}\int_{\Omega}(\underline{B}^k(\omega)-\underline{V}^k(\omega))^2q(d\omega)\\&\le M^0\end{aligned}$$

The second line of the above calculation follows by a change of variables. Fatou's lemma gives the third line, and the final inequality follows from the fact that $M(p^k)\le M°$ for all k and a change of variables. We have shown, therefore, that $M(p)\le M°$. An identical argument yields the result $N(p)\le N°$.

We next show that the local mean magnetic fields $\overline{B}^k$ corresponding to the measures p^k converges weakly to the local mean magnetic field $\overline{B}$ corresponding to the macrostate p. Let $g(x)=(g_1(x),g_2(x))$ be an arbitrary function in $C_b(D:\mathbb{R}^2)$, the space of bounded continuous functions on D taking values in $\mathbb{R}^2$. For any k, and any $a>0$, there holds

$$\int_{\{|g(X^k)\cdot\underline{B}^k|\ge a\}}|g(X^k)\cdot\underline{B}^k|q(d\omega)\le\frac{1}{a}\int_{\{|g(X^k)\cdot\underline{B}^k|\ge a\}}|g(X^k)\cdot\underline{B}^k|^2q(d\omega),$$

where X^k, $\underline{B}^k$ and q are defined above. Now,

$$\frac{1}{a}\int_{\{|g(X^k)\cdot\underline{B}^k|\geq a\}}|g(X^k)\cdot\underline{B}^k|^2 q(d\omega) \leq \frac{1}{a}\max_{x\in D}|g(x)|^2\int_\Omega|\underline{B}^k|^2 q(d\omega)$$

$$=\frac{1}{a}\max_{x\in D}|g(x)|^2\int_{D\times\mathbb{R}^4} b^2 dp^k$$

$$\leq\frac{2E^0}{a}\max_{x\in D}|g(x)|^2,$$

Consequently, the random variables $g(X^k)\cdot\underline{B}^k$ are uniformly integrable [5, 6]. Moreover, since $g(X^k)\cdot\underline{B}^k$ converges pointwise to $g(X)\cdot\underline{B}$ as well, there holds

$$\lim_{k\to\infty}\int_\Omega g(X^k)\cdot\underline{B}^k q(d\omega)=\int_\Omega g(X)\cdot\underline{B}q(d\omega).$$

Changing variables once again, we find that

$$\lim_{k\to\infty}\int_{D\times\mathbb{R}^4} g(x)\cdot b\,dp^k=\int_{D\times\mathbb{R}^4} g(x)\cdot b\,dp.$$

and therefore,

$$\lim_{k\to\infty}\int_D g(x)\cdot\overline{B^k}(x)\,dx=\int_D g(x)\cdot\overline{B}(x)\,dx. \tag{5.1}$$

Since (5.1) holds for all g in the dense subset $C_b(D:\mathbb{R}^2)$ of $L^2(D:\mathbb{R}^2)$, and since $\overline{B}^k$ is bounded in $L^2(D:\mathbb{R}^2)$, it follows that $\overline{B}^k$ converges weakly in $L^2(D:\mathbb{R}^2)$ to $\overline{B}$ [54]. (That $\overline{V}^k$ converges weakly to $\overline{V}$ in $L^2(D:\mathbb{R}^2)$ follows by an identical argument, but we will not make use of this fact).

Of course, $P_H\overline{B}^k$ also converges weakly to $P_H\overline{B}$ in $L^2(D:\mathbb{R}^2)$ as p^k converges weakly to p. The compactness of the operator $\mathrm{curl}^{-1}:H\to H_0^1(D)$ together with the Sobolev Embedding Theorem implies that $\overline{\Psi}^k$ converges strongly in $L^r(D)$ to $\overline{\Psi}$ for $1\leq r<\infty$, where $\overline{\Psi}^k=\mathrm{curl}^{-1}P_H\overline{B}^k$, and $\overline{\Psi}=\mathrm{curl}^{-1}P_H\overline{B}$ [1]. Thus, as $k\to\infty$, there holds

$$F_i^\circ = F_i(\overline{\psi^k}) \to F_i(\overline{\psi}), \; i=1,\dots,n.$$

We have shown, therefore, that W is closed.

∎

We now turn to the main result of this section, namely the proof of Theorem 4.1.

<u>Proof of Theorem 4.1</u>.

We remark that for $p \in M$, $K(p:\pi)$ can also be written in the more familiar form [13, 40]

$$K(p:\pi) = \int_{D \times \mathbb{R}^1} \log \frac{dp}{d\pi} dp, \; \textit{if } p \ll \pi, \textit{ and } |\log \frac{dp}{d\pi}| \in L^1(p)$$

$$K(p:\pi) = -\infty \qquad \textit{otherwise}.$$

It is well known [13, 40] that K is nonpositive upper semicontinuous, and has compact level sets in the weak topology, and therefore, K attains its maximum over any nonempty closed set. That W is closed has been established by Lemma 5.1. That it is nonempty follows because $p^t \in W$ for any t. Consequently, K attains its maximum over W.

When conditions (a) and either of the conditions (b), (c) of Theorem 4.1 hold, there are elements of W which have finite entropy. As a result, any solution of (MEP) must have finite entropy. Indeed, the macrostates constructed in Section (4.4) are such elements. (Whether or not there is a unique critical point of $E(B)$ subject to $F_i(B)=F_i^\circ, i=1,\dots n$. let $\overline{B}=B_m$, where B_m satisfies the flux constraints, as well as $E(B_m)=E_m$. Solve for μ by (4.25) and use (4.23) or (4.24) to find β. But note that when this critical point is not unique, the corresponding macrostate may not maximize the entropy).

∎

5.2 Large Deviations and the Concentration Property for Young Measures

While it may be considered conventional wisdom that the maximizers of Kullback-Liebler entropy are, in some sense, most probable, Robert's Concentration Theorem for Young measures [32, 33, 41] allows us to give a more precise mathematical meaning to the statement that solutions of (MEP) are the most probable elements of the set W of admissible macrostates. This theorem is simply a convenient restatement of some results from the theory of large deviations [13, 53]. Indeed, it is a consequence of Baldi's large deviations theorem [2], which in turn is a generalization of a theorem of Ellis [13]. We begin, therefore, with a statement of the large deviation property for a family of probability measures in an appropriately abstract setting for our purposes.

Let X be a locally convex Hausdorff topological vector space, and let μ_h, $h>0$ be a family of probability measures on X. We say that the family μ_h has the large deviation property [13, 53] with constants $a(h)$ and rate function I if and only if:

(a) $a(h)>0$, and $\lim_{h\to\infty} a(h)=\infty$.

(b) $0\leq I(x)\leq\infty$ for all $x\in X$, and $I\not\equiv\infty$.

(c) I is lower semicontinuous on X.

(d) I has compact level sets; that is, for each real number c, the set $\{x:I(x)\leq c\}$ is compact in X.

(e) For each open subset U of X, there holds

$$\liminf_{h\to\infty} a(h)^{-1} \log \mu_h(U) \geq -\inf_{x\in U} I(x).$$

(f) For each closed subset C of X, there holds

$$\limsup_{h\to\infty} a(h)^{-1} \log \mu_h(C) \leq -\inf_{x\in C} I(x).$$

Baldi's Theorem [2, 33] gives general conditions under which the large deviation property holds. The rate function I is often called the information functional, and the functional $K=-I$ is referred to as the entropy functional.

The notion of concentration is defined by means of measurable step functions. We denote by ξ an equipartition of D; that is, $\xi=\{D_1,D_2,\ldots,D_{n(\xi)}\}$, where $D_i\subset D$, $|D_i|=|D|/n(\xi)$, $\bigcup D_i=D$, and $D_i\cap D_j=\emptyset$ if $i\neq j$. Here, $n(\xi)$ is the number of elements in ξ. We also denote by $d(\xi)$ the diameter of ξ:

$$d(\xi) = \sup{}_i diam(D_i),$$

where $diam(D_i)=\sup\{|x-x'|:x,x'\in D_i\}$.

Given $y_1,y_2,\ldots,y_n\in\mathbb{R}^\ell$, where $n\equiv n(\xi)$, consider the step function

$$g_\xi(x)=\sum_{i=1}^{n} y_i I_{D_i}(x),$$

where I_A is the indicator of the set A. Of course, any microstate $Y\in L^2(D:\mathbb{R}^\ell)$ can be approximated in L^2 norm to within any given degree of accuracy by such a step funciton. Let δ_ξ be the Young measure which is a Dirac mass at g_ξ. Thus, δ_ξ is given by

$$\langle\delta_\xi,\varphi\rangle=\sum_{i=1}^{n}\int_{D_i}\varphi(x,y_i)\,dx.$$

Now, we think of the y_i as being chosen independently from $\mathbb{R}^\ell$ according to the density π°. Then, δ_ξ is a random element taking values in M. We denote by μ_ξ the distribution of δ_ξ, and write

$$Prob(\delta_\xi\in A)=\mu_\xi(A),$$

for a Borel subset A of M. We can now state:

Theorem 5.2

When $d(\xi)\to 0$, the family μ_ξ has the large deviation property with constants $n(\xi)$ and rate function
$I(p:\pi)=-K(p:\pi)$.

A general version of this theorem is proved in [33]. As the proof is very technical, we will not include it here. The Concentration Theorem below follows from Theorem 5.2. Let us now define precisely the notion of concentration [33, 40, 41, 42].

Let A, A^* be subsets of M, and let U, U', and U^* denote open neighborhoods of the origin (for the weak topology) in the space of bounded Radon measures on $D\times\mathbb{R}^{\ell}$. Let us ntroduce the notation $A_U=(A+U)\bigcap M$. A_U is an open neighborhood of A in M. We say that δ_ξ concentrates about A^* conditionally to A if:

(*i*) $\quad \liminf_{d(\xi)\to 0} n(\xi)^{-1}\log Prob(\delta_\xi \in A_{U'}) > -\infty,$

for all U',

(ii) Given U*, there exists $\alpha>0$ and U such that for all U',

$$\frac{Prob(\delta_\xi \in A_U \backslash A^*_{U^*})}{Prob(\delta_\xi \in A_{U'})} \le \exp\{-n(\xi)\,\alpha\},$$

for all equipartitions with $d(\xi)$ sufficiently small.

While this definition is quite technical, it has the intuitive interpretation that, if δ_ξ takes values in a neighborhood of A, then, with very high probability, δ_ξ will be in a neighborhood of A^*. Since Prob$(\delta_\xi \epsilon A)$ need not be defined for an arbitrary set A, we widen the sets into open neighborhoods. Even for a Borel set A, Prob$(\delta_\xi \epsilon A)$ could be zero. Condition (i) guarantees that Prob$(\delta_\xi \epsilon A_{U'})$ does not become too small as $d(\xi)\to 0$.

The Concentration Theorem [33, 40, 41, 42] can now be stated:

Theorem 5.3 (Concentration Theorem)

Let A be a nonempty closed subset of M, and let A^* be the (nonempty, closed) subset of A where $K(p:\pi)$ attains its maximum value on A; then δ_ξ concentrates about A^* conditionally to A.

Before proving Theorem 5.3, we state the following corollary, which is of particular interest to us.

Corollary 5.4

Let W^* be the subset of W where $K(p:\pi)$; then δ_ξ concentrates about W^* conditionally to W.

The corollary means, heuristically, that the vast majority of piecewise constant measurable functions that are in a neighborhood of W are actually in a neighborhood of the entropy maximizers(s) in W (in the sense that their corresponding Young measures lie in these neighborhoods). This lends a nice interpretation to the statement that maximizers of the entropy over the constraint set W are the most probable admissible macrostates. The notion of concentration depends upon the choice of the spatially homogeneous measure π. We have made the appealing choice of π to be the most probable spatially homogeneous macrostate consistent with the onservation of energy and quadratic cross-helicity.

Proof of Theorem 5.3

We follow the argument in [33].

Let us simply write $I(p)$ for $I(p:\pi)$. We also introduce the notation

$$I(A)=\inf\{I(p):p\epsilon A\}.$$

Let A and A^* be as in the statement of the theorem. As I is lower semicontinuous and has compact level sets, A^* is nonempty.

Assume, for now, that $I(A)<\infty$. Then, by condition (e) of the large deviation property, there holds

$$\begin{aligned}\liminf n(\xi)^{-1}\log Prob(\delta_\xi \in A_{U'}) &\geq -I(A_{U'})\\ &\geq -I(A)\\ &> -\infty,\end{aligned}$$

so that condition (i) of the concentration property holds.

Next, let U^* be given. Let r be a positive real number such that

$$I(A\backslash A^*_{U^*}) > I(A)+r,$$

(It is clear that such an r always exists. Note that we use the convention $I(\emptyset)=\infty$).

We now make use of the following lemma, the proof of which follows from Lemma 2.3 of [33].

Lemma 5.4

Let C,F be closed subsets of M; then as $U\to 0$, we have $I(C \cap \overline{F}_U)\to I(C \cap F)$.

We apply Lemma 5.4 to conclude that there exists $\alpha>0$ such that, for open neighborhoods U of the origin that are sufficiently small, there holds

$$I(\overline{A}_U\backslash A*_{U*})-I(A_{U'})>\alpha \text{ for all } U'.$$

By condition (f) of the large deviation property, we have

$$\begin{aligned}\limsup n(\xi)^{-1}\log Prob(\delta_\xi\in A_U\backslash A*_{U*}) &\le -I(\overline{A_U\backslash A*_{U*}})\\ &\le -I(\overline{A}_U\backslash A*_{U*})\end{aligned}$$

whereas condition (e) of the large deviation property gives

$$\liminf n(\xi)^{-1}\log Prob(\delta_\xi\in A_{U'})\ge -I(A_{U'}).$$

Now, we have

$$\begin{aligned}-\alpha &> I(A_{U'})-I(\overline{A}_U\backslash A*_{U*})\\ &\ge -\liminf n(\xi)^{-1}\log Prob(\delta_\xi\in A_{U'})\\ &\quad +\limsup n(\xi)^{-1}\log Prob(\delta_\xi\in A_U\backslash A*_{U*})\end{aligned}$$

From which there follows

$$\limsup n(\xi)^{-1}\log\frac{Prob(\delta_\xi\in A_U\setminus A*_{U*})}{Prob(\delta_\xi\in A_{U'})}<-\alpha.$$

This gives condition (ii) of the concentration property.

In the case when $I(A)=\infty$ it must be that $A*=A$. We may say, by convention, that δ_ξ concentrates about $A*$ conditionally to A when such is the case.

■

BIBLIOGRAPHY

[1] Adams, Robert A. *Sobolev Spaces*. New York: Academic Press, 1975.

[2] Baldi, Paolo. Large deviations and stochasitc homogenization. *Annali di Matematica Pura ed Applicata*, 151: 161-177, 1989.

[3] Balian, Roger. *From Microphysics to Macrophysics I*. Berlin: Springer-Verlag, 1991.

[4] Bauer, Heinz. *Probability Theory and Elements of Measure Theory*. London: Academic Press, 1981.

[5] Billingsley, Patrick. *Probability and Measure*. New York: John Wiley & Sons, 1986.

[6] Billingsley, Patrick. *Weak Convergence of Measures*. CBMS-NSF Regional Conference Series in Applied Mathematics, 5. Philadelphia: SIAM, 1971.

[7] Biskamp, D., and Bremer, U. Dynamics and statistics of inverse cascade processes in 2D magnetohydrodynamic turbulence. preprint, 1993.

[8] Biskamp, D., and Welter, H. Coalescence of magnetic islands. *Physical Review Letters*, 44: 1069-1071, 1980.

[9] Biskamp, D., and Welter, H. Dynamics of decaying two-dimensional magnetohydrodyhamic turbulence. *Physics of Fluids B*, 1: 1964-1979, 1989.

[10] Biskamp, D., and Welter, H. Magnetic field amplification and saturation in two-dimensional magnetohydrodynamic turbulence. *Physics of Fluids B*, 2: 1787-1793, 1990.

[11] Biskamp, D., Welter, H., and Walter, M. Statistical properties of two-dimensional magnetohydrodynamic turbulence. *Physics of Fluids B*, 2: 3024-3031, 1990.

[12] Burton, G. R. Rearrangements of functions, maximization of convex functionals, and vortex rings. *Mathematische Annalen*, 276: 225-253, 1987.

[13] Ellis, Richard S. *Entropy, Large Deviations, and Statistical Mechanics*. New York: Springer-Verlag, 1985.

[14] Evans, L. C. *Weak Convergence Methods for Nonlinear Partial Differential Equations*. Regional Conference Series in Mathematics, 74. Providence: AMS, 1990.

[15] Eydeland, Alexander, Spruck, Joel, and Turkington, Bruce. Multi-constrained variational problems of nonlinear eigenvalue type: New formulations and algorithms. *Mathematics of Computation*, 55: 509-535, 1990.

[16] Finn, J. M. and Kaw, P. K. Coalescence instability of magnetic islands. *Physics of Fluids,* 20: 72-78, 1977.

[17] Friedberg, Jeffrey P. *Ideal Magnetohydrodynamics*. New York: Plemium Press, 1987.

[18] Fyfe, David, Joyce, Glenn, and Montgomery, David. Magnetic dynamo action in two-dimensional turbulent magneto-hydrodynamics. *Journal of Plasma Physics*, 17: 317-335, 1977.

[19] Fyfe, David, and Montgomery, David. High-beta turbulence intwo-dimensional magnetohydrodynamics. *Journal of Plasma Physics*, 16: 181-191, 1976.

[20] Gruzinov, A. V. Gaussian free turbulence. *Journal of Experimental and Theoretical Physics*, 76: 241-244, 1993.

[21] Gruzinov, A. V. Gaussian free turbulence: Structures and relaxation in plasma models. *Comments on Plasma Physics and Controlled Fusion*, 15: 227-239, 1993.

[22] Holm, Darryl D., Marsden, Jerrold E., Ratiu, Tudor, and Weinstein, Alan. Nonlinear stability of fluid and plasma equilibria. *Physics Reports*, 123: 1-116, 1985.

[23] Ioffe, A. D., and Tihomirov, V. M. *Theory of Extremal Problems*. Amsterdam: North-Holland, 1979.

[24] Isichenko, M. B., and Gruzinov, A. V. Iso-topological relaxation, coherent structures, and Gaussian turbulence in two-dimensional magnetohydrodynamics. preprint, 1993.

[25] Jackson, J. D. *Classical Electrodynamics*. New York: John Wiley & Sons, 1975.

[26] Jaynes, E. T. Information theory and statistical mechanics. *Physical Review*, 106: 620-630, 1957.

[27] Jordan, Richard K. Ph.D Dissertation, University of Massachusetts, Amherst, 1994.

[28] Kraichnan, Robert H. Statistical dynamics of two-dimensional flow. *Journal of Fluid Mechanics*, 67: 155-175, 1975.

[29] Kraichnan, Robert H., and Montgomery, David. Two-dimensional turbulence. *Reports on Progess in Physics*, 43: 547-619, 1980.

[30] Kullback, S., Keegel, J. C., and Kullback, J. H. *Topics in Statistical Information Theory*. Berlin: Springer-Verlag, 1987.

[31] Matthaeus, W. H., and Montgomery, D. Dynamic alignment and selective decay in MHD. *Statistical Physics and Chaos in Fusion Plasmas*. pp. 285-291. New York: Wiley-Interscience, 1984.

[32] Michel, Julien. *Un Principe de Maximum d'Entropie pour les Mesures de Young, Applications*. These, l'Universite Claude Bernard Lyon I, 1993.

[33] Michel, Julien, and Robert, Raoul. Large deviations for Young measures and statistical mechanics of infintie dimensional dynamical systems with conservation law. *Communications in Mathematical Physics*, 159: 195-215, 1994.

[34] Miller, Jonathan. Statistical mechanics of Euler equations in two-dimensions. *Physical Review Letters*, 65: 2137-2140, 1990.

[35] Miller, Jonathan, Weichman, Peter B., and Cross, M. C. Statistical mechanics, Euler's equations, and Jupiter's Red Spot. *Physical Review A*, 45: 2328-2359, 1992.

[36] Pouquet, A. On two-dimensional magnetohydrodynamic turbulence. *Journal of Fluid Mechanics*, 88: 1-16, 1978.

[37] Priest, E. R. *Solar Magnetohydrodynamics*. Dordrecht: D. Reidel, 1982.

[38] Rao, M. M. *Measure Theory and Integration*. New York: John Wiley & Sons, 1987.

[39] Rao, M. M. *Probability Theory with Applications*. Orlando: Academic Press, 1984.

[40] Robert, Raoul. A maximum entropy principle for two-dimensional perfect fluid dynamics. *Journal of Statistical Physics*, 65: 531-553, 1991.

[41] Robert, Raoul. Concentration et entropie pour les mesures d'Young. *Comptes Rendus de l'Academie des Sciences*, 309, Série I: 757-760, 1989.

[42] Robert, Raoul, and Sommeria, J. Statistical equilibrium states for two-dimensional flows. *Journal of Fluid Mechanics*, 229: 291-310, 1991.

[43] Schumann, U. Numerical simulation of the transition from three-to two-dimensional turbulence under a uniform magnetic field. *Journal of Fluid Mechanics*, 74: 31-58, 1976.

[44] Shebalin, John V. Broken ergodicity and coherent structures in homogeneous turbulence. *Physica D*, 37: 173-191, 1989.

[45] Sommeria, J., and Denoix, M. A., to appear.

[46] Sommeria, J., Staquet, C., and Robert, R. Final equilibrium state of a two-dimensional shear layer. *Journal of Fluid Mechanics*, 233: 661-689, 1991.

[47] Sulem, Catherine. Quelques résultats de régularité pour les équations de la magnétohydrodynamique. *Comptes Rendus de l'Academie des Sciences*, 285: 365-368, 1977.

[48] Sweet, P. A. *Electromagnetic Phenomena in Cosmical Physics*. New York: Cambridge University Press, 1958.

[49] Temam, Roger. *Navier-Stokes Equations: Theory and Numerical Analysis*. Amsterdam: North-Holland, 1984.

[50] Turkington, Bruce, and Jordan, Richard. Turbulent relaxation of a magnetofluid: A statistical equilibrium model. preprint (prepared for *Proceedings of the International Conference on Advances in Geometric Analysis and Continuum Mechanics*), 1994.

[51] Turkington, Bruce, Lifschitz, Alexander, Eydeland, Alexander, and Spruck, Joel. Multiconstrained variational problems in magntohydrodynamics: Equilibrium and slow evolution. *Journal of Computational Physics*, 106: 269-285, 1993.

[52] Turkington, Bruce, and Whitaker, Nathaniel. Statistical equilibrium computations of coherent structures in turbulent shear layers. preprint, 1993.

[53] Varadhan, S. R. S. *Large Deviations and Applications*. CBMS-NSF Regional Conference Series in Applied Mathematics, 46. Philadelphia: SIAM, 1984.

[54] Weidmann, Joachim. *Linear Operators in Hilbert Spaces*. New York, Springer-Verlag, 1980.

[55] Whitaker, Nathaniel, and Turkington, Bruce. Maximum entropy states for rotating vortex patches. preprint, 1993.

[56] Woltjer, L. Hydromagnetic equilibrium, III. *Astrophysical Journal,* 130: 400-405, 1959.

[57] Zeidler, Eberhard. *Nonlinear Functional Analysis and its Applications III*. New York: Springer-Verlag, 1985.