博士論文

Generalized Conservative Dynamics in Topologically Constrained Phase Space: Macro-Hierarchy, Entropy Production, and Self-Organization

（トポロジー束縛を受けた位相空間における拡張保存型力学：
マクロ階層、エントロピー生成と自己組織化）

佐藤　直木

Doctoral Thesis

Generalized Conservative Dynamics in Topologically Constrained Phase Space: Macro-Hierarchy, Entropy Production, and Self-Organization

Naoki Sato

Department of Advanced Energy
Graduate School of Frontier Sciences
The University of Tokyo

July 2017

Contents

Introduction

0.1 Introduction

0.1.1 *Reduction of a microscopic system*

A physical system is characterized by the motion of energy through space. Energy manifests itself in various forms, such as matter, radiation, or a force field. In the presence of energy, dynamical motion may occur. The description of the dynamics requires two fundamental concepts: time, the parameter with respect to which we measure change, and space, the set of possible states that are dynamically accessible. The essential feature of the dynamics of a closed physical system is the conservation of total energy throughout dynamical evolution, while internal order is prohibited from increasing by the second law of thermodynamics.

The dynamics of microscopic physical systems, such as the motion of a point charge in an electrostatic field or that of a point mass in the gravitational field of the Earth, is naturally formulated in the language of canonical Hamiltonian mechanics [1]. The behavior of a canonical Hamiltonian system is determined by the Hamiltonian, the scalar function measuring the energy of the system. The resulting motion, which is prescribed on the level sets of the Hamiltonian, occurs in a 'flat' space, the so called phase space. The phase space does not have preferential directions or inaccessible regions, and is mathematically characterized as a symplectic manifold [2, 3].

A macroscopic system arises when the microscopic description of a certain physical system exhibits negligible degrees of freedom. The removal (or reduction) of such negligible degrees of freedom preserves the essential features of the original system, while providing a simpler description that can be used to understand and predict the relevant global behavior. A tangible example of reduction to a macroscopic system is the formulation of the collective motion of water molecules in the flow of a river as the motion of a fluid: the fluid elements no longer remember the complex interactions occurring among the molecules and their erratic motion caused by all sort of fluctuations, but contain the information necessary to construct a physical theory of fluid dynamics.

The constitutive property of macroscopic systems is that the process of reduction by which they are generated may destroy the canonical Hamiltonian form: the space where the motion of the macroscopic system takes place may not, in general, be qualified as a phase space. Consider, for example, the mechanical description of a rigid body. Rigidity is an ideal concept that acquires consistent physical meaning on appropriate length and time scales where the constituents of the rigid body can be considered as fixed with respect to each other. Such idealization imposes a set of *constraints* on the constituents of the rigid body, which are not allowed to move according to their microscopic dynamics and must occupy a prescribed position in space. Therefore, on the length and time scales where rigidity holds, the phase space of the microscopic constituents is no longer flat: certain regions are now inaccessible and preferential directions arise. The macroscopic model is then obtained by eliminating these inaccessible regions from the mathematical formulation of the dynamics. In the specific case of a rigid body that moves in the absence of external forces in a rotating reference frame determined by the principal axes of inertia, the reduced space is three dimensional with the components of the angular velocity (or angular momentum) as coordinates.

0.1.2 *Conservative dynamics and topological constraints*

If the cut performed on the microscopic phase space is 'sharp enough', the surviving slice of space may resemble a flat phase space. This occurs when the constraints acting on the system delineate a smooth surface, i.e. the macroscopic dynamics is forced on the level set of some integral manifold. Physically, these constraints represent additional constants of motion that do not depend on the energy (Hamiltonian) of the macroscopic system, but on the geometric properties of the reduced space. Whenever the reduced dynamics of a macroscopic system can be locally described as a canonical Hamiltonian system up to a change of coordinates, we speak of a non-canonical Hamiltonian system, which is mathematically characterized in terms of a Poisson algebra [4, 5]. The Poisson algebra is defined by a 'bracket' (a type of inner product whose action is represented by a Poisson 'operator') that satisfies certain mathematical properties, among which the so called Jacobi identity. The Jacobi identity ensures that space has the geometrical attributes of canonical phase space. In this formalism, one obtains the time evolution of a physical observable by taking the bracket between the observable itself and the Hamiltonian of the system. The Euler equations for the motion of a rigid body, as well as the Euler equations for the motion of an ideal fluid are examples of non-canonical Hamiltonian systems [6, 7, 8, 9]. Consider again the rigid body: if the body is made up of N particles, only 3 of the $6N$ phase space degrees of freedom are required to describe the dynamics. If we further note that total angular momentum is preserved, the actual degrees of freedom are just 2. By changing coordinates, the original $6N$ dimensional phase space can be transformed to a 2 dimensional phase space on a $6N-2$ dimensional integral manifold.

However, when the constraints representing the macroscopic properties of a physical system cannot be integrated to form an integral surface, the reduced dynamics is not, in general, Hamiltonian. In other words, there is no coordinate change by which the reduced space qualifies as a canonical phase space. In the Lagrangian formulation of mechanics, non-integrable constraints are usually called nonholonomic. The fundamental result on nonholonomically constrained mechanical systems is that they do not admit an Hamiltonian representation due to the failure of the Jacobi identity that characterizes the Poisson algebra of Hamiltonian mechanics [10, 11, 12, 13, 14]. It is known that the failure of the Jacobi identity has important consequences for the dynamics [15]. In general, a system that fails to be Hamiltonian due to the failure of the Jacobi identity is called almost Hamiltonian, and the associated space is characterized as an almost symplectic manifold [16] endowed with an almost Poisson bracket. We will avoid such terminology, and refer to those systems that possess an energy function but that violate the Jacobi identity simply as *conservative systems* with associated *antisymmetric brackets*. This choice is made to stress the fact that the failure of the Jacobi identity results in a strong departure from Hamiltonian mechanics which does not justify the word 'almost' in the context of the present study. We remark that both non-canonical and canonical Hamiltonian mechanics can be regarded as subclasses of conservative mechanics. A classical example of a nonholonomic mechanical system is the rolling of a disc without slipping on a horizontal surface, and, more generally, the rolling of a rigid body. This type of macroscopic dynamics, which displays peculiar and exotic properties, is yet unexploited and represents a promising source of applications.

Both non-canonical and conservative mechanics originate from certain constrains that are imposed on space. For this reason, they are named *topological constraints*. Object of the present study is the effect of such topological constraints on the statistical behavior of an ensemble of particles that lack (either in the non-canonical or conservative sense) a canonical phase space.

0.1.3 *The problem of statistical mechanics*

At present, there exists a satisfactory mathematical theory of reduction that relates the symmetry properties of the microscopic Hamiltonian function to transformations endowed with a Lie group structure [17], and effective analytical tools have been developed to describe infinite dimensional Hamiltonian systems through the algebraic construction provided by the Poisson bracket. However, a rigorous analysis of dynamical and geometrical properties resulting from inner products that fail to define a canonical phase space, as well as the formulation of the statistical mechanics of non-canonical and conservative systems, are physical and mathematical problems that deserve investigation.

So far, the progress in the field of conservative dynamics pertains to the integration of the equations of motion of particular systems (mainly rigid bodies affected by non-integrable constraints). The explicit integration of such systems, which finds its roots in the seminal work of [18, 19], is performed on a case by case basis by exploiting the Lagrangian representation of the dynamics, and specifically the existence of special classes of symmetries in the Hamiltonian function [20, 21, 22]. In recent years, with the growing interest in strongly reduced, macroscopic, and nonlinear systems, the need to study statistical properties of non-canonical and conservative ensembles has progressively emerged. The purpose here is evidently different from the integration of the dynamics, where the properties of matter (the Hamiltonian) play a crucial role. Statistical considerations rely on the nature of space-time, i.e. the geometrical properties of the inner product that determines the evolution of a physical system. For this reason, in the theory we develop, precise dynamical and statistical predictions are made only on the basis of the character of the inner product, while the Hamiltonian is not specified. Note that, in the standard formulation of statistical mechanics, the inner product is the canonical Poisson bracket of microscopic dynamics [23]. This ansatz, implying a flat phase space where Liouville's theorem holds [24], provides the preserved volume element (invariant measure) that is a prerequisite for the ergodic hypothesis [25, 26, 27, 28, 29] and the standard notion of entropy of a probability distribution. However, the brackets defined by the inner products we are concerned with are not canonical and thereby none of the results above can be used to construct statistical mechanics.

0.1.4 *Self-organization by topological constraints*

The physical motivation behind the present study is the understanding of certain phenomena where a system adjusts autonomously to a macroscopically ordered and organized form in a process called self-organization. These phenomena represent a conundrum when considering the second law of thermodynamics and pose a fundamental theoretical challenge. Aim of the present study is to formulate a rigorous statistical theory of those self-organizing phenomena that owe their internal structure to the existence of topological constraints affecting the phase space of microscopic dynamics. Examples of this kind of self-organization are the vortical shape of a galaxy [30], the accretion of a radiation belt [31, 32, 33], turbulent and vortical structures in fluid and plasma flow [34, 35, 36, 37, 38], the anisotropy in the magnetization of a ferromagnet [39, 40], the bizarre motion of a rattleback [41, 42]. We shall argue that such self-organizing phenomena can be ascribed to the same and basic physical principle: a topological constraint distorts the metric of the phase space and causes the emergence of a macroscopic structure. The emergent structure is, therefore, consistent with the second law of thermodynamics provided that the entropy principle is formulated on an appropriate metric reflecting the topology of the constraint.

The annular radiation belts (clumps of charged particles) that encircle the Earth are one among several heterogeneous structures that can be found in astronomical environments. Analogous formations are observed in other physical systems and across different scales. The common denominator is a non-trivial macroscopic structure that exhibit long term stability. For example, radiation belts (although subject to fluctuations in their shape) probably have the same age as the magnetosphere of the Earth, and the turbulent flows found in gaseous planets such as Saturn or Jupiter have conserved their shape since they were first observed some centuries ago.

The endurance of such self-organized structures seemingly violates the entropy principle dictated by the second law of thermodynamics [43, 44, 45, 46]. Indeed, according to the second law of thermodynamics, the internal disorder of a closed system, which is measured by the scattering of a probability distribution, must increase. The fundamental question we would like to answer is therefore how can those systems that exhibit self-organizing properties be consistent with the postulate of maximum entropy. So far, we have explained that the more macroscopic a system is, the more space-time departs from canonical phase space. Thus, we will show that the standard logic of Boltzmann statistics breaks down, and we will construct a new paradigm that takes into account the geometric properties of space-time. The statistical distribution of a macroscopic system will be determined not only by the Hamiltonian (as in the case of a microcanonical ensemble), but by the intrinsic metric dictated by the underlying inner product. It is therefore through this novel paradigm that entropy maximization leads to the emergence of macroscopic structures.

The type of self-organization discussed in the present study must be distinguished from biological, open, and 'non-physical' self-organization. The physical phenomena we are concerned with can be characterized as closed (in general non-equilibrium) systems affected by topological constraints (although the concept of topological constraint applies to open systems as well). In sharp contrast, biological, chemical, and certain physical systems display complex pattern formation driven by non-equilibrium thermodynamics of open systems [47, 48, 49]. Some biological examples pertain to the emergence of life and associated structures (from the deoxyribonucleic acid (DNA) to the organized behavior of groups of animals), while formation of ice crystals is a prototype of chemical pattern generation. 'Non-physical' self-organization include social and virtual systems, among which stock markets, traffic, internet and virtual networks [50, 51]. Although these 'non-physical' phenomena may be treated with the formalism of modern mathematical physics, they are substantially different in that they are not subjected to a conservation law such as conservation of energy in a closed physical system.

0.2 Objective and Outline of the Thesis

Aim of the present study is to formulate the statistical theory of those physical systems that lack the phase space of canonical Hamiltonian mechanics in virtue of topological constraints, and to construct the mathematical tools that are necessary to achieve this objective.

The present thesis is organized as follows. The first three chapters, constituting the first part (the mathematical background) of the thesis, are dedicated at reviewing the theoretical and mathematical tools that will be exploited in our investigation. In chapters 4 and 5, the second part of the thesis, we develop a geometrical characterization and categorization of antisymmetric operators by introducing novel classes of operators (measure preserving and Beltrami) and by defining the notion of current of an operator (the word current is used in analogy with electromagnetism). The concept of topological constraint in the context of conservative dynamics is also discussed. The statistical theory dictated by topological constraints is developed in chapter 6, where we explicitly calculate the equilibrium probability distribution for different classes of antisymmetric operators. The theory formulated in chapters $4-6$ is applied to the modeling of the diffusion of charged particles in magnetospheric plasmas in chapter 7, to the study of conformal mechanical systems in chapter 8, and extensively tested by detailed numerical simulations in chapter 9. Finally, chapter 10 is centered around the normal Laplacian, a novel (non-elliptic) differential operator whose properties determine the equilibrium probability distribution of certain conservative systems.

The order of the chapters follows a logical criterion that does not reflect the chronological development of the theory. In the following, we outline the main contents of the thesis by starting from the motivating physical problem, the self-organization of a radiation belt in a magnetospheric plasma.

0.2.1 *Inward diffusion and entropy production in magnetospheric plasmas*

Our investigation begins with a very practical problem: how to confine a hot plasma (simplistically an ionized gas, see [52, 53]) in a suitably designed 'magnetic bottle' in order to obtain thermonuclear fusion reactions that can be exploited as an energy source. In the core of a star, the temperature, density, and pressure required to sustain fusion reactions are achieved through the gravitational force. However, replicating the same thermodynamic conditions in the laboratory requires a different kind of 'trap', since using gravity would need an unfeasible mass distribution. The alternative trapping force is the magnetic field. Several magnetic confinement devices have been designed and constructed based on different magnetic configurations, such as linear traps [54, 55, 56], donut-shaped tokamaks [57], and stellarators, which use coils with more complicated (such as helical) geometries [58]. Although important progress has been made toward the realization of a thermonuclear fusion power plant, significant hurdles still remain. Namely, plasma instabilities [59, 60, 61, 62, 63], turbulence [64], material and divertor technology, cost-effective superconducting coil manufacturing, radiation (mainly neutrons) shielding, and so on.

An alternative configuration is the so called dipole confinement [65, 66]. In this magnetic configuration, a dipole magnetic field that mimics the magnetosphere of the Earth is generated by a levitated superconducting coil [67, 68, 69]. Such magnetospheric plasma, which is trapped in the same way charged particles accumulate in a radiation belt, shows superior stability properties and is expected to achieve sensibly higher temperatures than a conventional tokamak. These temperatures would allow the deuterium - helium-3 reaction, which does not produce neutron radiation, and therefore is preferable to the deuterium - tritium based fusion that will be exploited in the first generation of thermonuclear fusion power plants. In addition, the dipole magnetic field is the favorite candidate for optimal antimatter, single species and pair plasma confinement [70, 71] as it would not require external electric fields [72, 73].

The efficient trapping mechanism provided by the dipole magnetic field, whose complex dynamics is only partially understood, will be discussed in detail in chapter 7 and represents an important test for the theory we develop. Let us briefly summarize the content of this part.

In a dipole magnetic field, a charged particle performs three distinct periodic motions: the cyclotron gyration around the magnetic field, the bounce motion along the magnetic field, and the toroidal drift around the symmetry axis [74, 75]. Each periodic motion is characterized by an adiabatic invariant that is preserved on a time scale reflecting the strength of the constraint. It is the interplay between such hierarchy of constraints and self-induced electromagnetic fluctuations that generates a peculiar random walk, the so called inward diffusion of charged particles [76, 77, 67, 69, 78]. This process drives the system to the self-organized statistical equilibrium, whose heterogeneity seemingly violates the entropy principle. The naturally achieved equilibrium is characterized by strongly inhomogeneous particle density and temperature, a self-induced electric field, and a rigid rotation around the symmetry axis.

The creation of the density clump can be properly understood if we interpret the adiabatic invariants as topological constraints affecting the phase space of a noncanonical Hamiltonian system [79, 80, 81]: the electromagnetic fluctuations destroy the weaker invariants (magnetic flux and bounce action) but the charged particles are forced to move on the symplectic submanifold where the first adiabatic invariant (the magnetic moment) is constant. Due to the distorted metric on the symplectic submanifold, the concomitant inward diffusion pushes particles toward domains with higher magnetic field strength and entropy is maximized on the metric induced by the magnetic coordinates [82, 83, 84, 85, 86]. The plasma toroidal rotation is further explained as the necessary outcome of the relaxation process: the self-induced electric field is cancelled by the Lorentz-transformed electric field in a rigidly rotating coordinate system [83]. These results, which we generalize to arbitrary noncanonical Hamiltonian systems, represent the main achievement of this part of the study and are summarized in [82, 83, 84, 85].

The diffusion operator obtained by such rigorous geometrical construction exhibits an intrinsic complexity that reflects the non-trivial topology of the underlying magnetic field and clearly shows the inaccuracy of phenomenological diffusion equations that are not derived on the basis of solid statistical mechanical arguments.

0.2.2 *The geometrical hierarchy of conservative dynamics*

From a mathematical standpoint, the mechanical behavior of a charged particle in a dipole magnetic field is a consequence of the degeneracy of the Poisson operator (certain regions of the phase space are not accessible), and its non-zero current (the phase space metric is distorted by the constraints). Therefore, we can extend and apply the present theory to general systems (not necessarily Hamiltonian) affected by topological constraints that will not, in general, be integrable. A prototypical example of this scenario is the $\boldsymbol{E} \times \boldsymbol{B}$ drift motion of a charged particle in a non-integrable magnetic field, such as a Beltrami field (a vector field completely aligned with its own curl). $\boldsymbol{E} \times \boldsymbol{B}$ drift dynamics motivates the next part of the present work. Here, we investigate the nature of statistical processes in the presence of non-integrable topological constraints. Several complex mechanical systems exhibit analogous non-Hamiltonian behavior. In addition to nonholonomic dynamics of rigid bodies and dynamical equations in the context of molecular dynamics [87, 88, 89, 90] we will show that other systems of physical interest, such as the Landau-Lifshitz equation [39, 40] for the magnetization of ferromagnetic materials, fall in the category of conservative dynamics and owe their special properties to the failure of the Jacobi identity of Poisson algebras.

The main results of this second part can be summarized as follows. In our study of conservative dynamics, we categorize antisymmetric operators and the associated brackets: we find a hierarchical structure on top of which lies the symplectic matrix of canonical phase space. Immediately around canonical phase space, we encounter Poisson brackets of non-canonical Hamiltonian mechanics. Beyond Poisson brackets, the Hamiltonian nature of dynamics is lost, and we find the closest type of motion, governed by the conformal operator. Conformal dynamical systems, which are known in the literature [18, 20], are characterized by the fact that they can be transformed to a Hamiltonian form by a proper time reparametrization dictated by the failure of the Jacobi identity. Accordingly, we discuss the notion of conformal operator, and determine specific classes of systems that fall in this category. The tipping point of conservative dynamics is represented by the subsequent measure preserving operator: such novel operator guarantees the existence of an invariant measure for any choice of the Hamiltonian function. Using this property, we are able to prove a theorem on the form of the equilibrium distribution function for systems endowed with the measure preserving operator. We remark that this theorem, which does not require the existence of canonical phase space, applies to Poisson operators and reproduces the standard Maxwell-Boltzmann distribution in the special case of a canonical Hamiltonian system. In the specific case of a non-canonical Hamiltonian system we find that, on the invariant measure, the form of the distribution function at equilibrium is determined by the Hamiltonian function, as in the canonical case, and by the integrable kernel of the Poisson operator, which is the seed of what we call *type I self-organization.*

In this type of self-organization, the equilibrium probability distribution is, in general, a function of the Casimir invariants [5], the distinguished functions (first identified by Sophus Lie) whose gradients span the kernel of the Poisson operator of non-canonical Hamiltonian systems. Such functions are constants of motion that do not depend on the specific form of the Hamiltonian, but on the geometry of space dictated by the Poisson operator. The integrability (in the sense of Frobenius [91, 92, 93]) of the kernel of a Poisson operator, i.e. the existence of Casimir invariants, is guaranteed by the Darboux theorem of differential geometry [94, 95, 96]. Since the dynamics is constrained on the level sets of the Casimir invariants (the so called Casimir leaves), type I self-organization is characterized by the existence of inaccessible regions in the phase space.

Regarding measure preserving operators, we prove a second result: any antisymmetric operator can be extended to a measure preserving operator by adding a new degree of freedom to the dynamics. Thus, the probability distribution in the extended space can be obtained by application of the previous result on the equilibrium of measure preserving operators.

A drastic change in the structure of space (and therefore in the nature of the dynamics) occurs when we leave the realm of measure preserving operators. It is an essential achievement of the present study the proof that a measure preserving operator is represented by a closed differential form so that the operator does not generate current (which is given by the exterior derivative of the differential form associated to the operator) in the coordinates spanning the invariant measure. Hence, violation of measure preservation emerges when an operator causes a non-vanishing current in any coordinate system. Since operators determine the properties of space, operators with non-vanishing current impart flows and vortices to the metric of space. These 'metric' currents are the origin of what we call *type II self-organization.* In this case, phase space does not have inaccessible regions but preferential directions: a particle will tend to follow the current and fall toward the center of the eddies. See figure 0.1.

Immediately after measure preserving operators, we identify the new class of Beltrami operators. These operators, characterized by being completely aligned with their own current in an appropriate coordinate system, display peculiar properties from the standpoint of statistical mechanics: even if they do not define an invariant measure, a diffusion process driven by a Beltrami operator results in the complete flattening of the probability distribution. This fact can be rephrased by stating that they represent the largest class of operators whose diffusion admits, in a precise sense

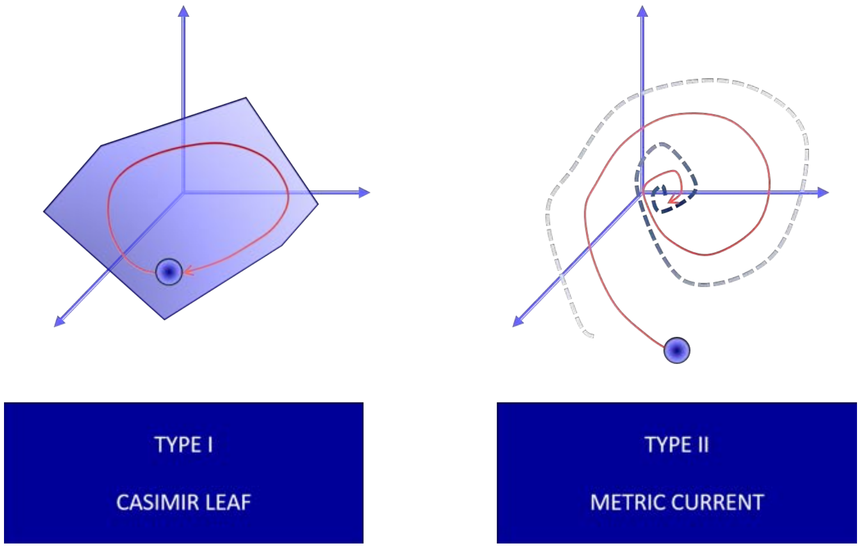

Figure 0.1: Types of self-organization. Left: type I self-organization driven by Casimir invariants. Right: type II self-organization driven by metric current.

we will specify later, the standard definition of differential entropy. Furthermore, Beltrami operators generalize the notion of Beltrami field, which plays a critical role in plasma and fluid relaxation theories [97, 98, 99, 100, 101, 102], from three to arbitrary dimensions.

The last class we encounter is that of antisymmetric operators that do not fall in any of the previous categories. See figure 0.2. To this last class belong $\boldsymbol{E} \times \boldsymbol{B}$ dynamics, the Landau-Lifshitz equation, and, in general, rigid body dynamics with nonholonomic constraints.

It is worth to mention that beyond antisymmetric operators, which identify conservative mechanics, we find dissipative systems. This transition is mathematically represented by the violation of operators antisymmetry. Dissipative systems will not be object of the present study.

0.2.3 *On the second law of thermodynamics in a topologically constrained phase space*

In our construction of statistical mechanics, we constantly deal with the implications of canonical phase space deficiency with respect to the second law of thermodynamics. Without the invariant measure of canonical phase space, the conventional entropy measure for continuous probability distributions breaks down [104, 105, 106, 107], and Boltzmann's H-theorem ceases to hold. Indeed, the conventional entropy measure of continuous probability distributions, which is derived

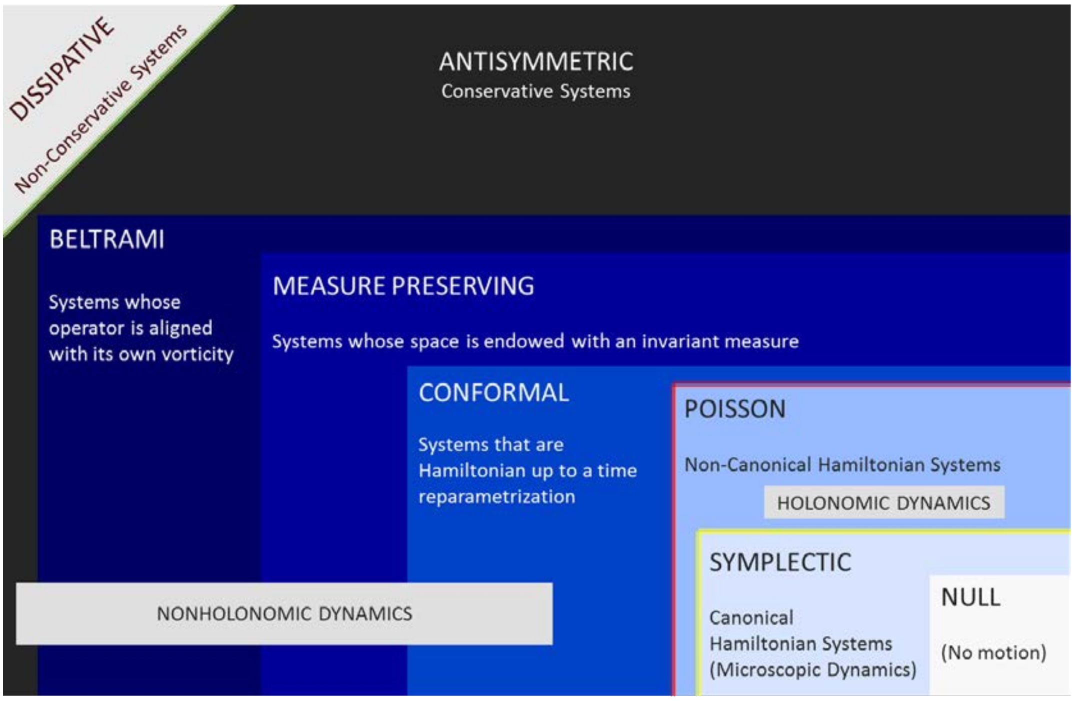

Figure 0.2: The hierarchical structure of antisymmetric operators. Each box is named by the corresponding operator. The yellow line indicates transition from microscopic to macroscopic dynamics. The red line indicates transition from Hamiltonian to conservative dynamics, with corresponding loss of phase space. Systems affected by integrable (holonomic) constraints possess a Poisson operator and lie within the red line. Systems affected by non-integrable (nonholonomic) constraints violate the Jacobi identity and fall outside the red line. The green line indicates transition from conservative to dissipative dynamics. The latter is not object of the present study.

from Shannon's entropy measure for discrete probability distributions [108], is not covariant and relies on the tacit assumption that space is canonical.

Exploiting the mathematical construction developed to characterize the geometric properties of antisymmetric operators, we construct a proper entropy measure that is not anymore determined by the probability distribution alone, but that explicitly depends on the degeneracy and the current of the operator. The new entropy measure, which displays a non-negative entropy production, implicitly defines a time-independent coordinate change dictated by the aforementioned geometric properties of space. Here, the laws of thermodynamics are fully respected.

The sharp departure from the conventional entropy measure is a direct consequence of the deviation from the flat metric of phase space. Thus, this study clearly shows that, in the general context of the statistical behavior of a macroscopic ensemble, the structure of space-time cannot be neglected in the determination of the thermodynamic behavior of the system. Moreover, the developed theory, which can account for the physics of complex phenomena such as the accretion of a radiation belt or the magnetization of a ferromagnet, is notably simple in that it derives its physical predictions only from two ingredients: the antisymmetric matrix representing the operator of the system, and its Hamiltonian function.

At this point, it is worth to add some general considerations on the perspective regarding the second law of thermodynamics developed here. First, we note that the essence of the question considered in this work lies in the non-trivial topology of macroscopic systems. This should not be confused with the infringement of the additive property in some entropy measures, such as the Tsallis entropy [109]. Indeed, the Tsallis entropy suffers the same problem of Shannon's discrete entropy when non-flat metrics are taken into account. In addition, we stress that the present theory gives an explicit expression of entropy production. Thereby, the problem addressed here is conceptually different from extremum principles for entropy production encountered in non-equilibrium thermodynamics of open systems [110, 111, 112, 113, 114].

From a mathematical standpoint, the principal hurdle affecting conservative dynamics is represented by the non-elliptic nature of the general stationary form of the Fokker-Planck equation describing the evolution of the probability distribution. At present, there is no satisfactory mathematical theory of non-elliptic partial differential equations. In our study we use elementary tools of functional analysis [115, 116, 117, 118] to examine existence and uniqueness of solutions to a special non-elliptic equation arising in the case of 3 dimensional antisymmetric operators. Exploiting the non-integrability property of the topological constraint affecting the conservative formulation of the dynamics, a novel type of inner product is defined. The associated Hilbert space provides the natural setting to determine a weak solution to the normal Laplace equation, the second order non-elliptic partial differential equation examined in the last chapter of this work. This result shows a clear relationship between the geometrical concept of integrability, and the mathematical theory (functional analysis) of non-elliptic partial differential equations.

Finally, we refer the reader to the following references concerning physical and mathematical tools that will be exploited in the present work: Brownian motion [119, 120], Langevin equation [121], probability theory [122], stochastic analysis [123, 124, 125, 126, 127, 128, 129, 130], Fokker-Planck equation [131], nonlinear and chaotic dynamics [132].

Part I

Mathematical Background

Chapter 1

Conservative Dynamics

Let $\mathcal{M}$ be a smooth manifold of dimension n.

***Def* 1.1.** *An antisymmetric operator is a bivector field $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$.*

Let $\left(x^1, ..., x^n\right)$ be a coordinate system on $\mathcal{M}$. Consider the tangent basis $(\partial_1, ..., \partial_n)$. We have:

$$\mathcal{J} = \sum_{i<j} \mathcal{J}^{ij} \partial_i \wedge \partial_j = \frac{1}{2} \mathcal{J}^{ij} \partial_i \wedge \partial_j, \qquad \mathcal{J}^{ij} = -\mathcal{J}^{ji}. \tag{1.1}$$

Here and throughout this study the summation convention on repeated indices is used. Note that the matrix $\mathcal{J}^{ij}$ is antisymmetric. Here and throughout this study we shall assume $\mathcal{J}^{ij} \in C^{\infty}(\mathcal{M})$.

***Def* 1.2.** *The pair $(\mathcal{M}, \mathcal{J})$ is called an antisymmetric manifold.*

We now introduce a notion of inner product on $\mathcal{M}$.

***Def* 1.3.** *An antisymmetric bracket on $\mathcal{M}$ is a binary operation $\{,\} : C^{\infty}(\mathcal{M}) \times C^{\infty}(\mathcal{M}) \to C^{\infty}(\mathcal{M})$ satisfying antisymmetry, bilinearity, and the Leibniz rule:*

$$\{f, g\} = -\{g, f\}, \tag{1.2}$$

$$\{\alpha f + g, \beta h + i\} = \alpha\beta\{f, h\} + \alpha\{f, i\} + \beta\{g, h\} + \{g, i\}, \tag{1.3}$$

$$\{fg, h\} = f\{g, h\} + \{f, h\} g, \tag{1.4}$$

for every $f, g, h, i \in C^{\infty}(\mathcal{M})$ and $\alpha, \beta \in \mathbb{R}$.

***Proposition* 1.1.** *Let $\mathcal{J}$ be an antisymmetric operator on $\mathcal{M}$. The inner product on pairs of functions $f, g \in C^{\infty}(\mathcal{M})$:*

$$\{f, g\} = \mathcal{J}(df, dg) = \mathcal{J}^{ij} f_i g_j, \tag{1.5}$$

is an antisymmetric bracket.

In this notation, lower indices applied to a function indicate derivation, i.e. $f_i = \frac{\partial f}{\partial x^i}$. Verification of proposition 1.1 is immediate. To clarify the mathematical formalism, let us review the passages leading to equation (1.5):

$$\begin{aligned} \mathcal{J}(df, dg) &= \frac{1}{2}\mathcal{J}^{ij}\partial_i \wedge \partial_j \left(f_k dx^k, g_l dx^l\right) \\ &= \frac{1}{2}\mathcal{J}^{ij} f_k g_l \left(\partial_i \left(dx^k\right) \otimes \partial_j \left(dx^l\right) - \partial_j \left(dx^k\right) \otimes \partial_i \left(dx^l\right)\right) \\ &= \frac{1}{2}\mathcal{J}^{ij}\left(f_i g_j - f_j g_i\right) = \mathcal{J}^{ij} f_i g_j. \end{aligned} \tag{1.6}$$

In the last passage, we used antisymmetry of $\mathcal{J}^{ij}$.

The definition of conservative vector field is the following:

***Def* 1.4.** *An antisymmetric operator $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ and an Hamiltonian function $H \in C^{\infty}(\mathcal{M})$ define a conservative vector field $X \in T\mathcal{M}$ as:*

$$X = \mathcal{J}(dH). \tag{1.7}$$

***Remark* 1.1.** *In this notation, the action of $\mathcal{J}$ on the 1-form dH is calculated as follows:*

$$X = \frac{1}{2}\mathcal{J}^{ij}\left(\partial_i \otimes \partial_j \left(H_k dx^k\right) - \partial_j \otimes \partial_i \left(H_k dx^k\right)\right) = \mathcal{J}^{ij} H_j \partial_i. \tag{1.8}$$

Thanks to antisymmetry, the result below holds:

***Proposition* 1.2.** *A conservative vector field X preserves the Hamiltonian H along the flow:*

$$\mathfrak{L}_X H = i_X dH = H_i dx^i \left(\mathcal{J}^{jk} H_k \partial_j\right) = \mathcal{J}^{ik} H_i H_k = 0. \tag{1.9}$$

In many practical situations a dual representation of (1.7) is useful. To introduce this alternative notation, first consider the following definition:

***Def* 1.5.** *A vorticity 2-form on $\mathcal{M}$ is a 2-form $\omega \in \bigwedge^2 T^*\mathcal{M}$. The pair $(\mathcal{M}, \omega)$ is called a vorticity manifold.*

Here and throughout this study we shall assume $\omega^{ij} \in C^{\infty}(\mathcal{M})$. Then, we have a second definition of conservative vector field:

***Def* 1.6.** *A vector field $X \in T\mathcal{M}$ is called a conservative vector field with Hamiltonian H and vorticity 2-form ω if:*

$$i_X \omega = -dH. \tag{1.10}$$

***Remark* 1.2.** *One can move from the representation* (1.7) *to* (1.10) *and vice versa when the matrix $\mathcal{J}^{ij}$ (or ω_{ij}) is invertible. In this case, we have $\mathcal{J}^{ij}\omega_{jk} = \delta^i_k$.*

***Proposition* 1.3.** *The conservative vector field defined by* (1.10) *preserves the Hamiltonian H along the flow:*

$$\mathfrak{L}_X H = i_X dH = -i_X^2 \omega = -\frac{1}{2}\omega_{ij} i_X^2 dx^i \wedge dx^j = \omega_{ij} X^i X^j = 0. \tag{1.11}$$

***Remark* 1.3.** *Note that in definitions 1.1 and 5.5 both $\mathcal{J}$ and ω do not have, in general, a constant rank. This implies that the matrices $\mathcal{J}^{ij}$ and ω_{ij} may have a non-trivial kernel whose dimensionality changes depending on the position $\boldsymbol{x}$ on $\mathcal{M}$.*

So far we have introduced the concept of conservative vector field. From a physical standpoint, this mathematical object characterizes the dynamics of conservative systems. The conserved quantity is the Hamiltonian H, which does not need to be the energy, but can represent any other constant of motion. Conservation is guaranteed by the antisymmetry of the matrices $\mathcal{J}^{ij}$ and ω_{ij}. In the next step, we add a special structure to the space where motion occurs, the Poisson bracket.

***Def* 1.7.** *Let $\mathcal{J}$ be an antisymmetric operator on $\mathcal{M}$. The antisymmetric bracket defined by $\mathcal{J}$ is called a Poisson bracket if it satisfies the Jacobi identity:*

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0, \tag{1.12}$$

for every $f, g, h \in C^\infty(\mathcal{M})$. In this case, $\mathcal{J}$ is called a Poisson operator and the associated vector field X a non-canonical Hamiltonian vector field.

***Remark* 1.4.** *In terms of the Poisson operator $\mathcal{J}$, the Jacobi identity reads:*

$$\mathcal{J}^{im}\frac{\partial \mathcal{J}^{jk}}{\partial x^m} + \mathcal{J}^{jm}\frac{\partial \mathcal{J}^{ki}}{\partial x^m} + \mathcal{J}^{km}\frac{\partial \mathcal{J}^{ij}}{\partial x^m} = 0 \quad \forall i, j, k = 1, ..., n. \tag{1.13}$$

If $\mathcal{J}$ is invertible with inverse ω, this condition is equivalent to requiring that $d\omega = 0$. This can be seen in the following manner:

$$d\omega = \sum_{i<j<k} \left(\frac{\partial \omega_{ij}}{\partial x^k} + \frac{\partial \omega_{jk}}{\partial x^i} + \frac{\partial \omega_{ki}}{\partial x^j}\right) dx^i \wedge dx^j \wedge dx^k. \tag{1.14}$$

Multiplying each component of this 3*-form by* $\mathcal{J}^{li}\mathcal{J}^{mj}\mathcal{J}^{nk}$ *and summing over* i, j, k*, we obtain:*

$$
\begin{aligned}
\mathcal{J}^{li}\mathcal{J}^{mj}\mathcal{J}^{nk}&\left(\frac{\partial\omega_{ij}}{\partial x^k}+\frac{\partial\omega_{jk}}{\partial x^i}+\frac{\partial\omega_{ki}}{\partial x^j}\right)\\
&=-\mathcal{J}^{mj}\mathcal{J}^{nk}\omega_{ij}\frac{\partial\mathcal{J}^{li}}{\partial x^k}-\mathcal{J}^{li}\mathcal{J}^{nk}\omega_{jk}\frac{\partial\mathcal{J}^{mj}}{\partial x^i}-\mathcal{J}^{li}\mathcal{J}^{mj}\omega_{ki}\frac{\partial\mathcal{J}^{nk}}{\partial x^j}\\
&=\delta^i_m\mathcal{J}^{nk}\frac{\partial\mathcal{J}^{li}}{\partial x^k}+\delta^n_j\mathcal{J}^{li}\frac{\partial\mathcal{J}^{mj}}{\partial x^i}+\delta^l_k\mathcal{J}^{mj}\frac{\partial\mathcal{J}^{nk}}{\partial x^j}\\
&=\mathcal{J}^{nk}\frac{\partial\mathcal{J}^{lm}}{\partial x^k}+\mathcal{J}^{li}\frac{\partial\mathcal{J}^{mn}}{\partial x^i}+\mathcal{J}^{mj}\frac{\partial\mathcal{J}^{nl}}{\partial x^j}.
\end{aligned}
\tag{1.15}
$$

Thus, if $d\omega = 0$*, the Jacobi identity is satisfied. With a similar procedure, one can show that the converse is also true.*

Let us now review the definition of non-canonical Hamiltonian vector field in terms of ω:

***Def* 1.8.** *Let* $(\mathcal{M}, \omega)$ *be a vorticity manifold and* H *the Hamiltonian function. If* $d\omega = 0$*,* ω *is called a symplectic* 2*-form. Furthermore, a vector field* X *such that* $i_X\omega = -dH$ *is a non-canonical Hamiltonian vector field.*

***Remark* 1.5.** *In the literature, both vorticity and symplectic* 2*-forms are defined with the additional requirement that* ω_{ij} *is an invertible matrix, i.e. it has constant and maximum rank* $2n$ *for some* $n \in \mathbb{N}$*. We remark that this assumption is not made here, since non-vanishing kernels impart non-trivial topologies that have important consequences for the construction of statistical ensembles.*

***Proposition* 1.4.** *A non-canonical Hamiltonian vector field* X *preserves the symplectic* 2*-form* ω*:*

$$\mathfrak{L}_X\omega = di_X\omega = -ddH = 0. \tag{1.16}$$

The physical properties of the Poisson operator can be understood through the Darboux theorem of differential geometry. Since this is a central result for the purpose of the present study, we shall give a complete proof and refer the reader to the literature [94, 133, 134, 135, 136] for additional details.

***Theorem* 1.1.** *(Darboux)*
Let $(\mathcal{M}, \omega)$ *be a symplectic manifold of dimension* $2n + r$*. Suppose that* ω *has rank* $2n$*. Then,* $\forall \boldsymbol{x} \in \mathcal{M}$ *one can find a coordinate neighborhood* $U \in \mathcal{M}$ *of* $\boldsymbol{x}$ *with coordinates* $\left(x^1, ..., x^{2n}, C^1, ..., C^r\right)$ *such that:*

$$\omega = \sum_{i=1}^{n} dx^i \wedge dx^{n+i} \quad on \quad U. \tag{1.17}$$

Proof. (Moser-Weinstein) The first step of the proof consists in showing that one can find a coordinate neighborhood $U \in \mathcal{M}$ with coordinates $\left(y^1, ..., y^{2n}, C^1, ..., C^r\right)$ such that:

$$\omega = \sum_{i<j}^{2n} \alpha_{ij} dy^i \wedge dy^j, \tag{1.18}$$

for some appropriate coefficients $\alpha_{ij} = \alpha_{ij}\left(y^1, ..., y^{2n}, C^1, ..., C^r\right)$. We will give an original proof of this statement in the next part of the present study (see proposition 5.1). Using this result, we can restrict our attention to the submanifold C^1 =constant, ..., C^r =constant and forget about the C^is.

Let ω_0 be the constant 2-form on $\mathbb{R}^{2n}$ defined by:

$$\omega_0\left(\boldsymbol{x}\right) = \sum_{i=1}^{n} dx^i \wedge dx^{n+i}, \qquad \boldsymbol{x} \in \mathbb{R}^{2n}. \tag{1.19}$$

Consider the following family of 2-forms:

$$\omega_t = \omega_0 + t\left(\omega - \omega_0\right), \qquad t \in [0,1]. \tag{1.20}$$

Note that at any point $p \in \mathcal{M}$ we can set $\omega_t(p) = \omega_0(p) = \omega(p)$. Hence, we can find a small neighborhood U of p where ω_t is non-degenerate for all $t \in [0,1]$ since the general linear group $Gl\left(2n, \mathbb{R}\right)$ is an open set of the set of $2n \times 2n$ matrices on $\mathbb{R}$, $gl\left(2n, \mathbb{R}\right)$.

Now consider the family of vector fields:

$$\frac{d}{dt}\phi_t(p) = X_t\left(\phi_t\right), \qquad \phi_0(p) = p. \tag{1.21}$$

Here, ϕ_t is a one-parameter group of diffeomorphisms. We want to determine ϕ_t such that the pull back of ω_t with respect to this transformation satisfies:

$$\phi_t^* \omega_t = \omega_0. \tag{1.22}$$

Using the Poincaré lemma, $\omega - \omega_0 = d\alpha$. Then, from (1.22):

$$\begin{aligned} \frac{d}{dt}\phi_t^*\omega_t &= \phi_t^*\left(\mathfrak{L}_{X_t}\omega_t + \frac{d}{dt}\omega_t\right) \\ &= \phi_t^*\left(di_{X_t}\omega_t + \omega - \omega_0\right) \\ &= \phi_t^*\left(di_{X_t}\omega_t + d\alpha\right) \\ &= 0. \end{aligned} \tag{1.23}$$

Setting $X_t = X_t^i \partial_i$, $\omega_t = \frac{1}{2}\omega_{t,ij} dy^i \wedge dy^j$, and $\alpha = \alpha_i dy^i$, we arrive at:

$$\omega_{t,ij} X_t^i + \alpha_j = 0. \tag{1.24}$$

Since ω_t is non-degenerate, this equation has always solution and we can determine in a unique manner X_t. Thus, we can find the desired transformation ϕ_t mapping ω to ω_0. □

***Def* 1.9.** *The symplectic* 2*-form:*

$$\omega_c = \sum_{i=1}^{n} dx^i \wedge dx^{n+i}, \tag{1.25}$$

is called canonical symplectic 2*-form of* $\mathbb{R}^{2n}$*. The inverse matrix:*

$$\mathcal{J}_c = \sum_{i=1}^{n} \partial_{n+i} \wedge \partial_i, \tag{1.26}$$

is called symplectic operator (or simplectic matrix). Let H be the Hamiltonian function. The vector field:

$$X = \mathcal{J}_c\left(dH\right), \tag{1.27}$$

or, equivalently,

$$i_X\omega_c = -dH, \tag{1.28}$$

is called a canonical Hamiltonian vector field.

In light of Darboux's theorem 1.1, we see that any symplectic 2-form ω of rank $2n$ on a $2n$-dimensional manifold $\mathcal{M}$ is locally diffeomorphic to the canonical symplectic 2-form ω_c on $\mathbb{R}^{2n}$. In physical terms, this means that, locally, one can always find an appropriate coordinate change such that $\omega = \omega_c$. If we further rename each pair $\left(x^i, x^{i+n}\right)$ as the canonical pair $\left(p^i, q^i\right)$, from (1.27) we find Hamilton's canonical equations of motion:

$$X = \dot{p}^i\partial_{p^i} + \dot{q}^i\partial_{q^i} = -\frac{\partial H}{\partial q^i}\partial_{p^i} + \frac{\partial H}{\partial p^i}\partial_{q^i}. \tag{1.29}$$

When ω is degenerate, i.e. its rank is $2n$ in a $2n+r$-dimensional manifold, on every integral surface C^1 =constant, ..., C^r =constant one can still find a local coordinate change sending ω to ω_c. It is also clear that, on such surface, Hamilton's canonical equations (1.29) preserve the value of the C^is.

As shown in remark 1.4, to every symplectic 2-form ω that is invertible, we can associate a Poisson operator $\mathcal{J}$, whose components are exactly the components of the inverse matrix $\left(\omega^{-1}\right)^{ij} = \mathcal{J}^{ij}$. Therefore, thanks to Darboux's theorem, an invertible Poisson operator admits a local coordinate change such that $\mathcal{J}$ is the symplectic operator (1.26). If $\mathcal{J}$ is a degenerate Poisson operator of rank $2n$ in a $2n+r$-dimensional manifold, it can be shown that $ker\left(\mathcal{J}\right)$ is spanned by the gradients of r functions $\left(C^1,...,C^r\right)$:

$$ker\left(\mathcal{J}\right) = span\left\{dC^1,...,dC^r\right\}. \tag{1.30}$$

A proof of this statement completely in terms of $\mathcal{J}$ will be given in the next part of the present study (see proposition 5.2). Therefore, as in Darboux's theorem 1.1, one can find a set of local coordinates $\left(x^1,...,x^{2n},C^1,...,C^r\right)$ such that $\mathcal{J}$ is the symplectic operator:

$$\mathcal{J} = \sum_{i=1}^{n} \partial_{n+i} \wedge \partial_i \tag{1.31}$$

The distinguished functions C^i play a crucial role in determining the topology of space. We have the following definition:

***Def* 1.10.** *Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator on a smooth manifold $\mathcal{M}$. A function $C : \mathcal{M} \to \mathbb{R}$ such that:*

$$\mathcal{J}\left(dC\right) = 0, \tag{1.32}$$

is called a Casimir invariant.

***Proposition* 1.5.** *Given an antisymmetric operator $\mathcal{J}$ with a Casimir invariant C, we have:*

$$\mathfrak{L}_X C = \{C,H\} = \mathcal{J}\left(dC,dH\right) = \mathcal{J}^{ij}C_iH_j = 0 \quad \forall H. \tag{1.33}$$

Therefore, regardless of the specific form of the Hamiltonian function H, a Casimir invariant is a constant of the flow generated by X. In particular, the orbit of a particle moving with velocity X always remain on the surface C =constant. Thus, if the shape of the level set (or *leaf*) C=constant is curved, the Casimir invariant induces a non-flat metric on this surface. It is clear that such topological constraint will have implications from a statistical standpoint, since the accessibility of space is limited.

***Remark* 1.6.** *In general, any 1-form $\theta \in ker\left(\mathcal{J}\right)$ is orthogonal to the conservative vector field $X = \mathcal{J}\left(dH\right)$ for any choice of H:*

$$i_X\theta = \theta_i\mathcal{J}^{ij}H_j = 0 \quad \forall H. \tag{1.34}$$

Given an antisymmetric operator $\mathcal{J}$ with a non-trivial kernel of dimension r, it turns out that the possibility of determining r Casimir invariants is a necessary condition for the validity of the Jacobi identity that defines Poisson operators (see proposition 5.2). This can be rephrased by saying that $ker\left(\mathcal{J}\right)$ must be *integrable.* Physically, $ker\left(\mathcal{J}\right)$ is the remnant of some process of reduction from microscopic dynamics, which is naturally described by canonical Hamiltonian vector fields. If $ker\left(\mathcal{J}\right)$ cannot be integrated, one cannot find a suitable coordinate change to recast the macroscopic dynamics in the form of a canonical Hamiltonian system. Therefore, one would like to know the mathematical conditions under which this is possible. This is the right juncture to introduce the notion of integrability. To do so, we need some new concepts. For the next definitions, we follow [91]. Denote with $\mathcal{M}^n$ a smooth manifold of dimension n. Consider two integers $r, k \in \mathbb{N}$ such that $r \leq n$ and $k \leq n - r$

***Def* 1.11.** *An r-dimensional distribution Δ_r on $\mathcal{M}^n$ assigns in a smooth fashion to each $\boldsymbol{x} \in \mathcal{M}^n$ an r-dimensional subspace $\Delta_r\left(\boldsymbol{x}\right)$ of the tangent space $T_{\boldsymbol{x}}\mathcal{M}^n$ to $\mathcal{M}^n$ at $\boldsymbol{x}$.*

***Def* 1.12.** *A k-dimensional integral manifold of Δ_r is a k-dimensional submanifold of $\mathcal{M}^n$ that is everywhere tangent to the distribution.*

***Def* 1.13.** *A distribution Δ_r is (completely) integrable if there are local coordinates $\left(x^1, ..., x^r, y^1, ..., y^{n-r}\right)$ on $\mathcal{M}^n$ such that the level sets y^1 =constant, ..., y^{n-r}=constant are $n-r$-dimensional integral manifolds of Δ_r.*

An r-dimensional distribution Δ_r can be conveniently represented through a set of $n-r$ linearly independent 1-forms θ_i:

$$\Delta_r = \left\{X \in T\mathcal{M}^n : \theta_i\left(X\right) = 0 \quad \forall i = 1, ..., n-r\right\}. \tag{1.35}$$

Here, $\theta_i\left(X\right) = i_X\theta_i$ is a contraction of the form θ_i with the vector field X. From now on, both notations (the bracket or the i symbol) will be used when dealing with contractions of forms.

***Def* 1.14.** *A distribution Δ_r is in involution if it is closed under brackets:*

$$[X, Y] \in \Delta_r \quad \forall X, Y \in \Delta_r. \tag{1.36}$$

In terms of the 1*-forms θ_i defining the distribution this is equivalent to:*

$$d\theta_i\left(X, Y\right) = X\theta_i\left(Y\right) - Y\theta_i\left(X\right) - \theta_i\left([X, Y]\right) = 0 \quad \forall X, Y \in \Delta_r, \qquad i = 1, ..., n-r. \tag{1.37}$$

The square bracket above $[X, Y] = XY - YX$ is the Lie bracket of vector fields. Finally, we arrive at the Frobenius integrability theorem:

***Theorem* 1.2.** *(Frobenius)*
In a local neighborhood $U \in \mathcal{M}$, the propositions below are equivalent:
i) Δ_r is in involution.
ii) Δ_r is integrable.
iii) The following identities hold:

$$d\theta_i\left(X, Y\right) = 0 \qquad \forall X, Y \in \Delta_r, \qquad i = 1, ..., n-r. \tag{1.38}$$

iv) There are 1-forms λ_{ij} such that locally:

$$d\theta_i = \sum_{j=1}^{r} \lambda_{ij} \wedge \theta_j \qquad \forall i = 1, ..., n-r. \tag{1.39}$$

v) The following identities hold:

$$\theta_1 \wedge ... \wedge \theta_{n-r} \wedge d\theta_i = 0 \qquad \forall i = 1, ..., n-r. \tag{1.40}$$

vi) There are functions ψ_{ij} and y_j such that locally:

$$\theta_i = \sum_{j=1}^{n-r} \psi_{ij} dy^j \qquad \forall i = 1, ..., n-r. \tag{1.41}$$

We refer the reader to [91, 92] for the proof of this theorem. In practice, the integrability of a distribution is checked by applying equation (1.40).

The following example is useful. Consider a smooth manifold $\mathcal{M}$ of dimension $2n + r$. Let $\mathcal{J}$ be an antisymmetric operator on $\mathcal{M}$ with rank $2n$. Construct a a cotangent basis $(\xi_1, ..., \xi_r)$ to span $ker\left(\mathcal{J}\right)$. The null space of these 1-forms defines a distribution $\Delta_\xi = \{X \in T\mathcal{M} : \xi_i = 0\}$. Construct a cotangent basis $(\theta_1, ..., \theta_{2n})$ with associate distribution $\Delta_\theta = \{X \in T\mathcal{M} : \theta_i = 0\}$ such that $T\mathcal{M} = \Delta_\xi \otimes \Delta_\theta$. Suppose that we verified integrability of Δ_ξ with (1.40). Then, since Δ_ξ is integrable, from (1.41) we can find a set of coordinates $\left(y^1, ..., y^{2n}, C^1, ..., C^r\right)$ such that the level sets C^1 =constant, ..., C^r =constant are integral manifolds of Δ_ξ, i.e. $\Delta^*_\theta = ker\left(\mathcal{J}\right) = span\left(\xi^1, ..., \xi^r\right) = span\left\{dC^1, ..., dC^r\right\}$. Here Δ^*_θ and Δ^*_ξ are the duals of Δ_θ and Δ_ξ.

We conclude this chapter with the statement of Carathéodory's theorem [137] relating the notion of accessibility and the integrability of a constraint. This theorem will be extremely useful in the final part of the present study.

***Theorem* 1.3.** *(Carathéodory)*
Let θ be a continuously differentiable non-vanishing 1-form on a smooth manifold

$\mathcal{M}$ of dimension n. Assume that the constraint $\theta = 0$ is not integrable. Thus, at some $\boldsymbol{x}_0 \in \mathcal{M}$ we have:

$$\theta \wedge d\theta \neq 0. \tag{1.42}$$

Then there is a neighborhood U of $\boldsymbol{x}_0$ such that any $\boldsymbol{y} \in U$ can be joined to $\boldsymbol{x}_0$ by a piecewise smooth path that is always tangent to the distribution.

Chapter 2

Statistical Mechanics in the Phase Space

The cornerstone of the standard formulation of statistical mechanics is Liouville's theorem, according to which the phase space volume is preserved by the canonical Hamiltonian flow. To understand the meaning of this assertion, we must define what phase space is.

***Def* 2.1.** *If it exists, the phase space is that set of coordinates* $\left(x^1, ..., x^{2n}\right)$ *such that* $\mathcal{J} = \mathcal{J}_c$ *is the symplectic operator* (1.26)*, or, equivalently,* $\omega = \omega_c$ *is the canonical symplectic* 2*-form* (1.25)*. The coupled coordinates* $\left(x^i, x^{n+i}\right) = \left(p^i, q^i\right)$ *are called canonical pairs.*

As an example, consider the phase space of a charged particle moving in an electromagnetic field. If $\boldsymbol{x} = (x, y, z)$ is the position in space and $\boldsymbol{v} = (v_x, v_y, v_z)$ the velocity of the point charge, it is well known that the three canonical pairs are $(\boldsymbol{p}, \boldsymbol{q}) = (m\boldsymbol{v} + q\boldsymbol{A}, \boldsymbol{x})$. Here, m is the particle mass, q the electric charge, and $\boldsymbol{A}$ the vector potential of the magnetic field $\boldsymbol{B}$, i.e. $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. If we write $\mathcal{J}$ in matrix notation by means of the canonical variables, we obtain:

$$\mathcal{J} = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.1}$$

Hence, in the tangent basis $\left(\partial_{p_x}, \partial_x, \partial_{p_y}, \partial_y, \partial_{p_z}, \partial_z\right)$, the components of $\mathcal{J}$ are constants, specifically the values 0, 1, or -1. Since $\mathcal{J}$ determines the geometrical properties of space, one then sees that the symplectic operator dictates a flat and homogeneous metric. The metric of the phase space is characterized by the following volume element:

***Def* 2.2.** *The phase space volume element is the volume 2n-form:*

$$vol_c = dp^1 \wedge dq^1 \wedge ... \wedge dp^n \wedge dq^n \tag{2.2}$$

Then, we have Liouville's theorem:

***Theorem* 2.1.** *(Liouville)*
A canonical Hamiltonian vector field X preserves the phase space volume vol_c.

Proof. Recalling that $\dot{p}^i = -\partial H/\partial q^i$ and $\dot{q}^i = \partial H/\partial p^i$,

$$\begin{aligned}\mathfrak{L}_X vol_c &= di_X vol_c \\ &= d\left\{\left[\dot{p}^i dp^1 \wedge ... \wedge dq^{i-1} \wedge dq^i \wedge ... \wedge dq^n - \dot{q}^i dp^1 \wedge ... \wedge dp^i \wedge dp^{i+1} \wedge ... \wedge dq^n\right]\right\} \\ &= \left[-\frac{\partial^2 H}{\partial p^i \partial q^i} + \frac{\partial^2 H}{\partial q^i \partial p^i}\right] vol_c \\ &= 0.\end{aligned} \tag{2.3}$$

□

The phase space volume vol_c is a special type of volume form called invariant measure:

***Def* 2.3.** *Let $\mathcal{M}$ be a smooth manifold of dimension n. Consider a real valued smooth function $g \neq 0$, $g \in C^\infty(\mathcal{M})$, $g: \mathcal{M} \to \mathbb{R}$. The n-dimensional volume form $vol^n = g dx^1 \wedge ... \wedge dx^n$ is called an invariant measure with respect to the vector field X if:*

$$\mathfrak{L}_X vol^n = div(X)\ vol^n = \frac{1}{g}\frac{\partial}{\partial x^i}\left(g X^i\right)\ vol^n = 0. \tag{2.4}$$

Thanks to Darboux's theorem 1.1, we know that any non-canonical Hamiltonian vector field locally admits canonical coordinates. It follows that, locally, any non-canonical Hamiltonian vector field is endowed with the invariant measure provided by Liouville's theorem.

An invariant volume form is needed for the standard formulation of statistical mechanics in virtue of 2 main reasons: 1. The ergodic theorem, which is used to invoke the ergodic hypothesis, holds for measure preserving transformations. 2. The conventional definition of entropy measure for continuous probability distributions requires an invariant volume form. First, let us review some basic facts about the ergodic theorem due to G. D. Birkhoff.

***Def* 2.4.** *A transformation $T : \mathcal{M} \to \mathcal{M}$ on a measure space $(\mathcal{M}, \mathcal{B}, \mu)$ is called measure preserving if:*

$$\forall \mathcal{A} \in \mathcal{B}, \qquad \mu\left(T^{-1}(\mathcal{A})\right) = \mu(\mathcal{A}). \tag{2.5}$$

In this notation, $\mathcal{B}$ is a so called σ-algebra on $\mathcal{M}$, and $\mu : \mathcal{B} \to [0,1]$ a probability measure on $\mathcal{M}$, i.e. $\mu(\mathcal{M}) = 1$ and $\mu(\emptyset) = 0$. Measure preserving transformations are at the basis of the concept of recurrence. A system is recurrent if it returns arbitrarily close to its original configuration after a sufficiently long time. J. H. Poincaré proved the following result:

***Theorem* 2.2.** *(Poincaré)*
Let $T : \mathcal{M} \to \mathcal{M}$ be a measure preserving transformation on a measure space $(\mathcal{M}, \mathcal{B}, \mu)$. Then, for all measurable sets $\mathcal{E} \in \mathcal{B}$ there exists $\mathcal{F} \subseteq \mathcal{E}$ with $\mu(\mathcal{F}) = \mu(\mathcal{E})$ such that $\forall \boldsymbol{x} \in \mathcal{F}$ we can find infinite many non-zero integers $n_i \in \mathbb{N}$ satisfying:

$$T^{n_i}\boldsymbol{x} \in \mathcal{E}. \tag{2.6}$$

This result can be summarized by saying that, if T is measure preserving, almost every point in $\mathcal{E}$ returns to $\mathcal{E}$ as many times as desired by a sufficient number of iterations.

***Def* 2.5.** *A measure preserving transformation $T : \mathcal{M} \to \mathcal{M}$ on a measure space $(\mathcal{M}, \mathcal{B}, \mu)$ is called ergodic if:*

$$\forall \mathcal{E} \in \mathcal{B}, \qquad T^{-1}(\mathcal{E}) = \mathcal{E} \quad \Longrightarrow \quad \mu(\mathcal{E}) = 1 \;\; \textit{or} \;\; \mu(\mathcal{E}) = 0. \tag{2.7}$$

The statement of the ergodic theorem is the following:

***Theorem* 2.3.** *(Ergodic theorem)*
Let $T : \mathcal{M} \to \mathcal{M}$ be an ergodic (and measure preserving) transformation on a measure space $(\mathcal{M}, \mathcal{B}, \mu)$ and $f \in L^1(\mathcal{M}, \mathcal{B}, \mu)$. Then:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f\left(T^j \boldsymbol{x}\right) = \int f \, d\mu \qquad \textit{a.e.} \tag{2.8}$$

Discussing the notion of ergodic transformation and measure space is beyond the scope of the present study. We refer the reader to [25, 26, 139] for proofs and additional details. Physically, the ergodic theorem states that, under appropriate hypothesis (among which measure preservation), the time average of a certain observable f converges to its spatial mean. This fact is a direct consequence of ergodicity, which ensures that trajectories explore the whole space. In the context of canonical Hamiltonian mechanics, the measure preserving transformation T is provided by the flow generated by the canonical vector field X, while the probability measure is given by the probability distribution in the phase space times the phase space volume, $d\mu = f vol_c$. However, we must remark that the transformation T generated by a canonical flow is not, in general, ergodic. This can be understood by observing that if the Hamiltonian function has some symmetry, the resulting orbit will lie on the level set of the corresponding constant of motion, and thus certain regions of the phase space would be precluded to the dynamics. Although in practical situations verifying whether a certain transformation satisfies the ergodic property may be challenging, one can easily check if a physical system has an invariant measure by solving equation (2.4) with respect to g. If such a function g can be found, one usually *assumes* the ergodic hypothesis on the volume form vol^n, and exchanges spatial means with time averages. This substitution may have great advantages: the ergodic hypothesis is at the basis of the procedure by which complex interactions in a physical system are represented through simplified random processes such as Gaussian white noise or Brownian motion. Clearly, such simplification, which ultimately enables the derivation of the evolution equation for the probability distribution of the system, cannot be made without an invariant measure. The importance of Liouville's theorem should now be very clear.

What about the relationship between entropy and Liouville's theorem? In order to clarify this point, a few considerations on the concept of entropy are now necessary. First of all, we must distinguish between thermodynamic entropy $\mathcal{S}$ and information entropy S. The former is the physical quantity appearing in the celebrated second law of thermodynamics. Various formulations of this fundamental law have been proposed. Below we report those due to Lord Kelvin, R. Clausius, and C. Carathéodory [140, 44]:

Lord Kelvin: *'In no quasi-static cyclic process can a quantity of heat be converted entirely into its mechanical equivalent of work'.*

R. Clausius: *'There exists no thermodynamic transformation whose sole effect is to extract a quantity of heat from a colder reservoir and deliver it to a hotter reservoir'.*

C. Carathéodory: *'In every neighborhood of every thermodynamic state x there are states y that are not accessible from x via quasi-static adibatic paths, that is, paths along which $Q = 0$'.*

In the last statement, Q is the heat 1-form. Together with the work 1-form W, representing the work done by the system, the first law of thermodynamics for the change in internal energy U reads $dU = Q - W$. It is worth to note that, while Kelvin's and Clausius' formulations are equivalent, Carathédory's ansatz is a weaker requirement. Using Carathédory's statement of the second law, it is possible to show that the heat 1-form Q defines a constraint $Q = 0$ (representing adiabatic transformations) that is integrable in the sense of the Frobenius theorem 1.2. Thus locally one can find appropriate functions T and $\mathcal{S}$ such that:

$$Q = Td\mathcal{S}. \tag{2.9}$$

Equation (2.9) locally defines thermodynamic entropy $\mathcal{S}$ and temperature T. Furthermore, it can be shown that the second law of thermodynamics determines the sign of the rate of change in entropy with respect to time in a closed system:

$$\frac{d\mathcal{S}}{dt} \geq 0. \tag{2.10}$$

The ineluctable nature of thermodynamic entropy increase describes the irreversibility of any isolated physical process.

Now consider a probabilistic set of events represented by a discrete probability distribution $p_1, ..., p_n$. The information entropy S measures the degree at which we are uncertain of the outcome of such probabilistic process. In his influential study, C. E. Shannon proved the following theorem [108]:

***Theorem* 2.4.** *(Shannon)*
Let $S = S(p_1, ..., p_n)$ be the entropy measure of a probabilistic set of events with discrete probability distribution $p_1, ..., p_n$. Suppose that:
1. S is continuous in the p_i.
2. When all the probabilities are equal with values $p_i = 1/n$, $i = 1, ..., n$, S is a monotonic increasing function of n.
3. If a choice is divided in multiple choices, the total entropy S is given by the weighted sum of the individual entropies.
Then, the only function of the p_i satisfying these requirements is:

$$S = -c \sum_{i=1}^{n} p_i \log p_i. \tag{2.11}$$

where $c \in \mathbb{R}_{>0}$ is an arbitrary normalization constant.

For the proof see [108]. From a mathematical standpoint the hypothesis made in Shannon's theorem are not necessary. However, they have important physical implications. On one hand, the function S defined in equation (2.11) vanishes only when all the probabilities except one are zero, i.e.:

$$S = 0 \iff p_i = 0 \quad \forall i \neq k, \quad p_k = 1. \tag{2.12}$$

This situation correspond to complete knowledge of the outcome of the probabilistic process, because the only possible event is that with probability p_k. Accordingly, the entropy is minimum. On the other hand, when all the events have the same probabilities $p_i = 1/n$, $i = 1, ..., n$, the function S takes its maximum value $\log n$, describing a state of maximal uncertainty. Finally, S is *additive*, i.e. the uncertainty $S(A, B)$ of a joint event (A, B) satisfies:

$$S(A, B) = S(A) + S(B|A) \leq S(A) + S(B). \tag{2.13}$$

with equality holding when the events A and B are independent.

A central question is if and when thermodynamic and information entropy coincide, that is if and when $\mathcal{S} = S$. As we will discuss later on, this problem is strictly related to how does the information entropy S look like when the probability distribution is continuous, i.e. when a notion of metric space is introduced. Unfortunately, $\mathcal{S} \neq S$ in general.

Consider a physical system described by a continuous probability distribution $P \in C^{\infty}(\Omega)$ on a volume element vol^n, with $P : \Omega \to \mathbb{R}_{\geq 0}$, $\Omega \subset \mathcal{M}^n$, and:

$$\int_{\Omega} P \, vol^n = 1. \tag{2.14}$$

The information entropy S is defined as follows:

***Def* 2.6.** *The information entropy of the probability distribution P is:*

$$S = -\int_{\Omega} P \log P \, vol^n. \tag{2.15}$$

One soon realizes that definition 2.6 is not covariant because the value of S depends on the specific measure vol^n chosen to define P [104, 105, 106]. To see this, we introduce a novel coordinate system with volume element vol_*^n related to the original one as:

$$vol_*^n = g vol^n. \tag{2.16}$$

Here, g is the Jacobian determinant of the transformation. Probability is a number, i.e. a scalar quantity. Therefore the probability distribution f on the new volume element must satisfy:

$$f vol_*^n = P vol^n. \tag{2.17}$$

Then, we obtain $fg = P$. Let Σ be the information entropy of f. It follows that:

$$\Sigma = -\int_{\Omega} f \log f \, vol_*^n = S + \int_{\Omega} P \log g \, vol^n = S + \langle \log g \rangle \neq S. \tag{2.18}$$

The angle bracket $\langle \, , \, \rangle$ denotes ensemble averaging. Note that the discrepancy between Σ and S is completely determined by the metric factor g. This result points to the fact that, while definition 2.6 is mathematically consistent as an information measure (i.e. as a measure of the scattering of a distribution), it is empty of any physical information. The entropy measure S can acquire physical significance if we specify the physically correct metric g containing the missing information on the structure of space where dynamics occurs. Defining the exact meaning of 'correct metric' will be a key topic of the present study. Below we give a concrete example of the unphysical results one can obtain by a wrong assumption of g.

Suppose that $vol_c = dp_x \wedge dq_x \wedge dp_y \wedge dq_y \wedge dp_z \wedge dq_z$ is the 6-dimensional canonical phase space of a molecule in an isolated neutral gas occupying a volume Ω. Let f be the probability distribution on vol_c of an ensemble of such particles. The second law of thermodynamics demands that the probability distribution of thermodynamic equilibrium maximizes the entropy of the system, while keeping the total energy E and the total particle number N constant. If H is the energy of a molecule, this condition can be written in the form of a variational problem with respect to the function f:

$$\delta \left(\mathcal{S} - \beta E - \gamma N \right) = \delta \left[\int_{\Omega} \left(-f \log f - \beta f H - \gamma f \right) vol_c \right] = 0. \tag{2.19}$$

Here, β and γ are Lagrange multipliers and we assumed that the thermodynamic entropy $\mathcal{S}$ coincides with the information entropy of f. The solution of the variational problem is:

$$f = \exp\{-\beta H - 1 - \gamma\}. \tag{2.20}$$

The multiplier γ serves as a normalization constant. One sees that (2.20) is the standard Maxwell-Boltzmann distribution of a neutral gas. Now, instead of the distribution f on vol_c, consider the distribution P on $vol_* = g vol_c$. Further assume that the information entropy of P coincides with the thermodynamic entropy $\mathcal{S}$. With the same procedure, one arrives at:

$$P = \exp\{-\beta H - 1 - \gamma\}. \tag{2.21}$$

But then:

$$f = g \exp\{-\beta H - 1 - \gamma\}. \tag{2.22}$$

Unless g is a constant, this function is not the expected Maxwell-Boltzmann distribution.

From experiments, we know that the correct result is (2.20). The fact is that the phase space volume vol_c where we defined f is 'special', in the sense that it is an invariant measure due to Liouville's theorem 2.1. We will see that this is the reason why the variation worked in the first case, while the second choice vol_* gave a wrong result.

Chapter 3

Stochastic Calculus and The Fokker-Planck Equation

This last introductory chapter is dedicated at reviewing the theory behind the Fokker-Planck equation, a second order partial differential equation describing the time evolution of a probability distribution. The derivation of the Fokker-Planck equation is based on the concept of Wiener process, a type of random process that exemplifies Brownian motion. For the results presented in this chapter, we refer the reader to [119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131].

The Wiener process W is defined as below:

***Def* 3.1.** *(Wiener process)*
The Wiener Process (or Brownian motion) is a real-valued stochastic process $W(\cdot)$ with the properties:

$$1. \quad W(0) = 0 \quad a.s., \tag{3.1a}$$

$$2. \quad W(t) - W(s) \sim N(0, t-s) \quad \forall t \geq s \geq 0, \tag{3.1b}$$

$$3. \quad \begin{array}{c} W(t_1), W(t_1) - W(t_2), ..., W(t_{n-1}) - W(t_n) \\ are \;\; independent \;\; \forall 0 < t_1 < t_2 < ... < t_n. \end{array} \tag{3.1c}$$

The first equation (3.1a) states that, at time $t = 0$, the Wiener process is 0 almost surely, i.e. with probability 1. The second equation (3.1b) implies that the increments $W(t) - W(s)$ are distributed as a normal distribution of mean zero and variance $t - s$. The third equation (3.1c) requires that the increments $W(t_{i-1}) - W(t_i)$ are independent random variables. If $\langle\cdot\rangle$ denotes the expectation value, we see that:

$$\langle W(t)\rangle = \langle W(t) - W(0)\rangle = 0, \tag{3.2a}$$

$$\left\langle W^2(t)\right\rangle = \left\langle (W(t) - W(0))^2\right\rangle = t, \tag{3.2b}$$

$$\langle W(t)\, W(s)\rangle = \langle W(s)\rangle \langle W(t) - W(s)\rangle + \left\langle W^2(s)\right\rangle = s \qquad \forall t \geq s \geq 0. \tag{3.2c}$$

If we identify the value W with the spatial position x of a Brownian particle, we obtain a random process in which the expectation value of x is the initial condition $x(0) = 0$ and the expectation value of x^2 the elapsed time t.

The Wiener process W can be conveniently related to Gaussian white noise Γ. Specifically, Γ can be thought as the time derivative of W in the sense we specify below.

***Def* 3.2.** *(White noise)*
Gaussian white noise is a real-valued stochastic process $\Gamma(\cdot)$ *satisfying the properties:*

$$\langle\Gamma\rangle = 0, \tag{3.3a}$$

$$\langle\Gamma(t)\,\Gamma(s)\rangle = \delta(t - s) \qquad \forall t, s \in \mathbb{R}_{\geq 0}. \tag{3.3b}$$

Then, we have an alternative definition of Wiener process in terms of Γ:

$$dW = W(t + dt) - W(t) = \Gamma dt. \tag{3.4}$$

This equation has to be interpreted in the sense that:

$$\lim_{h\to 0}\left\langle \frac{W(t+h) - W(t)}{h}\frac{W(s+h) - W(s)}{h}\right\rangle = \delta(t - s) \qquad \forall t, s \in \mathbb{R}_{\geq 0}. \tag{3.5}$$

A direct application of the Wiener process is stochastic calculus. As an example, we can write the equation of motion of a Brownian particle in the form of a stochastic differential equation:

$$dX = dW. \tag{3.6}$$

In this notation, the upper case letter X specifies that x is now a random variable. More generally, one can consider the following class of multivariate stochastic differential equations:

$$dX^i = F^i(\boldsymbol{X}, t)\, dt + G^{ij}(\boldsymbol{X}, t)\, dW_j, \qquad i = 1, ..., n, \quad j = 1, ..., m. \tag{3.7}$$

Here, $\boldsymbol{X} = \left(X^1, ..., X^n\right)$. F^i is a deterministic term corresponding to the Newtonian force (or velocity if X is a position) and mathematically represents ordinary calculus.

The (generally non-square) matrix G^{ij} accounts for the randomness of the system. Notice that the number of independent random processes appearing on the right-hand side is m. It is also worth to note that W_j does not need to be a Wiener process, but may describe other kinds of stochastic motion. However, in this study we will be concerned only with Brownian motion.

Equation (3.7) is not a complete mathematical definition of stochastic differential equation. Indeed, for (3.7) to make mathematically sense, a notion of stochastic integral has to be introduced. Depending on such definition, the result of integration of equation (3.7) will change. As one may expect, this arbitrariness is transferred to the chain rule of stochastic calculus, which varies according to a parameter $\alpha \in [0, 1]$. Physically, α reflects the properties of the actual process whose limiting representation is the Wiener process appearing in the stochastic differential equation. We will not discuss in depth this issue, and refer the reader to [123, 131, 124] and references therein for additional details. The chain rule of stochastic calculus is the following:

***Theorem* 3.1.** *(Generalized Ito's lemma)*
Let f be a real valued function. Define the stochastic integral of f according to the following mean-square limit:

$$\int_u^v f dW = ms\text{-}\lim_{n\to\infty} \sum_{i=1}^{n} f(t_{i-1} + \alpha \Delta t_i)[W(t_i) - W(t_{i-1})]. \tag{3.8}$$

Here $\alpha \in [0, 1]$. Let $X(\cdot)$ be a real-valued stochastic process obeying:

$$X(v) = X(u) + \int_u^v F dt + \int_u^v G_\alpha dW. \tag{3.9}$$

Here $F \in L^1(0, T)$, $G \in L^2(0, T)$ and $0 \leq u \leq v \leq T$. Let $Y = y(X(t), t)$ be a stochastic process depending on X, with $y : \mathbb{R} \times [0, T] \to \mathbb{R}$ and $y \in C^2$. Then, the stochastic differential of Y reads:

$$\begin{aligned} dY(X,t) &= \frac{\partial y}{\partial t}dt + \frac{\partial y}{\partial x}dX + \left(\frac{1}{2} - \alpha\right) G^2 \frac{\partial^2 y}{\partial x^2} dt \\ &= \left(\frac{\partial y}{\partial t} + F\frac{\partial y}{\partial x} + \left(\frac{1}{2} - \alpha\right) G^2 \frac{\partial^2 y}{\partial x^2}\right) dt + G_\alpha \left(\frac{\partial y}{\partial x}\right)_\alpha dW. \end{aligned} \tag{3.10}$$

Here, the subscript α denotes evaluation at $t_\alpha = t_{i-1} + \alpha \Delta t_i$.

See [82] for the proof of this result. Notice that, when $\alpha = 1/2$, the chain rule of standard calculus is recovered. This choice is known as Stratonovich integral. The case $\alpha = 0$ is referred to as Ito's integral. Equation (3.10) can be generalized to the multivariate case:

***Theorem* 3.2.** *(Multivariate Ito's Lemma)*
Let $y : \mathbb{R}^n \times [0,T] \to \mathbb{R}$ be a real-valued function of n random variables $\boldsymbol{X}(\cdot)$ and let y be of class C^2 on its domain. Let the stochastic integral be defined as in theorem 3.1 and take $F^i \in L^1(0,T)$, $G^{ij} \in L^2(0,T)$ such that:

$$dX^i = F^i dt + G^{ij}_{\alpha} dW_j, \tag{3.11}$$

where the subscript α denotes evaluation at $t_\alpha = t_{i-1} + \alpha \Delta t_i$. Then, the stochastic differential of y reads:

$$\begin{aligned} dY(\boldsymbol{X},t) &= \frac{\partial y}{\partial t} dt + \frac{\partial y}{\partial x^i} dX^i + \left(\frac{1}{2} - \alpha\right) G^{ik} G^{jk} \frac{\partial^2 y}{\partial x^i \partial x^j} dt \\ &= \left(\frac{\partial y}{\partial t} + F^i \frac{\partial y}{\partial x^i} + \left(\frac{1}{2} - \alpha\right) G^{ik} G^{jk} \frac{\partial^2 y}{\partial x^i \partial x^j}\right) dt + G^{ij}_{\alpha} \left(\frac{\partial y}{\partial x^i}\right)_{\alpha} dW_j. \end{aligned} \tag{3.12}$$

A standard result of stochastic analysis is the following:

***Theorem* 3.3.** *(Fokker-Planck equation) The time evolution of the probability distribution f of a set of random variables obeying the stochastic differential equation* (3.7) *is given by the Fokker-Planck equation:*

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial x^i} \left[-F^i f + \frac{1}{2} \frac{\partial}{\partial x^j} \left(G^{ik} G^{jk} f \right) - \alpha \frac{\partial G^{ik}}{\partial x^j} G^{jk} f \right]. \tag{3.13}$$

See [123, 131] for proofs.

Part II

Classification of Conservative Dynamics

Chapter 4

Geometrical Classification of Antisymmetric Operators

In this chapter we discuss the geometrical properties of antisymmetric operators. It is shown that antisymmetric operators can be categorized according to a hierarchical structure that reflects the departure from flat phase space. This hierarchy is at the basis of the statistical theory we develop.

4.1 The Jacobiator

In order to obtain the hierarchy, we need a mathematical measure of the degree at which the space defined by a certain antisymmetric operator departs from phase space. Clearly, the first discriminating factor is provided by the Jacobi identity of Poisson operators, equation (1.13). Therefore, the following measure[1] is useful:

***Def* 4.1.** *Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator. The trivector $\mathfrak{G} \in \bigwedge^3 T\mathcal{M}$ defined by:*

$$\mathfrak{G} = \mathfrak{G}\left(\mathcal{J}\right) = \frac{1}{2}\mathcal{J}^{im}\frac{\partial \mathcal{J}^{jk}}{\partial x^m}\partial_i \wedge \partial_j \wedge \partial_k, \tag{4.1}$$

is called the Jacobiator of $\mathcal{J}$.

***Proposition* 4.1.** *The Jacobiator $\mathfrak{G}$ measures the failure of the Jacobi identity.*

Proof. We have:

$$\mathfrak{G} = \sum_{i<j<k}\left(\mathcal{J}^{im}\frac{\partial \mathcal{J}^{jk}}{\partial x^m} + \mathcal{J}^{jm}\frac{\partial \mathcal{J}^{ki}}{\partial x^m} + \mathcal{J}^{km}\frac{\partial \mathcal{J}^{ij}}{\partial x^m}\right)\partial_i \wedge \partial_j \wedge \partial_k. \tag{4.2}$$

Thus, $\mathfrak{G} = 0$ if and only if $\mathcal{J}$ satisfies the Jacobi identity (1.13). □

[1] See [21, 138] for a definition in terms of the Schouten bracket.

The Jacobiator $\mathfrak{G}$ only measures the component of the current of $\mathcal{J}$ (which we will define afterward) 'aligned' with $\mathcal{J}$ itself. Even if the aligned component may be different from zero (and thus the system is not Hamiltonian), the 'normal' component of the current may vanish. Operators with this second property have notable features both from the mechanical and statistical point of view and will be discussed later on. We remark that here we used the words 'aligned' and 'normal' in reference to the 3-dimensional case (which will be discussed later) where the current of the operator can be decomposed in a parallel and a normal component.

4.2 The Conformal Operator

We already know that Hamiltonian systems, either canonical or not, fall in the category $\mathfrak{G} = 0$, because both symplectic and Poisson operators satisfy the Jacobi identity (remember definition 1.7). The next step is then to determine the type of dynamics that is the most similar to Hamiltonian mechanics. The answer to this problem is closely related to the possibility of transforming a non-Hamiltonian system with $\mathfrak{G} \neq 0$ to an Hamiltonian one by some appropriate method that is not limited to a spatial coordinate change (remember that a coordinate change is not enough for non-Hamiltonian systems because the Darboux theorem 1.1 locally assigns canonical coordinates only when $\mathfrak{G} = 0$). A system admitting such procedure would be closer to Hamiltonian mechanics than those systems to which canonical phase space cannot be assigned. Since a spatial coordinate change is not sufficient, we are left with two possibilities: performing a time reparametrization or increasing the number of variables.

In the following sections we will show that there are certain classes of operators that can be transformed to a Poisson operator by a combination of coordinate change, time reparametrization, and variable increase without altering the original dynamics.

Systems that can be transformed to a Poisson operator by a time reparametrization are already known in the literature and are called conformal. This idea originates from the work of [18] and is discussed in [20], where a time reparametrization is used to 'Hamiltonize' (i.e. transform to an Hamiltonian form) certain nonholonomic systems with symmetries and the notion of conformal bracket is introduced. In this section, we cast the concept of conformal system in the perspective of the geometrical hierarchy of antisymmetric operators, and define the conformal operator.

Consider a conservative vector field $X = \mathcal{J}(dH)$. Evidently, this representation does not contain any notion of time. The standard interpretation is that:

$$X = \frac{d\boldsymbol{x}}{dt} = \mathcal{J}\left(dH\right). \tag{4.3}$$

However, we may think that more generally:

$$X = r\left(\boldsymbol{x}\right)\frac{d\boldsymbol{x}}{d\tau} = \mathcal{J}\left(dH\right), \tag{4.4}$$

where $r : \mathcal{M} \to \mathbb{R}_{>0}$ is a $C^{\infty}\left(\mathcal{M}\right)$ function called *conformal factor* and τ is a *proper time*. It is worth noticing that in an even more general setting one could introduce multiple time variables. However, we will not explore this possibility here.

The proper time τ can be related to the standard time variable t by the differential equation:

$$\frac{dt}{d\tau} = r^{-1}. \tag{4.5}$$

If one integrates equation (4.4) and obtains $\boldsymbol{x}\left(\tau\right)$, the orbit in time t can then be calculated as $\boldsymbol{x}\left(\tau\left(t\right)\right)$ by integration of equation (4.5). Now the objective is clear: if we can find a conformal factor r such that the vector field:

$$Y = \frac{X}{r} = r^{-1}\mathcal{J}\left(dH\right), \tag{4.6}$$

is Hamiltonian, we are done. Therefore, we have the following definition:

***Def* 4.2.** *(Conformal operator)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator. $\mathcal{J}$ is called a conformal operator (or conformally Poisson operator) if there exists a non-zero function $r : \mathcal{M} \to \mathbb{R}_{>0}$, $r \in C^{\infty}\left(\mathcal{M}\right)$, such that:

$$\mathfrak{G}\left(r^{-1}\mathcal{J}\right) = r^{-1}\mathcal{J}^{im}\frac{\partial\left(r^{-1}\mathcal{J}^{jk}\right)}{\partial x^m}\partial_i \wedge \partial_j \wedge \partial_k = 0. \tag{4.7}$$

There is an equivalent concept in terms of the 2-form ω:

***Def* 4.3.** *(Conformal 2-form)*
Let $\omega \in \bigwedge^2 T^*\mathcal{M}$ be a vorticity 2-form. ω is called a conformal 2-form (or conformally symplectic 2-form) if it is conformally closed, i.e. there exists a non-zero function $r : \mathcal{M} \to \mathbb{R}_{>0}$, $r \in C^{\infty}\left(\mathcal{M}\right)$, such that:

$$d\left(r\omega\right) = 0. \tag{4.8}$$

The equivalence of these concepts can be understood by considering an invertible conformal operator $\mathcal{J}$ with conformal factor r^{-1}. Clearly, if ω is the inverse, $d\left(r\omega\right) = 0$.

We have the following:

***Example* 4.1.** *A Poisson operator is conformal with $r = 1$.*

The physical meaning of the conformal factor r will be discussed later. It should be noted that it is not always possible to find a conformal factor.

4.3 Covorticity and Cocurrent

In order to explore the next class of antisymmetric operators we need to introduce a new representation of the antisymmetric operator $\mathcal{J}$ in terms of differential forms. In the following, the prefix co- is used to distinguish quantities defined in terms of the antisymmetric operator $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ from those defined in terms of the vorticity 2-form $\omega \in \bigwedge^2 T^*\mathcal{M}$.

***Def* 4.4.** *(Covorticity)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator. Let $vol^n = gdx^1 \wedge ... \wedge dx^n$ be a volume element on $\mathcal{M}$. The covorticity $n-2$ form with respect to vol^n reads:

$$\mathcal{J}^{n-2} = i_{\mathcal{J}} vol^n. \tag{4.9}$$

The covorticity 2-form is defined as:

$$\mathcal{J}^2 = *\mathcal{J}^{n-2}. \tag{4.10}$$

In order to understand the meaning of this definition, it is useful to evaluate explicitly equation (4.9):

$$\begin{aligned}
\mathcal{J}^{n-2} &= \left(i_{\mathcal{J}} g dx^1 \wedge ... \wedge dx^n\right) \\
&= \sum_{i<j} (-1)^{i+j-1} g\mathcal{J}^{ij} \left(i_{\partial_i \wedge \partial_j} dx^i \wedge dx^j\right) \wedge dx^{n-2}_{ij} \\
&= \sum_{i<j} (-1)^{i+j-1} g\mathcal{J}^{ij} i_{(\partial_i \otimes \partial_j - \partial_j \otimes \partial_i)} \left(dx^i \otimes dx^j - dx^j \otimes dx^i\right) \wedge dx^{n-2}_{ij} \\
&= 2\sum_{i<j} (-1)^{i+j-1} g\mathcal{J}^{ij} dx^{n-2}_{ij}.
\end{aligned} \tag{4.11}$$

In this notation $dx^{n-2}_{ij} = dx^1 \wedge ... \wedge dx^{i-1} \wedge dx^{i+1} \wedge ... \wedge dx^{j-1} \wedge dx^{j+1} \wedge ... \wedge dx^n$. The last term in equation (4.11) can also be thought as an alternative definition of the covorticity $n-2$ form.

By analogy with electromagnetism where the Faraday tensor $\mathcal{F}$ acts as vorticity 2-form and $d * \mathcal{F} = J$ represents the current 3-form, we can now introduce a notion of current associated to both vorticity 2-forms and antisymmetric operators in the following manner:

***Def* 4.5.** *(Current of vorticity 2-form)*
Let $\omega \in \bigwedge^2 T^*\mathcal{M}$ be a vorticity 2-form on a smooth manifold $\mathcal{M}$ of dimension n. The current $n-1$ form is defined as:

$$J^{n-1} = d * \omega. \tag{4.12}$$

The current 1-form is defined as:

$$J^1 = *J^{n-1} = \delta\omega, \tag{4.13}$$

where $\delta = *d*$ is the codifferential.

***Def* 4.6.** *(Cocurrent of antisymmetric operators)*
The cocurrent $n-1$ form of an antisymmetric operator $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ with respect to the volume form vol^n on $\mathcal{M}$ is:

$$\mathcal{O}^{n-1} = d * \mathcal{J}^2 = d\mathcal{J}^{n-2}. \tag{4.14}$$

The cocurrent 1-form is defined as:

$$\mathcal{O}^1 = *\mathcal{O}^{n-1} = \delta\mathcal{J}^2, \tag{4.15}$$

where $\delta = *d*$ is the codifferential.

In the same way the closeness of the 2-form ω defines Hamiltonian mechanics, the closeness of the $n-2$ form $\mathcal{J}^{n-2}$ is a powerful condition. Indeed, we can prove the fundamental result:

***Proposition* 4.2.** *(Existence of invariant measure)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator. The conservative vector field $X = \mathcal{J}(dH)$ admits an invariant measure vol^n for any choice of the Hamiltonian function H if and only if $\mathcal{O}^{n-1} = d\mathcal{J}^{n-2} = 0$ on the volume form vol^n:

$$\mathfrak{L}_X vol^n = 0 \quad \forall H \quad \iff \quad \mathcal{O}^{n-1} = 0 \quad on \quad vol^n. \tag{4.16}$$

Proof. First, let us show that if $\mathcal{O}^{n-1} = 0$ then $\mathfrak{L}_X vol^n = 0$ for any choice of the function H. We have:

$$\begin{aligned}
\mathcal{O}^{n-1} &= d\mathcal{J}^{n-2} \\
&= d\left(i_{\mathcal{J}} g dx^1 \wedge ... \wedge dx^n\right) \\
&= d\sum_{i<j} (-1)^{i+j-1} g\mathcal{J}^{ij} \left(i_{\partial_i \wedge \partial_j} dx^i \wedge dx^j\right) \wedge dx_{ij}^{n-2} \\
&= d\sum_{i<j} (-1)^{i+j-1} g\mathcal{J}^{ij} i_{(\partial_i \otimes \partial_j - \partial_j \otimes \partial_i)} \left(dx^i \otimes dx^j - dx^j \otimes dx^i\right) \wedge dx_{ij}^{n-2} \\
&= 2d\sum_{i<j} (-1)^{i+j-1} g\mathcal{J}^{ij} dx_{ij}^{n-2} \\
&= (-1)^{i+j-1} \frac{\partial}{\partial x^k} \left(g\mathcal{J}^{ij}\right) dx^k \wedge dx_{ij}^{n-2} \\
&= 2(-1)^j \frac{\partial \left(g\mathcal{J}^{ij}\right)}{\partial x^i} dx_j^{n-1} \\
&= 0.
\end{aligned} \tag{4.17}$$

On the other hand:

$$\mathfrak{L}_X vol^n = \left[\frac{\partial}{\partial x^i}\left(g\mathcal{J}^{ij}H_j\right)\right] dx^1 \wedge ... \wedge dx^n = \frac{\partial \left(g\mathcal{J}^{ij}\right)}{\partial x^i} H_j dx^1 \wedge ... \wedge dx^n. \tag{4.18}$$

If we want this equation to vanish for every H, we must have:

$$\frac{\partial \left(g\mathcal{J}^{ij}\right)}{\partial x^i} = 0 \quad \forall j = 1, ..., n. \tag{4.19}$$

This is the same condition appearing in equation (4.17) and the first implication follows. Now it is also clear that if equation (4.18) vanishes for all H, we must have $\mathcal{O}^{n-1} = 0$ and the statement is proved. $\square$

We conclude this section by noting that in the proof of the previous proposition we have obtained the explicit expression of $\mathcal{O}^{n-1}$:

$$\mathcal{O}^{n-1} = d\mathcal{J}^{n-2} = 2(-1)^j \frac{\partial \left(g\mathcal{J}^{ij}\right)}{\partial x^i} dx_j^{n-1}. \tag{4.20}$$

4.4 The Measure Preserving Operator

The essential result of proposition 4.2 is that it introduces a notion of invariant measure that does not depend on the specific choice of the Hamiltonian H, but only on the geometrical properties of the operator $\mathcal{J}$. To know whether a certain operator $\mathcal{J}$ admits such kind of Hamiltonian-independent invariant measure it is therefore sufficient to determine whether a metric g can be found such that $\mathcal{O}^{n-1} = 0$. At this point, it is natural to define the *measure preserving operator*:

***Def* 4.7.** *An antisymmetric operator $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ on a smooth manifold $\mathcal{M}$ of dimension n is called measure preserving if there exists a volume form vol^n on $\mathcal{M}$ such that $\mathcal{O}^{n-1} = d\mathcal{J}^{n-2} = 0$.*

***Remark* 4.1.** *An antisymmetric operator can be measure preserving without satisfying the Jacobi identity* (1.13)*, i.e. without being a Poisson operator.*

We have the following:

***Example* 4.2.** *A constant rank conformal operator is measure preserving.*

Proof. Let $\mathcal{J}$ be conformal with conformal factor $r \in C^\infty(\mathcal{M})$. We must show that there is a Jacobian $g \in C^\infty(\mathcal{M})$ solving the divergence equation:

$$\frac{\partial\left(g\mathcal{J}^{ij}\right)}{\partial x^i} = 0 \quad \forall j = 1, ..., n. \tag{4.21}$$

Since $r^{-1}\mathcal{J}$ satisfies the Jacobi identity and has constant rank, the Darboux theorem 1.1 states that for every point $P \in \mathcal{M}$ there is a neighborhood $\mathcal{U} \subset \mathcal{M}$ such that the Hamiltonian vector field $X = r^{-1}\mathcal{J}(dH)$ takes the form:

$$X = \sum_{i=1}^{m} \dot{p}^i \partial_{p^i} + \dot{q}^i \partial_{q^i} + \sum_{j=1}^{k} \dot{C}^j \partial_{C^j} = \sum_{i=1}^{m} -H_{q^i}\partial_{p^i} + H_{p^i}\partial_{q^i}. \tag{4.22}$$

Here, the rank of $\mathcal{J}$ is $2m = n - k$ and the C^j are k constants of motion (Casimir invariants) with $\dot{C}^j = 0$. Due to Liouville's theorem 2.1 for Hamiltonian vector fields, one sees that the volume form:

$$vol_I^n = dp^1 \wedge dq^1 \wedge ... \wedge dp^m \wedge dq^m \wedge dC^1 \wedge ... \wedge dC^r, \tag{4.23}$$

is an invariant measure. Indeed, in light of (4.22):

$$\mathfrak{L}_X vol_I^n = \left(\sum_{i=1}^{m} \frac{\partial \dot{p}^i}{\partial p^i} + \frac{\partial \dot{q}^i}{\partial q^i}\right) vol_I^n = 0. \tag{4.24}$$

Let G be the Jacobian of the coordinate change $vol^n = dx^1 \wedge ... \wedge dx^n = G^{-1} vol_I^n$. In these coordinates equation (4.21) becomes:

$$\frac{\partial\left(Gr^{-1}\mathcal{J}^{ij}\right)}{\partial x^i} = 0 \quad \forall j = 1, ..., n. \tag{4.25}$$

Thus, the desired Jacobian is $g = Gr^{-1}$. □

Since a Poisson operator is conformal, it also follows that a constant rank Poisson operator is measure preserving. It is worth noticing that the invariant measure

$g = Gr^{-1}$ obtained in the previous example concerns the standard time t, and not the proper time τ. Indeed, we have:

$$\frac{\partial}{\partial x^i}\left(Gr^{-1}\frac{dx^i}{dt}\right) = \frac{\partial}{\partial x^i}\left(G\frac{dx^i}{d\tau}\right) = 0. \tag{4.26}$$

Thus, if we consider the proper time τ, the invariant measure is simply $g = G$.

In the next part of the present study it will be shown that on the invariant measure defined by a measure preserving operator the standard results of statistical mechanics can be recovered. Physically, this is because the metric associated to a measure preserving operator is 'current free', i.e. $\mathcal{O}^{n-1} = 0$. This means that, regardless of the form of the energy H, the dynamics can never encounter sinking or diverging points. Because of the special properties of the measure preserving operator, we would like to know if a general antisymmetric operator can be transformed to a measure preserving one through some procedure that is not necessarily limited to a spatial coordinate change. On this regard, we have the following extension method:

***Proposition* 4.3.** *(Extension)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator on a smooth manifold $\mathcal{M}$ of dimension n. Let x^{n+1} be a new variable with domain $\mathcal{D} \subseteq \mathbb{R}$. Then, the $n+1$ dimensional antisymmetric operator on $\bigwedge^2 T(\mathcal{M} \otimes \mathcal{D})$:

$$\mathfrak{J} = \mathcal{J} + x^{n+1}\frac{\partial \mathcal{J}^{ij}}{\partial x^i}\partial_j \wedge \partial_{n+1}, \tag{4.27}$$

is measure preserving.

Proof. We want to show that on the volume form $vol^{n+1} = dx^1 \wedge ... \wedge dx^n \wedge dx^{n+1}$, the cocurrent $\mathcal{O}^n = d\mathfrak{J}^{n-1}$ vanishes. Recalling the condition given by equation (4.17), we have:

$$\begin{aligned}
\sum_{i=1}^{n+1}\frac{\partial \mathfrak{J}^{ij}}{\partial x^i} &= \frac{\partial \mathfrak{J}^{n+1,j}}{\partial x^{n+1}} + \sum_{i=1}^{n}\frac{\partial \mathfrak{J}^{ij}}{\partial x^i} \\
&= -\left(1-\delta_{j,n+1}\right)\sum_{i=1}^{n}\frac{\partial}{\partial x^{n+1}}\left(x^{n+1}\frac{\partial \mathcal{J}^{ij}}{\partial x^i}\right) \\
&\quad + \delta_{j,n+1}\sum_{i=1}^{n}\frac{\partial \mathfrak{J}^{ij}}{\partial x^i} + \left(1-\delta_{j,n+1}\right)\sum_{i=1}^{n}\frac{\partial \mathcal{J}^{ij}}{\partial x^i} \\
&= x^{n+1}\sum_{i,k=1}^{n}\frac{\partial^2 \mathcal{J}^{ki}}{\partial x^i \partial x^k} \\
&= 0,
\end{aligned} \tag{4.28}$$

and the statement is proved. □

The meaning of this extension method can be understood as follows: at the price of increasing by one the degrees of freedom of a dynamical system, we can always find a preserved volume element that is independent of the Hamiltonian H.

***Remark* 4.2.** *Observe that if by chance $\mathcal{J}$ is already measure preserving, i.e. $\partial_i\left(g\mathcal{J}^{ij}\right)=0$ for some metric g, there is no need to perform the extension of proposition 4.3 (in fact, the extended term will vanish in the coordinate system $g=1$, see* (4.27)*).*

It turns out that all operators with $n=3$ can be extended not only to a measure preserving but to a conformal form. To see this, first we need the following:

***Proposition* 4.4.** *Let $\mathcal{J}\in\bigwedge^2 T\mathcal{M}$ be an antisymmetric operator on a smooth manifold $\mathcal{M}$ of dimension n. If $n=4$ and $\mathcal{J}$ is measure preserving with a non-zero Jacobiator $\mathfrak{G}\left(\mathcal{J}\right)\neq 0$, then $\mathcal{J}$ is conformal.*

Proof. First, we use the fact that $\mathcal{J}$ is measure preserving, and perform a change of coordinates $dx^1\wedge...\wedge dx^n=g^{-1}dy^1\wedge...\wedge dy^n$, where $dy^1\wedge...\wedge dy^n$ is the invariant measure and g the Jacobian of the coordinate change. Let $\mathcal{I}$ be the expression of $\mathcal{J}$ in the new coordinates. We need to prove that there is some non-zero function $r\in C^\infty\left(\mathcal{M}\right)$ such that $r^{-1}\mathcal{I}$ satisfies the Jacobi identity (1.13). Equation (1.13) becomes:

$$\begin{aligned}\mathfrak{G}\left(r^{-1}\mathcal{I}\right)&=r^{-1}\left(\mathcal{I}^{im}\frac{\partial}{\partial y^m}r^{-1}\mathcal{I}^{jk}+\mathcal{I}^{jm}\frac{\partial}{\partial y^m}r^{-1}\mathcal{I}^{ki}+\mathcal{I}^{km}\frac{\partial}{\partial y^m}r^{-1}\mathcal{I}^{ij}\right)\\&=r^{-1}\frac{\partial}{\partial y^m}\left[r^{-1}\left(\mathcal{I}^{im}\mathcal{I}^{jk}+\mathcal{I}^{jm}\mathcal{I}^{ki}+\mathcal{I}^{km}\mathcal{I}^{ij}\right)\right].\end{aligned}\tag{4.29}$$

In the above equation, we used the fact that $\partial_i\mathcal{I}^{ij}=0\ \forall j=1,...,n$. With the choice:

$$r=\left|\mathcal{I}^{21}\mathcal{I}^{34}+\mathcal{I}^{31}\mathcal{I}^{42}+\mathcal{I}^{41}\mathcal{I}^{23}\right|,\tag{4.30}$$

the right-hand side of equation (4.29) vanishes because the only nontrivial case is when i,j,k,m are all different and any of such combinations gives the quantity $\pm\left(\mathcal{I}^{21}\mathcal{I}^{34}+\mathcal{I}^{31}\mathcal{I}^{42}+\mathcal{I}^{41}\mathcal{I}^{23}\right)$. Note that r can never vanish because by hypothesis the Jacobiator is different from zero, i.e. $\mathfrak{G}\left(\mathcal{J}\right)\neq 0$ (and therefore $\mathfrak{G}\left(\mathcal{I}\right)\neq 0$). $\square$

This result implies that a constant rank measure preserving operator with $n=4$ that does not satisfy the Jacobi identity can always be transformed to a Poisson operator by operating the time reparametrization $d\tau/dt=r$. By combining propositions 4.3 and 4.4 we also have the following result regarding the 'Poissonization' of 3-dimensional antisymmetric operators:

***Proposition* 4.5.** *Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator on a smooth manifold $\mathcal{M}$ of dimension n. If $n = 3$ and $\mathfrak{G}(\mathcal{J}) \neq 0$, $\mathcal{J}$ can always be extended to a conformal operator of dimension $n + 1 = 4$.*

Proof. First, notice that if $\mathfrak{G}(\mathcal{J}) = 0$ there is no need to perform any extension since the system is already Hamiltonian. Assuming that $\mathfrak{G}(\mathcal{J}) \neq 0$ on $\mathcal{M}$, apply proposition 4.3 to obtain a 4-dimensional measure preserving bracket. Then, according to proposition 4.4, this extended bracket is conformal. □

***Remark* 4.3.** *Note that by operating the time reparametrization dictated by the conformal factor, the conformal operator can be converted to a Poisson operator. We also remark that if a 3-dimensional antisymmetric operator is measure preserving there is no need for extension since it will be shown that such operator is a Poisson operator (see below).*

***Proposition* 4.6.** *Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator with constant rank on a smooth manifold $\mathcal{M}$ of dimension $n = 3$. Let $X = \mathcal{J}(dH)$ be the conservative vector field generated by the Hamiltonian H. Then there exists a unique smooth vector field $\boldsymbol{w}$ such that $X = \boldsymbol{w} \times \nabla H$. Furthermore, the following conditions are locally equivalent:*

1. $\forall \boldsymbol{x} \in \mathcal{M} \;\; \exists\, U \subset \mathcal{M}: \quad \mathfrak{G}(\mathcal{J}) = 0,$ (4.31a)
2. $\forall \boldsymbol{x} \in \mathcal{M} \;\; \exists\, U \subset \mathcal{M},\;\; \lambda, C : U \to \mathbb{R}: \quad \boldsymbol{w} = \lambda \nabla C,$ (4.31b)
3. $\forall \boldsymbol{x} \in \mathcal{M} \;\; \exists\, U \subset \mathcal{M},\;\; g \neq 0,\;\; g : U \to \mathbb{R}: \quad \mathfrak{L}_X g dx \wedge dy \wedge dz = 0 \;\; \forall H.$ (4.31c)

Proof. One can verify that with the identification:

$$\begin{aligned} \mathcal{J} &= \mathcal{J}^{zy}\partial_z \wedge \partial_y + \mathcal{J}^{xz}\partial_x \wedge \partial_z + \mathcal{J}^{yx}\partial_y \wedge \partial_x \\ &= w_x\partial_z \wedge \partial_y + w_y\partial_x \wedge \partial_z + w_z\partial_y \wedge \partial_x, \end{aligned} \tag{4.32}$$

we have in a unique manner $X = \mathcal{J}(dH) = \boldsymbol{w} \times \nabla H$. The vector field $\boldsymbol{w}$ is also smooth because, by definition, the components of $\mathcal{J}$ are smooth.

(1 $\implies$ 2) is the Frobenius theorem 1.2. Indeed, using equation (4.1):

$$\begin{aligned}
&\mathfrak{G}\left(\mathcal{J}\right)\\
&=\left(\mathcal{J}^{xm}\frac{\partial\mathcal{J}^{yz}}{\partial x^m}+\mathcal{J}^{ym}\frac{\partial\mathcal{J}^{zx}}{\partial x^m}+\mathcal{J}^{zm}\frac{\partial\mathcal{J}^{xy}}{\partial x^m}\right)\partial_x\wedge\partial_y\wedge\partial_z\\
&=\left(\mathcal{J}^{xy}\frac{\partial\mathcal{J}^{yz}}{\partial y}+\mathcal{J}^{xz}\frac{\partial\mathcal{J}^{yz}}{\partial z}+\mathcal{J}^{yx}\frac{\partial\mathcal{J}^{zx}}{\partial x}+\mathcal{J}^{yz}\frac{\partial\mathcal{J}^{zx}}{\partial z}+\mathcal{J}^{zx}\frac{\partial\mathcal{J}^{xy}}{\partial x}+\mathcal{J}^{zy}\frac{\partial\mathcal{J}^{xy}}{\partial y}\right)\partial_x\\
&\quad\wedge\partial_y\wedge\partial_z\\
&=\\
&=\left(w_z\frac{\partial w_x}{\partial y}-w_y\frac{\partial w_x}{\partial z}-w_z\frac{\partial w_y}{\partial x}+w_x\frac{\partial w_y}{\partial z}+w_y\frac{\partial w_z}{\partial x}-w_x\frac{\partial w_z}{\partial y}\right)\partial_x\wedge\partial_y\wedge\partial_z\\
&=-\left(\boldsymbol{w}\cdot\nabla\times\boldsymbol{w}\right)\partial_x\wedge\partial_y\wedge\partial_z.
\end{aligned}\tag{4.33}$$

Therefore, the Jacobiator $\mathfrak{G}\left(\mathcal{J}\right)=\mathfrak{G}\left(\boldsymbol{w}\right)$ vanishes if and only if the vector field $\boldsymbol{w}$ is integrable in the sense of Frobenius 1.2, i.e. $\boldsymbol{w}\cdot\nabla\times\boldsymbol{w}=0$. Indeed, if we define the 1-form $\theta=w_xdx+w_ydy+w_zdz$ the condition $\boldsymbol{w}\cdot\nabla\times\boldsymbol{w}=0$ is exactly the integrability condition $\theta\wedge d\theta=0$ of equation (1.40). But then locally we can find two functions λ and C such that $\boldsymbol{w}=\lambda\nabla C$.

(2 $\implies$ 1) is trivial since:

$$\begin{aligned}
\mathfrak{G}\left(\boldsymbol{w}\right)&=-\left(\boldsymbol{w}\cdot\nabla\times\boldsymbol{w}\right)\partial_x\wedge\partial_y\wedge\partial_z\\
&=-\left(\lambda\nabla C\cdot\nabla\lambda\times\nabla C\right)\partial_x\wedge\partial_y\wedge\partial_z\\
&=0.
\end{aligned}\tag{4.34}$$

(2 $\implies$ 3) can be verified by observing that:

$$\begin{aligned}
\mathfrak{L}_X g dx\wedge dy\wedge dz=0\ \ \forall H\quad&\iff\quad\frac{\partial}{\partial x^i}\left(gX^i\right)=0\quad\forall H\\
&\iff\quad\nabla H\cdot\nabla\times\left(g\boldsymbol{w}\right)=0\quad\forall H.
\end{aligned}\tag{4.35}$$

If $\lambda=0$ or C=constant on U, $\boldsymbol{w}=\boldsymbol{0}$. Then, any function $g\neq 0$ will prove the statement. Otherwise, the implication follows by setting $g=\lambda^{-1}$.

(3 $\implies$ 2) If there is an invariant measure g for any choice of H, then $\nabla\times\left(g\boldsymbol{w}\right)=\boldsymbol{0}$. Therefore $\boldsymbol{w}=g^{-1}\nabla C$ on U. □

From the calculation of equation (8.11) we have learned that in 3-dimensions the Jacobiator $\mathfrak{G}\left(\boldsymbol{w}\right)$ vanishes if and only if the quantity $\boldsymbol{w}\cdot\nabla\times\boldsymbol{w}$ is zero. This fact will be used extensively in our study of 3-dimensional dynamics and diffusion. From a strictly geometrical point of view, $\boldsymbol{w}\cdot\nabla\times\boldsymbol{w}=0$ means that the curl $\nabla\times\boldsymbol{w}$ of the operator $\boldsymbol{w}$ is perpendicular to $\boldsymbol{w}$ itself. It is instructive to show that $\nabla\times\boldsymbol{w}$

corresponds to the cocurrent $\mathcal{O}^2$ on the metric $dx \wedge dy \wedge dz$ we have introduced in definition 4.14. Using the fact that $\mathcal{J}^{yx} = w_z$, $\mathcal{J}^{zy} = w_x$, and $\mathcal{J}^{xz} = w_y$:

$$\begin{aligned}\mathcal{O}^2 &= d\left(i_{\mathcal{J}} dx \wedge dy \wedge dz\right)\\ &= d\left(i_{\mathcal{J}^{zy}\partial_z\wedge\partial_y + \mathcal{J}^{xz}\partial_x\wedge\partial_z + \mathcal{J}^{yx}\partial_y\wedge\partial_x} dx \wedge dy \wedge dz\right)\\ &= 2d\left(-\mathcal{J}^{zy}dx - \mathcal{J}^{xz}dy - \mathcal{J}^{yx}dz\right)\\ &= 2d\left(-w_x dx - w_y dy - w_z dz\right)\\ &= 2\left[\left(\frac{\partial w_y}{\partial x} - \frac{\partial w_x}{\partial y}\right) dy \wedge dx + \left(\frac{\partial w_z}{\partial x} - \frac{\partial w_x}{\partial z}\right) dz \wedge dx + \left(\frac{\partial w_z}{\partial y} - \frac{\partial w_y}{\partial z}\right) dz \wedge dy\right]\\ &= -2\left[\left(\nabla \times \boldsymbol{w}\right)_x dy \wedge dz + \left(\nabla \times \boldsymbol{w}\right)_y dz \wedge dx + \left(\nabla \times \boldsymbol{w}\right)_z dx \wedge dy\right].\end{aligned} \tag{4.36}$$

Note that we have also shown that for $n = 3$ we have $\mathcal{J}^{n-2} = \mathcal{J}^1 = -2\left(w^i dx^i\right)$.

4.5 The Beltrami Operator

In the same way the condition $\boldsymbol{w} \cdot \nabla \times \boldsymbol{w} = 0$ defines Hamiltonian mechanics, we may wonder whether the vanishing of the component of $\nabla \times \boldsymbol{w}$ normal to $\boldsymbol{w}$ has relevant implications for the dynamics. The answer is positive, and we will examine the dynamical and statistical properties of such operators later on. The vanishing of the normal component is expressed by the Beltrami condition:

$$\boldsymbol{b} = \boldsymbol{w} \times \left(\nabla \times \boldsymbol{w}\right) = \boldsymbol{0}. \tag{4.37}$$

In general, a vector $\boldsymbol{w}$ that satisfies equation (4.37) is called a Beltrami field. A Beltrami field is characterized by the fact that it is completely aligned with its own curl $\nabla \times \boldsymbol{w}$ and therefore obeys the Beltrami equation:

$$\nabla \times \boldsymbol{w} = \mu \boldsymbol{w}. \tag{4.38}$$

Here μ is, in general, a scalar function called the Beltrami parameter.

As one may expect, the quantity $\boldsymbol{b}$ appearing in equation (4.37) does not vanish for a general $\boldsymbol{w}$. When $\boldsymbol{w}$ is the antisymmetric operator associated to the $\boldsymbol{E} \times \boldsymbol{B}$ drift motion (see chapter 8) of a charged particle in a magnetic field $\boldsymbol{B}$ of constant strength B =constant, the vector $\boldsymbol{b}$ is nothing but the magnetic force $\boldsymbol{B} \times \left(\nabla \times \boldsymbol{B}\right)$. Therefore, we shall name the quantity $\boldsymbol{b}$ the *field force vector* of $\boldsymbol{w}$. The more $\boldsymbol{b}$ approaches zero, the more $\boldsymbol{w}$ will resemble a Beltrami field. To understand the geometrical meaning of $\boldsymbol{b}$ the following vector identity is useful:

$$\boldsymbol{b} = \boldsymbol{w} \times \left(\nabla \times \boldsymbol{w}\right) = \frac{1}{2}\nabla\left(w^2\right) - \left(\boldsymbol{w} \cdot \nabla\right)\boldsymbol{w}. \tag{4.39}$$

Normalize $\boldsymbol{w}$ to $\hat{\boldsymbol{w}} = \boldsymbol{w}/w$. Then, using the previous formula for $\hat{\boldsymbol{w}}$:

$$\hat{\boldsymbol{b}} = -\left(\hat{\boldsymbol{w}} \cdot \nabla\right) \hat{\boldsymbol{w}} = -\hat{\boldsymbol{k}}, \tag{4.40}$$

where $\hat{\boldsymbol{k}}$ is the curvature vector measuring the rate of change of the unit vector $\hat{\boldsymbol{w}}$ along itself. Therefore, the vector $\boldsymbol{b}$ is geometrically related to the curvature of $\boldsymbol{w}$.

Nevertheless, since $\boldsymbol{b}$ is a vector, we would like to obtain a better (scalar) measure of the degree at which a certain $\boldsymbol{w}$ resembles a Beltrami field. For reasons that we will explain in the next part of the present study, one finds that such proper measure is:

$$\mathfrak{B} = 4\nabla \cdot \boldsymbol{b} = 4\nabla \cdot \left[\boldsymbol{w} \times \left(\nabla \times \boldsymbol{w}\right)\right]. \tag{4.41}$$

We call $\mathfrak{B}$ the *field force divergence* of $\boldsymbol{w}$.

The following result clarifies the relationship between the field force vector $\boldsymbol{b}$ and the cocurrent $\nabla \times \boldsymbol{w}$.

***Proposition* 4.7.** *(Beltrami-Jacobi decomposition of the curl operator)*
The curl of a vector field $\boldsymbol{w}$ admits the following decomposition:

$$\nabla \times \boldsymbol{w} = \frac{\boldsymbol{b} \times \boldsymbol{w} + \mathfrak{G}\boldsymbol{w}}{w^2}, \tag{4.42}$$

where $\boldsymbol{b} = \boldsymbol{w} \times \left(\nabla \times \boldsymbol{w}\right)$ is the field force vector of $\boldsymbol{w}$, and $\mathfrak{G} = \boldsymbol{w} \cdot \nabla \times \boldsymbol{w}$ its Jacobiator.

The task we are left with is the generalization of the concept of field force to arbitrary dimensions n. By consistency with equation (4.41), the field force divergence of a general antisymmetric operator $\mathcal{J}$ must be a 0-form. Furthermore, since $\mathfrak{B}$ is the divergence of the vector $\boldsymbol{b}$, the generalization of $\boldsymbol{b}$ must be an $n-1$ form. This is done in the following manner:

***Def* 4.8.** *(Field force)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator on a smooth manifold $\mathcal{M}$ of dimension n. The field force $n-1$ form of $\mathcal{J}$ is defined as:

$$b^{n-1} = \mathcal{J}^{n-2} \wedge *d\mathcal{J}^{n-2}. \tag{4.43}$$

The field force divergence of $\mathcal{J}$ is defined as:

$$\mathfrak{B} = *db^{n-1} = *d\left(\mathcal{J}^{n-2} \wedge *d\mathcal{J}^{n-2}\right). \tag{4.44}$$

Let us check that, for $n = 3$, these definitions reproduce $\boldsymbol{b}$ and $\mathfrak{B}$ of equation (4.41) on $vol^3 = dx \wedge dy \wedge dz$. First consider b^{n-1} for a general n:

$$\begin{aligned} b^{n-1} &= \mathcal{J}^{n-2} \wedge *d\mathcal{J}^{n-2} \\ &= 2 \sum_{i<j} (-1)^{i+j-1} g\mathcal{J}^{ij} dx_{ij}^{n-2} \wedge 2 (-1)^k \frac{\partial \left(g\mathcal{J}^{lk}\right)}{\partial x^l} * dx_k^{n-1} \\ &= 4 \sum_{i<j} (-1)^{i+j+k-1} g\mathcal{J}^{ij} \frac{\partial \left(g\mathcal{J}^{lk}\right)}{\partial x^l} dx_{ij}^{n-2} \wedge *dx_k^{n-1}. \end{aligned} \tag{4.45}$$

Here, we used equations (4.11) and (4.20). Now observe that in the case $n = 3$ of $\mathbb{R}^3$, $g = 1$ and that (x, y, z) is an orthonormal coordinate system:

$$\begin{aligned} b^{n-1} &= 4 \sum_{i<j} (-1)^{n+i+j-1} \mathcal{J}^{ij} \frac{\partial \mathcal{J}^{lk}}{\partial x^l} dx_{ij}^{n-2} \wedge dx^k \\ &= 4 \sum_{i<j} (-1)^{n+i+j-1} \left(\mathcal{J}^{ij} \frac{\partial \mathcal{J}^{li}}{\partial x^l} dx_{ij}^{n-2} \wedge dx^i + \mathcal{J}^{ij} \frac{\partial \mathcal{J}^{lj}}{\partial x^l} dx_{ij}^{n-2} \wedge dx^j \right) \\ &= 4 (-1)^j \mathcal{J}^{ij} \frac{\partial \mathcal{J}^{li}}{\partial x^l} dx_j^{n-1} \\ &= 4 \left[-\mathcal{J}^{yx} \left(\frac{\partial \mathcal{J}^{xy}}{\partial x} + \frac{\partial \mathcal{J}^{zy}}{\partial z} \right) - \mathcal{J}^{zx} \left(\frac{\partial \mathcal{J}^{xz}}{\partial x} + \frac{\partial \mathcal{J}^{yz}}{\partial y} \right) \right] dy \wedge dz \\ &\quad + 4 \left[\mathcal{J}^{xy} \left(\frac{\partial \mathcal{J}^{yx}}{\partial y} + \frac{\partial \mathcal{J}^{zx}}{\partial z} \right) + \mathcal{J}^{zy} \left(\frac{\partial \mathcal{J}^{xz}}{\partial x} + \frac{\partial \mathcal{J}^{yz}}{\partial y} \right) \right] dx \wedge dz \\ &\quad + 4 \left[-\mathcal{J}^{xz} \left(\frac{\partial \mathcal{J}^{yx}}{\partial y} + \frac{\partial \mathcal{J}^{zx}}{\partial z} \right) - \mathcal{J}^{yz} \left(\frac{\partial \mathcal{J}^{xy}}{\partial x} + \frac{\partial \mathcal{J}^{zy}}{\partial z} \right) \right] dx \wedge dy \\ &= 4 \left[-w_z \left(-\frac{\partial w_z}{\partial x} + \frac{\partial w_x}{\partial z} \right) + w_y \left(\frac{\partial w_y}{\partial x} - \frac{\partial w_x}{\partial y} \right) \right] dy \wedge dz \\ &\quad + 4 \left[-w_z \left(\frac{\partial w_z}{\partial y} - \frac{\partial w_y}{\partial z} \right) + w_x \left(\frac{\partial w_y}{\partial x} - \frac{\partial w_x}{\partial y} \right) \right] dx \wedge dz \\ &\quad + 4 \left[-w_y \left(\frac{\partial w_z}{\partial y} - \frac{\partial w_y}{\partial z} \right) + w_x \left(-\frac{\partial w_z}{\partial x} + \frac{\partial w_x}{\partial z} \right) \right] dx \wedge dy \\ &= 4 \Big[(\boldsymbol{w} \times (\nabla \times \boldsymbol{w}))_x \, dy \wedge dz + (\boldsymbol{w} \times (\nabla \times \boldsymbol{w}))_y \, dz \wedge dx + (\boldsymbol{w} \times (\nabla \times \boldsymbol{w}))_z \, dx \\ &\qquad \wedge dy \Big]. \end{aligned} \tag{4.46}$$

Now we return to equation (4.45) and evaluate $\mathfrak{B}$:

$$\begin{aligned}
\mathfrak{B} &= *db^{n-1} \\
&= 4\sum_{i<j}(-1)^{i+j+k-1}\frac{\partial}{\partial x^m}\left(g\mathcal{J}^{ij}\frac{\partial\left(g\mathcal{J}^{lk}\right)}{\partial x^l}\right)*\left(dx^m\wedge dx_{ij}^{n-2}\wedge *dx_k^{n-1}\right) \\
&= 4(-1)^{j+k}\frac{\partial}{\partial x^i}\left(g\mathcal{J}^{ij}\frac{\partial\left(g\mathcal{J}^{lk}\right)}{\partial x^l}\right)*\left(dx_j^{n-1}\wedge *dx_k^{n-1}\right) \\
&= 4(-1)^{j+k}\frac{\partial}{\partial x^i}\left(g\mathcal{J}^{ij}\frac{\partial\left(g\mathcal{J}^{lk}\right)}{\partial x^l}\right)*\left\langle dx_j^{n-1}, dx_k^{n-1}\right\rangle vol^n \\
&= 4\frac{\partial}{\partial x^i}\left(g\mathcal{J}^{ij}\frac{\partial\left(g\mathcal{J}^{lj}\right)}{\partial x^l}\right)*vol^n \\
&= 4\frac{\partial}{\partial x^i}\left(g\mathcal{J}^{ij}\frac{\partial\left(g\mathcal{J}^{lj}\right)}{\partial x^l}\right).
\end{aligned} \tag{4.47}$$

Recalling the calculation of (4.46), for the case $n=3$ with $g=1$ of $\mathbb{R}^3$:

$$\mathfrak{B} = 4\left\{\nabla\cdot\left[\boldsymbol{w}\times\left(\nabla\times\boldsymbol{w}\right)\right]\right\}dx\wedge dy\wedge dz, \tag{4.48}$$

which is the desired result.

A result analogous to the decomposition 4.7 can be obtained for a general n. For this purpose we need to define the norm of an antisymmetric operator. The standard Frobenius norm will serve our purposes:

***Def* 4.9.** *(Frobenius norm)*
Let $\mathcal{J}\in\bigwedge^2 T\mathcal{M}$ be an antisymmetric operator on a smooth manifold $\mathcal{M}$ of dimension n. Let $vol^n = gdx^1\wedge...\wedge dx^n$ be a volume form on $\mathcal{M}$ and $\mathcal{J}^{n-2}$ the associated covorticity $n-2$ form. The Frobenius norm of $\mathcal{J}$ on vol^n is:

$$|\mathcal{J}|^2 = \left|\mathcal{J}^{n-2}\right|^2 = \frac{1}{2}\sum_{ij}\left[\left(\mathcal{J}^{n-2}\right)^{ij}\right]^2 = \frac{1}{2}\sum_{ij}g^2\left(\mathcal{J}^{ij}\right)^2. \tag{4.49}$$

Let us check the value of this norm when $n=3$ and $g=1$ on $\mathbb{R}^3$:

$$|\mathcal{J}|^2 = \frac{1}{2}\sum_{i,j=1}^{3}\left(\mathcal{J}^{ij}\right)^2 = w_x^2+w_y^2+w_z^2 = |\boldsymbol{w}|^2 = w^2. \tag{4.50}$$

We can now obtain the following decomposition:

***Proposition* 4.8.** *(Beltrami-Jacobi decomposition of the cocurrent $n-1$ form)*
Let $\mathcal{J}\in\bigwedge^2 T\mathcal{M}$ be an antisymmetric operator on a smooth manifold $\mathcal{M}$ of dimension n. Let $\mathcal{O}^{n-1}$ be the cocurrent $n-1$ form of $\mathcal{J}$ on a volume form $vol^n =$

$g dx^1 \wedge ... \wedge dx^n$. *Then:*

$$\mathcal{O}^{n-1} = d\mathcal{J}^{n-2} = \frac{(-1)^k g\mathcal{J}^{ij}\left(\mathfrak{G}^{kij} + R_m^{ijmk}\right) dx_k^{n-1} - (-1)^n \frac{1}{4}\left(*b^{n-1}\right) \wedge \mathcal{J}^{n-2}}{|\mathcal{J}|^2}. \tag{4.51}$$

Here $\mathfrak{G}^{kij} = g\mathcal{J}^{km}\left(g\mathcal{J}^{ij}\right)_m + g\mathcal{J}^{im}\left(g\mathcal{J}^{jk}\right)_m + g\mathcal{J}^{jm}\left(\mathcal{J}^{ki}\right)_m$ *is the* kij *component of the Jacobiator* $\mathfrak{G}(g\mathcal{J})$ *and* $R_m^{ijmk} = \left[g^2\left(\mathcal{J}^{ij}\mathcal{J}^{mk} + \mathcal{J}^{im}\mathcal{J}^{kj} + \mathcal{J}^{jm}\mathcal{J}^{ik}\right)\right]_m$, *with the subscript meaning derivation.*

Proof. We have:

$$\begin{aligned}
2|\mathcal{J}|^2 \frac{\partial\left(g\mathcal{J}^{mk}\right)}{\partial x^m} &= g^2\mathcal{J}^{ij}\mathcal{J}^{ij}\frac{\partial\left(g\mathcal{J}^{mk}\right)}{\partial x^m} \\
&= g\mathcal{J}^{ij}\frac{\partial}{\partial x^m}\left(g^2\mathcal{J}^{ij}\mathcal{J}^{mk}\right) - g^2\mathcal{J}^{ij}\mathcal{J}^{mk}\frac{\partial\left(g\mathcal{J}^{ij}\right)}{\partial x^m} \\
&= g\mathcal{J}^{ij}\frac{\partial}{\partial x^m}\left(g^2\mathcal{J}^{ij}\mathcal{J}^{mk}\right) \\
&\quad + g\mathcal{J}^{ij}\left(g\mathcal{J}^{km}\frac{\partial\left(g\mathcal{J}^{ij}\right)}{\partial x^m} + g\mathcal{J}^{im}\frac{\partial\left(g\mathcal{J}^{jk}\right)}{\partial x^m} + g\mathcal{J}^{jm}\frac{\partial\left(g\mathcal{J}^{ki}\right)}{\partial x^m}\right) \\
&\quad - g^2\mathcal{J}^{ij}\mathcal{J}^{im}\frac{\partial\left(g\mathcal{J}^{jk}\right)}{\partial x^m} - g^2\mathcal{J}^{ij}\mathcal{J}^{jm}\frac{\partial\left(g\mathcal{J}^{ki}\right)}{\partial x^m} \\
&= g\mathcal{J}^{ij}\mathfrak{G}^{kij} + g\mathcal{J}^{ij}R_m^{ijmk} + g^2\mathcal{J}^{ij}\mathcal{J}^{jk}\frac{\partial\left(g\mathcal{J}^{im}\right)}{\partial x^m} + g^2\mathcal{J}^{ij}\mathcal{J}^{ki}\frac{\partial\left(g\mathcal{J}^{jm}\right)}{\partial x^m} \\
&= g\mathcal{J}^{ij}\left(\mathfrak{G}^{kij} + R_m^{ijmk}\right) + 2g^2\mathcal{J}^{ij}\mathcal{J}^{jk}\frac{\partial\left(g\mathcal{J}^{im}\right)}{\partial x^m}.
\end{aligned} \tag{4.52}$$

On the other hand:

$$\begin{aligned}
\frac{1}{4}\left(*b^{n-1}\right) \wedge \mathcal{J}^{n-2} = &\left[\sum_{i<j}(-1)^{i+j+k-1} g\mathcal{J}^{ij}\frac{\partial\left(g\mathcal{J}^{mk}\right)}{\partial x^m} * \left(dx_{ij}^{n-2} \wedge *dx_k^{n-1}\right)\right] \\
&\wedge 2\sum_{r<s}(-1)^{r+s-1} g\mathcal{J}^{rs}dx_{rs}^{n-2}.
\end{aligned} \tag{4.53}$$

Without loss of generality, we may assume that we have chosen the x^i to be orthonormal coordinates in $\mathbb{R}^n$. In such case we can conveniently evaluate the Hodge star and obtain:

$$\begin{aligned}
\frac{1}{4}\left(*b^{n-1}\right) \wedge \mathcal{J}^{n-2} &= 2\sum_{r<s}(-1)^{n+r+s-1} g^2\mathcal{J}^{ij}\mathcal{J}^{rs}\frac{\partial\left(g\mathcal{J}^{mi}\right)}{\partial x^m}dx^j \wedge dx_{rs}^{n-2} \\
&= -2(-1)^{n+s} g^2\mathcal{J}^{ir}\mathcal{J}^{rs}\frac{\partial\left(g\mathcal{J}^{im}\right)}{\partial x^m}dx_s^{n-1}.
\end{aligned} \tag{4.54}$$

Recalling the definition (4.17) and substituting (4.54) in (4.52), one obtains equation (4.51). □

To verify this result, we consider the case $n = 3$ with $g = 1$ of $\mathbb{R}^3$. First, observe that $R^{ijmk} = 0$ since at least 2 of the indices $ijmk$ are equal when $n = 3$. Secondly:

$$(-1)^k \mathcal{J}^{ij} \mathfrak{G}^{kij} dx_k^{n-1} = -2 (\boldsymbol{w} \cdot \nabla \times \boldsymbol{w}) (w_x dy \wedge dz + w_y dz \wedge dx + w_z dx \wedge dy) . \tag{4.55}$$

Finally, using equation (4.46) and recalling that $\boldsymbol{b} = \boldsymbol{w} \times (\nabla \times \boldsymbol{w})$:

$$\begin{aligned} -(-1)^n \frac{1}{4} \left(*b^{n-1}\right) \wedge \mathcal{J}^{n-2} &= * (b_x dy \wedge dz + b_y dz \wedge dx + b_z dx \wedge dy) \wedge \left(-w_i dx^i\right) \\ &= -b_i dx^i \wedge w_j dx^j \\ &= -b_i w_j dx^i \wedge dx^j \\ &= -2 (\boldsymbol{b} \times \boldsymbol{w})_x dy \wedge dz - 2 (\boldsymbol{b} \times \boldsymbol{w})_y dz \\ &\quad \wedge dx - 2 (\boldsymbol{b} \times \boldsymbol{w})_z dx \wedge dy. \end{aligned} \tag{4.56}$$

Since $\mathcal{O}^2 = -2 \left[(\nabla \times w)_x dy \wedge dz + (\nabla \times w)_y dz \wedge dx + (\nabla \times w)_z dx \wedge dy \right]$ we recover the decomposition of proposition 4.7.

Observe that from equation (4.51), when $n > 3$, we see that the cocurrent $\mathcal{O}^{n-1}$ on a certain metric g does not vanish even if both $\mathfrak{G}$ and b^{n-1} are zero. Furthermore, since for a measure preserving operator $\mathcal{O}^{n-1} = d\mathcal{J}^{n-2} = 0$, it follows that in this case $b^{n-1} = 0$ and therefore $\mathfrak{G}^{kij} = -R_m^{ijmk}$.

This is the right juncture to define the Beltrami operator:

***Def* 4.10.** *(Beltrami operator)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator on a smooth manifold $\mathcal{M}$ of dimension n. If a volume form $vol^n = g dx^1 \wedge ... \wedge dx^n$ can be found such that the field force divergence is zero, i.e. $\mathfrak{B} = *db^{n-1} = 0$, $\mathcal{J}$ is called a Beltrami operator on vol^n. If the field force $n-1$ form is zero, i.e. $b^{n-1} = 0$, $\mathcal{J}$ is called a strong Beltrami operator on vol^n.

The following result holds:

***Example* 4.3.** *Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be a measure preserving operator on a smooth manifold $\mathcal{M}$ of dimension n with invariant measure vol^n. Then, $\mathcal{J}$ is a strong Beltrami operator on the invariant measure, i.e. $b^{n-1} = 0$ on vol^n.*

Proof. A measure preserving operator satisfies $d\mathcal{J}^{n-2} = 0$ on the metric of the invariant measure (recall proposition 4.2). Therefore, the corresponding field force $n-1$ form $b^{n-1} = \mathcal{J}^{n-2} \wedge *d\mathcal{J}^{n-2}$ identically vanishes. □

4.6 The Hierarchy of antisymmetric operators

Figure 4.1 summarizes the geometrical categorization of antisymmetric operators developed in the present chapter. The available transformations among different categories are also shown. Figure 4.2 shows a similar summary for the special and instructive case $n = 3$. These figures should be compared with the expressions of the cocurrent $\mathcal{O}^{n-1}$ given in propositions 4.8 and 4.7: each category is characterized by a different form of $\mathcal{O}^{n-1}$.

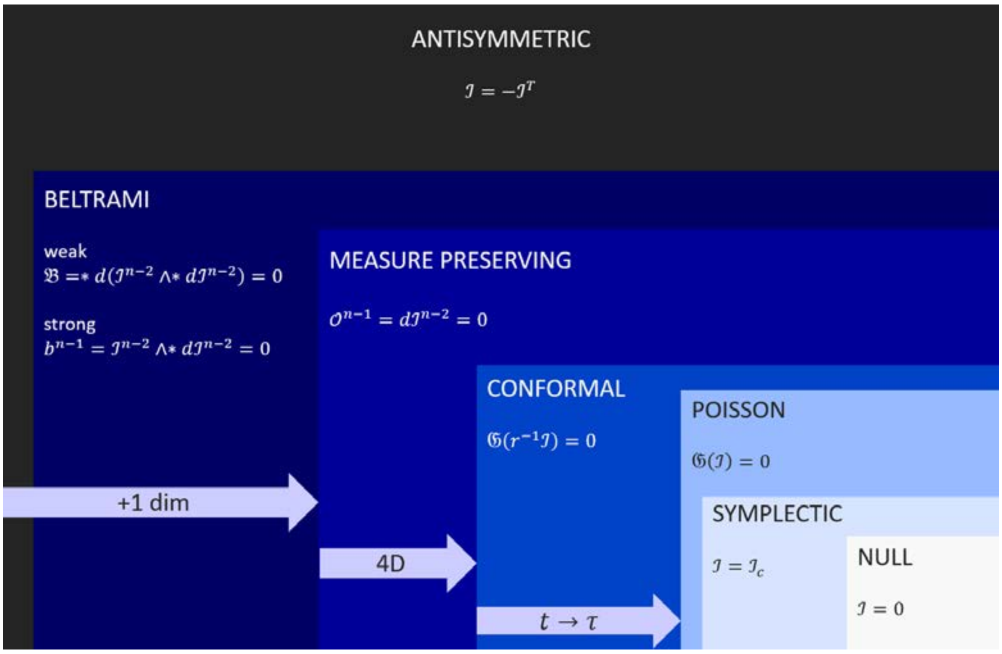

Figure 4.1: The hierarchical structure of antisymmetric operators. Each box is named by the corresponding operator. The arrows specify the available transformations. The operator $\mathcal{J}_c$ in the symplectic domain represents the canonical symplectic matrix.

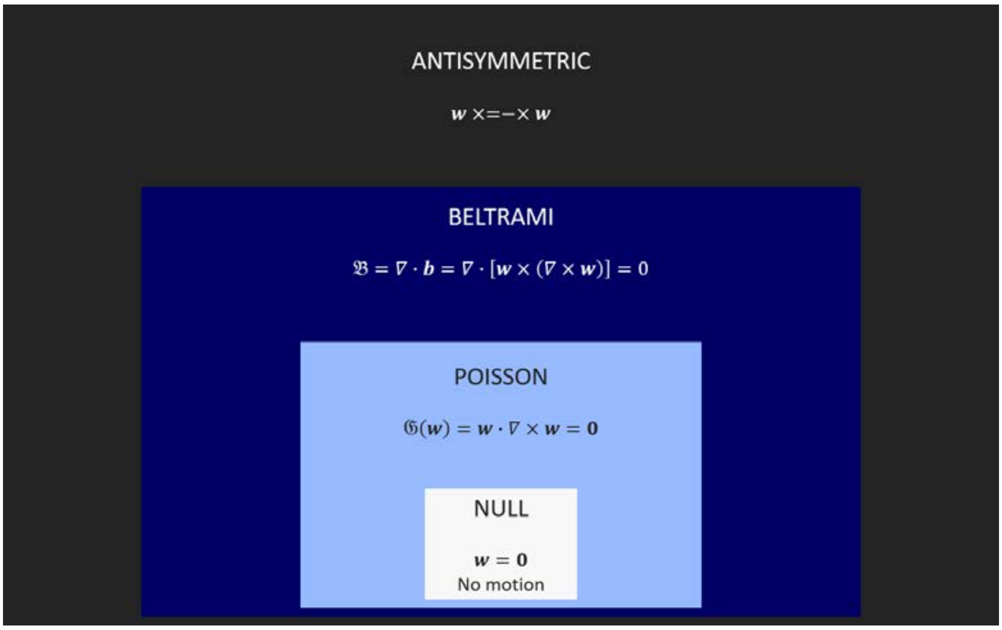

Figure 4.2: The hierarchical structure of antisymmetric operators for $n = 3$. Notice that measure preserving and conformal operators do not appear because they degenerate to Poisson operators when $n = 3$. Specifically, both the conformality condition $\mathfrak{G}\left(r^{-1}\boldsymbol{w}\right) = 0$ and the measure preserving condition $\nabla \times (g\boldsymbol{w}) = \boldsymbol{0}$ reduce to the integrability condition for $\boldsymbol{w}$. See proposition 4.6.

Chapter 5

Topological Constraints and Integrability

In this short chapter we discuss the relationship between the degeneracy of antisymmetric operators and the topological constraints affecting the phase space.

5.1 Topological Constraints on the Phase Space

In the introduction we have explained that the effect of constraints on the phase space of microscopic dynamics is the emergence of a macroscopic system. This transition can be summarized as the substitution of the canonical description of the dynamics in terms of the symplectic operator $\mathcal{J}_c$, with an antisymmetric operator $\mathcal{J}$. The constraints appear as a non trivial kernel $ker\left(\mathcal{J}\right)$ in the matrix $\mathcal{J}$. Such kernel can be truncated. The truncation results in the reduction of the number of degrees of freedom (the dimension of the original microscopic system).

We have also seen that, given an antisymmetric operator $\mathcal{J}$ on a smooth manifold $\mathcal{M}$ and a 1-form $\theta \in ker\left(\mathcal{J}\right)$, the conservative vector field $X = \mathcal{J}\left(dH\right)$ satisfies $\theta\left(X\right) = 0$ for any choice of the Hamiltonian function. Thus, we can introduce a formal definition of topological constraint in the following manner:

***Def* 5.1.** *(Topological constraint)*
A topological constraint is a 1-form $\theta \in ker\left(\mathcal{J}\right)$. The topological constraint is simply denoted as $\theta = 0$.

In the discussion of the Darboux theorem 1.1 we used, without proof, the following result relating the closeness of the symplectic 2-form ω of rank $2n$ to the integrability of $ker\left(\omega\right)$, i.e. the existence of $2n$ coordinates y^i such that $\omega = \sum_{i<j}^{2n} \alpha_{ij} dy^i \wedge dy^j$ and $ker\left(\omega\right) = \left\{X \in T\mathcal{M} : dy^i\left(X\right) = 0 \;\; \forall \;\; i = 1, ..., 2n\right\}$.

***Proposition* 5.1.** *Let $(\mathcal{M}, \omega)$ be a symplectic manifold of dimension $2n+r$. Suppose that ω has rank $2n$. Then, $\forall \boldsymbol{x} \in \mathcal{M}$ one can find a coordinate neighborhood $U \in \mathcal{M}$ with coordinates $\left(y^1, ..., y^{2n}, C^1, ..., C^r\right)$ such that:*

$$\omega = \sum_{i<j}^{2n} \alpha_{ij} \left(y^1, ..., y^{2n}, C^1, ..., C^r\right) dy^i \wedge dy^j \quad on \quad U. \tag{5.1}$$

Proof. Let $\left(x^1, ..., x^{2n+r}\right)$ be a Cartesian coordinate system and let $\boldsymbol{\xi}_i \in T\mathcal{M}$ be r orthonormal vectors that span $ker\,(\omega)$. Consider the associated 1-forms $\xi_i = \xi_i^j dx^j$ and construct a cotangent basis $(\theta_1, ..., \theta_{2n}, \xi_1, ..., \xi_r)$ of $T^*\mathcal{M}$ by adding $2n$ orthonormal vectors $\boldsymbol{\theta}_i$ with corresponding 1-forms $\theta_i = \theta_i^j dx^j$. Note that $\xi_i\left(\boldsymbol{\theta}_j\right) = \theta_j\left(\boldsymbol{\xi}_i\right) = 0\ \forall i, = 1, ..., r$ and $j = 1, ..., 2n$. In the new basis, ω has the expression:

$$\omega = \sum_{i<j}^{2n} \alpha_{ij}\theta_i \wedge \theta_j + \sum_{i,j}^{2n,r} \beta_{ij}\theta_i \wedge \xi_j + \sum_{i<j}^{r} \gamma_{ij}\xi_i \wedge \xi_j, \tag{5.2}$$

for some appropriate coefficients $\alpha_{ij}, \beta_{ij}, \gamma_{ij}$. Let $\boldsymbol{\theta} = \sum_{i=1}^{2n} \zeta^i \boldsymbol{\theta}_i$ be a vector orthogonal to $ker\,(\omega)$, i.e. such that $\xi_i\left(\boldsymbol{\theta}\right) = 0\ \forall i = 1, ..., r$. By definition, we have:

$$i^2_{\boldsymbol{\xi}_m, \boldsymbol{\theta}}\omega = \omega\left(\boldsymbol{\xi}_m, \boldsymbol{\theta}\right) = \sum_{i}^{2n} \beta_{im}\zeta^i = 0 \quad \forall m = 1, ..., r. \tag{5.3}$$

Since $\boldsymbol{\theta}$ is arbitrary, we must have $\beta_{im} = 0\ \forall i = 1, .., 2n$ and $m = 1, ..., r$. Similarly:

$$i^2_{\boldsymbol{\xi}_m, \boldsymbol{\xi}_n}\omega = \omega\left(\boldsymbol{\xi}_m, \boldsymbol{\xi}_n\right) = \gamma_{mn} = 0 \quad \forall m, n = 1, ..., r. \tag{5.4}$$

Therefore, we have shown that

$$\omega = \sum_{i<j}^{2n} \alpha_{ij}\theta_i \wedge \theta_j. \tag{5.5}$$

Since ω is symplectic, $d\omega = 0$. Then:

$$d\left(\alpha_{ij}\theta_i\right) \wedge \theta_j - \alpha_{ij}\theta_i \wedge d\theta_j = 0. \tag{5.6}$$

Multiply the above expression by the $2n-1$ form $\theta_k^{2n-1} = \theta_1 \wedge ... \wedge \theta_{k-1} \wedge \theta_{k+1} \wedge ... \wedge \theta_{2n}$:

$$(-1)^{2n-k}\, \theta_1 \wedge ... \wedge \theta_{2n} \wedge d\left(\alpha_{ik}\theta_i\right) - (-1)^{2n-k}\, \alpha_{kj}\theta_1 \wedge ... \wedge \theta_{2n} \wedge d\theta_j = 0. \tag{5.7}$$

Now observe that the matrix α_{ij} is invertible since ω has rank $2n$. If we multiply the last equation by $\left(\alpha^{-1}\right)^{kl}$ and sum over k, we arrive at:

$$2\theta_1 \wedge ... \wedge \theta_{2n} \wedge d\theta_l = 0 \qquad \forall l = 1, ..., 2n. \tag{5.8}$$

This is the Frobenius integrability condition for the distribution :

$$ker\,(\omega) = \Delta_\theta = \{X \in T\mathcal{M} : \theta_i\,(X) = 0\ \forall i = 1, ..., 2n\}\,. \tag{5.9}$$

Therefore we can find $2n + r$ coordinates $\left(y^1, ..., y^{2n}, C^1, ..., C^r\right)$ on an appropriate neighborhood $U \subset \mathcal{M}$ such that the coordinate slices y^1 =constant, ..., y^{2n} =constant are integral manifolds of Δ_θ, i.e. they are everywhere tangent to the distribution. Thus, we can write $\theta_i = \sum_{j=1}^{2n} \theta_i^j dy^j\ \forall i = 1, ..., 2n$. Substituting these expressions in (5.5), we obtain the desired result. □

An analogous result in terms of $\mathcal{J}$, relating the Jacobi identity to the existence of Casimir invariants, is the following.

***Proposition* 5.2.** *Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be an antisymmetric operator of rank $2m$ on a smooth manifold $\mathcal{M}$ of dimension $n = 2m + r$. $\mathcal{J}$ satisfies the Jacobi identity $\mathfrak{G}\,(\mathcal{J}) = 0$ if and only if the distribution $\Delta_\xi = \{X \in T\mathcal{M} : \xi_i\,(X) = 0\ \ \forall\ \ i = 1, ..., r\}$ such that $ker\,(\mathcal{J}) = span\,\{\xi_1, ..., \xi_r\}$ is integrable, and the reduced operator ϕ on the $2m$-dimensional integral manifold satisfies the Jacobi identity $\mathfrak{G}\,(\phi) = 0$.*

Proof. Let $\xi_i \in T^*\mathcal{M}$, $i = 1, ..., r$, be r orthonormal cotangent vectors that span $\Delta_\theta^* = ker\,(\mathcal{J}) = span\,(\xi_1, ..., \xi_r)$, i.e. $\mathcal{J}\,(\xi_i) = 0\ \forall i = 1, ..., r$. Define an orthonormal basis of cotangent vectors on $T^*\mathcal{M}$ as $(\theta_1, ..., \theta_{2m}, \xi_1, ..., \xi_r)$ and such that $\Delta_\xi^* = span\,(\theta_1, ..., \theta_{2m})$ is the complementary distribution to Δ_θ^*, i.e. $T^*\mathcal{M} = \Delta_\theta^* \otimes \Delta_\xi^*$. Construct the covorticity $n-2$ form $\mathcal{J}^{n-2}$ on the measure $vol^n = \theta_1 \wedge ... \wedge \xi_r$ on $\mathcal{M}$. The general expression of $*\mathcal{J}^{n-2}$ is:

$$*\mathcal{J}^{n-2} = \sum_{i<j}^{2m} \alpha_{ij}\theta_i \wedge \theta_j + \sum_{i,j}^{2m,r} \beta_{ij}\theta_i \wedge \xi_j + \sum_{i<j}^{r} \gamma_{ij}\xi_i \wedge \xi_j, \tag{5.10}$$

for some appropriate coefficients $\alpha_{ij}, \beta_{ij}, \gamma_{ij}$. Following the same argument of the previous proof, we must have $*\mathcal{J}^{n-2}\,(\boldsymbol{\theta}, \boldsymbol{\xi}_n) = 0$ and $*\mathcal{J}^{n-2}\,(\boldsymbol{\xi}_n, \boldsymbol{\xi}_o) = 0$ for all $\boldsymbol{\theta} \in \Delta_\xi$ and $(\boldsymbol{\xi}_n, \boldsymbol{\xi}_o) \in \Delta_\theta$. Therefore, $\beta_{ij} = \gamma_{ij} = 0$. Now, observe that since the basis $(\theta_1 \wedge ... \wedge \xi_r)$ is orthonormal:

$$*\,(\theta_i \wedge \theta_j) = (-1)^{i+j-1}\,\theta_{ij}^{2m-2} \wedge \xi_1 \wedge ... \wedge \xi_r. \tag{5.11}$$

Taking the Hodge star of equation (5.10), we have then:

$$\mathcal{J}^{n-2} = \xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2}, \tag{5.12}$$

where we defined the $2m-2$ form $\phi^{2m-2} = \frac{1}{2}\,(-1)^{i+j-1}\,\alpha_{ij}\theta_{ij}^{2m-2}$.

Let $\left(x^1,...,x^{2m+r}\right)$ be a coordinate system on $\mathcal{M}$ such that $vol^n = \theta_1 \wedge ... \wedge \xi_r = dx^1 \wedge ... \wedge dx^{2m+r}$. In such coordinates, $\mathcal{J}^{n-2}$ takes the expression given by equation (4.11). The next step is to redefine the bracket $\{,\}$ of the system in terms of differential forms. Consider two smooth 0-forms f and g on $\mathcal{M}$. We have:

$$\begin{aligned}
\{f,g\} &= *\left(df \wedge \mathcal{J}^{n-2} \wedge dg\right) \\
&= *\left(f_i dx^i \wedge 2\sum_{j<k} (-1)^{j+k-1}\,\mathcal{J}^{jk} dx^{n-2}_{jk} \wedge g_l dx^l\right) \\
&= 2\,(-1)^n\, f_j g_k \mathcal{J}^{jk} * vol^n \\
&= 2\,(-1)^n\, f_j g_k \mathcal{J}^{jk} \\
&= (-1)^{n-2} * \left(\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge df \wedge dg\right).
\end{aligned} \tag{5.13}$$

One can see that, apart from a constant factor, this new bracket corresponds to the standard bracket $\mathcal{J}\,(df, dg)$. In terms of the new bracket, the Jacobi identity 1.12 reads as:

$$\begin{aligned}
\mathfrak{G}\,(\mathcal{J}) &= \{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} \\
&= (-1)^{n-2} * \left[\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge d\left(fd\,\{g,h\} + gd\,\{h,f\} + hd\,\{f,g\}\right)\right] \\
&= *\left\{\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge d\left[*\left(\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge dh \wedge dg\right) df + \right.\right. \\
&\quad \left.\left. *\left(\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge df \wedge dh\right) dg + *\left(\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge dg \wedge df\right) dh\right]\right\}.
\end{aligned} \tag{5.14}$$

Here f, g, h are smooth 0-forms on $\mathcal{M}$. First, suppose that the distribution $\Delta_\xi = \{X \in T\mathcal{M} : \xi_i\,(X) = 0 \;\; \forall \;\; i = 1,...,r\}$ is integrable and that the $2m-2$ dimensional operator ϕ^{2m-2} satisfies the Jacobi identity $\mathfrak{G}\,(\phi) = 0$. Since Δ_θ is integrable, from the Frobenius theorem on integrable distributions 1.2, for each $\boldsymbol{x} \in \mathcal{M}$ we can find a neighborhood $U \subset \mathcal{M}$ of $\boldsymbol{x}$ such that $\xi_i = \lambda_{j_i} dC^{j_i}$, $i = 1,...,r$, $j_i = 1,...,r$, where λ_{j_i}, C^{j_i} are 0-forms. This implies:

$$\xi_1 \wedge ... \wedge \xi_k = \sum_{j_1,...,j_k} \lambda_{j_1} \cdots \lambda_{j_k} dC^{j_1} \wedge ... \wedge dC^{j_k} = L dC^1 \wedge ... \wedge dC^k \tag{5.15}$$

for some function L. The functions C^i, $i = 1,...,r$, define a local foliation of the manifold $\mathcal{M}$. Thus, we look at the $2m-2$ dimensional submanifold defined by the level sets C^1 =constant, ..., C^r=constant, and consider the new differential $\tilde{d}$, which does not variate the C^i, together with the new Hodge star $\tilde{*}$ operating on the reduced space as $\tilde{*}\,(\theta_1 \wedge ... \wedge \theta_{2m}) = 1$. Next, decompose the differentials of the functions f, g, h in the following way:

$$df = f_{\xi_i}\xi_i + f_{\theta_j}\theta_j = f_{C^i} dC^i + f_{\theta_j}\theta_j = f_{C^i} dC^i + \tilde{d}f, \tag{5.16}$$

and similarly for g and h. The Jacobi identity (5.14) becomes:

$$\begin{aligned}&\mathfrak{G}\left(\mathcal{J}\right) = *\left\{\mathcal{J}^{n-2}\wedge\tilde{d}\left[\alpha_{ij}h_{\theta_i}g_{\theta_j}\tilde{d}f+\alpha_{ij}f_{\theta_i}h_{\theta_j}\tilde{d}g+\alpha_{ij}g_{\theta_i}f_{\theta_j}\tilde{d}h\right]\right\}\\&= *\left\{\xi_1\wedge...\wedge\xi_r\wedge\tilde{*}\tilde{*}\left\{\phi^{2m-2}\right.\right.\\&\quad\left.\left.\wedge\tilde{d}\left[\tilde{*}\left(\phi^{2m-2}\wedge\tilde{d}h\wedge\tilde{d}g\right)\tilde{d}f+\tilde{*}\left(\phi^{2m-2}\wedge\tilde{d}f\wedge\tilde{d}h\right)\tilde{d}g+\tilde{*}\left(\phi^{2m-2}\wedge\tilde{d}g\wedge\tilde{d}f\right)\tilde{d}h\right]\right\}\right\}\\&= *\left\{\xi_1\wedge...\wedge\xi_r\wedge\tilde{*}\mathfrak{G}\left(\phi\right)\right\}.\end{aligned}\tag{5.17}$$

In this passages we used the fact that the terms involving the functions C^i in the differentials df, dg, and dh identically vanish because of $\xi_1\wedge...\wedge\xi_r = LdC^1\wedge...\wedge dC^r$ in $\mathcal{J}^{n-2}$ at the left of the expression. Since by hypothesis $\mathfrak{G}\left(\phi\right)=0$, we have proven the first implication.

Now assume that $\mathfrak{G}\left(\mathcal{J}\right)=0$. First we must show that the distribution Δ_ξ is integrable. According to the Frobenius theorem 1.2, this is true provided that:

$$[X,Y]\in\Delta_\xi \quad \forall \;\; X,Y\in\Delta_\xi.\tag{5.18}$$

Observe that the n vectors $\boldsymbol{J}^i = \mathcal{J}^{mi}\partial_m$, $i=1,...,n$ form a tangent basis of the $2m$ dimensional distribution Δ_ξ. Indeed, by definition, $\boldsymbol{J}^i\left(\xi_j\right) = \mathcal{J}^{mi}\xi_{jm} = 0$ $\forall$ $i=1,...,n$, $j=1,...,r$. Therefore, any vector $X\in\Delta_\xi$ can be expressed as a linear combination of the $\boldsymbol{J}^i$ and proving (5.18) amounts at showing that:

$$\left[\boldsymbol{J}^i,\boldsymbol{J}^j\right] = A^{ij}_k\boldsymbol{J}^k,\tag{5.19}$$

for some appropriate coefficients A^{ij}_k. Substituting the expressions of the vectors $\boldsymbol{J}^i$ inside this equation, we obtain:

$$\begin{aligned}\left[\boldsymbol{J}^i,\boldsymbol{J}^j\right] &= \left(\mathcal{J}^{mi}\frac{\partial\mathcal{J}^{kj}}{\partial x^m}-\mathcal{J}^{mj}\frac{\partial\mathcal{J}^{ki}}{\partial x^m}\right)\partial_k\\&= \left(\mathcal{J}^{im}\frac{\partial\mathcal{J}^{jk}}{\partial x^m}+\mathcal{J}^{jm}\frac{\partial\mathcal{J}^{ki}}{\partial x^m}\right)\partial_k\\&= \left(\mathfrak{G}^{ijk}\left(\mathcal{J}\right)-\mathcal{J}^{km}\frac{\partial\mathcal{J}^{ij}}{\partial x^m}\right)\partial_k\\&= -\mathcal{J}^{km}\frac{\partial\mathcal{J}^{ij}}{\partial x^m}\partial_k\\&= \boldsymbol{J}^m\frac{\partial\mathcal{J}^{ij}}{\partial x^m}.\end{aligned}\tag{5.20}$$

Here, we used the fact that, by hypothesis, $\mathfrak{G}\left(\mathcal{J}\right)=0$. Thus, we have shown that $A^{ij}_k = \mathcal{J}^{ij}_m$. It follows that the distribution Δ_ξ is integrable and there are r integral manifolds C^1 =constant, ..., C^r =constant that are always tangent to the distribution.

Substituting again the expressions of the differentials df, dg, and dh, the Jacobi identity (5.14) becomes:

$$
\begin{aligned}
0 &= *\left\{\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge d\left[\alpha_{ij} h_{\theta_i} g_{\theta_j} df + \alpha_{ij} f_{\theta_i} h_{\theta_j} dg + \alpha_{ij} g_{\theta_i} f_{\theta_j} dh\right]\right\} \\
&= *\left\{\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge d\left[\alpha_{ij}\left(h_{\theta_i} g_{\theta_j} f_{\xi_k} + f_{\theta_i} h_{\theta_j} g_{\xi_k} + g_{\theta_i} f_{\theta_j} h_{\xi_k}\right)\xi_k\right.\right. \\
&\qquad\qquad \left.\left. + \alpha_{ij}\left(h_{\theta_i} g_{\theta_j} f_{\theta_k} + f_{\theta_i} h_{\theta_j} g_{\theta_k} + g_{\theta_i} f_{\theta_j} h_{\theta_k}\right)\theta_k\right]\right\} \\
&= *\left\{\alpha_{ij}\left(h_{\theta_i} g_{\theta_j} f_{\xi_k} + f_{\theta_i} h_{\theta_j} g_{\xi_k} + g_{\theta_i} f_{\theta_j} h_{\xi_k}\right)\xi_1 \wedge ... \wedge \xi_r \wedge d\xi_k \wedge \phi^{2m-2} + \xi_1 \wedge ... \wedge \xi_r\right. \\
&\qquad\qquad \left. \wedge \phi^{2m-2} \wedge d\left[\alpha_{ij}\left(h_{\theta_i} g_{\theta_j} f_{\theta_k} + f_{\theta_i} h_{\theta_j} g_{\theta_k} + g_{\theta_i} f_{\theta_j} h_{\theta_k}\right)\theta_k\right]\right\}.
\end{aligned} \tag{5.21}
$$

Since the functions f, g, and h are arbitrary, the previous equation implies:

$$\xi_1 \wedge ... \wedge \xi_r \wedge d\xi_k \wedge \phi^{2m-2} = 0, \tag{5.22a}$$

$$\xi_1 \wedge ... \wedge \xi_r \wedge \phi^{2m-2} \wedge d\left[\alpha_{ij}\beta_{ijk}\theta_k\right] = 0. \tag{5.22b}$$

(5.22c)

Here we set $\beta_{ijk} = \left(h_{\theta_i} g_{\theta_j} f_{\theta_k} + f_{\theta_i} h_{\theta_j} g_{\theta_k} + g_{\theta_i} f_{\theta_j} h_{\theta_k}\right)$. Equation (5.22a) identically vanishes because the distribution Δ_ξ is integrable and satisfies the Frobenius integrability condition:

$$\xi_1 \wedge ... \wedge \xi_k \wedge d\xi_i = 0, \qquad \forall i = 1, ..., r. \tag{5.23}$$

Finally, with the same argument of equation (5.17), one sees that the second equation (5.22b) is just the Jacobi identity $\mathfrak{G}(\phi)$ for ϕ^{2m-2}. Therefore, we have proven the second implication.

In conclusion, $\mathcal{J}$ satisfies the Jacobi identity if and only if the distribution Δ_ξ is integrable and the reduced operator ϕ satisfies the identity itself.

□

Part III

Macro-Hierarchies and Entropy Production

Chapter 6

The Fokker-Planck Equation and Equilibrium

In this chapter we derive a set of general results regarding the statistical properties of the probability distribution f of dynamical ensembles endowed with antisymmetric operators. The results presented here will be employed in the next chapters where we discuss specific physical systems.

6.1 The Fokker-Planck Equation

In order to construct the evolution equation for the probability distribution f, we must first obtain the relevant stochastic differential equations governing particle dynamics. Consider an antisymmetric operator $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ on a smooth manifold $\mathcal{M}$ of dimension n and an Hamiltonian function $H_0 \in C^\infty(\mathcal{M})$, $H_0 : \mathcal{M} \to \mathbb{R}$. The motion of a single particle is described by the differential equation:

$$\frac{d\boldsymbol{x}}{dt} = X_0 = \mathcal{J}(dH_0). \tag{6.1}$$

Now, take an ensemble of non-interacting particles, each of them obeying equation (6.1). Then, if we switch on some interaction, the energy H_0 will change according to:

$$H = H_0(\boldsymbol{x}) + H_I(\boldsymbol{x}, t), \tag{6.2}$$

where H is the new Hamiltonian function accounting for the interaction energy $H_I(\boldsymbol{x}, t)$. We take H_I, and thus H, to be C^∞ on its domain $\mathcal{M} \otimes \mathbb{R}_{\geq 0}$. The interaction is therefore represented by the vector field X_I with components:

$$X_I^i = \mathcal{J}^{ij} H_{Ii}. \tag{6.3}$$

To complete the description of particle dynamics, we further assume that all perturbations caused by H_I are counterbalanced by a *friction force*:

$$X_F^i = -\gamma^{ij} H_{0j} = -\frac{1}{2}\beta \mathcal{J}^{ik}\mathcal{J}^{jk} H_{0j} = \frac{1}{2}\beta \mathcal{J}^{ik} X_0^k. \tag{6.4}$$

Here, $\gamma^{ij} = \frac{1}{2}\beta \mathcal{J}^{ik}\mathcal{J}^{jk}$ is the friction coefficient with $\beta \in \mathbb{R}_{>0}$ a constant. One can check that, if $\mathcal{J}$ is the symplectic matrix of canonical phase space (p, q), the friction force given by equation (6.4) is the classical result $\dot{p}_F = -\frac{1}{2}\beta \dot{q}_0$. More generally, since the gradient of the Hamiltonian physically represents force, equation (6.4) leads to a total force $H_{0i} + H_{Ii} - \frac{1}{2}\beta X_0^i$ where the friction term is proportional to the velocity as in the conventional definition.

We remark that the matrix $\mathcal{J}$ appearing in the expressions (6.3) and (6.4) does not need, in general, to be the same operator defining X_0 in equation (6.1). In particular, one can, for example, truncate $\mathcal{J}$ in (6.3) and (6.4) depending on the specific physical picture, i.e. the dynamical variables that are subject to change due to interaction and friction. In the present study, we consider only the case in which $\mathcal{J}$ is exactly the same operator for all the three terms X_0, X_I, and X_F.

In summary, the equation of motion governing the dynamics of a particle in the ensemble is:

$$\begin{aligned} X &= X_0 + X_I + X_F \\ &= \left[\mathcal{J}^{ij}\left(H_{0j} + H_{Ij}\right) - \frac{1}{2}\beta \mathcal{J}^{ik}\mathcal{J}^{jk} H_{0j}\right]\partial_i \\ &= \left[\left(\mathcal{J}^{ij} - \frac{1}{2}\beta \mathcal{J}^{ik}\mathcal{J}^{jk}\right) H_{0j} + \mathcal{J}^{ij}\Gamma_j\right]\partial_i. \end{aligned} \tag{6.5}$$

In the last passage we made the substitution:

$$\mathcal{J}^{ij} H_{Ij} = \mathcal{J}^{ij}\Gamma_j. \tag{6.6}$$

Here, we assumed that the j component of the gradient of H_I is represented by Gaussian white noise, i.e. $H_{Ij} = \Gamma_j$ (see definition 3.2). We will justify this assumption later.

In the following, we will need a slightly more general form of equation (6.5). Indeed, in equation (6.5) the white noise is applied in the same coordinate system $\boldsymbol{x} = \left(x^1, ..., x^n\right)$ used to describe the dynamics. However, we want to be able to perturb the ensemble so that the noise is white in a different coordinate system, say $\boldsymbol{y} = \left(y^1, ..., y^n\right)$. Restricting to the cases in which the map $\mathcal{T} : \boldsymbol{x} \to \boldsymbol{y}$ is a diffeomorphism, we introduce the tensor $R_j^m = \partial y^m / \partial x^j$ and generalize equation (6.5) as below:

$$X = \left[\left(\mathcal{J}^{ij} - \frac{1}{2}\beta \mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k\right) H_{0j} + \mathcal{J}^{ij} R_j^r \Gamma_r\right]\partial_i. \tag{6.7}$$

Here, the friction coefficient is $\gamma^{ij} = \frac{1}{2}\beta\mathcal{J}^{ir}R_r^k\mathcal{J}^{js}R_s^k$ and we used the formula $H_{Ij} = R_j^r\Gamma_r$. As one sees from this formula, now the noise is white in the new coordinate system $\boldsymbol{y}$ since $\partial H_I/\partial y^r = \Gamma_r$.

Observe that equation (6.7) is a stochastic differential equation (see equation (3.7)). Therefore, by application of equation (3.13), we can derive the corresponding Fokker-Planck equation for the probability distribution f on the volume element $vol^n = dx^1 \wedge ... \wedge dx^n$. We have:

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial x^i}\left[-\left(\mathcal{J}^{ij} - \frac{1}{2}\beta\mathcal{J}^{ir}R_r^k\mathcal{J}^{js}R_s^k\right)H_{0j}f + \frac{1}{2}\frac{\partial}{\partial x^j}\left(\mathcal{J}^{ir}R_r^k\mathcal{J}^{js}R_s^k f\right) - \alpha\frac{\partial\mathcal{J}^{ir}R_r^k}{\partial x^j}\mathcal{J}^{js}R_s^k f\right]. \tag{6.8}$$

Finally, we must assign a specific value to the parameter $\alpha \in [0,1]$ for the stochastic differential equation (6.7) and for the Fokker-Planck equation (6.8) to make mathematically sense. Assuming that the white noise Γ appearing in our equations is the limiting representation of a continuous perturbation, we take the value $\alpha = 1/2$. We shall not be concerned with other values of α, unless differently specified. When $\alpha = 1/2$, equation (6.8) reduces to:

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial x^i}\left[-\left(\mathcal{J}^{ij} - \frac{1}{2}\beta\mathcal{J}^{ir}R_r^k\mathcal{J}^{js}R_s^k\right)H_{0j}f + \frac{1}{2}\mathcal{J}^{ir}R_r^k\frac{\partial}{\partial x^j}\left(\mathcal{J}^{js}R_s^k f\right)\right]. \tag{6.9}$$

We conclude this section by noting that the matrix R_r^k can be interpreted as the square root of a generalized diffusion parameter.

6.2 Equilibrium with Measure Preserving Operators

A look at the derived Fokker-Planck equation (6.8) shows that the behavior of the probability distribution f depends on three main factors: the energy H representing the properties of matter, the metric of space characterized by the operator $\mathcal{J}$, and the type of perturbations described by the tensor R_r^k and the parameter α (notice that R_r^k accounts for the spatial properties and α for the type of time evolution of perturbations). In this section we examine the form of f in the limit $t \to \infty$. For this purpose, it is useful to introduce the following notation:

$$f^{eq} = \lim_{t\to\infty} f. \tag{6.10}$$

Furthermore, it is convenient to define the concept of Fokker-Planck velocity Z. Since the probability $f vol^n$ enclosed in each volume element must be preserved along the trajectories, if $Z \in T\mathcal{M}$ is the dynamical flow generating the evolution of such probability, we must have the following conservation law:

$$(\partial_t + \mathfrak{L}_Z) f vol^n = \left[\frac{\partial f}{\partial t} + \frac{\partial}{\partial x^i}\left(f Z^i\right)\right] vol^n = 0. \tag{6.11}$$

Comparing this equation with the Fokker-Planck equation (6.8), wee see that:

$$Z^i = \left(\mathcal{J}^{ij} - \frac{1}{2}\beta \mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k\right) H_{0j} - \frac{1}{2f}\frac{\partial}{\partial x^j}\left(\mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k f\right) + \alpha \frac{\partial \mathcal{J}^{ir} R_r^k}{\partial x^j} \mathcal{J}^{js} R_s^k. \tag{6.12}$$

The quantity Z is called the Fokker-Planck velocity of the system.

We anticipated that, in the absence of canonical phase space, the form of f^{eq} departs from the standard Maxwell-Boltzmann distribution of equation (2.20) and assumes a novel form that depends on the operator $\mathcal{J}$. On this regard, we begin with the following convergence theorem for measure preserving operators:

***Theorem* 6.1.** *(Equilibrium with measure preserving operators)*

Hypothesis:

• *Let $\mathcal{M}$ be a smooth manifold of dimension n.*

• *Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be a measure preserving operator with $\mathcal{J}^{ij} \in C^2(\mathcal{M})$ $\forall i, j = 1, ..., n$.*

• *Let $\boldsymbol{x} = \left(x^1, ..., x^n\right)$ be the coordinate system on $\mathcal{M}$ endowed with the invariant measure, i.e. $\partial_i \mathcal{J}^{ij} = 0$ $\forall j = 1, ..., n$.*

• *Let W_i, $i = 1, ..., n$ be n Wiener processes, with $dW_i = \Gamma_i dt$ and $\alpha = 1/2$ (Stratonovich stochastic integral).*

• *Define $R_k^j = \partial_k y^j$, $j, k = 1, ..., n$, where $\boldsymbol{y} = \left(y^1, ..., y^n\right)$ is a new coordinate system such that the map $\mathcal{T} : \boldsymbol{x} \to \boldsymbol{y}$ is a diffeomorphism.*

• *Let the equations of motion be:*

$$X^i = \left(\mathcal{J}^{ij} - \gamma^{ij}\right) H_{0j} + \mathcal{J}^{ik} R_k^j \Gamma_j, \tag{6.13}$$

where the function $H(\boldsymbol{x}, t) = H_0(\boldsymbol{x}) + y^i \Gamma_i(t)$ is the Hamiltonian of the system, $H_0 \in C^\infty(\mathcal{M})$, and $\gamma^{ij} = \frac{1}{2}\beta \mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k$ is the friction coefficient with $\beta \in \mathbb{R}$ a constant.

• *The corresponding transport equation for the probability distribution $f \in C^1(\mathbb{R}_{\geq 0}) \otimes C^2(\Omega)$ on a smoothly bounded compact domain $\Omega \subset \mathcal{M}$ with volume element $vol^n = dx^1 \wedge ... \wedge dx^n$ is:*

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial x^i}\left[-\left(\mathcal{J}^{ij} - \gamma^{ij}\right) H_{0j} f + \frac{1}{2}\frac{\partial}{\partial x^j}\left(\mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k f\right) - \alpha \frac{\partial \mathcal{J}^{ir} R_r^k}{\partial x^j} \mathcal{J}^{js} R_s^k f\right]. \tag{6.14}$$

Suppose that on the boundary $\partial\Omega$ the conditions $Z \cdot N = 0$ and $X_0 \cdot N = 0$ hold, with Z the Fokker-Planck velocity such that $\partial_t f = -\partial_i \left(f Z^i \right)$, $X_0 = \mathcal{J}^{ij} H_{0j} \partial_i$, and N the outward normal to $\partial\Omega$.

• *Assume that $f > 0$ on Ω.*

Thesis:

Then, the solution to (6.14) is such that:

$$\lim_{t\to\infty} \mathcal{J} \left(d\log f + \beta dH_0 \right) = 0 \qquad a.e., \tag{6.15}$$

for any choice of the diffeomorphic coordinates y^j, $j = 1, ..., n$.

Proof. Recalling the expression of the Fokker-Planck velocity Z, equation (6.12), and setting $\alpha = 1/2$ we obtain:

$$Z^i = \left(\mathcal{J}^{ij} - \gamma^{ij} \right) H_{0j} - \frac{1}{2} \mathcal{J}^{ir} R^k_r \mathcal{J}^{js} R^k_s \frac{\partial \log f}{\partial x^j}. \tag{6.16}$$

In going from (6.12) to this expression, we used the fact that $\mathcal{J}$ is measure preserving ($\partial_i \mathcal{J}^{ij} = 0$, $j = 1, ..., n$) and that the matrix $R^k_{sj} = \partial^2 y^k / \partial x^s \partial x^j$ is symmetric so that $\mathcal{J}^{sj} R^k_{sj} = 0$, $k = 1, ..., n$. Consider now the following entropy functional:

$$S = -\int_\Omega f \log f \, vol^n. \tag{6.17}$$

The rate of change of S is:

$$\begin{aligned} \frac{dS}{dt} &= -\int_\Omega \frac{\partial f}{\partial t} \left(1 + \log f \right) vol^n \\ &= \int_\Omega \frac{\partial \left(f Z^i \right)}{\partial x^i} \left(1 + \log f \right) vol^n \\ &= \int_\Omega f \frac{\partial Z^i}{\partial x^i} vol^n + \int_{\partial\Omega} f \log f \, Z^i N_i \, dS^{n-1} \\ &= \int_\Omega f \frac{\partial Z^i}{\partial x^i} vol^n \\ &= \int_{\partial\Omega} f Z^i N_i \, dS^{n-1} - \int_\Omega f_i Z^i \, vol^n \\ &= -\int_\Omega f_i Z^i \, vol^n. \end{aligned} \tag{6.18}$$

Here we used the fact that $Z^i N_i$ vanish on the boundary $\partial\Omega$. In this notation $N = N_i \partial_i$ is the outward normal to the bounding surface $\partial\Omega$ with surface element dS^{n-1}. Substituting (6.16) in (6.18) we get:

$$\begin{aligned} \frac{dS}{dt} &= -\int_\Omega f_i \mathcal{J}^{ij} H_{0j} \, vol^n + \frac{1}{2} \int_\Omega f_i \mathcal{J}^{ir} R^k_r \mathcal{J}^{js} R^k_s \left(\frac{\partial \log f}{\partial x^j} + \beta H_{0j} \right) vol^n \\ &= \frac{1}{2} \int_\Omega f_i \mathcal{J}^{ir} R^k_r \mathcal{J}^{js} R^k_s \left(\frac{\partial \log f}{\partial x^j} + \beta H_{0j} \right) vol^n. \end{aligned} \tag{6.19}$$

Here we used the fact that $\mathcal{J}$ is measure preserving and thus the term involving $f_i \mathcal{J}^{ij} H_{0j} = \frac{\partial}{\partial x^i}\left(f X_0^i\right)$ can be written as a vanishing surface integral. Consider now conservation of total energy $E = \int_\Omega f \beta H_0 \, vol^n$:

$$
\begin{aligned}
\frac{dE}{dt} &= -\int_\Omega \beta \frac{\partial\left(f Z^i\right)}{\partial x^i} H_0 \, vol^n \\
&= \int_\Omega \beta f Z^i H_{0i} \, vol^n \\
&= \int_\Omega \beta f \mathcal{J}^{ij} H_{0j} H_{0i} \, vol^n - \frac{1}{2}\int_\Omega f \mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k \left(\frac{\partial \log f}{\partial x^j} + \beta H_{0j}\right) \beta H_{0i} \, vol^n \\
&= -\frac{1}{2}\int_\Omega \mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k f_j \beta H_{0i} \, vol^n - \frac{1}{2}\int_\Omega f \left(\beta \mathcal{J}^{ir} R_r^k H_{0i}\right)^2 \, vol^n \\
&= 0.
\end{aligned}
\tag{6.20}
$$

Again, we used the fact that surface integrals vanish and the antisymmetry of $\mathcal{J}$. This implies:

$$
\int_\Omega \mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k f_j \beta H_{0i} \, vol^n = -\int_\Omega f \left(\beta \mathcal{J}^{ir} R_r^k H_{0i}\right)^2 \, vol^n. \tag{6.21}
$$

Substituting this result in (6.19), we arrive at:

$$
\frac{dS}{dt} = \frac{1}{2}\int_\Omega f \left[\left(\mathcal{J}^{ir} R_r^k \frac{\partial \log f}{\partial x^i}\right)^2 - \left(\beta \mathcal{J}^{ir} R_r^k H_{0i}\right)^2\right] vol^n \tag{6.22}
$$

$$
= \frac{1}{2}\int_\Omega f \left\{\left[\mathcal{J}^{ir} R_r^k \left(\frac{\partial \log f}{\partial x^i} + \beta H_{0i}\right)\right]^2 - 2\left(\beta \mathcal{J}^{ir} R_r^k H_{0i}\right)^2 - 2\mathcal{J}^{ir} R_r^k \mathcal{J}^{js} R_s^k f_j \beta H_{0i}\right\} vol^n.
$$

Using again conservation of energy (6.20), the last two terms in the final passage vanish, and we obtain:

$$
\frac{dS}{dt} = \frac{1}{2}\int_\Omega f \left[\mathcal{J}^{ir} R_r^k \left(\frac{\partial \log f}{\partial x^i} + \beta H_{0i}\right)\right]^2 vol^n. \tag{6.23}
$$

It is useful to rewrite equation (6.23) in matrix notation: defining $R^T = \left(R_r^k dx^r \otimes \partial_k\right)^T = R_r^k \partial_k \otimes dx^r$, we have:

$$
\frac{dS}{dt} = \frac{1}{2}\int_\Omega f \left|R^T \mathcal{J}\left(d\log f + \beta dH_0\right)\right|^2 vol^n. \tag{6.24}
$$

In the limit of thermodynamic equilibrium we must have:

$$
\lim_{t\to\infty} \frac{dS}{dt} = 0. \tag{6.25}
$$

In light of equation (6.23), this implies that for all non-zero f:

$$
\lim_{t\to\infty} \mathcal{J}\left(d\log f + \beta dH_0\right) = 0 \qquad a.e. \tag{6.26}
$$

for any choice of the diffeomorphic coordinates $\boldsymbol{y}$. Notice that the matrix R_r^k could be removed because the transformation $\mathcal{T} : \boldsymbol{x} \to \boldsymbol{y}$ is a diffeomorphism and is therefore invertible. $\square$

Let us make some considerations on the meaning and the physical implications of this result.

• The first aspect we want to stress is that the reason why equation (6.15) holds is that $\mathcal{J}$ is taken to be measure preserving. Without this assumption, there would not be any grounds to examine the entropy functional of equation (6.17). In fact, one can check that in the case of a general antisymmetric operator, the rate of change of this functional in not positively defined (and thus violates the second law of thermodynamics).

• Secondly, it is important to remark that f is the probability distribution *on* the invariant measure dictated by the measure preserving operator $\mathcal{J}$. Only in such coordinate system (6.17) has proper physical meaning, i.e. the entropy production represented by equation (6.23) has a definite sign and therefore an extremum principle (maximum entropy) applies to the functional (6.17). If g is the Jacobian of the coordinate change sending the invariant measure vol^n to a different reference system $vol_C^n = g^{-1} vol^n$, the probability distribution in the new frame would be $u = fg$. Here, the letter C stands for Cartesian, since usually one is interested in the probability distribution observed in the Cartesian coordinate system of the laboratory frame. Recalling the change of coordinates formula for information entropy, equation (2.18), and defining the information entropy of the new distribution u as $S_C = -\int_\Omega u \log u \, vol_C$, we have the following definition:

***Def* 6.1.** *(Proper entropy for measure preserving operators)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}$ be a measure preserving operator on a smooth manifold $\mathcal{M}$ of dimension n. Then, on the invariant measure vol^n, there is a proper and thermodynamically consistent information entropy measure Σ:

$$\begin{aligned} \Sigma &= -\int_\Omega f \log f \, vol^n \\ &= -\int_\Omega f \log(fg) \, vol^n + \int_\Omega f \log g \, vol^n \\ &= S_C + \langle \log g \rangle , \end{aligned} \tag{6.27}$$

where f is the probability distribution on vol^n, $\Omega \subset \mathcal{M}$, and S_C is the information entropy of the probability distribution $u = fg$ on $vol_C^n = g^{-1} vol^n$.

It is also useful to introduce the notion of entropy production. The rate of change in Σ is given by:

$$
\begin{aligned}
\frac{d\Sigma}{dt} &= -\int_{\Omega} \frac{\partial f}{\partial t} \left(1 + \log f\right) \, vol^n \\
&= \int_{\Omega} \frac{\partial \left(f Z^i\right)}{\partial x^i} \left(1 + \log f\right) \, vol^n \\
&= \int_{\Omega} f \frac{\partial Z^i}{\partial x^i} \, vol^n + \int_{\partial\Omega} f \log f \, Z^i N_i \, dS^{n-1} \\
&= \sigma - \Phi
\end{aligned}
\tag{6.28}
$$

The term:

$$
\sigma = \int_{\Omega} f \frac{\partial Z^i}{\partial x^i} \, vol^n, \tag{6.29}
$$

is called the entropy production rate of the system. The term:

$$
\Phi = -\int_{\partial\Omega} f \log f \, Z^i N_i \, dS^{n-1}, \tag{6.30}
$$

represents the flow of entropy across the boundary $\partial\Omega$. Notice that the entropy production σ is caused by the divergence of the Fokker-Planck velocity Z. Whenever the Fokker-Planck velocity Z is divergence free and the system is closed (i.e. $\Phi = 0$), the entropy Σ remains constant.

• The third remark is that, according to proposition 4.3, we can always take any antisymmetric operator of dimension n and extend it to a measure preserving form in a $n+1$ dimensional setting. Here, the result of theorem 6.1 apply. However, it must be noted that the equilibrium obtained from the extended operator has a different physical meaning from the equilibrium in the original system. Indeed, even though the particle Hamiltonian H_0 does not depend on the new variable x^{n+1}, the noise Γ_{n+1} associated to this variable affect the other coordinates through the interaction term $H_I = y^i \Gamma_i$. If $\mathfrak{J}$ is the extended $n+1$ dimensional measure preserving operator, we have:

$$
\dot{x}^i = \mathfrak{J}^{ij} H_j = \sum_{j=1}^{n} \mathcal{J}^{ij} \left(H_{0j} + H_{Ij}\right) + \mathfrak{J}^{i,n+1} R^k_{n+1} \Gamma_k. \tag{6.31}
$$

The second term on the right-hand side exists only in virtue of the extension and does not appear in the original n-dimensional dynamics.

• It is useful to add some explanation to the chosen boundary conditions $Z \cdot N = 0$ and $X_0 \cdot N = 0$ on $\partial\Omega$. Physically, they express the fact that probability does not escape from the domain Ω, and therefore the system can be considered as thermodynamically closed. The condition $X_0 \cdot N = 0$ can be thought as a definition of the boundary itself, and can be satisfied, for example, by taking an Hamiltonian H_0 that is constant on the boundary, $H_{0i} = 0$ on $\partial\Omega$. The condition $Z \cdot N = 0$ is rather a boundary condition for f. If $H_{0i} = 0$ on $\partial\Omega$ one can use the Neumann boundary condition $df = 0$ on $\partial\Omega$.

• If the matrix $\mathcal{J}$ is invertible, equation (6.15) becomes:

$$f^{eq} = \lim_{t\to\infty} f = A \exp\{-\beta H_0\} \qquad a.e., \tag{6.32}$$

where $A \in \mathbb{R}_{>0}$ is a normalization constant. Thereby, we can rephrase the result of theorem 6.1 in the following way: if the metric of space if vortex free, i.e. $\mathcal{O}^{n-1} = 0$, and space is accessible, i.e. $ker\,(\mathcal{J}) = 0$, the standard result of statistical mechanics apply on the invariant measure.

The effect of a non-trivial kernel $ker\,(\mathcal{J}) \neq 0$ can be understood with the next corollary of theorem 6.1:

***Corollary* 6.1.** *(Equilibrium with Poisson operators)*
Assume the hypothesis of theorem 6.1. In addition, assume that $\mathcal{J}$ has constant rank $2m = n-r$ and that it is a Poisson operator satisfying the Jacobi identity $\mathfrak{S}\,(\mathcal{J}) = 0$. Furthermore, assume that the limit $f^{eq} = \lim_{t\to\infty} f$ is itself of class $C^2\,(\Omega)$. Then, for almost every point $\boldsymbol{x} \in \Omega$ there exists a neighborhood $U \subset \Omega$ of $\boldsymbol{x}$ such that:

$$f^{eq} = \lim_{t\to\infty} f = A \exp\left\{-\beta H_0 - \gamma_i C^i\right\} \qquad on \;\; U, \tag{6.33}$$

where $\gamma_i \in \mathbb{R}$, $i = 1, ..., r$, are constants and the functions C^i are the r Casimir invariants whose gradients span the kernel of $\mathcal{J}$, i.e. $\mathcal{J}\left(dC^i\right) = 0$.

Proof. Thanks to Darboux's theorem 1.1, $\forall \boldsymbol{x} \in \Omega$ there exists a neighborhood $U \subset \Omega$ of $\boldsymbol{x}$ where we can find coordinates $\left(u^1, ..., u^{2m}, C^1, ..., C^r\right)$ such that the C^i are Casimir invariants. Thus, exception made for a set of measure zero, the local solution to equation (6.15) is of the form (6.33). □

In the case of a non-canonical Hamiltonian system, we see that statistical equilibrium, which is achieved on the invariant measure assigned by Liouville's theorem 2.1, is determined by the energy H_0 and the Casimir invariants C^i. In this way, the functions C^i impart a non-trivial structure to the probability distribution f:

***Def* 6.2.** *(Self-organization by Casimir invariants)*
The self-organized probability distribution caused by a Casimir invariant is called a type-I distribution. The associated self-organizing process is a type-I self-organization.

As we have explained in the introduction, this type of self-organization is caused by the existence of inaccessible regions in the phase space represented by the fact that $i_X dC^i = \mathcal{J}\left(dC^i, dH\right) = 0$.

• The last remark concerns the white noise assumption we made. This assumption must be justified on a case by case basis by showing that the perturbations affecting a certain ensemble statistically behave as Gaussian white noise in some appropriate coordinate system $\boldsymbol{y}$ (in the sense the gradient $\partial H_I/\partial y^r$ of the interaction Hamiltonian H_I with respect to the coordinates $\boldsymbol{y}$ can be considered as Gaussian white noise). In practical situations, using the invariant measure provided by the measure preserving operator, one invokes the ergodic hypothesis in virtue of the ergodic theorem 2.3, and exploits the fact that ensemble and time averages can be interchanged to show that the fluctuations can be linked to Gaussian white noise. In the next chapter we will discuss concrete examples of this procedure.

Finally, notice that equation (6.15) does not depend on the specific coordinates $\boldsymbol{y}$ where noise is white. This means that, regardless of the coordinate frame where a system is homogeneously perturbed, statistical equilibrium is achieved on the invariant measure.

6.3 Equilibrium with Beltrami Operators

So far, we have studied the equilibrium probability distribution for the class of measure preserving operators. We now move to operators that are not endowed with an invariant measure. Since the problem becomes mathematically more convoluted, we proceed with a gradual approach.

First, consider the case of pure diffusion, $H_0 = 0$. Then, from equation (6.7), the relevant equation of motion reads:

$$X = \left(\mathcal{J}^{ij} R^r_j \Gamma_r\right) \partial_i. \tag{6.34}$$

To further simplify the problem, set $R^r_j = \delta^r_j$ to obtain:

$$X = \left(\mathcal{J}^{ij} \Gamma_j\right) \partial_i. \tag{6.35}$$

Recalling the transport equation (6.8) and setting $\alpha = 1/2$, we arrive at the corresponding diffusion equation:

$$\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial}{\partial x^i}\left[\mathcal{J}^{ik}\frac{\partial\left(\mathcal{J}^{jk} f\right)}{\partial x^j}\right]. \tag{6.36}$$

It is instructive to rewrite this equation in terms of the covorticity $n-2$ form $\mathcal{J}^{n-2}$ given by (4.11). Using the same calculation of equation (4.47), one obtains:

$$\frac{\partial f}{\partial t}=\frac{1}{8}*d\left[\mathcal{J}^{n-2}\wedge *d\left(\mathcal{J}^{n-2}f\right)\right]. \tag{6.37}$$

Now assume that $\mathcal{J}$ is a Beltrami operator ($\mathfrak{B}=0$, see definition 4.10). We have the following:

***Theorem* 6.2.** *(Diffusion with Beltrami operators)*
Let $\mathcal{J}\in\bigwedge^2 T\mathcal{M}$ a Beltrami operator on a smooth manifold $\mathcal{M}$ of dimension n. Consider the diffusion equation (6.36) for the probability distribution $f\in C^1\left(\mathbb{R}_{\geq 0}\right)\otimes C^2\left(\Omega\right)$, $f>0$, on a smoothly bounded compact domain $\Omega\subset\mathcal{M}$. Assume the boundary conditions $Z\cdot N=0$ and $\boldsymbol{b}\cdot N=0$ on $\partial\Omega$, where Z is the Fokker-Planck velocity such that $\partial_t f=-\partial_i\left(fZ^i\right)$, $\boldsymbol{b}=\mathcal{J}^{ik}\mathcal{J}^{jk}_j\partial_i$ is the field force, and N the outward normal to $\partial\Omega$. Then,

$$\lim_{t\to\infty}\mathcal{J}\left(d\log f\right)=0\qquad a.e. \tag{6.38}$$

Proof. Consider the entropy functional:

$$S=-\int_\Omega f\log f\,vol^n. \tag{6.39}$$

The rate of change in S is:

$$\begin{aligned}
\frac{dS}{dt}&=-\int_\Omega\frac{\partial f}{\partial t}\left(1+\log f\right)vol^n\\
&=\int_\Omega\frac{\partial\left(fZ^i\right)}{\partial x^i}\left(1+\log f\right)vol^n\\
&=\int_\Omega f\frac{\partial Z^i}{\partial x^i}\,vol^n+\int_{\partial\Omega}f\log f\,Z^iN_i dS^{n-1}\\
&=-\int_\Omega Z^i f_i\,vol^n+\int_{\partial\Omega}fZ^iN_idS^{n-1}\\
&=\frac{1}{2}\int_\Omega f_i f^{-1}\mathcal{J}^{ik}\frac{\partial\left(\mathcal{J}^{jk}f\right)}{\partial x^j}\,vol^n\\
&=\frac{1}{2}\int_\Omega\left[f_i\mathcal{J}^{ik}\frac{\partial\mathcal{J}^{jk}}{\partial x^j}+f_i\mathcal{J}^{ik}\mathcal{J}^{jk}\frac{\partial\log f}{\partial x^j}\right]vol^n\\
&=\frac{1}{2}\int_{\partial\Omega}f\mathcal{J}^{ik}\frac{\partial\mathcal{J}^{jk}}{\partial x^j}N_i\,dS^{n-1}\\
&\quad+\frac{1}{2}\int_\Omega\left[-f\frac{\partial}{\partial x^i}\left(\mathcal{J}^{ik}\frac{\partial\mathcal{J}^{jk}}{\partial x^j}\right)+f\left|\mathcal{J}\left(d\log f\right)\right|^2\right]vol^n\\
&=\frac{1}{2}\int_\Omega\left[-\frac{f}{4}\mathfrak{B}+f\left|\mathcal{J}\left(d\log f\right)\right|^2\right]vol^n\\
&=\frac{1}{2}\int_\Omega f\left|\mathcal{J}\left(d\log f\right)\right|^2vol^n.
\end{aligned} \tag{6.40}$$

Here, we used the boundary conditions to eliminate surface integrals and the vanishing of $\mathfrak{B}$. We conclude that for any non-zero f:

$$\lim_{t\to\infty} \mathcal{J}\left(d\log f\right) = 0 \qquad a.e. \tag{6.41}$$

□

This result deserves comments.

• First, we must explain the chosen boundary conditions. As for theorem 6.1, the physical meaning of the requirements $Z \cdot N = 0$ and $\boldsymbol{b} \cdot N = 0$ on $\partial\Omega$ is that probability does not escape from the boundaries. If the diffusion equation is written in terms of the standard Cartesian coordinate system of $\mathbb{R}^n$, the components of the vector $\boldsymbol{b}$ correspond to the components of the field force $n-1$ form (4.45) and, when $n = 3$, one obtains the field force vector encountered in section 4.5. The vector $\boldsymbol{b}$ acts as an effective drift. Indeed, from equation (6.36), one sees that the Fokker-Plack velocity Z can be decomposed in the form below:

$$Z^i = \frac{1}{2f}\mathcal{J}^{ik}\frac{\partial\left(\mathcal{J}^{jk}f\right)}{\partial x^j} = \frac{1}{2}b^i + \frac{1}{2}\mathcal{J}^{ik}\mathcal{J}^{jk}\frac{\partial \log f}{\partial x^j}. \tag{6.42}$$

Thus, $\boldsymbol{b} \cdot N = 0$ on $\partial\Omega$ means that the boundary must be chosen so that the drift $\boldsymbol{b}$ does not transport any probability out of the domain Ω. The second condition $Z \cdot N = 0$ is a boundary condition for the probability distribution f. A possible way to satisfy these conditions is, for example, to assume that $\mathcal{J}$ is a strong Beltrami operator in a Cartesian coordinate system so that $\boldsymbol{b} = \boldsymbol{0}$ and then take the Neumann boundary condition $df = 0$ on $\partial\Omega$ for f.

• The result of equation (6.38) is remarkable from the standpoint of statistical mechanics. Let us explain why. If $\mathcal{J}$ happens to be measure preserving, the result (6.38) is expected: if $\mathcal{J}$ is invertible, we get a flat distribution $df = 0$ almost everywhere on the invariant measure. Now, suppose that $\mathcal{J}$ is still invertible but not measure preserving: we obtain a flat distribution $df = 0$ almost everywhere even if no invariant measure exists! In other words, the Beltrami operator is the largest class of antisymmetric operators such that the diffusion equation (6.36) admits the solution $f =$constant. This fact can be verified by substituting the solution $f =$constant in equation (6.36). One obtains:

$$0 = \frac{1}{8}\mathfrak{B}. \tag{6.43}$$

This is possible only if $\mathfrak{B} = 4\partial_i b^i = 0$. Therefore, in the case of diffusion with a Beltrami operator, the entropy functional (6.39) on the coordinate system where

$\mathfrak{B} = 0$ is physically consistent. Notice that such coordinate system is not, in general, the standard Cartesian coordinate system.

6.4 Equilibrium with Antisymmetric Operators

Unfortunately, beyond the case of pure diffusion with Beltrami operators presented in the last section, the determination of f in the limit $t \to \infty$ becomes a sensibly more complicated problem, because we need to solve a non-elliptic second order partial differential equation. As we have outlined in the introduction, there is no systematic mathematical theory for such kind of equations. In the final chapter of this study we will try to lay the ground for such systematic treatment, and discuss existence and uniqueness of solutions to the normal Laplacian, the non-elliptic second order differential operator we introduce to calculate the equilibrium distribution. Again, we restrict our attention to the easier case of pure diffusion, equation (6.36). Let us show why, for a general $\mathcal{J}$, this equation is non-elliptic. First, consider the linear differential equation with respect to the function u:

$$\alpha^{ij}\left(\boldsymbol{x}\right) u_{ij} + \beta^{i}\left(\boldsymbol{x}\right) u_{i} + \gamma\left(\boldsymbol{x}\right) u + \delta\left(\boldsymbol{x}\right) = 0. \tag{6.44}$$

Here, the subscripts mean derivation, i.e. $u_{ij} = \partial^2 u/\partial x^i \partial x^j$. This equation is elliptic if the matrix α^{ij} is positive definite in the domain of its arguments:

$$\alpha^{ij}\xi_i\xi_j > 0 \quad \forall \xi \neq 0. \tag{6.45}$$

A comparison with equation (6.36) shows that:

$$\frac{1}{2}\mathcal{J}^{ik}\mathcal{J}^{jk} f_{ij} + \frac{1}{2}\left[b^j + \frac{\partial}{\partial x^i}\left(\mathcal{J}^{ik}\mathcal{J}^{jk}\right)\right] f_j + \frac{1}{2}\frac{\partial b^i}{\partial x^i} f = 0. \tag{6.46}$$

Therefore, in the specific case of interest the matrix α^{ij} is:

$$\alpha^{ij} = \frac{1}{2}\mathcal{J}^{ik}\mathcal{J}^{jk}. \tag{6.47}$$

It is clear that if we take $\xi \in ker\left(\mathcal{J}\right)$, the condition of ellipticity, equation (6.45), cannot be satisfied.

In this section we will study the converge of the solution for a general antisymmetric operator (i.e. an operator that does not fall in any of the classes we have defined so far) for special cases that do not require the discussion of the solvability of the normal Laplace equation (which we will introduce in the final chapter). To simplify the problem, we set $n = 3$. Equation (6.36) reduces to:

$$\frac{\partial f}{\partial t} = \frac{1}{2}\nabla \cdot \left[\boldsymbol{w} \times \left(\nabla \times f\boldsymbol{w}\right)\right]. \tag{6.48}$$

We have the following:

Theorem 6.3. *(Diffusion with antisymmetric operators in 3D)*
Let $\boldsymbol{w}$ be a constant rank antisymmetric operator on a smooth manifold $\mathcal{M}$ of dimension $n = 3$. Consider the diffusion equation (6.48) on a domain $\Omega \subset \mathcal{M}$ and assume $f \in C^1(\mathbb{R}_{\geq 0}) \otimes C^2(\Omega)$, $f > 0$. We study the following three cases:

- *Suppose that $\mathfrak{S}(\boldsymbol{w}) = 0$ so that for each $\boldsymbol{x} \in \Omega$ there exists a neighborhood $U \subset \Omega$ of $\boldsymbol{x}$ where $\boldsymbol{w} = \lambda \nabla C$. Further assume that we can extend the representation $\boldsymbol{w} = \lambda \nabla C$ to the whole Ω. Consider the boundary condition $Z \cdot N = 0$ on $\partial\Omega$. Then:*

$$\lim_{t\to\infty} \nabla C \times \nabla(\lambda f) = 0 \quad a.e. \tag{6.49}$$

- *Suppose that $\mathfrak{S}(\boldsymbol{w}) \neq 0$ and $\mathfrak{B} = 0$. Consider the boundary conditions $Z \cdot N = 0$ and $\boldsymbol{b} \cdot N = 0$ on $\partial\Omega$. Then:*

$$\lim_{t\to\infty} \boldsymbol{w} \times \nabla f = \boldsymbol{0} \quad a.e. \tag{6.50}$$

- *Suppose that $\mathfrak{S}(\boldsymbol{w}) \neq 0$ and $\mathfrak{B} \neq 0$. Consider the boundary condition $Z \cdot N = \boldsymbol{0}$ on $\partial\Omega$. Further assume that the field force vector $\hat{\boldsymbol{b}} = \hat{\boldsymbol{w}} \times (\nabla \times \hat{\boldsymbol{w}})$ of the normalized vector field $\hat{\boldsymbol{w}} = \boldsymbol{w}/w$ can be written by means of a a scalar potential ζ as $\hat{\boldsymbol{b}} = \nabla\zeta$. Then:*

$$\lim_{t\to\infty} \boldsymbol{w} \times \nabla\left[\log(fw) + \zeta\right] = \boldsymbol{0} \quad a.e. \tag{6.51}$$

Proof. We begin with the first statement. Recalling the result of proposition 4.6, since $\boldsymbol{w} = \lambda\nabla C$ on Ω, $\lambda^{-1} dx \wedge dy \wedge dz$ is an invariant measure on Ω and $\boldsymbol{w}$ is a measure preserving operator on Ω. Therefore, on the invariant measure, we can apply theorem 6.1 for $H_0 = 0$ and obtain:

$$\lim_{t\to\infty} \boldsymbol{w} \times \nabla(f\lambda) = \boldsymbol{0} \quad a.e. \tag{6.52}$$

This is the desired result.

The second statement is the three dimensional version of theorem 6.2.

We move on to the third statement. Consider the following entropy functional:

$$S = -\int_\Omega f[\log(fw) + \zeta]\, vol^3. \tag{6.53}$$

The rate of change in entropy is:

$$\begin{aligned} \frac{dS}{dt} &= -\int_\Omega \partial_t f\left[1 + \log(fw) + \zeta\right] vol^3 \\ &= -\frac{1}{2}\int_{\partial\Omega} \boldsymbol{w} \times (\nabla \times f\boldsymbol{w}) \cdot \boldsymbol{n}\left[1 + \log(fw) + \zeta\right] dS^2 \\ &\quad + \frac{1}{2}\int_\Omega fw^2\left[\nabla\zeta + \nabla_\perp \log(fw)\right] \cdot \nabla[\log(fw) + \zeta]\, vol^3 \\ &= \frac{1}{2}\int_\Omega fw^2\left[\nabla\zeta + \nabla_\perp \log(fw)\right] \cdot \nabla[\log(fw) + \zeta]\, vol^3. \end{aligned} \tag{6.54}$$

The surface integral was eliminated by application of the boundary condition $Z \cdot N = 0$ on $\partial\Omega$. Now observe that, since $\hat{\boldsymbol{b}} \cdot \hat{\boldsymbol{w}} = 0$, $\hat{\boldsymbol{b}} = \nabla\zeta = \nabla_\perp\zeta$ and therefore the above equation becomes:

$$\begin{aligned} \frac{dS}{dt} &= \frac{1}{2}\int_\Omega f w^2 \left|\nabla_\perp \left[\log\left(fw\right) + \zeta\right]\right|^2 vol^3 \\ &= \frac{1}{2}\int_\Omega f \left|\boldsymbol{w} \times \nabla \left[\log\left(fw\right) + \zeta\right]\right|^2 vol^3. \end{aligned} \tag{6.55}$$

Therefore, for any $f > 0$, we must have:

$$\lim_{t\to\infty} \boldsymbol{w} \times \nabla[\log\left(fw\right) + \zeta] = \boldsymbol{0} \quad a.e. \tag{6.56}$$

$\square$

The following remarks are necessary.

• The remaining case where $\hat{\boldsymbol{b}}$ cannot be expressed as the gradient of some scalar potential will be outlined in the final chapter because, as we have anticipated, we need to define the normal Laplace operator.

• Notice that the geometric properties of the vector field $\boldsymbol{w}$ determine the specific coordinate system where entropy is maximized.

In the first case, $\boldsymbol{w}$ is integrable and defines a Poisson operator. The resulting equilibrium, equation (6.49), is a type-I self-organization on the invariant measure of the system. Specifically, assuming that the limit $\lambda f^{eq} = \lim_{t\to\infty} \lambda f$ is itself $C^2\left(\Omega\right)$, from equation (6.49) we see that:

$$f^{eq} = \frac{1}{\lambda}\mathcal{F}\left(C\right) \qquad a.e. \tag{6.57}$$

Here $\mathcal{F} = \mathcal{F}\left(C\right)$ is a function of the Casimir invariant C.

In the second case, $\boldsymbol{w}$ is a Beltrami operator. From equation (6.50) one obtains:

$$\lim_{t\to\infty} \nabla f = \lambda\boldsymbol{w} \qquad a.e., \tag{6.58}$$

for some appropriate function λ. Assuming that the limit $f^{eq} = \lim_{t\to\infty} f$ is itself $C^2\left(\Omega\right)$, since the components of $\boldsymbol{w}$ are smooth, λ is at least $C^1\left(\Omega\right)$. Then, we must have $\lambda = 0$ almost everywhere to guarantee that $\mathfrak{H}\left(\boldsymbol{w}\right) = \boldsymbol{w} \cdot \nabla \times \boldsymbol{w} \neq 0$ on Ω. We conclude that:

$$f^{eq} = c \qquad a.e. \tag{6.59}$$

Here $c \in \mathbb{R}_{>0}$ is some non-zero positive constant. This flat distribution tells us that a Beltrami operator prevents the emergence of any kind of self-organized structure.

We remind the reader that the distribution f is defined on the reference frame where $\mathfrak{B} = \nabla \cdot [\boldsymbol{w} \times (\nabla \times \boldsymbol{w})] = 0$.

The third case is characterized by an antisymmetric operator of the most general type with the condition that $\hat{\boldsymbol{b}} = \nabla\zeta$. The associated system is neither Hamiltonian, nor measure preserving, nor Beltrami. Assuming that the limit $\lim_{t\to\infty} fwe^{\zeta}$ is still $C^2(\Omega)$, with a similar reasoning to the previous case, it follows that:

$$f^{eq} = \frac{c}{w}e^{-\zeta} \quad a.e., \tag{6.60}$$

for some non-zero positive constant $c \in \mathbb{R}_{>0}$. Observe that:

$$\hat{\mathfrak{B}} = \nabla \cdot \hat{\boldsymbol{b}} = \nabla \cdot (\nabla\zeta) = \Delta\zeta. \tag{6.61}$$

Therefore, if we could define the inverse Laplacian Δ^{-1} in some appropriate sense, we would get $\zeta = \Delta^{-1}\hat{\mathfrak{B}}$. Then, the thermodynamically consistent entropy functional (6.53) would become:

$$S = -\int_{\Omega} f\left[\log(fw) + \Delta^{-1}\hat{\mathfrak{B}}\right] vol^3. \tag{6.62}$$

The novel term depending on $\hat{\mathfrak{B}}$ clearly reflects the fact that the metric of space is intrinsically affected by a non-vanishing cocurrent that cannot be removed by any coordinate change. The resulting self-organized distribution, equation (6.60), is the manifestation of a new type of self-organized structure that arises in virtue of such 'metric' current. It is the right juncture to introduce a second notion of self-organization by topological constraints:

***Def* 6.3.** *(Self-organization by metric current)*
The self-organized probability distribution caused by the cocurrent $\mathcal{O}^{n-1}$ is called a type-II distribution. The associated self-organizing process is a type-II self-organization.

***Remark* 6.1.** *In general, type-I and type-II self-organization can occur together. This could happen, for example, if $\mathcal{J}$ is not measure preserving and at least a certain part of its null space $ker(\mathcal{J})$ is spanned by Casimir invariants.*

***Remark* 6.2.** *Notice that whenever a certain system is observed in a coordinate system where $\mathcal{O} \neq 0$, the probability distribution on such reference frame will be a type-II distribution. This is true even if there exists another coordinate system where the associated cocurrent $\mathcal{O}'$ vanishes. In other words, even if a physical system is measure preserving, if one observes it in the wrong coordinates, the information entropy defined over such coordinates will not be maximized and a 'fictitious' structure will appear. Such structure disappears if one moves back to the invariant measure.*

All the results presented in this section will be verified with concrete physical examples and detailed numerical simulations in the following chapters.

• From the discussion of the previous section it should be clear that once we leave the realm of measure preserving dynamics, even a simple three dimensional diffusion process poses delicate physical and mathematical problems. Among these problems, the central one pertains to the mathematical identity of the normal Laplacian $\Delta_{\perp}$, which we will investigate in the last chapter. Other fundamental questions are the mathematical nature of the limit distribution f^{eq}, i.e. the function space to which f^{eq} belongs, and the feasibility of the non-trivial boundary conditions we assume, as well as their physical interpretation.

Finally, it is worth to mention that there is a slightly different approach that can be implemented to determine a suitable metric where to define an appropriate entropy functional: time dependent metrics (i.e. a time dependent coordinate change). In this approach, one looks for a time dependent metric which obeys a simpler evolution equation with respect to the Fokker-Planck equation for the probability distribution. Although this is only a partial solution of the problem because we still need to solve for the metric, this method may be useful especially when the evolution equation for the metric has more convenient mathematical properties. We will not pursue this possibility here.

Chapter 7

Self-Organization of a Radiation Belt

In this chapter, as a direct application of the theory developed so far, we study a concrete example of self-organization pertaining to non-canonical Hamiltonian mechanics, the creation of a radiation belt. The relevant operator is a Poisson operator with a non-trivial kernel.

For the results presented here we refer the reader to [82, 83, 84, 85].

7.1 Motion of a Charged Particle in a Dipole Magnetic Field

Magnetospheres are commonly observed throughout the universe. This kind of magnetic field typically encircles planets and stars, as well as the Earth itself. The magnetosphere of the Earth exhibits, in an approximate fashion, rotational symmetry around the axis going across the magnetic poles. Usually, the magnetosphere is mathematically represented in terms of a dipole magnetic field:

$$\boldsymbol{B} = \nabla\psi \times \nabla\theta, \tag{7.1}$$

where ψ is the so called flux function (also stream function, or simply magnetic flux) and θ the toroidal angle around the axis of symmetry. If (r, z, θ) is a cylindrical coordinate system centered at the center of the Earth, the flux function has the following expression:

$$\psi = \frac{r^2}{(r^2 + z^2)^{3/2}}, \tag{7.2}$$

where normalized units were used. Notice that $\boldsymbol{B} \cdot \nabla\psi = \boldsymbol{B} \cdot \nabla\theta = 0$. Thus, if we interpret ψ and θ as magnetic coordinates, the level set (ψ=constant, θ=constant) identifies a magnetic field line (a curve with tangent vector always aligned with $\boldsymbol{B}$). It is worth to remark that, since for a general magnetic field we always have

$\nabla \cdot \boldsymbol{B} = 0$, the vector field $\boldsymbol{B}$ is properly described as a closed 2-form $\mathcal{B} = d\mathcal{A}$, with $\mathcal{B} = B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy$. Therefore, a direct consequence of Darboux's theorem for closed two forms 1.1 is that, in regions where the rank of $\mathcal{B}$ is constant, we can always find two functions ψ and θ such that locally one has $\mathcal{B} = d\psi \wedge d\theta$ and $\mathcal{A} = \psi d\theta$.

It is convenient to introduce a magnetic coordinate ℓ along the magnetic field $\boldsymbol{B}$ itself, i.e. we seek for a variable ℓ satisfying $\boldsymbol{B} = B\partial_\ell$. The coordinate ℓ represents the length of a field line as measured from the equatorial plane $z = 0$. Since, by definition, the magnetic flux ψ is constant along a field line, differentiation of equation (7.2) gives the condition satisfied by the infinitesimal displacements dr and dz when moving along the magnetic field $\boldsymbol{B}$ by an infinitesimal amount $d\ell$:

$$dz = \pm \frac{\frac{2}{3} r^{\frac{1}{3}} \psi^{-\frac{2}{3}} - r}{\sqrt{r^{\frac{4}{3}} \psi^{-\frac{2}{3}} - r^2}} dr. \tag{7.3}$$

Then, the length of the field line starting at the equatorial point $(r, z) = (\tilde{r}, 0)$ as a function of (r, z) is:

$$\begin{aligned}
\ell\,(r, z) &= \int_{\tilde{r}}^{r} d\ell \\
&= \int_{\tilde{r}}^{r} \sqrt{1 + \left(\frac{dz}{dr}\right)^2}\, dr \\
&= \frac{1}{2\psi} \left[\frac{1}{\sqrt{3}} \log\left(\sqrt{3}\sqrt{1 - (r\psi)^{\frac{2}{3}}} + \sqrt{4 - 3(r\psi)^{\frac{2}{3}}} \right) \right. \\
&\qquad \left. + \sqrt{1 - (r\psi)^{\frac{2}{3}}}\sqrt{4 - 3(r\psi)^{\frac{2}{3}}} \right].
\end{aligned} \tag{7.4}$$

Here we used the fact that along a field line $\psi = \psi\,(r, z) = \psi\,(\tilde{r}, 0) = \tilde{r}^{-1}$ and evaluated the integral by the change of variable $\xi = (r\psi)^{2/3}$. Figure 7.1 shows the contour plots of the functions B, ψ, ℓ in the (r, z) plane (poloidal section) defined by θ=constant.

A schematic representation of the magnetosphere of the Earth and the 'natural' magnetic coordinate system (ℓ, ψ, θ) is given in figure 7.2.

The set of magnetic coordinates (ℓ, ψ, θ) can be related to the Cartesian coordinate

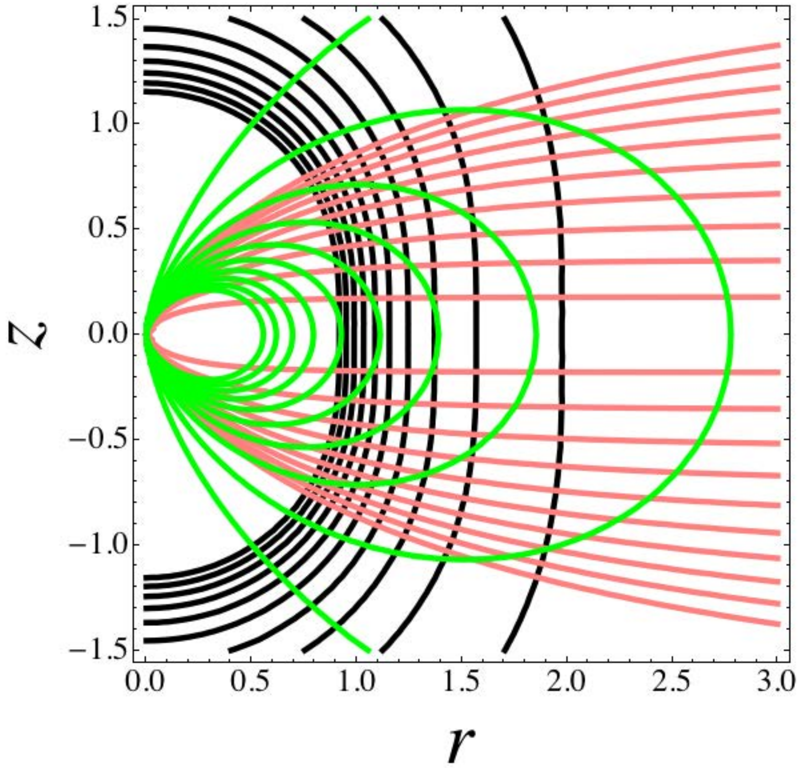

Figure 7.1: Contour plots of the magnetic field strength B (black lines), stream function ψ (green lines), and field line length ℓ (pink lines) in the poloidal section θ =constant.

system (x, y, z) according to:

$$\begin{aligned} d\ell \wedge d\psi \wedge d\theta &= \ell_i \psi_j \theta_k dx^i \wedge dx^j \wedge dx^k \\ &= \ell_i dx^i \wedge \sum_{j<k} (\psi_j \theta_k - \psi_k \theta_j)\, dx^j \wedge dx^k \\ &= (\nabla \ell \cdot \nabla \psi \times \nabla \theta)\, dx \wedge dy \wedge dz \\ &= (\nabla \ell \cdot \boldsymbol{B})\, dx \wedge dy \wedge dz \\ &= B\, (\nabla \ell \cdot \partial_\ell)\, dx \wedge dy \wedge dz \\ &= B\, dx \wedge dy \wedge dz. \end{aligned} \tag{7.5}$$

Recalling that $dx \wedge dy \wedge dz = r dr \wedge d\theta \wedge dz$, it also follows that $d\ell \wedge d\psi \wedge d\theta = B r dr \wedge d\theta \wedge dz$.

Magnetospheres are often populated by ‘clouds’ of charged particles that form a plasma. In the specific case of the Earth, these aggregations are called Van Allen radiation belts. The sharp density and temperature gradients characterizing the magnetospheric plasma represent a paradigmatic example of self-organization where an ordered structure is created and sustained in a process that apparently contradicts the entropy principle. Our objective in this chapter is to show that the mechanism

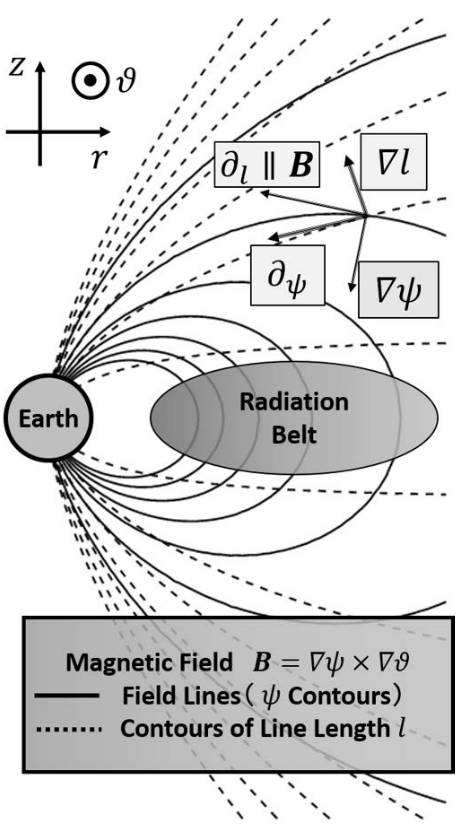

Figure 7.2: Schematic view of the Earth magnetosphere and the associated magnetic coordinate system (ℓ, ψ, θ).

behind the creation of a radiation belt is a type-I self-organization (see definition 6.2) on the metric induced by the magnetic coordinates. To see this, we need to determine the equations of motion of a charged particle in the new magnetic coordinate system (ℓ, ψ, θ).

Without entering into details, we recall that in a dipole magnetic field the motion of a charged particle is endowed with three adiabatic invariants which are preserved on distinct time scales typically separated by three orders of magnitude. The first and strongest invariant is the so called magnetic moment μ defined as the ratio between the kinetic energy $mv_c^2/2$ of the cyclotron gyration around the magnetic field lines and the field strength B:

$$\mu = \frac{mv_c^2}{2B}. \tag{7.6}$$

The phase of the cyclotron gyration is usually denoted by θ_c. The second adiabatic invariants is the bounce action $J_\parallel$ associated to the periodic motion along the magnetic field. In formulas:

$$J_{\parallel} = m \oint v_{\parallel} d\ell, \tag{7.7}$$

where the loop integral is carried along the bounce orbit and $v_{\parallel}$ is the velocity along a field line. The third and weakest adiabatic invariant is the flux function ψ which acts as the action variable associated to the periodic revolution around the symmetry axis with corresponding angle variable θ.

Due to the strength of the magnetic moment invariance, the motion of a magnetized particle can be accurately described by considering the dynamics of the geometrical center of the cyclotron gyration around the magnetic field (the so called guiding center). The resulting guiding center equations of motion read:

$$\dot{v}_{\parallel} = -\frac{1}{m} \left(\mu B + e\phi\right)_{\ell} + v_{\parallel} \boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}} \cdot \boldsymbol{k}, \tag{7.8a}$$

$$\boldsymbol{v} = \boldsymbol{v}_{\parallel} + \boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}} + \boldsymbol{v}_{\nabla B} + \boldsymbol{v}_{\boldsymbol{k}}, \tag{7.8b}$$

$$\dot{\mu} = 0, \tag{7.8c}$$

$$\dot{\theta}_c = \frac{eB}{m}. \tag{7.8d}$$

In these equations, e specifies the electric charge, ϕ the electric potential, $\boldsymbol{v}_{\parallel}$ the velocity along a magnetic field line, $\boldsymbol{k} = \partial_{\ell}\left(\boldsymbol{B}/B\right) = \partial_{\ell}\partial_{\ell} = \partial_{\ell}^2$ the curvature of the magnetic field, $\boldsymbol{v}$ the particle velocity, $\boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}}$, $\boldsymbol{v}_{\nabla B}$, and $\boldsymbol{v}_{\boldsymbol{k}}$ the $\boldsymbol{E}\times\boldsymbol{B}$, ∇B, and curvature drifts, and the subscript notation means derivation.

The $\boldsymbol{E}\times\boldsymbol{B}$ drift velocity $\boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}}$ plays a crucial role in the diffusion process at the basis of magnetospheric self-organization and has special mathematical properties that will serve us in the next chapters. It is therefore desirable to add some additional considerations on this drift velocity. The term 'drift' means that no acceleration is involved. In the specific case of the $\boldsymbol{E}\times\boldsymbol{B}$ drift, the associated velocity is obtained by requiring that the Lorentz force $\boldsymbol{F}_L$ vanishes:

$$\boldsymbol{F}_L = e\left(\boldsymbol{E} + \boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}} \times \boldsymbol{B}\right) = 0. \tag{7.9}$$

Noting that $\boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}}$ has to be orthogonal to the magnetic field, we can solve for $\boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}}$ and obtain:

$$\boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}} = \frac{\boldsymbol{E}\times\boldsymbol{B}}{B^2}. \tag{7.10}$$

The conservation of the first adiabatic invariant μ can be interpreted as a topological constraint affecting the 6-dimensional canonical phase space (x, y, z, p_x, p_y, p_z) of charged particle dynamics. To see this, let us 'separate' the topological constraint

μ from the phase space coordinates (x, y, z, p_x, p_y, p_z):

$$
\begin{aligned}
dx \wedge dy \wedge dz \wedge dp_x \wedge dp_y \wedge dp_z &= B^{-1} d\ell \wedge d\psi \wedge d\theta \wedge dp_\perp \wedge dp_\parallel \wedge dp_\theta \\
&= B^{-1} d\ell \wedge d\psi \wedge d\theta \wedge dmv_\perp \wedge dmv_\parallel \wedge d\left(mv_\theta + \frac{e\psi}{r}\right) \\
&= m^3 B^{-1} d\ell \wedge d\psi \wedge d\theta \wedge d\left(v_c \cos\theta_c + \boldsymbol{v}_d \cdot \partial_\perp\right) \\
&\quad \wedge dv_\parallel \wedge d\left(v_c \sin\theta_c + \boldsymbol{v}_d \cdot \frac{\partial_\theta}{|\partial_\theta|} + \frac{e\psi}{mr}\right) \\
&= -\frac{1}{2} m^3 B^{-1} d\ell \wedge d\psi \wedge d\theta \wedge dv_\parallel \wedge dv_c^2 \wedge d\theta_c \\
&= m^2 dv_\parallel \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu \wedge d\theta_c.
\end{aligned}
\tag{7.11}
$$

Here, we used equations (7.5) and (7.6), and defined the parallel moment $p_\parallel = mv_\parallel$, the perpendicular moment $p_\perp = mv_\perp = m(v_c cos\theta_c + \boldsymbol{v}_d \cdot \partial_\perp)$, and the toroidal moment $p_\theta = m(v_c sin\theta_c + \boldsymbol{v}_d \cdot \partial_\theta/|\partial_\theta|) + e\psi/r$, with $\boldsymbol{v}_d = \boldsymbol{v} - \boldsymbol{v}_c - \boldsymbol{v}_\parallel$ the drift velocity (assumed to be only a function of the spatial coordinates). In this notation $\partial_\perp = \nabla\psi/|\nabla\psi|$ denotes the unit normal vector to field lines and $\partial_\theta/|\partial_\theta|$ the unit vector in the toroidal direction.

Now we want to write the guiding center equations of motion (7.8) in the novel coordinate system $\left(v_\parallel, \ell, \psi, \theta, \mu, \theta_c\right)$. First, consider the rate of change in $v_\parallel$, equation (7.8a). The only term that needs to be expressed in the new coordinates is that involving the curvature $\boldsymbol{k} = \partial_\ell^2$ and the $\boldsymbol{E} \times \boldsymbol{B}$ drift velocity (7.10). Since we assume the magnetic field to be static, i.e. we take $\partial_t \boldsymbol{B} = \boldsymbol{0}$, we have $\boldsymbol{E} = -\nabla\phi$. This is reasonable as long as the background dipole magnetic field is much stronger than the self-induced magnetic field. Noting that due to θ-symmetry $\partial_\ell^2 \cdot \nabla\theta = 0$, and using $|\nabla\psi| = B/|\nabla\theta| = rB$, we have:

$$
v_\parallel \boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}} \cdot \boldsymbol{k} = v_\parallel \frac{\boldsymbol{B} \times \nabla\phi}{B^2} \partial_\ell^2 = v_\parallel \frac{\partial_l \times (\phi_\theta \nabla\theta)}{B} \cdot \partial_\ell^2 = -v_\parallel \phi_\theta \frac{\nabla\psi}{r^2 B^2} \cdot \partial_\ell^2. \tag{7.12}
$$

Now, define $q = -\partial_\ell \cdot \partial_\psi$ and observe that:

$$\begin{aligned}
q_\ell &= -\frac{\partial}{\partial \ell}\left(\partial_\ell \cdot \partial_\psi\right) \\
&= -\partial_\ell^2 \cdot \partial_\psi - \partial_\ell \cdot \partial_{\ell\psi} \\
&= -\partial_\ell^2 \cdot \partial_\psi - \frac{1}{2}\partial_\psi \left|\partial_\ell\right|^2 \\
&= -\partial_\ell^2 \cdot \partial_\psi \\
&= -\partial_\ell^2 \cdot \left(\frac{\partial_\psi \cdot \nabla\psi}{rB}\frac{\nabla\psi}{rB} + \left(\partial_\psi \cdot \partial_\ell\right)\partial_\ell\right) \\
&= -\partial_\ell^2 \cdot \left(\frac{\nabla\psi}{r^2B^2} - q\partial_\ell\right) \\
&= -\partial_\ell^2 \cdot \frac{\nabla\psi}{r^2B^2} + \frac{q}{2}\partial_\ell \left|\partial_\ell\right|^2 \\
&= -\partial_\ell^2 \cdot \frac{\nabla\psi}{r^2B^2}.
\end{aligned} \tag{7.13}$$

Putting this result in equation (7.12) one sees that:

$$v_\parallel \boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}} \cdot \boldsymbol{k} = -v_\parallel q_\ell \phi_\theta. \tag{7.14}$$

Now, consider the rate of change in ℓ:

$$\begin{aligned}
\dot{\ell} &= \nabla\ell \cdot \boldsymbol{v} \\
&= \nabla\ell \cdot \left(v_\parallel \partial_\ell + \boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}}\right) \\
&= v_\parallel + \phi_\theta \nabla\ell \cdot \left(\frac{\partial_\ell \times \nabla\theta}{B}\right) \\
&= v_\parallel - \phi_\theta \frac{\nabla\ell \cdot \nabla\psi}{r^2B^2}
\end{aligned} \tag{7.15}$$

Here we used the fact that, by symmetry, both the gradient and curvature drift $\boldsymbol{v}_{\nabla B}$ and $\boldsymbol{v}_{\boldsymbol{k}}$ are directed along $\nabla\theta$ and therefore do not affect $\dot{\ell}$. We also have:

$$\partial_\psi = \frac{\partial_\psi \cdot \nabla\psi}{r^2B^2}\nabla\psi - q\partial_\ell = \frac{\nabla\psi}{r^2B^2} - q\partial_\ell. \tag{7.16}$$

Then:

$$\nabla\ell \cdot \partial_\psi = \frac{\nabla\ell \cdot \nabla\psi}{r^2B^2} - q = 0. \tag{7.17}$$

Substituting this result in the previous equation, we arrive at:

$$\dot{\ell} = v_\parallel - q\phi_\theta. \tag{7.18}$$

We move on to the rate of change in ψ:

$$\dot{\psi} = \nabla\psi \cdot \boldsymbol{v} = \phi_\theta \nabla\psi \cdot \frac{\partial_\ell \times \nabla\theta}{B} = -\phi_\theta \frac{\nabla\psi \cdot \nabla\psi}{r^2B^2} = -\phi_\theta \tag{7.19}$$

Finally, the rate of change in θ is given by:

$$
\begin{aligned}
\dot{\theta} &= \nabla\theta \cdot \boldsymbol{v} \\
&= \nabla\theta \cdot \left(\frac{\partial_\ell \times \nabla\phi}{B} + \frac{\mu}{e} \frac{\boldsymbol{B} \times \nabla B}{B^2} + \frac{m v_\parallel^2}{e} \frac{\boldsymbol{B} \times \boldsymbol{k}}{B^2} \right) \\
&= \nabla\theta \cdot \frac{\partial_\ell \times (\phi_\psi \nabla\psi + \phi_\ell \nabla\ell)}{B} + \frac{\mu}{eB} \nabla\theta \cdot \partial_\ell \times \nabla B - \frac{m v_\parallel^2}{eB} \nabla\theta \cdot \partial_\ell^2 \times \partial_\ell \\
&= \nabla\theta \cdot \left(r^2 \phi_\psi \nabla\theta + B^{-1} \phi_\ell \partial_\ell \times q \nabla\psi \right) + \frac{\mu}{eB} \nabla B \cdot \frac{\nabla\psi}{r^2 B} + \frac{m v_\parallel^2}{eB} \partial_\ell^2 \cdot \frac{\nabla\psi}{r^2 B}.
\end{aligned}
\tag{7.20}
$$

Recalling equations (7.16) and (7.13) we obtain:

$$
\dot{\theta} = \phi_\psi + q\phi_\ell + \frac{\mu}{e} \left(B_\psi + q B_\ell \right) - \frac{m}{e} v_\parallel^2 q_\ell. \tag{7.21}
$$

Putting together these results, the guiding center equations of motion now read:

$$
\dot{v}_\parallel = -\frac{1}{m} \left(\mu B + e\phi \right)_\ell + v_\parallel q_\ell \phi_\theta, \tag{7.22a}
$$

$$
\dot{\ell} = v_\parallel - q\phi_\theta \tag{7.22b}
$$

$$
\dot{\psi} = -\phi_\theta, \tag{7.22c}
$$

$$
\dot{\theta} = \left(\partial_\psi + q\partial_\ell \right) \left(\frac{\mu}{e} B + \phi \right) - \frac{m}{e} v_\parallel^2 q_\ell, \tag{7.22d}
$$

$$
\dot{\mu} = 0, \tag{7.22e}
$$

$$
\dot{\theta}_c = \frac{eB}{m}. \tag{7.22f}
$$

7.2 The Poisson Operator of Guiding Center Dynamics in a Dipole Magnetic Field

Purpose of the present section is to show that system (7.22) is Hamiltonian. First, let us calculate the associated antisymmetric operator $\mathcal{J}$. Notice that, at this point, we do not ask $\mathcal{J}$ to satisfy the Jacobi identity. We only need to find the Hamiltonian H representing the energy of each charged particle. It is reasonable to expect H to be the sum of kinetic and potential energy:

$$
H = \frac{m}{2} \left(v_\parallel^2 + v_c^2 \right) + e\phi = \frac{m}{2} v_\parallel^2 + \mu B + e\phi. \tag{7.23}
$$

Observe that we did not include any term involving the drift velocities because they are derived by neglecting the particle mass m in the equations of motion perpendicular to the magnetic field and therefore they should not contribute to the kinetic energy. Now define the vector field:

$$
X_{gc} = \dot{v}_\parallel \partial_{v_\parallel} + \dot{\ell} \partial_\ell + \dot{\psi} \partial_\psi + \dot{\theta} \partial_\theta + \dot{\theta}_c \partial_{\theta_c}. \tag{7.24}
$$

The rate of change in H is then:

$$
\begin{aligned}
\frac{dH}{dt} &= \mathfrak{L}_{X_{gc}} H \\
&= i_{X_{gc}} dH \\
&= H_{v_\parallel} \dot{v}_\parallel + H_\ell \dot{\ell} + H_\psi \dot{\psi} + H_\theta \dot{\theta} \\
&= m v_\parallel \left[-\frac{1}{m} \left(\mu B + e\phi \right)_\ell + v_\parallel q_\ell \phi_\theta \right] + \left(\mu B + e\phi \right)_\ell \left(v_\parallel - q\phi_\theta \right) \\
&\quad - \left(\mu B + e\phi \right)_\psi \phi_\theta + e\phi_\theta \left[\left(\partial_\psi + q\partial_\ell \right) \left(\frac{\mu}{e} B + \phi \right) - \frac{m}{e} v_\parallel^2 q_\ell \right] \\
&= 0.
\end{aligned}
\tag{7.25}
$$

Therefore, H is the energy of the system. Since H is constant, we can find an antisymmetric operator $\mathcal{J}$ such that $X_{gc} = \mathcal{J}\left(dH\right)$. Solving for $\mathcal{J}$, one obtains:

$$
\mathcal{J} = \begin{bmatrix}
0 & -m^{-1} & 0 & e^{-1} v_\parallel q_\ell & 0 & 0 \\
m^{-1} & 0 & 0 & -e^{-1} q & 0 & 0 \\
0 & 0 & 0 & -e^{-1} & 0 & 0 \\
-e^{-1} v_\parallel q_\ell & e^{-1} q & e^{-1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -em^{-1} \\
0 & 0 & & 0 & em^{-1} & 0
\end{bmatrix}
\tag{7.26}
$$

Let us check whether this operator satisfies the Jacobi identity (1.13), i.e. whether $\mathcal{J}$ is a Poisson operator. First notice that the only non-constant components in $\mathcal{J}$ are $\mathcal{J}^{v_\parallel \theta} = e^{-1} v_\parallel q_\ell$ and $\mathcal{J}^{\ell\theta} = -e^{-1} q$. Secondly, remember that the component $\mathfrak{G}^{ijk} = \mathcal{J}^{im} \mathcal{J}^{jk}_m + \mathcal{J}^{jm} \mathcal{J}^{ki}_m + \mathcal{J}^{km} \mathcal{J}^{ij}_m$ of the Jacobiator $\mathfrak{G}\left(\mathcal{J}\right)$ is different from 0 only if all the indices (i, j, k) are different. Furthermore, any even permutation of the indices gives the same value, while odd permutations only result in a change of sign. For example $\mathfrak{G}^{ijk} = -\mathfrak{G}^{ikj}$. Finally, since all the components in $\mathcal{J}$ are independent of the pair (μ, θ_c) and the other variables $\left(v_\parallel, \ell, \psi, \theta\right)$ are not coupled with them in $\mathcal{J}$, we can neglect all the components of $\mathfrak{G}$ involving μ and θ_c. In light of these considerations one sees that the only non-trivial component of the Jacobiator $\mathfrak{G}$ is:

$$
\begin{aligned}
\mathfrak{G}\left(\mathcal{J}\right) &= \mathfrak{G}^{v_\parallel \ell \theta} \partial_{v_\parallel} \wedge \partial_\ell \wedge \partial_\theta \\
&= \left(\mathcal{J}^{v_\parallel m} \frac{\partial \mathcal{J}^{\ell\theta}}{\partial x^m} + \mathcal{J}^{\ell m} \frac{\partial \mathcal{J}^{\theta v_\parallel}}{\partial x^m} + \mathcal{J}^{\theta m} \frac{\partial \mathcal{J}^{v_\parallel \ell}}{\partial x^m} \right) \partial_{v_\parallel} \wedge \partial_\ell \wedge \partial_\theta \\
&= \left(-e^{-1} \mathcal{J}^{v_\parallel m} \frac{\partial q}{\partial x^m} - e^{-1} \mathcal{J}^{\ell m} \frac{\partial \left(v_\parallel q_\ell \right)}{\partial x^m} \right) \partial_{v_\parallel} \wedge \partial_\ell \wedge \partial_\theta \\
&= -e^{-1} \left(\mathcal{J}^{v_\parallel \ell} q_\ell + \mathcal{J}^{\ell v_\parallel} q_\ell \right) \partial_{v_\parallel} \wedge \partial_\ell \wedge \partial_\theta \\
&= 0.
\end{aligned}
\tag{7.27}
$$

We have thus shown that $\mathcal{J}$ is a Poisson operator. It is easy to see that all the columns of $\mathcal{J}$ are linearly independent vectors. Therefore, the Poisson operator $\mathcal{J}$ has maximum rank and is invertible with inverse ω such that $d\omega = 0$ and:

$$i_{X_{gc}}\omega = -dH. \tag{7.28}$$

Since we know both X_{gc} and H, we can solve the last equation with respect to ω and obtain:

$$\begin{aligned}
\omega &= m dv_\parallel \wedge d\ell + m v_\parallel q_\ell d\psi \wedge d\ell + e d\psi \wedge d\theta + m q d\psi \wedge dv_\parallel + \frac{m}{e} d\mu \wedge d\theta_c \\
&= m dv_\parallel \wedge d\ell + d\psi \wedge d\left(e\theta + m q v_\parallel\right) + \frac{m}{e} d\mu \wedge d\theta_c \\
&= d\left(m v_\parallel d\ell + \psi d\eta + \frac{m}{e}\mu d\theta_c\right),
\end{aligned} \tag{7.29}$$

where we defined $\eta = e\theta + mqv_\parallel$. It immediately follows that $\left(mv_\parallel, \ell\right)$, (ψ, η), and $\left(\mu, \frac{m}{e}\theta_c\right)$ are canonical pairs and that, in these coordinates, $\mathcal{J}$ is a symplectic operator, i.e.:

$$\mathcal{J} = \partial_{mv_\parallel} \wedge \partial_\ell + \partial_\psi \wedge \partial_\eta + \partial_\mu \wedge \partial_{\frac{m}{e}\theta_c}. \tag{7.30}$$

7.3 Reduction of Cyclotron Motion

As anticipated, the conservation of the first adiabatic invariant can be interpreted as a topological constraint affecting the canonical phase space of charged particle dynamics. Let us see how.

In light of the results of the last section, the canonical phase space is spanned by the magnetic coordinates $\left(v_\parallel, \ell, \psi, \theta, \mu, \theta_c\right)$. Note that these are not canonical variables, because the canonical set is $\left(mv_\parallel, \ell, \psi, \eta, \mu, \frac{m}{e}\theta_c\right)$. Nevertheless, we shall proceed with the magnetic coordinate system as it is easier to handle (the new variable η does not have a simple physical interpretation and its evaluation is more complicated) and because the Jacobian of the coordinate change is constant:

$$dv_\parallel \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu \wedge d\theta_c = m^{-2} d\left(mv_\parallel\right) \wedge d\ell \wedge d\psi \wedge d\eta \wedge d\mu \wedge d\left(\frac{m}{e}\theta_c\right). \tag{7.31}$$

Therefore, the volume element, which is the essential ingredient from the statistical mechanics standpoint, looks the same in both reference frames.

Recalling the definition 5.1 of topological constraint, we must 'reduce' the Poisson operator (7.26) so that $\theta = d\mu \in ker\left(\mathcal{J}\right)$. Typically, the frequency of the cyclotron gyration around the magnetic field can be considered as high enough with respect to all the characteristic frequencies determining dynamical change in the system. Thereby, the phase of the gyration θ_c is not a physically relevant quantity (on the time scale of interest the observables do not depend on θ_c) and can be integrated to give a factor 2π. Removing the constant 2π, the reduced volume element reads:

$$dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu. \tag{7.32}$$

The Poisson operator is accordingly reduced to the truncated form:

$$\mathcal{J} = \begin{bmatrix} 0 & -m^{-1} & 0 & e^{-1}v_{\parallel}q_{\ell} & 0 \\ m^{-1} & 0 & 0 & -e^{-1}q & 0 \\ 0 & 0 & 0 & -e^{-1} & 0 \\ -e^{-1}v_{\parallel}q_{\ell} & e^{-1}q & e^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{7.33}$$

It is now clear that the 1-form $\xi = d\mu$ belongs to the kernel of the truncated operator since $\mathcal{J}(d\mu) = 0$. Furthermore, the topological constraint ξ is clearly integrable, with integral μ=constant. One can verify that the truncated operator itself satisfies the Jacobi identity, and thus the magnetic moment μ is a Casimir invariant. Observe that, since the truncated operator is odd-dimensional, now there is no coordinate change by which one can obtain a canonical coordinate system.

On each Casimir leaf (the level set μ=constant) we can further truncate the Poisson operator and obtain the 4-dimensional Poisson operator:

$$\mathcal{J} = \begin{bmatrix} 0 & -m^{-1} & 0 & e^{-1}v_{\parallel}q_{\ell} \\ m^{-1} & 0 & 0 & -e^{-1}q \\ 0 & 0 & 0 & -e^{-1} \\ -e^{-1}v_{\parallel}q_{\ell} & e^{-1}q & e^{-1} & 0 \end{bmatrix} \tag{7.34}$$

This operator can be inverted to the symplectic 2-form:

$$\omega = d\left(mv_{\parallel} \wedge d\ell + \psi \wedge d\eta\right). \tag{7.35}$$

Thus, as prescribed by the Darboux's theorem 1.1, we can define a canonical Hamiltonian system on each symplectic submanifold corresponding to a level set of the Casimir invariant μ.

Note that, although the 4-dimensional canonical form (7.34) is useful to study the motion of a single particle, the operator that encloses all the relevant physical information is the 5-dimensional truncated operator (7.33). This is because even though μ is a dynamical constant, an ensemble of particles is characterized by a distribution of magnetic moments that determines the global behavior of the plasma.

7.4 Fokker-Planck Equation on a Foliated Phase Space

In this section we derive the Fokker-Planck equation of the diffusion mechanism responsible of magnetospheric self-organization. Such diffusion process is usually called inward (or up-hill) diffusion because of its peculiar property of creating density gradients. Before doing so, it is worth to make some additional considerations on the physical scenario we want to describe.

The separation of the gyro-phase θ_c was carried out on the basis of the assumption that the time and spatial scale of electromagnetic perturbations affecting the dynamics of a charged particle do not reach cyclotron motion. Whenever this hypothesis breaks down, a more careful treatment of guiding center dynamics may be necessary (see, for example, gyro-kinetic models [74, 141, 142]).

The plasma we are concerned with is charge neutral, that is the ensemble average of the electric field vanishes:

$$\langle \boldsymbol{E} \rangle = \int_\Omega f \boldsymbol{E} \, vol_c^6 = \boldsymbol{0}. \tag{7.36}$$

Here, f is the probability distribution on the canonical phase space $vol_c^6 = dx \wedge dy \wedge dz \wedge dp_x \wedge dp_y \wedge dp_z$ and Ω the domain occupied by the plasma. Physically, this condition can be achieved in a ion-electron plasma with a null total charge. Then, the guiding center equations of motion (7.22) describe the motion of the heavier ions, while the lighter electrons adjust their position accordingly to maintain charge neutrality (7.36). In the case of a single species plasma, or if one is interested in the detailed electromagnetic interaction between different species, the equations of motion of the electrons must be considered separately together with the self-induced electric and magnetic fields. For the purpose of the present study, this is not necessary and we will work under the hypothesis (7.36).

To fully exploit charge neutrality, we must determine the invariant measure of the system, which must exist (at least locally) because we have shown that guiding center dynamics (7.22) is Hamiltonian. Once we obtain the invariant measure, there are grounds for the ergodic hypothesis (remember that measure preservation is a necessary condition for the ergodic theorem 2.3 to hold) and we can exchange the ensemble average $\langle \phi \rangle$ with the time average:

$$\bar{\boldsymbol{E}} = \lim_{T \to \infty} \frac{1}{T} \int_0^T \boldsymbol{E} \, dt. \tag{7.37}$$

Using charge neutrality, we would have:

$$\langle \boldsymbol{E} \rangle = \bar{\boldsymbol{E}} = \boldsymbol{0}. \tag{7.38}$$

In light of the ergodic hypothesis (7.38), we can then represent each component of the electric field in terms of a random process with null time average. Taking the simplest Gaussian white noise $\boldsymbol{\Gamma}$ (remember definition 3.2):

$$\boldsymbol{E} = -\nabla\phi = \boldsymbol{\Gamma}. \tag{7.39}$$

Thanks to Liouville theorem 2.1 we know that the desired invariant measure is given by the volume element $dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu \wedge d\theta_c$ of equation (7.31) because the coordinate set $\left(mv_{\parallel}, \ell, \psi, \eta, \mu, \frac{m}{e}\theta_c\right)$ on the right-hand side is canonical. Observe that, the reduced volume element (7.32) as well as the further reduced 4-dimensional measure $dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta$ on each Casimir leaf μ =constant are also invariant measures with respect to the corresponding 5-dimensional and 4-dimensional dynamics. Therefore, equation (7.39) implies:

$$\boldsymbol{E} = \frac{m}{e} D_{\parallel}^{1/2} \Gamma_{\parallel} \nabla\ell + D_{\perp}^{1/2} \Gamma_{\perp} \nabla\theta + D_{\theta}^{1/2} \Gamma_{\theta} \nabla\psi. \tag{7.40}$$

Here we restored physical units, and $mD_{\parallel}^{1/2}\Gamma_{\parallel}/e = -\phi_{\ell}$, $D_{\perp}^{1/2}\Gamma_{\perp} = -\phi_{\theta}$, and $D_{\theta}^{1/2}\Gamma_{\theta} = -\phi_{\psi}$ are Gaussian white noises with $\Gamma dt = dW$. The parameters $D_{\parallel}$, $D_{\perp}$, and D_{θ} are constants (diffusion parameters) scaling the strength of perturbations. It is important to remark that the specific choice of the coordinate system where noise is white is arbitrary from the standpoint of statistical mechanics because the hypothesis of theorem 6.1 are all satisfied. Specifically, equation (7.40) corresponds to the choice $R^i_j = \delta^i_j$, that is the the gradient of the electrostatic potential ϕ is represented by Gaussian white noise in the coordinate system spanning the invariant measure.

Recalling the expression of the $\boldsymbol{E} \times \boldsymbol{B}$ drift velocity (7.10) and substituting equation (7.40), we have:

$$\boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}} = D_{\perp}^{1/2} \frac{\Gamma_{\perp}}{rB} \partial_{\perp} - \left(D_{\theta}^{1/2} \Gamma_{\theta} + \frac{m}{e} q D_{\parallel}^{1/2} \Gamma_{\parallel} \right) \partial_{\theta}. \tag{7.41}$$

Now, we can evaluate the displacements caused by the stochastic drift (7.41) along the perpendicular directions $\partial_{\perp} = \nabla\psi/|\nabla\psi|$ and $\partial_{\theta} = \partial_{\theta}/|\partial_{\theta}|$:

$$dX_{\theta} = -r \left(D_{\theta}^{1/2} dW_{\theta} + \frac{m}{e} q D_{\parallel}^{1/2} dW_{\parallel} \right), \tag{7.42a}$$

$$dX_{\perp} = \frac{D_{\perp}^{1/2}}{rB} dW_{\perp}. \tag{7.42b}$$

Remember that in this notation upper case letters are used to identify stochastic variables.

The next step involves the derivation of the stochastic differentials of the magnetic coordinates $\left(v_{\parallel}, \ell, \psi, \theta, \mu, \theta_c\right)$ by using equations (7.40) and (7.42) and by invoking the chain rule of stocahstic calculus (Ito's lemma 3.1) when necessary. The stochastic differential $dV_{\|}$ of the variable $v_{\|}$ is determined by a deterministic part (the first term in (7.22a)), the parallel component of $\boldsymbol{E}$ (the term proportional to $dW_{\|}$ in (7.40)), and the toroidal component of $\boldsymbol{E}$ (the term proportional to $dW_{\perp}$ in (7.40)):

$$dV_{\|} = -\left(\frac{\mu}{m}B_{\ell} + \gamma v_{\|} - \mathfrak{C}_{v_{\|}}\right) dt + D_{\|}^{1/2} dW_{\|} - D_{\perp}^{1/2} v_{\|} q_{\ell} dW_{\perp}. \tag{7.43}$$

In the above expression we introduced the friction force $\gamma v_{\|} - \mathfrak{C}_{v_{\|}}$ which is needed to preserve total energy when the system is closed. This term corresponds to the friction force of equation (6.4) when the parameter defining the stochastic integral is $\alpha = 1/2$.

The stochastic differential dL of the variable ℓ is determined by a deterministic part (the first term in (7.22b)) and additional terms caused by the displacement $dX_{\perp}$ of equation (7.42b). In this case dL is obtained by application of Ito's lemma, because ℓ is implicitly affected by the change $dX_{\perp}$ through the change of variables $dX_{\perp} \to dL$. We have:

$$\begin{aligned} dL &= v_{\|}dt + \left(\frac{1}{2} - \alpha\right) \frac{D_{\perp}}{r^2B^2} \frac{\partial^2 \ell}{\partial x_{\perp}^2} + \frac{D_{\perp}^{1/2}}{rB} \frac{\partial \ell}{\partial x_{\perp}} dW_{\perp} \\ &= v_{\|}dt + \left(\frac{1}{2} - \alpha\right) \frac{D_{\perp}}{r^2B^2} \left[\frac{\partial \psi}{\partial x_{\perp}} \frac{\partial}{\partial \psi} + \frac{\partial \ell}{\partial x_{\perp}} \frac{\partial}{\partial \ell}\right] \frac{\nabla \ell \cdot \nabla \psi}{rB} + D_{\perp}^{1/2} q dW_{\perp} \\ &= v_{\|}dt + \left(\frac{1}{2} - \alpha\right) \frac{D_{\perp}}{rB} \left[\frac{\partial}{\partial \psi} + q \frac{\partial}{\partial \ell}\right] rBq + D_{\perp}^{1/2} q dW_{\perp}. \end{aligned} \tag{7.44}$$

Here, we used the fact that $\partial_{\perp} = \partial/\partial x_{\perp} = \nabla \psi / |\nabla \psi| = \nabla \psi / rB$. Thus,

$$dL = \left\{v_{\|} + \mathfrak{C}_{\ell} + D_{\perp}\left(\frac{1}{2} - \alpha\right) \left[(\partial_{\psi} + q\partial_{\ell}) q + q (\partial_{\psi} + q\partial_{\ell}) \log (rB)\right]\right\} dt + qD_{\perp}^{1/2} dW_{\perp}. \tag{7.45}$$

Again, the friction term $\mathfrak{C}_{\ell}$ (corresponding to (6.4) when $\alpha = 1/2$) was added.

The stochastic differential $d\Psi$ of the variable ψ is calculated in the same way: $d\Psi$ is determined by $dX_{\perp}$ through the change of variables $dX_{\perp} \to d\Psi$. We have:

$$\begin{aligned} d\Psi &= \left(\frac{1}{2} - \alpha\right) \frac{D_{\perp}}{r^2B^2} \frac{\partial^2 \psi}{\partial x_{\perp}^2} + \frac{D_{\perp}^{1/2}}{rB} \frac{\partial \psi}{\partial x_{\perp}} dW_{\perp} \\ &= \left(\frac{1}{2} - \alpha\right) \frac{D_{\perp}}{rB} \left(\frac{\partial}{\partial \psi} + q \frac{\partial}{\partial \ell}\right) rB + D_{\perp}^{1/2} dW_{\perp}. \end{aligned} \tag{7.46}$$

Therefore:

$$d\Psi = \left[D_{\perp}\left(\frac{1}{2} - \alpha\right) (\partial_{\psi} + q\partial_{\ell}) \log (rB) + \mathfrak{C}_{\psi}\right] dt + D_{\perp}^{1/2} dW_{\perp}. \tag{7.47}$$

The friction term $\mathfrak{C}_\psi$ (corresponding to (6.4) when $\alpha = 1/2$) was also added.

With the same procedure as above, one can evaluate the stochastic differential $d\Theta$ of the variable θ. This time $d\Theta$ is determined by dX_θ. Observing that $x_\theta = r\theta$, Ito's lemma leads to:

$$d\Theta = \left[\frac{\mu}{e}\left(\partial_\psi + q\partial_\ell\right)B + \mathfrak{C}_\theta - \frac{m}{e}v_\parallel^2 q_\ell\right]dt - D_\theta^{1/2}dW_\theta - \frac{m}{e}qD_\parallel^{1/2}dW_\parallel. \quad (7.48)$$

Here, the friction term $\mathfrak{C}_\theta$ (corresponding to (6.4) when $\alpha = 1/2$) was added.

Finally, the equations of motion (7.22e) and (7.22f) are not affected by the random electric field (7.40) and are therefore unchanged.

At this point we invoke (3.13) to obtain the Fokker-Planck equation of the system of stochastic differential equations (7.43), (7.45), (7.47), and (7.48) for the probability distribution f on the reduced space $dv_\parallel \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu$ (see (7.32)). The deterministic current $\boldsymbol{F}$ appearing in equation (3.7) is given by:

$$\boldsymbol{F} = \begin{bmatrix} -\left(\frac{\mu}{m}B_\ell + \gamma v_\parallel - \mathfrak{C}_{v_\parallel}\right) \\ v_\parallel + \mathfrak{C}_\ell + D_\perp\left(\frac{1}{2} - \alpha\right)\left[\left(\partial_\psi + q\partial_\ell\right)q + q\left(\partial_\psi + q\partial_\ell\right)\log\left(rB\right)\right] \\ D_\perp\left(\frac{1}{2} - \alpha\right)\left(\partial_\psi + q\partial_\ell\right)\log\left(rB\right) + \mathfrak{C}_\psi \\ \frac{\mu}{e}\left(\partial_\psi + q\partial_\ell\right)B + \mathfrak{C}_\theta - \frac{m}{e}v_\parallel^2 q_\ell \\ 0 \end{bmatrix}. \quad (7.49)$$

Defining the 3-dimensional Wiener process $\boldsymbol{W}$ such that:

$$d\boldsymbol{W} = \begin{bmatrix} dW_\parallel \\ dW_\perp \\ dW_\theta \end{bmatrix}, \quad (7.50)$$

the matrix G appearing in the system of stochastic differential equations (3.7) can be calculated as:

$$G = \begin{bmatrix} D_\parallel^{1/2} & -D_\perp v_\parallel q_\ell & 0 \\ 0 & qD_\perp^{1/2} & 0 \\ 0 & D_\perp^{1/2} & 0 \\ -\frac{m}{e}D_\parallel^{1/2}q & 0 & -D_\theta^{1/2} \\ 0 & 0 & 0 \end{bmatrix}. \quad (7.51)$$

Note that in the expressions of $\boldsymbol{F}$ and G the lines are ordered according to the variables $\left(v_\parallel, \ell, \psi, \theta, \mu\right)$. The columns of the matrix G correspond instead to the lines of the vector $d\boldsymbol{W}$. The resulting Fokker-Planck equation is:

$$
\begin{aligned}
\frac{\partial f}{\partial t} = & -\frac{\partial}{\partial \ell}\left\{v_\parallel + \left(\frac{1}{2}-\alpha\right) D_\perp \left[\left(\partial_\psi + q\partial_\ell\right) q + q\left(\partial_\psi + q\partial_\ell\right)\log\left(rB\right)\right] + \mathfrak{C}_\ell\right\} f \\
& + \frac{\partial}{\partial v_\parallel}\left(\frac{\mu}{m}B_\ell + \gamma v_\parallel - \mathfrak{C}_{v_\parallel}\right) f - \frac{\partial}{\partial \theta}\left[\frac{\mu}{e}\left(\partial_\psi + q\partial_\ell\right) B + \mathfrak{C}_\theta - \frac{m}{e}v_\parallel^2 q_\ell\right] f \\
& - \frac{\partial}{\partial \psi}\left[D_\perp\left(\frac{1}{2}-\alpha\right)\left(\partial_\psi + q\partial_\ell\right)\log\left(rB\right) + \mathfrak{C}_\psi\right] f + \frac{1}{2}D_\perp\frac{\partial^2}{\partial \ell^2}q^2 f \\
& + \frac{1}{2}D_\parallel\frac{\partial^2 f}{\partial v_\parallel^2} + \frac{1}{2}D_\theta\frac{\partial^2 f}{\partial \theta^2} - \frac{m}{e}D_\parallel\frac{\partial^2}{\partial v_\parallel \partial\theta}qf + \frac{m^2}{2e^2}D_\parallel\frac{\partial^2}{\partial\theta^2}q^2 f + \frac{1}{2}D_\perp\frac{\partial^2 f}{\partial \psi^2} \\
& + D_\perp\frac{\partial^2}{\partial \ell\partial\psi}qf - \alpha D_\perp\frac{\partial}{\partial \ell}\left[\left(\partial_\psi + q\partial_\ell\right) q\right] f - D_\perp\frac{\partial^2}{\partial\ell\partial v_\parallel}v_\parallel q q_\ell f \\
& + \frac{1}{2}D_\perp\frac{\partial^2}{\partial v_\parallel^2}\left(v_\parallel q_\ell\right)^2 f - D_\perp\frac{\partial^2}{\partial\psi\partial v_\parallel}v_\parallel q_\ell f + \alpha D_\perp\frac{\partial}{\partial v_\parallel}\left[v_\parallel\left(\partial_\psi + q\partial_\ell\right) q_\ell - v_\parallel q_\ell^2\right] f
\end{aligned}
\tag{7.52}
$$

This is the Fokker-Planck equation of inward diffusion. Although the physical meaning of this expression is not immediate, notice that the first four terms are simply the divergence of the deterministic flow associated to the underlying stochastic differential equations. The remaining terms represent the diffusion operator of the system. Equation (7.52) acquires a simpler form if we choose $\alpha = 1/2$ (i.e. the white noise can be considered as the limit of a continuous perturbation), neglect the geometrical effect $q = -\partial_\ell \cdot \partial_\psi$ due to the non-orthogonality of tangent vectors in the magnetic coordinate system, omit the friction force, and assume toroidal symmetry $\partial_\theta = 0$. Then, (7.52) reduces to:

$$
\frac{\partial f}{\partial t} = -v_\parallel\frac{\partial f}{\partial \ell} + \frac{\mu}{m}B_\ell\frac{\partial f}{\partial v_\parallel} + \frac{1}{2}D_\parallel\frac{\partial^2 f}{\partial v_\parallel^2} + \frac{1}{2}D_\perp\frac{\partial^2 f}{\partial\psi^2}. \tag{7.53}
$$

Notice that the key feature of inward diffusion is the last term appearing in equation (7.53): the probability distribution f is progressively flattened with respect to the variable ψ.

7.5 Self-Organized Confinement in Magnetosphere

The last section of this chapter is dedicated to the study of the physical implications of the Fokker-Planck equation (7.52) for inward diffusion in a magnetosphere.

First, we define the *proper density* u as:

$$
u = \int_{\mathcal{V}} f\, d\mu \wedge dv_\parallel. \tag{7.54}
$$

Here, $\mathcal{V}$ is the plasma domain in the coordinates $\left(\mu, v_{\parallel}\right)$. Since the probability contained in each volume element is a scalar, we have the equivalence:

$$u\, d\ell \wedge d\psi \wedge d\theta = uB\, dx \wedge dy \wedge dz = \rho\, dx \wedge dy \wedge dz, \tag{7.55}$$

where we used equation (7.5) to perform the change of coordinates $(\ell, \psi, \theta) \rightarrow (x, y, z)$ and defined the *laboratory density* ρ such that:

$$\rho = uB. \tag{7.56}$$

From (7.54) we see that if f_{ψ} becomes progressively smaller as a consequence of the diffusion process described by equation (7.52), so will u_{ψ}. Then, in light of equation (7.56), ρ_{ψ} will approach the value $uB_{\psi} \neq 0$ whenever $u \neq 0$ and $B_{\psi} \neq 0$. In other words, inward diffusion flattens the gradient of the proper density u but steepens the profile of ρ. Physically, this is a direct consequence of the fact that the separation of the cyclotron motion induces the proper metric (7.32) which differs by the Jacobian factor B from the Cartesian metric on the Casimir leaf μ=constant:

$$dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu = B\, dv_{\parallel} \wedge dx \wedge dy \wedge dz \wedge d\mu. \tag{7.57}$$

Next, we introduce the definitions of parallel and perpendicular temperatures. The plasma parallel temperature is defined as the average in velocity space of the kinetic energy associated to the motion along the magnetic field $\boldsymbol{B}$:

$$T_{\parallel} = \frac{m}{2} u^{-1} \int_{\mathcal{V}} v_{\parallel}^2 f\, d\mu \wedge dv_{\parallel}. \tag{7.58}$$

Similarly, the plasma perpendicular temperature is defined as the average in velocity space of the kinetic energy associated to cyclotron motion:

$$T_{\perp} = B u^{-1} \int_{\mathcal{V}} \mu f\, d\mu \wedge dv_{\parallel} \tag{7.59}$$

The ratio:

$$\mathfrak{a} = \frac{T_{\perp}}{T_{\parallel}} = \frac{2B}{m} \frac{\int_{\mathcal{V}} \mu f\, d\mu \wedge dv_{\parallel}}{\int_{\mathcal{V}} v_{\parallel}^2 f\, d\mu \wedge dv_{\parallel}} \tag{7.60}$$

is called temperature anisotropy.

In order to understand the thermodynamic properties of the system, we need an entropy measure. Since the hypothesis of theorem 6.1 are verified with f the probability distribution on the invariant measure (7.32), we know that the proper entropy measure is given by the functional of definition 6.1:

$$\Sigma = - \int_{\Omega} f \log f\, dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu, \tag{7.61}$$

Here Ω is the plasma domain. The rate of change in entropy is thus:

$$\begin{aligned}
\frac{d\Sigma}{dt} &= -\int_{\Omega} \frac{\partial f}{\partial t} \left(1 + \log f\right) dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu \\
&= \int_{\Omega} \frac{\partial \left(f Z^i\right)}{\partial x^i} \left(1 + \log f\right) dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu \\
&= \int_{\Omega} f \frac{\partial Z^i}{\partial x^i} dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu + \int_{\partial\Omega} f \log f \, Z^i N_i dS \\
&= \sigma - \Phi
\end{aligned} \tag{7.62}$$

where σ the *entropy production* of the system:

$$\sigma = \int_{\Omega} f \frac{\partial Z^i}{\partial x^i} dv_{\parallel} \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu, \tag{7.63}$$

and Φ is the entropy flux across the boundary $\partial\Omega$:

$$\Phi = -\int_{\partial\Omega} f \log f \, Z^i N_i dS, \tag{7.64}$$

with dS the surface element and N the unit normal on $\partial\Omega$, and Z the Fokker-Planck velocity of the system. From equation (7.52) one sees that:

$$\begin{aligned}
Z^{v\,\parallel} = &-\left(\frac{\mu}{m} B_\ell + \gamma v_{\parallel} - \mathfrak{C}_{v_{\parallel}}\right) - \frac{1}{2} D_{\parallel} \frac{\partial \log f}{\partial v_{\parallel}} + \frac{m}{2e} D_{\parallel} q \frac{\partial \log f}{\partial \theta} - \frac{1}{2} D_{\perp} f^{-1} \frac{\partial}{\partial v_{\parallel}} v_{\parallel}^2 q_\ell^2 f \\
&+ \frac{1}{2} D_{\perp} f^{-1} \frac{\partial}{\partial \ell} v_{\parallel} q q_\ell f + \frac{1}{2} D_{\perp} f^{-1} \frac{\partial}{\partial \psi} v_{\parallel} q_\ell f + \alpha D_{\perp} v_{\parallel} \left[q_\ell^2 - \left(\partial_\psi + q \partial_\ell\right) q_\ell\right],
\end{aligned} \tag{7.65}$$

$$\begin{aligned}
Z^{\ell} = &\; v_{\parallel} + \mathfrak{C}_\ell + D_{\perp} \left(\frac{1}{2} - \alpha\right) \left[\left(\partial_\psi + q \partial_\ell\right) q + q \left(\partial_\psi + q \partial_\ell\right) \log\left(rB\right)\right] \\
&- \frac{1}{2} D_{\perp} f^{-1} \frac{\partial}{\partial \ell} q^2 f - \frac{1}{2} D_{\perp} f^{-1} \frac{\partial}{\partial \psi} q f + \alpha D_{\perp} \left(\partial_\psi + q \partial_\ell\right) q + \frac{1}{2} D_{\perp} f^{-1} \frac{\partial}{\partial v_{\parallel}} v_{\parallel} q q_\ell f,
\end{aligned} \tag{7.66}$$

$$\begin{aligned}
Z^{\psi} = &\; D_{\perp} \left(\frac{1}{2} - \alpha\right) \left(\partial_\psi + q \partial_\ell\right) \log\left(rB\right) + \mathfrak{C}_\psi \\
&- \frac{1}{2} D_{\perp} \frac{\partial \log f}{\partial \psi} - \frac{1}{2} D_{\perp} f^{-1} \frac{\partial}{\partial \ell} q f + \frac{1}{2} D_{\perp} f^{-1} \frac{\partial}{\partial v_{\parallel}} v_{\parallel} q_\ell f,
\end{aligned} \tag{7.67}$$

$$\begin{aligned}
Z^{\theta} = &\; \frac{\mu}{e} \left(\partial_\psi + q \partial_\ell\right) B + \mathfrak{C}_\theta - \frac{m}{e} v_{\parallel}^2 q_\ell - \frac{1}{2} D_\theta \frac{\partial \log f}{\partial \theta} \\
&+ \frac{m}{2e} D_{\parallel} f^{-1} \frac{\partial}{\partial v_{\parallel}} q f - \frac{m^2}{2e^2} D_{\parallel} f^{-1} \frac{\partial}{\partial \theta} q^2 f.
\end{aligned} \tag{7.68}$$

Then:

$$\sigma=\int_\Omega f\left\{\frac{\partial}{\partial\ell}\mathfrak{C}_\ell+\left(\frac{1}{2}-\alpha\right)D_\perp\frac{\partial}{\partial\ell}\left[\left(\partial_\psi+q\partial_\ell\right)q+q\left(\partial_\psi+q\partial_\ell\right)\log\left(rB\right)\right]\right.$$
$$-\frac{1}{2}D_\perp\frac{\partial}{\partial\ell}\left[\frac{1}{f}\left(\frac{\partial}{\partial\ell}q^2f+\frac{\partial}{\partial\psi}qf-qq_\ell\frac{\partial}{\partial v_\parallel}v_\parallel f\right)\right]+\alpha D_\perp\frac{\partial}{\partial\ell}\left[\left(q\partial_\ell+\partial_\psi\right)q\right]-\gamma$$
$$+\frac{\partial}{\partial v_\parallel}\mathfrak{C}_{v_\parallel}-\frac{D_\parallel}{2}\frac{\partial^2\log f}{\partial v_\parallel^2}+\frac{D_\perp}{2}\frac{\partial}{\partial v_\parallel}\left[\frac{1}{f}\left(-q_\ell^2\frac{\partial}{\partial v_\parallel}v_\parallel^2f+v_\parallel\frac{\partial}{\partial\ell}qq_\ell f+v_\parallel\frac{\partial}{\partial\psi}q_\ell f\right)\right]$$
$$+\frac{D_\parallel m}{2e}\frac{\partial}{\partial v_\parallel}\left(q\frac{\partial\log f}{\partial\theta}\right)-\alpha D_\perp\left[\left(\partial_\psi+q\partial_\ell\right)q_\ell-q_\ell^2\right]+\frac{\partial}{\partial\theta}\mathfrak{C}_\theta$$
$$+\frac{D_\parallel m}{2e}\frac{\partial}{\partial\theta}\left(q\frac{\partial\log f}{\partial v_\parallel}-\frac{m}{e}q^2\frac{\partial\log f}{\partial\theta}\right)-\frac{D_\theta}{2}\frac{\partial^2\log f}{\partial\theta^2}+\frac{\partial}{\partial\psi}\mathfrak{C}_\psi$$
$$+D_\perp\left(\frac{1}{2}-\alpha\right)\frac{\partial}{\partial\psi}\left[\left(\partial_\psi+q\partial_\ell\right)\log\left(rB\right)\right]+\frac{D_\perp}{2}\frac{\partial}{\partial\psi}\left[\frac{1}{f}\left(q_\ell\frac{\partial}{\partial v_\parallel}v_\parallel f-\frac{\partial}{\partial\ell}qf\right)\right]$$
$$\left.-\frac{D_\perp}{2}\frac{\partial^2\log f}{\partial\psi^2}\right\}dv_\parallel\wedge d\ell\wedge d\psi\wedge d\theta\wedge d\mu \tag{7.69}$$

Again, if we assume toroidal symmetry $\partial_\theta=0$, neglect the geometric factor q and friction, and take $\alpha=1/2$, the entropy production rate σ reduces to:

$$\sigma=-\frac{1}{2}\int_\Omega f\left[\left(D_\parallel\frac{\partial^2}{\partial v_\parallel^2}+D_\perp\frac{\partial^2}{\partial\psi^2}\right)\log f\right]dv_\parallel\wedge d\ell\wedge d\psi\wedge d\theta\wedge d\mu \tag{7.70}$$

In the following, we want to compare the behavior of Σ with the standard (and physically wrong) information measure $\tilde{S}$ in the Cartesian coordinate system $dv_\parallel\wedge dx\wedge dy\wedge dz\wedge d\mu$ with probability distribution $P=fB$:

$$\tilde{S}=-\int_\Omega f\log\left(fB\right)dv_\parallel\wedge d\ell\wedge d\psi\wedge d\theta\wedge d\mu. \tag{7.71}$$

Evidently, Σ and $\tilde{S}$ are related by:

$$\Sigma=\tilde{S}+\left\langle\log B\right\rangle. \tag{7.72}$$

Figures 7.3, 7.4, 7.5, and 7.6 show the results of the numerical simulation of the Fokker-Planck equation (7.52). The specific conditions for this simulation are listed below.

- Computational domain and boundary conditions: assuming toroidal symmetry $\partial_\theta=0$, the computational domain lies on a surface $\theta=$constant and is determined by the intersection between a level set of ψ (a field line) and a level set of B (see figure 7.3 to visualize the boundary). The range of the variables μ and $v_\parallel$ is taken to be such that it covers at least three times the standard deviation $\sqrt{1/(\beta m)} = \sqrt{k_B T/m}$ of the Maxwell-Boltzmann distribution chosen as initial condition for the simulation (see paragraph below for details on initial conditions). In this notation k_B is the Boltzmann constant, T the plasma temperature, and $\beta^{-1} = k_B T$. On the boundaries, we set Dirichlet boundary conditions for the probability distribution f. Physically, these boundary conditions reflect the fact that charged particles are lost once they hit the atmosphere of the planet or when they escape its magnetosphere. For technical reasons, Dirichlet boundary conditions are also used for the variables μ and $v_\parallel$. This is not a problem because their range is large enough so that particle loss at the corresponding boundaries does not affect the relevant physics. In practice, we limited the simulation to the upper half domain defined by $z \geq 0$ and assumed reflection symmetry around the $z = 0$ axis. This is physically consistent because the dipole magnetic field exhibits such symmetry.

• Initial conditions: the initial condition is a Maxwell-Boltzmann distribution:

$$\begin{aligned} f\left(t=0\right) &= 4\pi^2 \left(\frac{\beta m}{2\pi}\right)^{3/2} \exp\left\{-\beta\left(\frac{m}{2}v^2\right)\right\} \\ &= 4\pi^2 \left(\frac{\beta m}{2\pi}\right)^{3/2} \exp\left\{-\beta\left(\frac{m}{2}v_\parallel^2 + \mu B\right)\right\}. \end{aligned} \tag{7.73}$$

Here, the factor $4\pi^2$ is the result of the integration of the dummy variables θ and θ_c. We consider a plasma with initial temperature $T = 10\,eV$.

• Physical setting: in this simulation we choose $\alpha = 0$ for the definition of the stochastic integral. Then, the Wiener processes appearing in the stochastic differential equations can be thought as the limit of a discontinuous 'step' process. Physically, this means that the random electric field (7.40) perturbs the orbit of a particle for very short time intervals followed by long periods of unperturbed motion (see [82] on this point). As already noted, we also assume that the probability distribution f does not depend on the toroidal angle θ because of the symmetry of the dipole magnetic field $\boldsymbol{B}$ and set $f_\theta = 0$. Furthermore we neglect the friction terms $\mathfrak{C}_i$, with $i = v_\parallel, \ell, \psi$ since we assume that the source of the fluctuations (7.40) is external to the system and therefore we do not need to preserve total energy. This choice reflects the fact that a radiation belt is, typically, an open system where there is a constant and external energy supply. The diffusion parameters $D_\perp$ and $D_\parallel$ are chosen so that $\boldsymbol{E}_\perp \sim 100\,V/m$ and $\boldsymbol{E}_\parallel \sim 1\,V/m$ (since the electrons can move almost freely along the magnetic field, any potential hole in the direction of $\boldsymbol{B}$

is immediately adjusted and therefore we expect a sensibly smaller parallel electric field).

Consider now figure 7.3: in this picture the spatial profiles of density ρ, temperature anisotropy $\mathfrak{a}$, parallel temperature $T_{\|}$, and perpendicular temperature $T_{\perp}$ are reported for three distinct time shots $t = 5000$, $t = 7500$, and $t = 10000$. Here, time is given in arbitrary units. The contours appearing in the plots have the same meaning as those in figures 7.1 and 7.2.

Observe that due to the inward diffusion of particles, the density ρ becomes progressively peaked. At the same time, the temperature anisotropy $\mathfrak{a}$ grows at the equator where most of the particles accumulate. This is a direct consequence of the increase in the kinetic energy μB stored in the cyclotron gyration (since μ is constant and particles tend to move toward regions of higher B, $T_{\perp}$ has to increase). The opposite happens for $T_{\|}$: in this case particles with a high initial $v_{\|}$ are progressively lost at the boundaries with a consequent decrease in $T_{\|}$.

Figure 7.4 shows the time evolution of the spatially averaged parallel and perpendicular temperatures (i.e. the ensemble averages $\left\langle \frac{m}{2} v_{\|}^2 \right\rangle$ and $\langle \mu B \rangle$) for two different choices of the diffusion parameter $D_{\perp}$ scaling the strength of the inward diffusion across the magnetic field. Notice how a stronger inward diffusion causes a faster and more pronounced heating of $T_{\perp}$ with a concomitant loss of $T_{\|}$.

Figure 7.5 shows the radial profiles of parallel and perpendicular temperatures $T_{\|}$ and $T_{\perp}$ along the equator of the dipole magnetic field $z = 0$. The anisotropic heating associated to inward diffusion is evident.

The behavior of the entropy measures Σ and $\tilde{S}$ is given in figure 7.6. The thermodynamically consistent functional Σ is correctly maximized and the associated entropy production σ is positive. This is consistent with the progressive flattening, caused by inward diffusion, of the distribution f (as well as the proper density u) on the invariant measure (7.32) of the system. The wrong laboratory entropy measure $\tilde{S}$ is instead minimized with corresponding peaking of the laboratory density ρ. The creation of the density gradient is thus explained in terms of the Jacobian B of the coordinate change sending the invariant measure $\left(v_{\|}, \ell, \psi, \theta, \mu\right)$ to the Cartesian coordinate system $\left(v_{\|}, x, y, z, \mu\right)$.

We conclude this section with some considerations on the nature of magnetospheric self-organization. For this purpose it is useful to change the physical setting by putting $\alpha = 1/2$ and considering the friction term as expressed by equation (6.4). Then, in light of theorem 6.1 and recalling that the operator of the system is a Poisson operator with Casimir invariant μ, the solution to the Fokker-Planck equation for the distribution f on the invariant measure (7.32) is of the type:

$$f = f_0 \exp\left\{-\beta H - \gamma_\mu \mu\right\}, \tag{7.74}$$

where f_0 and γ_μ are constants. Therefore, the creation of a radiation belt is a type-I self-organization driven by the Casimir invariant μ. Indeed on the invariant measure (7.32) the cocurrent of the Poisson operator $\mathcal{O}^5 = d\mathcal{J}^4 = di_{\mathcal{J}} dv_\parallel \wedge d\ell \wedge d\psi \wedge d\theta \wedge d\mu$ vanishes. When observed in the Cartesian coordinates, a spurious cocurrent $\tilde{\mathcal{O}}^5 = d\tilde{\mathcal{J}}^4 = di_{\mathcal{J}} dv_\parallel \wedge dx \wedge dy \wedge dz \wedge d\mu$ appears and we see a combination of type-I and (spurious) type-II self-organization.

By integrating f over μ and $v_\parallel$, the laboratory density ρ can be evaluated explicitly:

$$\rho = Bu = B \int_0^\infty d\mu \int_{-\infty}^{+\infty} f dv_\parallel = \rho_0 \frac{B}{\gamma_\mu + \beta B}, \tag{7.75}$$

where ρ_0 is a constant.

Finally, while we have seen that the first adiabatic invariant μ is a constant of motion and that the third adiabatic invariant ψ obeys the stochastic differential equation 7.47, we did not discuss the behavior of the second adiabatic invariant $J_\parallel$ defined in equation (7.7). Without entering into details, it can be shown that the rate of change in bounce action is given by:

$$\frac{dJ_\parallel}{dt} = T\left[\phi_\theta \langle (\mu B + e\phi)_\psi \rangle - \langle \phi_\theta \rangle (\mu B + e\phi)_\psi - q\langle \phi_\theta \rangle (\mu B + e\phi)_\ell + \langle \phi_\theta \rangle m v_\parallel^2 q_\ell \right] \tag{7.76}$$

In this expression, $\langle\ \rangle / T$ stands for bounce orbit average, with $T = \oint ds/v_\parallel = 2\pi/\omega_b$ the period of the bounce oscillation. From equation (7.76) we can draw two main conclusions. When the period T of bounce motion is negligible if compared to the time scale of interest, $dJ_\parallel/dt$ can be neglected. If the electric field ϕ_θ is also small, the variations of $J_\parallel$ become even smaller. However, if the time scale τ_f of electromagnetic fluctuations is fast, that is $\tau_f \gg T$, and if their amplitude cannot be neglected ($e\phi \gg H$), the second adiabatic invariant breaks down. This second scenario amounts at considering a large diffusion parameter $D_\perp \sim \phi_\theta^2/\tau_f$ and the value $\alpha = 0$ in the Fokker-Planck equation.

By taking the bounce orbit average of the rate of change (7.76), the classical result $\langle dJ_\parallel/dt \rangle = 0$ of [143] can be recovered. Nevertheless, note that $dJ_\parallel/dt$ can be described in terms of its bounce orbit average only when electromagnetic fluctuations ϕ_θ are sufficiently slow and small.

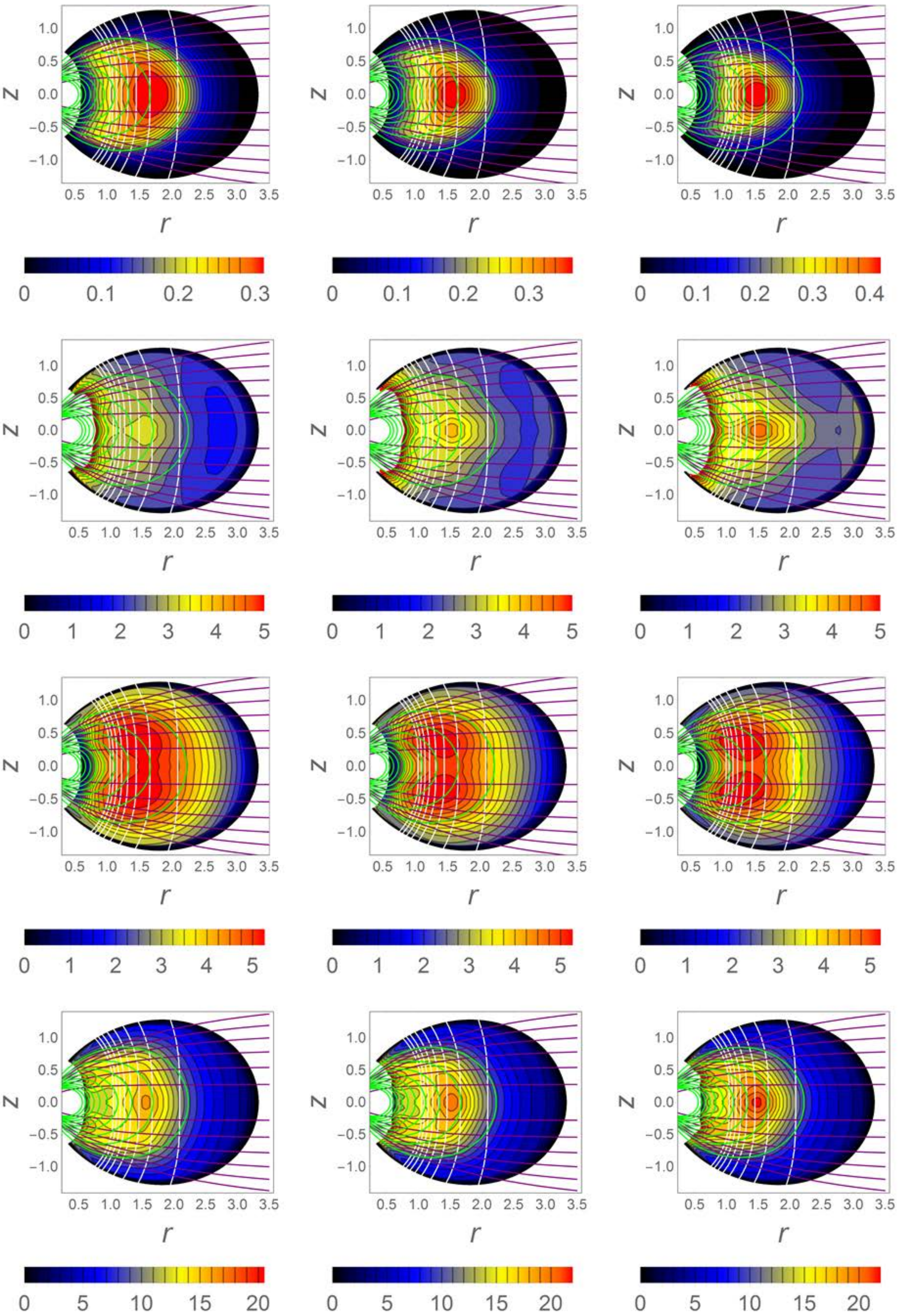

Figure 7.3: First row: time evolution of density $\rho\,(a.u.)$. Second row: temperature anisotropy $\mathfrak{a}$. Third row: parallel temperature $T_\parallel\,(eV)$. Fourth row: perpendicular temperature $T_\perp\,(eV)$. See main text for details.

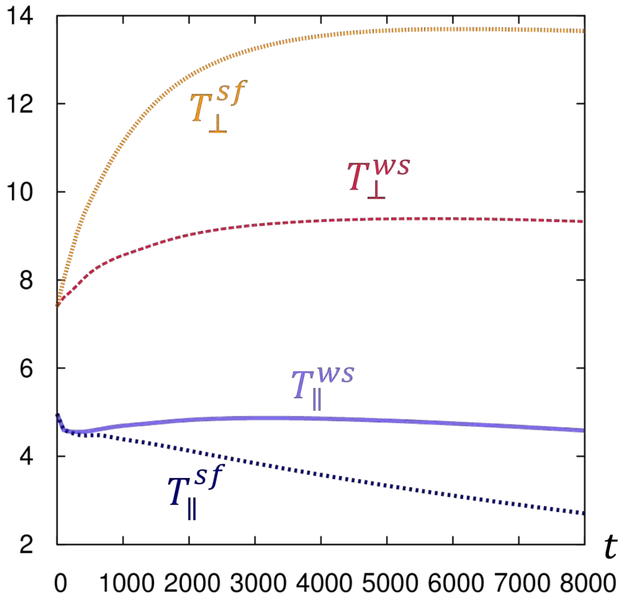

Figure 7.4: Time evolution of spatially averaged parallel and normal temperatures $T_\parallel\,(eV)$ and $T_\perp\,(eV)$. sf means strong and fast diffusion, while ws means weak and slow diffusion. In the first case the diffusion parameter $D_\perp$ is ten times greater than in the second one. Time is given in arbitrary units. This picture was taken from [84].

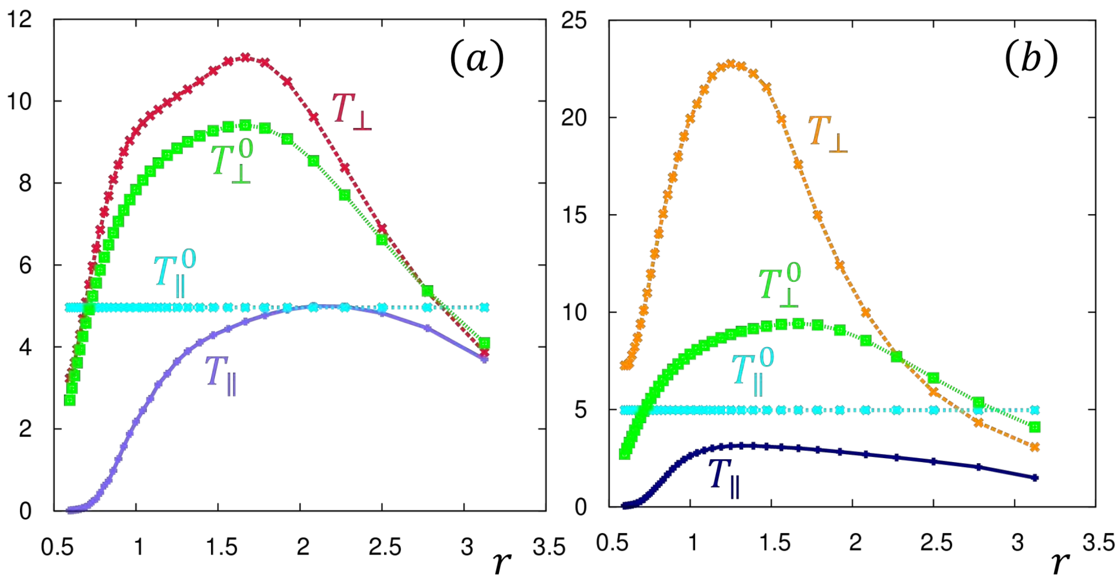

Figure 7.5: (a): Radial profiles at $z = 0$ of $T_\parallel\,(eV)$ and $T_\perp\,(eV)$ for the weaker diffusion parameter case. (b): Radial profiles at $z = 0$ of $T_\parallel\,(eV)$ and $T_\perp\,(eV)$ for the stronger diffusion parameter case. The 0 specifies the distribution at $t = 0$. This picture was taken from [84].

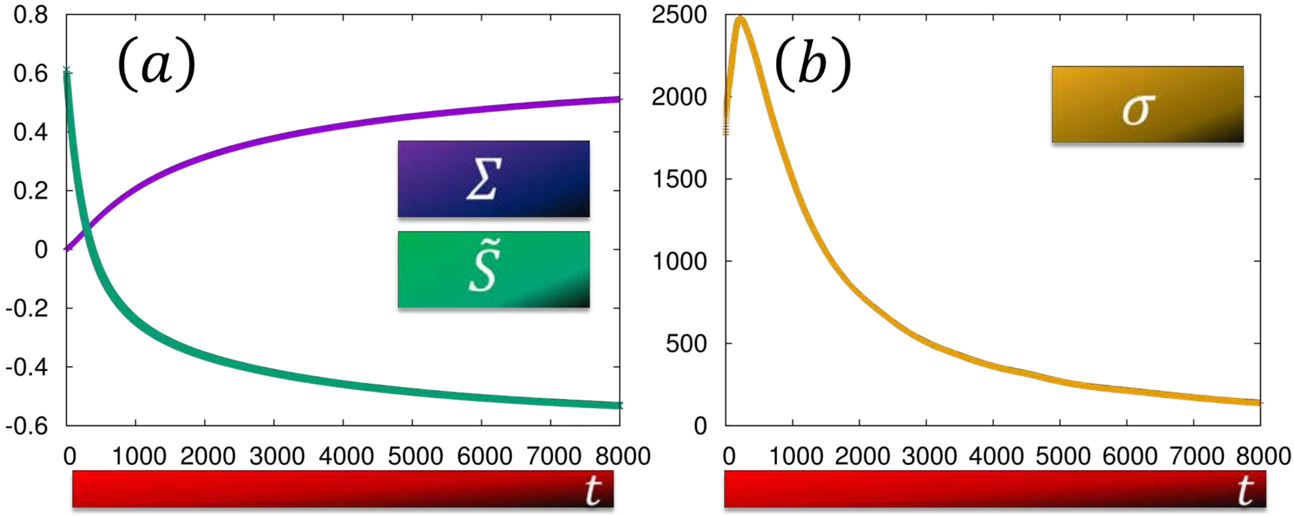

Figure 7.6: (a): Entropy measures Σ and $\tilde{S}$ as functions of time t. (b): Entropy production rate σ as a function of time t. The units are arbitrary. This picture was taken from [85].

Chapter 8

Poissonization of $\boldsymbol{E} \times \boldsymbol{B}$ Drift Dynamics

This chapter is dedicated to the study of conformal mechanical systems. The physical motivation behind the conformal operator introduced in definition 4.2 is the $\boldsymbol{E} \times \boldsymbol{B}$ drift velocity (7.10) already encountered in the previous chapter. Considered alone, this drift velocity represents a 3-dimensional conservative system with an associated antisymmetric operator. As already shown in proposition 4.5, any 3-dimensional antisymmetric operator can be extended to a 4-dimensional conformal operator. Then, a time reparametrization in terms of the conformal factor gives a 4-dimensional Hamiltonian system. Here, this 'Poissonization' procedure is worked out in detail and the statistical mechanics of the new extended system is investigated. Other physical examples of 3-dimensional conservative systems will be presented in the next chapter.

8.1 The Nonholonomic Plasma Particle

Consider a charged particle submerged in a magnetic field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ and subjected to an electric field $\boldsymbol{E} = -\nabla \phi$. The equation of motion is:

$$m\ddot{\boldsymbol{x}} = e\left(\dot{\boldsymbol{x}} \times \boldsymbol{B} + \boldsymbol{E}\right). \tag{8.1}$$

Here m is the particle mass, e its electric charge, and $\boldsymbol{x} \in \mathbb{R}^3$ its position. Suppose that m is sufficiently small so that the left-hand side of equation (8.1) can be neglected. We have already seen that if we further take the cross product with the magnetic field $\boldsymbol{B}$, (8.1) becomes:

$$X = \frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2}, \tag{8.2}$$

where $X = \dot{\boldsymbol{x}}_\perp$ is the velocity in the direction perpendicular to the magnetic field. We will also assume that the particle does not move along the magnetic field, giving $X = \dot{\boldsymbol{x}}$. The motion resulting from equation (8.2) goes under the name of $\boldsymbol{E} \times \boldsymbol{B}$ drift and will be the object of the present section. We refer the reader to [52, 74] for a more systematic derivation in the context of plasma physics.

The procedure leading to equation (8.2) can be made mathematically rigorous in terms of a reduction process. To see this, recall that the canonical phase space of a charged particle is described by the symplectic 2-form:

$$\begin{aligned}\omega &= dp_x \wedge dx + dp_y \wedge dy + dp_z \wedge dz \\ &= d\left(mv_x + eA_x\right) \wedge dx + d\left(mv_y + eA_y\right) \wedge dy + d\left(mv_z + eA_z\right) \wedge dz.\end{aligned} \tag{8.3}$$

The non-inertial reduction $m \to 0$ reduces ω to:

$$\omega^{'} = edA_x \wedge dx + edA_y \wedge dy + edA_z \wedge dz = ed\mathcal{A} = e\mathcal{B}, \tag{8.4}$$

where $\mathcal{B} = d\mathcal{A}$ is the magnetic field 2-form. Notice that the components of $\omega^{'}$ are now only functions of the spatial variables (x, y, z). Thus the reduction resulted in the contraction of the 6-dimensional canonical phase space to a 3-dimensional system. Furthermore, since $\omega^{'} = d\mathcal{A}$ is closed, it defines a non-canonical Hamiltonian system (it cannot be canonical because it is odd dimensional) with reduced equations of motion $X^{'} = \dot{\boldsymbol{x}}$ such that:

$$i_{X^{'}}\omega^{'} = -ed\phi. \tag{8.5}$$

Here, we used the fact that, since the particle mass is small, the energy of the system is the non-inertial Hamiltonian:

$$H = e\phi. \tag{8.6}$$

Solving equation (8.5) for $X^{'}$ one obtains:

$$X^{'} \times \boldsymbol{B} + \boldsymbol{E} = \boldsymbol{0}. \tag{8.7}$$

Since the first term on the left-hand side is perpendicular to $\boldsymbol{B}$, it follows that $E_{\parallel} = 0$. Then, no motion occurs along the magnetic field. We can thus perform the second reduction $v_{\parallel} \to 0$ so that $X^{'} = X^{'}_{\perp} = X$ and invert equation (8.7) to obtain the $\boldsymbol{E} \times \boldsymbol{B}$ drift equation of motion (8.2).

From this point we set $e = 1$ to simplify the notation. We want to show that, for an arbitrary magnetic field, equation (8.2) is conservative, i.e. it does not satisfy the Jacobi identity in general. Since the configuration space is 3-dimensional, the antisymmetric operator $\mathcal{J}$ associated to (8.2) will be a 3×3 matrix, whose action on the Hamiltonian can always be represented as the cross product of some vector $\boldsymbol{w}$, i.e. $\mathcal{J} = \boldsymbol{w}\times$. Then,

$$X = \mathcal{J}\left(dH\right) = \boldsymbol{w} \times \nabla H. \tag{8.8}$$

In our case $\boldsymbol{w} = \boldsymbol{B}/B^2$ and $\boldsymbol{E} = -\nabla\phi = -\nabla H$. One can verify that:

$$\mathcal{J} = w_x \partial_z \wedge \partial_y + w_y \partial_x \wedge \partial_z + w_z \partial_y \wedge \partial_x. \tag{8.9}$$

For the system (8.2) to be Hamiltonian, $\mathcal{J}$ has to be a Poisson operator, i.e. the Jacobi identity (1.13) has to be satisfied. Although equation (1.13) may be directly evaluated, we can avoid a lengthy calculation by recalling the Darboux theorem 1.1 according to which, in regions where the rank of $\boldsymbol{w}$ is constant, if the Jacobi identity holds the kernel of the Poisson operator has to be integrable. Since $\mathcal{J}$ is a 3-dimensional antisymmetric matrix, its maximum rank is 2. This implies that the dimension of the kernel is at least 1. In fact, noting that $\boldsymbol{w} \times \boldsymbol{w} = \boldsymbol{0}$, one sees that the covector associated to $\boldsymbol{w}$:

$$\theta = *i_{\boldsymbol{w}} vol^3 = w_x dx + w_y dy + w_z dz \tag{8.10}$$

belongs to $ker\left(\mathcal{J}\right)$, i.e. $\mathcal{J}\left(\theta\right) = 0$. In the above notation, $vol^3 = dx \wedge dy \wedge dz$, and $*$ is the Hodge star operator. Thus, assuming that the rank of $\boldsymbol{w}$ is two, a necessary condition for the Jacobi identity to be satisfied is:

$$\theta \wedge d\theta = \left(\boldsymbol{w} \cdot \nabla \times \boldsymbol{w}\right) vol^3 = 0. \tag{8.11}$$

Here, we used the Frobenius integrability condition for the covector θ (see theorem 1.2).

In definition 5.1 we have seen that the kernel of an antisymmetric operator defines topological constraints for the trajectory of the particle. Now, suppose that the constraint is integrable: $\theta = \lambda dC$, where the functions λ and C are integration factor and Casimir invariant respectively. This implies that the orbit of the particle lies on the integral manifold defined by C =constant. Then, one could reduce the equations of motion to the submanifold $\mathbb{R}^3/C$ and obtain a 2-dimensional Hamiltonian system with Hamiltonian H. Thus, equation (8.11) is not only a necessary condition, but also a sufficient condition for the Jacobi identity (1.13) to hold. Note that, since we have used neither a specific form of the Hamiltonian, nor an explicit expression for the antisymmetric operator $\mathcal{J}$, this argument is valid for all 3-dimensional antisymmetric operators. In fact, a direct evaluation of (1.13) would give exactly (8.11).

In the light of (8.11), $\boldsymbol{E}\times\boldsymbol{B}$ dynamics (8.2) is Hamiltonian only if locally $\boldsymbol{w}=\lambda\nabla C$, i.e. when the magnetic field is a local solution to the equation:

$$\frac{\boldsymbol{B}}{B^2}=\lambda\nabla C, \tag{8.12}$$

for some appropriate λ and C. The condition above is verified, for example, in the presence of an harmonic magnetic field $\boldsymbol{B}=\nabla\xi$. In this scenario $\lambda=B^{-2}$ and $C=\xi$. However, equation (8.12) does not hold in general, and the system is a degenerate antisymmetric algebra with the nonholonomic constraint $\theta=0$.

In order to understand the peculiar nature of nonholonomic dynamics, it is useful to compare it with a conventional holonomic system. First, consider the following nonholonomic particle performing $\boldsymbol{E}\times\boldsymbol{B}$ drift:

$$\theta=(\cos z+\sin z)\,dx+(\cos z-\sin z)\,dy, \tag{8.13a}$$

$$H=\frac{1}{2}\left(x^2+y^2+z^2\right). \tag{8.13b}$$

The Jacobi identity (8.11) reads:

$$\theta\wedge d\theta=(\boldsymbol{w}\cdot\nabla\times\boldsymbol{w})\,vol^3=(\boldsymbol{w}\cdot\boldsymbol{w})\,vol^3=2\,vol^3\neq 0, \tag{8.14}$$

which implies that this system is not Hamiltonian with nonholonomic constraint $\theta=0$. From (8.2), the equations of motion are:

$$\begin{aligned}X=(\cos z-\sin z)\,z\,\partial_x-(\cos z+\sin z)\,z\,\partial_y\\+\left[(\cos z+\sin z)\,y-(\cos z-\sin z)\,x\right]\partial_z.\end{aligned} \tag{8.15}$$

Now, consider the motion of a rigid body with angular momentum $\boldsymbol{x}$ and momenta of inertia I_x, I_y, I_z:

$$\theta=xdx+ydy+zdz, \tag{8.16a}$$

$$H=\frac{1}{2}\left(\frac{x^2}{I_x}+\frac{y^2}{I_y}+\frac{z^2}{I_z}\right). \tag{8.16b}$$

This time the Jacobi identity (8.11) is satisfied since $d\theta=0$ and the relevant Casimir invariant is the total angular momentum $C=\boldsymbol{x}^2/2$, with $\theta=dC$. Thus, this second system is Hamiltonian, with holonomic constraint $C=$constant. The equations of motion are written as:

$$X=yz\left(\frac{1}{I_z}-\frac{1}{I_y}\right)\partial_x+xz\left(\frac{1}{I_x}-\frac{1}{I_z}\right)\partial_y+xy\left(\frac{1}{I_y}-\frac{1}{I_x}\right)\partial_z. \tag{8.17}$$

Figure 8.1 shows the trajectory of the nonholonomic plasma particle (8.15) and that of the rigid body (8.17). Both of them lie on the integral surface of constant energy. However, while the orbit of the rigid body is closed and results from the intersection of the two integral manifolds defined by H and $C = \boldsymbol{x}^2/2$, the plasma particle spirals toward a sink and delineates an open path characterized by the non-zero divergence of the conservative vector field (8.15).

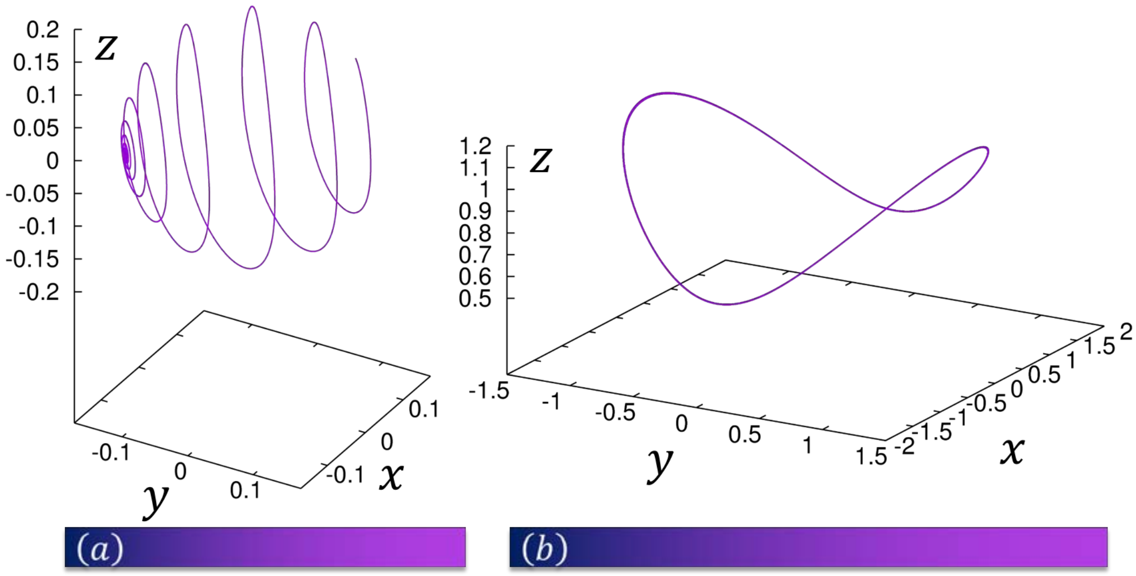

Figure 8.1: (a): numerical integration of (8.15). (b): numerical integration of (8.17).

This example shows that there is an important relationship between the existence of an invariant measure and the Hamiltonian nature of the system. Specifically, due to Liouville's theorem 2.1, the existence of an invariant measure is a necessary condition that a general system must satisfy to be Hamiltonian. For a 3-dimensional antisymmetric algebra it turns out that if the antisymmetric bracket is measure preserving, it is also a Poisson bracket (see proposition 4.6). Indeed, an invariant measure exists for any Hamiltonian provided that one can find a Jacobian g such that:

$$\mathfrak{L}_X g \, vol^3 = div\,(gX)\, vol^3 = 0 \qquad \forall H, \tag{8.18}$$

Recalling that in 3 dimensions $X = \boldsymbol{w} \times \nabla H$, we obtain the condition:

$$\nabla \cdot (g\boldsymbol{w} \times \nabla H) = \nabla H \cdot \nabla \times (g\boldsymbol{w}) = 0 \qquad \forall H, \tag{8.19}$$

which holds when $\boldsymbol{w} = g^{-1}\nabla C$ for some function C, i.e. the constraint $\theta = g^{-1}dC$ is integrable and the system is Hamiltonian.

The absence of an invariant measure may be interpreted as the consequence of missing degrees of freedom that would compensate the compressibility of the system. This is why we will need to 'extend' the system in order to recover an Hamiltonian structure.

8.2 Poissonization in Three Dimensions

The purpose of the present section is to develop a systematic procedure to 'repair' an arbitrary 3-dimensional antisymmetric bracket and obtain an equivalent Hamiltonian system describing the same dynamics.

8.2.1 *Extension*

The first step of the procedure consists of embedding the antisymmetric algebra in a larger space. The objective is to restore an invariant measure and a conformally Poisson structure. To do so in the three dimensional setting, it will be sufficient to add a single new variable s. We begin by extending the antisymmetric operator $\mathcal{J}$ of (8.9) in the following manner:

$$\mathfrak{J} = \mathcal{J} + a\partial_x \wedge \partial_s + b\partial_y \wedge \partial_s + c\partial_z \wedge \partial_s, \tag{8.20}$$

where $\mathfrak{J}$ is the 4×4 extended antisymmetric operator and the coefficients a, b, c have to be determined by requiring that the new operator is conformally Poisson. We remark that these new terms do not affect the original equations of motion since the Hamiltonian function does not depend on the new variable s, i.e. $H_s = 0$. Returning to our problem, we must show that the differential 2-form Ω satisfying:

$$i_X\Omega = i_{\mathfrak{J}(dH)}\Omega = -dH, \tag{8.21}$$

is conformally closed:

$$d\left(r\Omega\right) = 0, \tag{8.22}$$

for some conformal factor $r \neq 0$. First, let us evaluate Ω. From equation (8.21), we have:

$$i_X\Omega = \left(\sum_{k<l} \Omega_{kl} dx^k \wedge dx^l\right)\left(\mathfrak{J}^{ij} H_i \partial_j\right) = -\mathfrak{J}^{ik}\Omega_{kl} H_i dx^l = -H_l dx^l, \tag{8.23}$$

which amounts at finding the inverse matrix of $\mathfrak{J}$:

$$\mathfrak{J}^{ik}\Omega_{kl} = \delta^{il}. \tag{8.24}$$

Remembering (8.20), we find:

$$\begin{aligned}\Omega = &\frac{1}{aw_x + bw_y + cw_z}\left\{\left[a - s\left(\frac{\partial w_z}{\partial y} - \frac{\partial w_y}{\partial z}\right)\right] dy \wedge dz + \left[b - s\left(\frac{\partial w_x}{\partial z} - \frac{\partial w_z}{\partial x}\right)\right] dz \wedge dx\right.\\ &\left. + \left[c - s\left(\frac{\partial w_y}{\partial x} - \frac{\partial w_x}{\partial y}\right)\right] dx \wedge dy + d\left(w_x s\right) \wedge dx + d\left(w_y s\right) \wedge dy + d\left(w_z s\right) \wedge dz\right\}.\end{aligned} \tag{8.25}$$

With the choice:

$$a = D_x + s\left(\frac{\partial w_z}{\partial y} - \frac{\partial w_y}{\partial z}\right), \tag{8.26a}$$

$$b = D_y + s\left(\frac{\partial w_x}{\partial z} - \frac{\partial w_z}{\partial x}\right), \tag{8.26b}$$

$$c = D_z + s\left(\frac{\partial w_y}{\partial x} - \frac{\partial w_x}{\partial y}\right), \tag{8.26c}$$

the vorticity 2-form Ω becomes:

$$\Omega = \frac{\mathcal{D} + d\left(w_x s\right) \wedge dx + d\left(w_y s\right) \wedge dy + d\left(w_z s\right) \wedge dz}{\boldsymbol{w} \cdot \left(\boldsymbol{D} + s\nabla \times \boldsymbol{w}\right)} = \frac{d\left(\mathcal{E} + s\theta\right)}{*\left[\theta \wedge \left(d\mathcal{E} + sd\theta\right)\right]}, \tag{8.27}$$

Here $\mathcal{D} = d\mathcal{E} = D_x dy \wedge dz + D_y dz \wedge dx + D_z dx \wedge dy$ is an arbitrary closed 2-form that does not depend on s and $\boldsymbol{D} = (D_x, D_y, D_z)$:

$$d\mathcal{D} = \left(\nabla \cdot \boldsymbol{D}\right) dx \wedge dy \wedge dz = 0. \tag{8.28}$$

From equation (8.27) we see that Ω is conformally closed with conformal factor:

$$r = \boldsymbol{w} \cdot \left(\boldsymbol{D} + s\nabla \times \boldsymbol{w}\right) = *\left[\theta \wedge \left(d\mathcal{E} + sd\theta\right)\right]. \tag{8.29}$$

Notice that the vector $\boldsymbol{D}$ must be chosen so that $r \neq 0$ on the domain of interest.

It is useful to write down the explicit expression for the extended equations of motion $X = \mathfrak{J}\left(dH\right)$:

$$X = \boldsymbol{w} \times \nabla H - \left(\boldsymbol{D} + s\nabla \times \boldsymbol{w}\right) \cdot \nabla H \partial_s. \tag{8.30}$$

Finally, one can verify that the new equations are divergence free (the extended operator is measure preserving):

$$div\left(X\right) = \nabla \cdot \left(\boldsymbol{w} \times \nabla H\right) - \nabla \times \boldsymbol{w} \cdot \nabla H = 0. \tag{8.31}$$

8.2.2 *Time reparametrization*

The second step of the procedure involves a time reparametrization that will give us the desired Poisson structure. From chapter 4 we already know that this result can be achieved by introducing the new time variable (proper time) τ satisfying:

$$\frac{d\tau}{dt} = r. \tag{8.32}$$

Then, the operator $r^{-1}\mathfrak{J}$ satisfies the Jacobi identity and defines an Hamiltonian system with energy H and time τ.

In this paragraph we want to show that the time reparametrization can be understood in terms of a *gauge transformation* of the relevant vorticity 2-form (see [20] for the concept of gauge transformation). First, we shall treat the time variables on the same footing of the others and consider the vorticity 2-form:

$$\Gamma = \Omega + dt \wedge dH = \frac{d\eta}{r} + dt \wedge dH, \tag{8.33}$$

where we set $\eta = \mathcal{E} + s\theta$ and used (8.27). In terms of Γ, the equations of motion take the form ($\dot{t} = 1$):

$$i_{(X,t)}\Gamma = 0. \tag{8.34}$$

There are two types of gauge transformations that act on vorticity 2-forms and leave the dynamics unchanged. The first kind is called dynamical gauge transformation and consists of adding a gauge 2-form $\mathcal{U}$ such that $i_{(X,t)}\mathcal{U} = 0$:

$$\Gamma^{'} = \Gamma + \mathcal{U}, \tag{8.35}$$

where $\Gamma^{'}$ is the transformed vorticity 2-form. The second kind involves the multiplication of Γ by an arbitrary non-zero function $\mathcal{V}$:

$$\Gamma^{'} = \mathcal{V}\Gamma. \tag{8.36}$$

Now, set:

$$\mathcal{U} = \frac{d\tau \wedge dH}{r} - dt \wedge dH, \tag{8.37a}$$

$$\mathcal{V} = r. \tag{8.37b}$$

Using (8.32), one can verify that $i_{(X,\dot{\tau},t)}\mathcal{U} = 0$, and, by definition, $r \neq 0$. Then,

$$\Gamma^{'} = \mathcal{V}\left(\Gamma + \mathcal{U}\right) = d\eta + d\tau \wedge dH. \tag{8.38}$$

This 2-form is closed. Thus, we have obtained a Poisson structure where the new variable τ has the role of time parameter.

Then, the reparametrization procedure can be understood as follows. Suppose that we solved the new equations of motion:

$$i_{(X^{'},\tau^{'})}\Gamma^{'} = 0, \tag{8.39}$$

in the new time variable τ and obtained $\boldsymbol{x} = \boldsymbol{x}\left(\tau\right)$. In this notation $X^{'} = d\boldsymbol{x}/d\tau$ and $\tau^{'} = 1$. From (8.32) we calculate $t = t\left(\tau\right)$ and, since r does not change sign, we can invert the expression for t to obtain $\tau = \tau\left(t\right)$ and $\boldsymbol{x} = \boldsymbol{x}\left(t\right)$.

As an exercise, let us perform the reverse transformation from (8.38) to (8.33) by using different gauges. Define[1] $\psi = \tau/t$. Then,

$$\begin{aligned}\Gamma^{'} &= d\eta + \frac{\psi}{r}d\eta - \frac{\psi}{r}d\eta + \psi dt \wedge dH + t d\psi \wedge dH \\ &= \psi\left(\frac{d\eta}{r} + dt \wedge dH\right) + \left(1 - \frac{\psi}{r}\right)d\eta + t d\psi \wedge dH.\end{aligned} \tag{8.40}$$

From $i_X d\eta = -r dH$ and recalling (8.32):

$$i_{(X,\dot{\tau},t)}\left[\left(1 - \frac{\psi}{r}\right)d\eta + t d\psi \wedge dH\right] = 0. \tag{8.41}$$

Thus, by setting:

$$\mathcal{U} = \left(1 - \frac{\psi}{r}\right)d\eta + t d\psi \wedge dH, \tag{8.42a}$$

$$\mathcal{V} = \psi^{-1}, \tag{8.42b}$$

the gauge transformation gives $\mathcal{V}\left(\Gamma^{'} - \mathcal{U}\right) = \Gamma$.

8.3 Poissonization of the Nonholonomic Plasma Particle

Here we apply the procedure developed so far to Poissonize the nonholonomic plasma particle obeying the equation of motion (8.15).

[1] If τ or t are 0 one can use the gauge transformation (8.37).

8.3.1 *The physical meaning of* s *and* τ

First, let us spend some words on the physical meaning of the new variable s. From equation (8.30), and recalling that in this case $\boldsymbol{w} = \boldsymbol{B}/B^2$ with $B^2 = 1/2$, we have:

$$\begin{aligned}\dot{s} &= -\left(\boldsymbol{D} + s\nabla\times\boldsymbol{w}\right)\cdot\nabla H = -\left(\boldsymbol{D} + s\boldsymbol{w}\right)\cdot\nabla H \\ &= -\left(\boldsymbol{B} + s\frac{\boldsymbol{B}}{B^2}\right)\cdot\nabla H = -\left(\frac{1}{\sqrt{2}} + s\sqrt{2}\right)\frac{\partial H}{\partial \ell},\end{aligned} \tag{8.43}$$

Here we made the choice $\boldsymbol{D} = \boldsymbol{B}$ (we will justify this choice later) and ℓ measures the length along a field line ($\partial_\ell = \boldsymbol{B}/B$). We define:

$$m\tilde{v}_\parallel = \frac{1}{\sqrt{2}}\log\left[c_0\left(\frac{1}{\sqrt{2}} + s\sqrt{2}\right)\right], \tag{8.44}$$

with c_0 a constant. This implies:

$$m\frac{d\tilde{v}_\parallel}{dt} = -\frac{\partial H}{\partial \ell}. \tag{8.45}$$

Thus, the variable $\tilde{v}_\parallel$ can be interpreted as a pseudo-velocity in the direction parallel to $\boldsymbol{B}$: the new degree of freedom s compensates the motion along the magnetic field that was missing in the 3-dimensional system. Inverting equation (8.44) we also have (recall that m is small):

$$s = \frac{1}{2}\left(e^{\sqrt{2}m\tilde{v}_\parallel} - 1\right) = \frac{m\tilde{v}_\parallel}{\sqrt{2}} + o\left(\left(\sqrt{2}m\tilde{v}_\parallel\right)^2\right). \tag{8.46}$$

In the above equation we required that $s = 0$ when $\tilde{v}_\parallel = 0$ so that $c_0 = \sqrt{2}$ (we will justify this choice later).

What about the meaning of the proper time τ? Using the expression for r equation (8.29),

$$r = 1 + s\boldsymbol{w}\cdot\nabla\times\boldsymbol{w} = 1 + 2s. \tag{8.47}$$

Here we used the fact that $\boldsymbol{D}\cdot\boldsymbol{w} = \boldsymbol{B}\cdot\boldsymbol{w} = 1$. From (8.32) and defining the pseudo-length $\tilde{\ell}$ such that $\tilde{v}_\parallel = d\tilde{\ell}/dt$:

$$\frac{d\tau}{dt} = 1 + 2s = e^{\sqrt{2}m\tilde{v}_\parallel} = \exp\left\{\sqrt{2}m\frac{d\tilde{\ell}}{dt}\right\}. \tag{8.48}$$

Thus, the choices $\boldsymbol{D} = \boldsymbol{B}$ and $c_0 = \sqrt{2}$ are now physically justified because the conformal factor r must be 1 when $\boldsymbol{w}$ is integrable or $\tilde{v}_\parallel = 0$, i.e. we must have $d\tau/dt = 1$ when the Jacobi identity is satisfied or there is no motion along the magnetic field. The remarkable physical interpretation of the conformal factor is that a non-integrable magnetic field $\boldsymbol{B} = \boldsymbol{w}/w^2$ determines a distortion of time which depends on the non-zero helicity $\boldsymbol{w} \cdot \nabla \times \boldsymbol{w} = B^{-4}\boldsymbol{B} \cdot \nabla \times \boldsymbol{B} \neq 0$. If the mass m is small enough we can expand the exponential to obtain:

$$\frac{d\tau}{dt} = 1 + \sqrt{2}m\frac{d\tilde{\ell}}{dt} + o\left(\left(\sqrt{2}m\frac{d\tilde{\ell}}{dt}\right)^2\right). \tag{8.49}$$

Neglecting second order terms, the final result is:

$$\tau = \text{constant} + t + \sqrt{2}m\tilde{\ell}, \tag{8.50}$$

and the proper time τ can be interpreted as a measure of the distance traveled by the particle along the magnetic field.

8.3.2 *Poissonization in Cartesian coordinates*

We are now ready to write the canonical equations of motion for the nonholonomic plasma particle. Recalling (8.13a) and (8.38), the symplectic 2-form of interest is:

$$\begin{aligned}\Gamma' &= d\left[\mathcal{E} + s\left(\cos z + \sin z\right)dx + s\left(\cos z - \sin z\right)dy\right] + d\tau \wedge dH \\ &= -d\left\{xd\left[\left(s+\frac{1}{2}\right)\left(\cos z + \sin z\right)\right] + yd\left[\left(s+\frac{1}{2}\right)\left(\cos z - \sin z\right)\right]\right\} + d\tau \wedge dH \\ &= d\left(p_x dq_x + p_y dq_y\right) + d\tau \wedge dH,\end{aligned} \tag{8.51}$$

where we used the fact that $\mathcal{E} = \theta/2$ and introduced canonical variables:

$$q_x = \left(s + \frac{1}{2}\right)\left(\cos z + \sin z\right), \tag{8.52a}$$

$$p_x = -x, \tag{8.52b}$$

$$q_y = \left(s + \frac{1}{2}\right)\left(\cos z - \sin z\right), \tag{8.52c}$$

$$p_y = -y. \tag{8.52d}$$

In terms of these new variables we also have:

$$z = \arcsin\left[\frac{q_x - q_y}{\sqrt{2\left(q_x^2 + q_y^2\right)}}\right], \tag{8.53a}$$

$$s + \frac{1}{2} = \sqrt{\frac{q_x^2 + q_y^2}{2}}, \tag{8.53b}$$

$$H = \frac{1}{2}\left(p_x^2 + p_y^2 + \arcsin^2\left[\frac{q_x - q_y}{\sqrt{2\left(q_x^2 + q_y^2\right)}}\right]\right). \tag{8.53c}$$

Here, we chose the positive root for $s + 1/2$. Finally,

$$q_x^{'} = H_{p_x} = p_x, \tag{8.54a}$$

$$p_x^{'} = -H_{q_x} = \frac{-q_y}{q_x^2 + q_y^2}\arcsin\left[\frac{q_x - q_y}{\sqrt{2\left(q_x^2 + q_y^2\right)}}\right], \tag{8.54b}$$

$$q_y^{'} = H_{p_y} = p_y, \tag{8.54c}$$

$$p_y^{'} = -H_{q_y} = \frac{q_x}{q_x^2 + q_y^2}\arcsin\left[\frac{q_x - q_y}{\sqrt{2\left(q_x^2 + q_y^2\right)}}\right]. \tag{8.54d}$$

Figure 8.2 shows a numerical integration of the Hamiltonian system (8.54). The solution progressively approaches a 2-dimensional uniform rectilinear motion.

Finally, in the original time t, the equations of motion for the canonical variables p_x, q_x, p_y, and q_y take the form:

$$\dot{q}_x = r^{-1} H_{p_x}, \tag{8.55a}$$

$$\dot{p}_x = -r^{-1} H_{q_x}, \tag{8.55b}$$

$$\dot{q}_y = r^{-1} H_{p_y}, \tag{8.55c}$$

$$\dot{p}_y = -r^{-1} H_{q_y}. \tag{8.55d}$$

These equations, which are not canonical, imply that the ‘force’ acting on the particle is only proportional to the gradient of the Hamiltonian with proportionality factor r^{-1}. Therefore, the same energy gradient produces different forces depending on the position in space. Such behavior departs from the standard laws of physics and signals the importance of the Jacobi identity in determining the structure of the equations of motion. This inhomogeneity is also the reason why canonical equations can be obtained only by ‘adjusting’ the time variable.

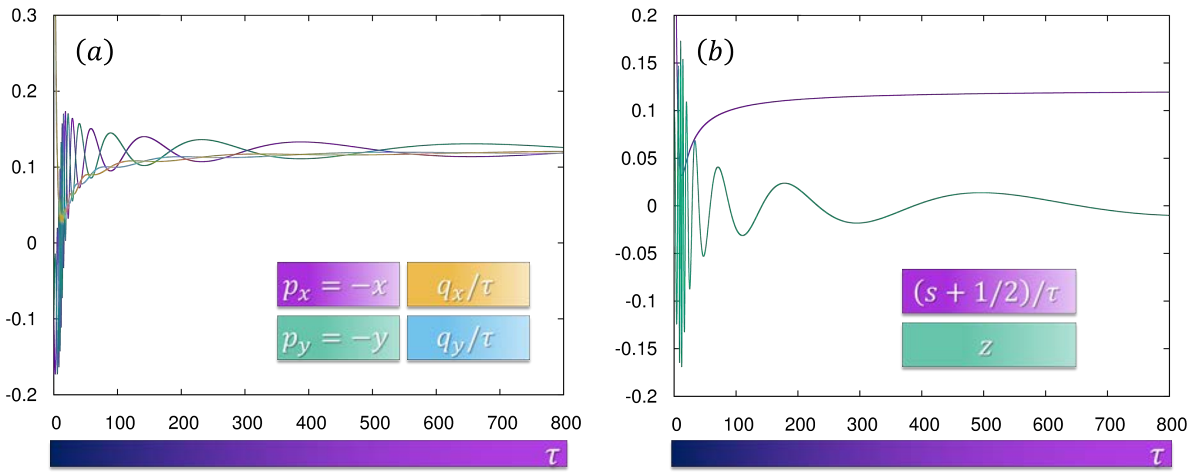

Figure 8.2: Numerical integration of system (8.54). (a): evolution with respect to the proper time τ of p_x, q_x/τ, p_y, and q_y/τ. (b): evolution with respect to the proper time τ of $\left(s + \frac{1}{2}\right)/\tau$ and z.

8.3.3 *Poissonization in magnetic coordinates*

Consider again the nonholonomic plasma particle of equation (8.15). By performing an appropriate change of coordinates before the Poissonization procedure, we can simplify the antisymmetric operator $\mathcal{J}$. The simplified form will offer us an insight into the relation between $\mathcal{J}$, which is the result of the abrupt reduction $\left(m, v_{\parallel}\right) \to 0$, and the full dynamics of a magnetized particle. The target coordinates are the magnetic coordinates (ℓ, ψ, ζ) (to avoid confusion with the constraint θ, the toroidal angle is now ζ). This time we shall consider a more general type of magnetic field[2]:

$$\boldsymbol{B} = \nabla\psi \times \nabla\zeta + i\left(\ell, \psi\right)\nabla\psi \times \nabla\ell. \tag{8.56}$$

Here i is an arbitrary function of ℓ and ψ. When $i = 0$ equation (8.56) can conveniently represent a dipole magnetic field. The flux function ψ is chosen so that $\psi = \psi\left(R, z\right)$, where (R, ζ, z) is a cylindrical coordinate system $dx \wedge dy \wedge dz = RdR \wedge d\zeta \wedge dz$ with R the radial coordinate in the (x, y) plane and ζ the toroidal angle. Here, the coordinate ℓ is defined to be the length along the field lines of the poloidal component $\boldsymbol{B}_p = \nabla\psi \times \nabla\zeta$ of the magnetic field $\boldsymbol{B}$. In formulae:

$$\partial_\ell = \frac{\boldsymbol{B}_p}{B_p}. \tag{8.57}$$

[2]Note that with the identification $\ell = y$, $\psi = \left(\sin z + \cos z\right)/2$, $\zeta = x$, $i = \left(\sin z + \cos z\right)/\left(\sin z - \cos z\right)$, where (x, y, z) are Cartesian coordinates, equation (8.56) also gives the magnetic field studied in the previous example.

Thus, if $i \neq 0$, the magnetic field has a toroidal component $\boldsymbol{B}_t = i(\ell, \psi) \nabla\psi \times \nabla\ell$. In terms of the new coordinates, the magnetic field 2-form $\mathfrak{B}$ reads as:

$$\mathfrak{B} = d\psi \wedge d\zeta + i d\psi \wedge dl. \tag{8.58}$$

Note that $d\mathfrak{B} = 0$. In order to express $\mathcal{J}$ in the new variables we need some geometrical relationships among tangent and cotangent vectors. We begin by calculating the Jacobian $\mathcal{Q}$ of the coordinate change:

$$\mathcal{Q} = \nabla\ell \cdot \nabla\psi \times \nabla\zeta = B_p. \tag{8.59}$$

Here we used $\nabla\ell \cdot \partial_\ell = 1$. Similarly, by using the reciprocity relationships $\nabla x^j \cdot \partial_k = \delta_{jk}$ with $x^j, x^k = \ell, \psi, \zeta$, one can obtain the following expressions:

$$\nabla\ell = \frac{1}{|\partial_\psi|^2 - q^2}\left(q\partial_\psi + |\partial_\psi|^2 \partial_\ell\right), \tag{8.60a}$$

$$\nabla\psi = \frac{1}{|\partial_\psi|^2 - q^2}\left(\partial_\psi + q\partial_\ell\right), \tag{8.60b}$$

$$|\nabla\ell|^2 = \frac{|\partial_\psi|^2}{|\partial_\psi|^2 - q^2}, \tag{8.60c}$$

$$|\partial_\psi|^2 = \frac{|\nabla\ell|^2}{R^2 B_p^2}, \tag{8.60d}$$

$$B_p^2 = \frac{1}{R^2\left(|\partial_\psi|^2 - q^2\right)}. \tag{8.60e}$$

Here $q = -\partial_\ell \cdot \partial_\psi$. Recalling that the $\boldsymbol{E} \times \boldsymbol{B}$ equations of motion are given by (8.2) and exploting (8.60) with the identities $\nabla x^j = (\partial_k \times \partial_m)/(\partial_j \cdot \partial_k \times \partial_m)$, $\partial_j = \left(\nabla x^k \times \nabla x^m\right)/\left(\nabla x^j \cdot \nabla x^k \times \nabla x^m\right)$ where $x^j, x^k, x^m = \ell, \psi, \zeta$ are all different:

$$\dot{\ell} = \nabla\ell \cdot X = \rho\left(iR^2 H_\psi - qH_\zeta\right), \tag{8.61a}$$

$$\dot{\psi} = \nabla\psi \cdot X = -\rho\left(H_\zeta + iR^2 H_\ell\right), \tag{8.61b}$$

$$\dot{\zeta} = \nabla\zeta \cdot X = \rho\left(H_\psi + qH_\ell\right), \tag{8.61c}$$

where $\rho = B_p^2/B^2$. Therefore, the bivector $\mathcal{J}$ takes the form:

$$\mathcal{J} = \rho\left(\partial_\zeta \wedge \partial_\psi + q\partial_\zeta \wedge \partial_\ell + iR^2 \partial_\ell \wedge \partial_\psi\right). \tag{8.62}$$

A straightforward application of the Poissonization procedure of section 4 (equations (8.27) and (8.29)) gives the 4-dimensional conformal 2-form:

$$\Omega = \frac{\mathcal{B} + d\left(s\theta\right)}{*\left[\theta \wedge \left(\mathcal{B} + sd\theta\right)\right]}, \tag{8.63}$$

where we set $\mathcal{D} = d\mathcal{A} = \mathcal{B}$ and the kernel of $\mathcal{J}$ is the covector:

$$\theta = \rho\left(d\ell - qd\psi - iR^2 d\zeta\right). \tag{8.64}$$

Thus, a time reparametrization $d\tau/dt = r$ with conformal factor:

$$\begin{aligned} r &= *\left[\theta \wedge \left(\mathfrak{B} + sd\theta\right)\right] \\ &= \rho\left\{1 + i^2R^2 + s\left[\rho\frac{\partial q}{\partial \zeta} + iR^2\left(\frac{\partial\left(q\rho\right)}{\partial \ell} + \frac{\partial \rho}{\partial \psi}\right) - \left(q\partial_\ell + \partial_\psi\right)\left(iR^2\rho\right)\right]\right\}, \end{aligned} \tag{8.65}$$

will give the symplectic 2-form:

$$\begin{aligned} \Gamma' &= \mathfrak{B} + d\left(s\theta\right) + d\tau \wedge dH \\ &= d\psi \wedge d\zeta + id\psi \wedge d\ell + d\left[s\rho\left(d\ell - qd\psi - iR^2d\zeta\right)\right] + d\tau \wedge dH. \end{aligned} \tag{8.66}$$

Now suppose that $\boldsymbol{B} = \boldsymbol{B}_p$, i.e. $i = 0$. Then, $\rho = 1$ (note that $q_\zeta = 0$ when $i = 0$ due to toroidal symmetry), $\theta \wedge d\theta = (d\ell - qd\psi) \wedge (q_\ell dl \wedge d\psi) = 0$ and $r = 1$. Furthermore, Ω becomes symplectic:

$$\Omega = d\left(\psi d\zeta + sd\ell - sqd\psi\right). \tag{8.67}$$

With the identification $s = v_{\|}$ of equation (8.46)), the expression (8.67) is exactly the symplectic 2-form for the motion of a magnetized particle in a magnetic field of the form $\nabla\psi \times \nabla\zeta$ (remember equation (7.35)). From this example we can see explicitly that the failure of the Jacobi identity is controlled by the non-integrability i of the magnetic field[3] $\boldsymbol{B}$, that the Poissonization procedure reproduces the correct physics, and that in the presence of a general magnetic field (note that (8.63) holds for any $\mathcal{B}$) canonical coordinates can be obtained by operating a time reparametrization.

We conclude this section by giving the antisymmetric operator and the conformal factor for a magnetic field written as:

$$\boldsymbol{B} = \alpha\nabla\psi \times \nabla\zeta + i\nabla\psi \times \nabla\ell + \beta\nabla\zeta \times \nabla\ell. \tag{8.68}$$

[3] One can verify that $\boldsymbol{B} \cdot \nabla \times \boldsymbol{B} = 0$ when $i = 0$ and $\psi = \psi\left(R, z\right)$.

Here α, β, and i are 3 arbitrary functions satisfying $\alpha_\ell - i_\zeta + \beta_\psi = 0$ to ensure that $d\mathcal{B} = 0$. Note that all magnetic fields with $\boldsymbol{B_p} \neq 0$ so that ℓ can be defined can be cast in the form (8.68). With the same procedure as above one obtains:

$$\mathcal{J} = \rho\left[(\alpha - \beta q)\,\partial_\zeta \wedge \partial_\psi + \left(\beta\,|\partial_\psi|^2 - \alpha q\right)\partial_\ell \wedge \partial_\zeta + iR^2 \partial_\ell \wedge \partial_\psi\right], \tag{8.69}$$

The kernel of this operator is the covector:

$$\theta = \rho\left[(\alpha - \beta q)\,d\ell + \left(\beta\,|\partial_\psi|^2 - \alpha q\right)d\psi - iR^2 d\zeta\right] \tag{8.70}$$

The conformal factor is:

$$\begin{aligned} r = \rho\Bigg\{ &\alpha\,(\alpha - \beta q) + \beta\left(\beta\,|\partial_\psi|^2 - \alpha q\right) + i^2 R^2 \\ &+ s\left[(\alpha - \beta q)\left(-\frac{\partial\left(iR^2\rho\right)}{\partial\psi} - \frac{\partial\rho\left(\beta\,|\partial_\psi|^2 - \alpha q\right)}{\partial\zeta}\right)\right. \\ &+ \left(\beta\,|\partial_\psi|^2 - \alpha q\right)\left(\frac{\partial\rho\,(\alpha - \beta q)}{\partial\zeta} + \frac{\partial\left(iR^2\rho\right)}{\partial\ell}\right) \\ &\left. - iR^2\left(\frac{\partial\rho\left(\beta\,|\partial_\psi|^2 - \alpha q\right)}{\partial\ell} - \frac{\partial\rho\,(\alpha - \beta q)}{\partial\psi}\right).\right]\Bigg\} \end{aligned} \tag{8.71}$$

Finally, the symplectic 2-form recovered after time reparametrization is:

$$\begin{aligned} \Gamma' &= \alpha d\psi \wedge d\zeta + i d\psi \wedge d\ell + \beta d\zeta \wedge d\ell \\ &+ d\left\{s\rho\left[(\alpha - \beta q)\,d\ell + \left(\beta\,|\partial_\psi|^2 - \alpha q\right)d\psi - iR^2 d\zeta\right]\right\} + d\tau \wedge dH. \end{aligned} \tag{8.72}$$

8.4 Statistical Mechanics in Extended Phase Space

In this section we apply the theory developed so far to the study of the statistical behavior of an ensemble of particles moving within the extended space obtained by the Poissonization of a three dimensional antisymmetric algebra.

8.4.1 *The Jacobian of the coordinate change*

Let x, y, z be a reference system. We have seen that the equations of motion take the form:

$$X = \boldsymbol{w} \times \nabla H, \tag{8.73}$$

where $\boldsymbol{w}$ is the antisymmetric operator and H the Hamiltonian function. Extending the system to 4-dimensions $\boldsymbol{x} = (x, y, z, s)$ according to equation (8.30), we obtain an extended bracket which is measure preserving with invariant measure:

$$vol_{\boldsymbol{x}}^4 = dx \wedge dy \wedge dz \wedge ds. \tag{8.74}$$

A further time reparametrization, equation (8.32), gives a symplectic manifold $\boldsymbol{y} = (p_x, q_x, p_y, q_y)$:

$$vol_{\boldsymbol{y}}^4 = dp_x \wedge dq_x \wedge dp_y \wedge dq_y. \tag{8.75}$$

The canonical variables $\boldsymbol{y}$ are determined by the specific form of the antisymmetric operator $\boldsymbol{w}$ so that $r\Omega = dp_x \wedge dq_x + dp_y \wedge dq_y$. We want to determine the Jacobian of the coordinate change going from (8.74) to (8.75). For this purpose, we need the following:

***Lemma* 8.1.** *Let $X = d\boldsymbol{x}/dt$ and $Y = d\boldsymbol{y}/d\tau$ be two vector fields with $\boldsymbol{x} = \left(x^1, ..., x^n\right)$ and $\boldsymbol{y} = \left(y^1, ..., y^n\right)$. Let g be the Jacobian of the coordinate change $vol_{\boldsymbol{y}}^n = dy^1 \wedge ... \wedge dy^n = g^{-1} dx^1 \wedge ... \wedge dx^n = vol_{\boldsymbol{x}}^n$. If*

$$\mathfrak{L}_X vol_{\boldsymbol{x}}^n = \mathfrak{L}_Y vol_{\boldsymbol{y}}^n = 0, \tag{8.76}$$

then,

$$g = \frac{dt}{d\tau}. \tag{8.77}$$

Proof. We have:

$$\begin{aligned}
0 &= \mathfrak{L}_X vol_{\boldsymbol{x}}^n \\
&= \mathfrak{L}_X g vol_{\boldsymbol{y}}^n \\
&= d i_X g vol_{\boldsymbol{y}}^n \\
&= (-1)^{m-1} d(g\dot{y}^m) \wedge dy^1 \wedge ... \wedge dy^{m-1} \wedge dy^{m+1} \wedge ... \wedge dy^n \\
&= \frac{1}{g} \frac{\partial}{\partial y^m} \left(g \frac{d\tau}{dt} \left(y^m\right)' \right) vol_{\boldsymbol{x}}^n \\
&= \frac{d\tau}{dt} \frac{\partial \left(y^m\right)'}{\partial y^m} vol_{\boldsymbol{x}}^n + \frac{\left(y^m\right)'}{g} \frac{\partial}{\partial y^m} \left(g \frac{d\tau}{dt} \right) vol_{\boldsymbol{x}}^n \\
&= \frac{\left(y^m\right)'}{g} \frac{\partial}{\partial y^m} \left(g \frac{d\tau}{dt} \right) vol_{\boldsymbol{x}}^n \\
&= \frac{1}{g} \frac{d}{d\tau} \left(g \frac{d\tau}{dt} \right) vol_{\boldsymbol{x}}^n.
\end{aligned} \tag{8.78}$$

Here, the apex $'$ indicates derivation with respect to time τ and we used the fact that $\mathfrak{L}_Y vol_{\boldsymbol{y}}^n = 0$ if and only if $\partial_{y^m} \left(y^m\right)' = 0$. The solution is, up to constants, $g = dt/d\tau$. ■ □

Applying lemma 8.1 to the specific case $\boldsymbol{x} = (x, y, z, s)$ and $\boldsymbol{y} = (p_x, q_x, p_y, q_y)$ we conclude that the Jacobian g of the coordinate change is:

$$g = \frac{dt}{d\tau} = r^{-1} = \frac{1}{\boldsymbol{w} \cdot (\boldsymbol{D} + s\nabla \times \boldsymbol{w})}. \tag{8.79}$$

In terms of the volume elements:

$$dx \wedge dy \wedge dz \wedge ds = \frac{dp_x \wedge dq_x \wedge dp_y \wedge dq_y}{\boldsymbol{w} \cdot (\boldsymbol{D} + s\nabla \times \boldsymbol{w})}. \tag{8.80}$$

Thus, the Jacobian of the coordinate change from the initial (extended) coordinates to the canonical phase space obtained by Poissonization is controlled by the Jacobiator, i.e. by the measure of the failure of the Jacobi identity. Indeed, recalling definition 4.1, we have:

$$\mathfrak{G} = \mathfrak{G}^{xyz} \partial_x \wedge \partial_y \wedge \partial_z = (\boldsymbol{w} \cdot \nabla \times \boldsymbol{w})\, \partial_x \wedge \partial_y \wedge \partial_z. \tag{8.81}$$

This implies $g^{-1} = \boldsymbol{w} \cdot \boldsymbol{D} + s\mathfrak{G}^{xyz}$.

8.4.2 *The distribution function and thermal equilibrium*

Let $P = P(\boldsymbol{y}, \tau)$ be the distribution function of an ensemble of particles on the canonical phase space $vol^4_{\boldsymbol{y}}$ at time τ. The value of the distribution function measures the probability of finding a particle in the unit volume at a given time. We want to know how the distribution function P is seen in the initial coordinates $vol^4_{\boldsymbol{x}}$. Using the result of equation (8.79), we have:

$$P vol^4_{\boldsymbol{y}} = P r vol^4_{\boldsymbol{x}}, \tag{8.82}$$

which implies that the distribution function $f(\boldsymbol{x}, \tau)$ on $vol^4_{\boldsymbol{x}}$ at the time τ is related to P as:

$$f = Pr = P\boldsymbol{w} \cdot (\boldsymbol{D} + s\nabla \times \boldsymbol{w}) = P(\boldsymbol{w} \cdot \boldsymbol{D} + s\mathfrak{G}^{xyz}). \tag{8.83}$$

From the result above we see that the distortion between f and P is dictated by the Jacobiator (8.81). Furthermore, by integrating over the variable s, we can calculate the shape of the distribution $\mathcal{F}(x, y, z, \tau)$ in the initial coordinates (x, y, z):

$$\mathcal{F} = \int f ds = \boldsymbol{w} \cdot \boldsymbol{D} \int P ds + \mathfrak{G}^{xyz} \int P s ds. \tag{8.84}$$

Let us make some considerations on thermal equilibrium. Since $vol^4_{\boldsymbol{y}}$ is the preserved volume element of a symplectic manifold spanned by canonical variables, we can exploit the usual formulation of statistical mechanics and define the differential entropy Σ of the distribution function P:

$$\Sigma = -\int_{V_{\boldsymbol{y}}} P \log P vol^4_{\boldsymbol{y}}. \tag{8.85}$$

Here the integral is performed on the whole phase space $V_{\boldsymbol{y}}$. The total number of particles and the total energy E of the ensemble are given by $N = \int_{V_{\boldsymbol{y}}} P vol^4_{\boldsymbol{y}}$ and $E = \int_{V_{\boldsymbol{y}}} H P vol^4_{\boldsymbol{y}}$ respectively. The form of the distribution function at equilibrium $P_{eq} = P(\tau \to \infty)$ is calculated my maximizing the entropy Σ under the constraints N and E with the variational principle:

$$\delta(\Sigma - \alpha N - \beta E) = 0. \tag{8.86}$$

Here α and β are the Lagrange multipliers associated to N and E. The result of the variation is:

$$P_{eq} = \frac{1}{Z} e^{-\beta H}. \tag{8.87}$$

In the above equation Z is a normalization constant. Thus, recalling equations (8.83) and (8.84), we arrive at the following formulas for the equilibrium $f_{eq} = f(\tau \to \infty)$ and $\mathcal{F}_{eq} = \mathcal{F}(\tau \to \infty)$ in the initial coordinates:

$$f_{eq} = \frac{r}{Z} e^{-\beta H} = \frac{1}{Z}\left(\boldsymbol{w} \cdot \boldsymbol{D} + s\mathfrak{G}^{xyz}\right) e^{-\beta H}, \tag{8.88a}$$

$$\mathcal{F}_{eq} = \int f_{eq} ds = \frac{\Delta s}{Z}\left(\boldsymbol{w} \cdot \boldsymbol{D} + \frac{\Delta s}{2}\mathfrak{G}^{xyz}\right) e^{-\beta H}. \tag{8.88b}$$

Here $\Delta s = \int ds$. The conclusion is that the thermal equilibrium of the extended system departs from the standard thermal equilibrium with homogeneous probability density on constant energy surfaces. The distortion is controlled by the Jacobiator, i.e. by the measure of the failure of the Jacobi identity.

8.4.3 *An example: thermal equilibrium by $\boldsymbol{E} \times \boldsymbol{B}$ drift in a magnetic field*

In this section we give a concrete example of how the theory developed so far can be applied to predict the thermal equilibrium of an ensemble resulting from the Poissonization of an antisymmetric algebra. We consider an ensemble of magnetized particles moving by $\boldsymbol{E}\times\boldsymbol{B}$ drift according to equation (8.2). The magnetic field $\boldsymbol{B}$ is assumed to be of the form:

$$\boldsymbol{B}=\partial_x+\left(\frac{y-\sin y\cos y}{2}-\sin x\right)\partial_z. \tag{8.89}$$

One can verify that $\mathcal{B}=d\left[\left(\frac{\sin^2 y-y^2}{4}+y\sin x\right)dx+ydz\right]$. Recalling that the antisymmetric operator is $\boldsymbol{w}=\boldsymbol{B}/B^2$, we have:

$$\boldsymbol{w}=\frac{\partial_x+\left(\frac{y-\sin y\cos y}{2}-\sin x\right)\partial_z}{1+\left(\frac{y-\sin y\cos y}{2}-\sin x\right)^2}, \tag{8.90}$$

and also,

$$\mathfrak{G}^{xyz}=\boldsymbol{w}\cdot\nabla\times\boldsymbol{w}=\frac{\boldsymbol{B}\cdot\nabla\times\boldsymbol{B}}{B^4}=\frac{\sin^2 y}{\left[1+\left(\frac{y-\sin y\cos y}{2}-\sin x\right)^2\right]^2}. \tag{8.91}$$

A typical scenario encountered in magnetized plasmas is quasi-neutrality, already discussed in the case of magnetospheric self-organization. In such situation, the time average of the electric potential $\phi(\tau)$ generating the electric field is zero, i.e. $\bar{\phi}=\lim_{T\to\infty}T^{-1}\int_0^T\phi d\tau=0$. Therefore, the Hamiltonian of each massless particle is itself zero $H=\bar{\phi}=0$. However, the random fluctuations in ϕ generated by the electromagnetic interactions among the charged particles drive the ensemble toward equilibrium, which according to (8.88b) is:

$$\mathcal{F}_{eq}=\frac{\Delta s}{Z}\left\{1+\frac{\Delta s}{2}\frac{\sin^2 y}{\left[1+\left(\frac{y-\sin y\cos y}{2}-\sin x\right)^2\right]^2}\right\}. \tag{8.92}$$

Here we used equation (8.91) and set $\boldsymbol{D}=\boldsymbol{B}=\boldsymbol{w}/w^2$ as required in the case of $\boldsymbol{E}\times\boldsymbol{B}$ drift. Figure 8.3 shows a plot of the predicted thermal equilibrium. The shape of the distribution sensibly departs from the flat profile one would expect by a naive application of the entropy principle in the initial noncanonical coordinates. This discrepancy is a consequence of the failure of the Jacobi identity.

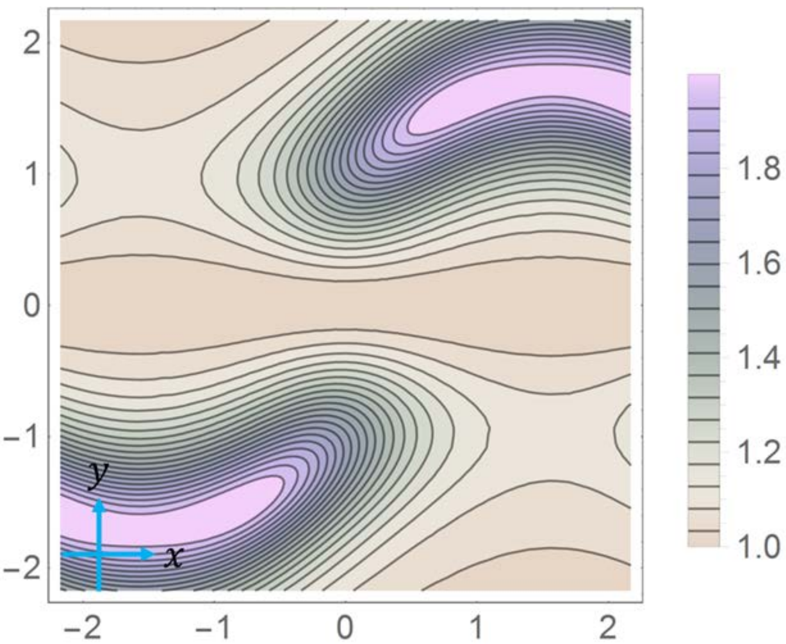

Figure 8.3: Thermal equilibrium $\mathcal{F}_{eq}(x, y)$ by $\boldsymbol{E} \times \boldsymbol{B}$ drift in the magnetic field (8.89). The inhomogeneous distribution is caused by the failure of the Jacobi identity.

Chapter 9

Non-Elliptic Diffusion in Three Dimensions

Here, the theory developed in chapters 4, 5, and 6 is put to the test of numerical simulations. This is done by comparing the analytical solution to the Fokker-Planck equation for the probability distribution (predicted by the results of chapter 6) with the direct integration of the stochastic differential equation governing the corresponding ensemble of particles endowed with the relevant antisymmetric operator. For the sake of simplicity, we limit our attention to the 3-dimensional case.

9.1 Constrained orbits

The equation of motion for a 3-dimensional system is of the form:

$$X = \boldsymbol{w} \times \nabla H_0, \tag{9.1}$$

with $\boldsymbol{w}$ the antisymmetric operator and H_0 the Hamiltonian of the system. It is useful to make qualitative considerations on how the orbit of a conservative particle obeying (9.1) is modified by the introduction of random noise. First, consider again the Euler rigid body with equations of motion given by (8.17). In this case $\boldsymbol{w} = \boldsymbol{x}$ is a Poisson operator because its Jacobiator $\mathfrak{G} = \mathfrak{G}^{xyz}\partial_x \wedge \partial_y \wedge \partial_z$ (remember definition 4.1) vanishes:

$$\mathfrak{G}^{xyz} = \boldsymbol{x} \cdot \nabla \times \boldsymbol{x} = 0. \tag{9.2}$$

In fact, we have already seen that $C = \boldsymbol{x}^2/2$ is a Casimir invariant. The unperturbed orbit of the rigid body, given by the intersection of the integral surfaces $H_0 = \left(x^2 I_x^{-1} + y^2 I_y^{-1} + z^2 I_z^{-1}\right)/2$ and C, is given in figure 9.1 (a). Now, we perturb the Hamiltonian H_0 so that the force acting on the particle becomes $\nabla H = \nabla H_0 + \boldsymbol{\Gamma}$, where $\boldsymbol{\Gamma} = (\Gamma_x, \Gamma_y, \Gamma_z)$ is 3-dimensional white noise (see figure 9.2). The resulting stochastic differential equation is:

$$X = \boldsymbol{x} \times \left(x^i I_{x^i}^{-1} \partial_i + \boldsymbol{\Gamma}\right). \tag{9.3}$$

Clearly, the energy H_0 is not anymore a constant of motion. However, the Casimir invariant C is unaffected by the perturbations. The result is a random process on the level set C =constant (see figure 9.1 (b)). This is exactly what happened in chapter 7 for the self-organization of a radiation belt where the inward diffusion occurred on the Casimir leaf μ=constant.

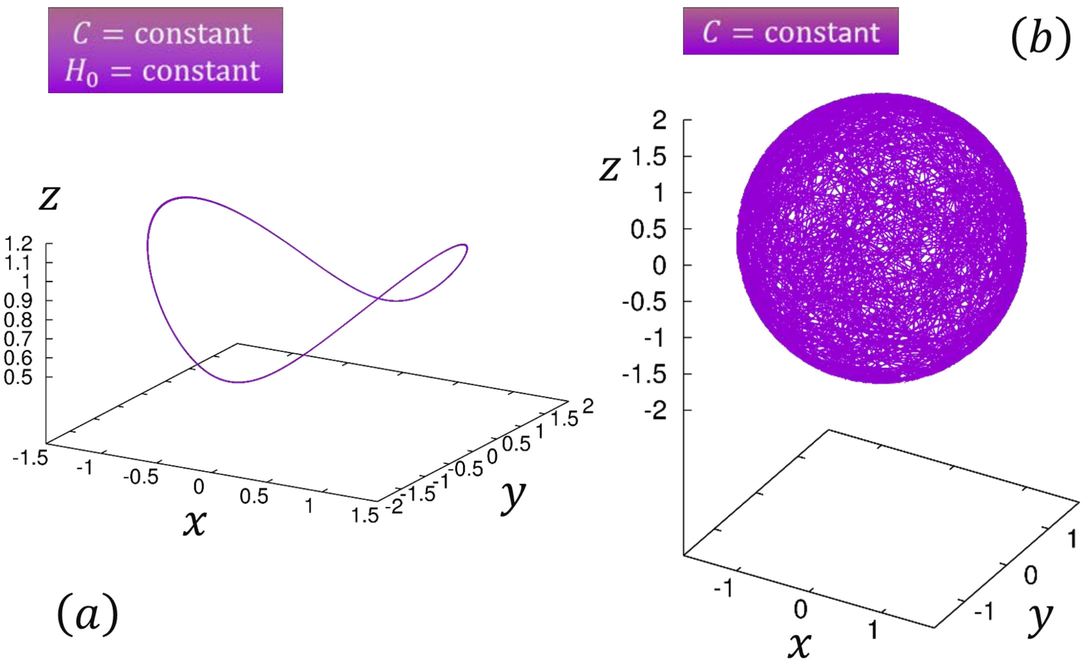

Figure 9.1: (a): numerical integration of (8.17). The orbit is the intersection of the surfaces C and H_0. (b): numerical integration of (9.3). If the Hamiltonian is perturbed $\nabla H = \nabla H_0 + \boldsymbol{\Gamma}$, the particle explores the surface C.

Next, consider the Beltrami operator $\boldsymbol{w} = (\cos z - \sin y, -\sin z, \cos y)$ with the same Hamiltonian H_0:

$$X = (\cos z - \sin y, -\sin z, \cos y) \times \left(x^i I_{x^i}^{-1} \partial_i\right). \tag{9.4}$$

One can check that $\mathfrak{G}^{xyz} = \boldsymbol{w}^2$ so that no Casimir invariant exists. The unperturbed orbit is shown in figure 9.3 (a). This time the trajectory is spiraling above the energy surface H_0. The absence of an invariant measure is also manifest. Again, perturb the Hamiltonian as $\nabla H = \nabla H_0 + \boldsymbol{\Gamma}$:

$$X = (\cos z - \sin y, -\sin z, \cos y) \times \left(x^i I_{x^i}^{-1} \partial_i + \boldsymbol{\Gamma}\right). \tag{9.5}$$

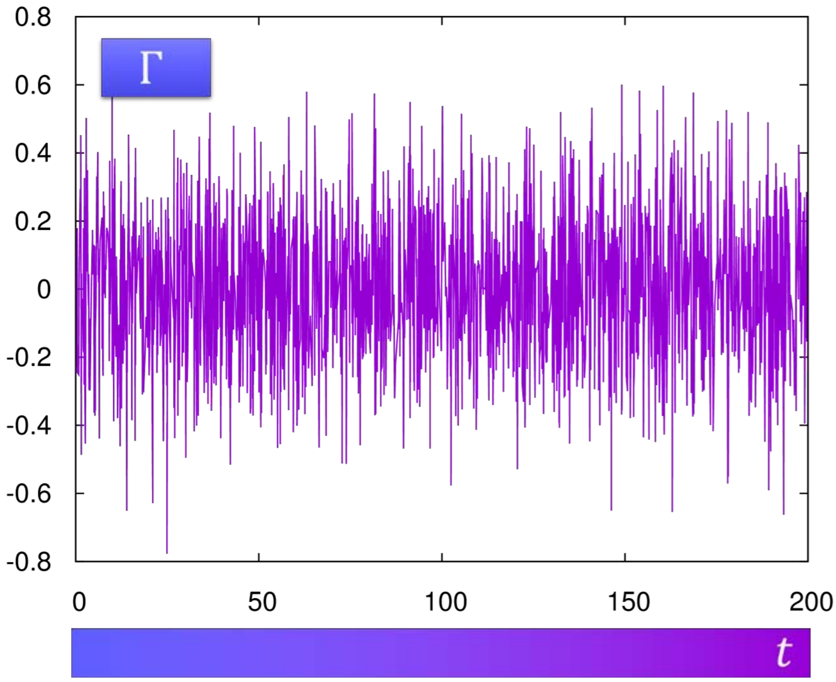

Figure 9.2: A typical numerical representation of the white noise process Γ as function of time t. Amplitude and time are given in arbitrary units.

The resulting orbit is shown in figure 9.3 (b). Notice that no integral surface exists anymore.

The remarkable fact about antisymmetric operators is that, even when they do not impart integrable constraints to the dynamics, the apparently disordered motion resulting by the breaking of the energy integral H_0 may hide an ordered macroscopic structure. The first step to see this is to superimpose the orbits of (statistically) many particles. As an example consider the Poisson operator:

$$\boldsymbol{w} = \left(\sqrt{1+\cos x^2}\right) \nabla \left(z - \cos x - \cos y\right). \tag{9.6}$$

To further simplify the problem, we take the purely random Hamiltonian H such that $\nabla H = \boldsymbol{\Gamma}$. The stochastic equation of motion becomes:

$$X = \left(\sqrt{1+\cos x^2}\right) \nabla \left(z - \cos x - \cos y\right) \times \boldsymbol{\Gamma}. \tag{9.7}$$

The superposition of the corresponding orbits of an ensemble of 250 particles is shown in figure 9.4. Notice that, while each orbit may individually appear as purely stochastic, vortices arise on the macroscopic level.

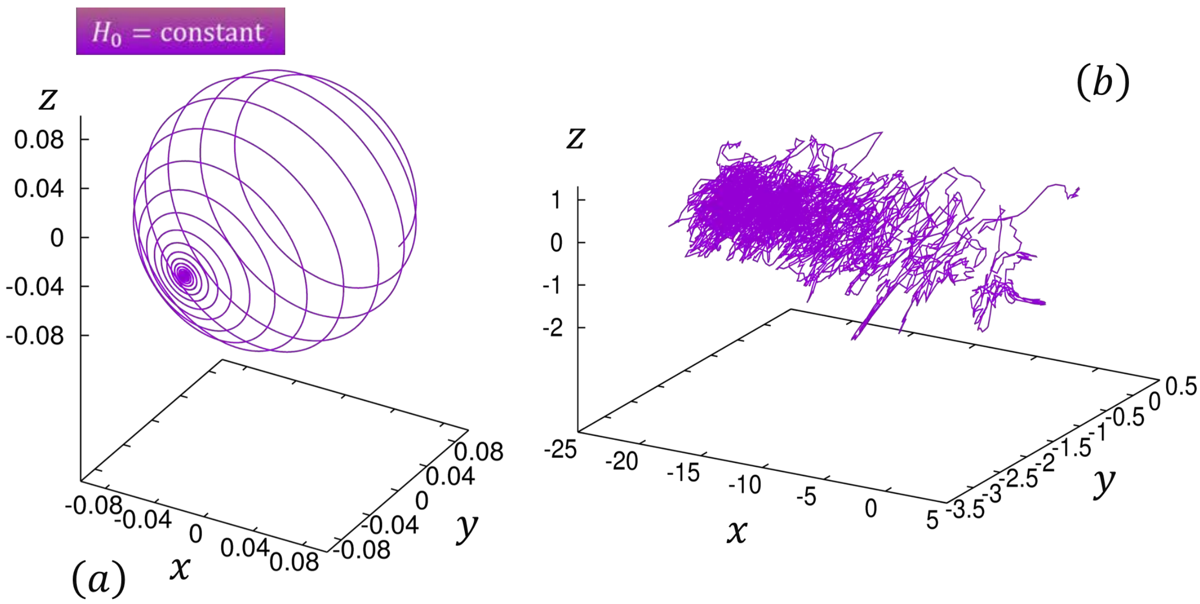

Figure 9.3: (a): numerical integration of (9.4). The orbit explores the energy surface H_0 and falls toward a sink. (b): numerical integration of (9.5). If the Hamiltonian is perturbed $\nabla H = \nabla H_0 + \mathbf{\Gamma}$, there are no integral manifolds.

9.2 The Diffusion Equation in Three Dimensions

In this section the results of numerical simulations are presented. We integrate the stochastic equation of motion:

$$X = \boldsymbol{w} \times \mathbf{\Gamma}, \tag{9.8}$$

for different choices of $\boldsymbol{w}$. In each simulation an ensemble of $8 \cdot 10^6$ particles is considered. The trajectory of each particle is tracked for the same period of time. A numerical probability distribution is obtained and compared with the expected stationary analytical form. Except when differently specified, the computational domain is a cube in (x, y, z) space with sides of size 6 and centered at $\boldsymbol{x} = \mathbf{0}$. The boundary conditions are periodic (except when differently specified) with the period given by the sides of the cube. The initial condition is a flat (or Gaussian when so specified) probability distribution as shown in figure 9.5.

The purely diffusive Fokker-Planck equation associated to (9.8) is given by the non-elliptic second order partial differential equation (6.48) with stationary form:

$$0 = \nabla \cdot \left[\boldsymbol{w} \times (\nabla \times f\boldsymbol{w})\right]. \tag{9.9}$$

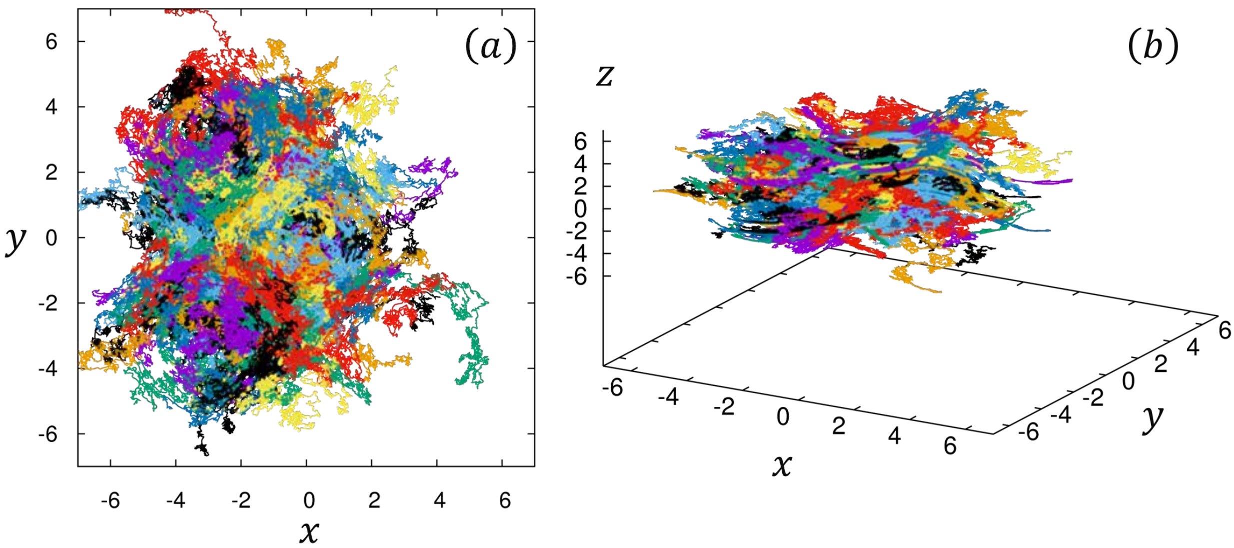

Figure 9.4: (a): Superposition of 250 sample paths generated by integration of (9.7) in the (x, y) plane. (b): Superposition of 250 sample paths generated by integration of (9.7).

The geometrical meaning of this equation can be made explicit by rewriting it in terms of field force divergence and field force vector $\mathfrak{B} = 4\nabla \cdot \boldsymbol{b}$ and $\boldsymbol{b} = \boldsymbol{w} \times (\nabla \times \boldsymbol{w})$ of $\boldsymbol{w}$ (remember equation (4.41)). We have:

$$0 = \Delta_{\perp} f + \nabla f \cdot \boldsymbol{b} + \frac{1}{4} f \mathfrak{B}. \tag{9.10}$$

Here, we introduced the *normal Laplace operator* $\Delta_{\perp} f = \nabla \cdot [\boldsymbol{w} \times (\nabla f \times \boldsymbol{w})]$. This novel differential operator is clearly non-elliptic because the component of ∇f aligned with $\boldsymbol{w}$ does not contribute to its value. The normal Laplace operator will be discussed from the standpoint of functional analysis in the next chapter, where we will show that equations of the type $\Delta_{\perp} f = \phi$ admit a weak and unique solution as long as $\boldsymbol{w}$ is not integrable. Since we already know the nature of the solution to (9.9) when $\boldsymbol{w}$ is integrable, a Beltrami operator, or $\hat{\boldsymbol{b}} = \nabla \zeta$ (see theorem 6.3), this result mathematically justifies and represents the fundamental step toward the determination of the general solution to (9.9).

9.2.1 *Constant operator*

The simplest possible situation is given by a constant operator $\boldsymbol{w}$. We choose $\boldsymbol{w} = \partial_z$. Obviously, the Jacobiator identically vanishes because $\nabla \times \boldsymbol{w} = \boldsymbol{0}$. Therefore, such $\boldsymbol{w}$ is a Poisson operator. The resulting dynamics $X = \partial_z \times \boldsymbol{\Gamma}$ can be thought as

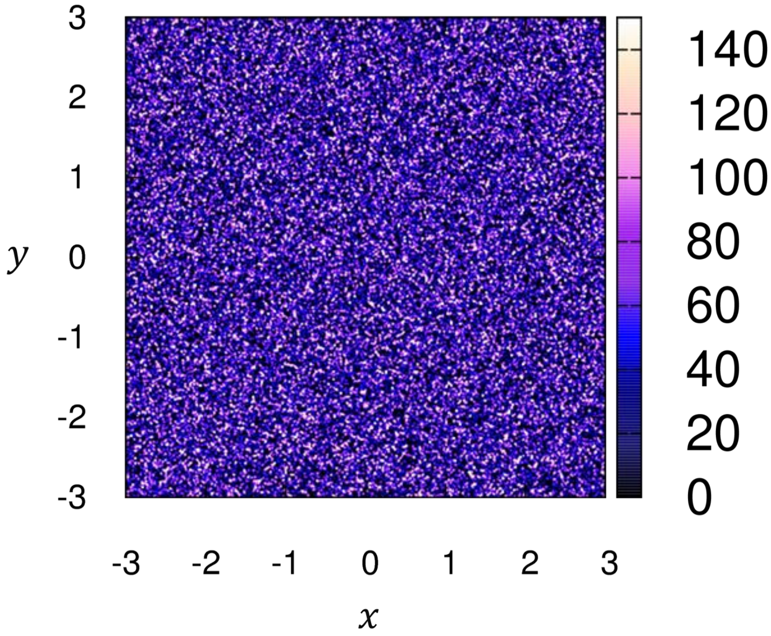

Figure 9.5: The initial flat probability distribution on the slice $z = 0$ of the cubic computational domain.

the $\boldsymbol{E} \times \boldsymbol{B}$ motion of a charged particle in a constant magnetic field $B = w^{-1} = 1$ (remember that in the case of $\boldsymbol{E} \times \boldsymbol{B}$ drift $\boldsymbol{w} = \boldsymbol{B}/B^2$). It is also clear that the coordinate system $dx \wedge dy \wedge dz$ is an invariant measure for any choice of the Hamiltonian function (the operator $\boldsymbol{w}$ is measure preserving). This can be verified by showing that the cocurrent $n-1$ form $\mathcal{O}^{n-1}$ vanishes on this volume form:

$$\mathcal{O}^{n-1} = d\mathcal{J}^{n-2} = 2d\left(\mathcal{J}^{xy}dz + \mathcal{J}^{yz}dx + \mathcal{J}^{zx}dy\right) = -2d\left(w^i dx^i\right) = -2ddz = 0. \tag{9.11}$$

Here we used equations (4.11) and (4.32). The analytical form of the equilibrium probability distribution is then determined by corollary 6.1. In our case $H_0 = 0$ and the kernel of $\boldsymbol{w}$ is spanned by the vector ∂_z itself. Such kernel is also integrable in terms of an arbitrary function (Casimir invariant) $C(z)$ of the coordinate z, since ∇C is aligned with ∂_z. Then, we expect f to be of the type:

$$\lim_{t\to\infty} f = A \exp\left\{-\gamma C(z)\right\} \qquad a.e., \tag{9.12}$$

Here, A and γ are positive real constants. Furthermore, since the initial distribution is flat, the diffusion process $X = \partial_z \times \boldsymbol{\Gamma}$, which is constrained in the (x, y) plane, cannot generate any inhomogeneity in the ∂_z direction. In conclusion, f must remain constant throughout the simulation. The result of the numerical simulation is shown in figure 9.6.

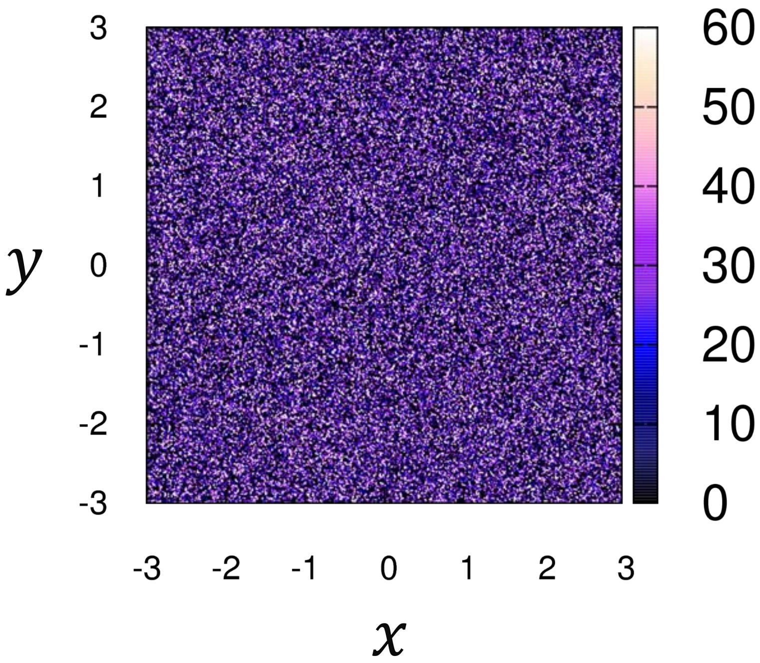

Figure 9.6: Calculated equilibrium probability distribution f in the (x, y) plane at $z = 0$ with constant Poisson operator $\boldsymbol{w} = \partial_z$. The initial condition at $t = 0$ is the flat distribution of figure 9.5. Observe that the distribution remains flat.

In figure 9.7 we report the result of the numerical simulation corresponding to a different initial condition, a Gaussian distribution in the (x, y) plane.

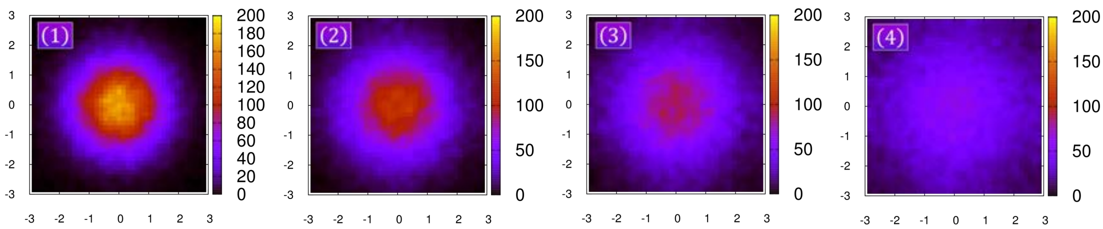

Figure 9.7: Time evolution of the probability distribution f in the (x, y) plane at $z = 0$ with constant Poisson operator $\boldsymbol{w} = \partial_z$. The initial condition at $t = 0$ is the Gaussian distribution in the (x, y) plane of figure (1). Each plot number i corresponds to the state of time evolution $t = i\Delta t$, where Δt is a fixed time interval. The particle sample is $\sim 10^5$. Observe that the distribution converges to a flat profile.

9.2.2 *Poisson operator on an invariant measure*

Next, we consider the following Poisson operator:

$$\boldsymbol{w} = \nabla C = \nabla \left(z - \cos x - \cos y \right). \tag{9.13}$$

Again, the Jacobi identity $\mathfrak{S} = 0$ is identically satisfied because $\nabla\times\boldsymbol{w} = \nabla\times\nabla C = \boldsymbol{0}$. If we interpret the resulting dynamics as the motion of a charged particle in the magnetic field $\boldsymbol{B} = \boldsymbol{w}/w^2$, the magnetic field strength is:

$$B = \left(1 + \sin x^2 + \sin y^2\right)^{-1/2}. \tag{9.14}$$

See figure 9.8 for the plot of B.

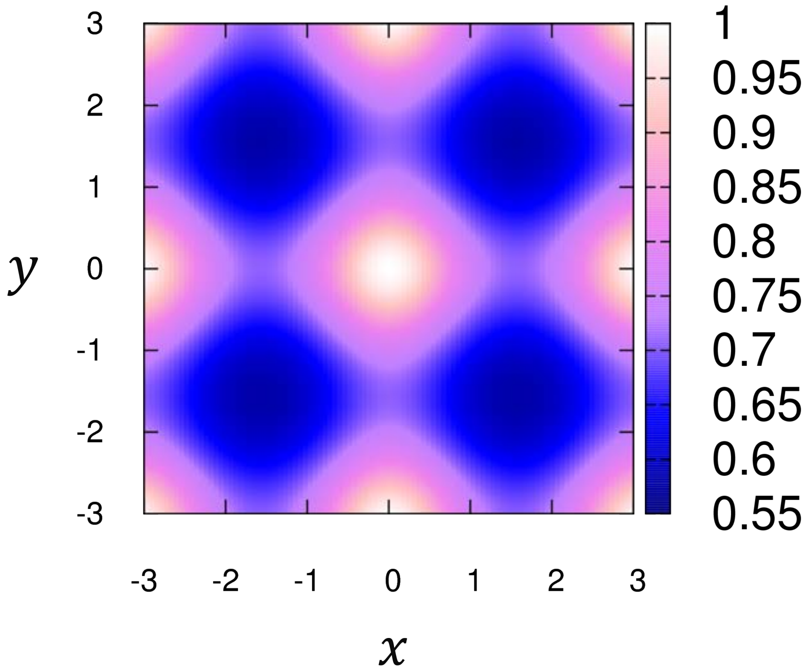

Figure 9.8: Magnetic field strength (9.14) in the (x, y) plane.

This time the Casimir invariant whose gradient spans the kernel of $\boldsymbol{w}$ is the function $C = z - \cos x - \cos y$. Using proposition 4.6, we also know that $dx \wedge dy \wedge dz$ is an invariant measure for any choice of the Hamiltonian function (i.e. $\boldsymbol{w}$ is measure preserving). Therefore, in light of corollary 6.1 we expect the equilibrium probability distribution to be of the type:

$$\lim_{t\to\infty} f = A \exp\{-\gamma \mathcal{F}(C)\} \quad a.e., \tag{9.15}$$

where A and γ are positive real constants, and $\mathcal{F}$ an arbitrary function of the Casimir invariant C. Since the initial distribution is flat in the whole (x, y, z) space, and the diffusion process $\nabla C \times \mathbf{\Gamma}$ flattens the distribution on each level set C=constant, no inhomogeneity can arise along ∇C. More precisely, the 2-dimensional volume form vol_C^2 on each reduced space $\mathbb{R}^3/C$ is itself an invariant measure for any choice of the Hamiltonian function because $dx \wedge dy \wedge dz = dC \wedge vol_C^2$ is a preserved volume element and C is a dynamical constant: $\mathfrak{L}_X \left(dx \wedge dy \wedge dz\right) = dC \wedge \mathfrak{L}_X vol_C^2 = 0$ $\forall H_0$. It follows that, if $X_C^2 = \mathcal{J}_C^2 \left(dH_0\right)$ is the 2-dimensional flow on the level set C =constant, the reduced operator $\mathcal{J}_C^2$ is measure preserving. Therefore, to each distribution f_C on the volume element vol_C^2 on each leaf C =constant the result of corollary 6.1 applies: $\lim_{t\to\infty} f_C$ =constant, exception made for a set of measure zero. In conclusion, the three dimensional distribution f must remain constant throughout the simulation. Figure 9.9 shows the results of the numerical simulation. In particular, notice that the distribution remains flat regardless of the fact that the random process under consideration is spatially inhomogeneous: $|X| = |\nabla C| = w = B^{-1}$.

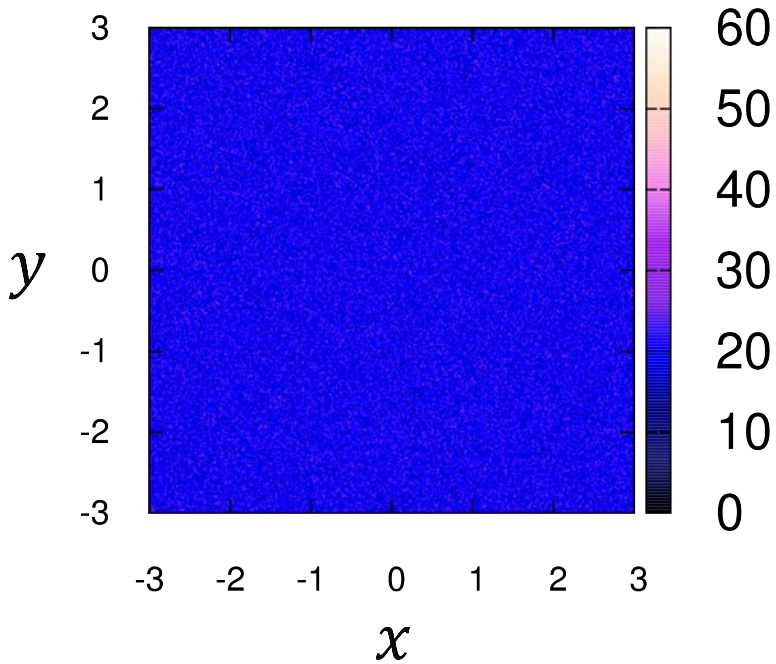

Figure 9.9: Calculated equilibrium probability distribution f in the (x, y) plane at $z = 0$ with Poisson operator (9.13). The initial condition $t = 0$ is the flat distribution of figure 9.5. Observe that the distribution remains flat.

In figure 9.10 we report the result of the numerical simulation corresponding to a different initial condition, a Gaussian distribution in the (x, y) plane.

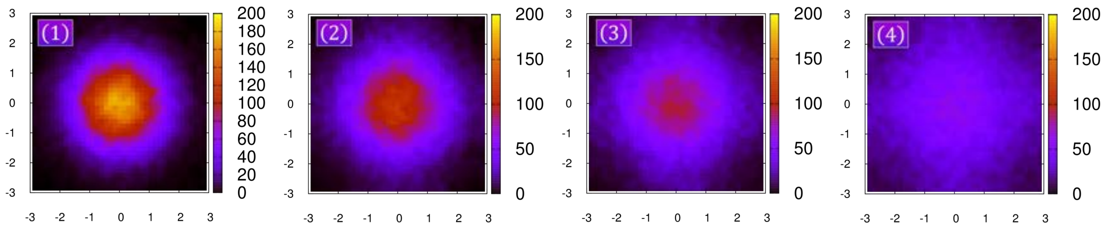

Figure 9.10: Time evolution of the probability distribution f in the (x, y) plane at $z = 0$ with Poisson operator (9.13). The initial condition at $t = 0$ is the Gaussian distribution in the (x, y) plane of figure (1). Each plot number i corresponds to the state of time evolution $t = i\Delta t$, where Δt is a fixed time interval. The particle sample is $\sim 10^5$. Observe that the distribution converges to a flat profile.

9.2.3 *Poisson operator in arbitrary coordinates*

Consider now the Poisson operator:

$$\boldsymbol{w} = \lambda \nabla C = \left(\sqrt{1 + \cos x^2}\right) \nabla \left(z - \cos x - \cos y\right). \tag{9.16}$$

Here $\lambda = \sqrt{1 + \cos x^2} \neq 0$ and $C = z - \cos x - \cos y$. The Jacobi identity is easily verified: $\mathfrak{G}^{xyz} = \lambda \nabla C \cdot \nabla \times \lambda \nabla C = 0$. Furthermore, C is evidently a Casimir invariant. The corresponding magnetic field strength:

$$B = \frac{1}{\sqrt{\left(1 + \cos^2 x\right)\left(1 + \sin^2 x + \sin^2 y\right)}}, \tag{9.17}$$

is shown in figure 9.11.

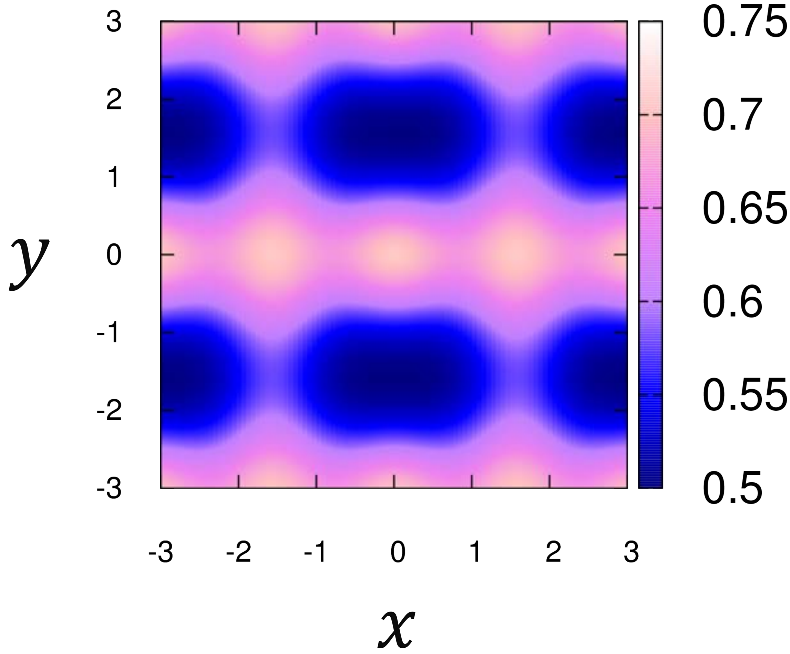

Figure 9.11: Magnetic field strength (9.17) in the (x, y) plane.

According to proposition 4.6, this time the invariant measure is given by the volume element $\lambda^{-1}dx \wedge dy \wedge dz$. In light of corollary 6.1 the probability distribution f on our coordinate system $dx \wedge dy \wedge dz$ must satisfy:

$$\lim_{t\to\infty} f = \frac{A}{\lambda}\exp\{-\gamma\mathcal{F}(C)\} \quad a.e. \tag{9.18}$$

Here A and γ are positive real constants and $\mathcal{F}$ is an arbitrary function of the Casimir invariant C. Applying the same reasoning of the previous case on the invariant measure $\lambda^{-1}dx \wedge dy \wedge dz$, since the initial distribution is spatially flat, we expect the solution to converge to $f \propto \lambda^{-1}$. Figure 9.12 shows the density plot of λ^{-1}. Figure 9.13 shows the result of the numerical simulation.

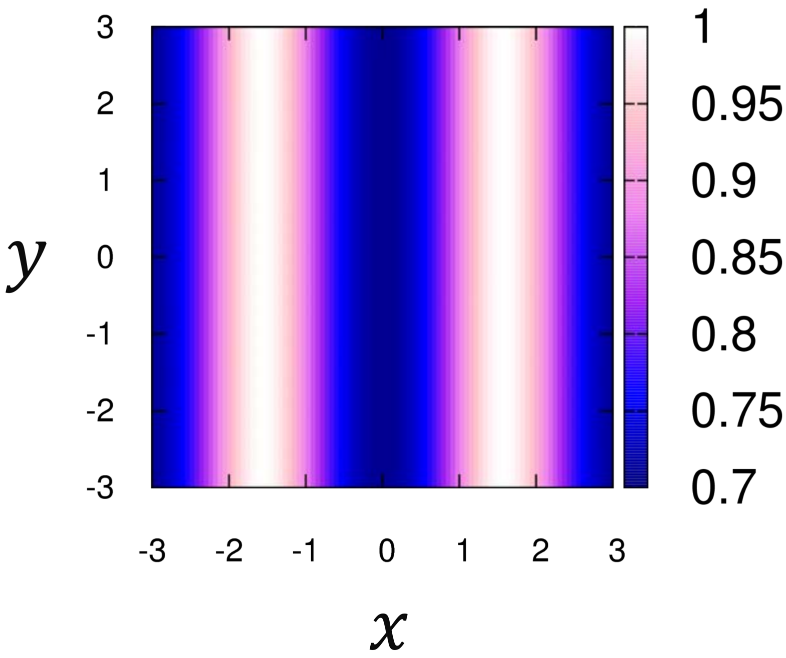

Figure 9.12: Spatial profile of λ^{-1} in the (x, y) plane.

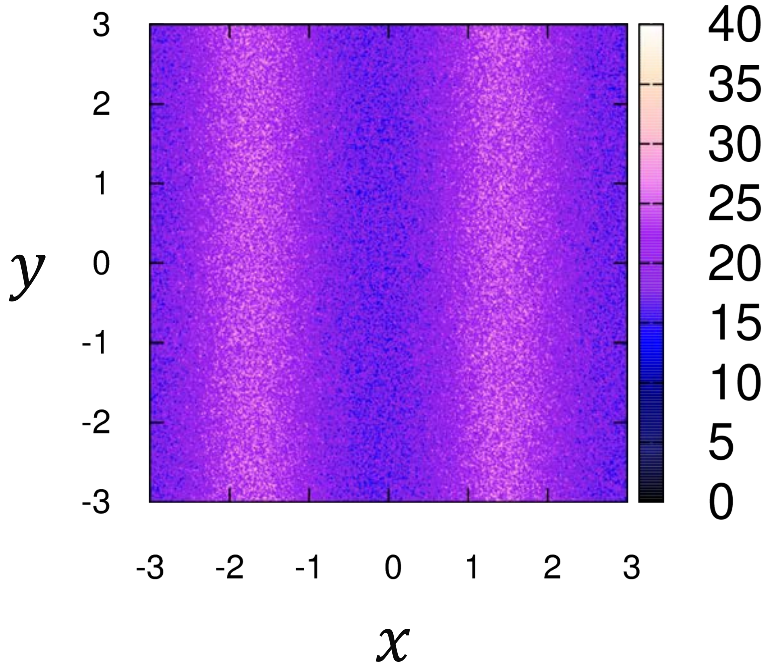

Figure 9.13: Calculated equilibrium probability distribution f in the (x, y) plane at $z = 0$ with Poisson operator (9.16). The initial condition $t = 0$ is the flat distribution of figure 9.5. Observe that the distribution converges to the profile $f \propto \lambda^{-1}$.

In figure 9.14 we report the result of the numerical simulation corresponding to a different initial condition, a Gaussian distribution in the (x, y) plane.

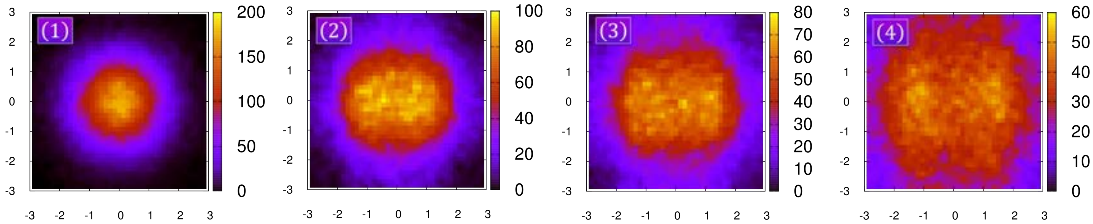

Figure 9.14: Time evolution of the probability distribution f in the (x, y) plane at $z = 0$ with Poisson operator (9.16). The initial condition at $t = 0$ is the Gaussian distribution in the (x, y) plane of figure (1). Each plot number i corresponds to the state of time evolution $t = i\Delta t$, where Δt is a fixed time interval. The particle sample is $\sim 10^5$. Observe that the distribution converges to the profile $f \propto \lambda^{-1}$.

9.2.4 *Beltrami operator*

The next case we study is that of a Beltrami operator:

$$\boldsymbol{w} = (\cos z + \sin z)\,\partial_x + (\cos z - \sin z)\,\partial_y. \tag{9.19}$$

One can verify that the Jacobiator is $\mathfrak{S}^{xyz} = \boldsymbol{w}^2 = 2 \neq 0$. Therefore, this operator is not a Poisson operator. Furthermore, the field force vector is $\boldsymbol{b} = \boldsymbol{w} \times \nabla \times \boldsymbol{w} = \boldsymbol{w} \times \boldsymbol{w} = \boldsymbol{0}$. This means that $\boldsymbol{w}$ is a strong Beltrami operator. Notice that the corresponding magnetic field strength is constant: $B = w^{-1} = 1/\sqrt{2}$. In this case the statement of theorem 6.2 applies:

$$\lim_{t\to\infty} \boldsymbol{w} \times \nabla f = \boldsymbol{0} \qquad a.e. \tag{9.20}$$

Observe that, since $\boldsymbol{w}$ does not satisfy the Jacobi identity, there is no Casimir invariant C whose gradient spans the kernel of the operator. That is, if we could satisfy $\nabla f = \alpha \boldsymbol{w}$ *a.e.* for some function $\alpha \neq 0$, this would contradict non-integrability of $\boldsymbol{w}$. Therefore, equation (9.20) implies $\nabla f = 0$ *a.e.*. In the next chapter we will make this reasoning more rigorous, and show that due to the non-integrability of $\boldsymbol{w}$, we must have $f = c$ on Ω, with $c \in \mathbb{R}_{>0}$. This is confirmed by the result of the numerical simulation, figure 9.15.

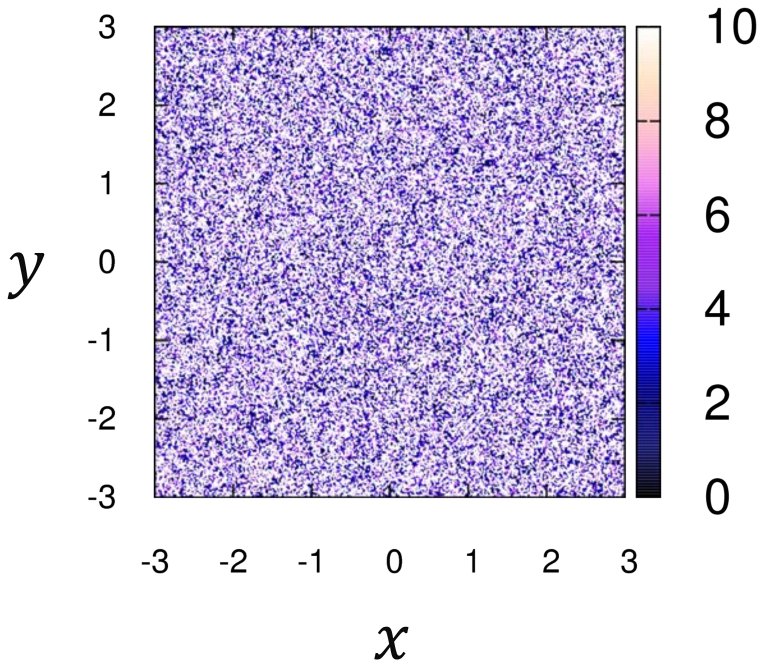

Figure 9.15: Calculated equilibrium probability distribution f in the (x, y) plane at $z = 0$ with Beltrami operator (9.19). The initial condition $t = 0$ is the flat distribution of figure 9.5. Observe that the distribution remains flat.

In figure 9.16 we report the result of the numerical simulation corresponding to a different initial condition, a constant distribution along the diagonal of the (x, y) plane.

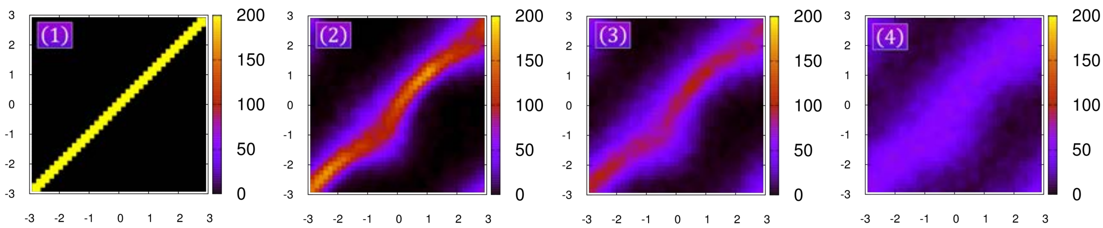

Figure 9.16: Time evolution of the probability distribution f in the (x, y) plane at $z = 0$ with Beltrami operator (9.19). The initial condition at $t = 0$ is the constant distribution along the diagonal of the the (x, y) plane of figure (1). Each plot number i corresponds to the state of time evolution $t = i\Delta t$, where Δt is a fixed time interval. The particle sample is $\sim 10^5$. Observe that the distribution converges to a flat profile.

9.2.5 *Antisymmetric operator*

Consider the antisymmetric operator:

$$\boldsymbol{w} = \partial_x + (\sin x + \cos y)\,\partial_y + (\cos x)\,\partial_z. \tag{9.21}$$

The Jacobiator is:

$$\mathfrak{G}^{xyz} = 1 + \sin x \cos y \geq 0. \tag{9.22}$$

Observe that $\mathfrak{G} \neq 0$ exception made for the set of measure zero $\{(x,y) \in \mathbb{R}^2 : \sin x \cos y = -1\}$, that is the Jacobi identity is violated almost everywhere. Furthermore, the field force divergence is given by:

$$\mathfrak{B} = -4 \sin x \cos y, \tag{9.23}$$

which is different from zero exception made for the set of measure zero $\{(x,y) \in \mathbb{R}^2 : \sin x \cos y = 0\}$. Therefore, this operator is neither a Poisson operator, nor a Beltrami operator in the chosen coordinate system. The corresponding magnetic field strength is:

$$B = w^{-1} = \frac{1}{\sqrt{1 + (\sin x + \cos y)^2 + \cos^2 x}}. \tag{9.24}$$

A density plot of B is given in figure 9.17.

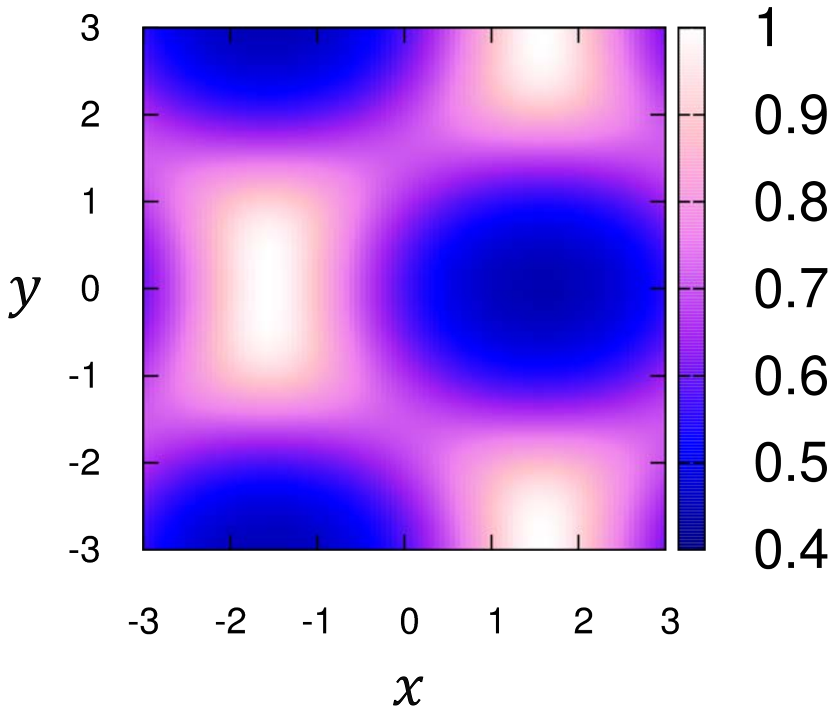

Figure 9.17: Magnetic field strength (9.24) in the (x, y) plane.

The nature of the stationary solution to equation (9.10) in the case of an antisymmetric operator like (9.21) is the motivation behind the theory developed in the next chapter. The result of the corresponding numerical simulation is given in figure 9.18. Notice that there is a strong similarity between the profile of magnetic field strength B and that of the equilibrium probability distribution f. This fact should be compared with the analytic result obtained in theorem 6.3 for the special class of antisymmetric operators such that $\hat{\boldsymbol{b}} = \nabla\zeta$ where $f \propto w^{-1}e^{-\zeta} = Be^{-\zeta}$ (although the operator (9.21) is such that $\hat{\boldsymbol{b}} \neq \nabla\zeta$). The tendency of the equilibrium probability distribution to approach the shape of the magnetic field is observed also in the next simulation.

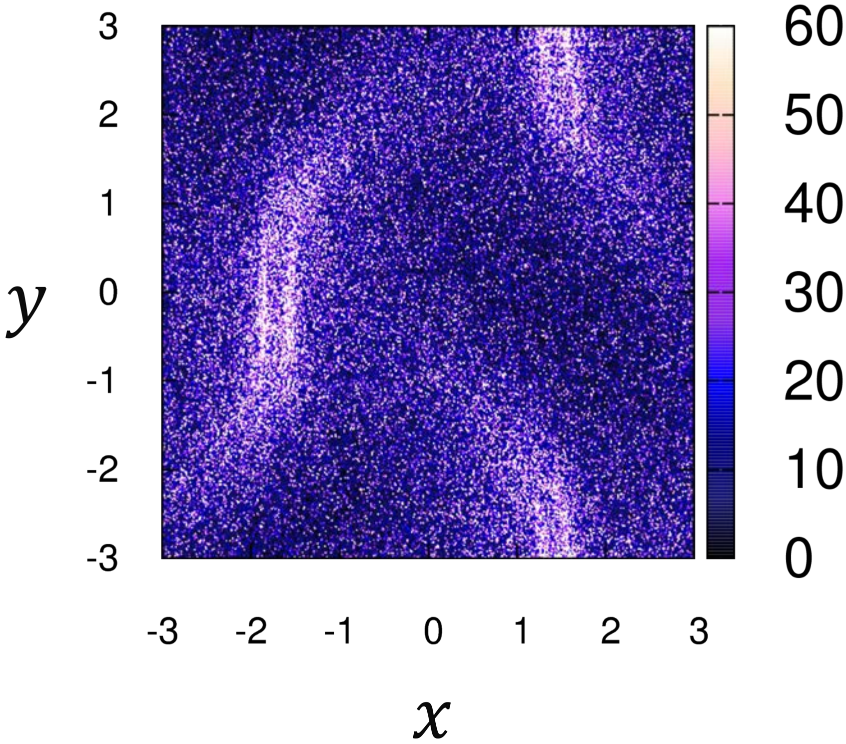

Figure 9.18: Calculated equilibrium probability distribution f in the (x, y) plane at $z = 0$ with antisymmetric operator (9.21). The initial condition $t = 0$ is the flat distribution of figure 9.5. Observe that the distribution resembles the profile of the magentic field strength B (compare with figure 9.17).

9.2.6 *Antisymmetric operator without boundaries*

In this simulation we change the boundary conditions. More precisely, we follow the trajectories of the particles as far as they go. The antisymmetric operator is chosen to be:

$$\boldsymbol{w} = \frac{\partial_x + \left(\frac{y-\sin y\cos y}{2} - \sin x\right)\partial_z}{1+\left(\frac{y-\sin y\cos y}{2} - \sin x\right)^2}. \tag{9.25}$$

To prove that $\boldsymbol{w}$ is not integrable, we evaluate the Jacobiator of the associated magnetic field $\boldsymbol{B} = \boldsymbol{w}/w^2$, which has a simpler expression. We find that:

$$\mathfrak{G}^{xyz}\left(\boldsymbol{B}\right) = \sin^2 y \geq 0, \tag{9.26}$$

and equality holds on a set of measure zero. Therefore $\boldsymbol{B}$ is not integrable, and thus $\boldsymbol{w}$ is not a Poisson operator. The field force divergence $\mathfrak{B}$ of the operator $\boldsymbol{w}$ does not vanish (the lengthy expression of $\mathfrak{B}$ is omitted). Therefore, $\boldsymbol{w}$ is not a Beltrami operator in the chosen coordinate system. The magnetic field strength has the expression:

$$B = w^{-1} = \sqrt{1+\left(\frac{y-\sin y\cos y}{2} - \sin x\right)^2}. \tag{9.27}$$

The profile of the magnetic field is shown in figure 9.19.

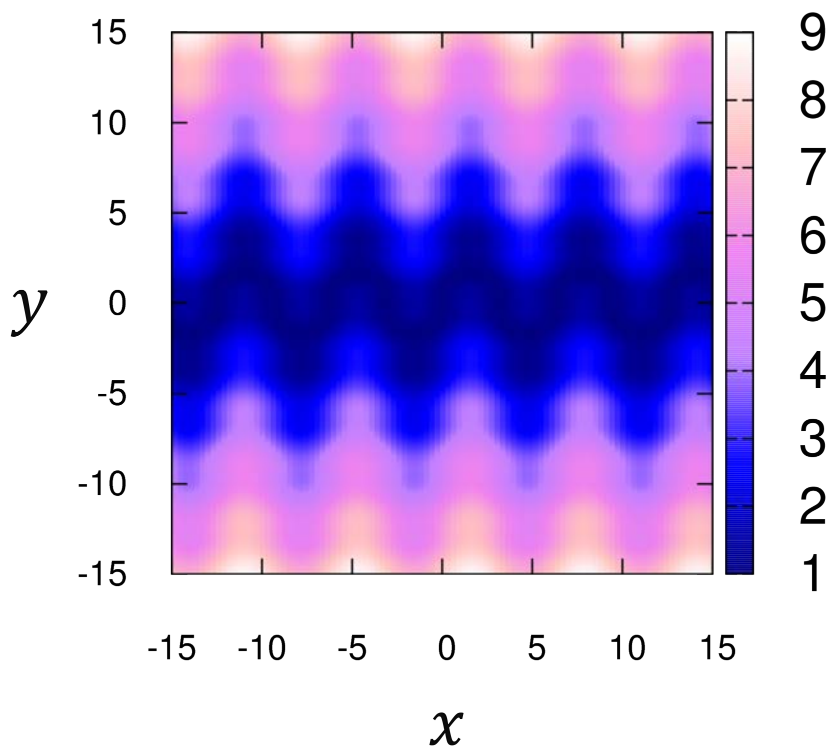

Figure 9.19: Magnetic field strength (9.27) in the (x, y) plane.

The results of the numerical simulations are given in figures 9.20 and 9.21. The initial conditions is still the flat distribution of figure 9.5. Observe the creation of an ordered structure. In figure 9.21 the probability distribution is shown on a domain of similar size to that of plot 9.19. Notice the similarity between the profiles of f and B.

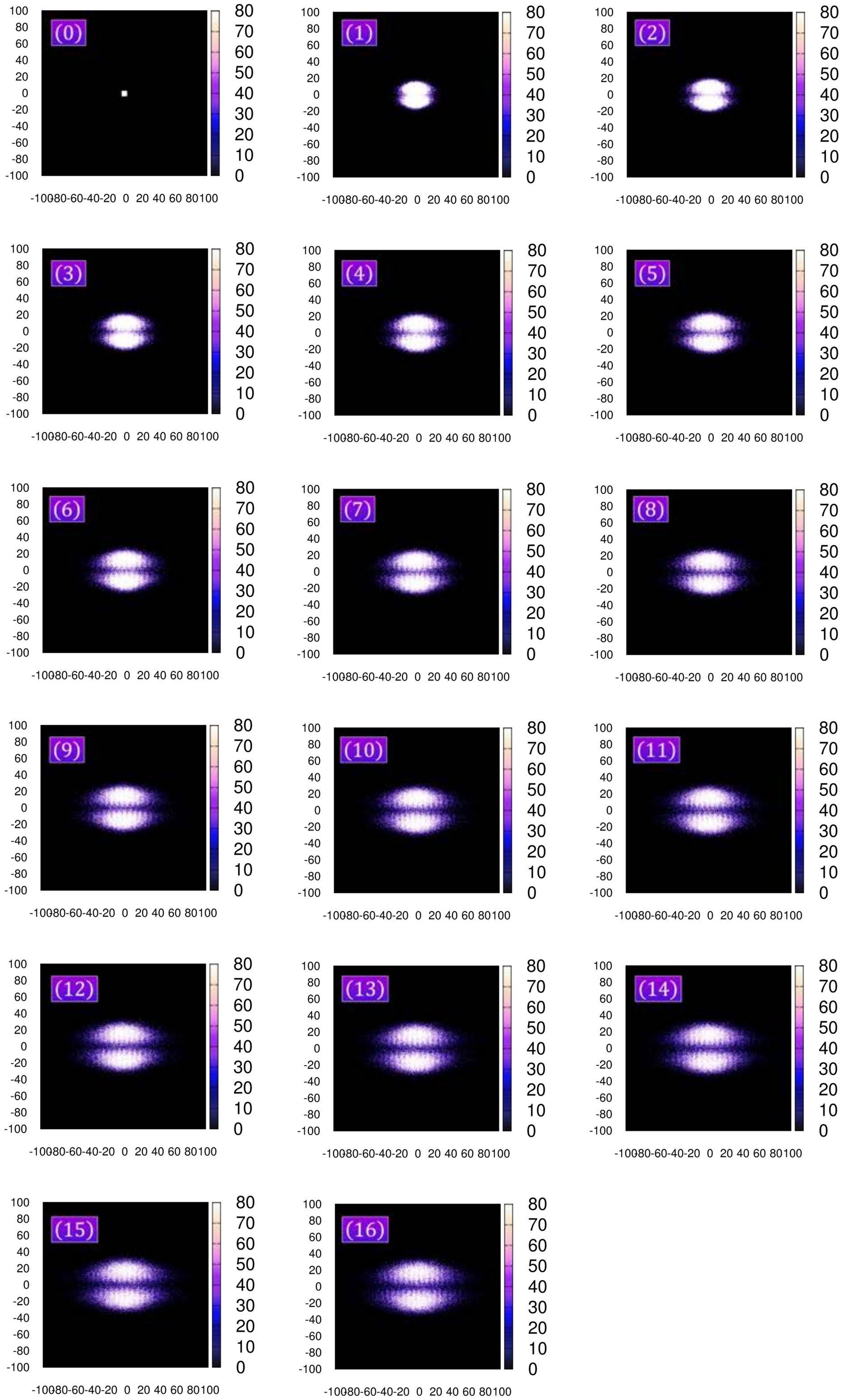

Figure 9.20: Time evolution of the probability distribution f in the (x, y) plane at $z = 0$. Each plot number i corresponds to the state of time evolution $t = i\Delta t$, where Δt is a fixed time interval.

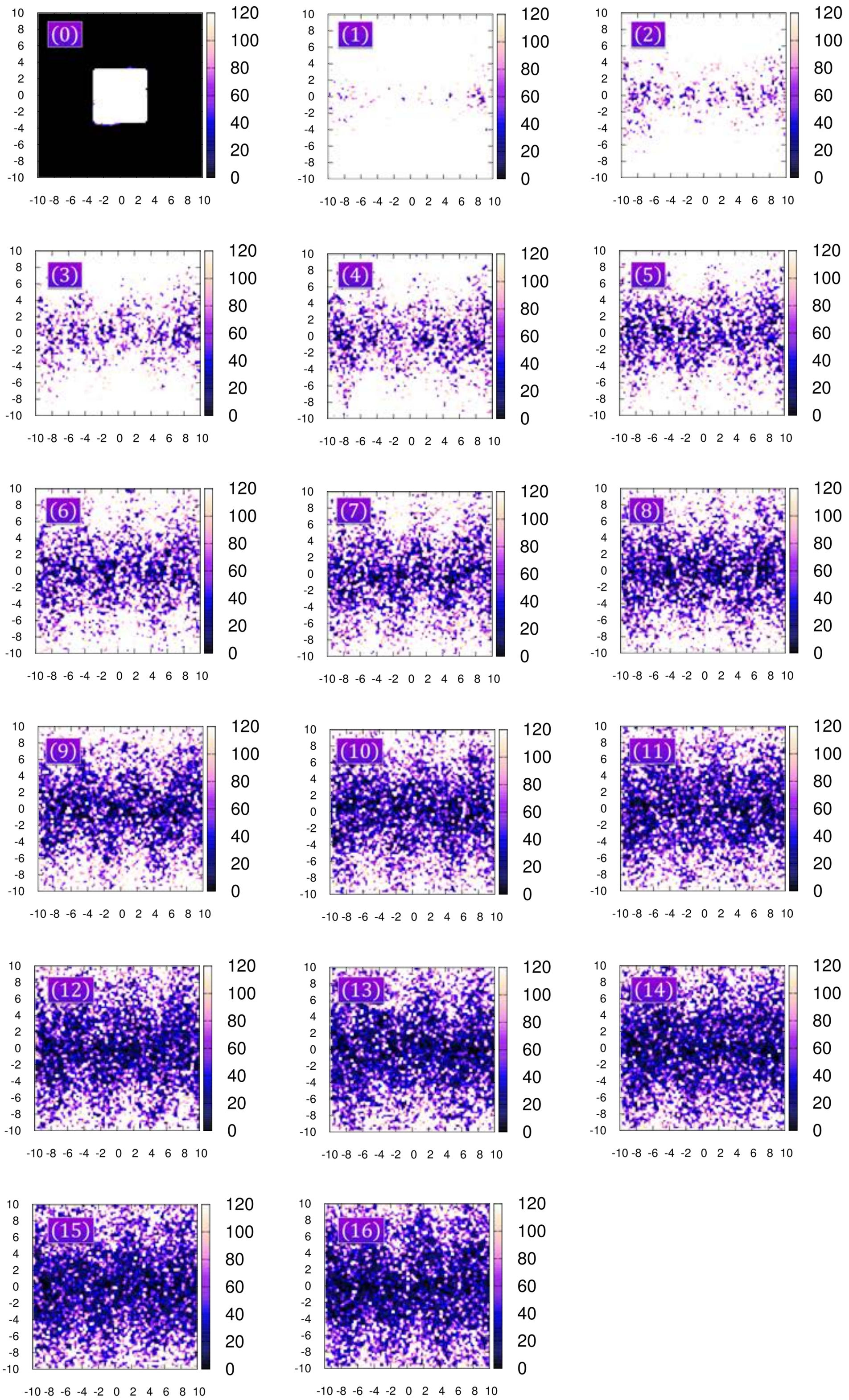

Figure 9.21: Time evolution of the probability distribution f on a square of side 20 in the (x, y) plane centered at $\boldsymbol{x} = \mathbf{0}$. Observe the similarity between the calculated profile and the shape of the magnetic field B as shown in figure 9.19.

9.2.7 *Antisymmetric operator with unit norm*

So far, we have analyzed two cases of antisymmetric operators and observed that the profile of the probability distribution tends to resemble that of the magnetic field strength $B = w^{-1}$. The natural question is then what happens to the probability distribution if the magnetic field strength $B = w^{-1}$ is constant. To answer this problem we consider the antisymmetric operator:

$$\hat{\boldsymbol{w}} = \frac{1}{\sqrt{1+\cos^2 x}} \left(\cos y, \cos x, \sin y\right). \tag{9.28}$$

Observe that $B = \hat{w}^{-1} = 1$ (and thus $\boldsymbol{B} = \hat{\boldsymbol{w}}$). The Jacobiator is evaluated more easily for the rescaled vector field $\boldsymbol{w}^{'} = (\cos y, \cos x, \sin y)$:

$$\mathfrak{S}^{xyz}\left(\boldsymbol{w}^{'}\right) = 1 - \sin y \sin x \geq 0, \tag{9.29}$$

with equality holding on a set of measure zero. Therefore $\boldsymbol{w}^{'}$ is not integrable, and thus $\hat{\boldsymbol{w}}$ is not a Poisson operator. The field force divergence $\hat{\mathfrak{B}}$ of the operator $\hat{\boldsymbol{w}}$ does not vanish (the lengthy expression of $\hat{\mathfrak{B}}$ is omitted). Therefore, $\hat{\boldsymbol{w}}$ is not a Beltrami operator in the chosen coordinate system.

The density profile obtained from the numerical simulation is shown in figure 9.22. Notice that, regardless of the fact that $B = \hat{w}^{-1} = 1$, an heterogeneous structure is self-organized. What is the determinant of such structure? The answer to this problem will be outlined in the next chapter. At this point we observe that the essential ingredient is the non-vanishing field force divergence $\hat{\mathfrak{B}}$. In fact, there is a strong similarity between the profile of the probability distribution and that of $\hat{\mathfrak{B}}$ (compare figure 9.22 with figure 10.4 showing the profile of $\hat{\mathfrak{B}}$).

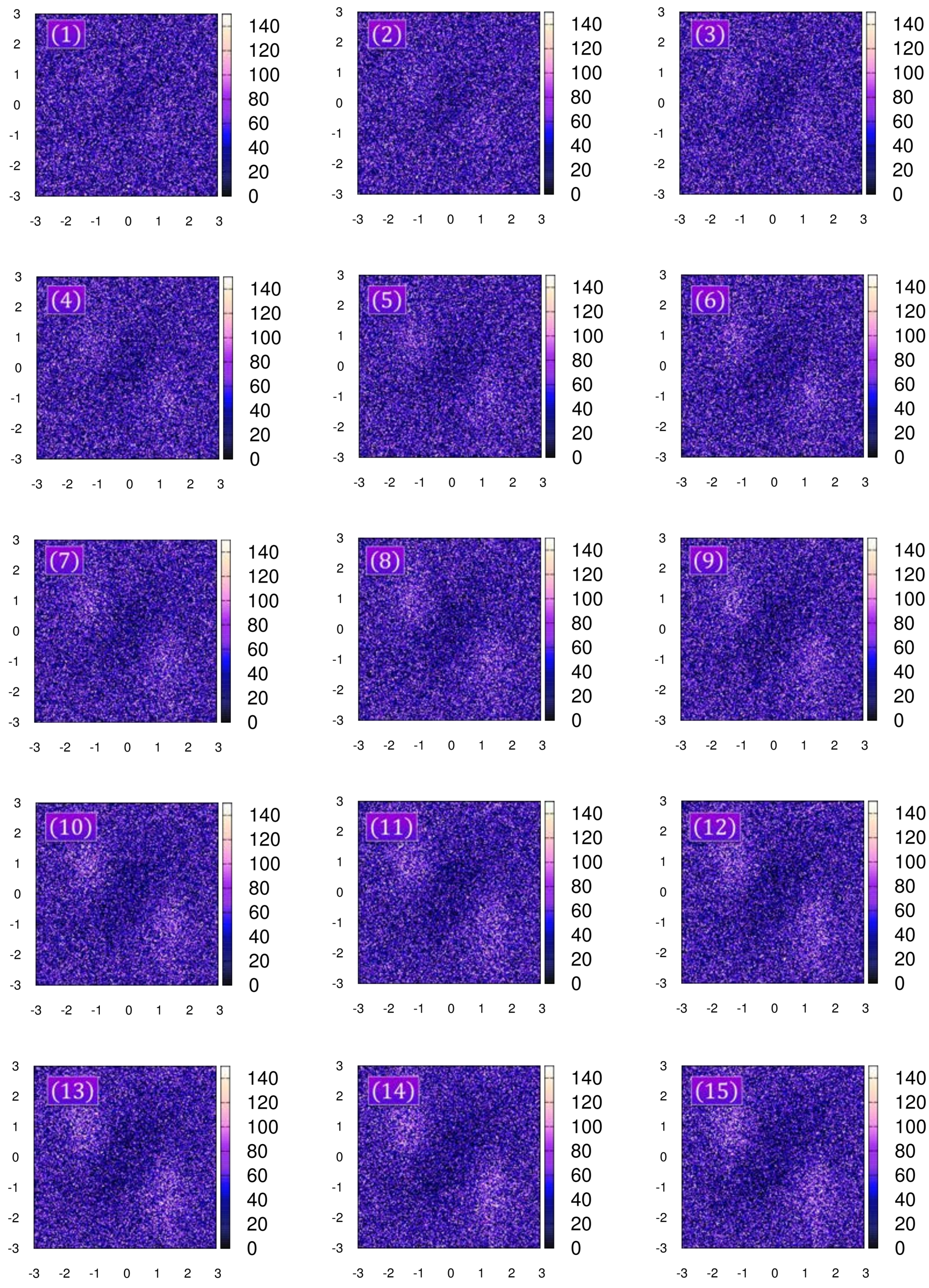

Figure 9.22: Time evolution of the probability distribution f in the (x, y) plane at $z = 0$. Each plot number i corresponds to the state of time evolution $t = i\Delta t$, where Δt is a fixed time interval.

9.2.8 *The Landau-Lifshitz Equation*

The last case we consider is the Landau-Lifshitz equation describing the time evolution of the magnetization $\boldsymbol{x}$ in a ferromagnet (specifically, we study equation (35) of [39]). Without entering into details, the Hamiltonian of the system, physically corresponding to the total magnetization, is given by:

$$H_0 = \frac{\boldsymbol{x}^2}{2}. \tag{9.30}$$

Therefore, in this simulation the perturbed Hamiltonian H is such that $\nabla H = \nabla H_0 + \boldsymbol{\Gamma}$. The antisymmetric operator of the system is:

$$\boldsymbol{w} = \gamma \mathcal{H} - \frac{\lambda}{x^2} \mathcal{H} \times \boldsymbol{x}. \tag{9.31}$$

Here, γ is the so called damping parameter, λ a physical constant, and $\mathcal{H}$ the effective magnetic field. The effective magnetic field $\mathcal{H}$ is chosen to be:

$$\mathcal{H} = (h, 0, z), \tag{9.32}$$

where h represents a constant external magnetic field. Then, equation (9.31) can be rewritten as:

$$\boldsymbol{w} = \left(h + \lambda \frac{zy}{\boldsymbol{s}^2}\right) \partial_x + \lambda \frac{z\,(h - x)}{\boldsymbol{s}^2} \partial_y + \left(z - \lambda \frac{hy}{\boldsymbol{s}^2}\right) \partial_z. \tag{9.33}$$

One can verify that this operator violates the Jacobi identity:

$$\mathfrak{G}^{xyz} = \lambda \frac{-2h^2x^2 - 2z^4 + h\left[x^3 + \lambda yz + x\left(y^2 - 3z^2\right)\right]}{\boldsymbol{s}^4} \tag{9.34}$$

Therefore, $\boldsymbol{w}$ is not a Poisson operator. The field force divergence $\mathfrak{B}$ can be calculated to give:

$$\mathfrak{B} = 4\lambda \frac{2h^2\lambda\left(3x^2 - \boldsymbol{s}^2\right) + \lambda\left(\boldsymbol{s}^4 - 11z^2\boldsymbol{s}^2 + 14z^4\right) - 4h\left[\lambda x\left(\boldsymbol{s}^2 - 5z^2\right) + yz\boldsymbol{s}^2\right]}{\boldsymbol{s}^6}. \tag{9.35}$$

Thus, in the chosen reference frame, the operator $\boldsymbol{w}$ of the Landau-Lifshitz equation is not of the Beltrami type.

In figure 9.23 the results of the numerical simulation are shown. This time, the initial condition is a Maxwell-Boltzmann distribution centered at $\boldsymbol{x} = (0, 0, z_0)$. As in the previous case, the trajectory of each magnetization is followed as far as it goes. Notice how the probability distribution becomes strongly anisotropic, with preferential alignment of the magnetization along the z-axis (representing the direction of easiest magnetization of the ferromagnetic crystal).

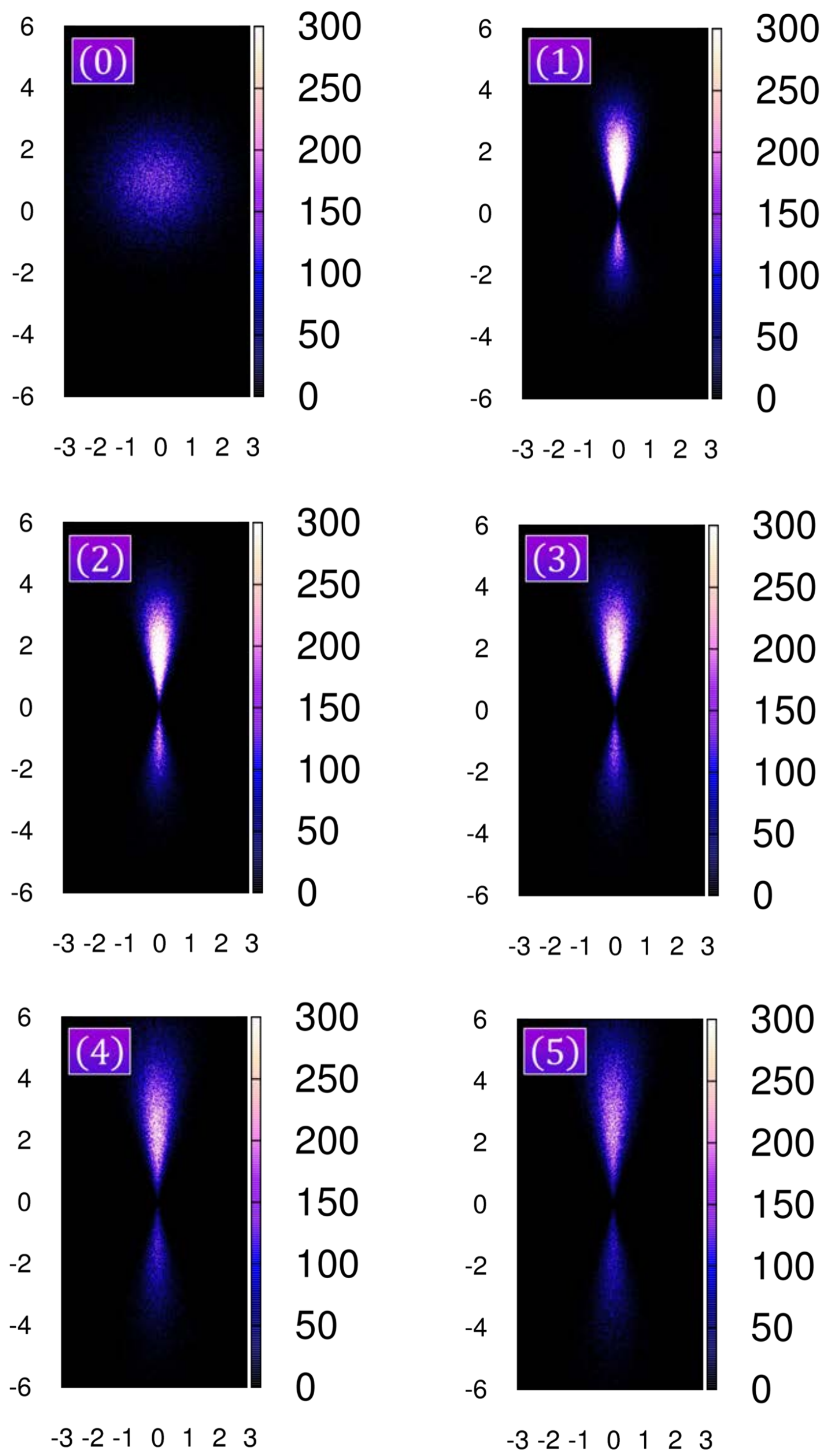

Figure 9.23: Time evolution of the probability distribution f in the (x, z) plane at $y = 0$. Each plot number i corresponds to the state of time evolution $t = i\Delta t$, where Δt is a fixed time interval.

Chapter 10

The Normal Laplacian

In this last chapter we examine the *normal Laplace operator* $\Delta_\perp$ encountered in equation (9.10), and study existence and uniqueness of solution to the normal Laplace equation, which we will define shortly. To simplify the problem, the discussion is mainly limited to the 3-dimensional case.

Exploiting the non-integrability of the vector field $\boldsymbol{w}$, we show that a novel norm $||\cdot||_\perp$ can be defined. Then, using Riesz's representation theorem, the existence of a weak unique solution is proven.

10.1 The Normal and Parallel Laplacian Operators

As usual, consider a smooth manifold $\mathcal{M}^n$ of dimension n.

***Def* 10.1.** *(Normal gradient in 3D)*
Let $\boldsymbol{w} \in T\mathcal{M}^3$ be a vector field. The normal gradient of a function $f \in C^1\left(\mathcal{M}^3\right)$, $f : \mathcal{M}^3 \to \mathbb{R}$ with respect to $\boldsymbol{w}$ is defined as:

$$\nabla_\perp f = \frac{\boldsymbol{w} \times (\nabla f \times \boldsymbol{w})}{w^2}. \tag{10.1}$$

***Def* 10.2.** *(Parallel gradient in 3D)*
Let $\boldsymbol{w} \in T\mathcal{M}^3$ be a vector field. The parallel gradient of a function $f \in C^1\left(\mathcal{M}^3\right)$, $f : \mathcal{M}^3 \to \mathbb{R}$ with respect to $\boldsymbol{w}$ is defined as:

$$\nabla_\parallel f = \frac{\boldsymbol{w}}{w^2}\left(\boldsymbol{w} \cdot \nabla f\right). \tag{10.2}$$

***Def* 10.3.** *(Normal Laplacian in 3D)*
The normal Laplacian of a function $f \in C^2\left(\mathcal{M}^3\right)$, $f : \mathcal{M}^3 \to \mathbb{R}$ with normal gradient given by definition 10.1 is defined as:

$$\Delta_\perp f = \nabla \cdot \left(w^2 \nabla_\perp f\right). \tag{10.3}$$

***Def* 10.4.** *(Parallel Laplacian in 3D)*
The parallel Laplacian of a function $f \in C^2\left(\mathcal{M}^3\right)$, $f : \mathcal{M}^3 \to \mathbb{R}$ with parallel gradient given by definition 10.2 is defined as:

$$\Delta_\| f = \nabla \cdot \left(w^2 \nabla_\| f\right). \tag{10.4}$$

These definitions can be generalized to higher dimensions in the following fashion.

***Def* 10.5.** *(Normal gradient)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}^n$ be an antisymmetric operator. The normal gradient of a function $f \in C^1\left(\mathcal{M}^n\right)$, $f : \mathcal{M}^n \to \mathbb{R}$ with respect to $\mathcal{J}$ is defined as:

$$\nabla_\perp f = \frac{\mathcal{J}^{ik}\mathcal{J}^{jk} f_j}{|\mathcal{J}|^2}\partial_i. \tag{10.5}$$

***Def* 10.6.** *(Parallel gradient)*
Let $\mathcal{J} \in \bigwedge^2 T\mathcal{M}^n$ be an antisymmetric operator. The parallel gradient of a function $f \in C^1\left(\mathcal{M}^n\right)$, $f : \mathcal{M}^n \to \mathbb{R}$ with respect to $\mathcal{J}$ is defined as:

$$\nabla_\| f = \left(f_i - \frac{\mathcal{J}^{ik}\mathcal{J}^{jk} f_j}{|\mathcal{J}|^2}\right)\partial_i. \tag{10.6}$$

***Def* 10.7.** *(Normal Laplacian)*
The normal Laplacian of a function $f \in C^2\left(\mathcal{M}^n\right)$, $f : \mathcal{M}^n \to \mathbb{R}$ with normal gradient given by definition 10.5 is defined as:

$$\Delta_\perp f = \nabla \cdot \left(|\mathcal{J}|^2 \nabla_\perp f\right). \tag{10.7}$$

***Def* 10.8.** *(Parallel Laplacian)*
The parallel Laplacian of a function $f \in C^2\left(\mathcal{M}^n\right)$, $f : \mathcal{M}^n \to \mathbb{R}$ with parallel gradient given by definition 10.6 is defined as:

$$\Delta_\| f = \nabla \cdot \left(|\mathcal{J}|^2 \nabla_\| f\right). \tag{10.8}$$

Next, we define the normal Laplace equation:

***Def* 10.9.** *(Normal Laplace equation)*
Let $\Omega \subset \mathcal{M}^n$ be a smoothly bounded compact domain with boundary $\partial\Omega$. Let $\phi : \Omega \to \mathbb{R}$ and $a : \Omega \to \mathbb{R}_{\geq 0}$ be known functions. The normal Laplace equation with respect to the function $f \in C^2(\Omega)$, $f : \Omega \to \mathbb{R}$ is the non-elliptic second order partial differential equation:

$$\begin{aligned} &\Delta_\perp f - af = \phi \quad in \;\; \Omega, \\ &f = 0 \quad on \;\; \partial\Omega. \end{aligned} \tag{10.9}$$

We refer the reader to section 6.4 for the discussion of the non-ellipticity of this differential operator.

The parallel Laplace equation can be defined in a similar way. We omit it because our attention will be focused exclusively on equation (10.9).

10.2 Existence and Uniqueness of Solution to the Normal Laplace Equation

From this point we set $n = 3$. Let $\Omega \subset \mathcal{M}^3$ be a smoothly bounded compact domain. We denote by $\partial\Omega$ the boundary of Ω, and by $\boldsymbol{n}$ the unit outward normal vector on $\partial\Omega$. For a given smooth vector field $\boldsymbol{w} \in T\mathcal{M}^3$ such that

$$\boldsymbol{w} \cdot \nabla \times \boldsymbol{w} \neq 0 \; in \; \Omega, \quad \boldsymbol{n} \cdot \boldsymbol{w} = 0 \; on \; \partial\Omega,$$

we have (denoting $\hat{\boldsymbol{w}} = \boldsymbol{w}/|\boldsymbol{w}|$)

$$\nabla_\perp u = \hat{\boldsymbol{w}} \times (\nabla u \times \hat{\boldsymbol{w}}), \quad \nabla_\| u = \nabla u - \nabla_\perp u.$$

The direction of $\boldsymbol{w}$ is said *parallel*, and the others *normal* or *perpendicular*. We consider the following spaces of scalar functions (firstly, without topology):

***Def* 10.10.** *(Normal space)*
The normal space $H^\perp(\Omega)$ is the function space of $L^2(\Omega)$-measurable functions u such that $|\nabla_\perp u| \in L^2(\Omega)$:

$$H^\perp(\Omega) = \left\{u \in L^2(\Omega)\,; \;\; |\nabla_\perp u| \in L^2(\Omega)\right\}. \tag{10.10}$$

***Def* 10.11.** *(Parallel space)*
The parallel space $H^\|(\Omega)$ is the function space of $L^2(\Omega)$-measurable functions u such that $\left|\nabla_\| u\right| \in L^2(\Omega)$:

$$H^\|(\Omega) = \left\{u \in L^2(\Omega)\,; \;\; \left|\nabla_\| u\right| \in L^2(\Omega)\right\}. \tag{10.11}$$

Recall that $L^2(\Omega)$ is the set of functions u such that $|u|^2$ is Lebesgue-integrable on Ω. We also denote with $H^1(\Omega) = \left\{u \in L^2(\Omega);\ |\nabla u| \in L^2(\Omega)\right\}$ the standard Sobolev space. Evidently $H^1(\Omega) \subset H^{\perp}(\Omega) \subset L^2(\Omega)$ and $H^1(\Omega) \subset H^{\parallel}(\Omega) \subset L^2(\Omega)$.

In order to define a topology, we consider the following bilinear product:

***Def* 10.12.** *(Normal product)*
Let $\boldsymbol{w} \in T\mathcal{M}^3$ be a smooth vector field with $\mathfrak{H}^{xyz} = \boldsymbol{w} \cdot \nabla \times \boldsymbol{w} \neq 0$ on a smoothly bounded compact domain $\Omega \subset \mathcal{M}^3$. Let $a : \Omega \to \mathbb{R}_{\geq 0}$ be a non-negative smooth function. The normal product of $u, v \in H^{\perp}(\Omega)$ with respect to $\boldsymbol{w}$ is defined as:

$$(u,v)_{\perp,a} = (u, av) + \left(\nabla_{\perp} u, w^2 \nabla_{\perp} v\right) = \int_{\Omega} \left[auv + w^2 \left(\nabla_{\perp} u \cdot \nabla_{\perp} v\right)\right] vol^3. \quad (10.12)$$

where (f,g) denotes the standard L^2 inner product (f and g may be scalar or vector valued functions). We define:

$$||u||^2_{\perp,a} = (u,u)_{\perp,a}. \quad (10.13)$$

For the case $a = 0$, we will use the notation $(u,v)_{\perp} = (u,v)_{\perp,0}$ and $||u||_{\perp} = ||u||_{\perp,0}$.

Evidently, we have

***Proposition* 10.1.** *For $a > 0$,*

1. *The bilinear form $(u,v)_{\perp,a}$ satisfies the axioms of inner-product on $H^{\perp}(\Omega)$; hence, $||u||_{\perp,a}$ is a norm on $H^{\perp}(\Omega)$,*
2. *$H^{\perp}(\Omega)$ is complete for the norm $||u||_{\perp,a}$,*
3. *Hence, $H^{\perp}(\Omega)$ can be regarded as a Hilbert space endowed with the inner-product $(u,v)_{\perp,a}$.*

Proof. We begin by verifying that definition 10.12 is consistent with the requirements of an inner product on $H^{\perp}(\Omega)$.

(proof of 1)

First, we must check that $\forall u \in H^{\perp}(\Omega)$:

$$(u,u)_{\perp,a} \geq 0, \quad (u,u)_{\perp,a} = 0 \iff u = 0. \quad (10.14)$$

$(u,u)_{\perp,a} \geq 0$ is immediate. Note that $(u,u)_{\perp,a} = 0$ is possible only if $u = 0$ *a.e.*. The remaining requirements are trivial:

$$(u,v)_{\perp,a} = (v,u)_{\perp,a} \quad \forall u, v \in H^{\perp}(\Omega), \quad (10.15a)$$

$$(c_1 u_1 + c_2 u_2, v)_{\perp,a} = c_1 (u_1, v)_{\perp,a} + c_2 (u_2, v)_{\perp,a} \quad \forall u_1, u_2, v \in H^{\perp}(\Omega),\ \forall c_1, c_2 \in \mathbb{R}. \quad (10.15b)$$

In a similar fashion, one can verify that the functional:

$$||u||^2_{\perp,a} = (u,u)_{\perp,a}\,, \tag{10.16}$$

satisfies the norm properties $\forall u, v \in H^{\perp}(\Omega)$:

$$||u||_{\perp,a} \geq 0, \quad ||u||_{\perp,a} = 0 \iff u = 0, \tag{10.17a}$$
$$||u+v||_{\perp,a} \leq ||u||_{\perp,a} + ||v||_{\perp,a}\,, \tag{10.17b}$$
$$||cu||_{\perp,a} = |c|\,||u||_{\perp,a}\,, \quad \forall c \in \mathbb{R}. \tag{10.17c}$$

(proof of 2)

Suppose that $u_1, u_2, \dots \in H^{\perp}(\Omega)$ is a Cauchy sequence with respect to the norm $\|u\|_{\perp,a}$. Then $u_j \to \exists u_\infty \in L^2(\Omega)$ and $\nabla_\perp u_j \to \exists \boldsymbol{g} \in L^2(\Omega)$. Necessarily $\boldsymbol{g} = \nabla_\perp u_\infty$, so $u_\infty \in H^{\perp}(\Omega)$, and $u_j \to u_\infty \in H^{\perp}(\Omega)$.

□

***Theorem* 10.1.** *(Uniqueness of solution for the normal Laplacian with $a > 0$)*
Let $\boldsymbol{w} \in T\mathcal{M}^3$ be a smooth vector field with $\mathfrak{G}^{xyz} = \boldsymbol{w}\cdot\nabla\times\boldsymbol{w} \neq 0$ on a smoothly bounded compact domain $\Omega \subset \mathcal{M}^3$. Suppose that $a > 0$. Then, if it exists, the solution $u \in H^{\perp}(\Omega)$ to the normal Laplace equation:

$$\begin{aligned} &\Delta_\perp u - au = \phi \quad in \;\; \Omega, \\ &u = 0 \quad on \;\; \partial\Omega, \end{aligned} \tag{10.18}$$

is unique.

Proof. Consider the functional:

$$\mathcal{F} = \int_\Omega \left[au^2 + w^2\,|\nabla_\perp u|^2\right] vol^3. \tag{10.19}$$

Evidently $\mathcal{F} \geq 0$. Suppose that u_1 and u_2 are two distinct solutions to system (10.18). Set $u = u_1 - u_2$. We have $\Delta_\perp u - au = 0$ in Ω and $u = 0$ on $\partial\Omega$. The following identity holds:

$$\nabla\cdot\left(uw^2\nabla_\perp u\right) = w^2\,|\nabla_\perp u|^2 + u\Delta_\perp u = w^2\,|\nabla_\perp u|^2 + au^2. \tag{10.20}$$

Therefore:

$$\mathcal{F} = \int_\Omega \nabla\cdot\left(uw^2\nabla_\perp u\right) vol^3 = \int_{\partial\Omega} uw^2\nabla_\perp u\cdot\boldsymbol{n}\,dS^2 = 0. \tag{10.21}$$

The last passage can be obtained by using the boundary condition. However, due to equation (10.19), $\mathcal{F} = 0$ is possible only if $u = 0$ *a.e.* on Ω. It follows that $u_1 = u_2$ *a.e.* on Ω. □

***Theorem* 10.2.** *(Existence of weak solution for the normal Laplacian with $a > 0$) Let $H^{\perp}(\Omega)$ be the normal (Hilbert) space equipped with the normal inner product* (10.12) *with $a > 0$ defined by a non-integrable smooth vector field $\boldsymbol{w}$ with $\mathfrak{H}^{xyz} = \boldsymbol{w} \cdot \nabla \times \boldsymbol{w} \neq 0$. Then, the normal Laplace equation* (10.18) *admits a unique and weak solution $u \in H^{\perp}(\Omega)$ for any $\phi \in H^{\perp}(\Omega)$.*

Proof. We apply Riesz's representation theorem: since $H^{\perp}(\Omega)$ is a Hilbert space, we can find $u \in H^{\perp}(\Omega)$ such that:

$$\int_{\Omega} v\phi \, vol^3 = -(v,u)_{\perp,a} \quad \forall v \in H^{\perp}(\Omega). \tag{10.22}$$

Now, observe that:

$$\begin{aligned}(v,u)_{\perp,a} &= \int_{\Omega} \left[auv + w^2 \left(\nabla_{\perp} v \cdot \nabla_{\perp} u\right)\right] vol^3 \\ &= \int_{\partial\Omega} w^2 v \nabla_{\perp} u \cdot \boldsymbol{n} \, dS^2 + \int_{\Omega} v \left[au - \Delta_{\perp} u\right] vol^3 \\ &= \int_{\Omega} v \left[au - \Delta_{\perp} u\right] vol^3.\end{aligned} \tag{10.23}$$

Therefore u is a weak solution to the normal Laplace equation. In virtue of theorem 10.1, the solution is also unique. □

Notice that with theorem 10.2 we have shown that we can solve equations of the type:

$$\Delta_{\perp} f + \frac{1}{4}\mathfrak{B} f = 0, \tag{10.24}$$

with $\mathfrak{B} < 0$. The equation above should be compared with the original equation (9.10).

Now we want to consider the case $a = 0$. Define the function space:

$$C_0^1(\Omega) = \left\{u \in C^1(\Omega) : \ u = 0 \ on \ \partial\Omega\right\}, \tag{10.25}$$

***Proposition* 10.2.** *For $a = 0$,*

1. *The bilinear form $(u,v)_{\perp} = (u,v)_{\perp,0}$ satisfies the axioms of inner-product on $C_0^1(\Omega)$; hence, $||u||_{\perp} = (u,u)_{\perp,0}$ is a norm on $C_0^1(\Omega)$.*

2. *The function space $C_0^1(\Omega)$ equipped with the norm $||u||_{\perp}$ is a pre-Hilbert space and can be completed to the Hilbert space $\mathcal{H}_0^{\perp}(\Omega)$. We also have $H_0^{\perp}(\Omega) \subset \mathcal{H}_0^{\perp}(\Omega)$.*

Proof. We begin by verifying that $(u,v)_\perp$ is consistent with the requirements of an inner product on $C_0^1(\Omega)$. First, we must check that $\forall u \in C_0^1(\Omega)$:

$$(u,u)_\perp \geq 0, \quad (u,u)_\perp = 0 \iff u = 0. \tag{10.26}$$

$(u,u)_\perp \geq 0$ is immediate. Note that $(u,u)_\perp = 0$ is possible only if $u = 0$. Indeed, by vector identity, $|\nabla_\perp u|^2 = w^{-2}|\boldsymbol{w} \times \nabla u|^2$ and since $\boldsymbol{w}$ is a non-integrable vector field, no function u with $\nabla u \neq \mathbf{0}$ on Ω, $u \in C_0^1(\Omega)$ can be found such that $\boldsymbol{w} \times \nabla u = \mathbf{0}$ *a.e.* on Ω (otherwise we could find $\alpha \in C^1(\Omega)$, $\alpha \neq 0$ such that $\boldsymbol{w} = \alpha \nabla u$, contradicting the non-integrability of $\boldsymbol{w}$). The remaining requirements are trivial and can be verified as in proof of proposition 10.1. It follows that the functional $||u||_\perp$ satisfies the norm properties in $C_0^1(\Omega)$. With this norm, $C_0^1(\Omega)$ is a pre-Hilbert that can be completed to $\mathcal{H}_0^\perp(\Omega)$. The inclusion of the normed space $H_0^\perp$ with norm $||\cdot||_{\perp,a}$, $(a \geq 0)$, into $\mathcal{H}_0^\perp(\Omega)$:

$$H_0^\perp(\Omega) \subset \mathcal{H}_0^\perp(\Omega), \tag{10.27}$$

also follows. $\square$

Although the new Hilbert space $\mathcal{H}_0^\perp$ is an abstract object, Riesz's representation theorem can still be applied to prove existence and uniquencess of solution in $\mathcal{H}_0^\perp$ to the normal Laplace equation with $a = 0$:

***Theorem* 10.3.** *(Existence of weak solution for the normal Laplacian with $a = 0$) Let $\mathcal{H}_0^\perp$ be the normal (Hilbert) space equipped with the normal inner product $||\cdot||_\perp$ defined by a non-integrable vector field $\boldsymbol{w}$ with $\mathfrak{G}^{xyz} = \boldsymbol{w} \cdot \nabla \times \boldsymbol{w} \neq 0$. Then, the normal Laplace equation* (10.9) *with $a = 0$ admits a unique and weak solution $u \in \mathcal{H}_0^\perp$ for any $\phi \in \mathcal{H}_0^\perp$.*

Proof. We apply Riesz's representation theorem: since $\mathcal{H}_0^\perp$ is a Hilbert space, we can find a unique $u \in \mathcal{H}_0^\perp$ such that:

$$\int_\Omega v\phi \, vol^3 = -(v,u)_\perp \quad \forall v \in \mathcal{H}_0^\perp. \tag{10.28}$$

Now, observe that:

$$\begin{aligned}
(v,u)_\perp &= \int_\Omega w^2 (\nabla_\perp v \cdot \nabla_\perp u) \, vol^3 \\
&= \int_{\partial\Omega} w^2 v \nabla_\perp u \cdot \boldsymbol{n} \, dS^2 - \int_\Omega v \Delta_\perp u \, vol^3 \\
&= -\int_\Omega v \Delta_\perp u \, vol^3.
\end{aligned} \tag{10.29}$$

Therefore u is a unique and weak solution to the normal Laplace equation. $\square$

At this point, we want to push our results further by showing that if a solution $u \in H_0^{\perp}(\Omega)$ to the normal Laplace equation (10.9) with $a = 0$ exists, then this solution is also unique. For this purpose, we need the following lemma:

***Lemma* 10.1.** *The solution $u \in H^{\perp}(\Omega)$ of the equations*

$$\boldsymbol{w} \times \nabla u = \mathbf{0} \quad \textit{a.e. in } \Omega, \tag{10.30a}$$

$$u = 0 \quad \textit{on } \partial\Omega \tag{10.30b}$$

is uniquely $u = 0$ a.e. in Ω.

Proof. The proof involves three main steps:

Step 1: We show that the solution in $C^1(\Omega)$ is only $u = 0$. Equation (10.30a) implies that $\nabla u = \alpha \boldsymbol{w}$ with some scalar function α. We find that α must be zero everywhere in Ω, and thus, by (10.30b), $u = 0$ in Ω. In fact, if $\alpha \neq 0$ in any open set $U \subset \Omega$, we can write $\boldsymbol{w} = \alpha^{-1}\nabla u$ in U, which contradicts with the non-integrability of $\boldsymbol{w}$ (Frobenius theorem 1.2).

Step 2: The solution in $H^1(\Omega)$ is also only $u = 0$ (*a.e.* in Ω); this is evident because $C^1(\Omega)$ is dense in $H^1(\Omega)$.

Step 3: When we extend the set of candidates for the solution to $H^{\perp}(\Omega)$, we have to care for the possibility of u such that $\nabla_{\parallel} u$ is not definable in $L^2(\Omega)$. Suppose that a solution in $H^{\perp}(\Omega)$ has a 'singularity' $\boldsymbol{x}_s \in \Omega$ at which the variation of u in the parallel direction is finite, but it cannot be evaluated by $\nabla_{\parallel} u$ (i.e., u is discontinuous in the parallel direction). Solving

$$\frac{d\boldsymbol{\xi}}{d\tau} = \boldsymbol{w}, \quad \boldsymbol{\xi}(0) = \boldsymbol{x}_s,$$

we define a characteristic curve $\boldsymbol{\xi}(\tau)$ of $\boldsymbol{w}$ including $\boldsymbol{x}_s$ (see figure 10.1). In a neighborhood U of $\boldsymbol{x}_s = \boldsymbol{\xi}(0)$, we choose two points

$$\boldsymbol{p} = \boldsymbol{\xi}(\tau_-), \quad \boldsymbol{q} = \boldsymbol{\xi}(\tau_+)$$

with $\tau_- < 0$ and $\tau_+ > 0$. Since $\boldsymbol{x}_s$ is a singularity (in the sense of the foregoing definition), $u(\boldsymbol{p}) \neq u(\boldsymbol{q})$ [1]. On the other hand, by Caratheodory's theorem 1.3, there is a piecewise smooth connection $\Gamma \subset U$ (see figure 10.1), between $\boldsymbol{p}$ and $\boldsymbol{q}$, such that

$$\hat{\boldsymbol{w}} \cdot \frac{d\boldsymbol{\eta}}{d\tau} = 0, \quad \boldsymbol{\eta}(\tau) \in \Gamma.$$

[1] $u \in H^{\perp}(\Omega)$ is not necessarily a continuous function, so we may not evaluate the local values of u. Here, by $u(\boldsymbol{x})$, we mean the volume average in some neighborhood of $\boldsymbol{x}$.

In the direction parallel to Γ, we may evaluate $\nabla_\perp u$ which must be zero if u satisfies (10.30a). Integrating the variation of u along Γ, we obtain $u(\boldsymbol{p}) = u(\boldsymbol{q})$ [2], contradicting with the previous assumption. Therefore, the variation is eliminated also in the parallel direction when $\nabla_\perp u = \mathbf{0}$; this is because the non-integrable $\boldsymbol{w}$ cannot define a boundary separating any sub-domain in which $u \neq 0$.

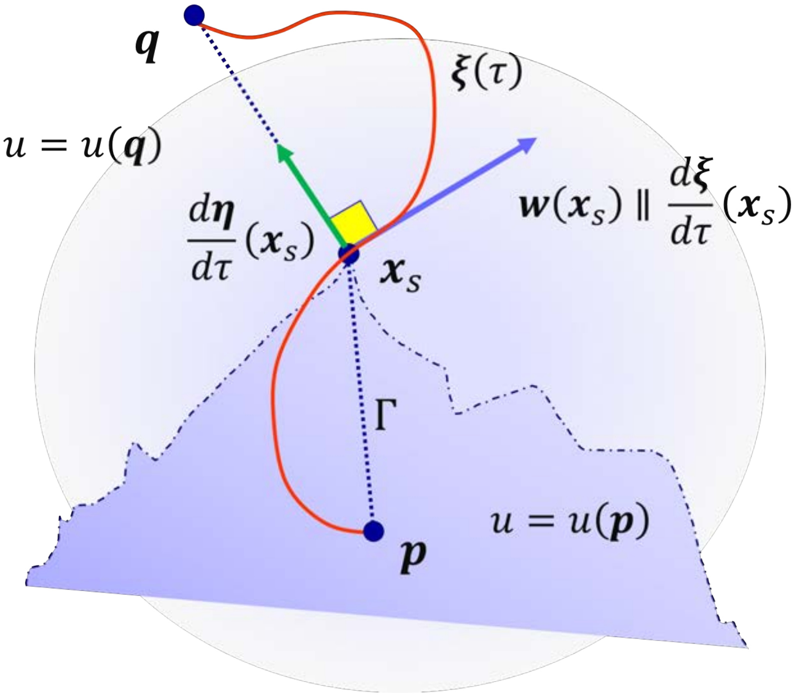

Figure 10.1: The singularity $\boldsymbol{x}_s$ and the curves $\boldsymbol{\xi}(\tau)$ and Γ.

$\square$

***Theorem* 10.4.** *(Uniqueness of solution for the normal Laplacian with $a = 0$)* *Let $\boldsymbol{w} \in T\mathcal{M}^3$ be a smooth vector field with $\mathfrak{S}^{xyz} = \boldsymbol{w} \cdot \nabla \times \boldsymbol{w} \neq 0$ on a smoothly bounded compact domain $\Omega \subset \mathcal{M}^3$. Suppose that $a = 0$. Then, if it exists, the solution $u \in H^\perp(\Omega)$ to the normal Laplace equation:*

$$\begin{aligned} \Delta_\perp u &= \phi \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{10.31}$$

is unique.

Proof. Consider the functional:

$$\mathcal{F} = \int_\Omega w^2 \left|\nabla_\perp u\right|^2 vol^3. \tag{10.32}$$

[2] In $H^\perp(\Omega)$, $\nabla_\perp u$ is not necessarily an absolutely continuous function. However, for the solution u, $\nabla_\perp u = 0$ a.e. in U; hence the total variation of u along Γ is zero.

Evidently $\mathcal{F} \geq 0$. Suppose that u_1 and u_2 are two distinct solutions to system (10.31). Set $u = u_1 - u_2$. We have $\Delta_\perp u = 0$ on Ω and $u = 0$ on $\partial\Omega$. The following identity holds:

$$\nabla \cdot \left(uw^2 \nabla_\perp u\right) = w^2 \left|\nabla_\perp u\right|^2 + u\Delta_\perp u = w^2 \left|\nabla_\perp u\right|^2. \tag{10.33}$$

Therefore:

$$\mathcal{F} = \int_\Omega \nabla \cdot \left(uw^2 \nabla_\perp u\right) vol^3 = \int_{\partial\Omega} uw^2 \nabla_\perp u \cdot \boldsymbol{n}\, dS^2 = 0. \tag{10.34}$$

The last passage can be obtained by using the boundary condition. However, due to equation (10.32), $\mathcal{F} = 0$ is possible only if $w^2 \left|\nabla_\perp u\right|^2 = \left|\boldsymbol{w} \times \nabla u\right|^2 = 0$ *a.e.* on Ω. Notice that $\boldsymbol{w}$ is always different from zero since by hypothesis $\mathfrak{G}^{xyz} \neq 0$. Then, the conditions of lemma 10.1 are satisfied and we must have $u = 0$ *a.e.* on Ω. It follows that $u_1 = u_2$ *a.e.* on Ω. □

In virtue of lemma 10.1, we also have the following:

***Proposition* 10.3.** *For $a = 0$,*

1. *The bilinear form $(u,v)_\perp = (u,v)_{\perp,0}$ satisfies the axioms of inner-product on $H_0^\perp(\Omega)$; hence, $||u||_\perp = (u,u)_{\perp,0}$ is a norm on $H_0^\perp(\Omega)$.*
2. *The function space $H_0^\perp(\Omega)$ equipped with the norm $||u||_\perp$ is a pre-Hilbert space and can be completed to a Hilbert space.*

Proof. The proof is analogous to that of proposition 10.2, except that the axioms of inner-product are verified by using lemma 10.1. □

10.3 Estimates for the Normal and Parallel Gradients

It is worth to notice that the operator $\Delta_\perp$ does not involve any derivation of the function u along the vector field $\boldsymbol{w}$:

$$\begin{aligned}\Delta_\perp u &= \nabla \cdot \left[\boldsymbol{w} \times (\nabla u \times \boldsymbol{w})\right] \\ &= (\nabla_\perp u \times \boldsymbol{w}) \cdot \nabla \times \boldsymbol{w} - \boldsymbol{w} \cdot \nabla \times (\nabla_\perp u \times \boldsymbol{w}) \\ &= (\nabla_\perp u \times \boldsymbol{w}) \cdot \nabla \times \boldsymbol{w} - \boldsymbol{w} \cdot \left[\hat{\boldsymbol{w}}\,(\hat{\boldsymbol{w}} \cdot \nabla) - \hat{\boldsymbol{w}} \times (\hat{\boldsymbol{w}} \times \nabla)\right] \times (\nabla_\perp u \times \boldsymbol{w}) \\ &= (\nabla_\perp u \times \boldsymbol{w}) \cdot \nabla \times \boldsymbol{w} - \boldsymbol{w} \cdot \nabla_\perp \times (\nabla_\perp u \times \boldsymbol{w}).\end{aligned} \tag{10.35}$$

Here, we decomposed the differential operator ∇ in parallel and normal components as:

$$\begin{aligned}\nabla &= \nabla_\parallel + \nabla_\perp \\ &= \hat{\boldsymbol{w}}\,(\hat{\boldsymbol{w}} \cdot \nabla) - \hat{\boldsymbol{w}} \times (\hat{\boldsymbol{w}} \times \nabla),\end{aligned} \tag{10.36}$$

with $\hat{\boldsymbol{w}} = \boldsymbol{w}/w$.

Using the representation of the differential operator ∇ of equation (10.36) and the decomposition of the curl of a vector field (4.51), we want to determine estimates for the normal and parallel gradients $\nabla_\perp$ and $\nabla_\parallel$. Take $u \in L_0^2(\Omega)$ and let $\hat{\boldsymbol{w}} = \boldsymbol{w}/w$. Futhermore, let $||\cdot||$ be the $L^2(\Omega)$ norm. Then, the following identity holds:

$$\begin{aligned}
||\nabla \cdot (u\hat{\boldsymbol{w}})||^2 &= \int_\Omega |\nabla \cdot (u\hat{\boldsymbol{w}})|^2 \, vol^3 \\
&= \int_\Omega \left[(\nabla u \cdot \hat{\boldsymbol{w}})^2 + 2u (\nabla \cdot \hat{\boldsymbol{w}}) (\nabla u \cdot \hat{\boldsymbol{w}}) + u^2 (\nabla \cdot \hat{\boldsymbol{w}})^2 \right] vol^3 \\
&= \left|\left|\nabla_\parallel u\right|\right|^2 + \int_\Omega (\nabla \cdot \hat{\boldsymbol{w}}) \left(\hat{\boldsymbol{w}} \cdot \nabla u^2\right) vol^3 + ||u (\nabla \cdot \hat{\boldsymbol{w}})||^2 \\
&= \left|\left|\nabla_\parallel u\right|\right|^2 - \int_\Omega u^2 \nabla \cdot [\hat{\boldsymbol{w}} (\nabla \cdot \hat{\boldsymbol{w}})] \, vol^3 + ||u (\nabla \cdot \hat{\boldsymbol{w}})||^2 \\
&= \left|\left|\nabla_\parallel u\right|\right|^2 - \int_\Omega u^2 \hat{\boldsymbol{w}} \cdot \nabla (\nabla \cdot \hat{\boldsymbol{w}}) \, vol^3.
\end{aligned} \tag{10.37}$$

Here we used the fact that since $u \in L_0^2(\Omega)$, $u = 0$ on $\partial\Omega$. Therefore:

$$\left|\left|\nabla_\parallel u\right|\right|^2 = ||\nabla \cdot (u\hat{\boldsymbol{w}})||^2 + \int_\Omega u^2 \hat{\boldsymbol{w}} \cdot \nabla (\nabla \cdot \hat{\boldsymbol{w}}) \, vol^3. \tag{10.38}$$

Now suppose that $\forall \boldsymbol{x} \in \Omega$ we have $\hat{\boldsymbol{w}} \cdot \nabla (\nabla \cdot \hat{\boldsymbol{w}}) > 0$. Then, we obtain the preliminary estimate:

$$\left|\left|\nabla_\parallel u\right|\right|^2 \geq \min_{\boldsymbol{x} \in \Omega} [\hat{\boldsymbol{w}} \cdot \nabla (\nabla \cdot \hat{\boldsymbol{w}})] \, ||u||^2 . \tag{10.39}$$

The quantity $\hat{h} = \hat{\boldsymbol{w}} \cdot \nabla (\nabla \cdot \hat{\boldsymbol{w}})$ can be related to the field force divergence $\hat{\mathfrak{B}} = \nabla \cdot \hat{\boldsymbol{b}}$ of $\hat{\boldsymbol{w}}$ in the following manner:

$$\begin{aligned}
\hat{h} &= \hat{\boldsymbol{w}} \cdot \nabla (\nabla \cdot \hat{\boldsymbol{w}}) \\
&= \nabla \cdot [\hat{\boldsymbol{w}} (\nabla \cdot \hat{\boldsymbol{w}})] - |\nabla \cdot \hat{\boldsymbol{w}}|^2 \\
&= \nabla \cdot [\nabla \cdot (\hat{\boldsymbol{w}}\hat{\boldsymbol{w}})] - \nabla \cdot [\hat{\boldsymbol{w}} \cdot \nabla \hat{\boldsymbol{w}}] - |\nabla \cdot \hat{\boldsymbol{w}}|^2 \\
&= \nabla \cdot [\hat{\boldsymbol{w}} \times (\nabla \times \hat{\boldsymbol{w}})] + \nabla \cdot [\nabla \cdot (\hat{\boldsymbol{w}}\hat{\boldsymbol{w}})] - |\nabla \cdot \hat{\boldsymbol{w}}|^2 \\
&= \mathfrak{B} + \nabla \cdot [\nabla \cdot (\hat{\boldsymbol{w}}\hat{\boldsymbol{w}})] - |\nabla \cdot \hat{\boldsymbol{w}}|^2 .
\end{aligned} \tag{10.40}$$

In figure 10.2 we show the plot of $\hat{h}$ for:

$$\hat{\boldsymbol{w}} = \frac{1}{\sqrt{1 + \cos x^2}} (\cos y, \cos x, \sin y) . \tag{10.41}$$

Note that this vector field is such that $\hat{w} = 1$.

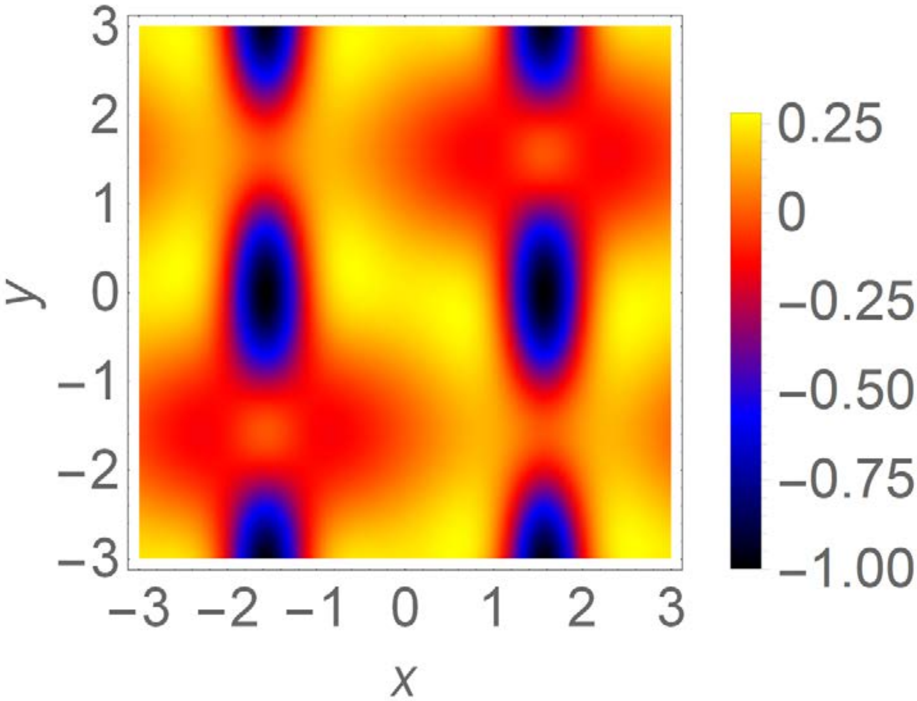

Figure 10.2: Plot of $\hat{h} = \hat{\boldsymbol{w}} \cdot \nabla (\nabla \cdot \hat{\boldsymbol{w}})$ for $\hat{\boldsymbol{w}}$ given by equation (10.41).

Consider now the normal gradient. Again, take $u \in L^2_0(\Omega)$. The following identity holds:

$$\begin{aligned}
||\nabla \times (u\hat{\boldsymbol{w}})||^2 &= \int_\Omega u\hat{\boldsymbol{w}} \cdot \nabla \times \nabla \times (u\hat{\boldsymbol{w}})\, vol^3 \\
&= \int_\Omega \left[u\hat{\boldsymbol{w}} \cdot \nabla \times (\nabla u \times \hat{\boldsymbol{w}} + u\nabla \times \hat{\boldsymbol{w}})\right] vol^3 \\
&= \int_\Omega u\hat{\boldsymbol{w}} \cdot \left[\nabla \times (\nabla u \times \hat{\boldsymbol{w}}) + \nabla u \times (\nabla \times \hat{\boldsymbol{w}}) + u\nabla \times \nabla \times \hat{\boldsymbol{w}}\right] vol^3 \\
&= \int_\Omega \left[(\nabla u \times \hat{\boldsymbol{w}}) \cdot \nabla \times (u\hat{\boldsymbol{w}}) + u^2 \left(\frac{\hat{\mathfrak{B}}}{2} + \hat{\boldsymbol{w}} \cdot \nabla \times \nabla \times \hat{\boldsymbol{w}}\right)\right] vol^3 \\
&= ||\nabla_\perp u||^2 + \int_\Omega u^2 \left(\hat{\boldsymbol{w}} \cdot \nabla \times \nabla \times \hat{\boldsymbol{w}}\right) vol^3 \\
&= ||\nabla_\perp u||^2 + \int_\Omega u^2 \left\{\nabla \cdot [(\nabla \times \hat{\boldsymbol{w}}) \times \hat{\boldsymbol{w}}] + |\nabla \times \hat{\boldsymbol{w}}|^2\right\} vol^3 \\
&= ||\nabla_\perp u||^2 + \int_\Omega u^2 \left[-\hat{\mathfrak{B}} + |\nabla \times \hat{\boldsymbol{w}}|^2\right] vol^3
\end{aligned} \tag{10.42}$$

Therefore:

$$||\nabla_\perp u||^2 = ||\nabla \times (u\hat{\boldsymbol{w}})||^2 + \int_\Omega u^2 \left(\hat{\mathfrak{B}} - |\nabla \times \hat{\boldsymbol{w}}|^2\right) vol^3. \tag{10.43}$$

By combining equations (10.38) and (10.43) one obtains the identity:

$$||\nabla u||^2 = ||\nabla \cdot (u\hat{\boldsymbol{w}})||^2 + ||\nabla \times (u\hat{\boldsymbol{w}})||^2 + \int_\Omega u^2 \hat{\boldsymbol{w}} \cdot \Delta \hat{\boldsymbol{w}} \, vol^3. \tag{10.44}$$

This equation clearly shows the geometrical relationship between the $L_0^2(\Omega)$ norm of the gradient of a function, and the integrability of a vector field $\boldsymbol{w}$. Indeed, the first two terms on the right-hand side vanish when $u\hat{\boldsymbol{w}} = \nabla \times \boldsymbol{\xi}$ for some appropriate vector $\boldsymbol{\xi}$ and $u\hat{\boldsymbol{w}} = \nabla C$ for some appropriate function C respectively. Furthermore, note that $||\nabla u||^2$ is made up of three terms which reflect the Hodge decomposition of differential forms in terms of closed, co-closed and harmonic components.

Using equation (10.43) we can prove the following estimate:

***Proposition* 10.4.** *(Estimates for the normal gradient)*
Take $u \in L_0^2(\Omega)$ and let $\boldsymbol{w}$ be a vector field with components in $C^\infty(\Omega)$ and $\mathfrak{G}^{xyz} = \boldsymbol{w} \cdot \nabla \times \boldsymbol{w} \neq 0$. Suppose that $\forall \boldsymbol{x} \in \Omega$ we have $\hat{\mathfrak{B}} - \hat{\boldsymbol{b}}^2 = \nabla \cdot [\hat{\boldsymbol{w}} \times (\nabla \times \hat{\boldsymbol{w}})] - |\hat{\boldsymbol{w}} \times (\nabla \times \hat{\boldsymbol{w}})|^2 > 0$. Then, the following estimate holds:

$$||\nabla_\perp u||^2 \geq \min_{\boldsymbol{x} \in \Omega} \left(\hat{\mathfrak{B}} - \hat{\boldsymbol{b}}^2 \right) ||u||^2 . \tag{10.45}$$

Proof. First, note that if $\mathfrak{G}^{xyz} \neq 0$ also $\hat{\mathfrak{G}}^{xyz} \neq 0$:

$$\mathfrak{G}^{xyz} = \boldsymbol{w} \cdot \nabla \times \boldsymbol{w} = \frac{\hat{\boldsymbol{w}}}{w} \cdot \nabla \times \left(\frac{\hat{\boldsymbol{w}}}{w} \right) = \frac{\hat{\mathfrak{G}}^{xyz}}{w^2}. \tag{10.46}$$

Thus, the non-integrability of $\boldsymbol{w}$ implies the non-integrability of $\hat{\boldsymbol{w}}$. Thereby, we can perform all calculations in terms of $\hat{\boldsymbol{w}}$ without loss of generality. To simplify the notation we shall omit the apex xyz in $\hat{\mathfrak{G}}^{xyz}$, and write just $\hat{\mathfrak{G}}$. The estimate (10.45) is a direct consequence of equation (10.43). To see this, observe that from (4.51):

$$|\nabla \times \hat{\boldsymbol{w}}|^2 = \hat{\boldsymbol{b}}^2 + \hat{\mathfrak{G}}^2. \tag{10.47}$$

Furthermore:

$$u\hat{\boldsymbol{w}} \cdot [\nabla \times (u\hat{\boldsymbol{w}})] = u^2 \hat{\mathfrak{G}} = u \, |\nabla \times (u\hat{\boldsymbol{w}})| \cos\theta. \tag{10.48}$$

Here θ is an angle depending on the choice of u. Then:

$$|\nabla \times (u\hat{\boldsymbol{w}})|^2 = \frac{u^2 \hat{\mathfrak{G}}^2}{\cos^2\theta}. \tag{10.49}$$

Combining this result with (10.43) and (10.47), we arrive at:

$$||\nabla_\perp u||^2 = \int_\Omega u^2 \left[\hat{\mathfrak{G}}^2 \left(\frac{1}{\cos^2\theta} - 1 \right) + \hat{\mathfrak{B}} - \hat{\boldsymbol{b}}^2 \right] vol^3. \tag{10.50}$$

Note that, for any θ, $(\cos\theta)^{-2} - 1 \geq 0$. Therefore, if $\hat{\mathfrak{B}} - \hat{\boldsymbol{b}}^2 > 0 \ \forall \boldsymbol{x} \in \Omega$, we arrive at (10.45). $\square$

In figure 10.3 we show a plot of $\hat{\mathfrak{B}} - \hat{\boldsymbol{b}}^2$ (with $\hat{\boldsymbol{w}}$ given by equation (10.41)) concerning the estimate for the normal gradient, equation (10.45). In figure 10.4 we show a plot of the field force divergence $\hat{\mathfrak{B}}$ for $\hat{\boldsymbol{w}}$ given by equation (10.41). The shape of this distribution should be compared to figure 10.5, where a plot of the density resulting by solving the equation of motion $X = \hat{\boldsymbol{w}} \times \Gamma$ for a system of $8 \cdot 10^6$ particles is shown (for the details of this simulation see section 9.2.7). This profile corresponds to the stationary solution of the diffusion equation (9.10). As already noted, there is a good resemblance between the profile of $\hat{\mathfrak{B}}$ and the calculated density profile. This is a notable fact, since the absolute value of $\hat{\boldsymbol{w}}$, representing the 'strength' of the fluctuations, is unity $\hat{w} = 1$. This does not prevent the creation of an ordered structure, which is determined by the non-vanishing field force divergence $\hat{\mathfrak{B}}$.

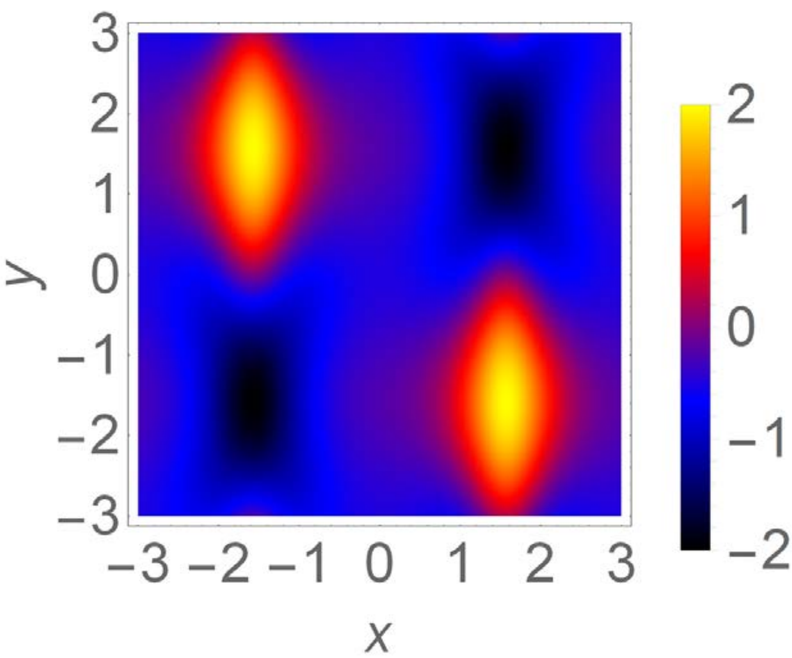

Figure 10.3: Plot of $\hat{\mathfrak{B}} - \hat{b}^2$ for $\hat{\boldsymbol{w}}$ given by equation (10.41). Note that there are several regions where $\hat{\mathfrak{b}} - \hat{b}^2 > 0$. In such domains the estimate (10.45)holds.

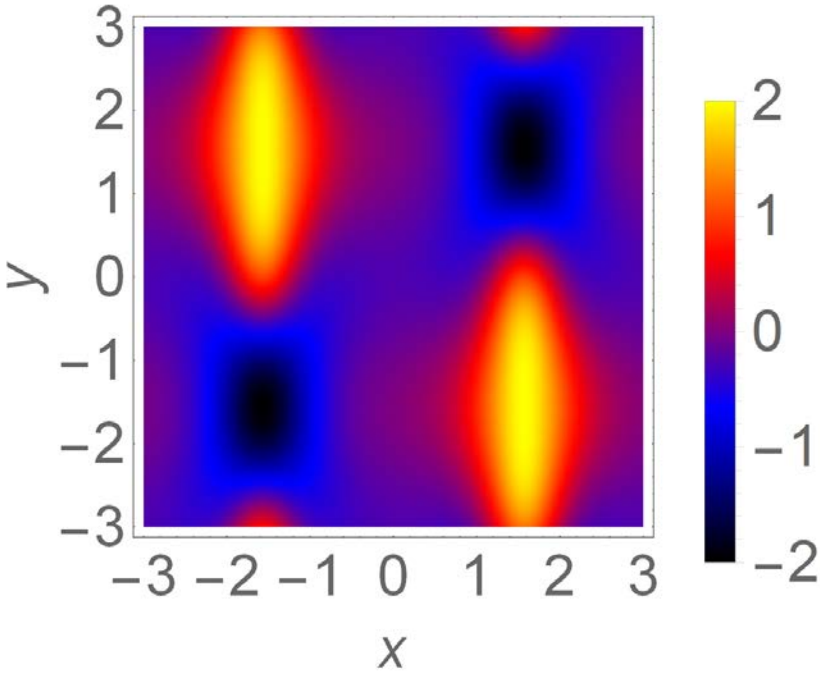

Figure 10.4: Plot of $\hat{\mathfrak{B}}$ for $\hat{\boldsymbol{w}}$ given by equation (10.41).

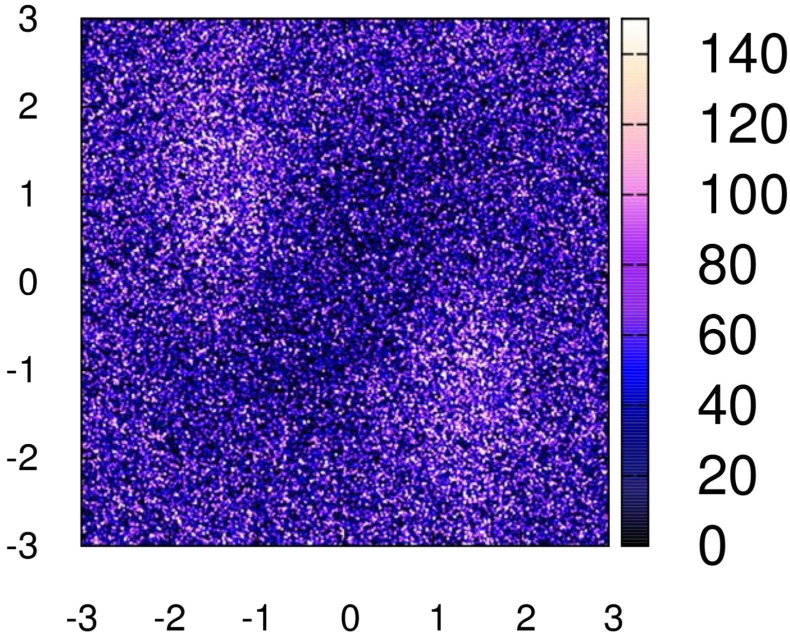

Figure 10.5: Calculated particle density in (x, y) plane.

Conclusion

In the present study, dynamical and statistical properties of conservative mechanical systems were investigated.

The orbit of a conservative system lies on the level set of the energy (Hamiltonian function) and represents the building block of *macroscopic* phenomena. The special nature of a macroscopic system originates from the process of *reduction* in which the underlying *microscopic* dynamics is portrayed by 'selecting' the (dynamically) relevant degrees of freedom and by removing redundancies. The macroscopic description of a physical system must not be intended as an artificial mathematical construction: it represents the result of the superposition of microscopic degrees of freedom on the spatial and times scales that define the identity of the macroscopic system itself.

The dynamics of a microscopic system is naturally written in the language of canonical Hamiltonian mechanics. A macroscopic system emerges when constraints are imposed on the 'flat' phase space of microscopic degrees of freedom. Such *topological constraints* may destroy the canonical Hamiltonian form. Integrable constraints foliate the phase space and dictate a non-canonical Hamiltonian structure represented by a Poisson operator. Non-integrable constraints impart 'current' to the metric of space and induce a conservative structure with an associated antisymmetric operator.

Thereby, a statistical theory of macroscopic phenomena must take into account the non-trivial topology of space-time dictated by topological constraints. The achievement of the present work is the development of such statistical theory and the construction of the mathematical tools required to attain this novel formulation.

It is shown that the topology of space is directly reflected in the thermodynamically consistent entropy measure, which now explicitly depends on the geometrical properties (degeneracy and current) of the antisymmetric operator that acts on the

Hamiltonian function to generate the dynamics. The probability distribution resulting from the maximum entropy principle changes accordingly and therefore departs from the standard Maxwell-Boltzmann distribution of canonical Hamiltonian mechanics. The physics and thermodynamics of self-organizing phenomena arising from topological constraints are thus explained.

The present research offers new perspectives from the mathematical point of view. It is shown that the dynamics of conservative mechanical systems is endowed with a geometrical hierarchy reflecting the properties of the antisymmetric operator. Each of the new operators (measure preserving and Beltrami) introduced in this study exhibits peculiar dynamical and statistical properties. In particular, it is found that the standard results of statistical mechanics can be extended to the class of measure preserving operators. This fact is remarkable, because such operators do not posses an Hamiltonian structure and the canonical phase space dictated by Liouville's theorem. Furthermore, it is demonstrated that entropy maximization can occur even in the absence of an invariant measure, as is in the case of diffusion driven by a Beltrami operator. These results suggest that, beyond Hamiltonian mechanics, an entire unexplored world deserving further investigation exists.

The normal Laplace operator discussed in the last chapter is also a novel object of special mathematical interest: this operator shows a clear interplay between integrability in the context of differential geometry and the study of non-elliptic partial differential equations.

List of Publications

Reviewed Journal Papers:

1. N. Sato and Z. Yoshida, *A Stochastic Model of Inward Diffusion in Magnetospheric Plasmas*, J. Phys. A: Math. Theor., **48**, 205501 (2015), doi:10.1088/1751-8113/48/20/205501.

2. N. Sato, N. Kasaoka, and Z. Yoshida, *Thermal Equilibrium of Non-Neutral Plasma in Dipole Magnetic Field*, Phys. Plasmas, **22**, 042508 (2015), doi:10.1063/1.4917474.

3. N. Sato, Z. Yoshida, and Y. Kawazura, *Self-Organization and Heating by Inward Diffusion in Magnetospheric Plasmas*, Plasma and Fusion Research, **11**, 2401009 (2016), doi:10.1585/pfr.11.2401009.

4. N. Sato and Z. Yoshida, *Up-Hill Diffusion, Creation of Density Gradients: Entropy Measure for Systems with Topological Constraints*, Phys. Rev. E, **93**, 062140 (2016), doi:10.1103/PhysRevE.93.062140.

5. Y. Ushida, Y. Kawazura, N. Sato, and Z. Yoshida, *Inward Diffusion and Acceleration of Particles Driven by Turbulent Fluctuations in Magnetosphere*, Phys. Plasmas **23**, 114501 (2016), doi:10.1063/1.4967281.

Conferences and Talks:

Invited Talk:

1. N. Sato and Z. Yoshida, *Hamiltonian and Non-Hamiltonian Reductions of Charged Particle Dynamics: Diffusion and Self-Organization*, 12th International Workshop on Non-Neutral Plasmas, Lawrence University, Appleton, Wisconsin, USA, 11th July 2017.

Oral Talks:

2. N. Sato and R. L. Dewar, *Relaxation and Stability of Compressible Euler Flow in a Toroidal Domain*, The Japan Society of Fluid Mechanics 2017 annual meeting, Tokyo University of Science, Tokyo, Japan, 31th August-1st September 2017 (to be held).

3. N. Sato and Z. Yoshida, *Self-Organization of Macroscopic Structures and Entropy Production in Conservative Systems with Topological Constraints*, 9th Festival de Thorie: Avalanching and Self-Organization in Plasmas: 30 Years of BTW, Aix-en-Provence, France, 29th June 2017.

4. N. Sato and Z. Yoshida, *Statistical Mechanics of Almost Hamiltonian Systems with Measure Preserving Brackets*, 19pB31-9, JPS 72nd annual meeting, Osaka University, Osaka, Japan, 19th March 2017.

5. R. L. Dewar and N. Sato, *Recirculating flow of an Euler Fluid*, 35, ANZIAM 2017, The Adelaide Hills Convention Centre, Hahndorf, Australia, 5th February 2017.

6. N. Sato and R. L. Dewar, *Relaxed States in MRxMHD Plasmas*, MHD workshop 2016, National Institute for Fusion Science, Toki, Japan, 13th December 2016.

7. N. Sato and Z. Yoshida, *Entropy Measure for Systems with Topological Constraints*, Centre for Plasmas and Fluids, The Australian National University, Canberra, Australia, 27th September 2016.

8. N. Sato and Z. Yoshida, *Poisson Bracket Extension of Bracket Failing to Satisfy the Jacobi Identity*, 13aAK-11, JPS 71st annual meeting, Kanazawa University,

Kanazawa, Japan, 13th September 2016.

9. N. Sato and Z. Yoshida, *Generalized Hamiltonian Mechanics and Invariant Measure of Nonholonomic 3D dynamics*, 19aBC-8, JPS 71st annual meeting, Tohoku Gakuin University, Sendai, Japan, 19th March 2016.

10. N. Sato, Z. Yoshida, and Y. Kawazura, *Self-Organization and Heating by Inward Diffusion in Magnetospheric Plasmas*, O-8, 25th International Toki Conference (ITC-25), Ceratopia Toki, Toki, Gifu, Japan, 6th November 2015.

11. N. Sato and Z. Yoshida, *Diffusion and Entropy in Self-Organized Magnetospheric Plasmas*, 17pCN-7, JPS 70th annual meeting, Kansai University, Osaka, Japan, 17th September 2015.

12. N. Sato, N. Kasaoka, and Z. Yoshida, *Thermal Equilibrium of Non-Neutral Plasma in Dipole Magnetic Field*, 22pCM-7, JPS 70th annual meeting, Waseda University, Tokyo, Japan, 22nd March 2015.

13. N. Sato, Z. Yoshida, and Y. Kawazura, *A Stochastic Model of Inward Diffusion in Magnetospheric Plasmas*, Culham Center for Fusion Energy, Culham (UK), 12th December 2014.

14. N. Sato, Z. Yoshida, and Y. Kawazura, *A Stochastic Model of Inward Diffusion in Magnetospheric Plasmas*, Plasma Seminars and Group Meetings, Rudolf Peierls Centre for Theoretical Physics, University of Oxford (UK), 4th December 2014.

15. N. Sato, Z. Yoshida, and Y. Kawazura, *A Stochastic Model of Inward Diffusion in Magnetospheric Plasmas*, 18aD2-2, Plasma Conference 2014, Tokimesse, Niigata, Japan, 18th November 2014.

16. N. Sato, Z. Yoshida, Y. Kawazura, *A Stochastic Model of Inward Diffusion in Magnetospheric Plasmas*, 29pAT-5, JPS 69th annual meeting, Tokai University, Japan, 29th March 2014.

Awards:

1. *Outstanding Performance Award in Advanced Energy*, Department of Advanced Energy, Graduate School of Frontier Sciences, The University of Tokyo (March 2015).

Fellowships, Research Grants and Scholarships

1. JSPS research fellow (DC2, April 2016-March 2018).

2. *Musha-shugyou program*: scholarship provided by The University of Tokyo to conduct overseas training at Rudolf Peierls Centre for Theoretical Physics, University of Oxford (December 2014).

3. Scholarship provided by the British Institute of Florence to conduct research at Magdalene College, University of Cambridge (July 2012).

Acknowledgments

I would like to express my gratitude to Professor Z. Yoshida for his guidance throughout the present study. I am deeply grateful for his expert advice and continuous support, without which this project would have been impossible.

The research of N. Sato was supported by the JSPS Research Fellowship for Young Scientists, grant number $16J01486$.

I would like to thank Professor M. Nishiura and my colleagues Dr. Y Kawazura, Dr. Y. Ohno, Dr. H. M. Abdelhamid, and N. Kasaoka who have helped me greatly in different ways during this research.

I would like to thank Professor R. L. Dewar for his helpful collaboration and kind hospitality.

My gratitude goes to all present and past members of the plasma physics laboratory at the University of Tokyo. I would like to especially thank K. Kitayama for her help.

Next, I would like to thank Professor P. J. Morrison and Dr. H. Saitoh for their insightful suggestions.

I would like to thank Professor Z. Yoshida, Professor Y. Ogawa, Professor K. Suzuki, Professor M. Nishiura, and Professor N. Ito for the effort of reviewing this thesis.

Finally, I want to express my gratitude to my family, to A. Iwanaga for her continuous support, to Dr. D. Rossi, and to my friends.

Bibliography

[1] L. D. Landau and E. M. Lifshitz, *Mechanics*, 3rd ed. (Butterworth-Heinenann, Oxford, 1993), pp. 131-3.

[2] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed. (Springer, New York, 1989), pp. 201-4.

[3] V. I. Arnold, *Mathematical Aspects of Classical and Celestial Mechanics*, 3rd ed. (Springer, 2006), pp. 30-41.

[4] P. J. Morrison, *Poisson Brackets for Fluids and Plasmas*, in Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems, edited by M. Tabor and Y. Treve, AIP Conference Proceedings No. 88 (AIP, New York,1982), pp. 13-46.

[5] R. Littlejohn, *Singular Poisson Tensors*, in Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems, edited by M. Tabor and Y. Treve, AIP Conference Proceedings No. 88 (AIP, New York, 1982), pp. 47-66.

[6] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*, 2nd ed. (Springer,New York, 1999), pp. 6-8.

[7] P. J. Morrison, *Hamiltonian Description of the Ideal Fluid*, Rev. Mod. Phys., **70**, 467 (1998).

[8] P. J. Morrison and J. M. Greene, *Noncanonical Hamiltonian Density Formulation of Hydrodynamics and Ideal Magnetohydrodynamics*, Phys. Rev. Lett., **45**, 10 (1980).

[9] V. I. Arnold and B. A. Khesin, *Topological Methods in Hydrodynamics*, chap. 1 (Springer, New York, 1998).

[10] A. M. Bloch, *Nonholonomic Mechanics and Control*, 2nd ed. (Springer, New York, 2015), pp. 11-5.

[11] A. M. Bloch, J. E. Marsden, and D. V. Zenkov, *Nonholonomic Dynamics*, Notices of the AMS, **52**, pp. 320-9 (2005).

[12] M. de León, *A Historical Review on Nonholonomic Mechanics*, RACSAM, **106**, pp. 191-224 (2012).

[13] A. J. Van Der Shaft, B. M. Maschke, *On the Hamiltonian Formulation of Nonholonomic Mechanical Systems*, Reports on Mathematical Physics, **34**, 225-233 (1994).

[14] L. Bates and J. Sniatycki, *Nonholonomic Reduction*, Reports on Mathematical Physics, **32**, pp. 99-115 (1993).

[15] C. E. Caligan and C. Chandre, *Conservative Dissipation: How Important is the Jacobi Identity in the Dynamics?*, Chaos, **26**, 053101 (2016).

[16] M. de León, *Methods of Differential Geometry in Analytical Mechanics*, (Elsevier, New York, 1989), p. 234.

[17] N. Bourbaki, *Elements of Mathematics, Lie Groups and Lie Algebras, Chapters 1-3*, (Springer, New York, 1989), pp. 209-10.

[18] S. A. Chaplygin, *On the Theory of Motion of Nonholonomic Systems. The Reducing Multiplier Theorem*, Regular and Chaotic Dynamics, **13**, pp. 369-76 (2008).

[19] A. V. Borisov and I. S. Mamaev, *Chaplygin's Ball Rolling Problem Is Hamiltonian*, Mathematical Notes, **70**, 5, pp. 720-3 (2001).

[20] P. Balseiro, L. C. Garcia-Naranjo, *Gauge Transformations, Twisted Poisson Brackets and Hamiltonization of Nonholonomic Systems*, Arch. Rational Mech. Anal., **205**, pp. 267-310 (2012).

[21] P. Balseiro, *The Jacobiator of Nonholonomic Systems and the Geometry of Reduced Nonholonomic Brackets*, Arch. Rational Mech. Anal., **214**, pp. 453-501 (2014).

[22] P. Balseiro, and N. Sansonetto, *A Geometric Characterization of Certain First Integrals for Nonholonomic Systems with Symmetries*, Symmetry, Integrability and Geometry: Methods and Applications, **12**, 018 (2016).

[23] K. Huang, *Statistical Mechanics*, 2nd ed., (Wiley and Sons, 1987), pp. 52-65.

[24] J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, chap. 1 (Scribner's Sons, New York, 1902), pp. 3-19.

[25] G. D. Birkhoff, *Proof of the Ergodic Theorem*, Proc. Nat. Acad. Sci. U.S.A., **17**, 12, pp. 656-60 (1931).

[26] YA. G. Sinai, *Topics in Ergodic Theory*, (Princeton University Press, Princeton, 1994), pp. 3-15.

[27] J. V. Neumann, *Proof of the Quasi-Ergodic Hypothesis*, Proc. Nat. Ac. Sci., **18**, pp. 70-82 (1932).

[28] J. V. Neumann, *Physical Applications of the Ergodic Hypothesis*, Proc. Nat. Ac. Sci., **18**, pp. 263-6 (1932).

[29] G. D. Birkhoff, *What is the Ergodic Theorem?*, The American Mathematical Monthly, **49**, 4 (1942).

[30] P. H. Chavanis, *Statistical Mechanics of Two-Dimensional Vortices and Stellar Systems*, in Dynamics and Thermodynamics of Systems with Long-Range Interactions, (Springer, New York, 2002), pp. 208-89.

[31] A. M. Persoon, D. A. Gurnett, W. S. Kurth, G. B. Hospodarsky, J. B. Groene, P. Canu, and M. K. Dougherty, *Equatorial Electron Density Measurements in Saturn's Inner Magnetosphere*, Geophys. Res. Lett., **32**, L23105 (2005).

[32] A. M. Persoon *et al.*, *A Diffusive Equilibrium Model for the Plasma Density in Saturn's Magnetosphere*, J. Geophys. Res., **114**, A04211 (2009).

[33] A. M. Persoon *et al.*, *The Plasma Density Distribution in the Inner Region of Saturn's Magnetosphere*, J. Geophys. Res.: Space Phys., **118**, pp. 2970-4 (2013).

[34] A. N. Kolmogorov, *The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds Numbers*, Proc. R. Soc. Lond. A, **434**, pp. 9-13 (1991).

[35] A. N. Kolmogorov, *A Refinement of Previous Hypothesis Concerning the Local Structure of Turbulence in a Viscous Incompressible Fluid at High Reynolds Number*, J. Fluid Mech., **13**, pp. 82-5 (1962).

[36] A. N. Kolmogorov, *Dissipation of Energy in the Locally Isotropic Turbulence*, Proc. R. Soc. Lon. A, **434**, pp. 15-7 (1991).

[37] H. K. Moffatt, *Helicity and Singular Structures in Fluid Dynamics*, Proc. Nat. Acad. Sci., **111**, 10, pp. 3663-70 (2014).

[38] A. Hasegawa, *Self-Organization Processes in Continuous Media*, Adv. Phys. **34**, (1985).

[39] L. D. Landau and E. M. Lifshitz, *On the Theory of the Dispersion of Magnetic Permeability in Ferromagnetic Bodies*, Ukr. J. Phys., **53**, pp. 14-22 (2008).

[40] T. L. Gilbert, *A Phenomenological Theory of Damping in Ferromagnetic Materials*, IEEE Trans. Magnetics, **40**, 6 (2004).

[41] H. Bondi, *The Rigid Body Dynamics of Unidirectional Spin*, Proc. R. Soc. Lond. A, **405**, pp. 255-74 (1986).

[42] H. K. Moffatt and T. Tokieda, *Celt Reversal: a Prototype of Chiral Dynamics*, Proc. R. Soc. Edin., **138A**, pp. 361-8 (2008).

[43] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. Part 1 (Butterworth-Heinenann, Oxford, 1980), pp. 29-33.

[44] K. Huang, *Statistical Mechanics*, 2nd ed., (Wiley and Sons, 1987), pp. 9-11.

[45] D. L. Goodstein, *States of Matter*, (Dover, New York, 1985), p. 12.

[46] J. B. Boyling, *An Axiomatic Approach to Classical Thermodynamics*, Proc. R. Soc. Lond. A., **329**, pp. 35-70 (1972).

[47] S. Camazine *et al.*, *Self-Organization in Biological Systems*, 1st ed. (Princeton University Press, 2003).

[48] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems-From Dissipative Structures to Order Through Fluctuations*, (Wiley & Sons, 1977).

[49] S. Kaufmann, *Investigations*, (Oxford University Press, 2002).

[50] M. Kyan *et al.*, *Unsupervised Learning-A Dynamic Approach*, (Wiley & Sons, 2014).

[51] T. Ikegami and M. Oka, *Massive Data Flows*, Proc. of the 6th International Conference on Agents and Artificial Intelligence, pp. 237-42 (2014).

[52] R. J. Goldston and P. H. Rutherford, *Introduction to Plasma Physics*, (Institute of Physics, Bristol and Philadelphia, 1995), pp. 1-3.

[53] F. F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, 2nd ed. (Plenum Press, New York, 1984), pp. 3-4.

[54] D. H. E. Dubin and T. M. O'Neil, *Trapped Nonneutral Plasmas, Liquids, and Crystals (The Thermal Equilibrium States)*, Rev. Mod. Phys., **71**, 87 (1999).

[55] J. H. Malmberg and J. S. de Grassie, *Properties of Nonneutral Plasma*, Phys. Rev. Lett., **35**, 577 (1975).

[56] J. H. Malmberg and C. F. Driscoll, *Long-Time Containment of a Pure Electron Plasma*, Phys. Rev. Lett., **44**, 654 (1980).

[57] J. Wesson, *Tokamaks*, 4th ed. (Oxford University Press, Oxford, 2011).

[58] K. Miyamoto, *Controlled Fusion and Plasma Physics*, (Taylor & Francis, 2007), pp. 131-8.

[59] A. Hasegawa, *Theory of Longitudinal Plasma Instabilities*, Physical Review, **169**, 1 (1968).

[60] R. J. Hosking and R. L. Dewar, *Fundamental Fluid Mechanics and Magnetohydrodynamics*, chap. 6 (Springer, 2016).

[61] I. B. Bernstein *et al.*, *An Energy Principle for Hydromagnetic Stability Problems*, Proc. R. Soc. Lon. A **244**, pp. 17-40 (1958).

[62] T. H. Stix, *Waves in Plasmas*, (American Institute of Physics, New York, 1992), pp. 217-8.

[63] S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Stability*, (Dover, New York, 1981).

[64] S. B. Pope, *Turbulent Flows*, (Cambridge University Press, Cambridge, 2000), pp. 182-91.

[65] A. Hasegawa, L. Chen, and M. E. Mauel, *A* $D-^3He$ *Fusion Reactor Based on a Dipole Magnetic Field* , Nuc. Fus., **30**, 11 (1990).

[66] A. Hasegawa, *Motion of a Charged Particle and Plasma Equilibrium in a Dipole Magnetic Field - Can a Magnetic Field Trap a Charged Particle? Can a Magnetic Field Having Bad Curvature Trap a Plasma Stably?*, Phys. Scripta, **T116**, pp. 72-4 (2005).

[67] Z. Yoshida *et al.*, *Magnetospheric Vortex Formation: Self-Organized Confinement of Charged Particles*, Phys. Rev. Lett., **104**, 235004 (2010).

[68] Z. Yoshida *et al.*, *Self-Organized Confinement by Magnetic Dipole: Recent Results From RT-1 and Theoretical Modeling*, Plasma Phys. Control. Fusion, **55**, 014018 (2013).

[69] A. C. Boxer *et al.*, *Turbulent Inward Pinch of Plasma Confined by a Levitated Dipole Magnet*, Nature Physics, **6**, pp. 207-12 (2010).

[70] T. S. Pedersen and A. H. Boozer, *Confinement of Nonneutral Plasmas on Magnetic Surfaces*, Phys. Rev. Lett., **88**, 20 (2002).

[71] T. S. Pedersen *et al.*, *Prospects for the Creation of Position-Electron Plasmas in a Non-Neutral Stellarator*, J. Phys. B: At. Mol. Opt. Phys., **36**, pp. 1029-39 (2003).

[72] H. Saitoh *et al.*, *Confinement of Pure-Electron Plasmas in a Toroidal Magnetic-Surface Configuration*, Phys. Rev. Lett., **92**, 25 (2004).

[73] H. Saitoh *et al.*, *Efficient Injection of an Intense Positron Beam into a Dipole Magnetic Field*, New J. Phys., **17** 103038 (2015).

[74] J. R. Cary and A. J. Brizard, *Hamiltonian Theory of Guiding-Center Motion*, Rev. Mod. Phys., **81**, 693 (2009).

[75] R. G. Littlejohn, *Hamiltonian Formulation of Guiding Center Motion*, Phys. Fluids, **24**, 1730 (1981).

[76] M. Schulz and L. J. Lanzerotti, *Particle Diffusion in the Radiation Belts*, (Springer, New York, 1974).

[77] T. J. Birmingham, T. G. Northrop, and C. G. Fälthammar, *Charged Particle Diffusion by Violation of the Third Adiabatic Invariant*, The Physics of Fluids, **10**, 11 (1967).

[78] T. G. Northrop, *Stability of the Adiabatic Motion of Charged Particles in the Earth's Field*, Phys. Rev., **117**, 1 (1960).

[79] Z. Yoshida and S. M. Mahajan, *Self-Organization in Foliated Phase Space: Construction of a Scale Hierarchy by Adiabatic Invariants of Magnetized Particles*, Prog. Theor. Exp. Phys., **2014**, 073J01 (2014).

[80] Z. Yoshida, *Self-Organization by Topological Constraints: Hierarchy of Foliated Phase Space*, Adv. Phys. X, **1**, pp. 2-19 (2016).

[81] Z. Yoshida and P. J. Morrison, *A Hierarchy of Noncanonical Hamiltonian Systems: Circulation Laws in an Extended Phase Space*, Fluid Dyn. Res., **46**, 031412 (2014).

[82] N. Sato and Z. Yoshida, *A Stochastic Model of Inward Diffusion in Magnetospheric Plasmas*, J. Phys. A: Math. Theor., **48**, 205501 (2015), doi:10.1088/1751-8113/48/20/205501.

[83] N. Sato, N. Kasaoka, and Z. Yoshida, *Thermal Equilibrium of Non-Neutral Plasma in Dipole Magnetic Field*, Phys. Plasmas, **22**, 042508 (2015), doi:10.1063/1.4917474.

[84] N. Sato, Z. Yoshida, and Y. Kawazura, *Self-Organization and Heating by Inward Diffusion in Magnetospheric Plasmas*, Plasma and Fusion Research, **11**, 2401009 (2016), doi:10.1585/pfr.11.2401009.

[85] N. Sato and Z. Yoshida, *Up-Hill Diffusion, Creation of Density Gradients: Entropy Measure for Systems with Topological Constraints*, Phys. Rev. E, **93**, 062140 (2016), doi:10.1103/PhysRevE.93.062140.

[86] Y. Ushida, Y. Kawazura, N. Sato, and Z. Yoshida, *Inward Diffusion and Acceleration of Particles Driven by Turbulent Fluctuations in Magnetosphere*, Phys. Plasmas **23**, 114501 (2016).

[87] A. Sergi and M. Ferrario, *Non-Hamiltonian Equations of Motion With a Conserved Energy*, Phys. Rev. E, **64**, 056125 (2001).

[88] A. Sergi, *Non-Hamiltonian Equilibrium Statistical Mechanics*, Phys. Rev. E, **67**, 021101 (2003).

[89] M. E. Tuckerman *et al.*, *Non-Hamiltonian Molecular Dynamics: Generalizing Hamiltonian Phase Space Principles to Non-Hamiltonian Systems*, J. Chem. Phys., **115**, 4 (2001).

[90] G. S. Erza, *On the Statistical Mechanics of Non-Hamiltonian Systems: the Generalized Liouville Equation, Entropy, and Time-Dependent Metrics*, J. Math. Chem., **35**, 1 (2004).

[91] T. Frankel, *The Geometry of Physics, An Introduction*, 3rd ed. (Cambridge University Press, Cambridge, 2012), pp. 165-78.

[92] I. Agricola and T. Friedrich, *Global Analysis, Differential Forms in Analysis, Geometry and Physics*, (American Mathematical Society, Providence, 2002), pp. 111-20.

[93] S. Kobayashi and K. Nomizu, *Foundations of Differential Geomtry*, (John Wiley & Sons, 1969), pp. 321-4.

[94] M. de León, *Methods of Differential Geometry in Analytical Mechanics*, (Elsevier, New York, 1989), p. 250-3.

[95] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed. (Springer, New York, 1989), pp. 230-2.

[96] R. Abraham and J. E. Marsden, *Foundations of Mechanics*, 2nd ed. (Addison-Wesley, 1978), p. 175.

[97] S. Chandrasekhar and L. Woltjer, *On Force-Free Magnetic Fields*, Proc. Nat. Ac. Sci., **44**, 4 (1958).

[98] L. Woltjer, *A Theorem on Force Free Magnetic Fields*, Proc. Nat. Ac. Sci., **44**, 6 (1958).

[99] J. B. Taylor, *Relaxation of Toroidal Plasma and Generation of Reverse Magnetic Fields*, Phys. Rev. Lett., **33**, 19 (1974).

[100] J. B. Taylor, *Relaxation and Magnetic Reconnection in Plasmas*, Rev. Mod. Phys., **58**, 3 (1986).

[101] H. K. Moffatt, *Magnetostatic Equilibria and Analogous Euler Flows of Arbitrarily Complex Topology. Part 1. Fundamentals*, J. Fluid Mech., **159**, pp. 359-78 (1985).

[102] H. K. Moffatt, *Magnetostatic Equilibria and Analogous Euler Flows of Arbitrarily Complex Topology. Part 2. Stability Considerations*, J. Fluid Mech., **166**, pp. 359-78 (1986).

[103] L. C. Steinhauer and A. Ishida, *Relaxation of a Two-Specie Magnetofluid*, Phys. Rev. Lett., **79**, 18 (1997).

[104] E. T. Jaynes, *Probability Theory, The Logic of Science*, (Cambridge University Press, Cambridge, 2003), pp. 374-8.

[105] E. T. Jaynes, *Information Theory and Statistical Mechanics*, Phys. Rev., **106**, 4, pp. 620-30 (1957).

[106] E. T. Jaynes, *Information Theory and Statistical Mechanics. II*, Phys. Rev., **108**, 2, pp. 171-90 (1957).

[107] J. D. Ramshaw, *Remarks on Entropy and Irreversibility in Non-Hamiltonian Systems*, Phys. Lett. A,**116**, 3 (1986).

[108] C. E. Shannon, *A Mathematical Theory of Communication*, **27**, 3 (1948).

[109] C. Tsallis, *Possible Generalization of Boltzmann-Gibbs Statistics*, J. Stat. Phys., **52**, 1/2 (1988).

[110] E. T. Jaynes, *The Minimum Entropy Principle*, Ann. Rev. Phys. Chem. **31**, pp. 579-601 (1980).

[111] R. C. Dewar, *Maximum Entropy Production and the Fluctuation Theorem*, J. Phys. A: Math. Gen., **38**, pp. L371-L381 (2005).

[112] R. C. Dewar, *A Theoretical Basis for Maximum Entropy Production* in Beyond the Second Law, Entropy Production and Non-Equilibrium Systems, (Springer, 2014), pp. 49-71.

[113] Z. Yoshida and S. M. Mahajan, *Variational Principles and Self-Organization in Two-Fluid Plamsas*, Phys. Rev. Lett., **88**, 9 (2002).

[114] Z. Yoshida and S. M. Mahajan, *'Maximum' Entropy Production in Self-Organized Plasma Boundary Layer: A Thermodynamic Discussion About Turbulent Heat Transport*, Phys. Plasmas, **15**, 032307 (2008).

[115] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, (Springer, 2001).

[116] P. D. Lax, *Functional Analysis*, (John Wiley & Sons, 2002).

[117] G. Bachman and L. Narici, *Functional Analysis*, (Dover, New York, 1966).

[118] W. Rudin, *Functional Analysis*, 2nd ed. (McGraw-Hill, 1991).

[119] A. Einstein, *On the Movement of Small Particles Suspended in Stationary Liquids Required by the Molecular-Kinetic Theory of Heat*, Annalen der Physic, **17**, pp. 549-60 (1905).

[120] A. Einstein, *On the Theory of the Brownian movement*, Annalen der Physic, **4**, 19 (1906).

[121] P. Langevin, *On The Theory Of Brownian Motion*, C. R. Acad. Sci. (Paris) **146**, pp. 530533 (1908).

[122] K. Ito, *Introduction to Probability Theory*, (Cambridge University Press, Cambridge, 1984).

[123] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 2nd ed., (Springer, 1985).

[124] L. C. Evans, *An Introduction to Stochastic Differential Equations*, (American Mathematical Society, 2014).

[125] K. Ito, *Stochastic Integral*, Proc. Imp. Acad., **20**, 8, pp. 519-24 (1944).

[126] K. Ito, *A Kinematic Theory of Turbulence*, Proc. Imp. Acad., **20**, 3, pp. 120-2 (1944).

[127] K. Ito, *Stochastic Differential Equations in a Differentiable Manifold*, Nagoya Math. J., **1**, pp. 35-47 (1950).

[128] K. Ito, *On a Formula Concerning Stochastic Differentials*, Nagoya Math. J., **3**, pp. 55-65 (1951).

[129] F. C. Klebaner, *Introduction to Stochastic Calculus with Applications*, 2nd ed., (Imperial College Press, 2005).

[130] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, (Chapman & Hall, 2004).

[131] H. Risken, *The Fokker-Planck Equation. Methods of Solution and Applications*, 2nd ed., (Springer, 1989).

[132] Z. Yoshida, *Nonlinear Science, The Challenge of Complex Systems*, (Springer, 2010).

[133] C. Godbillon, *Géométrie Différentielle et Mécanique Analytique*, (Hermann, Paris, 1969).

[134] R. C. Robinson, *Lectures on Hamiltonian Systems*, (Monografias De Matemática, **7**, Guanabara, 1971), pp. 51-3.

[135] J. Moser, *On the Volume Elements on a Manifold*, Trans. Am. Math. Soc., **120**, pp. 286-94 (1965).

[136] A. Weinstein, *Lectures on Symplectic Manifolds*, (CBMS Conf. Series, **29**, American Mathematical Society, 1977).

[137] T. Frankel, *The Geometry of Physics, An Introduction*, 3rd ed. (Cambridge University Press, Cambridge, 2012), pp. 182.

[138] I. M. Gelfand and I. Y. Dorfman, *The Schouten Bracket and Hamiltonian Operators*, Funct. Anal. Its Appl., **14**, pp. 223-6 (1980).

[139] M. Einsielder and T. Ward, *Ergodic Theory, with a View Towards Number Theory*, chap. 2 (Springer, London, 2011).

[140] T. Frankel, *The Geometry of Physics, An Introduction*, 3rd ed. (Cambridge University Press, Cambridge, 2012), pp. 181.

[141] A. J. Brizard and T. S. Ham, *Foundations of Nonlinear Gyrokinetic Theory*, Rev. Mod. Phys., **79**, 421 (2007).

[142] T. S. Ham, *Nonlinear Gyrokinetic Equations for Tokamak Microturbulence*, Phys. Fluids **31**, 2670 (1988).

[143] T. G. Northrop and E. Teller, *Stability of the Adiabatic Motion of Charged Particles in the Earth's Field*, Phys. Rev., **117**, 1 (1960).