Self-organization in foliated phase space: Construction of a scale hierarchy by adiabatic invariants of magnetized particles

Z. Yoshida[1],* and S. M. Mahajan[2]

[1]*Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba 277-8561, Japan*
[2]*Institute for Fusion Studies, The University of Texas at Austin, Austin, Texas 78712, USA*
*E-mail: yoshida@k.u-tokyo.ac.jp, mahajan@mail.utexas.edu

. . . .

Adiabatic invariants foliate phase space, and impart a macro-scale hierarchy by separating microscopic variables. On a macroscopic leaf, long-scale ordered structures are created while maximizing entropy. A plasma confined in a magnetosphere is invoked to unveil the organizing principle—in the vicinity of a magnetic dipole, the plasma self-organizes to a state with a steep density gradient. The resulting nontrivial structure has maximum entropy in an appropriate, constrained phase space. One could view such a phase space as a leaf foliated in terms of Casimir invariants—adiabatic invariants measuring the number of quasi-particles (macroscopic representation of periodic motions) are identified as the relevant Casimir invariants. The density clump is created in response to the inhomogeneity of the energy levels (frequencies) of the quasi-particles.

. . . .

Subject Index   A30, A47, J18, J21, J24

1. Introduction

The process of self-organization of a structure may appear to be an antithesis of the maximum entropy *ansatz*, yet various nonlinear systems display what may be viewed as the simultaneous existence of order and disorder. This co-existence begins to make sense if the self-organization processes and the entropy principle were to manifest on different scales; disorder can still develop at a microscopic scale while an ordered structure emerges on some appropriate macroscopic scale. Writing a theory of self-organization, then, will be an exercise in delineating and understanding the characteristic *scale hierarchy* of the physical system.

A biological body is a typical example in which an evident hierarchical structure is preprogrammed enabling effective consumption of energy and materials as well as emission of entropy and waste. A physical macro-system—a collective system of simple elements (a gravitational system, a plasma, etc.)—is anchored on a different framework. The automatic emergence of a scale hierarchy is not programmed; yet the controlling nonlinear dynamics can mimic a fundamental process of *creation*.

In this paper we develop a new framework to expose the creation process in action. The ordering principle is generally epitomized in a constraint—a possible conservation law—that, by restricting the class of motions available to the system, limits its ability to degenerate into general disorder. The *effective* phase space (ensemble) limited by such a constraint is the relevant *macro-hierarchy* on which nontrivial structures emerge. By invoking the geometrical notion of Hamiltonian mechanics,

we formulate a macro-hierarchy as a *Casimir leaf* of foliated phase space, i.e., the level-set of a *Casimir invariant* [1]. The connection between the notion of scale hierarchy and a Casimir invariant (an a priori geometrical structure of the phase space) is built by identifying a Casimir invariant as an *adiabatic invariant*; the adiabaticity criterion, then, determines what is macro. After the microscopic action is separated, the macroscopic object, which we call a *quasi-particle*, resides on a Casimir leaf. We will construct the Boltzmann distribution of quasi-particles on a Casimir leaf. The Casimir invariant represents the number of quasi-particles, which is the determinant of the corresponding *grand canonical ensemble*. Interestingly, heterogeneity (structure) is created by the distortion of the metric (invariant measure) dictating *equipartition* on the leaf.

Basic physical mechanisms and processes that embody our general framework of describing macro-hierarchy and self-organization will be brought to light via investigations of a magnetospheric plasma. Magnetospheric plasmas (naturally occurring ones such as the planetary magnetospheres [2–4], as well as their laboratory simulations [5–8]) are self-organized around the dipole magnetic fields in which charged particles cause a variety of interesting phenomena: the often observed *inward diffusion* (or uphill diffusion) of particles injected from the outer region is of particular interest. This process is driven by some spontaneous fluctuations (symmetry breaking) that violate the constancy of angular momentum. In a strong enough magnetic field, the canonical angular momentum P_θ is dominated by the magnetic part $q\psi$: the charge multiplied by the flux function [in the (r, θ, z) cylindrical coordinates, $\psi = r A_\theta$, where A_θ is the θ component of the vector potential]. The conservation of $P_\theta \approx q\psi$, therefore, restricts the particle motion to the magnetic surface (level-set of ψ). It is only via randomly phased fluctuations that the particles can diffuse across magnetic surfaces. Although the diffusion is normally a process that diminishes gradients, numerical experiments do exhibit preferential inward shifts through random motions of test particles [9,10]. Detailed specification of the fluctuations or the microscopic motion of particles is not the subject of the present effort. We plan to construct, instead, a clear-cut description of equilibria that maximize entropy while simultaneously bearing steep density gradients. Such an equilibrium will be formulated as a grand canonical distribution on a leaf of foliated phase space that represents a macro-hierarchy. In a strongly inhomogeneous magnetic field (typically a dipole magnetic field), the phase-space metric of *magnetized particles* is distorted; thus the projection of the equipartition distribution onto the flat space of the laboratory frame yields a peaked profile because of the connecting inhomogeneous Jacobian weight.

2. General Framework

2.1. Preliminaries: Poisson algebra and Casimir invariants

A general Hamiltonian system is endowed with a Poisson bracket $\{a, b\}$ satisfying antisymmetry $\{a, b\} = -\{b, a\}$, Jacobi's identity $\{\{a, b\}, c\} + \{\{b, c\}, a\} + \{\{c, a\}, b\} = 0$, and Leibniz's property $\{ab, c\} = a\{b, c\} + b\{a, c\}$. Let $z \in X = \mathbb{R}^n$ denote a state vector (here we assume that the phase space X is an n-dimensional Euclidean space), and ∂_z the gradient in X. An observable is a real function on X. We may represent a Poisson bracket as

$$\{a, b\} = \langle \partial_z a, \mathcal{J} \partial_z b \rangle, \tag{1}$$

where $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ is the inner product of X, and $\mathcal{J}$ (Poisson operator) is an antisymmetric $n \times n$ matrix (then the antisymmetry and Leibniz's property are satisfied, while Jacobi's identity is conditional [1]). Given a Hamiltonian H, the evolution of an observable $f(z)$ is described by

$$\frac{\mathrm{d}}{\mathrm{d}t} f = \{f, H\}. \tag{2}$$

In a *canonical* Hamiltonian system, the Poisson operator is a symplectic matrix; writing the state vector as $z = (q^1, p^1, \ldots, q^m, p^m)$,

$$\mathcal{J}_c := \begin{pmatrix} J_c & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_c \end{pmatrix}, \quad J_c := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3}$$

Because $\mathrm{Ker}(\mathcal{J}_c) = \{0\}$, the equilibrium point is given by $\partial_z H(z) = 0$. As we see in many examples of so-called weakly coupled systems, Hamiltonians are rather simple—they are often norms of the phase space—and thus the equilibrium points are at most trivial (remember the example of a harmonic oscillator).

A general Hamiltonian system may allow the Poisson operator $\mathcal{J}$ to be nontrivial; it may be a function of z, and, moreover, may have a nontrivial kernel $\mathrm{Ker}(\mathcal{J}) = \{u \in X; \; \mathcal{J}u = 0\}$. A nontrivial kernel introduces an essential *noncanonicality* to the system, and brings about interesting structures (Sect. 2.2). If

$$\partial_z C \in \mathrm{Ker}(\mathcal{J}), \tag{4}$$

such a $C(z)$ is called a *Casimir invariant* (or a *center* of the Poisson algebra). Evidently, $\{C, G\} = 0$ for every $G(z)$. Hence, by (2), $\mathrm{d}C/\mathrm{d}t = 0$, i.e. $C(z)$ is a constant of motion.

Notice that the constancy of $C(z)$ is independent of the choice of Hamiltonian, a clear contrast to the more usual invariant that is related to a symmetry of a Hamiltonian. In later discussion, however, we will connect a Casimir invariant to an *adiabatic invariant*, and then the constancy of a Casimir invariant will be interpreted as a result of a micro-scale (coarse-grained) symmetry of a Hamiltonian.

Remark 1. *Obviously, if* $\mathrm{Rank}\,\mathcal{J}(z) = n$ *(the dimension of the phase space), (4) has only a trivial solution ($C = \mathrm{constant}$). If the dimension ν of* $\mathrm{Ker}(\mathcal{J}(z))$ *does not change, the solution of (4) may be constructed by "integrating" the elements of* $\mathrm{Ker}(\mathcal{J}(z))$*—then the Casimir leaves are symplectic manifolds. This expectation turns out to be true as far as the Poisson bracket satisfies Jacobi's identity and $n - \nu$ is an even number (Lie–Darboux theorem). However, the point where* $\mathrm{Rank}\,\mathcal{J}(z)$ *changes is the singularity of PDE (4), from which singular Casimir elements are generated [11–13].*

2.2. Energy-Casimir function

When we have a Casimir invariant $C(z)$ in a noncanonical Hamiltonian system, a transformation of the Hamiltonian $H(z)$ such as (with an arbitrary real constant μ)

$$H(z) \mapsto H_\mu(z) = H(z) - \mu C(z) \tag{5}$$

does not change the dynamics. In fact, the equation of motion (2) is invariant under this transformation. We call the transformed Hamiltonian $H_\mu(z)$ an *energy-Casimir* function [1].

Interpreting the parameter μ as a Lagrange multiplier of variational principle, $H_\mu(z)$ is the effective Hamiltonian with the constraint restricting the Casimir element $C(z)$ to a fixed value (since $C(z)$ is a constant of motion, its value is fixed at the initial value). Even when a Hamiltonian is simple, an energy-Casimir functional may have a nontrivial structure. Geometrically, $H_\mu(z)$ is the distribution of $H(z)$ on a Casimir leaf (a surface of $C(z) = \mathrm{constant}$). If Casimir leaves are distorted with respect to the energy norm, the effective Hamiltonian $H_\mu(z)$ may have complex distribution on the leaf, which is, in fact, the origin of various interesting structures in noncanonical Hamiltonian systems.

2.3.   Grand canonical ensemble

The foliated phase space of a noncanonical Hamiltonian system can be viewed as an ensemble of a constrained system—a Casimir invariant, representing the constraint, is often regarded as a "charge" of the system (the conservation of charge is an a priori condition of dynamics, which is independent of the Hamiltonian). To formulate the statistical mechanics for such a system, we consider a *grand canonical ensemble* determined by a total charge M, in addition to the standard determinants, the total particle number N, and the total energy E. The equilibrium is, then, the maximizer of the entropy $S = -\int f \log f \, \mathrm{d}^n z$ under the constraints on the particle number $N = \int f \mathrm{d}^n z$, the energy $E = \int H f \mathrm{d}^n z$, and the charge $M = \int C f \mathrm{d}^n z$; the variational principle

$$\delta(S - \alpha N - \beta E - \gamma M) = 0 \tag{6}$$

yields a Boltzmann distribution

$$f(z) = Z^{-1} \mathrm{e}^{-\beta H - \gamma C}, \tag{7}$$

where $Z \, (= \mathrm{e}^{\alpha+1})$ is the normalization factor, β is the inverse temperature, and γ/β is the *chemical potential* measuring the energy brought about by a change in the charge.

Remark 2. *One may interpret (7) as a Boltzmann distribution with two different energies H and C (with the corresponding inverse temperatures β and γ). Here, we remember the pioneering work of Nambu [14], in which a similar grand canonical distribution function was derived for a "generalized Hamiltonian system" with two Hamiltonians on an* SO(3) *configuration space—the second Hamiltonian corresponds to a Casimir invariant in the present framework.*

2.4.   Diffusion on distorted phase space

How a density f depends on the metric of the phase space (n dimensional) is formulated by identifying it as a differential n-form (or an n-covector). It is essential to distinguish an extensive quantity f and an intensive quantity ϕ; the former (latter) is an n-form (a 0-form); the former transforms as [with a Jacobian weight $\mathrm{D}(y^1, \ldots, y^n)/\mathrm{D}(x^1, \ldots, x^n)$]

$$f(x^1, \ldots, x^n) = f(y^1, \ldots, y^n) \frac{\mathrm{D}(y^1, \ldots, y^n)}{\mathrm{D}(x^1, \ldots, x^n)},$$

while the latter is independent of the coordinate transformation. This is because $f(x^1, \ldots, x^n)\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$ (instead of f alone) represents a physical number, and is exactly at the core of the calculus performed in the foregoing subsections.

   In the theoretical foundation of statistical mechanics, the *invariant measure* based on Liouville's theorem (corresponding to the Poisson bracket of the system) is of fundamental significance. A diffusion equation (or a collision operator), therefore, must be formulated to be consistent with the invariant measure. The equilibrium state is, then, given by maximizing entropy with respect to the invariant measure. Given an invariant measure $\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$, the diffusion equations governing f and ϕ are, respectively,

$$\partial_t f = \mathrm{d}(\mathcal{D}\delta f), \tag{8}$$

$$\partial_t \phi = \delta(\mathcal{D}\mathrm{d}\phi), \tag{9}$$

where $\mathcal{D}$ is a diffusion coefficient, d is the exterior derivative (gradient), and $\delta := (-1)^{n+1} *\mathrm{d}*$ is the codifferential ($*$ is the Hodge star operator). Thus, the diffusion of f is a process of flattening $*f = f * \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$, while that of ϕ is simply the flattening of ϕ. When we observe the

diffusion on some reference frame with coordinates $y^1, \ldots, y^n$, the volume element $dx^1 \wedge \cdots \wedge dx^n$ may be inhomogeneous, and then the diffusion results in creating an inhomogeneous density $f(y^1, \ldots, y^n)$. In the next section, we will see such an example of a *distorted metric* caused by the inhomogeneity in the magnetic field; instead of the conventional Lebesgue measure of the flat Galilean space, the flux-tube volume is invariant, and this is the root cause of the inward diffusion observed in magnetospheres (for example, [15] formulates a diffusion equation for the flux-tube density of a magnetospheric plasma).

The diffusion of f is caused by fluctuations that violate conservation of microscopic data (the initial conditions of each particle) while conserving the macroscopic invariants that serve as the determinants of a statistical ensemble. What is highly nontrivial is that the diffusion (occurring "inward" as demonstrated, for example, in an electron plasma [6,7]) is a process creating an inhomogeneous structure. In a flat (homogeneous metric) space, the equilibrium state is just trivially stable, while the equilibrium associated with a distorted (inhomogeneous) metric remains stable because the *free energy* is constrained by the macroscopic constants. The free energy of a grand canonical system (the logarithm of the grand canonical partition function) is the sum of the internal energy E and the coupled "external" energies; the latter are measured by "particle numbers" multiplied by chemical potentials.

3. Foliation by adiabatic invariants

In the foregoing argument, a Casimir invariant was considered as an abstract constraint on a Hamiltonian system; while we called $M = \int C d^n z$ a total charge of a grand canonical ensemble, the physical meaning of such a charge has not been identified. In this section, we study a concrete example in which a Casimir invariant is equivalent to an adiabatic invariant. The physical meanings of the foliated phase space and the Boltzmann distribution on it then become clear.

3.1. *Hamiltonian of charged particle*

As an example of Hamiltonian system that has a hierarchical structure in terms of adiabatic invariants [16], we study a plasma confined by a magnetic field, and by which we relate Casimir invariants to adiabatic invariants.

3.1.1. Magnetic coordinates.   Here we consider an axisymmetric system with a poloidal (but no toroidal) magnetic field that can be written as

$$\boldsymbol{B} = \nabla\psi \times \nabla\theta, \tag{10}$$

where θ is the toroidal angle and ψ is the magnetic flux function (the Gauss potential of $\boldsymbol{B}$). Let ζ be the parallel coordinate along each magnetic surface (the level-set of ψ). We can choose (ψ, ζ, θ) as the coordinates of the configuration space (θ is ignorable in an axisymmetric system). For example, a point-dipole magnetic field is represented by

$$\begin{cases} \psi(r, z) = Mr^2(r^2 + z^2)^{-3/2}, \\ \zeta(r, z) = Mz(r^2 + z^2)^{-3/2}, \end{cases} \tag{11}$$

where (r, z, θ) are the cylindrical coordinates and M is the magnetic moment.

3.1.2. Hierarchy of adiabatic invariants.   The magnetized particles have three different adiabatic invariants, i.e., the magnetic moment μ, the action $J_\parallel$ of bounce motion, and the action (canonical

angular momentum) P_θ of the toroidal drift [16]. When the magnetic field is sufficiently strong, the corresponding frequencies define a hierarchy: ω_c (cyclotron frequency) $\gg \omega_b$ (bounce frequency) $\gg \omega_d$ (drift frequency). Hence, μ is the most robust adiabatic invariant. On the other hand, the constancy of P_θ is easily broken by large-scale ($\sim$ system size), slow ($\lesssim \omega_d$) perturbations destroying the azimuthal symmetry. In a quasi-neutral plasma ($\phi = 0$), we may estimate $|v_d|/|v_c| \sim \rho_c/L \ll 1$ (v_c is the gyration velocity, v_d is the toroidal drift velocity, ρ_c is the gyro-radius, and L is the macroscopic system size). Neglecting v_d in $P_\theta = mrv_d + q\psi$, we may approximate $P_\theta = q\psi$.

3.1.3. Hamiltonian.   The Hamiltonian of a charged particle is the sum of the kinetic energy and the potential energy:

$$H = \frac{m}{2}v^2 + q\phi, \tag{12}$$

where $v := (P - qA)/m$ is the velocity, P is the canonical momentum, (ϕ, A) is the electromagnetic 4-potential, m is the particle mass, and q is the charge. In the present work, we may treat electrons and ions equally. In a non-neutral plasma, ϕ includes the self-electric field that plays an essential role in determining the equilibrium [6,17].

   In order to take into account the aforementioned hierarchy of actions, we invoke a canonical phase space of action–angle pairs (denoting the gyro angle by ϑ_c and the bounce angle by ϑ_b):

$$z = (\mu, \vartheta_c; J_\parallel, \vartheta_b; \psi, \theta), \tag{13}$$

and write the Hamiltonian of a particle as

$$H_{gc} = \omega_c\mu + \omega_b J_\parallel + q\phi. \tag{14}$$

Here, we have omitted the kinetic energy of the toroidal drift velocity by approximating $P_\theta = q\psi$.[1] The gyro angle is coarse grained (averaged out), so it is eliminated in H_{gc} (i.e., H_{gc} dictates the motion of the guiding center of the gyrating particle). In the standard interpretation, in analogy with the Landau levels in quantum theory, ω_c is the *energy level* and μ is the *number* of quasi-particles (quantized guiding center) at the corresponding energy level;[2] the term $\omega_c\mu$ in H_{gc} represents the macroscopic (classical) energy of the quasi-particles.

3.2.   Foliation by adiabatic invariants

3.2.1.   Foliation by μ.   To extract the macro-hierarchy, we separate the microscopic variables (ϑ_c, μ) by modifying the Poisson matrix as

$$\mathcal{J}_\mu := \begin{pmatrix} 0 & 0 & 0 \\ 0 & J_c & 0 \\ 0 & 0 & J_c \end{pmatrix}. \tag{15}$$

---

[1] There is a subtlety in the representation of the energy of the drift velocity if it were retained in the Hamiltonian formalism in order to analyze the dynamics (in response to the electromagnetic perturbations) of the guiding center [18]. While we omit the energy of the drift velocity, we still have a drift velocity; the drift frequency (including all grad-B, curvature, and E×B drifts) is given by bounce-averaging the toroidal angular velocity, i.e., $\omega_d = d\theta/dt = \partial H_{gc}/\partial\psi$. In a homogeneous magnetic field, both ω_c and ω_b are constant, and then ω_d evaluates the E×B drift frequency.

[2] To be correct about the dimension of μ, $\mu/\hbar$ is the particle number, and $\hbar\omega_c$ is the energy level, where $\hbar$ is a certain constant having the dimension of action (which may not be Planck's constant in the classical regime).

The Poisson bracket

$$\{F, G\}_\mu := \langle \partial_z F, \mathcal{J}_\mu \partial_z G \rangle$$

determines the kinematics on the macro-hierarchy that separates the canonical pair (μ, ϑ_c).

The nullity of $\mathcal{J}_\mu$ makes the Poisson bracket $\{ , \}_\mu$ *noncanonical*. Evidently, μ is a Casimir invariant (more generally, every $C = g(\mu)$ with g being any smooth function is a Casimir invariant). The level-set of μ, a leaf of the Casimir foliation, identifies what we may call the *macro-hierarchy*.

Remark 3. *The adiabatic invariant μ appears in several manifestations; it has been called "Casimir invariant," "charge," and "particle number." Although these names are used synonymously, their specific conservations carry different implications. As noted above, the Landau level analogy allows us to assign an adiabatic invariant with a particle number, and this interpretation plays an essential role in formulating the "grand canonical ensemble" (Sect. 2.3) in which μ plays the role of a quasi-particle number (Sect. 3.3). The equivalent christening of μ as "charge" or a Casimir invariant is rather profoundly motivated. In a noncanonical Hamiltonian system, the nullity of the Poisson bracket yields a topological charge, i.e., a Casimir invariant (Sect. 2.1). Such a topological charge can be related to a Noether charge pertinent to a symmetry of some appropriate action principle. When a Casimir invariant is an adiabatic invariant (as constructed here), it is a consequence of symmetry with respect to a coarse-grained angle ϑ. The following mathematical formality provides content to the preceding statement. Let $S = \int \Lambda - H\mathrm{d}t$ be a microscopic (canonical) action, where Λ is a canonical 1-form and H is a Hamiltonian. Suppose that ϑ is a microscopic angle of some periodic motion, which gives an adiabatic invariant $J = \oint P_\vartheta \mathrm{d}\vartheta/2\pi$. Then a macroscopic action S' can be defined by separating, from S, a microscopic action $S_J = \int (J\dot{\vartheta} - \omega J)\mathrm{d}t$ (ωJ is the energy of the microscopic periodic motion with frequency ω). Averaging over the periodic motion, S' is made to be independent of ϑ. The symmetry $\partial_\vartheta = 0$ of the coarse-grained action $\overline{S} = S' + S_J$ yields a Noether charge $\partial_{\dot{\vartheta}} \overline{S} = J$, i.e., the Casimir invariant (the variation of the macroscopic part $S' = \int \Lambda' - H'\mathrm{d}t$ yields a degenerate 2-form $\mathrm{d}\Lambda'$ whose nullity is spanned by the Casimir invariant).*

3.2.2. Foliation by $J_\parallel$.

We may define a more macroscopic hierarchy by separating the second canonical pair $(\vartheta_b, J_\parallel)$ from the phase space. In comparison with the previous process of defining $\{ , \}_\mu$ we need a somewhat complicated procedure, because the bounce angle (ϑ_b) is not ignored in H_{gc}; the frequencies ω_c and ω_b (as well as ϕ in a non-neutral plasma) are functions of the spatial coordinates including ϑ_b ($\zeta = \ell_\parallel \sin \vartheta_b$ with the bounce orbit length $\ell_\parallel$). We have

$$\dot{J}_\parallel = \frac{\partial H_{gc}}{\partial \vartheta_b}. \tag{16}$$

For the periodic bounce motion, $\oint (\partial H_{gc}/\partial \vartheta_b) d\vartheta_b = \oint dH_{gc} = 0$. Integrating (16) over the cycle of bounce motion yields the bounce-average $\langle J_\parallel \rangle =$ constant. When we calculate macroscopic quantities (like the total energy or the total action), we evaluate $J_\parallel$ as the adiabatic invariant $\langle J_\parallel \rangle$, and then the second action–angle pair $(\vartheta_b, J_\parallel)$ is separated from the dynamical variables; the corresponding Poisson matrix is

$$\mathcal{J}_{\mu J_\parallel} := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & J_c \end{pmatrix}, \tag{17}$$

and the Poisson bracket is

$$\{F, G\}_{\mu J_\parallel} := \langle \partial_z F, \mathcal{J}_{\mu J_\parallel} \partial_z G \rangle.$$

Now, the dynamical variables are only θ and ψ. The drift frequency is given by bounce-averaging the toroidal angular velocity:

$$\omega_{\mathrm{d}} = \dot{\theta} = \frac{\partial H_{\mathrm{gc}}}{\partial \psi} = \mu \frac{\partial \omega_{\mathrm{c}}}{\partial \psi} + J_{\parallel} \frac{\partial \omega_{\mathrm{b}}}{\partial \psi} + q \frac{\partial \phi}{\partial \psi}. \tag{18}$$

As long as the system maintains the toroidal symmetry $\partial/\partial\theta = 0$, the third action ψ remains constant, and the orbit of the guiding center is completely integrable. A slow perturbation, however, may break the constancy of ψ, allowing the guiding center to cross magnetic surfaces. As shown in the next section, the Boltzmann distribution on this minimum (most macroscopic) phase space has an interesting structure.

3.3. Boltzmann distributions

3.3.1. Microscopic phase space. The standard Boltzmann distribution function is derived when we assume that $\mathrm{d}^6 z = \mathrm{d}^3 v \mathrm{d}^3 x$ is an invariant measure and the Hamiltonian H is the determinant of the ensemble. Maximizing the entropy S while keeping the total energy E and the total particle number N constant, we obtain

$$f(\boldsymbol{x}, \boldsymbol{v}) = Z^{-1} \mathrm{e}^{-\beta H}. \tag{19}$$

The corresponding configuration-space density is

$$\rho(\boldsymbol{x}) = \int f \mathrm{d}^3 v \propto \mathrm{e}^{-\beta q \phi}, \tag{20}$$

which becomes constant for an electrically neutral system ($\phi = 0$).

Needless to say, the Boltzmann distribution or the corresponding configuration-space density, with an appropriate Jacobian multiplication, is independent of the choice of phase-space coordinates. Moreover, the density is invariant no matter whether we coarse-grain the cyclotron motion or not. Let us confirm this fact by a direct calculation. The Boltzmann distribution of the *guiding-center plasma* is

$$\begin{aligned} f(\boldsymbol{z}) &= Z^{-1} \mathrm{e}^{-\beta H_{\mathrm{gc}}} \\ &= Z^{-1} \mathrm{e}^{-\beta \left[m(v_{\perp}^2 + v_{\parallel}^2)/2 + q\phi \right]}, \end{aligned} \tag{21}$$

where $\boldsymbol{v}_{\perp}$ and $\boldsymbol{v}_{\parallel}$ are the perpendicular and parallel (with respect to the local magnetic field) components of the velocity. Here we neglected the kinetic energy of the drift motion to approximate $\omega_{\mathrm{c}} \mu \approx m v_{\perp}^2 / 2$. The corresponding density reproduces (20).

3.3.2. Boltzmann distribution on the μ leaf. Now we calculate the Boltzmann distribution on the *macro-hierarchy*. We start with the Casimir leaf of μ. The adiabatic invariance of μ imposes a topological constraint on the motion of particles; this constraint is the root cause of a macro-hierarchy and of structure formation.

By applying Liouville's theorem to the Poisson bracket $\{\ ,\ \}_\mu$, the invariant measure on the macro-hierarchy (the Γ-space of quasi-particles) is

$$\prod_j \mathrm{d}^4 z_j = \prod_j \mathrm{d}^6 z_j / (2\pi \mathrm{d}\mu_j),$$

i.e., the total phase-space measure modulo the microscopic measure (suffix j is the index of each particle). The most probable state (statistical equilibrium) on the macroscopic ensemble must maximize the entropy with respect to this invariant measure.
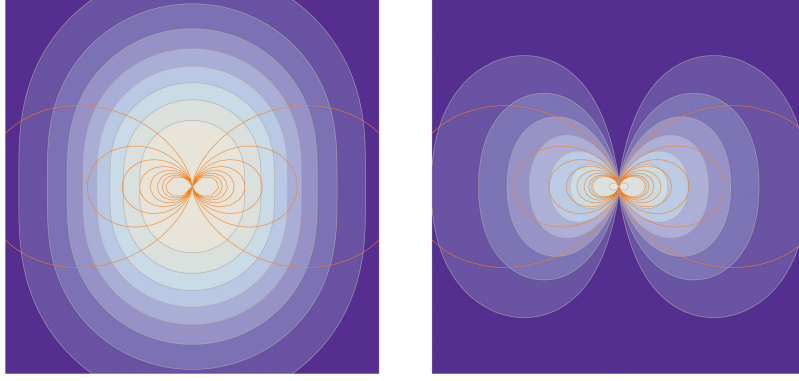
Fig. 1. Density distribution (contours) and the magnetic field lines (level-sets of ψ) in the neighborhood of a point dipole. Left: The equilibrium on the leaf of μ-foliation. Right: The equilibrium on the leaf of μ- and $J_\parallel$-foliation.

To determine the distribution function, the variational principle is set up immersing the ensemble into the general phase space, and incorporating the constraints through the Lagrange multipliers: We maximize entropy $S = -\int f \log f \, \mathrm{d}^6 z$ for a given particle number $N = \int f \mathrm{d}^6 z$, a quasi-particle number $M_1 = \int \mu f \mathrm{d}^6 z$, and an energy $E = \int H_{\mathrm{gc}} f \mathrm{d}^6 z$, to obtain the distribution function (see Sect. 2.3)

$$f = f_\gamma := Z^{-1} \mathrm{e}^{-(\beta H_{\mathrm{gc}} + \gamma \mu)}. \tag{22}$$

The factor $\mathrm{e}^{-\gamma \mu}$ in f_γ yields a direct ω_{c} dependence of the coordinate-space density:

$$\rho = \int f_\gamma \frac{2\pi \omega_{\mathrm{c}}}{m} \mathrm{d}\mu \mathrm{d}v_\parallel \propto \frac{\omega_{\mathrm{c}}(\boldsymbol{x})}{\beta \omega_{\mathrm{c}}(\boldsymbol{x}) + \gamma}. \tag{23}$$

Here, we are assuming electric neutrality to put $\phi = 0$. Notice that the Jacobian $(2\pi \omega_{\mathrm{c}}/m)\mathrm{d}\mu$ multiplying the macroscopic measure $\mathrm{d}^4 z$ reflects the distortion of the macroscopic phase space (Casimir leaf) caused by the magnetic field. Figure 1 (left) shows the density distribution and the magnetic field lines.

3.3.3.   Boltzmann distribution on the μ-$J_\parallel$ leaf.   We may further restrict the second action $J_\parallel$ and calculate the Boltzmann distribution of the μ-$J_\parallel$ leaf. Imposing another constraint on $M_2 = \int J_\parallel f \mathrm{d}^6 z$, we modify (22) as

$$f_{\gamma_1, \gamma_2} = Z^{-1} \mathrm{e}^{-(\beta H_{\mathrm{gc}} + \gamma_1 \mu + \gamma_2 J_\parallel)}. \tag{24}$$

To find explicit expressions for the parallel action–angle variables, let us solve the equation of parallel motion under some approximations. Neglecting the curvature of magnetic field lines and putting $\phi = 0$,

$$m\ddot{\zeta} = -\mu \nabla_\parallel \omega_{\mathrm{c}}, \tag{25}$$

where $\nabla_\parallel := \boldsymbol{b} \cdot \nabla$ with the magnetic unit vector $\boldsymbol{b} := \boldsymbol{B}/B$. In the vicinity of $\zeta = 0$, where ω_{c} has a minimum on each magnetic surface, we may approximate

$$\omega_{\mathrm{c}} = \Omega_{\mathrm{c}}(\psi) + \Omega_{\mathrm{c}}''(\psi) \frac{\zeta^2}{2},$$

where $\Omega_c(\psi)$ is the minimum of ω_c on each contour of ψ, and $\Omega_c''(\psi) := d^2\omega_c/d\zeta^2|_\psi$. Integrating (25), we obtain a harmonic oscillation with the bounce frequency

$$\omega_b = \sqrt{\frac{\Omega_c''(\psi)\mu}{m}} = \frac{v_\perp}{L_\parallel(\psi)}, \tag{26}$$

where $L_\parallel(\psi) := \sqrt{2\Omega_c(\psi)/\Omega_c''(\psi)}$ is the length scale of the variation of ω_c along ζ. The amplitude of the oscillation, i.e., the bounce orbit length, is

$$\ell_\parallel = \sqrt{\frac{2E_\parallel}{m\omega_b^2}}, \tag{27}$$

where $E_\parallel := (mv_\parallel^2)/2|_{\zeta=0}$ is the kinetic energy of the parallel motion. Assuming $E_\parallel \approx E_\perp := \mu\Omega_c$, we may estimate $\ell_\parallel \approx L_\parallel$. By $E_\parallel = \omega_b J_\parallel$, we obtain

$$dv_\parallel = \frac{\omega_b}{mv_\parallel}dJ_\parallel = \sqrt{\frac{\omega_b}{2mJ_\parallel}}dJ_\parallel.$$

Using the relation $\omega_b/(mv_\parallel) = v_\perp/(L_\parallel mv_\parallel) \approx 1/(mL_\parallel)$, we may write

$$dv_\parallel \approx \frac{dJ_\parallel}{mL_\parallel}. \tag{28}$$

The density is given by

$$\rho = \int f_{\gamma_1,\gamma_2} \frac{2\pi\omega_c d\mu}{m} \frac{dJ_\parallel}{mL_\parallel},$$

$$\propto \frac{\omega_c(\boldsymbol{x})}{m^2} \int_0^\infty \frac{e^{-(\beta\omega_c+\gamma_1)\mu}d\mu}{\beta\sqrt{2\omega_c\mu/m} + \gamma L_\parallel(\psi)}. \tag{29}$$

Numerical integration of (29) gives the density profile depicted in Fig. 1 (right) [19].

Remark 4. *The derived distribution function f_{γ_1,γ_2} is a particular solution of the stationary kinetic equation $\{H_{gc}, f\}_{\mu J_\parallel} = 0$. This thermodynamic equilibrium, however, has a deeper meaning than the arbitrary solutions such as $f = F(\mu, J_\parallel, \psi)$ that are often invoked in drift-kinetic calculations. For instance, $f = F(\mu, J_\parallel)$ yields a density $\rho \propto \omega_c/L_\parallel$ (implying that the particle number per unit flux tube distributes homogeneously), which in a dipole magnetic field scales as $\propto r^{-4}$, precisely the density profile given by Hasegawa [20] (for example, see [21] for a realization of a similar distribution in Saturn's magnetosphere). Choosing F to be a Gaussian, $Z^{-1}e^{-(\gamma_1\mu+\gamma_2 J_\parallel)}$ is the asymptotic form of (29) in the limit $r \to \infty$ ($\omega_c \to 0$ so that $\beta H_{gc} \ll \gamma_1\mu + \gamma_2 J_\parallel$). Such a solution is also the $\beta \to 0$ (infinite temperature) limit of f_{γ_1,γ_2}. For finite temperatures, the energy constraint prevents the particle distributing homogeneously on the ensemble foliated by μ and $J_\parallel$. Notice that, for the distribution f_{γ_1,γ_2}, the density ρ remains finite, while for the solution $f = F(\mu, J_\parallel)$, it diverges as $\omega_c \to \infty$. For experimental evidence of density limitation, see [8].*

Remark 5. *In the foregoing analysis, we did not pay attention to the field equation (Maxwell's equation), and dealt with the magnetic flux function ψ as a given function of space. However, when the plasma pressure becomes comparable to the pressure of the dipole magnetic field (i.e., the so-called beta ratio is of order unity; see [19]), we have to adjust the magnetic field to take into account the spontaneous component. This can be done by solving the Grad–Shafranov equation for ψ with the plasma pressure given by the distribution function (to take into account the pressure anisotropy,*

we have to use the generalized Grad–Shafranov equation [22,23]). We also assumed charge neutrality, and put the electric potential $\phi = 0$. In a non-neutral plasma [6], the spontaneous electric potential ϕ must be determined by solving the Poisson equation (the grand canonical ensemble of a non-neutral plasma is also constrained by the third adiabatic invariant $P_\theta \approx q\psi$, because the total angular momentum plays an important role in "neutralizing" the spontaneous electric field on the comoving frame; see also [24]). The field equations (the Grad–Shafranov equation for ψ, and the Poisson equation for ϕ) pose a nonlinear problem, because we have to find a self-consistent distribution function that depends on ψ and ϕ. Numerical analyses of these equations will be presented elsewhere.

4. Summary and concluding remarks

In this paper, we have developed a conceptual framework for delineating and understanding the advanced notion of self-organization simultaneous with entropy production. An appropriate *scale hierarchy*, encompassing large-scale order and small-scale disorder, is established by exploiting phase-space foliation provided by the adiabatic invariants of the system; the corresponding invariant measure is also specified. A leaf of the foliated phase space is identified as a grand canonical ensemble of macroscopic quasi-particles representing coarse-grained (averaged over microscopic angle of periodic motion) orbits.

As an explicit example, we have constructed a foliated phase space representing the scale hierarchy of magnetized particles in a magnetospheric plasma. The Boltzmann distribution is obtained by maximizing the entropy for a given particle number and a quasi-particle number as well as a total energy. The system is driven to such a Boltzmann distribution by some entropy production mechanism that, inherently, preserves the adiabatic invariants (Casimir invariants). The spatiotemporal scales of associated fluctuations must be larger than the scales on which the conjugate coarse-grained angle variables vary. Under the same condition on possible perturbations, the Boltzmann distribution on the macro-hierarchy is absolutely stable, because it is the minimizer of the energy (as an isolated system). It is interesting that the steep density gradient predicted by the distribution (29) is stable against macroscopic modes such as interchange modes (cf. [15]); an intuitive explanation is that the magnetized particles reside in the magnetic-coordinate space, where the actual density distribution is flat, leaving no free energy for instabilities.

The derived grand canonical distribution function opens a new channel for extracting interesting properties of magnetized plasmas. For example, μ could be boosted by cyclotron heating (for magnetically confined particles), and the resulting increase in the total magnetic moment of the system could cause macroscopic motion of the levitated magnet accompanied by the plasma [25]. In our model, an increase of μ means injection of quasi-particles; an increased quasi-particle number M, in turn, automatically increases the macroscopic magnetic moment. This simple picture is beyond the reach of the conventional canonical (or micro-canonical) distributions that are unaware of any direct relation between the macroscopic magnetic moment and μ [even if we write $mv_c^2/2 = \mu\omega_c$ as in (21)]. Heating, therefore, could not create or destroy magnetic moment (or any axial vector); the coupling of heat and mechanical energy could manifest only through the pressure force. Our distribution function, on the other hand, is capable of delineating such connections since it embodies magneto–fluid–thermo couplings. Many other applications such as estimate of fluctuations, phase equilibrium relations, possible condensation at low temperature, etc. will become accessible through the grand canonical distribution.

The framework we have developed will apply to general systems with nontrivial *topologies*. Viewing from a different angle, our work has a common perspective with Nambu's "generalized Hamiltonian system" that has two Hamiltonians (one of which is a Casimir invariant in the present terminology); see [14] and Remark 2. In the present theory, connecting a Casimir invariant further to an adiabatic invariant, we have written a kinetic–thermodynamic theory with a built-in *scale hierarchy*. Structure formation is a direct consequence of embedding the Casimir leaf (where the microscopic actions are abstracted as quasi-particle numbers) into the laboratory flat space.

We end this paper with some comments on self-organization in fluid-mechanical systems. The dynamics in some fluid models can be cast into a unified Hamiltonian form [1]

$$\partial_t \omega = \mathcal{J} \partial_\omega H(\omega), \tag{30}$$

where ω is a state vector belonging to a Hilbert space X, $H(\omega)$ is the Hamiltonian which is a real-valued functional on X, and $\mathcal{J}$ is the Poisson operator. The Poisson bracket is defined by $\langle \partial_\omega F, \mathcal{J} \partial_\omega G \rangle$, where $\langle \, , \, \rangle$ is the inner-product of X. The vortex equation of two-dimensional Eulerian (inviscid, incompressible) flow is the simplest example; with a Poisson operator $\mathcal{J} = [\omega, \circ]$ (where ω is the vorticity, and $[a, b] = \partial_y a \cdot \partial_x b - \partial_x a \cdot \partial_y b$), and a Hamiltonian $H(\omega) = \int \omega \cdot (-\Delta)^{-1} \omega \, d^2 x / 2$ (where Δ is the two-dimensional Laplacian, and Δ^{-1} is its inverse), (30) reads $\partial_t \omega = [\omega, \phi]$ (where $\phi = (-\Delta)^{-1} \omega$ is the Gauss potential of the flow). Notice that $\mathcal{J}$ depends on the dynamical variable ω. Evidently, $C_f = \int f(\omega) \, d^2 x$ (f is an arbitrary C^2-class function) is a Casimir invariant. Slightly modifying the Poisson operator as $\mathcal{J} = [\omega - g, \circ]$ with an inhomogeneous term g (cf. [26]), and the Hamiltonian as $H = \int \omega \cdot L^{-1} \omega \, d^2 x / 2$ with $L = -\Delta + 1$ (the term $+1$ in the operator L reflects the compressibility of the fluid), Hamilton's equation (30) becomes $\partial_t \omega = [\omega - g, \phi]$ with $\phi = L^{-1} \omega$, which is formally the Hasegawa–Mima equation [27] of drift waves in a magnetized plasma (g represents the inhomogeneity of the equilibrium plasma density), or the Charney equation [28] of Rossby waves (g represents the inhomogeneity of the Coriolis force and the depth of the atmospheric fluid). The Casimir invariant is $C_f = \int f(\omega - g) \, d^2 x$.

A theory of self-organization can be described by invoking the scenario of *selective dissipation* which compares different constants of motion, the Hamiltonian (energy), and some Casimir invariants (choosing $f(\omega) = \omega^2 / 2$, C_f is the *enstrophy*); a functional including higher-order spatial derivatives is more fragile in comparison with a lower-order one, because small-scale turbulence can dissipate it more easily, thus the energy conserves better than the enstrophy. Minimizing the enstrophy for a fixed energy, we obtain a "relaxed state." The review paper by Hasegawa [29] describes a list of successful applications of this model, including the creation of zonal flows by Rossby wave turbulence, and the Taylor relaxed state [30] by magnetohydrodynamic (MHD) turbulence (which dissipates the energy while conserving the magnetic helicity, a Casimir invariant of an MHD system).

The same minimization principle has a different connotation, which aligns with the present theory. The target functional to be minimized is nothing but the energy-Casimir functional (see Sect. 2.2), thus the minimizer is an equilibrium point of the Hamiltonian mechanics (by the duality of well-posed variational principles [31], the minimizer of the enstrophy for a given energy is equivalent to the maximizer of the energy for a given enstrophy, and the latter is also an equilibrium point of the Hamiltonian system). To put this equilibrium point into the perspective of the entropy principle, we have to consider statistical mechanics on a function space. In [32], a grand canonical ensemble of MHD is formulated by considering a "magnetic-helicity quantum." The Taylor relaxed state is, then, the *low-temperature limit* (ground state) of the Boltzmann distribution (here, the *temperature* measures the strength of the turbulence).

While the present theory of self-organization has wide applications encompassing particle models to fluid models, the connection between the two formalisms awaits further exploration. For example, the helicities are the determinants of the macroscopic hierarchy (Casimir leaf) of the MHD system, which control the bifurcation of various equilibrium points by shifting the leaf in the phase space [11]. However, we have not yet unearthed the coarse-grained "angle variable" corresponding to the magnetic helicity. If we can specify the origin of the fluid Casimir invariants in the particle model, we will be able to write a kinetic theory of far richer structures that have various helicities (or vorticities and currents twisting the stream lines and magnetic field lines).

The foliation is also an interesting subject to be explored in the framework of a new variable—the entropy production rate. Since, for a driven (or open) system, the standard entropy is no longer an effective state variable to characterize long-lived structures, the entropy production rate (σ) has been deemed to be an alternative determinant of the state. Debates have raged whether the organizing principle is the minimum or maximum of σ. While the minimum σ principle[3] applies to linear systems, typical nonlinear fluid-mechanical systems instead prefer maximum σ states. Sawada [38] proposed a nonlinear mechanism that maximizes σ; for other models of fluids and plasmas, see [39–41]. Dewar [42] proposed to evaluate the probability of phase-space trajectories (instead of points), which is shown to have a Boltzmann distribution by replacing the minus energy by the entropy production along each trajectory; hence the maximum entropy production is most probable (see also [43,44] for the thermodynamic duality of the maximum and minimum principles that switch depending on whether the system is flux driven or force driven). When the phase space is foliated, such trajectories must be restricted on a Casimir leaf, and then the structures self-organized by maximizing σ will have topological complexity charged by the Casimir invariants.

Acknowledgements

References

[1] P. J. Morrison, Rev. Mod. Phys. **70**, 467 (1998).
[2] M. Schulz and L. J. Lanzerotti, *Particle Diffusion in the Radiation Belts* (Springer, New York, 1974).
[3] D. H. Brautigam and J. M. Albert, J. Geophys. Res. **105**, 291 (2000).
[4] Y. Chen, G. D. Reeves, and R. H. W. Friedel, Nature Phys. **3**, 614 (2007).
[5] A. C. Boxer, R. Bergmann, J. L. Ellsworth, D. T. Garnier, J. Kesner, M. E. Mauel, and P. Woskov, Nature Phys. **6**, 207 (2010).
[6] Z. Yoshida, H. Saitoh, J. Morikawa, Y. Yano, S. Watanabe, and Y. Ogawa, Phys. Rev. Lett. **104**, 235004 (2010).

---

[3] One often assumes σ to be equivalent to energy dissipation rate. The history of the latter concept goes back to Helmholtz's minimum dissipation principle, by which the velocity distribution of a slow stationary incompressible fluid can be determined [33]. Generalizing the notion of force and current (or flux), we may describe thermodynamic non-equilibrium in parallel with mechanical non-equilibrium, and then the mechanical minimum-dissipation principle (Helmholtz's principle) extends to the thermodynamic minimum-dissipation principle (Onsager's principle) [34,35]. Prigogine's minimum σ principle has an independent origin: first it was applied to linear irreversible processes in discontinuous systems; it was then extended to some nonlinear and continuous systems [36]. For some physical systems, σ becomes equivalent to the energy dissipation function [37].

[7] H. Saitoh, Z. Yoshida, J. Morikawa, Y. Yano, H. Hayashi, T. Mizushima, Y. Kawai, M. Kobayashi, and
     H. Mikami, Phys. Plasmas **17**, 112111 (2010).
[8] H. Saitoh, Z. Yoshida, J. Morikawa, Y. Yano, T. Mizushima, Y. Ogawa, M. Furukawa, Y. Kawai,
     K. Harima, Y. Kawazura, Y. Kaneko, K. Tadachi, S. Emoto, M. Kobayashi, T. Sugiura, and G. Vogel,
     Nucl. Fusion **51**, 063034 (2011).
[9] T.J. Birmingham, T.G. Northrop, and C.-G. Falthammar, Phys. Fluids **10**, 2389 (1967).
[10] M. Walt, Space Sci. Rev. **12**, 446 (1971).
[11] Z. Yoshida and R. L. Dewar, J. Phys. A: Math. Theor. **45**, 365502 (2012).
[12] Z. Yoshida, P. J. Morrison, and F. Dobarro, J. Math. Fluid Mech. **16**, 41 (2014).
[13] Z. Yoshida, Procedia IUTAM **7**, 141 (2013).
[14] Y. Nambu, Phys. Rev. D **7**, 2405 (1973).
[15] G. L. Siscoe and D. Summers, J. Geophys. Res. Space Phys. **86**, 8471 (1981).
[16] A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics*, (Springer-Verlag, New York,
     1992), 2nd ed., Sect. 2.3b.
[17] D. H. E. Dubin and T. M. O'Neil, Rev. Mod. Phys. **71**, 87 (1999).
[18] J. R. Cary and A. J. Brizard, Rev. Mod. Phys. **81**, 693 (2009).
[19] Z. Yoshida, H. Saitoh, Y. Yano, H. Mikami, N. Kasaoka, W. Sakamoto, J. Morikawa, M. Furukawa, and
     S. M. Mahajan, Plasma Phys. Control. Fusion **55**, 014018 (2013).
[20] A. Hasegawa, Phys. Scr. **T116**, 72 (2005).
[21] P. Schippers, M. Moncuquet, N. Meyer-Vernet, and A. Lecacheux, J. Geophys. Res. Space Phys. **118**,
     7170 (2013).
[22] H. Grad, Phys. Fluids **10**, 137 (1967).
[23] M. Furukawa, Phys. Plasmas **21**, 012511 (2014).
[24] Yu. I. Gubar and A. R. Moszuhina, Kosmicheskie issledovanija (Cosmic Research) **27**, 135 (1989).
[25] Z. Yoshida, Y. Yano, J. Morikawa, and H. Saitoh, Phys. Plasmas **19**, 072303 (2012).
[26] A. Weinstein, Phys. Fluids **26**, 388 (1983).
[27] A. Hasegawa and K. Mima, Phys. Rev. Lett. **39**, 205 (1977).
[28] J. C. Charney, J. Atmos. Sci. **28**, 1087 (1971).
[29] A. Hasegawa, Advances Phys. **34**, 1 (1985).
[30] J. B. Taylor, Phys. Rev. Lett. **33**, 1139 (1974).
[31] Z. Yoshida and S. M. Mahajan, Phys. Rev. Lett. **88**, 095001 (2002).
[32] N. Ito and Z. Yoshida, Phys. Rev. E **53**, 5200 (1996).
[33] H. Helmholtz, Wiss. Abh **1**, 223 (1868).
[34] L. Onsager, Phys. Rev. **37**, 405 (1931).
[35] L. Onsager, Phys. Rev. **38**, 2265 (1931).
[36] G. Nicolis and I. Prigogine, *Self-organization in Nonequilibrium Systems: From Dissipative Structures
     to Order through Fluctuations* (Wiley, New York, 1977).
[37] I. Gyarmati, *Nonequilibrium Thermodynamics* (North-Holland, Amsterdam, 1962).
[38] Y. Sawada, Prog. Theor. Phys. **66**, 68 (1981).
[39] H. Ozawa, S. Shimokawa, and H. Sakuma, Phys. Rev. E **64**, 026303 (2001).
[40] Z. Yoshida and S. M. Mahajan, Phys. Plasmas **15**, 032307 (2008).
[41] R. K. Niven, Phys. Rev. E **80**, 021113 (2009).
[42] R. Dewar, J. Phys. A **36**, 631 (2003).
[43] Y. Kawazura and Z. Yoshida, Phys. Rev. E **82**, 066403 (2010).
[44] Z. Yoshida and Y. Kawazura, in *Beyond The Second Law: Entropy Production and Non-Equilibrium
     Systems*, (Springer, New York, 2014), Chap. 15.