Commun Nonlinear Sci Numer Simulat 161 (2026) 110076

Contents lists available at ScienceDirect

Communications in Nonlinear Science and Numerical Simulation

journal homepage: www.elsevier.com/locate/cnsns

Invited Article

Metriplectic relaxation to equilibria

C. Bressan [1], M. Kraus [a], O. Maj [a], P.J. Morrison [b,*]

[a] *Max Planck Institute for Plasma Physics, Boltzmannstrasse 2, 85747, Garching, Germany*
[b] *Department of Physics, Institute for Fusion Studies, University of Texas at Austin, 78712-1060, Austin, TX, USA*

ARTICLE INFO

Keywords:
Dissipative dynamical systems
Metriplectic systems
Hamiltonian systems
Lyapunov theorem
Landau collision operator
Magnetohydrodynamics
Fluid dynamics

ABSTRACT

Metriplectic dynamical systems consist of a special combination of a Hamiltonian and a (generalized) entropy-gradient flow, such that the Hamiltonian is conserved and entropy is dissipated/produced (depending on a sign convention). It is natural to expect that, in the long-time limit, the orbit of a metriplectic system should converge to an extremum of entropy restricted to a constant-Hamiltonian surface. In this paper, we discuss sufficient conditions for this to occur. Then, we construct a class of metriplectic systems inspired by the Landau operator for Coulomb collisions in plasmas, which is included as special case. For this class of brackets, checking the conditions for convergence reduces to checking two usually simpler conditions, and we discuss examples in detail. We apply these results to the construction of relaxation methods for the solution of equilibrium problems in fluid dynamics and plasma physics.

1. Introduction

There are two main purposes of this paper: investigate sufficient conditions for metriplectic relaxation (reviewed in Section 2.1) to occur and use metriplectic relaxation to find equilibria of fluid dynamics and plasma physics systems. Metriplectic dynamical systems, as introduced in [1–3], are designed to formally converge to an extremum of entropy while being restricted to a constant Hamiltonian surface. Here we more rigorously examine the conditions for such relaxation and then construct and investigate metriplectic systems that achieve the relaxation for finding equilibria of a collection of fluid and plasma physical systems. In the remainder of this section, in Section 1.1 we first give an overview of the challenges and previous methods for calculating equilibria, followed in Section 1.2 by our and others' previous metriplectic relaxation methods.

1.1. Overview of equilibrium calculations

The calculation of equilibria of physical systems often leads to ill-posed nonlinear problems, where the ill-posedness is due to the nonuniqueness of the solution. Additional constraints are needed to define uniquely the equilibrium of interest, depending on the application at hand. In some situations, prescribing enough constraints to determine a unique equilibrium may not be straightforward. This lack of uniqueness for equilibrium problems is precisely discussed below in Section 2.2 for examples taken from fluid dynamics and magnetohydrodynamics (MHD): equilibria of the Euler equations in vorticity form reduced to two dimensions [4, p.488], axisymmetric MHD equilibria [5], linear and nonlinear Beltrami fields.

* Corresponding author.
E-mail addresses: michael.kraus@ipp.mpg.de (M. Kraus), omar.maj@ipp.mpg.de (O. Maj), morrison@physics.utexas.edu (P.J. Morrison).
[1] Formerly at the Max Planck Institute for Plasma Physics, Boltzmannstrasse 2, 85747, Garching, Germany

https://doi.org/10.1016/j.cnsns.2026.110076
Received 17 July 2025; Received in revised form 4 April 2026; Accepted 23 April 2026
Available online 26 April 2026
1007-5704/

In some cases, after providing additional physical constraints, the equilibrium conditions can be reformulated as a well-posed mathematical problem. This is the case, for instance, for the Euler equations, axisymmetric MHD equilibria, and linear Beltrami fields.

In more complicated situations, such as for nonlinear Beltrami fields and full MHD equations in three dimensions, the problem of computing an equilibrium point has no good solution. The difficulties were shown in [6] to be related to the Kolmogorov-Arnold-Moser (KAM) theorem (see for example [7]). A mathematical perspective on these difficulties can be found in the introduction of the paper by Bruno and Laurence [8], cf. also the recent developments by Enciso et al. [9]. A large fraction of MHD equilibrium calculations in three dimensions are based on a reformulation of the problem in which one assumes that the magnetic field is tangent to a family of nested toroidal surfaces [10,11]. On the one hand, such a configuration is a natural generalization to three-dimensions of the confined region in axisymmetric MHD equilibria. In addition, by basic considerations of topology, the confinement of a plasma in a volume bounded by a closed orientable pressure isosurface, where the pressure gradient is balanced by the electromagnetic force, requires the surface to be a torus in the simplest case [12]. Therefore searching for equilibria with nested toroidal flux surfaces appears to be the simplest and most appealing approach to the three-dimensional equilibrium problem. On the other hand, Grad conjectured that such equilibria may not exist [13] unless they are axisymmetric or we allow for weak solutions characterized by singular current sheets localized on specific flux surfaces, namely the resonant surfaces. The non-existence of smooth non-axisymmetric equilibria with nested flux surfaces is referred to as the *Grad's conjecture*, and (to the best of our knowledge) it is still an open question. Nevertheless, the variational formulation adopted for equilibria with nested flux surfaces [10,11], in principle at least, allows for weak solutions [14,15], although usually this possibility is not exploited in state-of-the-art codes such as VMEC [11], DESC [16], and GVEC [17], which use a highly regular representation of the magnetic and pressure fields. (For sake of completeness, we note that the GVEC code has the built-in possibility of relaxing the regularity of the magnetic field allowing for current sheets on prescribed surfaces, but this possibility has not been exploited yet.) The singular current layers that are expected according to the Grad's conjecture cannot be considered physical; these equilibria are regarded as computationally efficient proxies for equilibria that may have a complicated field-line topology in a neighborhood of some resonant surfaces, but have nested toroidal surfaces elsewhere for good confinement. This strategy has been extremely successful for the design of stellarators [18], since it reduces significantly the complexity of the problem. While this approach gives an acceptable representation of the magnetic field in the core of a stellarator with modest computational cost, it cannot account for more complicated magnetic field configurations, such as those with magnetic islands and chaotic field lines, due to the built-in foliation of the domain by toroidal surfaces. Yet magnetic-field islands and chaotic regions are relevant in practice and calculations of MHD equilibria beyond the paradigm of nested toroidal surfaces are needed. An iterative procedure for the calculation of general MHD equilibria has been outlined by Grad [19], and a similar iterative procedure is implemented in the PIES code [20]. Such iteration schemes are purely heuristics: there is no theoretical control on the convergence.

Another approach to the computation of equilibria is based on artificial relaxation. Relaxation methods solve the Cauchy problem for a fictitious dissipative evolution law that contains a tailored dissipation mechanism. If the dissipation mechanism is well designed, the solution of the Cauchy problem, with a given (well-prepared) initial condition, exists globally in time, has a limit for $t \to +\infty$, and the limit is an equilibrium of the considered physical system. The dynamical evolution itself might not be physical, but the solution should converge to a physical equilibrium as fast as possible. Some care may be taken to preserve important properties of the solution. For the specific case of MHD, for instance, one can evolve the magnetic field B according to Faraday's equation,

$$\partial_t B + c\,\mathrm{curl}\,E = 0, \quad B(0,x) = B_0(x),$$

but with a properly chosen effective electric field E, which does not need to have physical meaning. (Note, Gaussian units are used throughout this paper, with c being the speed of light in free space.) If the initial condition B_0 satisfies $\mathrm{div}\,B_0 = 0$, then $\mathrm{div}\,B = 0$ for all $t \geq 0$. Probably the most intuitive relaxation method can be obtained by choosing $E = -(U \times B)/c$, where the advecting velocity field U solves the viscous MHD momentum balance equation [21]. The idea of this method is physically intuitive: magnetic energy is converted into kinetic energy and dissipated by viscosity. Since there is no resistivity, magnetic helicity is preserved and this provides a lower bound for the dissipation of magnetic energy [22,23]. The evolution of the system is not physically consistent (the viscosity term is usually very simple and resistivity is zero), but the relaxed state, if it is reached and it is smooth enough, is guaranteed to be an ideal MHD equilibrium. However, while a relaxation method in general seeks to find an ideal MHD equilibrium as the long-time limit of the evolution of a given initial configuration, in most applications, we ask for the answer to a different question: we seek an equilibrium that is compatible with given data. For instance one might need to impose given pressure and current profiles, i.e., the constant value of the pressure and the concatenated plasma current on the surfaces tangent to the magnetic field (flux surfaces). This raises the question of relaxing to an equilibrium compatible with the given data starting from a suitable initial condition. This problem is related to the concept of *accessibility* of an equilibrium since the relaxation mechanisms usually entail constraints: the solution of a relaxation method, formally at least, evolves on the constrained submanifold that contains the initial condition, but this submanifold may not contain equilibria compatible with the given data. One therefore needs either to prepare the initial condition appropriately (by making assumption on the targeted equilibrium) or to adapt the solution during the evolution, using the available data.

For instance, with the choice of electric field $E = -(U \times B)/c$ mentioned above, the Faraday's equation reduces to Lie-dragging of the magnetic field and therefore, smooth solutions preserve the magnetic flux and the field-line topology of the initial condition (frozen-in law [24], cf. also general MHD textbooks [5,25]). In this case, Moffatt has introduced two different concepts [21,23]:

- *Topological equivalence.* Two vector fields B_0 and B_1 are topologically equivalent if there is a diffeomorphism that maps one field into the other via push-forward. We may think of topologically equivalent fields as one being a smooth deformation of the other.

- *Topological accessibility*. A vector field B_1 is topologically accessible from the vector field B_0 if $B_1(x) = B(t, x)$ for some $t \geq 0$, where B is the solution of the Faraday's equation with $E = -(U \times B)/c$ and some velocity field U, i.e., the MHD induction equation. We note that in its original definition, Moffatt restricted U to be solenoidal [21].

If B_1 is topologically accessible from B_0 with U sufficiently smooth (e.g., of class C^1 with a C^2 flow, as functions of (t, x)), then B_0 and B_1 are also topologically equivalent. In general however, the solution B of the induction equation may develop a singularity in finite time. More precisely, the magnetic field may develop tangential discontinuities at certain surfaces that correspond to current sheets. In fact, according to an argument put forward by Parker [26], the formation of current sheets is a general occurrence for braided fields, i.e., "most" braided initial conditions should develop current sheets. This is known as the *Parker's conjecture*. Recently, Enciso and Peralta-Salas [27] have proven that, on axisymmetric toroidal domains, there exists a set of smooth braided solenoidal vector fields that are *not* topologically equivalent to any MHD equilibrium. Furthermore, this set is rather large, in the sense that it is dense in a nonempty open subset of the space of smooth braided solenoidal fields (equipped with the C^∞ topology). This suggests that a relaxation method based on the MHD induction equation should either allow for low-regularity solutions with the possible formation of current sheets, as conjectured by Parker, or be complemented with a way to prepare a suitable initial condition. We note that equilibria with current sheets can be acceptable in some applications as discussed above in relation to the Grad's conjecture.

In summary, from the applications point of view, relaxation methods constructed in this way are not fully satisfactory because of the following drawbacks: (1) not all equilibrium points are accessible from a given initial condition. (2) The method does not offer any mechanism to control important properties of the equilibrium such as the pressure profile and current profiles. (3) The relaxation mechanisms based solely on viscosity do not necessarily lead to the shortest path from the initial condition down to an equilibrium point.

An example of a relaxation method based on viscosity is implemented in the HINT code [28,29]. In HINT, pressure is relaxed with an ad hoc algorithm, in a separate step, during the magnetic field relaxation. If resistivity is accounted for in the relaxation of the magnetic field, then the topology of the magnetic field lines can change, but with finite resistivity helicity is no longer preserved and there is no lower bound for the magnetic energy.

Another relaxation method that seeks a faster way to relax the magnetic energy is based on the variational principle for the equilibrium conditions (reviewed in Appendix C for the case of Beltrami fields). This method has been proposed by Chodura and Schlüter [30], cf. also Moffatt [23, sec. 8.2], and it can be specialized to the case of nonlinear Beltrami fields [31]. The idea is to Lie drag both the magnetic field and the pressure with an advecting velocity field U chosen to guarantee the maximum decay rate of the magnetic energy. These ideas are strictly related to the modern theory of optimal transport of differential forms [32].

The idea of Lie dragging both the magnetic field and the pressure has been exploited in the SIESTA code as well [33], where each displacement of both magnetic field and pressure is generated by an infinitesimal "Lie dragging step".

1.2. Overview of metriplectic relaxation

Metriplectic dynamics is a class of dynamical systems. Its mathematical structure has associated relaxation methods with desirable properties for calculating equilibria. Metriplectic dynamics and its concomitant variational principles for equilibria will be thoroughly reviewed in Section 2. In this subsection we give a overview of the paper while describing the usage in this work, where we explore using *artificially* constructed metriplectic dynamical systems in order to construct relaxation methods for the calculation of equilibrium points. Metriplectic dynamics was introduced by Morrison [1–3] as a generalization of noncanonical Hamiltonian dynamics with the aim of including dissipative phenomena. (See [34–39] for recent developments.) The equation of evolution is constructed in terms of two algebraic structures: a Poisson bracket [4], which is antisymmetric and defines the Hamiltonian part of the equation, and a metric bracket, which is symmetric and accounts for dissipation. In addition to the brackets, a Hamiltonian function $\mathcal{H}$ and an entropy function $\mathcal{S}$ are given, satisfying appropriate compatibility conditions. As a direct consequence of the construction, the Hamiltonian $\mathcal{H}$ is conserved and the entropy $\mathcal{S}$ is dissipated (more precisely, it is nonincreasing). The fact that entropy is a monotonic function of time quantifies the dissipation in the system. Defining entropy to be nonincreasing is inconsistent with its usual physical interpretation as a measure of uncertainty or "disorder", which would require it to be nondecreasing. In this work however, entropy is treated as a Lyapunov function [40], and thus we prefer to reverse the sign and work with a nonincreasing entropy. Many physically relevant mathematical models have been found to possess a metriplectic structure. For instance the Vlasov-Maxwell-Landau system [3], various fluid mechanical systems [2,36,39,41], visco-resistive MHD [42], and dissipative extended MHD [34]. This structure can also be exploited in order to design numerical schemes that preserve the key features of the physical model [43,44]. Here instead we shall use metric brackets as equilibrium solvers.

Not to be confused with metriplectic relaxation is a relaxation method that uses the Hamiltonian structure. This method, which is based on squaring the Poisson bracket, was introduced in [45,46] for two-dimensional vortical motion of neutral fluids; it was later generalized so as to make it more effective and work in a broader context in Flierl and Morrison [47] and it was applied to reduced MHD problems by Furukawa and coworkers [48–51] (see [52] for recent review). This approach, which has been named "double bracket" dynamics or simulated annealing, has different properties as compared to metriplectic dynamics: with double brackets, the Hamiltonian is dissipated while the system evolves on the a specific hypersurface (a symplectic leaf) determined by the constants of motion built into the Poisson brackets (Casimir invariants). Another early approach is that of Brockett who used a version of the double bracket for matrices constructed out of the commutator [53] (see also the work of Bloch and coauthors [54,55]).

Since metriplectic dynamics dissipates entropy but preserves the Hamiltonian, one might expect that a global solution, if it exists, will approach a minimum of the entropy function restricted to the level set of the Hamiltonian that contains the initial condition.

Specifically, if we denote by u a point in the phase space V of the system, we might expect that the solution of a generic metriplectic system with Hamiltonian $\mathcal{H}(u)$, entropy $\mathcal{S}(u)$, and initial condition $u(0) = u_0 \in V$ would converge to a solution of the variational principle

$$\min\{\mathcal{S}(u) : u \in V, \ \mathcal{H}(u) = \mathcal{H}(u_0)\}. \tag{1}$$

Variational principles of the form (1) are physically relevant since, the resulting state of a physical relaxation mechanism can, in certain cases, be well approximated by the solution of such a minimum entropy principle. Typically one envisages a situation where the physical evolution of the system is nearly ideal, that is, dissipation mechanisms are very small, but such small nonideal effects are sufficient to dissipate one of the ideal constant of motion, while causing negligible variations in the other constants of motion. This is referred to as selective decay and plays an important role in the relaxation of fluids and plasmas to a self-organized state [56,57]. In MHD for instance, linear Beltrami fields are found as a result of processes that dissipate magnetic energy, while (approximately) preserving magnetic helicity as argued by Woltjer [58]. The precise physical relaxation mechanism has been discussed in detail by Taylor and it is referred to as Woltjer-Taylor relaxation [59–61]. In Section 2.2 we shall see that equilibrium problems can often be reformulated as a variational principle of the form (1): among the many possible solutions of the equilibrium conditions, the variational principle (1) selects only those that minimize entropy on a constant energy hypersurface, and thus reduces significantly the issue of nonuniqueness of the equilibrium. If the equilibrium problem of interest is formulated as a variational principle of the form (1), then a relaxation method for such a problem should converge to a minimum of entropy on the constant-energy surface. A significant part of this work is dedicated to understanding when metriplectic systems have this long-time convergence property.

When the solution of a metriplectic system has a limit for $t \to +\infty$ and the limit is a solution of (1), we say that the system has *completely relaxed*. Unfortunately, complete relaxation does not always happen. It depends on the null space of the metric brackets, which is defined precisely in Section 2.1. We shall demonstrate complete relaxation (or lack thereof) by means of numerical experiments. We shall propose and test a particular class of brackets modeled upon Morrison's brackets for the Landau collision operator [1,3] and show by means of numerical experiments that these new brackets completely relax an initial condition. These new brackets are referred to as collision-like metric bracket.

Collision-like brackets have the disadvantage of generating integro-differential evolution equations, that are usually computationally expensive (although efficient methods exist [62]). In an attempt to reduce the computational cost of the relaxation methods, we have introduced a simplified version of the collision brackets that are local and thus lead to pure partial differential equations that have the structure of diffusion equations. We refer to these simplified brackets as diffusion-like. We are able to recover known relaxation methods, such as the metriplectic bracket based on Nambu dynamics of [55], which was also obtained and proposed in the context of vortex dynamics in [63,64], and the method of Chodura and Schlüter [30], as special cases of diffusion-like brackets. All of the brackets used in this paper follow naturally from the inclusive 4-bracket construction given in [35].

The remainder of the paper is structured as follows. As noted above, in Section 2 we cover review material: we recall the precise definition of metriplectic systems in Section 2.1 and describe the equilibrium problems that we consider as test cases together with corresponding variational principles in Section 2.2. Section 3 presents some mathematical results on relaxation and the relaxation rate for metriplectic systems, including results that we are unable to find in the literature. In Sections 3.1 and 3.2 we prove extensions of the Lyapunov stability theorem and the Polyak–Łojasiewicz condition for the rate of relaxation, for finite-dimensional systems with the inclusion of constraints, respectively, while in Section 3.3 we make some comments on the extension of these results to infinite-dimensional systems. In Section 4, we address simple examples of metric brackets and study the issue of complete relaxation, both analytically and numerically. Section 4.1 describes brackets, which we refer to as metric double brackets, that fail to completely relax, while Section 4.2 describes projection based brackets that do completely relax. Collision-like brackets are introduced in Section 5 with theory presented in Sections 5.1–5.3 and numerical experiments in Sections 5.4 and 5.5 for the Euler equations in vorticity form and for the Grad-Shafranov MHD equilibria, respectively. For these applications complete relaxation of the solution is critical. Section 6 is dedicated to diffusion-like brackets, with their general construction and various forms given in Sections 6.1–6.3. We shall see that their properties make them most suitable for applications such as the calculation of *nonlinear* Beltrami fields, which are considered in Section 6.4, for which the property of complete relaxation is not needed. Finally, we conclude in Section 7.

2. Metriplectic dynamics and variational principles for equilibria of fluids and plasmas

In this section, we review the necessary background material. First we briefly recall the definition and basic properties of metriplectic dynamics. Then we formulate and discuss examples of equilibrium problems, and their variational formulations. These problems will be used as a test bed for the metriplectic relaxation method proposed in this work.

2.1. Metriplectic dynamics

Metriplectic dynamics, being a special kind of continuous-time dynamical system, is determined by a phase space and a law governing the evolution in time of a point in the phase space. In this work, we are mainly concerned with infinite-dimensional metriplectic systems with a phase space given by a Banach space V of functions over a domain $\Omega \subset \mathbb{R}^d$ with values in $\mathbb{R}^N$, where $d, N \in \mathbb{N}$. (We use "domain" as a shorthand for "open, connected set".) We shall however briefly discuss finite-dimensional examples for the sake of clarity and simplicity. In the finite-dimensional case the phase space is chosen to be an open connected subset $\mathcal{Z} \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, with coordinates $z = (z^i)_{i=1}^n$.

For any $\mathcal{F} \in C^1(V)$, we denote by $D\mathcal{F}(u)$ its Fréchet derivative [65] at the point $u \in V$. We shall also need the functional derivative of $\mathcal{F}$. If W is another Banach space with a nondegenerate pairing $\langle \cdot, \cdot \rangle_{V\times W} : V \times W \to \mathbb{R}$, we can define the functional derivative $\delta\mathcal{F}/\delta u$ of $\mathcal{F}$ in W as the unique element of W, if it exists, such that [66]

$$D\mathcal{F}(u)v = \left\langle v, \frac{\delta \mathcal{F}(u)}{\delta u} \right\rangle_{V\times W}, \quad \forall v \in V. \tag{2}$$

Unless otherwise stated, in this work we assume $V \subseteq L^2(\Omega, \mu; \mathbb{R}^N)$, $W = L^2(\Omega, \mu; \mathbb{R}^N)$, and the pairing is the L^2 product with respect to a given measure $d\mu = m(x)dx$, where m is a smooth and integrable function and dx the Lebesgue measure (volume element) on Ω,

$$(u, v)_{L^2} = \int_\Omega u(x) \cdot v(x) d\mu(x), \quad u, v \in L^2(\Omega, \mu; \mathbb{R}^N).$$

(The nontrivial measure is needed in order to accommodate cases such as Grad-Shafranov equilibria discussed below.)

A metriplectic dynamical system is specified by giving two functions $\mathcal{H}, \mathcal{S} \in C^\infty(V)$, namely the Hamiltonian and the entropy, respectively, together with compatible Poisson and metric brackets on V. We recall the definitions [1,3,4,66,67].

A *Poisson bracket* on V is a bilinear *antisymmetric* map

$$\{\cdot, \cdot\} : C^\infty(V) \times C^\infty(V) \to C^\infty(V),$$

such that, for any $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{H}$ in $C^\infty(V)$,

$$\{\mathcal{F}, \mathcal{G}\mathcal{H}\} = \{\mathcal{F}, \mathcal{G}\}\mathcal{H} + \mathcal{G}\{\mathcal{F}, \mathcal{H}\}, \tag{3a}$$

$$\big\{\mathcal{F}, \{\mathcal{G}, \mathcal{H}\}\big\} + \big\{\mathcal{G}, \{\mathcal{H}, \mathcal{F}\}\big\} + \big\{\mathcal{H}, \{\mathcal{F}, \mathcal{G}\}\big\} = 0. \tag{3b}$$

Eqs. (3a) and (3b) are referred to as the Leibniz identity and the Jacobi identity, respectively. Then a Poisson bracket defines a Lie algebra structure on $C^\infty(V)$ which in addition is a derivation in each argument.

A *metric bracket* on V is a bilinear *symmetric* map

$$(\cdot, \cdot) : C^\infty(V) \times C^\infty(V) \to C^\infty(V),$$

such that, for any $\mathcal{F}$ in $C^\infty(V)$,

$$(\mathcal{F}, \mathcal{F}) \geq 0.$$

By definition, the Poisson bracket must satisfy Leibniz and Jacobi identities. Leibniz identity, in particular, implies at least formally that the bracket can be written in term of the functional derivatives of its arguments, cf. Appendix A for a precise definition. Usually, the symmetric bracket does not need to satisfy any condition other then bilinearity, symmetry, and positive semidefiniteness. However, if one requires the symmetric bracket to satisfy Leibniz identity,

$$(\mathcal{F}, \mathcal{G}\mathcal{H}) = (\mathcal{F}, \mathcal{G})\mathcal{H} + \mathcal{G}(\mathcal{F}, \mathcal{H}), \tag{4}$$

then both the Poisson and the symmetric brackets have a similar representation, i.e.,

$$\{\mathcal{F}, \mathcal{G}\} = \sum_{i,j=1}^N \int_\Omega \int_\Omega \frac{\delta \mathcal{F}(u)}{\delta u_i}(x) \mathcal{J}_{ij}(u; x, x') \frac{\delta \mathcal{G}(u)}{\delta u_j}(x')\, d\mu(x')\, d\mu(x), \tag{5a}$$

and analogously

$$(\mathcal{F}, \mathcal{G}) = \sum_{i,j=1}^N \int_\Omega \int_\Omega \frac{\delta \mathcal{F}(u)}{\delta u_i}(x) \mathcal{K}_{ij}(u; x, x') \frac{\delta \mathcal{G}(u)}{\delta u_j}(x')\, d\mu(x')\, d\mu(x), \tag{5b}$$

where the functional derivatives are computed with respect to the L^2 product with a given measure μ on Ω. The kernels $\mathcal{J}(u)$ and $\mathcal{K}(u)$ define an anti-symmetric and a symmetric, positive semidefinite operator, $J(u)$ and $K(u)$, respectively. In finite dimensions, Eqs. (5) take the form

$$\{\mathcal{F}, \mathcal{G}\} = J^{ij}(z) \frac{\partial \mathcal{F}(z)}{\partial z^i} \frac{\partial \mathcal{G}(z)}{\partial z^j},$$
$$(\mathcal{F}, \mathcal{G}) = K^{ij}(z) \frac{\partial \mathcal{F}(z)}{\partial z^i} \frac{\partial \mathcal{G}(z)}{\partial z^j},$$

where here the sum over repeated indices ranges to n, $J(z)$ is an antisymmetric contravariant tensor, and $K(z)$ is a symmetric positive semidefinite contravariant tensor over the domain $\mathcal{Z} \subseteq \mathbb{R}^n$. In particular J is referred to as the Poisson tensor.

The evolution equation for a metriplectic system $u(t) \in V$ is formulated as an evolution equation for arbitrary functions of $u(t)$, that is,

$$\frac{d\mathcal{F}}{dt} = \{\mathcal{F}, \mathcal{H}\} - (\mathcal{F}, \mathcal{S}), \quad \text{for all } \mathcal{F} \in C^\infty(V), \tag{6a}$$

where $\mathcal{H}, \mathcal{S} \in C^\infty(V)$ are the Hamiltonian and entropy functions, respectively, whereas $\{\cdot, \cdot\}$ and $(\cdot, \cdot)$ are the Poisson and metric bracket on V, respectively, satisfying the compatibility conditions

$$\{\mathcal{F}, \mathcal{S}\} = 0, \quad (\mathcal{F}, \mathcal{H}) = 0, \quad \text{for all } \mathcal{F} \in C^\infty(V). \tag{6b}$$

If both brackets satisfy the Leibniz identity, the evolution equation then reads

$$\partial_t u = J(u)\frac{\delta \mathcal{H}(u)}{\delta u} - K(u)\frac{\delta \mathcal{S}(u)}{\delta u}.$$

In general, both $J(u)$ and $K(u)$ have nontrivial null spaces. Per definition, the null space of a bracket is identified with that the corresponding operators $J(u)$ and $K(u)$. We note that, in general, the null space of a bracket depends on the phase-space point u, since J and K depend on u. The null space of the Poisson bracket is due to the noncanonical form that often originates from a reduction procedure based on the symmetries of the system [68,69] (see for example the texts [70,71]), while the null space of the metric bracket is due to the requirement that at least energy is preserved, cf. Eq. (6b). In finite dimensions, the null space of $J(z)$ is spanned by the gradients of Casimir invariants and equilibria from the variational principle align with those of the equations of motion [4]. We note, however, that there are subtleties at points where the rank of $J(z)$ changes [72] and equilibria at such points can possess nearby behavior that is not Hamiltonian [73,74]. The null space of $K(z)$ contains at least the gradient of the Hamiltonian, due to the compatibility condition (6b); similar remarks on rank changing could apply. For a metric bracket defined on a Banach space $V \subseteq L^2(\Omega, \mu; \mathbb{R}^N)$ and corresponding to a bounded operator $K(u)$ on $L^2(\Omega, \mu; \mathbb{R}^N)$, the null space at $u \in V$ can be equivalently characterized as the space of functions $\mathcal{F}$ over V such that $\big(\mathcal{F}, \mathcal{F}\big)(u) = 0$, since this is equivalent to $\delta\mathcal{F}(u)/\delta u \in \ker K(u)$.

In general, the vector field $J(u)\delta\mathcal{H}/\delta u$ can be viewed as a generalization of Hamiltonian flow $\omega(u)^{-1}\delta\mathcal{H}/\delta u$ where the inverse of the symplectic operator $\omega(u)$ is replaced by a possibly noninvertible Poisson operator $J(u)$. On the other hand, the vector field $-K(u)\delta\mathcal{H}/\delta u$ can be viewed as generalization of a gradient flow $-G(u)^{-1}\delta\mathcal{H}/\delta u$ where the inverse of the metric operator $G(u)$ is replaced by a symmetric, positive *semi*definite operator. Metriplectic dynamics combines the (generalized) symplectic and gradient flows. Typically, the symplectic part describes the ideal dynamics, while the gradient flow accounts for a nonideal relaxation mechanism. For this reason, accepting a slight abuse of terminology, we refer to symmetric brackets $(\cdot,\cdot)$ with the Leibniz property as *metric brackets*. We shall always tacitly assume that $\mathcal{H}$ and $\mathcal{S}$ are not functionally dependent, i.e., $\delta\mathcal{S}(u)/\delta u$ is not everywhere parallel to $\delta\mathcal{H}(u)/\delta u$, otherwise the metric bracket part vanishes identically.

We are mostly interested in the dynamical systems generated by metric brackets, i.e., we shall drop the Poisson bracket part,

$$\frac{d\mathcal{F}}{dt} = -(\mathcal{F}, \mathcal{S}), \quad \text{for all } \mathcal{F} \in C^\infty(V), \tag{7a}$$

where $\mathcal{S}$ is the entropy function and the metric bracket satisfies the compatibility condition

$$(\mathcal{F}, \mathcal{H}) = 0, \quad \text{for all } \mathcal{F} \in C^\infty(V), \tag{7b}$$

where $\mathcal{H}$ is the Hamiltonian.

A solution of either (6) or (7) satisfies

$$\frac{d\mathcal{H}(u)}{dt} = 0, \qquad \frac{d\mathcal{S}(u)}{dt} \le 0, \tag{8}$$

that is, both system (6) and (7) dissipate entropy on the surface of constant energy (Hamiltonian).

Because of (8) one may expect that solutions of the variational principle (1) are necessarily equilibria of a metriplectic system. Indeed this is the case and it follows from the method of Lagrange multipliers [66], which gives a necessary condition for (1): if u is a solution of (1), then there is a constant $\lambda \in \mathbb{R}$ (the Lagrange multiplier) such that

$$D\mathcal{S}(u) - \lambda D\mathcal{H}(u) = 0, \quad \mathcal{H}(u) = \mathcal{H}_0\,. \tag{9}$$

Alternatively, one can write the Lagrange condition using the functional derivative, if they exist,

$$\frac{\delta\mathcal{S}(u)}{\delta u} - \lambda\frac{\delta\mathcal{H}(u)}{\delta u} = 0, \quad \mathcal{H}(u) = \mathcal{H}_0\,.$$

Eq. (9) constitute a system of two equations for the pair $(u, \lambda) \in V \times \mathbb{R}$. Let us assume that the set of solutions

$$\mathfrak{C}_{\mathcal{H}_0} := \{u \in V : \exists \lambda \in \mathbb{R} \text{ such that } (u, \lambda) \text{ solves } (9)\},$$

is nonempty ($\mathfrak{C}_{\mathcal{H}_0} \neq \emptyset$), the restriction $\mathcal{S}$ to $\mathfrak{C}_{\mathcal{H}_0}$ is bounded from below, and the minimum is attained, that is, there are points $u_e \in \mathfrak{C}_{\mathcal{H}_0}$ where $\mathcal{S}(u_e) = \min\{\mathcal{S}(u) : u \in \mathfrak{C}_{\mathcal{H}_0}\}$. Then, the solutions of (1) correspond to those points u_e of $\mathfrak{C}_{\mathcal{H}_0}$ where $\mathcal{S}$ attains its minimum. The set $\mathfrak{C}_{\mathcal{H}_0}$ is the set of constrained critical points of $\mathcal{S}$, that is, of critical points of $\mathcal{S}$ restricted to the energy surface $\mathcal{H}(u) = \mathcal{H}_0$.

If the symmetric bracket satisfies the Leibniz identity, the Lagrange condition together with either the compatibility condition (6a) or (7b) imply that any point in $\mathfrak{C}_{\mathcal{H}_0}$, i.e., any constrained critical point of $\mathcal{S}$, is an equilibrium point of the metriplectic system, and thus, in particular, any solution u_e of (1) is necessarily an equilibrium point [55, sec. 4.1]. The converse, however, is not true, since all constrained critical points are equilibria, not just the minima. In addition, there can be equilibrium points of either (6) or (7) that are not constrained critical points of $\mathcal{S}$. One example is given in Section 4.1 below.

The fact that the set of equilibrium points of a metriplectic system can, in general, be (much) larger than the set $\mathfrak{C}_{\mathcal{H}_0}$ of constrained critical points of $\mathcal{S}$ can be an obstruction to convergence of an orbit $u(t)$ to a solution of (1), and this is important in some (but not all) applications. In the next section, we review a few physically relevant equilibrium problems and discuss their relation to variational principles of the form (1). The main result of the paper is the construction of appropriate metric brackets that relax a given initial condition to a solution of such equilibrium problems.

2.2. Examples of equilibrium problems

In this section, we review the examples of equilibrium problems that we shall use as test cases for metriplectic relaxation. All considered test problems are mathematically ill-posed, because they admit multiple solutions. In some cases, the ill-posedness can be mitigated by adding additional physics constraints. We shall also discuss the variational principles for the considered equilibrium problems.

2.2.1. Reduced Euler equations

We begin with the Euler equations reduced to two dimensions [4, p.488 and references therein], which is the simplest of a hierarchy of models including the reduced MHD model [75]. Let $x = (x_1, x_2)$ be Cartesian coordinates in a bounded domain $\Omega \subset \mathbb{R}^2$ with a sufficiently regular boundary $\partial\Omega$. In Ω, we consider an incompressible flow $U = (U_1, U_2) = \big(\partial_2\phi, -\partial_1\phi\big)$, given in terms of a stream function $\phi(x)$ with $\partial_i = \partial/\partial x_i$. Then $\operatorname{div} U = 0$ and from the definition of the scalar vorticity $\omega := \partial_1 U_2 - \partial_2 U_1$ one obtains the Poisson equation

$$-\Delta\phi = \omega \text{ in } \Omega\,, \quad \phi = 0 \text{ on } \partial\Omega\,, \tag{10}$$

where Δ is the Laplace operator in $\mathbb{R}^2$. Vice versa, given ω, we can solve (10) for ϕ and reconstruct the flow U. Hence, the incompressible Euler equations in two-dimensions

$$\partial_t U + U \cdot \nabla U = -\nabla p, \quad \operatorname{div} U = 0,$$

with p being the pressure field, amount to an evolution equation for the scalar vorticity ω,

$$\begin{cases} \partial_t\omega + [\omega,\phi] = 0, & \text{in } \Omega, \\ -\Delta\phi - \omega = 0, & \text{in } \Omega, \\ \phi = 0, & \text{on } \partial\Omega, \end{cases}$$

where $[\omega,\phi] := \partial_1\omega\,\partial_2\phi - \partial_2\omega\,\partial_1\phi$ is the canonical Poisson bracket in $\mathbb{R}^2$. This model is referred to as the reduced Euler equations.

The phase space V of the reduced Euler equations is the space of vorticity fields, i.e., $u = \omega$, and $\phi = -\Delta^{-1}_{\Omega,0}\,\omega$ is regarded as a function of ω, given by the inverse of the Laplacian on Ω with homogeneous Dirichlet boundary conditions.

The equilibrium problem for the reduced Euler equations then reads

$$\begin{cases} [\omega,\phi] = 0, & \text{in } \Omega, \\ -\Delta\phi - \omega = 0, & \text{in } \Omega, \\ \phi = 0, & \text{on } \partial\Omega. \end{cases} \tag{11}$$

Problem (11) admits many solutions and therefore is mathematically ill-posed. One can in fact construct a large class of solutions upon noticing that $[\omega,\phi] = 0$ implies that ω is constant on the isolines (contours) of the potential ϕ. The contours may have many connected components, each one with a possibly different topology (i.e., homeomorphic to a different model space) and the constant value of ω on different connected components may be different. Given a function $f \in C^1(\mathbb{R})$, we may set $\omega = \lambda f(\phi)$ with a normalization factor $\lambda \in \mathbb{R}$ to be determined; in this way we assign the same value of ω to all connected components of the same contour of ϕ. This is a special case which is considered here for sake of simplicity. Then we consider the problem: find (ϕ,λ), with $\lambda \neq 0$, such that

$$\begin{cases} -\Delta\phi = \lambda f(\phi), & \text{in } \Omega, \\ \phi = 0, & \text{on } \partial\Omega. \end{cases} \tag{12}$$

Any solution (ϕ,λ) of (12) yields a solution $\omega = \lambda f(\phi)$ of the equilibrium problem (11). The case $\lambda = 0$ leads to the trivial equilibrium $\omega = 0$ and it is not considered.

Problem (12) is a "eigenvalue problem" for a semilinear elliptic equation. If $f'(y) \leq 0$ (respectively, $f'(y) \geq 0$) for all y, the equation has a unique solution for any $\lambda \geq 0$ (respectively, $\lambda \leq 0$) [76]. In the other cases, the solution may not exist for all λ; if a solution exists, uniqueness is not guaranteed, e.g., for degenerate eigenvalues. For instance, when $f(y) = y$, problem (12) reduces to the standard eigenvalue problem for the Laplace operator with Dirichlet boundary conditions; then, we have discrete, possibly degenerate, positive eigenvalues $\lambda_n > 0$, $n \in \mathbb{N}$, each with a corresponding finite set of eigenfunctions $\phi_{n,k}$ depending on the multiplicity of λ_n. For $\lambda \leq 0$, the trivial solution $\phi = 0$ is the unique solution.

This, in particular, shows that problem (11) is ill-posed, since there is a rich set of solutions for each choice of f, and many choices of f are possible. In order to mitigate the nonuniqueness problem, in practice the function f, which will be referred to as the *equilibrium profile*, is prescribed, and among the solutions of (12), the one with the lowest λ is considered. The reformulation of the equilibrium problem (11) into the nonlinear eigenvalue problem (12) with fixed f is good enough in practice. There are efficient iterative algorithms [77] for the solution of this type of eigenvalue problem with the lowest λ.

An alternative reformulation of the equilibrium problem (11) is possible, based on a variational principle of the form (1). For $\mathcal{H}_0 > 0$ and $s \in C^2(\mathbb{R})$ satisfying $s''(y) \neq 0$, $y \in \mathbb{R}$, let us consider the problem

$$\min\{\mathcal{S}(\omega) : \mathcal{H}(\omega) = \mathcal{H}_0\}. \tag{13}$$

where

$$\mathcal{S}(\omega) = \int_\Omega s(\omega)dx, \quad \mathcal{H}(\omega) = \frac{1}{2}\int_\Omega |\nabla\phi|^2 dx, \tag{14}$$

are the entropy and Hamiltonian functions, respectively, with ϕ depending on ω via the Poisson Eq. (10). The condition on the entropy profile $s(y)$ implies that $s'(y)$ is a strictly monotonic function, either decreasing or increasing. In general, $s(y)$ will be chosen *ad hoc* and does not necessarily have a physical meaning. The Hamiltonian $\mathcal{H}$ is the kinetic energy of the incompressible fluid, since $|\nabla\phi|^2 = |U|^2$, where U is the flow velocity.

Problem (9) in this case reads: Find (ω, λ), such that

$$\begin{cases} s'(\omega) - \lambda\phi = 0, \quad \mathcal{H}(\omega) = \mathcal{H}_0, \\ \text{with } \phi \text{ solution of (10).} \end{cases} \tag{15}$$

The set $\mathfrak{C}_{\mathcal{H}_0}$ of constrained critical points, cf. Section 2.1, amounts to

$$\mathfrak{C}_{\mathcal{H}_0} = \{\omega : \ (\omega, \lambda) \text{ solves (15) for some } \lambda \in \mathbb{R}\}.$$

All elements of the set $\mathfrak{C}_{\mathcal{H}_0}$ are solutions of the original equilibrium problem (11). In fact, for $\lambda \neq 0$, we have $[\omega, \phi] = \lambda^{-1}[\omega, s'(\omega)] = 0$. The case $\lambda = 0$ is somewhat special, since $s'(\omega) = 0$ with $s''(y) \neq 0$ implies that $\omega(x) = \omega_c(x) = y_c =$ constant, where y_c is the unique zero of $s'(y)$; the corresponding potential ϕ_c is given by the solution of problem (10) with constant right-hand side. There is therefore only one solution for $\lambda = 0$ and this carries the energy $\frac{1}{2}\|\nabla\phi_c\|^2_{L^2} = \mathcal{H}_c$ and it is always an equilibrium, since ω_c is constant. If $\mathcal{H}_0 = \mathcal{H}_c$, this solution belongs to $\mathfrak{C}_{\mathcal{H}_0}$, otherwise $\lambda = 0$ is not a possible value for the Lagrange multiplier.

Under the hypothesis $s''(y) \neq 0$, $s'(y)$ is monotonic, and thus an invertible function of $y \in \mathbb{R}$. Problem (15) is related to the eigenvalue problem (12) with equilibrium profile given by $f(y) = (s')^{-1}(y)$. Precisely, if (ω, λ) is a solution of (15) with $\lambda \neq 0$, we can define $\tilde{\omega} = \lambda\omega$ and $\tilde{\phi} = \lambda\phi$, and obtain from (15)

$$-\Delta\tilde{\phi} = \lambda\omega = \lambda(s')^{-1}(\lambda\phi) = \lambda f(\tilde{\phi}),$$

which shows that $\tilde{\phi}$ solves (12) with the eigenvalue being the same as the Lagrange multiplier λ.

Among all these equilibria, the minimization in (13) selects those with minimum entropy, thus mitigating the nonuniqueness problem, as shown in the following special case. As in Yoshida and Mahajan [57], we choose $s(y) = y^2/2$, hence $s'(y) = y$ and solutions of (15) must necessarily solve the eigenvalue problem for the Laplace operator on Ω with homogeneous Dirichlet boundary conditions. From standard theory [78], we know that there is an orthonormal basis $\{\phi_{j,k}\}$ in $L^2(\Omega)$ of eigenfunctions, $-\Delta\phi_{j,k} = \lambda_j\phi_{j,k}$, labeled by $j \in \mathbb{N}_0$ with $k = 1, \ldots, d_j$ counting the multiplicity and d_j being the dimension of the eigenspace corresponding to the eigenvalue $\lambda_j > 0$. Then, the set $\mathfrak{C}_{\mathcal{H}_0}$ comprises all and only the vorticity fields of the form

$$\omega_j = \lambda_j\phi_j, \quad \phi_j = \sum_{k=1}^{d_j} a_{j,k}\phi_{j,k},$$

for any $j \in \mathbb{N}_0$ and $a_j = (a_{j,k}) \in \mathbb{R}^{d_j}$ satisfying the energy constraint

$$\lambda_j a_j^2 = 2\mathcal{H}_0.$$

The energy constraint fixes the length of the vector a_j but not its direction. The function $\mathcal{S}$ restricted to $\mathfrak{C}_{\mathcal{H}_0}$ is given by

$$\mathcal{S}(\omega_j) = \mathcal{H}_0\lambda_j,$$

which is bounded from below, since the spectrum of $-\Delta$ is bounded from below and $\mathcal{H}_0$ is a constant. The solutions of (13) correspond to the eigenfunctions with minimum eigenvalue (the ground states). Usually the eigenspace corresponding to the lowest eigenvalue is one-dimensional hence we have two solutions that differ only by the sign of the vorticity. In this example, the variational principle (13) picks the solution of (12) with the lowest λ.

In this paper, we use metriplectic dynamics in order to solve the variational principle (13). For this particular application it is essential that the orbit of the chosen metriplectic dynamical system relaxes completely to a constrained entropy minimum.

2.2.2. *Axisymmetric MHD equilibria*

A similar equilibrium problem arises from axisymmetric ideal magnetohydrodynamic (MHD) equilibria of electrically conducting fluids. For MHD, the general equilibrium condition with zero flow amounts to [5,8,12,19,79]

$$J \times B = c\nabla p, \quad \operatorname{curl} B = 4\pi J/c, \quad \operatorname{div} B = 0, \tag{16}$$

where B and J are vector fields and p is a scalar field. Physically B and J are the magnetic field and the electric current density, respectively, while p is the fluid pressure. As noted above, Gaussian units are used with c being the speed of light in free space. The first equation expresses the force balance between the Lorentz force $J \times B/c$ and the pressure gradient ∇p. The force balance implies the necessary conditions

$$B \cdot \nabla p = 0, \quad J \cdot \nabla p = 0, \tag{17}$$

that is, p is constant on the field lines of both B and J.

For axisymmetric solutions, i.e., solutions that have rotational symmetry around an axis, we introduce cylindrical coordinates (r, φ, z) around the symmetry axis z, and from $\operatorname{div} B = 0$ it follows that [5],

$$B = \chi\nabla\varphi + \nabla\psi \times \nabla\varphi,$$

$$4\pi J/c = \operatorname{curl} B = -\Delta^*\psi\nabla\varphi + \nabla\chi\times\nabla\varphi,$$

where $\Delta^* = r[\partial_r(r^{-1}\partial_r)] + \partial_z^2$ is a linear elliptic second-order differential operator in (r,z) coordinates, the Grad-Shafranov operator, whereas $\chi(r,z)$, $p(r,z)$, and $\psi(r,z)$ are real-valued scalar functions. The operator ∇ is the full three-dimensional gradient. Then axisymmetric equilibria with zero flow must satisfy the conditions

$$\begin{aligned} &-\Delta^*\psi\nabla\psi - \chi\nabla\chi - 4\pi r^2\nabla p = [\psi,\chi]r\nabla\varphi,\\ &[\psi,p]=0, \quad [\chi,p]=0, \end{aligned}$$

where the brackets $[\cdot,\cdot]$ are the canonical Poisson brackets in the (r,z)-plane, e.g., $[\chi,\psi] = r\nabla\varphi\cdot(\nabla\psi\times\nabla\chi) = \partial_r\chi\partial_z\psi - \partial_r\psi\partial_z\chi$. The first equilibrium condition expresses the force balance of (16), while the latter two follow from the necessary conditions (17), respectively. With homogeneous Dirichlet boundary conditions, we can formulate the problem

$$\begin{cases} u\nabla\psi - \chi\nabla\chi - 4\pi r^2\nabla p = 0, & \text{in } \Omega,\\ [\psi,p]=0, \quad [\psi,\chi]=0, & \text{in } \Omega,\\ -\Delta^*\psi - u = 0, & \text{in } \Omega,\\ \psi = 0, & \text{on } \partial\Omega, \end{cases} \tag{18}$$

where $\Omega\subset\mathbb{R}_+\times\mathbb{R}$ is a domain in the (r,z) plane satisfying $r>0$ in the closure $\overline{\Omega}$ (so that Ω is bounded away from the singularity of cylindrical coordinates at $r=0$). The auxiliary variable u is related to the φ component of the current density, since $J_\varphi = J\cdot\nabla\varphi/|\nabla\varphi| = -c\Delta^*\psi/(4\pi r) = cu/(4\pi r)$. Without specifying other constraints this problem is ill-posed in the same way as problem (11): the vanishing of the two Poisson brackets in (18) implies that, if $\nabla\psi$, ∇p, and $\nabla\chi$ are all nonzero, both χ and p are constant on the level sets of ψ. However, the functional relation between χ, p and ψ is undetermined. Fortunately, providing such information is straightforward. If we prescribe $\chi = \sqrt{\lambda}F(\psi)$ and $p = \lambda G(\psi)$ for given functions $F,G\in C^1(\mathbb{R})$ and $\lambda>0$, we obtain $u = \lambda[(F^2/2)'(\psi) + 4\pi r^2 G'(\psi)] = \lambda f(r,\psi)$ and the equilibrium problem reduces to

$$\begin{cases} -\Delta^*\psi = \lambda f(r,\psi), & \text{in } \Omega,\\ \psi = 0, & \text{on } \partial\Omega, \end{cases} \tag{19}$$

which is referred to as the Grad-Shafranov equation, and is the analog of (12). (We remark that in realistic applications proper care must be taken to assign physically meaningful values of χ and p to different connected components of the ψ-contours.) Eq. (19) is an "eigenvalue problem" for a semilinear elliptic equation and thus the same remarks about the well-posedness of Eq. (12) are valid here. Solving this eigenvalue problem is the standard way of computing axisymmetric MHD equilibria in tokamaks [77,80,81]. As for the Euler equations, the Grad-Shafranov problem can be considered solved, and it is used in this work as a benchmark problem.

As for the case of reduced Euler equilibria, solutions of the axisymmetric MHD equilibrium conditions (18) can be characterized by a variational principle of the form (1). For a state variable we choose $u(r,z) = (4\pi/c)rJ_\varphi(r,z)$ defined over the bounded domain $\Omega\subset\mathbb{R}_+\times\mathbb{R}$, with $r>0$ on $\overline{\Omega}$, and we consider the measure $d\mu = r^{-1}drdz$ on Ω. We assume that the profiles $F,G\in C^2(\mathbb{R})$ are given so that $f(r,y) := (F^2/2)'(y) + 4\pi r^2 G'(y)$ satisfies $\partial_y f(r,y)\neq 0$. Therefore the map $y\mapsto f(r,y)$ is monotonic and has an inverse $y\mapsto \partial_y s(r,y)$ for any fixed r. After integration in y we find a function $s\in C^2(\mathbb{R}_+\times\mathbb{R})$, with $\partial_y^2 s(r,y)\neq 0$ and such that $\partial_y s(r,\cdot)^{-1} = f(r,\cdot)$. Then we consider the problem

$$\min\{\mathcal{S}(u):\ \mathcal{H}(u) = \mathcal{H}_0\}, \tag{20}$$

with entropy and Hamiltonian

$$\mathcal{S}(u) = \int_\Omega s(r,u)d\mu, \quad \mathcal{H}(u) = \frac{1}{2}\int_\Omega |\nabla_{r,z}\psi|^2 d\mu, \tag{21}$$

where ψ is regarded as a function of u given by the solution of the linear elliptic problem

$$-\Delta^*\psi = u, \quad \psi|_{\partial\Omega} = 0. \tag{22}$$

Here $\mathcal{H}$ amounts to the magnetic energy stored in the poloidal component $\nabla\psi\times\nabla\varphi$ of the magnetic field.

The functional derivatives are defined with respect to the L^2-product with the measure μ on Ω, so that

$$\frac{\delta\mathcal{S}(u)}{\delta u} = \partial_y s(r,u), \quad \frac{\delta\mathcal{H}(u)}{\delta u} = \psi.$$

Eq. (9) gives

$$\begin{cases} \partial_y s(r,u) - \lambda\psi = 0, \quad \mathcal{H}(u) = \mathcal{H}_0,\\ \text{with } \psi \text{ solution of (22)}. \end{cases} \tag{23}$$

The set $\mathfrak{C}_{\mathcal{H}_0}$ of constrained critical points is then

$$\mathfrak{C}_{\mathcal{H}_0} = \{u:\ (u,\lambda) \text{ solves (23) for some } \lambda>0\},$$

and among the elements of $\mathfrak{C}_{\mathcal{H}_0}$, those with minimum entropy are solutions of the entropy principle (1).

Under the assumption on $s(r, y)$, each element $u \in \mathfrak{C}_{\mathcal{H}_0}$ with $\lambda \neq 0$ corresponds to an axisymmetric equilibrium. In order to see this, we have to find the fields p and χ corresponding to u. Given the profiles F and G from which s has been derived, let us define

$$\chi = \sqrt{\lambda} F(\lambda\psi), \quad p = \lambda G(\lambda\psi),$$

together with the scaled variables $\tilde{u} = \lambda u$ and $\tilde{\psi} = \lambda\psi$. One can check that $\tilde{u}, \tilde{\psi}$ together with χ and p given above solve conditions (18). We also have that $\tilde{\psi}$ is a solution of the Grad-Shafranov equation.

We remark that in general, the variational principle (20) can be formulated for generic profiles $s(r, y)$ and as long as $\partial_y^2 s(r, y) \neq 0$, we can still find $f(r, \cdot) = \partial_y s(r, \cdot)^{-1}$ and a correspondence between (23) and (19). However, for general profiles, $f(r, y)$ cannot be written in terms of $F(y)$ and $G(y)$, since f may not be a quadratic function of r; therefore, one cannot always find χ and p and thus an equilibrium in the sense of (18).

As in the case of the reduced Euler equations, for the application of metriplectic dynamics to the solution of (20) it is essential that the orbits of the metriplectic system completely relax to a constrained entropy minimum.

2.2.3. Beltrami fields

Let $\Omega \subset \mathbb{R}^3$ be a bounded simply connected domain, with sufficiently regular boundary $\partial\Omega$, and let $n : \partial\Omega \to \mathbb{R}^3$ be the outward unit normal to Ω. A vector field $B : \Omega \to \mathbb{R}^3$ is called a *Beltrami field* (also known as *nonlinear* or *weak* Beltrami field) if it satisfies the Beltrami conditions

$$(\mathrm{curl}\, B) \times B = 0, \quad \mathrm{div}\, B = 0,$$

which, with the addition of the natural homogeneous boundary condition for a divergence-free field, lead to [82,83]

$$\begin{cases} (\mathrm{curl}\, B) \times B = 0, \quad \mathrm{div}\, B = 0, & \text{in } \Omega, \\ \qquad\qquad\qquad\quad\;\; n \cdot B = 0, & \text{on } \partial\Omega. \end{cases} \tag{24}$$

Solutions of (24) satisfy the ideal MHD equilibrium conditions (16) with constant pressure, hence Beltrami fields are force-free MHD equilibria [31]. The force-free condition can be equivalently rewritten as $\mathrm{curl}\, B = f B$ for a real scalar function f, referred to as the proportionality factor. From the divergence-free condition, $\mathrm{div}\, B = 0$, it follows that $B \cdot \nabla f = 0$, i.e., if B is a smooth nonlinear Beltrami field corresponding to a smooth non-constant scalar multiplier f, then f is a first integral of B. Enciso and Peralta-Salas have shown that the Beltrami conditions constrain the field in such a way that nontrivial solutions with a sufficiently regular proportionality factor f exist only if f satisfies a very restrictive condition, so that smooth nonlinear Beltrami fields are "rare" [84]. On the other hand, one can search for solutions with low regularity requirements, e.g., $B \in H^1(\Omega)^3$, the Sobolev space of fields $B : \Omega \to \mathbb{R}^3$ such that $B \in L^2$ and $\nabla B \in L^2$ componentwise. With the nonhomogeneous boundary condition $n \cdot B = g$ on $\partial\Omega$, an existence result has been proven for $B \in H^1(\Omega)^3$ and $f \in L^\infty(\Omega)$ by using a fixed-point argument [82], but on a simply connected domain this solution reduces to $B = 0$ if $g = 0$. For the purposes of this work, we are interested in a weaker formulation of the Beltrami condition, which will become relevant in Section 6.4. For this formulation, we need to introduce the Sobolev space $H_0(\mathrm{div}, \Omega)$ of L^2 vector fields w with $\mathrm{div}\, w$ in L^2 and $w \cdot n = 0$ on $\partial\Omega$. We also need the space $H(\mathrm{curl}, \Omega)$ of L^2 vector fields w with $\mathrm{curl}\, w$ in L^2, and its subspace $H_0(\mathrm{curl}, \Omega)$ of vector fields $w \in H(\mathrm{curl}, \Omega)$ satisfying the homogeneous boundary condition $w \times n = 0$ on $\partial\Omega$. Specifically, we consider the problem of finding $B \in H_0(\mathrm{div}, \Omega) \cap H(\mathrm{curl}, \Omega)$, such that

$$(\mathrm{curl}\, B) \times B = 0, \quad \mathrm{div}\, B = 0.$$

In this formulation the current $J = (4\pi/c)\,\mathrm{curl}\, B$ is only required to be in L^2; thus, it can have singularities as long as they are squared-integrable. Since the space $H_0(\mathrm{div}, \Omega) \cap H(\mathrm{curl}, \Omega)$ is not convenient for the purposes of a finite element discretization, in our numerical experiment in Section 6.4, we consider an even weaker formulation: find $B \in H_0(\mathrm{div}, \Omega)$, such that

$$j \times H = 0, \quad \mathrm{div}\, B = 0,$$

where the current j (different than $J = (4\pi/c)\,\mathrm{curl}\, B$) and H are the unique elements in $H_0(\mathrm{curl}, \Omega)$ such that $(j, k)_{L^2} = (B, \mathrm{curl}\, k)_{L^2}$ and $(H, G)_{L^2} = (B, G)_{L^2}$ for any $k, G \in H_0(\mathrm{curl}, \Omega)$. This formulation, in principle allows for even stronger current singularities. We are not aware of any existence result for either of these two formulations of the Beltrami problem. While looking for solutions with low regularity might appear physically obscure, there are two reasons to consider them. The first is that existence of smooth equilibria is an issue, and equilibria with singular currents are acceptable in some applications as discussed in Section 1.1 in the context of the Grad's conjecture. The second reason is that these spaces of functions are natural for modern numerical methods in MHD [85,86, and references therein].

A special class of Beltrami fields is given by the solutions of the eigenvalue problem for the curl operator: find (B, λ), $\lambda \in \mathbb{R}$, such that

$$\begin{cases} \mathrm{curl}\, B = \lambda B, & \text{in } \Omega, \\ n \cdot B = 0 & \text{on } \partial\Omega. \end{cases} \tag{25}$$

The eigenfunctions of this problem will be referred to as *linear* Beltrami fields and they are necessarily divergence-free. Since there is a countable family of eigenvalues of the curl operator, each corresponding to a finite-dimensional space of eigenfunctions [87,88], both problem (24) and (25) are mathematically ill-posed due to nonuniqueness of the solution (although for (25), this is standard, since it is an eigenvalue problem).

Linear Beltrami fields can be characterized by a variational principle of the form (1), which has been proposed by Woltjer [58] and later applied to self-organized states in fusion plasmas by Taylor [59]. This variational principle is also central in multi-region relaxed MHD [89,90]. Generalizations of Woltjer's variational principle have been proposed by Dixon and co-workers, including the free-boundary case [91].

For the formulation of the variational principle, let $\mathcal{H}_0 \in \mathbb{R}$, and let V be the space of L^2 vector fields $u = B$ on Ω such $\operatorname{div} B = 0$ and $n \cdot B = 0$ on $\partial\Omega$. Then we consider the problem

$$\min\{\mathcal{S}(B) : \mathcal{H}(B) = \mathcal{H}_0\}, \tag{26}$$

with the entropy and Hamiltonian functions given by

$$\mathcal{S}(B) = \frac{1}{2}\int_\Omega |B|^2 dx, \quad \mathcal{H}(B) = \frac{1}{2}H_m(B) := \frac{1}{2}\int_\Omega A \cdot B dx, \tag{27}$$

where A is the vector potential for B defined as the unique solution of the problem

$$\begin{cases} \operatorname{curl} A = B, & \text{in } \Omega, \\ \operatorname{div} A = 0, & \text{in } \Omega, \\ n \times A = 0, & \text{on } \partial\Omega. \end{cases} \tag{28}$$

The necessary condition for solutions of (26) reads

$$\begin{cases} B - \lambda A = 0, \quad \mathcal{H}(B) = \mathcal{H}_0, \\ \text{with } A \text{ given by (28)}, \end{cases} \tag{29}$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier. The set of constrained critical points of the entropy function is

$$\mathfrak{C}_{\mathcal{H}_0} = \{B : (B, \lambda) \text{ solves (29) for some } \lambda \in \mathbb{R}\}.$$

From (28) and (29), one has that each constrained critical point $B \in \mathfrak{C}_{\mathcal{H}_0}$ is a linear Beltrami field since $\operatorname{curl} B = \lambda \operatorname{curl} A = \lambda B$, and the same holds true for the corresponding potential, $\operatorname{curl} A = \lambda A$.

The formal analysis of the variational problem proceeds as in the example of the reduced Euler equations with a linear profile $s'(y)$. Condition (29) implies that $B \in V$ satisfies the eigenvalue problem (25). From the theory of this eigenvalue problem [87], we know that there is an orthonormal basis of eigenfunctions corresponding to the eigenvalue $\lambda = \lambda_j \in \mathbb{R} \setminus \{0\}$ labeled by $j \in \mathbb{N}_0$ and with $k = 1, \dots, d_j$ counting the multiplicity. As in the case of the Euler equations, the set $\mathfrak{C}_{\mathcal{H}_0}$ comprises all and only the vector fields $B_j = \lambda_j A_j$ with $A_j = \sum_k a_{j,k} A_{j,k}$, $a_j = (a_{j,k})_k \in \mathbb{R}^{d_j}$, and that satisfy the constraint $\mathcal{H}(B) = \mathcal{H}_0$, which is $2\mathcal{H}_0\lambda_j = |a_j|^2$. Different from the case of the reduced Euler equations, both $\mathcal{H}_0$ and the eigenvalues of the curl operator can be negative. Since $|a_j|^2 > 0$, if $\mathcal{H}_0 < 0$ (resp. $\mathcal{H}_0 > 0$), only the eigenfunctions B_j with $\lambda_j < 0$ (resp. $\lambda_j > 0$) belong to $\mathfrak{C}_{\mathcal{H}_0}$. The entropy function restricted to $\mathfrak{C}_{\mathcal{H}_0}$ amounts to

$$\mathcal{S}(B_j) = \mathcal{H}_0\lambda_j > 0,$$

and it is always positive even if $\mathcal{H}_0$ and λ_j can be negative. It is possible to show that the minimum is attained [92]. Hence the solutions of (26) are the ground states for the curl operator.

The variational principle (26) asks for divergence-free fields B with minimum energy $(1/2)\|B\|^2_{L^2} = \mathcal{S}(B)$ subject to the constraint that magnetic helicity $H_m(B) = 2\mathcal{H}(B)$ is held constant. Since magnetic helicity is a global constraint, i.e. $\mathcal{H}(B)$ takes values in $\mathbb{R}$, λ is a constant in $\mathbb{R}$ and thus the variational principle selects *linear* Beltrami fields. Magnetic helicity H_m is related to the topology of the field lines of B [93]. Specifically magnetic helicity is the average asymptotic linking number of the field lines [94,95]. Nonlinear Beltrami fields (24) on the other hand can be obtained from a variational principle with a much stronger constraint, i.e., Beltrami fields are energy minima constrained to configurations of the field B that are continuous deformation of a prescribed field B_0. This constraint preserves the topology of the field lines, and thus magnetic helicity as well [96]. An overview of this variational principle is given in Appendix C for sake of completeness.

In this work, we discuss metriplectic dynamical systems on the space V of divergence-free fields. We address in particular convergence of an orbit to *nonlinear* Beltrami fields. This is an example for which complete relaxation of the orbit is not a desired property, since complete relaxation would leads to linear Beltrami fields. We shall address instead brackets that have a class of invariants much richer than just the Hamiltonian.

From a computational point of view, the direct numerical solution for Beltrami fields is possible by a variety of techniques [83,90,97]. Here, we shall not attempt to compare the performance of these methods with that of the metriplectic relaxation. Our aim is rather to study the convergence of metriplectic systems on a physically relevant problem. We note however that relaxation methods for force-free equilibria are common in several applications [31] and our study eventually aims at improving the rate of convergence of such methods.

3. On the relaxation of metriplectic systems

In this section, we present some remarks and mathematical results pertaining to equilibrium points of metriplectic systems, their stability, and sufficient conditions for convergence of nearby orbits.

As described above, metriplectic dynamical systems dissipate entropy at constant energy, cf. (8). It is therefore central to ask whether a solution $u(t)$ of a metriplectic system, defined for $t \in [0, +\infty)$, has a limit for $t \to +\infty$, and whether the limit, when it exists, is a minimum of entropy on the surface of constant Hamiltonian $\mathcal{H}(u) = \mathcal{H}_0 = \mathcal{H}(u_0)$, $u_0 = u(0)$ being the initial state. This is the variational principle (1). For some applications, e.g., the equilibrium problems introduced in Sections 2.2.1 and 2.2.2, complete relaxation of the orbit is essential. Recall, this means that the solution of the metriplectic system has a limit as $t \to +\infty$ and the limit is a solution of (1). In general for applications it is also useful to know the rate of convergence to the limit.

In this section we address implications of general metriplectic structure, i.e., properties (8), for the complete relaxation of the orbit. A first implication is that, since $\mathcal{S}$ is monotonically nonincreasing along an orbit, it is a candidate for a Lyapunov function [40,98,99]. This argument is standard for nondegenerate gradient flows, but adaptation is needed for metriplectic systems, which have degeneracy. In Section 3.1 we first recall some standard arguments for Lyapunov stability, and then discuss their adaptation to metriplectic systems. Then, in Section 3.2, we consider another classical tool in the theory of nondegenerate gradient flows, the Polyak–Łojasiewicz condition. We develop the details in the finite-dimensional setting, but make some comments on extension to infinite dimensions in Section 3.3.

3.1. Finite-dimensional systems: Lyapunov stability

Let us start by recalling the Lyapunov stability theorem for finite-dimensional dynamical systems. Consider a generic vector field $X : \mathcal{Z} \to \mathbb{R}^n$ on a domain $\mathcal{Z} \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$. We only assume that X is locally Lipschitz continuous (locally Lipschitz for short), that is, every point $z_0 \in \mathcal{Z}$ has a neighborhood $\mathcal{U}_{z_0}$ where

$$|X(z) - X(z')| \leq L_{z_0}|z - z'|, \quad z, z' \in \mathcal{U}_{z_0},$$

for a constant $L_{z_0} > 0$, possibly depending on z_0, thus we have existence and uniqueness for the ordinary differential equation system $dz/dt = X(z)$.

Let $z_* \in \mathcal{Z}$ be an equilibrium point, i.e., $X(z_*) = 0$. A continuous function $\mathcal{L} : \mathcal{O} \to \mathbb{R}$ defined on an open subset $\mathcal{O} \subseteq \mathcal{Z}$ containing z_*, and differentiable in $\mathcal{O} \setminus \{z_*\}$ is a Lyapunov function for X, if

$$X(z) \cdot \nabla\mathcal{L}(z) \leq 0, \quad z \in \mathcal{O} \setminus \{z_*\}; \tag{L1}$$

$$\mathcal{L}(z_*) = 0 \text{ and } \mathcal{L}(z) > 0, \quad z \neq z_*. \tag{L2}$$

Such an $\mathcal{L}$ is a *strict* Lyapunov function for X, if it is a Lyapunov function and satisfies

$$X(z) \cdot \nabla\mathcal{L}(z) < 0, \quad z \in \mathcal{O} \setminus \{z_*\}. \tag{L3}$$

Condition (L2) in particular implies that z_* is an *isolated* minimum of $\mathcal{L}$ in $\mathcal{O}$.

The Lyapunov stability theorem [40,66,100,101] states that, if a Lyapunov function exists, then $z_* \in \mathcal{O}$ is a *stable* equilibrium point. By definition this means that for any $\varepsilon > 0$ there is $\delta > 0$, such that $|z_0 - z_*| < \delta$ implies $|z(t) - z_*| < \varepsilon$ for all $t \geq 0$, with $z(t)$ being an integral curve of X with initial condition $z(0) = z_0$. In addition, if the Lyapunov function is strict, z_* is an *asymptotically stable* equilibrium point, that is, z_* is a stable equilibrium point, in the sense defined above, and $\delta > 0$ can be chosen so that $\lim_{t\to+\infty} z(t) = z_*$, for all initial conditions z_0 satisfying $|z_0 - z_*| < \delta$.

Although the Lyapunov stability theorem is well known [40,66,100,101], we recall the proof for sake of completeness, since the other results in Section 3 rely on the same ideas. The various arguments available in the literature differ essentially only in the final step in the proof of asymptotic stability. Here we follow Moretti [101], which we find particularly clear. Recall that $B_r(z) = \{z' \in \mathbb{R}^n : |z' - z| < r\}$ denotes the open ball of radius $r > 0$ in $\mathbb{R}^n$.

Theorem 1 (Lyapunov stability). *Let $X : \mathcal{Z} \to \mathbb{R}^n$ be a locally Lipschitz vector field, $z_* \in \mathcal{Z}$ an equilibrium point of X, $\mathcal{O} \subseteq \mathcal{Z}$ an open subset containing z_*, and $\mathcal{L} : \mathcal{O} \to \mathbb{R}$ continuous in $\mathcal{O}$ and differentiable in $\mathcal{O} \setminus \{z_*\}$.*

(i) If $\mathcal{L}$ is a Lyapunov function for X, then z_* is a stable equilibrium point.
(ii) If $\mathcal{L}$ is a strict Lyapunov function for X, then z_* is an asymptotically stable equilibrium point.

Proof. (i) Stability. Since $\mathcal{O}$ is open, we can choose $\varepsilon > 0$ so small that the ball $B_\varepsilon(z_*)$ is contained in the neighborhood $\mathcal{O}$. On the boundary ∂B_ε, $\mathcal{L}(z) > 0$ because of (L2). Let $\mathcal{L}_\varepsilon := \min_{z\in\partial B_\varepsilon(z_*)} \mathcal{L}(z)$. The minimum exists since $\partial B_\varepsilon(z_*)$ is compact. Let us now choose $\delta > 0$ so small that $\mathcal{L}(z) < \mathcal{L}_\varepsilon$ for $z \in B_\delta(z_*)$. This is possible since $\mathcal{L}$ is continuous and $\mathcal{L}(z_*) = 0$. For any $z_0 \in B_\delta(z_*)$, let $z(t)$, $t \in [-\tau_0, +\tau_0]$ be an integral curve of the vector field X with initial condition $z(0) = z_0$. Assumption (L1) implies that,

$$\mathcal{L}\big(z(t)\big) \leq \mathcal{L}(z_0) < \mathcal{L}_\varepsilon, \text{ for all } t \in [0, \tau_0].$$

The function $t \mapsto |z(t) - z_*|$ is continuous; hence, if there is a time $t_o \in [0, \tau_0]$ at which $|z(t_o) - z_*| \geq \varepsilon$, the intermediate value theorem implies that there is $t_e \in (0, t_o]$ at which $|z(t_e) - z_*| = \varepsilon$ and thus $\mathcal{L}\big(z(t_e)\big) \geq \mathcal{L}_\varepsilon$, and this is a contradiction. Therefore $|z(t) - z_*| < \varepsilon$ for all $t \in [0, \tau_0]$. Since the integral curve stays in a bounded subdomain, the solution can be extended to the interval $[-\tau_0, +\infty)$ and $|z(t) - z_*| < \varepsilon$ for all $t \in [0, +\infty)$.

(ii) Asymptotic stability. If $\mathcal{L}$ is a strict Lyapunov function, then in particular, it is a Lyapunov function and thus z_* is a stable equilibrium point. We can choose $\varepsilon_0 > 0$ and a $\delta_0 > 0$ such that any integral curve $z(t)$ with $z(0) \in B_{\delta_0}(z_*)$ stays in $B_{\varepsilon_0}(z_*)$.

We want to show that for any $z_0 \in B_{\delta_0}(z_*)$ and for any $\varepsilon \in (0, \varepsilon_0)$ there is a time $T_\varepsilon > 0$ such that the integral curve $z(t)$ with initial condition $z(0) = z_0$ satisfies $|z(t) - z_*| < \varepsilon$ for all $t > T_\varepsilon$.

Uniqueness of the orbit passing through a given point implies that, if $z_0 \neq z_*$, then $|z(t) - z_*| > 0$ (since $z(t) = z_*$ is a solution) and thus $d\mathcal{L}\big(z(t)\big)/dt < 0$ for all $t > 0$. Therefore function $t \mapsto \mathcal{L}\big(z(t)\big)$ is strictly monotonically decreasing, and it is bounded form below, hence it has a limit

$$\mathcal{L}\big(z(t)\big) \to \ell = \inf_{t\geq 0} \mathcal{L}\big(z(t)\big) \geq 0,$$

The limit must be $\ell = 0$. If not, $\mathcal{L}\big(z(t)\big) \geq \ell > 0$ for all $t > 0$, and continuity of $\mathcal{L}$ implies that there a radius $r \in (0, \epsilon_0)$ such that $|z(t) - z_*| > r$. This leads to a contradiction, since, if $z(t)$ stays in the compact region $\mathscr{R} = \{z : r \leq |z - z_*| \leq \epsilon\}$ for all $t \geq 0$, then with $-M = \max_{z\in\mathscr{R}} X(z) \cdot \nabla\mathcal{L}(z) < 0$,

$$\mathcal{L}\big(z(t)\big) = \mathcal{L}(z_0) + \int_0^t X\big(z(s)\big) \cdot \nabla\mathcal{L}\big(z(s)\big) ds \leq \mathcal{L}(z_0) - Mt.$$

For $t > \mathcal{L}(z_0)/M > 0$, we have $\mathcal{L}\big(z(t)\big) < 0$, which is impossible. Hence, the limit must be $\ell = 0$, that is, for every $\lambda > 0$ there is $T_\lambda > 0$ such that $\mathcal{L}\big(z(t)\big) < \lambda$ for $t > T_\lambda$.

We claim that for any $\epsilon > 0$ we can find $\lambda = \lambda_\epsilon > 0$ such that

$$A_\lambda := \{z \in B_{\epsilon_0}(z_*) : \mathcal{L}(z) < \lambda\},$$

is contained in the ball $B_\epsilon(z_*)$, i.e. $A_\lambda \subset B_\epsilon(z_*)$. If this is the case, corresponding to λ_ϵ, there is a time $T_\epsilon > 0$ such that, for all $t > T_\epsilon$, $z(t) \in A_{\lambda_\epsilon} \subset B_\epsilon(z_*)$, which is the thesis. Therefore it remains to prove the claim. Let us assume that the claim is false, i.e. there is a value $\epsilon_* > 0$ such that, for all λ, there is at least one point $\tilde{z}_\lambda$ that satisfies the conditions $\mathcal{L}(\tilde{z}_\lambda) < \lambda$ and $\epsilon_* \leq |\tilde{z}_\lambda - z_*| \leq \epsilon_0$. Upon choosing $\lambda = 1/n$ for $n \in \mathbb{N}$ we obtain a sequence $z_n = \tilde{z}_{1/n}$, which belongs to a compact set. Hence there is a converging subsequence $z_{n_k} \to \tilde{z}_*$, and $\epsilon_* \leq |\tilde{z}_* - z_*| \leq \epsilon_0$ so that necessarily $\mathcal{L}(\tilde{z}_*) > 0$. On the other hand, we have $\mathcal{L}(z_{n_k}) < 1/n_k \to 0$ and, by continuity of $\mathcal{L}$, $\mathcal{L}(\tilde{z}_*) = 0$, which is a contradiction. □

We want to apply Theorem 1 to metriplectic vector fields of the following form:

$$X(z) = J(z)\nabla\mathcal{H}(z) - K(z)\nabla\mathcal{S}(z). \tag{30}$$

For comparison, we also address the case of a standard nondegenerate gradient flow

$$X(z) = -\nabla\mathcal{S}(z), \tag{31}$$

with entropy $\mathcal{S} \in C^2(\mathcal{Z})$.

First, recall that in the case of nondegenerate gradient flows (31), with $\mathcal{S} \in C^2(\mathcal{Z})$, the function $\mathcal{L}(z) = \mathcal{S}(z) - \mathcal{S}(z_*)$ is a strict Lyapunov function in a neighborhood of any *isolated*, local minimum z_* of $\mathcal{S}$ [100, Proposition 15.0.2]. Therefore, the Lyapunov stability theorem implies that any integral curve of the gradient flow with initial condition near an isolated entropy minimum converges to that entropy minimum for $t \to +\infty$ (asymptotic stability). This result is a *local version* of the property we called complete relaxation.

Now, consider the case of a metriplectic vector field (30). If $z_* \in \mathcal{Z}$ is an equilibrium point, the function $\mathcal{L}(z) = \mathcal{S}(z) - \mathcal{S}(z_*)$ satisfies condition (L1), because of the general properties of metriplectic systems, cf. (8). If the entropy function $\mathcal{S}$ has an isolated minimum at $z_* \in \mathcal{Z}$, $\mathcal{L}(z)$ is a Lyapunov function of the system in a neighborhood $\mathcal{O}$ of z_*. Therefore z_* is a stable equilibrium. However, an orbit $z(t)$ can converge to z_* only if the initial condition $z_0 = z(0)$ and the local entropy minimum z_* belong to the same energy isosurface, i.e., $\mathcal{H}(z_0) = \mathcal{H}(z_*)$. This is a necessary condition that follows from the continuity of $\mathcal{H}$: if $z(t) \to z_*$, passing to the limit in $\mathcal{H}(z_0) = \mathcal{H}(z(t))$ yields $\mathcal{H}(z_0) = \mathcal{H}(z_*)$. This observation leads to the following conclusion:

Proposition 1. *Let $X : \mathcal{Z} \to \mathbb{R}^n$ be a locally Lipschitz vector field on a domain $\mathcal{Z} \subseteq \mathbb{R}^n$, $z_* \in \mathcal{Z}$ an equilibrium point of X, and $\mathcal{H}, \mathcal{S} \in C^1(\mathcal{Z})$ such that*

1. *$X(z) \cdot \nabla\mathcal{H}(z) = 0$ and $X(z) \cdot \nabla\mathcal{S}(z) \leq 0$, $z \in \mathcal{Z}$;*
2. *$\nabla\mathcal{H}(z_*) \neq 0$.*

Then, if $\mathcal{L} = \mathcal{S} - \mathcal{S}(z_*)$ satisfies (L2) in a neighborhood $\mathcal{O} \subseteq \mathcal{Z}$ of z_*, there is at least one point $z' \in \mathcal{O} \setminus \{z_*\}$ such that $X(z') \cdot \nabla\mathcal{S}(z') = 0$.

Proof. By contradiction, let us assume that $X \cdot \nabla\mathcal{S} < 0$ in $\mathcal{O} \setminus \{z_*\}$. Then hypothesis 1 implies that $\mathcal{L}(z) = \mathcal{S}(z) - \mathcal{S}(z_*)$ satisfies the conditions for a strict Lyapunov function for X. For the Lyapunov stability theorem, there exists $\delta' > 0$ such that for any z_0 with $|z_0 - z_*| < \delta'$ the orbit $z(t)$ of the dynamical system $dz/dt = X(z)$ with initial condition $z(0) = z_0$ exists for all $t \geq 0$ and $z(t) \to z_*$ as $t \to +\infty$. Hypothesis 1 also implies that $\mathcal{H}$ is a constant of motion and it is continuous, therefore $\mathcal{H}(z_0) = \mathcal{H}(z_*)$.

From hypothesis 2 and the continuity of the derivative $\nabla\mathcal{H}$ we can find a ball of radius $\delta'' > 0$ around z_* where $\nabla\mathcal{H} \neq 0$.

We choose $\delta < \min\{\delta', \delta''\}$. In the ball of radius δ centered at z_* there is at least one point z_0 such that $\mathcal{H}(z_0) \neq \mathcal{H}(z_*)$. If not, then $\mathcal{H}$ is constant in the ball, and this is not possible since $\nabla\mathcal{H} \neq 0$. On the other hand, since $|z_0 - z_*| < \delta < \delta'$ we must have $\mathcal{H}(z_0) = \mathcal{H}(z_*)$, which is a contradiction. □

For metriplectic systems, hypothesis 1 is verified, cf. (8). Hypothesis 2 holds away from critical points of $\mathcal{H}$, which are usually isolated. Therefore, this proposition implies that the entropy $\mathcal{S}$ of a metriplectic system in most cases (specifically under hypothesis 2) cannot be used as a strict Lyapunov function, since it is not strictly decaying everywhere in $\mathcal{O} \setminus \{z_*\}$. Without a strict Lyapunov function, asymptotic stability of local entropy minima does not follows directly from the usual Lyapunov theorem. This is in stark contrast with the case of pure gradient flows (31) discussed above.

Asymptotic stability with Lyapunov functions that are not strictly dissipated have been addressed by De Salle and Lefschetz [102], who showed that any integral curve, defined for $t \geq 0$, in a bounded strict sublevel set of the Lyapunov function must approach the largest invariant set contained in the region where $X(z) \cdot \nabla \mathcal{L}(z) = 0$. More specific results for systems with a conserved energy and a dissipated entropy were considered by Beretta [103] with applications to quantum thermodynamics. Here we apply similar ideas to metriplectic systems. Suppose a locally Lipschitz vector field $X : \mathcal{Z} \to \mathbb{R}^n$ on a domain $\mathcal{Z} \subseteq \mathbb{R}^n$ has k constants of motion $\mathcal{I}^1, \ldots, \mathcal{I}^k \in C^\infty(\mathcal{Z})$, for $1 \leq k < n$, which means $X(z) \cdot \nabla \mathcal{I}^\alpha(z) = 0$, for $\alpha \in \{1, \ldots, k\}$. The functions $\mathcal{I}^\alpha$ are independent at $z \in \mathcal{Z}$ if

$$\nabla \mathcal{I}^1(z), \ldots, \nabla \mathcal{I}^k(z) \text{ are linearly independent in } \mathbb{R}^k. \tag{32}$$

It is convenient to define the function $\mathcal{I} := \left(\mathcal{I}^1, \ldots, \mathcal{I}^k\right) \in C^\infty(\mathcal{Z}, \mathbb{R}^k)$. Then (32) is equivalent to $\operatorname{rank} \nabla \mathcal{I}(z) = k$.

If the constants of motion are independent at an equilibrium point $z_* \in \mathcal{Z}$, where $X(z_*) = 0$, then the local submersion theorem [66, Theorem 2.5.13] allows us to find a neighborhood $\mathcal{U}$ of z_* where the level sets

$$\mathcal{U}_\eta := \{z \in \mathcal{U} : \mathcal{I}(z) = \eta\},$$

are closed submanifolds of $\mathcal{U}$, for η is a neighborhood of $\eta_* := \mathcal{I}(z_*)$. Since $\mathcal{I}^\alpha$ are constants of motion, X is tangent to $\mathcal{U}_\eta$, so that the dynamical system can be reduced locally to each submanifold $\mathcal{U}_\eta$. This leads to the following straightforward result, which is essentially a finite-dimensional version of Theorem 2 in [103].

Proposition 2. *Let $X : \mathcal{Z} \to \mathbb{R}^n$ be a locally Lipschitz vector field on a domain $\mathcal{Z} \subseteq \mathbb{R}^n$, $\mathcal{I} = (\mathcal{I}^1, \ldots, \mathcal{I}^k) \in C^\infty(\mathcal{Z}, \mathbb{R}^k)$ be constants of motion with $1 \leq k < n$, $z_* \in \mathcal{Z}$ be an equilibrium point of X, $\operatorname{rank} \nabla \mathcal{I}(z_*) = k$, $\eta_* := \mathcal{I}(z_*)$, and let $\mathcal{U}_{\eta_*} = \{z \in \mathcal{U} : \mathcal{I}(z) = \eta_*\}$, where $\mathcal{U}$ is the neighborhood of z_* given by the local submersion theorem. If there is $\mathcal{L} \in C^1(\mathcal{U})$ satisfying*

$$X(z) \cdot \nabla \mathcal{L}(z) \leq 0, \qquad z \in \mathcal{U}, \tag{L1$'$}$$

$$\mathcal{L}(z_*) = 0 \text{ and } \mathcal{L}(z) > 0, \qquad z \in \mathcal{U}_{\eta_*} \setminus \{z_*\}, \tag{L2$'$}$$

then for any sufficiently small $\varepsilon > 0$, there is a $\delta > 0$ such that the integral curve $z(t)$ of X with initial condition $z(0) = z_0 \in B_\delta(z_*) \cap \mathcal{U}_{\eta_*}$ is defined for all $t \geq 0$, and $z(t) \in B_\varepsilon(z_*) \cap \mathcal{U}_{\eta_*}$. If in addition

$$X(z) \cdot \nabla \mathcal{L}(z) < 0, \quad z \in \mathcal{U}_{\eta_*} \setminus \{z_*\}, \tag{L3$'$}$$

then $z(t) \to z_*$ for $t \to +\infty$.

Proof. Since $\operatorname{rank} \nabla \mathcal{I}(z_*) = k$, the local submersion theorem [66] allows us to find open subsets $\mathcal{N} \subseteq \mathcal{I}(\mathcal{Z})$ and $\mathcal{Z}' \subseteq \mathbb{R}^{n-k} \cong \ker \nabla \mathcal{I}(z_*)$, together with local coordinates $(\eta, \zeta) \in \mathcal{N} \times \mathcal{Z}'$, defined in an open, connected subset $\mathcal{U} \subseteq \mathcal{Z}$ containing the point z_*, such that the inverse coordinate map $\varphi : \mathcal{N} \times \mathcal{Z}' \to \mathcal{U}$ is a C^∞-diffeomorphism and satisfies

$$\mathcal{I}\big(\varphi(\eta, \zeta)\big) = \eta.$$

Therefore, the sets $\mathcal{U}_\eta = \{z \in \mathcal{U} : \mathcal{I}(z) = \eta\}$, with $\eta \in \mathcal{N}$, are closed submanifolds parameterized by $\zeta \mapsto \varphi(\eta, \zeta)$. Let us denote by $(\eta_*, \zeta_*) \in \mathcal{N} \times \mathcal{Z}'$ the point corresponding to $z_* = \varphi(\eta_*, \zeta_*)$.

In this local coordinate system the integral curves of the vector field X in $\mathcal{U}$ solve

$$\frac{d\eta}{dt} = X \cdot \nabla \eta = 0, \quad \frac{d\zeta}{dt} = X_\eta(\zeta) := (X \cdot \nabla \zeta) \circ \varphi(\eta, \zeta),$$

since $X \cdot \nabla \eta = X \cdot \nabla \mathcal{I} = 0$. The field X_η defines a tangent vector on $\mathcal{U}_\eta$ for any $\eta \in \mathcal{N}$. Since $\nabla \zeta$ is C^∞, one can check that X_η is a locally Lipschitz continuous function of ζ. Then, with $\mathcal{L}_\eta := \mathcal{L} \circ \varphi(\eta, \cdot)$, one has

$$(X \cdot \nabla \mathcal{L}) \circ \varphi = X_\eta^a \frac{\partial \mathcal{L}_\eta}{\partial \zeta^a},$$

where the sum over $a \in \{1, \ldots, n-k\}$ is implied. Hence hypotheses (L1$'$)–(L3$'$) are equivalent to Lyapunov conditions (L1)–(L3) for the function $\mathcal{L}_{\eta_*}$ and for the dynamical system $d\zeta/dt = X_{\eta_*}$ on $\mathcal{U}_{\eta_*}$, and the claim follows from Theorem 1. $\square$

Proposition 2 establishes stability and asymptotic stability for orbits with initial conditions in $\mathcal{U}_{\eta_*}$, which being a lower dimensional set has zero Lebesgue measure in the phase space. This issue was addressed in [103] by assuming z_* is part of a continuous family of equilibria.

At last, we address metriplectic vector fields. We know that there is at least one constant of motion, the Hamiltonian $\mathcal{H}$. More generally, let X be a metriplectic vector field with k constants of motion $\mathcal{I}^1, \ldots, \mathcal{I}^k$, $1 \leq k < n$, and for definiteness $\mathcal{I}^k = \mathcal{H}$. Unlike the case of Proposition 2, we assume that (32) holds in a whole subdomain $\mathcal{U}$. Under these conditions, the submersion theorem [66, Theorem 3.5.4] establishes that the set

$$\mathcal{U}_\eta := \{z \in \mathcal{U} : \mathcal{I}(z) = \eta\}, \tag{33}$$

is a closed submanifold of $\mathcal{U}$ for any $\eta \in \mathcal{I}(\mathcal{U})$, and the restriction of the entropy to such submanifolds, i.e., $\mathcal{S}|_{\mathcal{U}_\eta}$, is smooth.

We make the following hypothesis on the entropy in $\mathcal{U}$:

$$\mathcal{U} \text{ is bounded}, \overline{\mathcal{U}} \subset \mathcal{Z}, \ \exists\, z_m \in \mathcal{U}, \ \text{ where } \mathcal{S}(z_m) < \inf\{\mathcal{S}(z) : z \in \partial \mathcal{U}\}. \tag{34}$$

This ensures that there is a minimum of the entropy in the interior of the subdomain $\mathcal{U}$.

Lemma 1. *Let $X : \mathcal{Z} \to \mathbb{R}^d$ be a vector field on a domain $\mathcal{Z} \subseteq \mathbb{R}^d$, and let $\mathcal{U} \subset \mathcal{Z}$ and $\mathcal{S} \in C^1(\mathcal{Z})$ satisfy (34). If in the subdomain $\mathcal{U}$, X is locally Lipschitz and $X \cdot \nabla \mathcal{S} \leq 0$, then there is an open, non-empty subset $\mathcal{O} \subseteq \mathcal{U}$ such that, for any $z_0 \in \mathcal{O}$ the integral curve $z(t)$ of X with initial condition $z(0) = z_0$ can be prolonged to the interval $t \in [0, +\infty)$ and $z(t) \in \mathcal{O}$ for all $t \geq 0$.*

Proof. Per hypotheses $\mathcal{S} \in C^1(\mathcal{Z})$, therefore $\mathcal{S} \in C^1(\overline{\mathcal{U}})$, and we have $\mathcal{S}_b := \inf_{\partial\mathcal{U}} \mathcal{S} > \mathcal{S}(z_m) > -\infty$. The set

$$\mathcal{O} := \{z \in \mathcal{U} : \ \mathcal{S}(z) < \mathcal{S}_b\}$$

is non-empty, since $z_m \in \mathcal{O}$, and open, since $\mathcal{S}(z)$ is continuous. Let $z : (\tau_1, \tau_2) \to \mathcal{U}$, $\tau_1 < 0 < \tau_2$, be the maximal solution of the initial value problem

$$dz/dt = X(z), \quad z(0) = z_0 \in \mathcal{O}.$$

The solution exists given that X is locally Lipschitz continuous in $\mathcal{U}$. Since $X \cdot \nabla\mathcal{S} \leq 0$, the function $t \mapsto \mathcal{S}\big(z(t)\big)$ is C^1 and non-increasing, hence $\mathcal{S}\big(z(t)\big) \leq \mathcal{S}(z_0) < \mathcal{S}_b$, and thus $z(t) \in \mathcal{O}$ for all $t \in [0, \tau_2)$.

We show that $\tau_2 = +\infty$. With this aim we rely on a standard argument from the theory of ordinary differential equations, which we report in full for sake of completeness. If τ_2 is finite, let $\{t_n\}_{n\in\mathbb{N}}$ be a sequence, $t_n \in (\tau_1, \tau_2)$, and $t_n \to \tau_2$ as $n \to +\infty$. Then

$$|z(t_n) - z(t_m)| \leq \max_{z\in\overline{\mathcal{U}}} |X(z)||t_n - t_m|,$$

for all $m, n \in \mathbb{N}$. Since $\{t_n\}$ is a convergent sequence, $\{z(t_n)\} \subset \mathcal{O}$ is a Cauchy sequence and must have a limit $\bar{z} \in \overline{\mathcal{O}}$ as $n \to \infty$. Passing to the limit in the inequality $\mathcal{S}\big(z(t_n)\big) \leq \mathcal{S}(z_0) < \mathcal{S}_b$ yields $\mathcal{S}(\bar{z}) \leq \mathcal{S}(z_0) < \mathcal{S}_b$, hence $\bar{z} \in \mathcal{O}$. We can use the limit point $\bar{z}$ as an initial condition, and we can extend the solution $z(t)$ beyond τ_2 contradicting the fact that $z(t)$ is the maximal solution. Therefore $\tau_2 = +\infty$ as claimed. □

The set $\mathcal{O}$ is given by

$$\mathcal{O} = \left\{z \in \mathcal{U} : \ \mathcal{S}(z) < \inf_{\partial\mathcal{U}} \mathcal{S}\right\}.$$

This lemma shows that, if $\mathcal{S}$ and $\mathcal{U}$ satisfy (34) and $\mathcal{S}$ is non-increasing along the integral curves of a vector field X, then we can find a positively invariant subset $\mathcal{O}$. This lemma can be applied directly to both standard gradient flows and metriplectic systems.

Proposition 3. *Let $X = J\nabla\mathcal{H} - K\nabla\mathcal{S}$ be a metriplectic vector field on a domain $\mathcal{Z} \subseteq \mathbb{R}^n$, with $1 \leq k < n$ constants of motion $\mathcal{I} = (\mathcal{I}^1, \dots, \mathcal{I}^k) \in C^\infty(\mathcal{Z}, \mathbb{R}^k)$, $\operatorname{rank} \nabla\mathcal{I} = k$ in a subdomain $\mathcal{U}$ satisfying (34), and let $\mathcal{O}$ be the open set given by Lemma 1. If in addition it holds that*

$$\begin{aligned}&\forall\eta \in \mathcal{I}(\mathcal{O}),\ \text{on } \mathcal{U}_\eta := \{z \in \mathcal{U} : \ \mathcal{I}(z) = \eta\},\ \mathcal{S}|_{\mathcal{U}_\eta} \text{ has a unique critical point } z_\eta,\\ &\text{which is a strict local minimum,}\end{aligned} \tag{L2''}$$

and

$$\ker K = \operatorname{span}\{\nabla\mathcal{I}^1, \dots, \nabla\mathcal{I}^k\} \ \text{ in } \ \overline{\mathcal{U}}. \tag{L3''}$$

then for any integral curve $z(t)$ of X with $z(0) = z_0 \in \mathcal{O}$, $\lim_{t\to+\infty} z(t) = z_\eta \in \mathcal{O}$, where $\eta = \mathcal{I}(z_0)$.

Proof. First we show that $z_\eta \in \mathcal{U}_\eta \cap \mathcal{O}$. Given that $\eta \in \mathcal{I}(\mathcal{O})$, the set $\mathcal{U}_\eta \cap \mathcal{O}$ is non-empty. If $z_\eta \notin \mathcal{U}_\eta \cap \mathcal{O}$, there cannot be any extremum of $\mathcal{S}|_{\mathcal{U}_\eta}$ in $\mathcal{U}_\eta \cap \mathcal{O}$ since z_η is the only critical point in $\mathcal{U}_\eta$. This implies that both the minimum and the maximum of $\mathcal{S}$ restricted to $\mathcal{U}_\eta \cap \overline{\mathcal{O}}$ are attained on the boundary $\mathcal{U}_\eta \cap \partial\mathcal{O}$, where we have $\mathcal{S} = \mathcal{S}_b = \inf\{\mathcal{S}(z) : \ z \in \partial\mathcal{U}\}$. We deduce that $\mathcal{S}|_{\mathcal{U}_\eta}$ is constant and equal to $\mathcal{S}_b$ in $\mathcal{U}_\eta \cap \overline{\mathcal{O}}$, but per definition, points in $\mathcal{O}$ satisfy $\mathcal{S} < \mathcal{S}_b$. Therefore, the only critical point z_η of $\mathcal{S}|_{\mathcal{U}_\eta}$ must be in $\mathcal{U}_\eta \cap \mathcal{O}$.

The function $\mathcal{S}|_{\mathcal{U}_\eta}$ is smooth and its critical points necessarily satisfy the Lagrange condition

$$\nabla\mathcal{S}(z) = \sum_{\alpha=1}^{k} \lambda_\alpha \nabla\mathcal{I}^\alpha(z), \quad \mathcal{I}(z) = \eta, \tag{35}$$

for $\lambda_\alpha \in \mathbb{R}$ [66, Theorem 3.5.27]. Hypothesis (L2″), states, in particular, that this can only happen at the point z_η.

Condition (L2″) also requires z_η to be a strict local minimum of $\mathcal{S}|_{\mathcal{U}_\eta}$. Then, z_η must be an equilibrium point of X. In order to see this, let us consider the integral line $\hat{z}_\eta(t) \in \mathcal{O}$ of the vector field X with initial condition $\hat{z}_\eta(0) = z_\eta$. Given that $z_\eta \in \mathcal{O}$, $\hat{z}_\eta(t)$ is defined for $t \geq 0$ (Lemma 1). Since $\mathcal{I}$ is a constant of motion, $\hat{z}_\eta(t) \in \mathcal{U}_\eta$, and continuity implies that for any $\varepsilon > 0$, there is $\delta > 0$ such that $|t| \leq \delta$ implies $|\hat{z}_\eta(t) - z_\eta| < \varepsilon$. If $\hat{z}_\eta$ is not identically equal to z_η when $t \in [0, \delta)$, then at some $t' \in (0, \delta)$, $\hat{z}_\eta(t') \neq z_\eta$ and $\mathcal{S}\big(\hat{z}_\eta(t')\big) > \mathcal{S}(z_\eta)$ since z_η is an entropy local minimum. On the other hand, dissipation of entropy requires $\mathcal{S}\big(\hat{z}_\eta(t')\big) \leq \mathcal{S}\big(\hat{z}_\eta(0)\big) = \mathcal{S}(z_\eta)$. Therefore, it must be $\hat{z}_\eta(t) = z_\eta$ identically for $t \in [0, \delta)$, which implies $X(z_\eta) = 0$ and z_η is an equilibrium.

Given $z_0 \in \mathcal{O}$, the integral curve of X with initial condition $z(0) = z_0$ exists for all $t \geq 0$ (Lemma 1). Let us define the function

$$\mathcal{L}(z) = \mathcal{S}(z) - \mathcal{S}(z_\eta),$$

where $\eta = \mathcal{I}(z_0)$. Then $\nabla\mathcal{L}(z) = \nabla\mathcal{S}(z)$ and

$$X(z) \cdot \nabla\mathcal{L}(z) = -\nabla\mathcal{S}(z) \cdot K(z)\mathcal{S}(z) \leq 0, \tag{36}$$

so that $\mathcal{L}$ satisfies condition (L1′) of Proposition 2 in the open set $\mathcal{O}$ defined above.

Since z_η is a local minimum of $\mathcal{S}|_{\mathcal{U}_\eta}$ there is a neighborhood of z_η on $\mathcal{U}_\eta$ where $\mathcal{S}(z) > \mathcal{S}(z_\eta)$ for $z \neq z_\eta$. But by definition of the submanifold topology, this neighborhood has the form $\mathcal{U}_\eta \cap \mathcal{V}$ with $\mathcal{V}$ an open subset of $\mathcal{U}$. Then, condition (L2′) holds true with $\mathcal{U}$ replaced by $\mathcal{V}$.

At last we show that condition (L3′) is also satisfied. If not, one could find a point $\tilde{z}_\eta \in \mathcal{U}_\eta$ such that $\tilde{z}_\eta \neq z_\eta$ and $\nabla\mathcal{S}(\tilde{z}_\eta) \in \ker K(\tilde{z}_\eta)$, but then assumption (L3″) implies that $\tilde{z}_\eta$ satisfies the Lagrange conditions (35) and this contradicts the uniqueness of the critical point z_η. Therefore, inequality (36) must be strict, which is condition (L3′).

In summary, the point $z_* = z_\eta$, $\eta = \mathcal{I}(z_0)$, and the function $\mathcal{L}(z) = \mathcal{S}(z) - \mathcal{S}(z_\eta)$ on the neighborhood $\mathcal{V}$ satisfy all the hypotheses of Proposition 2, including (L3′). Therefore, there is a $\delta > 0$ such that any integral curve of X with initial condition in $\mathcal{U}_\eta \cap B_\delta(z_\eta)$ converges to z_η as $t \to +\infty$.

We claim that, for any integral curve, $z(t) \in \mathcal{O}$, $t \geq 0$ and $z(0) = z_0 \in \mathcal{U}_\eta \cap \mathcal{O}$, there must be a time t_δ such that $z(t_\delta) \in \mathcal{U}_\eta \cap B_\delta(z_\eta)$. If this is the case, then necessarily $z(t) \to z_\eta$, which is the thesis. Therefore it remains to prove this last claim.

The sets $\overline{\mathcal{U}_\eta}$ and $\overline{\mathcal{U}_\eta} \setminus B_\delta(z_\eta)$ are compact in $\mathbb{R}^n$. In view of conditions (L2″) and (L3″), on the set $\overline{\mathcal{U}_\eta} \setminus B_\delta(z_\eta)$ we have $X \cdot \nabla\mathcal{S} < 0$. Let $-M := \sup\{X(z) \cdot \nabla\mathcal{S}(z) : z \in \overline{\mathcal{U}_\eta} \setminus B_\delta(z_\eta)\}$. Then $M > 0$ since the supremum is attained, and if the integral curve $z(t)$ is such that $z(t) \in \overline{\mathcal{U}_\eta} \setminus B_\delta(z_\eta)$ for all $t \geq 0$,

$$\mathcal{S}\big(z(t)\big) = \mathcal{S}\big(z_0\big) + \int_0^t (X \cdot \nabla\mathcal{S})\big(z(s)\big)ds \leq \mathcal{S}\big(z_0\big) - Mt.$$

This leads to a contradiction as in the proof of the classical Lyapunov stability theorem. Hence there must be a time t_δ at which $z(t_\delta) \in \mathcal{U}_\eta \cap B_\delta(z_\eta)$. □

Condition (L2″) ensures that (L2′) holds for the function $\mathcal{L}(z) = \mathcal{S}(z) - \mathcal{S}(z_\eta)$ on $\mathcal{U}_\eta$. The uniqueness of the critical point of $\mathcal{S}$ together with (L3″) ensures that entropy is strictly dissipated by requiring that the metric bracket generated by the tensor K is "specifically degenerate", by which we mean it preserves the invariants $\mathcal{I}^\alpha$ only. Recall we use the terminology "minimally degenerate" to mean that the only degeneracy is that associated with $\mathcal{H}$.

We remark that the limit point $z_\eta \in \mathcal{U}_\eta$ is the unique solution of the optimization problem

$$\min\{\mathcal{S}(z) : z \in \mathcal{O}, \ \mathcal{I}(z) = \mathcal{I}(z_0)\};$$

hence Proposition 3 establishes a local version of the desired complete-relaxation property for all orbits starting in the set $\mathcal{O}$.

3.2. *Finite-dimensional systems: the Polyak–Łojasiewicz inequality*

Nondegenerate gradient flows and gradient descent methods have been studied under the assumptions that the entropy function $\mathcal{S}$ satisfies the classical Polyak–Łojasiewicz (PL) condition [104,105] and $\nabla\mathcal{S}$ is Lipschitz continuous, the latter being a natural hypothesis.

A function $\mathcal{S} \in C^1(\mathcal{Z})$ satisfies the PL condition in a non-empty subset $Q \subseteq \mathcal{Z}$ if $\mathcal{S}_* := \inf\{\mathcal{S}(z) : z \in Q\} > -\infty$ and there exists a constant $\kappa > 0$ such that

$$\frac{1}{\kappa}|\nabla\mathcal{S}(z)|^2 \geq \mathcal{S}(z) - \mathcal{S}_*, \quad z \in Q. \tag{PL}$$

If $\mathcal{S}$ satisfies the PL inequality, the associated gradient flow has the following properties:

- If $z(t) \in Q$, $t \in [0, +\infty)$, is an integral curve of the gradient flow (31), then $\mathcal{S}\big(z(t)\big)$ converges exponentially to $\mathcal{S}_*$ as $t \to +\infty$. This follows from
$$\frac{d}{dt}\big[\mathcal{S}\big(z(t)\big) - \mathcal{S}_*\big] = -|\nabla\mathcal{S}\big(z(t)\big)|^2 \leq -\kappa\big[\mathcal{S}\big(z(t)\big) - \mathcal{S}_*\big],$$
which implies
$$\big[\mathcal{S}\big(z(t)\big) - \mathcal{S}_*\big] \leq \big[\mathcal{S}\big(z_0\big) - \mathcal{S}_*\big]e^{-\kappa t},$$
where $z_0 \in Q$ is the initial condition at $t = 0$. The exponential rate of convergence is given by the constant κ in the inequality. Unfortunately convergence of the entropy values alone does not directly imply convergence of the orbit $z(t)$ to a limit for $t \to +\infty$ (cf. the counterexample of Palis and de Melo [106]).
- All equilibrium points of the gradient flow (31) are global minima of $\mathcal{S}$ (not just critical points). In fact, if $z_e \in Q$ is an equilibrium point, $\nabla\mathcal{S}(z_e) = 0$ and the PL condition implies $0 = |\nabla\mathcal{S}(z_e)|^2 \geq \kappa\big(\mathcal{S}(z_e) - \mathcal{S}_*\big) \geq 0$, hence $\mathcal{S}(z_e) = \mathcal{S}_*$.

Any strongly convex function satisfies condition (PL), cf. the short proof reported below, and thus are included as a special case. In general, (PL) is weaker than strong convexity. In fact, it has been shown that the classical PL condition is weaker than several other conditions introduced in order to address the convergence of gradient descent methods [107]. In addition, no assumption is made on the global minimum of $\mathcal{S}$, which, in particular, does not need to be an isolated point. For instance, the function $\mathcal{S}(z) = (|z|^2 - 1)^2$ satisfies the PL condition with $\kappa = 16r_0^2$ in the domain $S = \{z : |z| > r_0\}$ for any $r_0 \in (0, 1/2)$, and attains its minimum on the sphere $|z| = 1$; hence there is no isolated minimum. On the other hand, this function does not satisfy the PL condition on the whole space $\mathbb{R}^n$, because of the critical point at $z = 0$, which is a local maximum.

Polyak [104] under the additional (natural) hypothesis that the vector field $X = -\nabla\mathcal{S}$ is Lipschitz continuous (*not* just locally Lipschitz), established convergence of the gradient flow trajectories to the global entropy minimum. Specifically, Polyak's result in the notation used here amounts to the following.

Theorem 2 (Polyak 1963, Theorem 9 [104]). *Let $z_0 \in \mathcal{Z}$, $\rho, \kappa, L > 0$, and $\mathcal{S} \in C^1(\mathcal{Z})$ be such that (PL) is satisfied and $\nabla\mathcal{S}$ is Lipschitz continuous in the closed ball $\overline{B_\rho(z_0)} \subset \mathcal{Z}$ with Lipschitz constant L, and $\gamma = \sqrt{8L\varphi_0}/(\rho\kappa) \le 1$, where $\varphi_0 = \mathcal{S}(z_0) - \mathcal{S}_*$. Then there is $z_* \in \overline{B_{\gamma\rho}(z_0)}$ such that the integral curve $z(t)$ of the gradient flow (31) with $z(0) = z_0$ is defined for all $t \ge 0$ and $\left|z(t) - z_*\right| \le \gamma\rho e^{-\kappa t/2}$.*

We shall now give the details concerning the Polyak–Łojasiewicz condition. Although well-known, we recall for convenience of the reader the short proof of the fact that a strongly convex entropy satisfies inequality (PL).

Proof: Strongly convex imply PL. A strongly convex function $\mathcal{S} \in C^1(\mathcal{Z})$, $\mathcal{Z} \subseteq \mathbb{R}^n$ with parameter $\alpha > 0$ satisfies

$$\mathcal{S}(z') \ge \mathcal{S}(z) + (z' - z) \cdot \nabla\mathcal{S}(z) + \frac{\alpha}{2}|z' - z|^2,$$

for any $z, z' \in \mathcal{Z}$. The right-hand side is bounded from below by its infimum over z', which is attained at $z' = z - \alpha^{-1}\nabla\mathcal{S}(z)$ and thus

$$\mathcal{S}(z') \ge \mathcal{S}(z) - \frac{1}{2\alpha}|\nabla\mathcal{S}(z)|^2.$$

This inequality holds for any z', and thus implies that $\mathcal{S}_* = \inf\{\mathcal{S}(z) : z \in \mathcal{Z}\} > -\infty$. Taking the infimum over z' yields

$$\mathcal{S}_* \ge \mathcal{S}(z) - \frac{1}{2\alpha}|\nabla\mathcal{S}(z)|^2.$$

This can be rearranged to give the PL inequality with constant $\kappa = 2\alpha$. $\square$

In Polyak's original formulation of the theorem, the size ρ of the ball is determined by the condition $\gamma \le 1$ in terms of the entropy at the initial position z_0. For instance, $\mathcal{S}(z) = z^2/2$ satisfies the hypothesis with $\kappa = 2$ and $L = 1$; given any $z_0 \in \mathbb{R}^n$, $\gamma \le 1$ is equivalent to $\rho \ge |z_0|$, hence $\overline{B_\rho(z_0)}$ is large enough to contain the unique global minimum $z_* = 0$. We give a slightly different statement of Polyak's result.

Proposition 4. *Let $\mathcal{S} \in C^1(\mathcal{Z})$, $X = -\nabla\mathcal{S}$ Lipschitz continuous in a subdomain $\mathcal{U} \subset \mathcal{Z}$ satisfying (34), and let $\mathcal{O}$ be the open, non-empty subset given by Lemma 1. If $\mathcal{S}$ satisfies (PL) in $\mathcal{U}$, then for any integral curve $z(t)$ of X with initial condition $z(0) = z_0 \in \mathcal{O}$ there is $z_* \in \mathcal{O}$ and a constant $\theta > 0$ depending on z_0 such that $\mathcal{S}(z_*) = \mathcal{S}_*$,*

$$|z(t) - z_*| \le \theta e^{-\kappa t/2}, \quad \left[\mathcal{S}\big(z(t)\big) - \mathcal{S}_*\right] \le \left[\mathcal{S}\big(z_0\big) - \mathcal{S}_*\right]e^{-\kappa t}, \quad \textit{for } t \ge 0.$$

Next we give a proof of Proposition 4, which is based on the original argument due to Polyak [104]. The only difference consists in the use of Lemma 1 to establish the existence of a global solution. We give the proof in details since similar ideas are then needed below for the metriplectic case.

Proof of Proposition 4. Lemma 1 ensures that for any $z_0 \in \mathcal{O}$ there is an integral curve $z(t) \in \mathcal{O}$ of the gradient flow $X = -\nabla\mathcal{S}$ passing through $z(0) = z_0$ and this is defined for all $t \ge 0$. As a consequence of (PL), $\varphi(t) = \mathcal{S}\big(z(t)\big) - \mathcal{S}_*$ decays to zero exponentially, i.e., $\varphi(t) \le \varphi_0 e^{-\kappa t}$, with $\varphi_0 = \varphi(0)$.

If there is a finite time $0 \le \bar{t} < +\infty$ at which $\varphi(\bar{t}) = 0$, then $z(\bar{t})$ is a minimum for the entropy $\mathcal{S}$ and thus $\nabla\mathcal{S}\big(z(\bar{t})\big) = 0$ so that $z(\bar{t})$ is an equilibrium point. We deduce $z(t) = z(\bar{t}) = z_0$ for all $t \ge 0$ and the thesis holds true with $z_* = z_0$.

As for the non-trivial case $\varphi(t) > 0$ for all $t \ge 0$, we distinguish two key steps.

Step 1. For any $t_1, t_2 \ge 0$, $t_1 < t_2$, we have

$$\mathcal{S}\big(z(t_1)\big) - \mathcal{S}\big(z(t_2)\big) = \int_{t_1}^{t_2} |\nabla\mathcal{S}\big(z(t)\big)|^2 dt.$$

Following Polyak, we estimate the right-hand side from below. First the triangular inequality gives

$$\left|\nabla\mathcal{S}\big(z(t)\big)\right| \ge \left|\left|\nabla\mathcal{S}\big(z(t_1)\big)\right| - \left|\nabla\mathcal{S}\big(z(t)\big) - \nabla\mathcal{S}\big(z(t_1)\big)\right|\right|.$$

The second term on the right-hand side is bounded by

$$\left|\nabla\mathcal{S}\big(z(t)\big) - \nabla\mathcal{S}\big(z(t_1)\big)\right| \le L\left|z(t) - z(t_1)\right|,$$

where $L > 0$ is the Lipschitz constant of $\nabla\mathcal{S}$. In addition,

$$\begin{aligned}\frac{d}{dt}\left|z(t) - z(t_1)\right| &\le \left|\frac{d}{dt}\left|z(t) - z(t_1)\right|\right| \le \left|\nabla\mathcal{S}\big(z(t)\big)\right| \\ &\le \left|\nabla\mathcal{S}\big(z(t_1)\big)\right| + \left|\nabla\mathcal{S}\big(z(t)\big) - \nabla\mathcal{S}\big(z(t_1)\big)\right| \\ &\le \left|\nabla\mathcal{S}\big(z(t_1)\big)\right| + L\left|z(t) - z(t_1)\right|.\end{aligned}$$

Grönwall's inequality then gives

$$\left|z(t) - z(t_1)\right| \le \frac{1}{L}\left|\nabla\mathcal{S}\big(z(t_1)\big)\right|\left[e^{L(t-t_1)} - 1\right],$$

and thus

$$\left|\nabla\mathcal{S}\big(z(t)\big) - \nabla\mathcal{S}\big(z(t_1)\big)\right| \le \left|\nabla\mathcal{S}\big(z(t_1)\big)\right|\left[e^{L(t-t_1)} - 1\right].$$

If $L(t_2 - t_1) \le \log 2$, the term on the right-hand side is $\le \left|\nabla\mathcal{S}\big(z(t_1)\big)\right|$ and thus

$$\left|\nabla\mathcal{S}\big(z(t)\big)\right| \ge \left|\nabla\mathcal{S}\big(z(t_1)\big)\right|\left[2 - e^{L(t-t_1)}\right].$$

With $t_2 = t_1 + (\log 2)/L$ and t_1 arbitrary, we obtain

$$\mathcal{S}\big(z(t_1)\big) - \mathcal{S}\big(z(t_2)\big) \ge \frac{1}{\alpha L}\left|\nabla\mathcal{S}\big(z(t)\big)\right|^2,$$

where

$$\frac{1}{\alpha} := \int_0^{\log 2}\left[2 - e^s\right]^2 ds$$

is a positive numerical constant. At last, we use the fact that $\mathcal{S}\big(z(t)\big)$ is non-increasing and $t_1 < t_2$ so that

$$\begin{aligned}\mathcal{S}\big(z(t_1)\big) - \mathcal{S}\big(z(t_2)\big) &= \left[\mathcal{S}\big(z(t_1)\big) - \mathcal{S}_*\right] - \left[\mathcal{S}\big(z(t_2)\big) - \mathcal{S}_*\right]\\ &\le \left[\mathcal{S}\big(z(t_1)\big) - \mathcal{S}_*\right],\end{aligned}$$

and deduce that, for any $t_1 \ge 0$,

$$\left|\nabla\mathcal{S}\big(z(t_1)\big)\right|^2 \le \alpha L\left[\mathcal{S}\big(z(t_1)\big) - \mathcal{S}_*\right] = \alpha L\varphi(t_1). \tag{37}$$

We have established that, under the hypotheses, $\nabla\mathcal{S}$ along the orbit is controlled by the entropy decay.

Step 2. Given arbitrary points in time $t_1, t_2 \ge 0$, $t_1 < t_2$,

$$|z(t_1) - z(t_2)| \le \int_{t_1}^{t_2}\left|\nabla\mathcal{S}\big(z(t)\big)\right|dt,$$

and (37) together with $\varphi(t) \le \varphi_0 e^{-\kappa t}$ yields

$$|z(t_1) - z(t_2)| \le \frac{\sqrt{4\alpha L\varphi_0}}{\kappa}\left[e^{-\kappa t_1/2} - e^{-\kappa t_2/2}\right]. \tag{38}$$

This inequality can be used to show that for any sequence $\{t_n\}_{n\in\mathbb{N}}$ with $t_n \to +\infty$ as $n \to +\infty$, $z(t_n)$ is a Cauchy sequence on $\mathcal{O}$. Therefore there is a point $z_* \in \overline{\mathcal{O}}$ such that $z(t) \to z_*$ as $t \to +\infty$. Continuity of $\mathcal{S}$ implies $\mathcal{S}(z_*) = \mathcal{S}_*$ as claimed. At last, $z_* \in \mathcal{O}$, for, if not, then $z_* \in \partial\mathcal{O}$ and $\mathcal{S}(z_*) > \mathcal{S}(z_0) \ge \mathcal{S}(z_*)$, which is a contradiction (the first inequality is strict). Passing to the limit $t_2 \to +\infty$ in inequality (38) yields

$$|z(t_1) - z_*| \le \theta e^{-\kappa t_1/2},$$

with $\theta = \sqrt{4\alpha L\varphi_0}/\kappa$ and any $t_1 \ge 0$. □

We now generalize this result to the case of finite-dimensional metriplectic systems. Under the same conditions as in Proposition 3, a simple generalization of (PL) for metriplectic vector fields reads: for any $\eta \in \mathcal{I}(\mathcal{U})$, $\inf\{\mathcal{S}(z) :\ z \in \mathcal{U}_\eta\} =: \mathcal{S}_\eta > -\infty$ and there is a constant $\kappa_\eta > 0$, depending on η, such that

$$\frac{1}{\kappa_\eta}\nabla\mathcal{S}(z)\cdot K(z)\nabla\mathcal{S}(z) \ge \mathcal{S}(z) - \mathcal{S}_\eta, \quad z \in \mathcal{U}_\eta. \tag{PL$'$}$$

Differently from (PL), which is a condition on $\mathcal{S}$, inequality (PL$'$) involves both the entropy function and the bracket $(\mathcal{S},\mathcal{S}) = \nabla\mathcal{S}\cdot K\nabla\mathcal{S}$, which gives the entropy decay rate. Similarly to (PL), if inequality (PL$'$) holds, then the metriplectic vector field has the following properties:

- If $z(t) \in \mathcal{U}_\eta$, $t \in [0,+\infty)$, is an integral curve of X with initial condition $z_0 = z(0)$, then $\mathcal{S}\big(z(t)\big) \to \mathcal{S}_\eta$ as $t \to +\infty$, with exponential convergence. In fact, (PL$'$) implies

  $$\left[\mathcal{S}\big(z(t)\big) - \mathcal{S}_\eta\right] \le \left[\mathcal{S}\big(z_0\big) - \mathcal{S}_\eta\right]e^{-\kappa t},$$

  which follows as in the case of standard gradient flows.
- If $z_e \in \mathcal{U}_\eta$ is an equilibrium point, then it is necessarily a global minimum of $\mathcal{S}|_{\mathcal{U}_\eta}$ that is, $\mathcal{S}(z_e) = \mathcal{S}_\eta$. This follows from the fact that at an equilibrium point necessarily $\nabla\mathcal{S}(z_e)\cdot K(z_e)\nabla\mathcal{S}(z_e) = 0$, and (PL$'$) implies $\mathcal{S}(z_e) = \mathcal{S}_\eta$.

We can now state the analog of Proposition 4 for the case of metriplectic vector fields. However, we consider only the dissipative part, that is, a metriplectic vector field of the form (7), without the symplectic part, and make an additional assumption that is sufficient to ensure that the orthogonal projection onto $\ker K(z)$ is smooth in z. Specifically, we assume that

$$\begin{aligned}&\text{there is a constant } r > 0, \text{ such that, for any } z \in \mathcal{Z},\\ &\text{zero is the only eigenvalue of } K(z) \text{ in the interval } [-r, r].\end{aligned} \tag{39}$$

Hypothesis (39) implies that, for any z, only one eigenvalue of $K(z)$, the zero eigenvalue, belongs in the disk of radius $r > 0$ centered at zero in the complex plane.

Concerning the application of metriplectic dynamics to the calculation of equilibria of fluids and plasmas, the restriction to systems of the form (7) is not a significant limitation, since we often construct the relaxation method from the metric bracket only, as we do in the examples of Sections 4–6 below. In general however, it could be convenient to account for the ideal dynamics of the considered system. In that case Polyak's argument fails since the entropy gradient alone is not sufficient to control the time derivative of $|z(t) - z(t_1)|$ in the first step of the proof. We are not aware of any generalization of the PL inequality to completely general metriplectic fields.

Proposition 5. *Let $X = -K\nabla\mathcal{S}$ be a vector field of the form (7) on a domain $\mathcal{Z} \subseteq \mathbb{R}^n$, with K satisfying (39) and let $\mathcal{I} = (\mathcal{I}^1, \dots, \mathcal{I}^k) \in C^\infty(\mathcal{Z}, \mathbb{R}^k)$, $1 \le k < n$, be such that $K\nabla\mathcal{I}^\alpha = 0$. Assume that $\operatorname{rank}\nabla\mathcal{I} = k$ in a subdomain $\mathcal{U}$ satisfying (34) and let $\mathcal{O}$ be the open set given by Lemma 1. If (PL′) holds in $\mathcal{U}$, then for any integral curve $z(t)$ of X with initial condition $z(0) = z_0 \in \mathcal{O}$ there is a point $z_\eta \in \mathcal{O}$ and a constant θ_η depending on z_0, such that $\eta = \mathcal{I}(z_\eta) = \mathcal{I}(z_0)$, $\mathcal{S}(z_\eta) = \mathcal{S}_\eta$ and*

$$|z(t) - z_\eta| \le \theta_\eta e^{-\kappa_\eta t/2}, \quad \big[\mathcal{S}\big(z(t)\big) - \mathcal{S}_\eta\big] \le \big[\mathcal{S}\big(z_0\big) - \mathcal{S}_\eta\big] e^{-\kappa_\eta t}, \quad \textit{for } t \ge 0.$$

The proof of Proposition 4 can be adapted to this case. The key point is replacing $\nabla\mathcal{S}$ with $(I - \pi_0)\nabla\mathcal{S}$, where π_0 is the orthogonal projector onto the $\ker K$.

Proof of Proposition 5. Per hypothesis, both K and $\mathcal{S}$ are smooth on $\mathcal{Z}$ and thus X is Lipschitz continuous on any bounded subset and in particular on $\mathcal{U}$. Then Lemma 1 gives an open subset of $\mathcal{O} \subseteq \mathcal{U}$ such that for any point $z_0 \in \mathcal{O}$ there is an integral curve $z(t) \in \mathcal{O}$ of X through the point $z_0 = z(0)$, defined for all $t \ge 0$. Along the orbit $\mathcal{I}^\alpha\big(z(t)\big) = \mathcal{I}^\alpha(z_0) = \eta^\alpha$, $\alpha \in \{1, \dots, k\}$, and thus $z(t) \in \mathcal{U}_\eta \cap \mathcal{O}$.

At a point $z_\eta \in \mathcal{U}_\eta$ where $\mathcal{S}(z_\eta) = \mathcal{S}_\eta$ the function $\mathcal{S}|_{\mathcal{U}_\eta}$ attains its minimum and thus $\mathcal{S}$ must satisfy (35). It follows from the assumption $K\nabla\mathcal{I}^\alpha = 0$ that $X(z_\eta) = 0$. Therefore, if there is $\bar{t} \ge 0$ such that $\mathcal{S}\big(z(\bar{t})\big) = \mathcal{S}_\eta$, then $X\big(z(\bar{t})\big) = 0$ and $z(t) = z(\bar{t}) = z_0$ for all $t \ge 0$. In this case the statement of the proposition is trivially true.

Let us now consider the non-trivial case $\mathcal{S}\big(z(t)\big) > \mathcal{S}_\eta$, $t \ge 0$. We shall follow the same two steps as in Polyak original proof, with the necessary changes to account for the degeneracy of the metriplectic flow. This requires a preliminary step in which we establish the needed properties of the matrix K.

Step 0. For any point $z \in \mathcal{Z}$, let $K_i(z) > 0$ be the i-th non-zero eigenvalue of $K(z)$ and $\pi_i(z)$ the orthogonal projector on $\ker\big(K(z) - K_i(z)I\big)$, i.e. on the eigenspace of $K(z)$ corresponding to the eigenvalue $K_i(z)$. Then

$$K(z) = \sum_i K_i(z)\pi_i(z).$$

Since $K(z)$ is symmetric any eigenvalue (including zero) is semisimple [108, Appendix 3.I]. We assumed that the closed disk of radius $r > 0$ centered in $0 \in \mathbb{C}$ contains only the zero eigenvalue of $K(z)$ for all $z \in \mathcal{Z}$, hence $K_i(z) > r$. It follows that [108, Theorem 3.I.1] the orthogonal projector $\pi_0(z)$ onto $\ker K(z)$ is a smooth function of $z \in \mathcal{Z}$. In addition we have that, for any vector $Z \in \mathbb{R}^n$,

$$\begin{aligned} Z \cdot K(z)Z &= \sum_i K_i(z) Z \cdot \pi_i(z) Z \\ &\ge r \sum_i Z \cdot \pi_i(z) Z = r Z \cdot (I - \pi_0) Z. \end{aligned}$$

If $Z = \nabla\mathcal{S}(z)$, we deduce

$$\nabla\mathcal{S}(z) \cdot K(z)\nabla\mathcal{S}(z) \ge r\Big|\big(I - \pi_0(z)\big)\nabla\mathcal{S}(z)\Big|^2. \tag{40}$$

We also have $X = -K\nabla\mathcal{S} = -K(I - \pi_0)\nabla\mathcal{S}$, and $|X(z)| \le \|K(z)\|_F\Big|\big(I - \pi_0(z)\big)\nabla\mathcal{S}(z)\Big|$, where $\|K(z)\|_F$ is the Frobenius norm, which is a continuous function of z. For any z in the compact set $\overline{\mathcal{U}}$, $\|K(z)\|_F \le R = \max\{\|K(z')\|_F : z' \in \overline{\mathcal{U}}\}$, so that

$$|X(z)| \le R\Big|\big(I - \pi_0(z)\big)\nabla\mathcal{S}(z)\Big|. \tag{41}$$

Since π_0 is smooth, $Y(z) = \big(I - \pi_0(z)\big)\nabla\mathcal{S}(z)$ is smooth. This is the component of $\nabla\mathcal{S}$ orthogonal to $\ker K$.

From (41) we can also deduce that $Y\big(z(t)\big) \ne 0$ for, if not, then (PL′) implies $\mathcal{S}\big(z(t)\big) = \mathcal{S}_\eta$, which is the trivial case.

We shall show that Polyak's argument can be repeated with Y instead of $\nabla\mathcal{S}$.

Step 1. Upon using (40), given $0 \le t_1 < t_2$,

$$\mathcal{S}\big(z(t_1)\big) - \mathcal{S}\big(z(t_2)\big) \ge r\int_{t_1}^{t_2}\Big|Y\big(z(t)\big)\Big|^2 dt.$$

On the other end, (41) yields

$$\frac{d}{dt}\Big|z(t) - z(t_1)\Big| \le \Big|X\big(z(t)\big)\Big| \le R\Big|Y\big(z(t)\big)\Big|.$$

Since $Y \in C^\infty(\mathcal{Z}, \mathbb{R}^n)$, it is in particular Lipschitz continuous on $\mathcal{O}$. Let $L > 0$ be the Lipschitz constant of Y. The same argument of step 1 in the proof of Proposition 4 can be repeated leading to

$$\Big|Y\big(z(t)\big)\Big|^2 \le \frac{\alpha R L}{r}\big[\mathcal{S}\big(z(t)\big) - \mathcal{S}\big], \tag{42}$$

where α is the same constant defined in the proof of Proposition 4.

Step 2. We have already shown that

$$\varphi_\eta(t) \le \varphi_{0,\eta} e^{-\kappa_\eta t},$$

where $\varphi_\eta(t) = \mathcal{S}\big(z(t)\big) - \mathcal{S}_\eta$, and $\varphi_{0,\eta} = \varphi_\eta(0)$. For any $0 \le t_1 < t_2$, inequality (42) gives

$$\begin{aligned}
|z(t_1) - z(t_2)| &\le R \int_{t_1}^{t_2} \left|Y\big(z(t)\big)\right| dt \\
&\le R\sqrt{\frac{\alpha R L}{r}} \int_{t_1}^{t_2} \sqrt{\varphi_\eta(t)} dt \\
&\le \frac{1}{\kappa_\eta}\sqrt{\frac{4\alpha R^3 L\varphi_{0,\eta}}{r}}\left[e^{-\kappa_\eta t_1/2} - e^{-\kappa_\eta t_2/2}\right].
\end{aligned}$$

It follows that $z(t)$ has a limit $z_\eta \in \overline{\mathcal{O}}$ for $t \to +\infty$ and the limit satisfies $\mathcal{S}(z_\eta) = \lim_{t\to+\infty} \mathcal{S}\big(z(t)\big) = \mathcal{S}_\eta$, hence $z_\eta \in \mathcal{O}$. In addition, $\mathcal{I}(z_\eta) = \lim_{t\to+\infty} \mathcal{I}\big(z(t)\big) = \mathcal{I}(z_0)$, hence $z_\eta \in \mathcal{U}_\eta \cap \mathcal{O}$. At last, passing to the limit $t_2 \to +\infty$ yields

$$|z(t_1) - z_\eta| \le \theta_\eta e^{-\kappa_\eta t_1/2},$$

with constant

$$\theta_\eta = \frac{1}{\kappa_\eta}\sqrt{\frac{4\alpha R^3 L\varphi_{0,\eta}}{r}}.$$

This is the claimed inequality. □

This exponential convergence result becomes less useful when the constant θ_η is large, which can happen when either r or κ_η are small. In the former case, an eigenvalue of K becomes small at least in some region of the domain. In the latter case, the metric bracket is small even where entropy is far from the minimum.

This result follows from a minimal modification of Polyak's argument for standard gradient flows. One should note that here $\nabla\mathcal{I}^\alpha$ are assumed to be in the kernel of K, and this assumption is consistent with the requirement (7b) for $\mathcal{H}$.

Unlike the results of Section 3.1, convergence results based on Polyak inequalities do not require the uniqueness of the minimum entropy state. On the other hand, the precise point on the set of minima at which each orbit converges depends on the specific orbit. We illustrate inequality (PL′) with a few examples.

Example 1.

Let the phase space be $\mathcal{Z} = \mathbb{R}^n$, $n \in \mathbb{N}$ and $n \ge 2$, with coordinates $z = (z^i)_{i=1}^n$, $s = (s_i)_{i=1}^n$ and $h = (h_i)_{i=1}^n \in \mathbb{R}^n$ be covariant vectors, $K = (K^{ij})$ be a symmetric, positive-semidefinite, contravariant tensor on $\mathbb{R}^n$ such that $K^{ij}h_j = 0$, and let $\sigma = (\sigma_{ij})_{ij}$ be a symmetric positive definite, covariant tensor. We define the dissipative part of the metriplectic vector field $X = -K\nabla\mathcal{S}$, with Hamiltonian $\mathcal{H}(z) = h_i z^i$, and entropy $\mathcal{S}(z) = s_i z^i + \frac{1}{2}\sigma_{ij} z^i z^j$. We further assume that the null space of K coincides with the line spanned by h, i.e., $K\omega = 0$ implies $\omega = \lambda h$ for some $\lambda \in \mathbb{R}$. Then the metric bracket defined by K is "minimally degenerate", in the sense defined above.

We claim that this metriplectic system satisfies condition (PL′), and thus Proposition 5 applies.

In order to show this, let us first observe that the change of variables $z \mapsto \tilde{z} = \sigma^{1/2} z$ transforms the system into an analogous one with σ_{ij} replaced by δ_{ij} and with h and s replaced by $\sigma^{-1/2}h$ and $\sigma^{-1/2}s$, respectively. Hence it is enough to discuss the case $\sigma_{ij} = \delta_{ij}$. We can also assume $|h|^2 = 1$, because the normalization of h only changes the value of the Hamiltonian but not its isosurfaces. Then, for any $\eta \in \mathbb{R}$, $\mathcal{U}_\eta = \{z \in \mathbb{R}^n : \mathcal{H}(z) = \eta\}$ is the plane given by $h_i z^i = \eta$. A point $z \in \mathcal{U}_\eta$ can be written as $z = \eta h + z_\perp$, with $z_\perp = z - (h\cdot z)h$. Given $\eta \in \mathbb{R}$, we can use Lagrange multipliers to compute the constrained entropy minima $\mathcal{S}_\eta$: we search for (λ, z) such that (with $\partial_i = \partial/\partial z^i$) $\partial_i\mathcal{S}(z) = \lambda\partial_i\mathcal{H}(z)$ with $\mathcal{H}(z) = \eta$, which is equivalent to

$$\begin{cases} z = \lambda h - s, \\ h\cdot z = \eta. \end{cases}$$

The solution (λ_η, z_η) is readily found,

$$\lambda_\eta = \eta + h\cdot s, \quad z_\eta = \lambda_\eta h - s = \eta h - s_\perp,$$

where $s_\perp = s - (h\cdot s)h$. Therefore, if $z \in \mathcal{U}_\eta$,

$$\mathcal{S}(z) = s\cdot(\eta h + z_\perp) + \frac{1}{2}|\eta h + z_\perp|^2 = \mathcal{S}(z_\eta) + \frac{1}{2}|z_\perp + s_\perp|^2.$$

On the other hand, since $Kh = 0$, for any $z \in \mathcal{U}_\eta$,

$$\nabla\mathcal{S}(z)\cdot K\nabla\mathcal{S}(z) = (z+s)\cdot K(z+s) = (z_\perp + s_\perp)\cdot K(z_\perp + s_\perp) \ge K_1|z_\perp + s_\perp|^2,$$

where $K_1 > 0$ is the smallest eigenvalue of K restricted to $(\ker K)^\perp$, the orthogonal of its kernel. We deduce

$$\nabla\mathcal{S}(z)\cdot K\nabla\mathcal{S}(z) \ge 2K_1\left[\mathcal{S}(z) - \mathcal{S}_\eta\right], \quad z \in \mathcal{U}_\eta,$$

where $\mathcal{S}_\eta = \mathcal{S}(z_\eta)$ is the constrained minimum of the entropy. This is condition (PL′) with $\kappa_\eta = 2K_1$.

We remark that, at least in this case, the condition on the kernel of K being "minimal" is crucial for the modified PL condition. In fact, if there is a vector $h' \in \mathbb{R}^n$, orthogonal to h, and such that $Kh' = 0$, then $(z_\perp + s_\perp) \cdot K(z_\perp + s_\perp) = 0$ for any non-zero $z_\perp + s_\perp \propto h'$.

This example is simple enough that an analytical solution of the integral curves of X can be obtained. In fact, the equation for the new variable $y = s + z$ amounts to the linear system $dy/dt = -Ky$ with initial condition $y_0 = s + z_0$. If $\mathcal{H}(z_0) = h \cdot z_0 = \eta$, we must have $h \cdot y_0 = \eta + h \cdot s$. Upon representing y on the basis of the unit eigenvectors $\{e_i\}_{i=0}^{n-1}$ of K, with $e_0 = h$ being the eigenvector that corresponds to the zero eigenvalue, we obtain

$$y(t) = (h \cdot y_0)h + \sum_{i\geq 1} e^{-t\lambda_i(K)} c_i e_i,$$

where $\lambda_i(K) > 0$ are the positive eigenvalues of K and $c_i = e_i \cdot y_0$. We deduce

$$|z(t) - z_\eta| \leq |z_0 - z_\eta| e^{-tK_1},$$

with $K_1 = \min_{i\geq 1}\{\lambda_i(K)\}$.

Hence, in this case we have exponential convergence to the equilibrium point, with convergence rate being half of the constant in (PL′).

Example 2.

With the same metric bracket and Hamiltonian as in Example 1, let us consider the entropy function

$$\mathcal{S}(z) = \frac{|z|^2}{1+|z|^2}, \quad z \in \mathcal{Z} = \mathbb{R}^n.$$

As before, this entropy is rotationally symmetric, with a global minimum at $z = 0$, but it is not a convex function.

Since $\mathcal{H}$ is the same as in Example 1, $z \in \mathcal{U}_\eta$ if and only if $z = \eta h + z_\perp$ and we compute

$$\nabla\mathcal{S}(z) \cdot K\nabla\mathcal{S}(z) = 4\frac{z_\perp \cdot Kz_\perp}{(1+|z|^2)^4} \geq 4K_1\frac{|z_\perp|^2}{(1+|z|^2)^4},$$

where K_1 is defined in Example 1. We can use Lagrange multipliers in order to compute minima of the entropy constrained to $\mathcal{U}_\eta$ with the result that there is a unique minimum at $z_\eta = \eta h$ and

$$\mathcal{S}_\eta = \mathcal{S}(z_\eta) = \frac{\eta^2}{1+\eta^2}.$$

Then, we compute

$$\mathcal{S}(z) - \mathcal{S}_\eta = \frac{|z_\perp|^2}{(1+\eta^2)(1+|z|^2)},$$

from which we deduce

$$\frac{|z_\perp|^2}{(1+|z|^2)^4} = \frac{1+\eta^2}{(1+|z|^2)^3}\left[\mathcal{S}(z) - \mathcal{S}_\eta\right], \quad z \in \mathcal{U}_\eta,$$

hence, for any $R > 0$, on the ball $|z| < R$, we have

$$\nabla\mathcal{S}(z) \cdot K\nabla\mathcal{S}(z) \geq \kappa_\eta\left[\mathcal{S}(z) - \mathcal{S}_\eta\right], \quad z \in \mathcal{U}_\eta \cap B_R(0),$$

with constant $\kappa_\eta = 4K_1(1+\eta^2)/(1+R^2)^3$. Therefore the modified PL condition is satisfied on balls of arbitrary large radius R, even though the entropy is not convex, but the constant as a function of the radius R is not uniformly bounded away from zero.

Example 3.

As a last example, we consider a strongly nonlinear case with a bracket built from an orthogonal projection onto the hyper-plane perpendicular to the gradient of the Hamiltonian. This particular metric structure will play a key role in the following, even in the infinite-dimensional cases in fluid and plasma dynamics.

Given $s \in \mathbb{R}^n$, on the open half-space $\mathcal{Z} = \{z \in \mathbb{R}^n : z \cdot s < 0\}$, let us consider the field $X(z) = -K(z)\nabla\mathcal{S}(z)$, with

$$K(z) := |z|^2 I - z \otimes z, \quad \mathcal{H}(z) := \tfrac{1}{2}|z|^2, \quad \text{and} \quad \mathcal{S}(z) := s \cdot z\,.$$

Since $K(z)$ is proportional to the projector onto the subspace normal to $\nabla\mathcal{H}(z)$, we have $K(z)\nabla\mathcal{H}(z) = 0$ and $K(z)$ is a symmetric positive semidefinite tensor; hence, X is metriplectic with a trivial symplectic part. We also stress that $K(z)$ is minimally degenerate since $K(z)$ is strictly positive definite on the subspace normal to $\nabla\mathcal{H}(z)$.

The constant-energy surfaces are spheres, for any $\eta > 0$,

$$z \in \mathcal{U}_\eta \iff z = \sqrt{2\eta}\,\zeta, \quad \zeta \in S^{n-1}, \quad \zeta \cdot s < 0,$$

where points ζ on the $(n-1)$-dimensional sphere S^{n-1} are identified with unit vectors in $\mathbb{R}^n$. The entropy restricted to $\mathcal{U}_\eta$ amounts to $\mathcal{S}|_{\mathcal{U}_\eta}(\zeta) = \sqrt{2\eta}s \cdot \zeta$ and the minimum $\mathcal{S}_\eta = -\sqrt{2\eta s^2}$ is attained at $\zeta = -s/|s|$. The same result is of course obtained by means of Lagrange multipliers that lead to the system

$$\begin{cases} s = \lambda z, \\ \tfrac{1}{2}|z|^2 = \eta, \end{cases} \quad \text{and} \quad z \cdot s < 0.$$

Then we compute, for $z = \sqrt{2\eta}\zeta \in \mathcal{U}_\eta$,

$$\nabla \mathcal{S}(z) \cdot K(z)\nabla \mathcal{S}(z) = |z|^2|s|^2 - (z \cdot s)^2 = \big(\sqrt{2\eta}|s| - \sqrt{2\eta}s \cdot \zeta\big)\big(\sqrt{2\eta}|s| + \sqrt{2\eta}s \cdot \zeta\big)$$
$$\geq \sqrt{2\eta}|s|\big[\mathcal{S}(z) - \mathcal{S}_\eta\big],$$

which is inequality (PL′).

It should be noted that, if one drops the condition $z \cdot s < 0$, so that $\mathcal{Z} = \mathbb{R}^n$, then the metric system cannot satisfy (PL′) since $\mathcal{S}|_{\mathcal{U}_\eta}$ has two critical points: a minimum z_η^- with $z_\eta^- \cdot s < 0$ and a maximum z_η^+ with $z_\eta^+ \cdot s > 0$. The metric bracket $(\mathcal{S}, \mathcal{S}) = \nabla \mathcal{S} \cdot K\nabla \mathcal{S}$ vanishes at both points, but $\mathcal{S}(z_\eta^+) - \mathcal{S}_\eta > 0$.

3.3. Infinite-dimensional systems: tentative generalizations

A version of the Lyapunov stability theorem, valid for the case of infinite-dimensional systems, is available under suitable coercivity assumptions on the Lyapunov function [66]. Such assumptions are needed to compensate for the lack of compactness. For instance, the closed unit ball is not compact in a Banach space. Compactness of closed and bounded sets in $\mathbb{R}^n$ is used repeatedly in the classical proofs in finite dimensions (cf. Sections 3 and 3.2). Analogously, the (PL) condition can be extended to infinite-dimensional systems. A more difficult point is the existence of a global-in-time solution to the equation defining the dynamical system, under reasonable hypothesis [99]. In the infinite-dimensional setting, this means proving the existence of a global solution for highly nonlinear partial differential equations, which is often difficult and requires special treatment for each individual case. Nonetheless, under the assumption that a global-in-time solution exists, one can think of extending Propositions 3 and 5 to infinite dimensions, but we leave the details for future work. Here we merely state the infinite-dimensional version of condition (L3″) and inequality (PL′).

Consider a metriplectic system on a Banach space V as introduced in Section 2.1. We assume that this system has a finite (for simplicity) family of constants of motion $\mathcal{I} \in C^\infty(V, \mathbb{R}^k)$, that satisfies the hypotheses of the submersion theorem [66, Theorem 3.5.4] in an open set $\mathcal{U} \subseteq V$. In particular, the operator $D\mathcal{I}(u)$ must be surjective with split kernel for any $u \in \mathcal{U}$. Then $\mathcal{U}_\eta = \{u \in V : \mathcal{I}(u) = \eta\}$ are closed submanifolds of $\mathcal{U}$ for any $\eta \in \mathcal{I}(\mathcal{U})$, as in the finite-dimensional case. Since we consider systems of the form (7), satisfying in particular condition (7b), there is at least one invariant, namely the Hamiltonian, and thus we have $k \geq 1$. Then condition (L3″) can be generalized by

$$(\mathcal{F}, \mathcal{F})(u) = 0 \iff D\mathcal{F}(u) = \sum_\alpha \lambda_\alpha D\mathcal{I}^\alpha(u), \tag{43}$$

for some constant $\lambda_\alpha \in \mathbb{R}$. This means that if, for a given function $\mathcal{F}$, the bracket $(\mathcal{F}, \mathcal{F})$ vanishes at a point u_0, then u_0 must be a critical point of $\mathcal{F}$ restricted to the manifold $\mathcal{I}(u) = \mathcal{I}(u_0) =$ constant. We referred to brackets with this property as specifically degenerate brackets. If the only invariant is the Hamiltonian, then we called them minimally degenerate.

The equivalent of condition (PL′) reads: $\mathcal{S}_\eta := \inf\{\mathcal{S} : z \in \mathcal{U}_\eta\} > -\infty$ and there exists a constant $\kappa_\eta > 0$ depending on η, such that

$$\frac{1}{\kappa_\eta}(\mathcal{S}, \mathcal{S}) \geq \mathcal{S} - \mathcal{S}_\eta, \quad \text{on } \mathcal{U}_\eta. \tag{PL″}$$

If inequality (PL″) is fulfilled, the exponential convergence of the entropy follows as in the finite-dimensional case. Also, $(\mathcal{S}, \mathcal{S})(u) = 0$ on $\mathcal{U}_\eta$ only if u is a global minimum of $\mathcal{S}$ restricted to $\mathcal{U}_\eta$, i.e. $\mathcal{S}(u) = \mathcal{S}_\eta$.

A necessary condition for (PL″) can be stated for the special class of specifically degenerate metric brackets, i.e., when (43) is satisfied. Then condition (PL″) is satisfied *only if* critical points of $\mathcal{S}|_{\mathcal{U}_\eta}$ are global minima. In fact, if $u \in \mathcal{U}_\eta$ is a critical point of $\mathcal{S}|_{\mathcal{U}_\eta}$, the theory of Lagrange multiplier [66, Theorem 3.5.27] gives $D\mathcal{S}(u) = \sum_\alpha \lambda_\alpha D\mathcal{I}^\alpha(u)$, hence $(\mathcal{S}, \mathcal{S})(u) = 0$. But if $\mathcal{S}(u) > \mathcal{S}_\eta$, inequality (PL″) is violated.

Beyond these preliminary considerations, the mathematical analysis of (PL″) exceeds the scope of this paper. We conclude with an example of a infinite-dimensional metric bracket that is specifically degenerate and satisfies (PL″). We shall give physically relevant examples in Section 4.

Example 4.

In this example (cf. [35]) we proceed formally. On the space V of smooth functions from $\mathbb{T} \to \mathbb{R}$, where the torus $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ is identified with the interval $\Omega = [0, 2\pi]$ with periodic boundary conditions, we consider the metric bracket given by

$$(\mathcal{F}, \mathcal{G}) = \int_0^{2\pi} \left(\frac{\delta \mathcal{F}(u)}{\delta u}\right)' \left(\frac{\delta \mathcal{G}(u)}{\delta u}\right)' dx, \tag{44}$$

where $v'(x) = dv(x)/dx$ denotes the derivative of $v : \mathbb{T} \to \mathbb{R}$, and the functional derivatives are computed with respect to the standard L^2 product (cf. Section 2.1). We assume that the functions $\mathcal{F}(u)$ and $\mathcal{G}(u)$ are regular enough for their functional derivative to exist and be sufficiently smooth. The Hamiltonian and entropy functions are given by

$$\mathcal{H}(u) = \int_0^{2\pi} u(x)dx \quad \text{and} \quad \mathcal{S}(u) = \frac{1}{2}\int_0^{2\pi} |u(x)|^2 dx = \frac{1}{2}\|u\|^2_{L^2(\Omega)}.$$

Condition (7b) is satisfied since $\delta\mathcal{H}(u)/\delta u = 1$. After integration by parts, the strong form of (7a) amounts to the heat equation

$$\begin{cases} \partial_t u = \partial_x^2 u, & (t, x) \in [0, +\infty) \times [0, 2\pi], \\ u(t, 0) = u(t, 2\pi), & t \in [0, +\infty), \\ u(0, x) = u_0(x), & x \in [0, 2\pi]. \end{cases} \tag{45}$$

First we show that (44) is a minimally degenerate bracket, i.e., the null space is spanned by $\delta\mathcal{H}(u)/\delta u$. In fact, $(\mathcal{F},\mathcal{F}) = 0$ implies $(\delta\mathcal{F}(u)/\delta u)' = 0$ and thus $\delta\mathcal{F}(u)/\delta u = \lambda$ where $\lambda \in \mathbb{R}$ is constant. Hence, $\delta\mathcal{F}(u)/\delta u = \lambda\delta\mathcal{H}(u)/\delta u$, which proves property (43), with the Hamiltonian being the only invariant.

The manifolds of constant Hamiltonian,

$$\mathcal{U}_\eta = \{u : \mathcal{H}(u) = \eta \in \mathbb{R}\},$$

consist of functions with the same average over $[0, 2\pi]$. They are affine spaces, rather than generic manifolds. The critical points of entropy restricted to the constant-Hamiltonian spaces are determined by

$$u(x) = \lambda, \quad \int_0^{2\pi} u(x)dx = \eta.$$

Therefore, for any $\eta \in \mathbb{R}$ there is only one critical point, that is, the constant function

$$u_\eta(x) = \eta/(2\pi).$$

The entropy of u_η is $\mathcal{S}_\eta = \mathcal{S}(u_\eta) = \eta^2/(4\pi)$. The Fourier series representation,

$$u(x) = \sum_{n\in\mathbb{Z}} u_n e^{inx}, \quad u_n \in \mathbb{C},$$

yields that $\mathcal{H}(u) = \eta$ if and only if $u_0 = \eta/(2\pi)$, hence

$$\mathcal{S}(u) - \mathcal{S}_\eta = \pi \sum_{n\neq 0} |u_n|^2 \geq 0, \quad u \in \mathcal{U}_\eta,$$

with equality only if $u = u_\eta$. This shows that u_η is a global minimum of $\mathcal{S}$ restricted to $\mathcal{U}_\eta$. Upon using again the Fourier series representation, one finds

$$\big(\mathcal{S},\mathcal{S}\big)(u) = \|u'\|^2_{L^2(\Omega)} = 2\pi \sum_{n\neq 0} n^2|u_n|^2 \geq 2\pi \sum_{n\neq 0} |u_n|^2,$$

and

$$\big(\mathcal{S},\mathcal{S}\big)(u) \geq 2\big[\mathcal{S}(u) - \mathcal{S}_\eta\big], \quad u \in \mathcal{U}_\eta,$$

which is condition (PL$''$) with $\kappa_\eta = 2$.

In fact, the solution of (45) can be readily written in terms of a Fourier series as

$$u(t,x) = \sum_{n\in\mathbb{Z}} e^{inx-n^2t} u_{0,n},$$

where $u_{0,n} \in \mathbb{C}$ are the Fourier coefficients of the initial condition. With $\eta = \mathcal{H}(u_0) = 2\pi u_{0,0}$, we deduce

$$\|u(t) - u_\eta\|_{L^2(\Omega)} \leq e^{-t}\|u(0) - u_\eta\|_{L^2(\Omega)},$$

which shows exponential relaxation toward the entropy minimum on $\mathcal{U}_\eta$, with the energy η being determined by the initial condition.

In conclusion, this metric system satisfies the generalized Polyak–Łojasiewicz condition (PL$''$) and all orbits completely relax exponentially to a solution of (1) with exponential convergence rate given by $\kappa_\eta/2$, where κ_η is the constant in (PL$''$). We observe that the convergence rate is the same as the one predicted in Proposition 5 for finite-dimensional systems.

4. Two examples: metric double brackets and projectors

In this section, we discuss two special cases of metriplectic systems of the form (7), which we call metric double bracket and projector bracket systems. We shall see that for the metric double brackets treated in Section 4.1, the dissipation mechanism does not completely relax the state of the system (in the sense of Section 3), while for the projector brackets of Section 4.2 complete relaxation is achieved. We shall discuss and compare the properties of these two metric brackets on the basis of the insights gained in Section 3.

We select two benchmark equilibrium problems, and we attempt to construct a relaxation method to solve them by using the two considered metric brackets. The benchmark problems are: the reduced Euler equations (cf. Section 2.2.1), but with Dirichlet boundary conditions replaced by periodic boundary conditions, and an analytically solvable model derived from the reduced Euler equations. In both cases periodic boundary conditions give rise to an additional invariant other than the Hamiltonian, and this allows us to examine cases where specifically degenerate brackets are not minimally degenerate. We shall return to the original problem with Dirichlet boundary conditions later.

Therefore in both cases, the domain is $\mathbb{T}^2 := (\mathbb{R}/2\pi\mathbb{Z})^2$ with coordinates $x = (x_1, x_2)$, and the phase space V is the space of smooth functions $v : \mathbb{T}^2 \to \mathbb{R}$. As usual, $\mathbb{T}^2$ is identified with the square $\Omega = [0, 2\pi]^2$ with periodic boundary conditions. On such a periodic domain, the scalar vorticity $\omega = -\Delta\phi$ (cf. Section 2.2.1 for the definitions) must have zero average, i.e.,

$$\omega_\Omega := \frac{1}{4\pi^2}\int_\Omega \omega(x)dx = 0.$$

We use systematically the subscript Ω to denote the average over the domain Ω. We choose the whole space V as the phase space and define the vorticity by

$$\omega = u - u_\Omega,$$

but we impose the additional constraint

$$\mathcal{M}(u) = \int_\Omega u(x)dx = 4\pi^2 u_\Omega = \mathcal{M}_0 \in \mathbb{R},$$

as well as energy conservation $\mathcal{H}(u) = \mathcal{H}_0 \in \mathbb{R}$. Given $u \in V$, and thus ω, the Poisson equation determines the stream function ϕ modulo a constant, which we set to zero; hence (10) is replaced by

$$-\Delta\phi = u - u_\Omega, \quad \phi_\Omega = 0. \tag{46}$$

(Equivalently, we could have chosen the phase space to be the subspace of functions satisfying $u_\Omega = 0$ and $u = \omega$.)

Both the considered benchmark problems can be formulated as variational problems: given a Hamiltonian function $\mathcal{H}$, and a regular value $\eta = (\mathcal{M}_0, \mathcal{H}_0)$ for the two invariants $\mathcal{I} = (\mathcal{I}^1, \mathcal{I}^2) := (\mathcal{M}, \mathcal{H})$, find

$$\min\{\mathcal{S}(u) : \mathcal{I}(u) = \eta\}, \tag{47a}$$

with entropy

$$\mathcal{S}(u) = \frac{1}{2}\int_\Omega \omega^2 dx = \frac{1}{2}\|u - u_\Omega\|^2_{L^2(\Omega)}. \tag{47b}$$

The two test cases differ by the choice of the Hamiltonian. In summary,

- Analytical test case: Given a function $h : \mathbb{T}^2 \to \mathbb{R}$, let

  $$\mathcal{H}(u) = \int_\Omega h\,\omega\,dx = (h - h_\Omega, u)_{L^2(\Omega)}. \tag{48}$$

  The solutions of (47) with (48) are equilibria of the linear advection equation

  $$\partial_t u + [h, u] = 0,$$

  with $[f, g] = \partial_1 f \partial_2 g - \partial_1 g \partial_2 f$.
- Reduced Euler test case: We again use $\mathcal{S}$ as in (47b), but now

  $$\mathcal{H}(u) = \frac{1}{2}\int_\Omega |\nabla\phi|^2 dx = \frac{1}{2}(\phi, u)_{L^2(\Omega)}, \tag{49}$$

  i.e., we assume (14) with $s(y) = y^2/2$, and we assume ϕ depends on u via (46). Solutions of (47) with (49) are equilibria of the reduced Euler equations on the flat torus $\mathbb{T}^2$.

In both cases, problem (47) can be solved analytically. In order to compute the solutions, we first find the set of critical points of entropy restricted to $\mathcal{U}_\eta = \{u : \mathcal{I}(u) = \eta\}$, i.e.,

$$\mathfrak{C}_\eta := \{u : D\mathcal{S}(u) = \sum_\alpha \lambda_\alpha D\mathcal{I}^\alpha(u),\ \mathcal{I}(u) = \eta\}. \tag{50}$$

Then we find the minimum of $\mathcal{S}$ on $\mathfrak{C}_\eta$. We summarize here the results, which will be used to assess the metriplectic relaxation methods.

Solution for the analytical test case: For the case of Eq. (48), the set $\mathfrak{C}_\eta$ is given by

$$u - u_\Omega = \lambda_1 + \lambda_2(h - h_\Omega), \quad \mathcal{I}(u) = \eta. \tag{51}$$

Then $\lambda_1 = 0$ and $u_\Omega = \mathcal{M}_0/(4\pi^2)$. Upon multiplying by $h - h_\Omega$ and integrating over Ω, we deduce

$$\mathcal{H}_0 = \lambda_2 \|h - h_\Omega\|^2_{L^2(\Omega)},$$

from which we can compute λ_2. Hence, the set $\mathfrak{C}_\eta$ contains the following single point:

$$u_\eta = \frac{\mathcal{M}_0}{4\pi^2} + \frac{\mathcal{H}_0}{\|h - h_\Omega\|^2_{L^2(\Omega)}}(h - h_\Omega) \tag{52}$$

and the value of entropy on this unique critical point is

$$\mathcal{S}_\eta = \min\{\mathcal{S}(u) : \mathcal{I}(u) = \eta\} = \frac{\mathcal{H}_0^2/2}{\|h - h_\Omega\|^2_{L^2(\Omega)}}. \tag{53}$$

One can check that this is the minimum of the entropy on $\mathcal{U}_\eta$.

Solution for the reduced Euler equations: For the case of (49), elements of $\mathfrak{C}_\eta$ satisfy

$$\begin{cases} u - u_\Omega = \lambda_1 + \lambda_2\phi, \quad \mathcal{I}(u) = \eta, \\ \text{with } \phi \text{ solution of (46).} \end{cases} \tag{54}$$

Since $\phi_\Omega = 0$, we have $\lambda_1 = 0$ and $u_\Omega = \mathcal{M}_0/(4\pi^2)$, as before. Then (ϕ, λ_2) must be a solution of the eigenvalue problem

$$-\Delta\phi = \lambda_2\phi, \qquad \phi_\Omega = 0,$$

which is readily solved in terms of Fourier series. We find that the set $\mathfrak{C}_\eta$ consists of vorticity fields $u - u_\Omega = \omega = \lambda_2\phi$, with ϕ being an eigenfunction of $-\Delta$ corresponding to the eigenvalue $\lambda_2 > 0$, and with norm $\|\phi\|^2_{L^2(\Omega)} = 2\mathcal{H}_0/\lambda_2$. Then the entropy evaluated on the constrained critical points amounts to $\lambda_2\mathcal{H}_0$. It follows that the entropy minimum on $\mathcal{U}_\eta$ corresponds to the lowest non-trivial eigenvalue, which is $\lambda_2 = 1$. The corresponding stream function must be of the form

$$\phi(x) = a_1\cos(x_1 + \theta_1) + a_2\cos(x_2 + \theta_2),$$

with arbitrary phase shifts θ_1, θ_2, and with coefficient a_1, a_2 determined by the condition $\|\phi\|^2_{L^2(\Omega)} = 2\mathcal{H}_0$. Since $\lambda_2 = 1$, this amounts to $a_1^2 + a_2^2 = \mathcal{H}_0/\pi^2$. Thus,

$$\omega(x) = \phi(x) = \frac{\sqrt{\mathcal{H}_0}}{\pi}\big[\cos\theta_0\cos(x_1 + \theta_1) + \sin\theta_0\cos(x_2 + \theta_2)\big], \tag{55}$$

with arbitrary phases θ_0, θ_1, and θ_2 in $[0, 2\pi)$.

From the analytical solution, (55), we deduce that the entropy minimum constrained to $\mathcal{I}(u) = \eta$ is not attained at an isolated point, but on a family of points parameterized by three phases. The constrained minimum value of the entropy is given by

$$\mathcal{S}_\eta = \min\{\mathcal{S}(u) : \mathcal{I}(u) = \eta\} = \mathcal{H}_0, \tag{56}$$

since $\lambda_2 = 1$.

4.1. Metric double brackets

The first example is given by the metric double bracket [55,63,64] (not to be confused with the double bracket of Flierl and Morrison [47] discussed in the introduction). We recall the general definition first, but quickly restrict the discussion to the examples.

In general, metric double brackets originate from a Lie algebra. If $\mathfrak{g}$ is a Lie algebra with Lie brackets $[\cdot,\cdot]$, let the vector space V be its dual, $V = \mathfrak{g}^*$. The functional derivative of a function $f \in C^\infty(\mathfrak{g}^*)$ is computed with respect to the duality pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$, that is, $\delta f(u)/\delta u \in \mathfrak{g}$ is the unique element of $\mathfrak{g}$ such that $Df(u)v = \langle \delta f(u)/\delta u, v\rangle$ for all $v \in V = \mathfrak{g}^*$. Under these conditions, it is well-known [66] that the Lie bracket in $\mathfrak{g}$ induces two Poisson brackets in $C^\infty(\mathfrak{g}^*)$, namely, $\{f,g\}_\pm = \pm\langle u, [\frac{\delta f(u)}{\delta u}, \frac{\delta g(u)}{\delta u}]\rangle$, so that both $(\mathfrak{g}^*, \{\cdot,\cdot\}_\pm)$ are Poisson manifolds. However, if in addition $\mathfrak{g}$ is equipped with a positive definite bilinear form $\gamma : \mathfrak{g}\times\mathfrak{g} \to \mathbb{R}$, on the space of smooth functions $C^\infty(\mathfrak{g}^*)$, we can also define the symmetric bracket

$$(f,g) = \gamma\Big(\Big[\frac{\delta f(u)}{\delta u}, \frac{\delta h(u)}{\delta u}\Big], \Big[\frac{\delta g(u)}{\delta u}, \frac{\delta h(u)}{\delta u}\Big]\Big),$$

for any fixed function $h \in C^\infty(\mathfrak{g}^*)$. One can readily check that this is a metric bracket preserving the Hamiltonian function h.

Formally at least, this construction can be extended to infinite-dimensional systems. Let V and W be Banach spaces, with a nondegenerate duality pairing $\langle\cdot,\cdot\rangle_{V\times W} : V \times W \to \mathbb{R}$, and let W be equipped with (i) a symmetric positive definite bilinear form $\gamma : W \times W \to \mathbb{R}$ and (ii) a bilinear antisymmetric operation $[\cdot,\cdot] : W \times W \to W$. (For the purpose of defining the metric bracket we do not need to require $[\cdot,\cdot]$ to be a Lie bracket, that is, we can relax the Jacobi identity.) Then, given a fixed Hamiltonian $\mathcal{H} \in C^\infty(V)$, we can construct the bilinear form $(\cdot,\cdot) : C^\infty(V) \times C^\infty(V) \to C^\infty(V)$ given by [63, Eq. (2.9)]

$$(\mathcal{F},\mathcal{G}) := \gamma\Big(\Big[\frac{\delta\mathcal{F}}{\delta u}, \frac{\delta\mathcal{H}}{\delta u}\Big], \Big[\frac{\delta\mathcal{G}}{\delta u}, \frac{\delta\mathcal{H}}{\delta u}\Big]\Big), \tag{57}$$

where the functional derivative are evaluated with respect to the duality pairing between V and W, and thus are elements of W. We remark that when $W = V'$ is the (topological) dual of V, that is, the space of continuous linear functionals on V, $\delta\mathcal{F}(u)/\delta u$ exists and it is equal to $D\mathcal{F}(u)$ for any $\mathcal{F} \in C^1(V)$ and for all $u \in V$. In general, however, $\delta\mathcal{F}(u)/\delta u$ does not always exists for all $\mathcal{F}$.

As an example, let $V = W$ be the space of smooth functions from $\mathbb{T}^d \to \mathbb{R}$, identified with functions over $\Omega = [0, 2\pi]^d$ with periodic boundary conditions. If $[u, v]_J = \nabla u \cdot J\nabla v$ is a Poisson bracket on $\mathbb{R}^d$ (not necessarily canonical), $\mathcal{H}(u)$ is a given Hamiltonian function, and γ is given by the standard product in $L^2(\Omega)$, then (57) reduces to

$$(\mathcal{F},\mathcal{G}) := \int_\Omega \Big[\frac{\delta\mathcal{F}}{\delta u}, \frac{\delta\mathcal{H}}{\delta u}\Big]_J \Big[\frac{\delta\mathcal{G}}{\delta u}, \frac{\delta\mathcal{H}}{\delta u}\Big]_J dx. \tag{58}$$

In the following, let $d = 2$, $x = (x_1, x_2) \in \Omega = [0, 2\pi]^2 \subset \mathbb{R}^2$ with periodic boundary conditions, and for J, we choose the canonical Poisson tensor in $\mathbb{R}^2$, so that $[\cdot,\cdot]_J = [\cdot,\cdot]$ is the canonical bracket defined after Eq. (11).

We construct a relaxation method based on this bracket for the solution of the two test problems introduced at the beginning of this section. In both cases we consider a field $u \in V$ evolving from an initial condition u_0 according to Eqs. (7) and with bracket given by (58).

We start from problem (47) with (48) for the linear advection equation. With those choices of bracket, Hamiltonian and entropy, Eq. (7a) amounts to

$$\int_\Omega \frac{\delta \mathcal{F}}{\delta u}\frac{\partial u}{\partial t}dx = -\int_\Omega \Big[\frac{\delta \mathcal{F}}{\delta u}, h\Big]\big[u, h\big]dx,$$

for all $\mathcal{F}$. This can be viewed as the weak form of the evolution equation, with $\delta\mathcal{F}/\delta u$ being the test function. After integration by parts, one obtains the evolution equation for u in strong form, namely,

$$\partial_t u = \big[h, [h, u]\big].$$

We observe that $[u, h] = \mathrm{div}(X_h u) = X_h \cdot \nabla u$, where $X_h = {}^t(\partial_2 h, -\partial_1 h)$ is the Hamiltonian vector field generated by h with canonical Poisson bracket in $\mathbb{R}^2$. Hence

$$\partial_t u = \mathrm{div}(X_h \otimes X_h \nabla u), \tag{59}$$

which shows that this particular combination of metric bracket and entropy describes anisotropic diffusion, parallel to the field lines of X_h, or equivalently along the contours of the function h.

The Cauchy problem associated to (59) with periodic boundary conditions can be solved analytically, at least under suitable conditions. We consider the equation in a region of the domain where the contours of h are closed simple curves and $\nabla h \neq 0$ (as relevant in the example discussed below). Let ξ be a function such that

$$[\xi, h] = 1,$$

in the considered subdomain. Since $[\xi, h] \neq 0$, the pair of functions (ξ, h) defines a local coordinate system with inverse Jacobian determinant $\mathcal{J}^{-1} = [\xi, h] = 1$ and such that $X_h^1 := X_h \cdot \nabla\xi = [\xi, h] = 1$ and $X_h^2 := X_h \cdot \nabla h = [h, h] = 0$. Then, in these coordinates the contravariant components X_h^i, $i = 1, 2$, of the vector field X_h as well as the Jacobian determinant $\mathcal{J}$ are constant. In order to compute ξ, we observe that, along a contour we must have $dx/ds = X_h/|X_h|$ where s is the arclength (with the Euclidean metric, $ds^2 = dx_1^2 + dx_2^2$), because the field X_h is tangent to $h =$ constant contours. Hence

$$\frac{d\xi}{ds} = \frac{dx}{ds}\cdot\nabla\xi = \frac{X_h\cdot\nabla\xi}{|X_h|} = \frac{1}{|\nabla h|},$$

so that $d\xi = |\nabla h|^{-1}ds$. With some abuse of notation, let $x(\xi, h)$ be the coordinate map $(\xi, h) \mapsto x$. Then we also have, with h fixed,

$$\frac{\partial x(\xi, h)}{\partial \xi} = \frac{dx}{ds}\frac{ds}{d\xi} = X_h\big(x(\xi, h)\big),$$

hence the coordinate map $x(\xi, h)$ is essentially related to the flow of the vector field X_h. This conclusion also follows from the fact that $\partial_\xi x$ is a vector of the covariant basis, hence it must hold that $\partial_\xi x \cdot \nabla h = 0$ and $\partial_\xi x \cdot \nabla \xi = 1$, which imply $\partial_\xi x = X_h$.

Eq. (59) in the coordinates (ξ, h) takes the form of a heat equation,

$$\partial_t \tilde{u} - \partial_\xi^2 \tilde{u} = 0,$$

where the new unknown is given by $u(t, x) = \tilde{u}(t, \xi, h)$. Since, per assumption, the contours of h are closed, the function $\tilde{u}$ must be periodic in ξ with period possibly depending on h, i.e., there is ℓ_h such that $\tilde{u}(t, \xi + \ell_h, h) = \tilde{u}(t, \xi, h)$. The period ℓ_h is given by the variation of ξ over a full loop around the considered contour of h, that is,

$$\ell_h = \int_{C_h} d\xi = \int_{C_h}\frac{ds}{|\nabla h|},$$

where C_h is the considered contour of h. This gives a way to compute ℓ_h from $|\nabla h|$ on a contour C_h.

We rescale the variable ξ to an angle $\vartheta \in [0, 2\pi]$, i.e. $\vartheta = 2\pi\xi/\ell_h$. In terms of the angle ϑ, Eq. (59) amounts to

$$\partial_t v - \kappa_h \partial_\vartheta^2 v = 0,$$

where $\kappa_h = (2\pi/\ell_h)^2$, and the rescaled unknown is given by $u(t, x) = \tilde{u}(t, \ell_h\vartheta/(2\pi), h) = v(t, \vartheta, h)$. This is the classic heat equation on $[0, 2\pi]$ with periodic boundary conditions, and it can be readily solved by Fourier series. The solution is

$$v(t, \vartheta, h) = \sum_{n\in\mathbb{Z}} \hat{v}_n(0)e^{-n^2\kappa_h t + in\vartheta},$$

where $\hat{v}_n(0)$ are the Fourier coefficients of the initial condition for v, which is explicitly given by $v(0, \theta, h) = u_0\big(x(\xi, h)\big)$. Each Fourier mode with $n \neq 0$ decays exponentially with exponential decay time given by $1/(n^2\kappa_h)$. The relaxation time is identified with the decay time of the slowest modes ($n = \pm 1$),

$$\tau_h = 1/\kappa_h = (\ell_h/2\pi)^2, \qquad \ell_h = \int_{C_h} d\xi = \int_{C_h}\frac{ds}{|\nabla h|}. \tag{60}$$

This expression allows us to estimate numerically the relaxation time from a sample of points on the considered contour of h, which can be obtained by integrating the ordinary differential equation for the flow of X_h. The limit for $t \to +\infty$ of the solution exists and is equal to the average of the initial condition on the contours of h. Explicitly this can be computed as

$$u_\infty(h) := \hat{v}_0(0) = \frac{1}{\ell_h}\int_0^{\ell_h} u_0\big(x(\xi, h)\big)d\xi = \frac{1}{\ell_h}\int_{C_h}\frac{u_0\,ds}{|\nabla h|}.$$

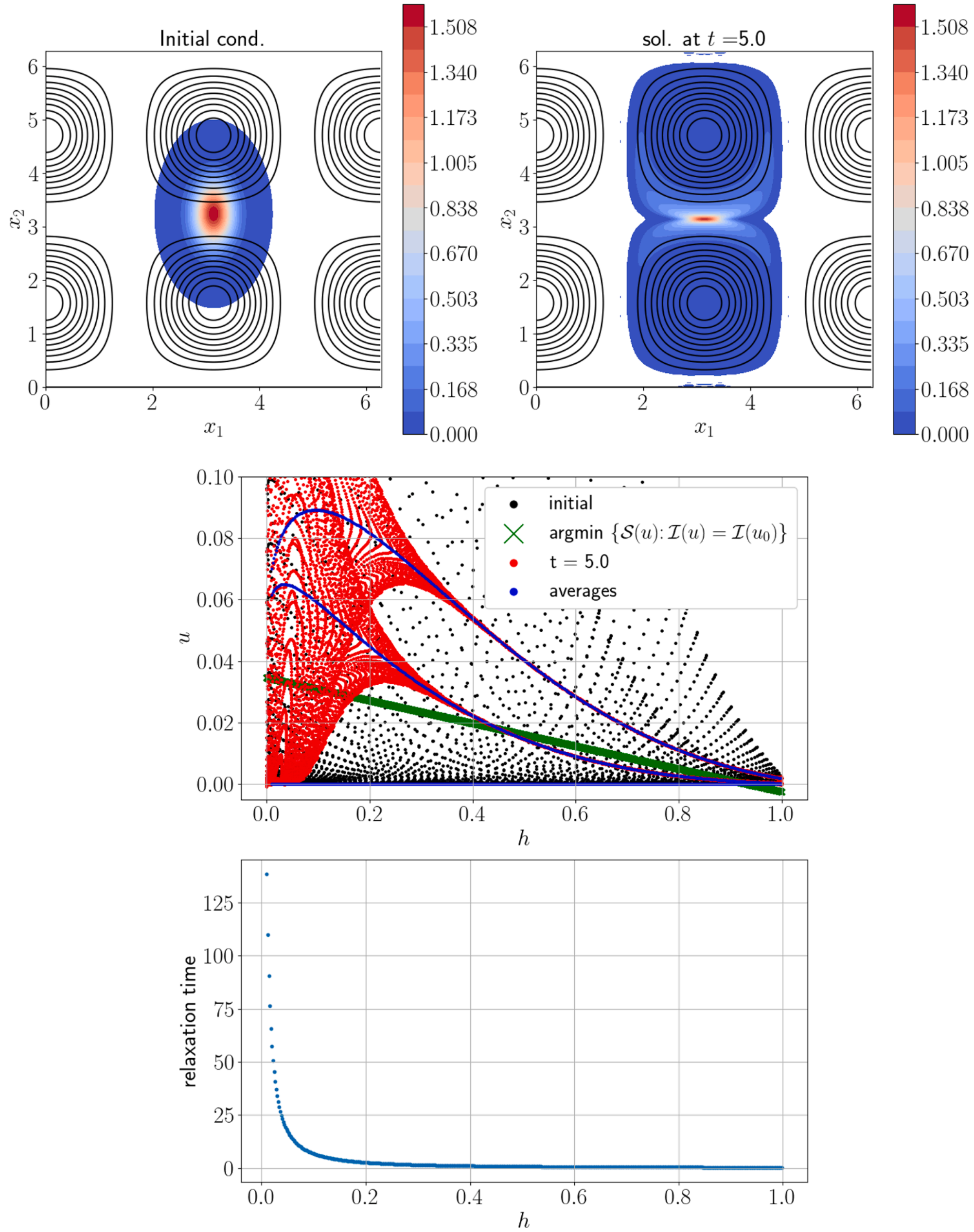

Fig. 1. Example of solution of Eq. (59) with Hamiltonian (61). Upper panels: initial condition and final state, compared to the contours of h (black circular curves). Middle panel: visualization of the functional relation between h and u obtained by plotting the points (h_{ij}, u_{ij}), with h_{ij} and u_{ij} being the values of h and u, at the node (i, j) of the computational grid. Lower panel: relaxation time τ_h, computed from Eq. (60) on the contours of the two central (full) islands, as a function of h. (For clarity, in the color maps we display the solution u only where $u \geq 10^{-4}$).

Therefore, in general the limit of the solution retains some information of the initial condition, while the completely relaxed solutions (52) only depends on the energy $\mathcal{H}_0 = \mathcal{H}(u_0)$ of the initial condition u_0. This implies that for generic initial conditions, the limit of the solution of a metric dynamical system is in general not a solution of the variational principle (47).

Fig. 1 shows the result of the numerical solution of (59) with Hamiltonian

$$h(x) = \cos^2(x_1)\sin^2(x_2). \tag{61}$$

The contours of h form a periodic array of islands, cf. the black contours in the upper panels of Fig. 1. In each island, h takes the same values. The initial condition, represented as a color map, is an anisotropic Gaussian centered between two islands, i.e., $u_0(x) = u_G(x)$

with

$$u_G(x) = \frac{1}{N}\exp\Big[-\frac{(x_1-x_{0,1})^2}{w_1^2} - \frac{(x_2-x_{0,2})^2}{w_2^2}\Big], \tag{62}$$

with center $x_0 = (\pi, \pi + 0.1)$, $w_1 = 0.25$, $w_2 = 0.4$, and $N = 2\pi w_1 w_2$. The solution is obtained with a standard spectral method with Fourier basis, on a 256×256 uniform grid. The time integrator is the standard 4th order explicit Runge-Kutta method with time step $\Delta t = 10^{-4}$. The parallel diffusion Eq. (59) tends to equalize the solution on the contours of h, but the dynamics at the boundary of the islands is very slow: the Hamiltonian vector field X_h at the boundary of the islands is zero and the solution remains constant on those boundary contours (referred to as separatrices). The color map of the relaxed state is shown in Fig. 1 upper panel, while the middle panel represents the functional relation between h and the solution u by marking on the $h - u$ plane a point (h_{ij}, u_{ij}) for each grid node x_{ij}. At the initial time, (black markers) there is no relation between the values of u and those of h, showing that the initial condition (62) is far from an equilibrium. As the solution evolves, all values of u sampled on the same contour of h tend to a common value, but the "condensation" of points on a line is slower for h small, that is, near the separatrices. Blue markers show the average of the initial condition on each contour of h: one can see that the solution tends to the averages as predicted by the analytical solution. In the limit $t \to +\infty$ the relation between h and u is multi-valued with a countable set of branches, one for each island. In Fig. 1 one can distinguish the upper branch (larger values of u), corresponding to the island that contains the maximum of the initial condition. (In order to separate the two branches the center x_0 of the Gaussian has been shifted up in the direction x_2.) A second branch with lower values of u corresponds to the neighboring island. All the other islands do not overlap with the initial condition significantly and therefore appear as a line of points $u = 0$ for all h. In Fig. 1 the analytical solution (52) for the completely relaxed state is also shown (green crosses), and it is clearly different from the obtained equilibrium. Fig. 1, lower panel, shows the relaxation time τ_h computed according to Eq. (60). This result confirms that the relaxation becomes progressively slower as h approaches $h = 0$, that is, near the separatrices of the islands. In the limit $h \to 0$ we have $\tau_h \to +\infty$ consistently with the fact that $\nabla h = 0$ and $X_h = 0$ on the separatrices, cf. the denominator in Eq. (60).

We now consider problem (47) with (49) for the Euler equations. For this choice of entropy and Hamiltonian, Eq. (7a) with (58) amounts to the anisotropic diffusion Eq. (59), but with h replaced by ϕ, which depends on the state variable u, and thus on the vorticity $\omega = u - u_\Omega$, via Eq. (46). Hence the problem is nonlinear with a cubic nonlinearity, and in general no analytical solution is known (to the best of our knowledge).

Fig. 2 shows an example of relaxation of an initially anisotropic vortex. The initial state is again given by (62), but now with $x_0 = (\pi, \pi)$, $w_1 = 0.3$, $w_2 = 1.0$, and $N = 1$, on $[0, 2\pi]^2$ with a uniform mesh of 256×256 nodes. The time integrator is the standard 4th-order explicit Runge-Kutta method with time step $\Delta t = 10^{-3}$. The solution relaxes to a symmetric vortex, which is an equilibrium of the Euler equations. During the evolution, the Hamiltonian is constant and the entropy is monotonically dissipated, consistently with (8). However, the entropy appears to converge to a value that is higher than its constrained minimum $\mathcal{S}_\eta$, given in Eq. (56), and indicated by the thick horizontal line in Fig. 2. The fact that the final state is (a numerical approximation of) an equilibrium of the Euler equations can be deduced from the plot of the final state: the solution for the vorticity ω appears to be constant on the contours of the potential ϕ. A more quantitative indication is provided in Fig. 3, where the relation between ϕ and ω is represented. At the initial time $t = 0$, there is no functional relation between ω and ϕ: the values (ϕ, ω) on the computational grid do not belong to a curve. As the state relaxes, the scatter of points is reduced, and at the final time, one finds a clear functional relation.

For a completely relaxed state, we expect a linear relation between the potential and the vorticity, Eq. (55), and this is indicated by green crosses in Fig. 3. The obtained relationship is however very different, proving that the dynamical system reaches an equilibrium that does not satisfy the variational principle of minimum constrained entropy (1). From the numerical experiments in Fig. 3, one can see that the relaxed state is close to the average of the initial condition on the contours of the initial potential. The average of the initial condition gives the exact long-time limit of the solution in the case of the linear problem (59), and it is not expected to give an accurate prediction of the relaxed state in general. Yet we observe that the solution converges to a state close to the average of the initial condition.

We conclude that the relaxation mechanism of (58) fails to capture the linear profile encoded in the choice of the entropy function. The relaxed state is an equilibrium of the reduced Euler equations, but corresponding to a profile that differs from the target one and that depends on the initial condition in a complicated way.

We can try to understand the behavior of this metric system in terms of the ideas put forward in Section 3.3. Specifically, for both the analytical case of Fig. 1 and the reduced Euler case of Figs. 2 and 3, we shall show that the metric bracket is not specifically degenerate, and the generalization of the PL inequality, Eq. (PL″), is not satisfied. We recall that in both cases the bracket is given by the metric double bracket defined in Eq. (58), but the Hamiltonian is different, and thus the null space of the bracket is different. Therefore we treat the two cases separately even though there are some similarities.

- Analytical test case. We begin by showing that the bracket is *not* specifically degenerate. With mass $\mathcal{I}^1 = \mathcal{M}$ and the energy $\mathcal{I}^2 = \mathcal{H}$ as the only two invariants, we want to show that $(\mathcal{F}, \mathcal{F})(u) = 0$ at a point u does not imply $\delta\mathcal{F}(u)/\delta u = \lambda_1 + \lambda_2 h$. With this aim, we observe that

$$(\mathcal{F}, \mathcal{F})(u) = 0 \iff \Big[\frac{\delta\mathcal{F}(u)}{\delta u}, \frac{\delta\mathcal{H}(u)}{\delta u}\Big] = \Big[\frac{\delta\mathcal{F}(u)}{\delta u}, h\Big] = 0.$$

This condition is satisfied at any point u for functions of the form

$$\mathcal{F}(u) = (f(h), u)_{L^2(\Omega)},$$

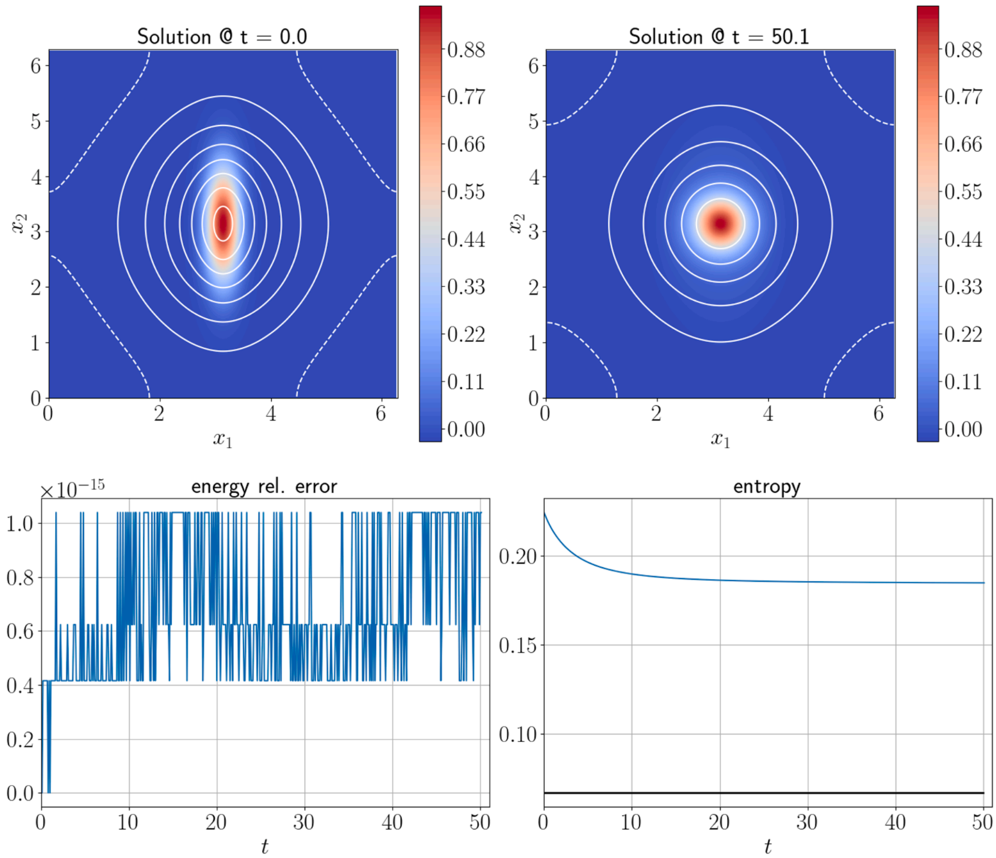

Fig. 2. Metriplectic relaxation of a vortex toward an equilibrium of the reduced Euler equations, using (58) with (14) and $s(y) = y^2/2$. Top row: initial and final state of the system; the color scheme represents the vorticity $\omega = u - u_\Omega$; white lines represent the contours of the potential ϕ. Bottom row: relative error of the Hamiltonian and the value of entropy during the evolution. The thick horizontal line indicates the constrained entropy minimum $\mathcal{S}_\eta = \mathcal{H}_0$, cf. Eq. (56).

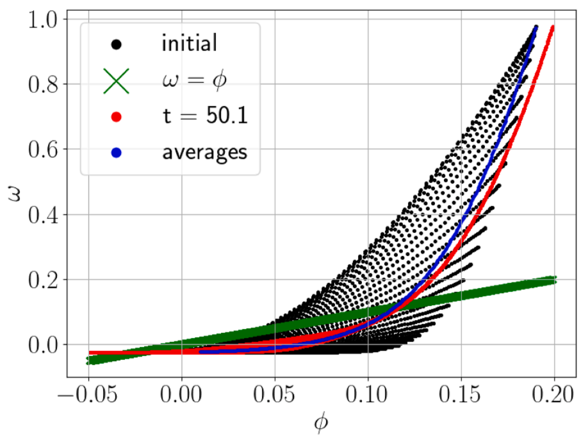

Fig. 3. Visualization of the relation between the potential ϕ and the vorticity ω, for the initial and final state of the calculation. The green crosses mark the linear relation for a minimum entropy state, Eq. (55). The data marked "averages" represent the average of the initial condition on the contours of the corresponding potential ϕ.

for any sufficiently regular function $f : \mathbb{R} \to \mathbb{R}$. We see that the condition $\delta\mathcal{F}(u)/\delta u = \lambda_1 + \lambda_2 h$, corresponds to the special case $f(h) = \lambda_1 + \lambda_2 h$. Therefore there are functions $\mathcal{F}$ for which $(\mathcal{F},\mathcal{F})(u) = 0$, but $\delta\mathcal{F}(u)/\delta u \neq \lambda_1 + \lambda_2 h$, and thus the bracket is *not* specifically degenerate, in the sense of Eq. (43).

As for inequality (PL″), we have that

$$\big(\mathcal{S},\mathcal{S}\big)(u) = 0 \iff \left[\frac{\delta\mathcal{S}(u)}{\delta u}, \frac{\delta\mathcal{H}(u)}{\delta u}\right] = [u,h] = 0.$$

Therefore any phase-space point of the form $u = f(h)$, with $f : \mathbb{R} \to \mathbb{R}$ a sufficiently regular function, is a zero of the bracket $(\mathcal{S},\mathcal{S})$. Constrained entropy minima, on the other hand, are affine functions $u - u_\Omega = \lambda_1 + \lambda_2 h$ with specific values of the multiplier λ_1 and λ_2 (the exact formula has been given in Eq. (52) but it is not needed here); hence condition (PL″) is false.

Therefore, neither one of the conditions of Section 3.3 holds true in this case. In fact the analytical solution and the numerical experiment show that the relaxation method finds a point of the (rather large) set $\{u : (\mathcal{S},\mathcal{S})(u) = 0\}$, instead of the unique entropy minimum (52).

- Reduced Euler test case. We first show that the bracket is not specifically degenerate. With this aim we construct a similar counterexample to the one used in the analytical case above, the only difference being that now $\delta\mathcal{H}(u)/\delta u = \phi$ is related to u via the Poisson equation $-\Delta\phi = u - u_\Omega$, Eq. (46). We consider a point u in phase space given by a solution of the problem

$$-\Delta\phi = u - u_\Omega, \qquad u = f(\phi),$$

for a given smooth function $f : \mathbb{R} \to \mathbb{R}$. The existence of nontrivial solutions is guaranteed for a large class for functions f [109]. Then, for any smooth function $g : \mathbb{R} \to \mathbb{R}$, and for $\mathcal{F}(u) = \int_\Omega g(u)dx$, we have that

$$(\mathcal{F},\mathcal{F})(u) = \int_\Omega \big[g'(u),\phi\big]dx = \int_\Omega \big[g'\circ f(\phi),\phi\big]dx = 0,$$

where u is the phase-space point defined above. On the other hand, $\delta\mathcal{F}(u)/\delta u = g'\circ f(\phi)$ in general is not a linear combination of the derivatives $\frac{\delta\mathcal{M}(u)}{\delta u}$ and $\frac{\delta\mathcal{H}(u)}{\delta u}$ of the two invariants, i.e.,

$$\frac{\delta\mathcal{F}(u)}{\delta u} \neq \lambda_1\frac{\delta\mathcal{M}(u)}{\delta u} + \lambda_2\frac{\delta\mathcal{H}(u)}{\delta u} = \lambda_1 + \lambda_2\phi,$$

therefore the bracket is not specifically degenerate.

As for condition (PL″), we observe that the entropy $\mathcal{S}$ is a special case of the class of functions discussed above, corresponding the choice $g(u) = u^2/2$, and, if u is the same phase-space point used above, we have $(\mathcal{S},\mathcal{S})(u) = 0$, but $\delta\mathcal{S}(u)/\delta u = f(\phi)$, which shows that in general u is not a constrained entropy minimum, since for a constrained entropy minimum we should have $f(\phi) = \phi$, cf. Eq. (54).

While we have rigorous results in the finite-dimensional case only, these observations show that, at least in these two cases, failure to relax the system to a (local) constrained entropy minimum occurs for a bracket that is neither specifically degenerate, nor satisfies the generalized PL inequality. This supports the idea that the convergence results obtained in Section 3.2 for finite-dimensional systems may also hold in general. For comparison, below in Section 4.2, we shall discuss a bracket that is specifically degenerate, and for this bracket we observe complete relaxation.

It is worth noting that in both the analytical case and the reduced Euler case the bracket defined in Eq. (58) does find a valid equilibrium of the system, but this equilibrium, in general, is not a constrained critical point of the entropy function, and in particular it cannot be a local constrained minimum of entropy. This implies that the bracket (58) could not be used to solve problems like the Grad-Shafranov equation as the resulting equilibrium would not be consistent with the imposed profiles that are encoded in the entropy function, cf. Section 2.2.2. Yet they can be useful in another way as we shall see below for the Beltrami fields.

4.2. Projector-based metric bracket

We address now a construction of metric brackets based on L^2-orthogonal projectors, which are patterned after that given for finite-dimensional systems in [3]. As before let $\Omega = [0,2\pi]^d$ and V be the space of functions on $\mathbb{R}^d$, 2π-periodic in each direction. Given a Hamiltonian function $\mathcal{H}$, the L^2 orthogonal projector onto the direction of $\delta\mathcal{H}(u)/\delta u$ is

$$\begin{aligned}
&\Pi_{\mathcal{H}}(u)v := v - c(u,v)\frac{\delta\mathcal{H}(u)}{\delta u},\\
&c(u,v) := \left\|\frac{\delta\mathcal{H}(u)}{\delta u}\right\|^{-2}_{L^2(\Omega)}\left(\frac{\delta\mathcal{H}(u)}{\delta u}, v\right)_{L^2(\Omega)}.
\end{aligned} \tag{63}$$

With the projector, let us define

$$(\mathcal{F},\mathcal{G}) := \left(\frac{\delta\mathcal{F}}{\delta u}, \Pi_{\mathcal{H}}\frac{\delta\mathcal{G}}{\delta u}\right)_{L^2(\Omega)}. \tag{64}$$

We claim that the symmetric bi-linear form (64) satisfies (7b) and the Leibniz identity. Since $\Pi_{\mathcal{H}}$ is a projector $\Pi_{\mathcal{H}}(\delta\mathcal{H}/\delta u) = 0$, hence $(\mathcal{F},\mathcal{H}) = 0$ for all $\mathcal{F}$; in addition,

$$(\mathcal{F},\mathcal{F}) = \left(\frac{\delta\mathcal{F}}{\delta u}, \Pi_{\mathcal{H}}\frac{\delta\mathcal{F}}{\delta u}\right)_{L^2(\Omega)} \geq 0,$$

since projectors are symmetric and nonnegative definite. The Leibniz identity is straightforward.

We utilize bracket (64) in (7) in order to obtain a relaxation method for the variational problems (47). For the case of the linear advection equation, energy (48) yields the evolution equation

$$\partial_t u = -\left[u - u_\Omega - \frac{\mathcal{H}(u)}{\|h - h_\Omega\|^2_{L^2(\Omega)}}(h - h_\Omega) \right],$$

where $u_\Omega = \mathcal{M}(u)/(4\pi^2)$. Since both $\mathcal{M}$ and $\mathcal{H}$ are constants of motion, the affine transformation $u \mapsto w = u - \mathcal{M}_0/(4\pi^2) - \left[\mathcal{H}_0/\|h - h_\Omega\|^2_{L^2(\Omega)}\right](h - h_\Omega)$ with $\mathcal{M}_0 = \mathcal{M}(u_0)$ and $\mathcal{H}_0 = \mathcal{H}(u_0)$, u_0 being the initial condition, transforms the equation into $\partial_t w = -w$, which leads to the analytical solution

$$u(t, \cdot) = \left[\frac{\mathcal{M}_0}{4\pi^2} + \frac{\mathcal{H}_0}{\|h - h_\Omega\|^2_{L^2(\Omega)}}(h - h_\Omega) \right](1 - e^{-t}) + u_0 e^{-t}.$$

The term in square brackets is exactly the unique entropy minimum (52) on the manifold $\mathcal{U}_\eta$ with $\eta = (\mathcal{M}_0, \mathcal{H}_0)$ being determined by the initial condition. Therefore for any initial condition u_0, this metriplectic system relaxes completely to the accessible entropy minimum and with exponential convergence rate. This is the desired behavior. From the analytical solution one can see how the initial condition is quickly "forgotten", leaving only the fully relaxed state.

Let us now move to the test case of the reduced Euler equations. Bracket (64) with Hamiltonian (49) and Eq. (7) yields the evolution equation

$$\partial_t u = -\left[u - u_\Omega - \frac{\mathcal{H}(u)}{\|\phi\|^2_{L^2(\Omega)}}\phi \right],$$

with ϕ depending on u via the Poisson Eq. (46). The evolution of u is therefore governed by a nonlinear integral operator with a nonpolynomial nonlinearity. We begin by considering a numerical experiment. Fig. 4 shows the initial and final state of a solution of the initial value problem for this equation, and Fig. 5 gives the representation of the functional relation between the potential ϕ and the vorticity $\omega = u - u_\Omega$. The initial condition as well as the numerical method and the numerical parameters (grid size and time steps) are the same as in Fig. 2.

Fig. 4 shows the relative error in energy conservation and the entropy as a function of time. The thick horizontal line denotes the minimum entropy value in Eq. (56). This time the minimum entropy value is quickly reached by the system.

From the scatter plot in Fig. 5 one can see that the final state is characterized by a linear relation between ϕ and ω. This suggests that the projector-based metric bracket (64) relaxes the state of the system completely. This nice property comes at the price of a larger dissipation of entropy, cf. Fig. 4. As a consequence, the vorticity of the final state is significantly lower than in the initial condition. The scatter plot in Fig. 5 also shows the average of the initial condition on the contours of the initial potential (which is the same as in Fig. 3); in this case, the relaxed state bears little or no similarity to the initial condition.

In order to check if the relaxed vorticity is in agreement with the analytical solution (55), we have computed the best fit of the analytical solution (55) to the final state in Fig. 4, varying the three phases θ_0, θ_1, and θ_2. The difference between the best fit and the relaxed state gives an estimate of the distance of the latter from the entropy minimum and it is shown in Fig. 5, right-hand-side panel.

In the terminology of Section 3.3, we claim that bracket (64) is *minimally degenerate*. In fact, for any function $\mathcal{F}$, we have $(\mathcal{F}, \mathcal{F}) = 0$ if and only if $\delta\mathcal{F}(u)/\delta u \in \ker \Pi_{\mathcal{H}}(u)$, or equivalently

$$\frac{\delta\mathcal{F}(u)}{\delta u} = \lambda\phi,$$

which is condition (43) with $\lambda_\alpha \neq 0$ only if $\mathcal{I}^\alpha = \mathcal{H}$.

If $\mathcal{F} = \mathcal{S}$, this condition is satisfied at any point of the set $\mathfrak{C}_\eta$ defined in Eq. (50), i.e., for any constrained critical point of the entropy, not just at the minimum. It follows that the generalization (PL″) of the PL condition cannot be true on the whole phase space. Nonetheless, with (58) we have

$$(\mathcal{S}, \mathcal{S})(u) = 2\mathcal{S}(u) - \frac{4\mathcal{H}_0^2}{\|\phi\|^2_{L^2(\Omega)}}. \tag{65}$$

We claim that for any u such that $\mathcal{H}(u) = \mathcal{H}_0$, the following inequalities hold true:

$$2\mathcal{S}(u) - \frac{4\mathcal{H}_0^2}{\|\phi\|^2_{L^2(\Omega)}} \geq 0, \tag{66a}$$

$$\mathcal{S} \geq \mathcal{H}_0, \tag{66b}$$

$$\|\phi\|^2_{L^2(\Omega)} \leq \|\nabla\phi\|^2_{L^2(\Omega)} = 2\mathcal{H}_0. \tag{66c}$$

Specifically, Eq. (66a) follows from the Cauchy-Schwarz inequality (which also ensure the positivity of the projector). Eq. (66b) is a consequence of (56), while (66c) is a Poincaré inequality on $\mathbb{T}^2$, obtained from the solution of the Poisson problem (46) via Fourier

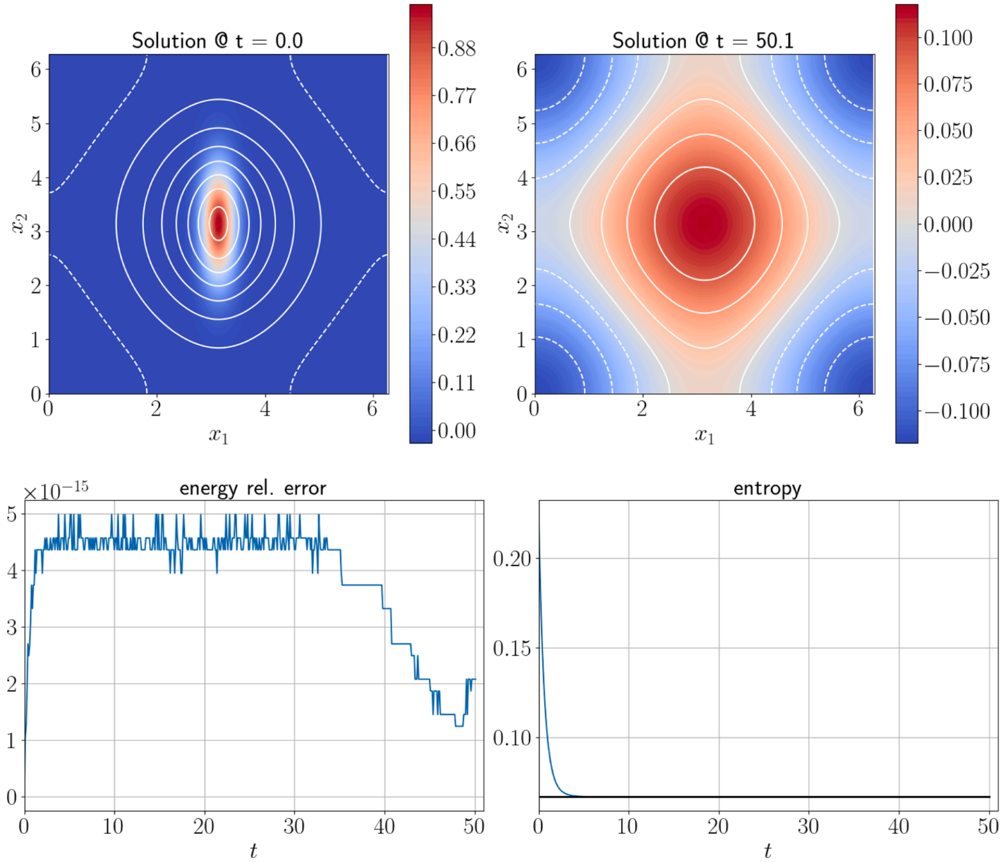

Fig. 4. The same as in Fig. 2 but for the metric bracket (64).

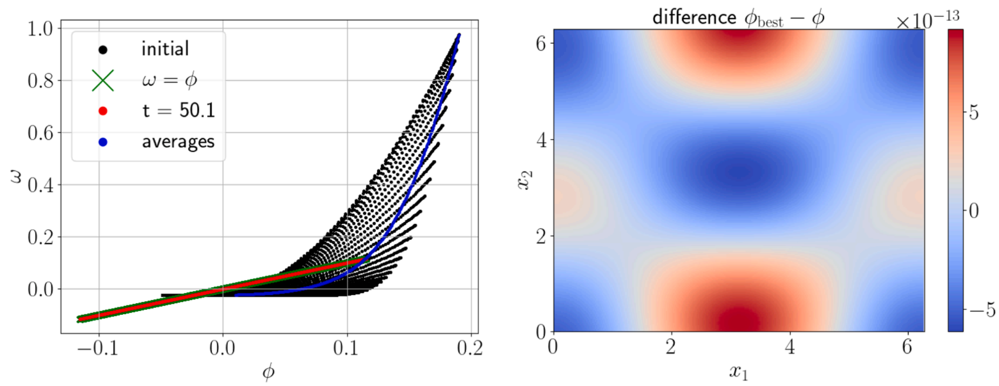

Fig. 5. Left-hand-side panel: the relation between ϕ and $\omega = u - u_\Omega$ for the case of Fig. 4. Right-hand-side panel: difference between the best fit of the exact solution (55) and the relaxed state.

series. Inequalities (66) imply that the evolution of the solution of the metriplectic system must be such that $\left(\mathcal{S}(u), 1/\|\phi\|^2_{L^2(\Omega)}\right) \in \mathbb{R}^2_+$ remains within the cone

$$\mathcal{S}(u) \geq \mathcal{H}_0 = \mathcal{S}_\eta, \quad \frac{1}{2\mathcal{H}_0} \leq \frac{1}{\|\phi\|^2_{L^2(\Omega)}} \leq \frac{1}{2\mathcal{H}_0} + \frac{\mathcal{S}(u) - \mathcal{S}_\eta}{2\mathcal{H}_0^2}.$$

We observe that the set $\mathfrak{C}_\eta$ is contained in the "upper boundary" of this cone, i.e., it satisfies

$$\frac{1}{\|\phi\|^2_{L^2(\Omega)}} = \frac{1}{2\mathcal{H}_0} + \frac{\mathcal{S}(u) - \mathcal{S}_\eta}{2\mathcal{H}_0^2}.$$

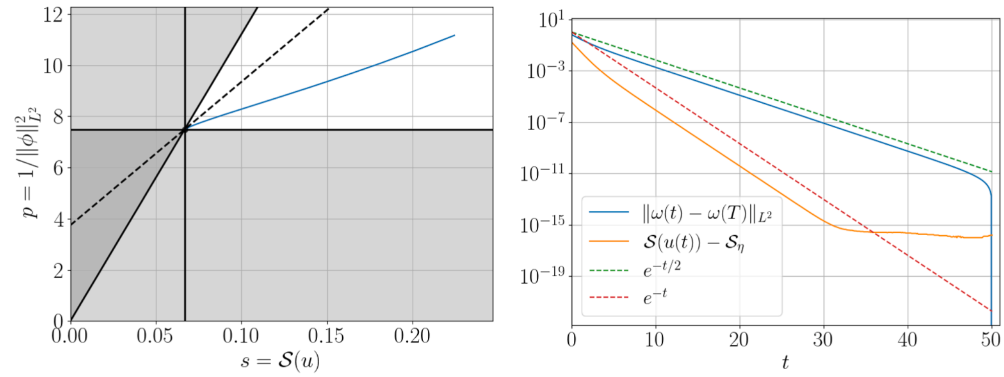

Fig. 6. Left-hand-side panel: trajectory in the plane $(\mathcal{S}(u), 1/\|\phi\|_{L^2(\Omega)})$ for the case of Fig. 4; the shaded area indicates the region of the plane excluded by inequalities (66). The dashed line indicates the upper boundary of the shrunk cone (67) with $a = 1/2$. The minimum entropy state corresponds to the vertex of the cone. Right-hand-side panel: evolution of the norm of the distance of $\omega(t)$ from the relaxed state, using the vorticity ω at the last point in time $t = T$ as an approximation of the latter, and the "excess entropy" $\mathcal{S}\big(u(t)\big) - \mathcal{S}_\eta$. The semi-log scale shows exponential relaxation of entropy with exponential rate ≈ 1. This is consistent with the inequality (PL$''$). The fact that ω has relaxation rate $\approx 1/2$ is a consequence of the simple choice of the entropy function, cf. Eq. (47).

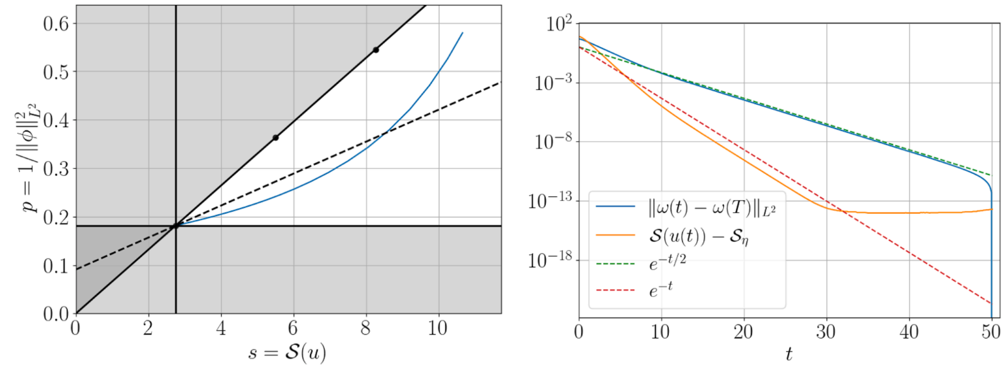

Fig. 7. The same as in Fig. 6, but for the initial condition (68). The black dots marked on the boundary of the cone correspond to the points of the set $\mathfrak{C}_\eta$.

For any $a \in (0, 1)$, and for any $u \in \mathcal{U}_\eta$ such that

$$\frac{1}{\|\phi\|^2_{L^2(\Omega)}} = \frac{1}{2\mathcal{H}_0} + (1-a)\frac{\mathcal{S}(u) - \mathcal{S}_\eta}{2\mathcal{H}_0^2}, \tag{67}$$

we have

$$\big(\mathcal{S}, \mathcal{S}\big) \geq 2a\big[\mathcal{S}(u) - \mathcal{S}_\eta\big],$$

which is inequality (PL$''$) with constant $\kappa_\eta = 2a$. Hence, if the solution stays within the shrunk cone (67), we expect exponential convergence with exponent related the angle of the cone (67).

Fig. 6 shows the trajectory of the solution in the plane $\big(\mathcal{S}(u), 1/\|\phi\|^2_{L^2(\Omega)}\big)$ for the case of Fig. 4. The solution indeed stays within the cone defined by (66). The figure also demonstrates exponential convergence of both entropy and vorticity with exponential rate consistent with the generalized PL condition. We note that, cf. Fig. 4, the system traverses most of the trajectory in Fig. 6 quickly, and reaches the vertex of the cone already at $t \approx 5$.

Fig. 7 shows the same trajectory for a different initial condition, viz.,

$$u_0(x) = \cos(2x_2) + u_G(x), \tag{68}$$

where u_G is the Gaussian defined in (62) with $1/N = 1.8$, $x_0 = (\pi, 3\pi/2)$, $w_1 = 0.3$, and $w_2 = 1$. The cosine terms shift the initial condition closer to the boundary. As a consequence the initial entropy relaxation rate is slower, but approaches ≈ 1 as the trajectory approaches the vertex of the cone.

5. Collision-like metric brackets

The specific structure of the metric bracket for the Landau collision operator was introduced and generalized in [1,3]. Here we propose a further generalization that we use as a "template" for the construction of relaxation methods. This generalized bracket will be referred to as the collision-like metric bracket [110], as it originates from Morrison's bracket for the Landau collision operator. The resulting evolution equation is integro-differential. We demonstrate the use of collision-like brackets for the solution of the variational principles in Eqs. (13) and (20) for equilibria of the reduced Euler equations and axisymmetric MHD, respectively. In both these applications the specification of *equilibrium profiles*, i.e., the relation between the state variable u and the corresponding potential (either ϕ for the Euler equations or ψ for axisymmetric MHD), is essential. In the variational principle, the profile is encoded in the choice of the entropy; therefore, for the result of metriplectic relaxation to be consistent with the imposed profile, the metric bracket must *completely relax* the state of the system, in the sense made precise in Section 3. Preliminary, numerical results on these two problems have been reported in the proceedings of the joint Varenna-Lausanne workshop on "The Theory of Fusion Plasmas" [111] and in the Ph.D. thesis by Bressan [110].

5.1. General construction of collision-like brackets

We introduce the class of collision-like metric brackets in a fairly abstract way, but we make no attempt to give a mathematically rigorous definition, except for a few basic considerations. The construction given here is somewhat more general than the one proposed earlier [110,111].

As in Section 2.1, the phase space V is a space of sufficiently regular functions $v : \Omega \to \mathbb{R}^N$ over a bounded domain $\Omega \subset \mathbb{R}^d$ with $N, d \in \mathbb{N}$. We always assume $V \subseteq L^2(\Omega, \mu; \mathbb{R}^N) =: W$, and the functional derivative of a function $\mathcal{F} \in C^1(V)$ is computed with respect to the standard inner product in W, hence $\delta\mathcal{F}(u)/\delta u \in W$, when it exists.

For the construction of the bracket, we consider a *bounded* domain $\mathcal{O} \subset \mathbb{R}^n$ equipped with a measure ν, and the space $\tilde{W} := L^2(\mathcal{O}, \nu; \mathbb{R}^{\tilde{N}})$ with $n, \tilde{N} \in \mathbb{N}$. For our purposes it is sufficient to assume that $d\nu(z) = \tilde{m}(z)dz$ with $\tilde{m} \in C^\infty(\overline{\mathcal{O}})$, dz being the Lebesgue measure on $\mathcal{O}$. Next, we choose a (possibly unbounded) linear operator

$$P : W \to \tilde{W}, \qquad \mathrm{dom}(P) = \Phi,$$

where the domain Φ is a subspace of W with a finer topology, so that $V \subseteq \Phi \subseteq W \subseteq \Phi'$, with Φ' being the dual of Φ (the space of continuous linear functionals on Φ), and with continuous inclusions. Then, W has the structure of a rigged Hilbert space, with the finer space Φ containing the phase space V. In this way, both quadratic functionals like $\mathcal{F}(u) = \int_\Omega u^2 d\mu$ and linear functionals like $\mathcal{F}(u) = \int_\Omega w \cdot u d\mu$ for $w \in \Phi$ are such that $\delta\mathcal{F}(u)/\delta u \in \Phi$. In addition, given a Hamiltonian $\mathcal{H}$ such that $\delta\mathcal{H}(u)/\delta u \in \Phi$, we choose

$$\mathsf{T} : V \to \mathcal{B}(\tilde{W}),$$

with values in the space $\mathcal{B}(\tilde{W})$ of bounded linear operators from $\tilde{W}$ into $\tilde{W}$, such that $T(u)$ is symmetric, positive semidefinite, and

$$\mathsf{T}(u) P \frac{\delta\mathcal{H}(u)}{\delta u} = 0. \tag{69}$$

In terms of the operator P and the function T, collision-like brackets that preserve $\mathcal{H}$ are defined by

$$(\mathcal{F}, \mathcal{G}) := \int_{\mathcal{O}} P\frac{\delta\mathcal{F}(u)}{\delta u} \cdot \mathsf{T}(u) P \frac{\delta\mathcal{G}(u)}{\delta u}\, d\nu. \tag{70}$$

Symmetry and positive semidefiniteness of the bracket is ensured by the fact that $\mathsf{T}(u)$ is symmetric and positive semidefinite, while Eq. (69) implies $(\mathcal{F}, \mathcal{H}) = 0$ for any $\mathcal{F}$. Hence, Eq. (70) defines a metric bracket on the class of functions $\mathcal{F}$ such that $\delta\mathcal{F}(u)/\delta u \in \Phi$. In most cases, $\mathsf{T}(u)$ is the operator of multiplication by a function $\mathsf{T}(u; x)$ with values in the space of real symmetric, positive semidefinite $\tilde{N} \times \tilde{N}$ matrices.

Upon introducing the dual operator $P' : \tilde{W} \to \Phi'$ defined by

$$\left\langle w, P'\tilde{w} \right\rangle = \int_{\mathcal{O}} \tilde{w} \cdot Pw\, d\nu, \qquad \text{for all } w \in \Phi, \tag{71}$$

where $\langle\cdot,\cdot\rangle : \Phi \times \Phi' \to \mathbb{R}$ is the duality pairing between Φ and Φ', bracket (70) can be equivalently written as

$$(\mathcal{F}, \mathcal{G}) = \left\langle \frac{\delta\mathcal{F}(u)}{\delta u}, P'\left[\mathsf{T}(u) P \frac{\delta\mathcal{G}(u)}{\delta u}\right] \right\rangle.$$

If $u \in C^1([0,T], V)$ is a trajectory in V and $\mathcal{F} \in C^1(V)$ has a functional derivative $\delta\mathcal{F}(u)/\delta u$ in Φ, then $t \mapsto \mathcal{F}\big(u(t)\big)$ is differentiable and

$$\frac{d}{dt}\mathcal{F}\big(u(t)\big) = \left\langle \frac{\delta\mathcal{F}\big(u(t)\big)}{\delta u}, \partial_t u \right\rangle.$$

Therefore, given an entropy $\mathcal{S}$ such that $\delta\mathcal{S}(u)/\delta u \in \Phi$, Eq. (7a) for $u(t)$ amounts to

$$\partial_t u = -P'\left[\mathsf{T}(u) P \frac{\delta\mathcal{S}(u)}{\delta u}\right] \quad \text{in } \Phi'. \tag{72}$$

This construction is summarized in Fig. 8.

The operator P maps an $\mathbb{R}^N$-valued function over the d-dimensional domain Ω into an $\mathbb{R}^{\tilde{N}}$-valued function over the n-dimensional domain $\mathcal{O}$, and we are particularly interested in the case $n > d$, $\tilde{N} \geq N$. As we shall see, increasing the number of dimensions has

$$
\begin{tikzcd}
 & & \tilde{W} \arrow[loop above, "\mathsf{T}(u)"] \arrow[dr, "P'"] & \\
V \arrow[r, hook] & \Phi \arrow[r, hook] \arrow[ur, "P"] & W \arrow[r, hook] & \Phi' \\
\cup & \cup & \parallel & \\
u(t),\ \partial_t u(t) & \frac{\delta \mathcal{F}(u)}{\delta u} & L^2(\Omega, \mu; \mathbb{R}^N) &
\end{tikzcd}
$$

Fig. 8. Construction of the operators P and P' in Eqs. (70) and (72).

some advantages that, however, come at the price of a higher computational cost. The function T will be referred to as the *kernel of the bracket* and, in general, it depends on both P and $\mathcal{H}$, because of (69), but for simplicity, this dependence is not explicitly indicated in the notation.

There are of course many ways to choose T such that condition (69) is satisfied and we give two particularly relevant examples below. Among the various choices, we are interested in those that satisfy:

$$\|w\|_W \leq C_P \|Pw\|_{\tilde{W}}, \qquad w \in \mathrm{dom}(P), \tag{73a}$$

$$\tilde{w} \in \ker \mathsf{T}(u) \cap \operatorname{ran} P \iff \tilde{w} = \lambda P \frac{\delta \mathcal{H}(u)}{\delta u}, \quad \text{for } \lambda \in \mathbb{R} \text{ constant}, \tag{73b}$$

for a constant C_P. Here $\ker \mathsf{T}(u)$ and $\operatorname{ran} P$ denote the null space of the operator $\mathsf{T}(u)$ and the range of the operator P, respectively, and they are both subspaces of $\tilde{W}$. When $P = \nabla$, Eq. (73a) is the Poincaré inequality [112].

Formally at least, conditions (73) imply that the bracket defined in Eq. (70) is *minimally degenerate* in the sense of Eq. (43). In fact, $(\mathcal{F}, \mathcal{F}) = 0$ is equivalent to

$$P\frac{\delta \mathcal{F}(u)}{\delta u} \in \ker \mathsf{T}(u) \cap \operatorname{ran}(P),$$

and condition (73b) implies that there is a constant λ such that

$$P\frac{\delta \mathcal{F}(u)}{\delta u} = \lambda P \frac{\delta \mathcal{H}(u)}{\delta u}.$$

Then condition (73a) yields

$$\Big\| \frac{\delta \mathcal{F}(u)}{\delta u} - \lambda \frac{\delta \mathcal{H}(u)}{\delta u} \Big\|_W \leq C_P \Big\| P\Big[\frac{\delta \mathcal{F}(u)}{\delta u} - \lambda \frac{\delta \mathcal{H}(u)}{\delta u}\Big] \Big\|_{\tilde{W}} = 0,$$

which implies condition (43) with $\lambda_\alpha \neq 0$ only if $\mathcal{I}^\alpha = \mathcal{H}$.

It follows that metric brackets of the form (70), with the defining operator P and the kernel T satisfying (73) could be used to construct a relaxation method for variational problems of the form (1).

We now give examples of (70) for which conditions (73) are, at least formally, satisfied.

Example 5. Morrison's brackets for the Landau collision operator [1] and its generalization [3] can be obtained as special cases of (70). This motivates our choice of the name *collision-like*.

The configuration space V is the space of particle distribution functions $u(x) = f(\mathsf{x}, \mathsf{v})$, where x is the spatial position and v is the velocity of the particles. We assume that $\mathsf{x} \in \Omega_{\mathsf{x}}$ and $\mathsf{v} \in \Omega_{\mathsf{v}}$, with both domains Ω_{x}, $\Omega_{\mathsf{v}} \subset \mathbb{R}^D$ being bounded, $D \geq 2$, and $f(\mathsf{x}, \cdot)$ compactly supported in Ω_{v}, i.e., the velocity-space domain is large enough to contain all particle velocities. Then, $\Omega = \Omega_{\mathsf{x}} \times \Omega_{\mathsf{v}}$, $d = 2D$, and $N = 1$, since f is a scalar field. The measure on Ω is the Lebesgue measure $d\mu(x) = d\mathsf{x}d\mathsf{v}$.

We choose $\mathcal{O} = \Omega \times \Omega_{\mathsf{v}}$, $n = 3D$, $d\nu(\mathsf{x}, \mathsf{v}, \mathsf{v}') = d\mathsf{x}d\mathsf{v}d\mathsf{v}'$,

$$Pg(\mathsf{x}, \mathsf{v}, \mathsf{v}') = \nabla_{\mathsf{v}} g(\mathsf{x}, \mathsf{v}) - \nabla_{\mathsf{v}'} g(\mathsf{x}, \mathsf{v}'),$$

with $\mathrm{dom}\, P$ given by the functions $g \in L^2(\Omega)$ such that $\nabla_{\mathsf{v}} g \in L^2(\Omega; \mathbb{R}^D)$; hence $Pg(\mathsf{x}, \mathsf{v}, \mathsf{v}') \in \mathbb{R}^{\tilde{N}}$ with $\tilde{N} = D$. The Hamiltonian is

$$\mathcal{H}(f) = \int_\Omega \big[\frac{1}{2} m \mathsf{v}^2 + \mathrm{V}(\mathsf{x})\big] f(\mathsf{x}, \mathsf{v}) d\mathsf{x} d\mathsf{v},$$

where m is the mass of the considered particle species and V is a potential energy. We have $P(\delta\mathcal{H}(f)/\delta f) = m(\mathsf{v} - \mathsf{v}')$. Therefore, a possible choice of the operator $\mathsf{T}(u) = \mathsf{T}(f)$ is the multiplication by the matrix-valued kernel

$$\mathsf{T}_{\mathrm{L}}(f; \mathsf{x}, \mathsf{v}, \mathsf{v}') = \frac{\nu_c}{2} M\big(f(\mathsf{x}, \mathsf{v})\big) M\big(f(\mathsf{x}, \mathsf{v}')\big) U_{\mathrm{L}}(\mathsf{v} - \mathsf{v}'), \tag{74}$$

where $\nu_c > 0$ is a constant collision frequency, $M : \mathbb{R}_+ \to \mathbb{R}_+$ is arbitrary, and

$$U_{\mathrm{L}}(\mathsf{v} - \mathsf{v}') := \frac{1}{|\mathsf{v} - \mathsf{v}'|}\Big(I - \frac{(\mathsf{v} - \mathsf{v}') \otimes (\mathsf{v} - \mathsf{v}')}{|\mathsf{v} - \mathsf{v}'|^2}\Big).$$

Condition (69) holds, and bracket (70) reduces to

$$(\mathcal{F},\mathcal{G}) = \frac{\nu_c}{2}\int_{\mathcal{O}} M\big(f(\mathrm{x},\mathrm{v})\big)M\big(f(\mathrm{x},\mathrm{v}')\big)\Big[\nabla_{\mathrm{v}}\frac{\delta\mathcal{F}(f)}{\delta f}(\mathrm{x},\mathrm{v}) - \nabla_{\mathrm{v}'}\frac{\delta\mathcal{F}(f)}{\delta f}(\mathrm{x},\mathrm{v}')\Big] \cdot U_{\mathrm{L}}(\mathrm{v}-\mathrm{v}')\Big[\nabla_{\mathrm{v}}\frac{\delta\mathcal{G}(f)}{\delta f}(\mathrm{x},\mathrm{v}) - \nabla_{\mathrm{v}'}\frac{\delta\mathcal{G}(f)}{\delta f}(\mathrm{x},\mathrm{v}')\Big]d\mathrm{x}d\mathrm{v}d\mathrm{v}', \tag{75}$$

which is the bracket of Eq. (44) in Morrison's paper [3]. One can notice that in Eq. (74) the orthogonal projection onto the direction of

$$m(\mathrm{v}-\mathrm{v}') = \nabla_{\mathrm{v}}\frac{\delta\mathcal{H}(f)}{\delta f}(\mathrm{x},\mathrm{v}) - \nabla_{\mathrm{v}'}\frac{\delta\mathcal{H}(f)}{\delta f}(\mathrm{x},\mathrm{v}')$$

ensures property (69). In most examples considered in this paper, the kernel of the bracket is constructed from a similar projector. We also note that, in the construction given here, there is no need to use a singular distributional kernel, cf. the Dirac's delta in Eq. (45) of [3].

The operator P' defined in Eq. (71) acting on a function $\tilde{w}(z) = \tilde{g}(\mathrm{x},\mathrm{v},\mathrm{v}')$ can be formally computed after integration by parts, with the result that

$$P'\tilde{g}(\mathrm{x},\mathrm{v}) = -\operatorname{div}_{\mathrm{v}}\Big[\int_{\Omega_{\mathrm{v}}}\big(\tilde{g}(\mathrm{x},\mathrm{v},\mathrm{w}) - \tilde{g}(\mathrm{x},\mathrm{w},\mathrm{v})\big)d\mathrm{w},\Big],$$

and the evolution Eq. (72) takes the form

$$\partial_t f = \operatorname{div}_{\mathrm{v}}\Big[\nu_c\int_{\Omega_{\mathrm{v}}} M\big(f(\mathrm{x},\mathrm{v})\big)M\big(f(\mathrm{x},\mathrm{w})\big) \times U_{\mathrm{L}}(\mathrm{v}-\mathrm{w})\Big(\nabla_{\mathrm{v}}\frac{\delta\mathcal{S}(f)}{\delta f}(\mathrm{x},\mathrm{v}) - \nabla_{\mathrm{w}}\frac{\delta\mathcal{S}(f)}{\delta f}(\mathrm{x},\mathrm{w})\Big)d\mathrm{w}\Big],$$

which reduces to the Landau operator for Coulomb collisions when $\mathcal{S}(f) = \int f\log f\, d\mathrm{x}d\mathrm{v}$ and $M(f) = f$.

As for conditions (73), given $g(\mathrm{x},\mathrm{v})$ in the domain of the operator P, we have

$$\int_{\Omega_{\mathrm{v}}}\nabla_{\mathrm{v}} g(\mathrm{x},\mathrm{v})d\mathrm{v} = 0,$$

provided that $g(\mathrm{x},\cdot)$ vanishes near the boundary of Ω_{v}. Then,

$$\|Pg\|^2_{\tilde{W}} = 2|\Omega_{\mathrm{v}}|\int_{\Omega}\big|\nabla_{\mathrm{v}} g(\mathrm{x},\mathrm{v})\big|^2 d\mathrm{x}d\mathrm{v},$$

where $|\Omega_{\mathrm{v}}| = \int_{\Omega_{\mathrm{v}}} d\mathrm{v}'$. The minimum of $\|Pg\|^2_{\tilde{W}}$ subject to the constraint $\|g\|_W = 1$ is attained for

$$-\Delta_{\mathrm{v}} g = \lambda_0 g, \qquad \int_{\Omega}|g(\mathrm{x},\mathrm{v})|^2 d\mathrm{x}d\mathrm{v} = 1,$$

where λ_0 is the minimum eigenvalue of the Laplace operator Δ_{v} on the velocity domain Ω_{v} with homogeneous Dirichlet boundary conditions. Hence

$$\|Pg\|^2_{\tilde{W}} = 2|\Omega_{\mathrm{v}}|\int_{\Omega}\big|\nabla_{\mathrm{v}} g(\mathrm{x},\mathrm{v})\big|^2 d\mathrm{x}d\mathrm{v} \geq 2|\Omega_{\mathrm{v}}|\lambda_0\|g\|^2_W,$$

which shows (formally) that condition (73a) holds. We shall use this type of argument to study the brackets considered below, for which assuming homogeneous Dirichlet boundary conditions is natural. For the Landau collision operator, however, there is no physical reason to assume that $g = 0$ on the boundary of Ω_{v}. If we drop this unphysical requirement, then P no longer satisfies condition (73a) as it has a nontrivial null space, $\ker P$, given by the functions $g \in \operatorname{dom} P$ such that

$$\nabla_{\mathrm{v}} g(\mathrm{x},\mathrm{v}) - \nabla_{\mathrm{v}'} g(\mathrm{x},\mathrm{v}') = 0.$$

More specifically, the velocity gradient of an element of $\ker P$ is necessarily constant in velocity v for almost all x, and thus

$$g \in \ker P \iff g(\mathrm{x},\mathrm{v}) = a(\mathrm{x}) + b(\mathrm{x})\cdot\mathrm{v},$$

where $a : \Omega_{\mathrm{x}} \to \mathbb{R}$ and $b : \Omega_{\mathrm{x}} \to \mathbb{R}^D$ are arbitrary functions. As for condition (73b), a function $\tilde{g}(\mathrm{x},\mathrm{v},\mathrm{v}')$ belongs to $\ker \mathsf{T}_{\mathrm{L}}$ when

$$\tilde{g}(\mathrm{x},\mathrm{v},\mathrm{v}') = \Lambda(\mathrm{x},\mathrm{v},\mathrm{v}')\big(\mathrm{v}-\mathrm{v}'\big),$$

almost everywhere (a.e.) in $\mathcal{O}$, where $\Lambda(\mathrm{x},\mathrm{v},\mathrm{v}') \in \mathbb{R}$. If, in addition, $\tilde{g} \in \operatorname{ran}(P)$, we must have

$$\nabla_{\mathrm{v}} g(\mathrm{x},\mathrm{v}) - \nabla_{\mathrm{v}'} g(\mathrm{x},\mathrm{v}') = \Lambda(\mathrm{x},\mathrm{v},\mathrm{v}')\big(\mathrm{v}-\mathrm{v}'\big), \quad \text{a.e. in } \mathcal{O}.$$

If g and Λ satisfying this condition exist, then necessarily $\Lambda(\mathrm{x},\mathrm{v},\mathrm{v}') = \Lambda(\mathrm{x},\mathrm{v}',\mathrm{v})$ and, upon fixing a point $\mathrm{v}' = \mathrm{a} \in \Omega_{\mathrm{v}}$,

$$\nabla_{\mathrm{v}} g(\mathrm{x},\mathrm{v}) = \nabla_{\mathrm{v}'} g(\mathrm{x},\mathrm{a}) + \Lambda(\mathrm{x},\mathrm{v},\mathrm{a})(\mathrm{v}-\mathrm{a}),$$

and

$$\nabla_{\mathrm{v}} g(\mathrm{x},\mathrm{v}) - \nabla_{\mathrm{v}'} g(\mathrm{x},\mathrm{v}') = \Lambda(\mathrm{x},\mathrm{v},\mathrm{a})(\mathrm{v}-\mathrm{a}) - \Lambda(\mathrm{x},\mathrm{v}',\mathrm{a})(\mathrm{v}'-\mathrm{a})$$

$$= \Lambda(\mathsf{x},\mathsf{v},\mathsf{v}')\big((\mathsf{v}-\mathsf{a})-(\mathsf{v}'-\mathsf{a})\big).$$

Then we must have

$$\big[\Lambda(\mathsf{x},\mathsf{v},\mathsf{v}')-\Lambda(\mathsf{x},\mathsf{v},\mathsf{a})\big](\mathsf{v}-\mathsf{a})-\big[\Lambda(\mathsf{x},\mathsf{v},\mathsf{v}')-\Lambda(\mathsf{x},\mathsf{a},\mathsf{v}')\big](\mathsf{v}'-\mathsf{a})=0.$$

The two vectors $\mathsf{v}-\mathsf{a}$ and $\mathsf{v}'-\mathsf{a}$ in $\mathbb{R}^D$ are linearly dependent only if they are proportional, which can only happen for a set of points $(\mathsf{x},\mathsf{v},\mathsf{v}')$ of measure zero in $\mathcal{O}$. Hence, we deduce that Λ must satisfy the necessary conditions

$$\begin{cases} \Lambda(\mathsf{x},\mathsf{v},\mathsf{v}')-\Lambda(\mathsf{x},\mathsf{v},\mathsf{a})=0, \\ \Lambda(\mathsf{x},\mathsf{v},\mathsf{v}')-\Lambda(\mathsf{x},\mathsf{a},\mathsf{v}')=0, \end{cases} \quad \text{a.e. in } \mathcal{O},$$

and this must hold for almost any choice of the arbitrary point a. This is possible only if there is a function $mc(\mathsf{x})$ of position x only (we factor out the mass m for convenience) such that

$$\Lambda(\mathsf{x},\mathsf{v},\mathsf{v}')=mc(\mathsf{x})= \text{ a.e. in } \mathcal{O}.$$

We deduce that a function $\tilde{g}\in\ker T(f)\cap\operatorname{ran}P$ must be of the form

$$\tilde{g}(\mathsf{x},\mathsf{v},\mathsf{v}')=mc(\mathsf{x})(\mathsf{v}-\mathsf{v}')=c(\mathsf{x})P\frac{\delta\mathcal{H}(f)}{\delta f}(\mathsf{x},\mathsf{v},\mathsf{v}').$$

This in not exactly Eq. (73b), since c is not a constant on the whole extended domain $\mathcal{O}$, but only in velocity space. This residual dependence of x should be expected since the collision operator only acts in velocity space, pointwise in x. Therefore, conditions (73) are not satisfied for this bracket. We can however obtain a complete characterization of the null space of this bracket by using the results on $T(f)$ and P obtained above. In fact, we have

$$\big(\mathcal{F},\mathcal{F}\big)(f)=0 \iff P\frac{\delta\mathcal{F}(f)}{\delta f}\in\ker T(f),$$

hence

$$P\frac{\delta\mathcal{F}(f)}{\delta f}=mc(\mathsf{x})(\mathsf{v}-\mathsf{v}')=c(\mathsf{x})P\frac{\delta\mathcal{H}(f)}{\delta f},$$

and, since c is independent on (v,v'), we have

$$P\Big(\frac{\delta\mathcal{F}(f)}{\delta f}-c\frac{\delta\mathcal{H}(f)}{\delta f}\Big)=0.$$

The elements of the null space of P have been computed above, and can conclude that

$$\big(\mathcal{F},\mathcal{F}\big)(f)=0 \iff \frac{\delta\mathcal{F}(f)}{\delta f}=a\frac{\delta\mathcal{N}(f)}{\delta f}+b\cdot\frac{\delta\mathcal{P}(f)}{\delta f}+c\frac{\delta\mathcal{H}(f)}{\delta f},$$

where a, b, and c are functions of x only and

$$\mathcal{N}(f)=\int_\Omega f(\mathsf{x},\mathsf{v})d\mathsf{x}d\mathsf{v},\quad \mathcal{P}(f)=\int_\Omega f(\mathsf{x},\mathsf{v})\mathsf{v}d\mathsf{x}d\mathsf{v},$$

together with $\mathcal{H}(f)$ constitute the three collision invariants [113,114], namely, the total particle number, momentum (per unit mass), and energy. In summary, the null space of Morrison's bracket is spanned by a linear combination of the derivatives of the three collision invariants, but with coefficients depending on x. Therefore the bracket is not minimally degenerate, since energy is not the only invariant, and it is not even specifically degenerate since the coefficients a, b, and c are functions of space. This is due to the well known fact that the collision operator acts on velocity space only, and the relaxation of the full distribution function $f(\mathsf{x},\mathsf{v})$ in both space and velocity requires the interaction of the collision operator with the ideal phase-space transport dynamics [115]. This bracket does become specifically degenerate with respect to the three invariants if we restrict the phase space to functions of velocity only.

Example 6. The bracket based on the L^2-orthogonal projector discussed in Section 4.2 can be obtained as a special case of (70). In fact, we can utilize Eq. (70) in order to generalize (64) to the case of $\mathbb{R}^N$-valued fields.

With this aim, let $\Omega\subset\mathbb{R}^d$, $N\in\mathbb{N}$, $\mathcal{O}=\Omega\times\Omega$, and $\tilde{N}=2N$, that is, we double both the dimension of the domain and of the field. Then, with $\Phi=W=L^2(\Omega,\mu;\mathbb{R}^N)=\Phi'$, $\tilde{W}=L^2(\mathcal{O},\nu;\mathbb{R}^{\tilde{N}})$, and $d\nu(x,x')=d\mu(x)d\mu(x')$, we define

$$P:W\to\tilde{W},\qquad Pw(x,x')=\begin{pmatrix} w(x)\\ w(x')\end{pmatrix}\in\mathbb{R}^{2N}.$$

Given a Hamiltonian function $\mathcal{H}(u)$ with $\delta\mathcal{H}(u)/\delta u\in W$, we choose $\mathsf{T}(u)$ to be the multiplication operator by the kernel

$$\mathsf{T}(u;x,x')=\kappa(u)\begin{pmatrix} |h(x')|^2 I_N & -h(x)\otimes h(x')\\ -h(x')\otimes h(x) & |h(x)|^2 I_N\end{pmatrix},$$

where I_N is the $N\times N$ identity block, and $h=\delta\mathcal{H}(u)/\delta u$ is a short-hand notation for the functional derivative (h may still depend on u). One can check that condition (69) is satisfied. Inserting these choices into Eq. (70) yields

$$(\mathcal{F},\mathcal{G})=2\kappa(u)\int_{\mathcal{O}}\frac{\delta\mathcal{F}(u)}{\delta u}(x)\cdot\Big[\Big|\frac{\delta\mathcal{H}(u)}{\delta u}(x')\Big|^2\frac{\delta\mathcal{G}(u)}{\delta u}(x)$$

$$- \frac{\delta \mathcal{H}(u)}{\delta u}(x)\Big(\frac{\delta \mathcal{H}(u)}{\delta u}(x') \cdot \frac{\delta \mathcal{G}(u)}{\delta u}(x')\Big)\Big] d\mu(x) d\mu(x')$$
$$= 2\kappa(u) \Big\| \frac{\delta \mathcal{H}(u)}{\delta u} \Big\|_{L^2}^2 \int_\Omega \frac{\delta \mathcal{F}(u)}{\delta u}(x) \cdot \Pi_{\mathcal{H}} \frac{\delta \mathcal{G}(u)}{\delta u}(x) d\mu(x),$$

where $\Pi_{\mathcal{H}}$ is the L^2-orthogonal projector defined in Eq. (63), but generalized to the case of $\mathbb{R}^N$-valued fields. If 2κ is set to the inverse of $\|h\|_{L^2}^2$, this reduces to (64) when $N = 1$.

Condition (73a) follows from

$$\|Pw\|^2_{L^2(\mathcal{O},\nu;\mathbb{R}^{2N})} = \int_{\Omega\times\Omega} \big(|w(x)|^2 + |w(x')|^2\big) d\mu(x) d\mu(x') = 2\mu(\Omega) \|w\|^2_{L^2(\Omega,\mu;\mathbb{R}^N)},$$

where $\mu(\Omega) = \int_\Omega d\mu$. As for condition (73b), if $h \neq 0$, $\tilde{w} \in \ker \mathsf{T}(u) \cap \operatorname{ran}(P)$ implies that there is $w \in W$ such that $\tilde{w} = Pw$, and

$$|h(x')|^2 w(x) - h(x)\big(h(x') \cdot w(x')\big) = 0 \in \mathbb{R}^N,$$

for almost all (x, x'). Upon multiplying by $h(x)$, we have

$$|h(x')|^2 \big(h(x) \cdot w(x)\big) = |h(x)|^2 \big(h(x') \cdot w(x')\big).$$

This is equivalent to saying that $\big(-|h(x')|^2, |h(x)|^2\big)$ and $\big(h(x)\cdot w(x), h(x')\cdot w(x')\big)$ are orthogonal in $\mathbb{R}^2$, or that

$$\begin{pmatrix} h(x) \cdot w(x) \\ h(x') \cdot w(x') \end{pmatrix} = \Lambda(x, x') \begin{pmatrix} |h(x)|^2 \\ |h(x')|^2 \end{pmatrix},$$

for a function Λ, which in general depends on (x, x') and this holds for almost all (x, x'). We deduce that Λ must be constant almost everywhere in $\mathcal{O}$, i.e. $\Lambda(x, x') = \lambda$ for almost all (x, x') and thus

$$\lambda |h(x)|^2 = h(x) \cdot w(x) \iff w(x) = \lambda h(x),$$

which implies condition (73b). Hence, metriplectic brackets constructed by means of an L^2-orthogonal projection are minimally degenerate as already shown in Section 4.2 for the scalar case ($N = 1$).

Example 7. As in Example 6, let $\mathcal{O} = \Omega \times \Omega$, $n = 2d$, with coordinates $z = (x, x') \in \Omega \times \Omega$, and $d\nu(x, x') = d\mu(x) d\mu(x')$. The operator P is given by

$$Pw(x, x') = P_L w(x, x') = Lw(x) - Lw(x'),$$

where $L : W \to L^2(\Omega, \mu; \mathbb{R}^{\tilde{N}})$ is possibly an unbounded linear operator with domain $\operatorname{dom}(L) = \Phi \subseteq W$ and taking values in the space of $\mathbb{R}^{\tilde{N}}$-valued, squared-integral functions, with $\tilde{N} \geq 2$. In general, we allow $\tilde{N}$ to be different from N. Specific cases will be considered in Sections 5.2 and 5.3 below.

As for the kernel, a way to satisfy condition (69) utilizes the matrix

$$Q_k(z) := |z|^2 I_k - z \otimes z, \qquad z \in \mathbb{R}^k, \tag{76}$$

where $k \in \mathbb{N}$ and I_k is the $k \times k$ identity block. Specifically, let $\mathsf{T}(u)$ be the multiplication operator by the matrix

$$\mathsf{T}(u; x, x') = \frac{1}{2} \kappa(u; x, x') Q_{\tilde{N}} \Big(P_L \frac{\delta \mathcal{H}(u)}{\delta u}(x, x') \Big), \tag{77}$$

where $\kappa(u; \cdot, \cdot)$ is a positive scalar weight function satisfying the symmetry condition $\kappa(u; x, x') = \kappa(u; x', x)$. (This choice of the kernel is trivial if $\tilde{N} = 1$, which is excluded.) With the foregoing choices, Eq. (70) becomes

$$\begin{aligned}(\mathcal{F}, \mathcal{G}) := &\frac{1}{2} \int_\Omega \int_\Omega \kappa(u; x, x') \Big[P_L \frac{\delta \mathcal{F}(u)}{\delta u}(x, x') \Big] \\ &\cdot Q_{\tilde{N}} \Big(P_L \frac{\delta \mathcal{H}(u)}{\delta u}(x, x') \Big) \Big[P_L \frac{\delta \mathcal{G}(u)}{\delta u}(x, x') \Big] d\mu(x) d\mu(x').\end{aligned} \tag{78}$$

This is the form of collision-like bracket as originally formulated by Bressan et al. [110,111]. The dual operator P' can be determined in terms of the dual of L, which is the linear operator $L' : L^2(\Omega, \mu; \mathbb{R}^{\tilde{N}}) \to \Phi'$ defined for all $\varpi \in L^2(\Omega, \mu; \mathbb{R}^{\tilde{N}})$ by

$$\langle w, L' \varpi \rangle = \int_\Omega \varpi(x) \cdot Lw(x) d\mu(x), \qquad \text{for all } w \in \Phi, \tag{79}$$

and we find

$$\begin{aligned}\int_{\mathcal{O}} \tilde{w}(x, x') \cdot P_L w(x, x') d\nu(x, x') &= \int_\Omega Lw(x) \int_\Omega \big[\tilde{w}(x, x') - \tilde{w}(x', x)\big] d\mu(x') d\mu(x) \\ &= \Big\langle w, L' \Big[\int_\Omega \big[\tilde{w}(\cdot, x') - \tilde{w}(x', \cdot)\big] d\mu(x') \Big] \Big\rangle \\ &= \big\langle w, P_L' \tilde{w} \big\rangle.\end{aligned}$$

In terms of the operator L', the evolution Eq. (72) reads

$$\partial_t u = -L' \Big[\mathbb{D}(u) L \frac{\delta \mathcal{S}(u)}{\delta u} - \mathbb{F}\Big(u, \frac{\delta \mathcal{S}(u)}{\delta u}\Big) \Big], \tag{80}$$

where we have defined

$$\mathbb{D}(u;x) := \int_\Omega \kappa(u;x,x')Q_{\tilde N}\left(P_L \frac{\delta\mathcal{H}(u)}{\delta u}(x,x')\right)d\mu(x'), \tag{81a}$$

$$\mathbb{F}(u,v;x) := \int_\Omega \kappa(u;x,x')Q_{\tilde N}\left(P_L \frac{\delta\mathcal{H}(u)}{\delta u}(x,x')\right)Lv(x')d\mu(x'). \tag{81b}$$

Eq. (80) has the same structure as the Landau operator for Coulomb collisions discussed in Example 5, with ∇_{v} being replaced by L.

At last, we check conditions (73). Since $d\mu(x) = m(x)dx$, with m continuous on the closed domain $\overline{\Omega}$, and $d\nu(x,x') = d\mu(x)d\mu(x')$, we have

$$\|Pw\|^2_{\tilde W} = \int_{\mathcal{O}} \left|Lw(x) - Lw(x')\right|^2 d\nu(x,x') \geq m_0^2 \int_{\mathcal{O}} \left|Lw(x) - Lw(x')\right|^2 dx dx',$$

where $m_0 = \min\{m(x) : x \in \overline{\Omega}\}$. For condition (73a) to hold, it is therefore sufficient that L satisfies

$$\|w\|_W \leq C_L \|Lw\|_{L^2(\Omega)}, \qquad \int_\Omega Lw(x)dx = 0, \tag{82}$$

so that

$$\int_{\mathcal{O}} \left|Lw(x) - Lw(x')\right|^2 dx dx' = 2|\Omega| \int_\Omega \left|Lw(x)\right|^2 dx - 2\left|\int_\Omega Lw(x)dx\right|^2 \geq 2|\Omega| C_L^{-2}\|w\|^2_W,$$

which give Eq. (73a). All considered cases in the applications below satisfy condition (82).

As for the kernel of the bracket, we argue as in the case of Example 5. With the kernel given in Eq. (77), a function $\tilde w \in \tilde W$ belongs to $\ker \mathsf{T}(u) \cap \operatorname{ran}(P)$ only if there is a function $w \in W$ and $\Lambda : \Omega \times \Omega \to \mathbb{R}$, such that

$$Lw(x) - Lw(x') = \Lambda(x,x')\left[Lh(x) - Lh(x')\right],$$

where $h = \delta\mathcal{H}(u)/\delta u$ and necessarily $\Lambda(x,x') = \Lambda(x',x)$ for $x \neq x'$. If we fix an arbitrary point $x' = a \in \Omega$, we obtain an explicit expression for $Lw(x)$,

$$Lw(x) = Lw(a) + \Lambda(x,a)X_a(x), \qquad X_a(x) := Lh(x) - Lh(a).$$

Therefore

$$\begin{aligned} Lw(x) - Lw(x') &= \Lambda(x,a)X_a(x) - \Lambda(a,x')X_a(x') \\ &= \Lambda(x,x')\left[X_a(x) - X_a(x')\right], \end{aligned}$$

or equivalently

$$\left[\Lambda(x,x') - \Lambda(x,a)\right]X_a(x) - \left[\Lambda(x,x') - \Lambda(a,x')\right]X_a(x') = 0.$$

Under the assumption that $X_a(x)$ and $X_a(x')$ are linearly independent for almost all $(x,x') \in \Omega\times\Omega = \mathcal{O}$, we deduce $\Lambda(x,x') = \lambda =$ constant as in Example 5. Hence, if the operator L satisfies (82) and the Hamiltonian is such that $X_a(x)$ and $X_a(x')$ are linearly independent almost everywhere in $\Omega\times\Omega$, the bracket (78) is minimally degenerate.

In general we observe that the brackets of the form discussed in Example 7 are all special cases of the metriplectic 4-bracket introduced recently by Morrison and Updike [35], since the kernel depends quadratically on $\delta\mathcal{H}/\delta u$. Specifically, they can be obtained from the metriplectic 4-bracket constructed by utilizing the Kulkarni-Nomizu product as discussed in Section III.D.1 of Ref. [35].

Example 7 will be used as a template for the construction of dissipative operators for the solution of some of the variational problems introduced in Section 2.2. Before addressing the applications, we specialize bracket (78) for two relevant choices of the operator L.

5.2. Collision-like brackets based on div–grad operators

In Eq. (78), let us consider the case of a scalar field ($N = 1$) and choose

$$Lw = \nabla w,$$

with domain $\operatorname{dom}(L) = H^1_0(\Omega)$, the space of functions in $L^2(\Omega)$ with weak derivatives also in $L^2(\Omega)$ and satisfying homogeneous Dirichlet boundary conditions on $\partial\Omega$. This operator satisfies condition (82), and $\tilde N = d$ (the inequality, in particular, is the classical Poincaré inequality for H^1_0, cf. Theorem 3 in Section 5.6.1 of Ref. [112]). The dual operator L' can be obtained from Eq. (79). Given $\varpi \in L^2(\Omega,\mu;\mathbb{R}^d)$, sufficiently regular, we have

$$\langle w, L'\varpi\rangle = \int_\Omega \varpi(x)\cdot\nabla w(x)d\mu(x) = -\int_\Omega w(x)\operatorname{div}_\mu \varpi(x)d\mu(x),$$

and thus $-L'\varpi = \operatorname{div}_\mu \varpi = \frac{1}{m}\sum_i \partial_{x_i}\left[m\varpi^i\right]$ is the divergence operator associated to the volume form $d\mu(x) = m(x)dx$. For a generic ϖ, we write $L' = -\widetilde{\operatorname{div}}_\mu$ to denote that the divergence is defined weakly. As for the kernel, we choose

$$\kappa(u;x,x') = M\left(x,u(x)\right)M\left(x',u(x')\right),$$

where $M(x, y) > 0$ is a positive function over $\Omega \times \mathbb{R}$, which can be used to simplify the bracket [3] as in Example 5.

As for the entropy, we shall restrict to functions of the form

$$\mathcal{S}(u) = \int_\Omega s\big(x, u(x)\big)\,dx, \tag{83}$$

where $s(x, y)$ is a given profile, convex in y, and possibly depending on position x. Then

$$\frac{\delta \mathcal{S}(u)}{\delta u}(x) = \partial_y s\big(x, u(x)\big),$$

and

$$\nabla \frac{\delta \mathcal{S}(u)}{\delta u}(x) = \partial_y^2 s\big(x, u(x)\big)\nabla u(x) + \partial_x \partial_y s\big(x, u(x)\big).$$

The arbitrary function M can be chosen so that [3]

$$M(x, y)\partial_y^2 s(x, y) = 1, \tag{84}$$

under the condition $\partial_y^2 s(x, y) > 0$, consistently with the convexity assumption for the profile. Eq. (84) introduces a dependence of the kernel of the bracket on the entropy function. With this choice, Eq. (80) becomes

$$\partial_t u = \widetilde{\operatorname{div}}_\mu\Big[\mathbb{D}_s(u)\big(\nabla u + M(x, u)\partial_x\partial_y s(x, u)\big) - M(x, u)\mathbb{F}_s(u, \nabla u)\Big], \tag{85}$$

and

$$\begin{aligned}
\mathbb{D}_s(u; x) &= \int_\Omega Q_2\Big(\nabla \frac{\delta \mathcal{H}(u)}{\delta u}(x) - \nabla \frac{\delta \mathcal{H}(u)}{\delta u}(x')\Big) M\big(x', u(x')\big)\,d\mu(x'),\\
\mathbb{F}_s(u; x) &= \int_\Omega Q_2\Big(\nabla \frac{\delta \mathcal{H}(u)}{\delta u}(x) - \nabla \frac{\delta \mathcal{H}(u)}{\delta u}(x')\Big)\\
&\qquad \big[\nabla u(x') + M\big(x', u(x')\big)\partial_x\partial_y s\big(x', u(x')\big)\big]\,d\mu(x').
\end{aligned}$$

Eq. (85) still depends on the choice of the measure μ, the entropy profile s, and the Hamiltonian function $\mathcal{H}$. This choices will be specified below separately for the cases of the reduced Euler and Grad-Shafranov equations in two-dimensions ($d = \tilde{N} = 2$).

5.3. Collision-like brackets based on curl–curl operators

With magnetic fields in mind, we consider an example of bracket (70) for divergence-free fields. Then, Ω is bounded, simply connected domain in $\mathbb{R}^3$ with sufficiently regular connected boundary, $d = N = 3$, and $d\mu(x) = dx$. We choose

$$Lw = \operatorname{curl} w,$$

with domain

$$\Phi = \{w \in H(\operatorname{curl}, \Omega) \cap H(\operatorname{div}, \Omega) :\ \operatorname{div} w = 0 \text{ in } \Omega,\ n \times w = 0 \text{ on } \partial\Omega\},$$

where $H(\operatorname{curl}, \Omega)$ and $H(\operatorname{div}, \Omega)$ are the spaces of vector fields $w \in L^2(\Omega; \mathbb{R}^3)$ such that $\operatorname{curl} w \in L^2(\Omega; \mathbb{R}^3)$ and $\operatorname{div} w \in L^2(\Omega)$, respectively. Specifically, $w \in \Phi = \operatorname{dom}(L)$ is a divergence-free field with zero tangential component $n \times w$ on the boundary, and $n : \partial\Omega \to \mathbb{R}^3$ is the outward unit normal on $\partial\Omega$. It follows that condition (82) is satisfied: the inequality amounts to the Poincaré inequality for divergence-free fields on a simply connected domain with connected boundary (cf. Corollary 3.51 in Ref. [116]), while

$$\int_\Omega \operatorname{curl} w\,dx = \int_{\partial\Omega} (n \times w)\,d\sigma = 0,$$

where $d\sigma$ is the surface element on $\partial\Omega$.

If $\varpi \in L^2(\Omega; \mathbb{R}^3)$ is sufficiently regular, Eq. (79) and integration by parts gives

$$\langle w, L'\varpi\rangle = \int_\Omega \varpi(x)\cdot \operatorname{curl} w(x)\,dx = \int_\Omega w(x)\cdot \operatorname{curl}\varpi(x)\,dx,$$

hence $L'\varpi = \operatorname{curl}\varpi$, while for a generic ϖ, $L' = \widetilde{\operatorname{curl}}$ is the weak curl operator.

For the kernel (77), we choose $\kappa = 1$, and the corresponding evolution equation can be written as

$$\partial_t u = -\widetilde{\operatorname{curl}}\Big[\mathbb{D}_v(u)\operatorname{curl}\frac{\delta \mathcal{S}(u)}{\delta u} - \mathbb{F}_v\Big(u, \frac{\delta \mathcal{S}(u)}{\delta u}\Big)\Big], \tag{86}$$

where

$$\begin{aligned}
\mathbb{D}_v(u; x) &= \int_\Omega Q_3\Big(\operatorname{curl}\frac{\delta \mathcal{H}(u)}{\delta u}(x) - \operatorname{curl}\frac{\delta \mathcal{H}(u)}{\delta u}(x')\Big)\,dx',\\
\mathbb{F}_v(u, v; x) &= \int_\Omega Q_3\Big(\operatorname{curl}\frac{\delta \mathcal{H}(u)}{\delta u}(x) - \operatorname{curl}\frac{\delta \mathcal{H}(u)}{\delta u}(x')\Big)\operatorname{curl} v(x')\,dx'.
\end{aligned}$$

A non-trivial factor κ of the form used in Section 5.2 can easily be accounted for [110].

The evolution Eq. (86) preserves the divergence-free constraint. In fact, for any $\varphi \in H_0^1(\Omega)$ the function $\mathcal{F}_\varphi(u) = \int_\Omega u \cdot \nabla\varphi\,dx = -\int_\Omega \varphi \operatorname{div} u\,dx$ is a constant of motion.

5.4. Application to the reduced Euler equations

In order to demonstrate the properties of the collision-like brackets, we construct a relaxation method for the solution of the variational principle (13) for equilibria of the reduced Euler equations, with entropy and Hamiltonian functions given in Eq. (14).

This is the same problem addressed in Section 4, with the difference here being we consider a bounded domain Ω with homogeneous Dirichlet boundary conditions as discussed in Section 2.2.1. Specifically, the domain is the unit square $\overline{\Omega} = [0,1] \times [0,1]$.

When, in Eq. (14), $s(\omega) = \omega^2/2$, we known from Section 2.2.1 that the solution of the variational principle (13) is related to the eigenfunction of the Laplacian operator on Ω corresponding to the smallest eigenvalue. On the unit square with homogeneous Dirichlet boundary conditions, the eigenfunctions and the corresponding eigenvalues are given by

$$\phi_{m,n}(x) = \mathcal{N}_{m,n} \sin(m\pi x_1)\sin(n\pi x_2), \qquad \lambda_{m,n} = \pi^2(m^2+n^2),$$

labeled by two non-zero integers $m, n \in \mathbb{N}$ and normalized by a non-zero constant $\mathcal{N}_{m,n}$. The smallest eigenvalue corresponds to the function $\phi_{1,1}$, and energy conservation allows us to determine the normalization constant $\mathcal{N}_{1,1}$, that is

$$2\mathcal{H}_0 = \int_\Omega \phi\omega dx = \lambda_{1,1}\|\phi_{1,1}\|^2_{L^2(\Omega)} = \lambda_{1,1}\mathcal{N}^2_{1,1}/4,$$

from which we deduce the unique solution of (13) with this entropy, that is

$$\begin{aligned} \phi(x) &= \left(2\sqrt{\mathcal{H}_0}/\pi\right)\sin(\pi x_1)\sin(\pi x_2), \\ \omega(x) &= \lambda_{1,1}\phi(x) = 4\pi\sqrt{\mathcal{H}_0}\sin(\pi x_1)\sin(\pi x_2), \end{aligned} \qquad \lambda_{1,1} = 2\pi^2 \approx 19.7392. \tag{87}$$

More general choices of the entropy density $s(\omega)$ lead to the nonlinear eigenvalue problems of the form (15). Then no analytical solution is known, but an estimate of the eigenvalue λ can be found from energy conservation in a similar way. Upon multiplying by ω Eq. (15) and integrating over Ω, we find

$$\int_\Omega \omega s'(\omega) dx = \lambda \int_\Omega \omega\phi dx = 2\lambda\mathcal{H}_0,$$

from which we can obtain λ, provided that the integral on the left-hand side can be evaluated, e.g. from the numerical solution. Specifically, if $s(\omega) = \omega\log\omega$ we have

$$\lambda = \frac{1}{2\mathcal{H}_0}\big(\mathcal{M}(\omega) + \mathcal{S}(\omega)\big), \tag{88}$$

where, in particular, $\mathcal{M}(\omega) = \int_\Omega \omega dx$. These analytical results can be used to assess the result of the proposed relaxation method.

As for the construction of the method itself, we consider the collision-like metriplectic system of Section 5.2. The state variable $u(t) \in V \subseteq \Phi$ is identified with vorticity, $u(t,x) = \omega(t,x)$, and evolved according to Eq. (85) with $\mathcal{S}$ and $\mathcal{H}$ given in Eq. (14).

The numerical scheme for the solution of Eq. (85) is based on H^1-conforming finite elements, i.e., the discrete solution $u_h(t)$ belongs to the same space $H^1_0(\Omega)$ that contains the phase space V. The scheme preserves the discrete Hamiltonian $\mathcal{H}(u_h)$ (modulo round-off errors) and we use the Crank-Nicolson scheme in time, which preserves the property of monotonic dissipation of entropy in the quadratic case ($s(\omega) = \omega^2/2$). More details on the derivation of the scheme can be found in Ref. [110].

The obtained numerical method has been implemented in the finite-element library FEniCS [117,118], in which the weak form of the operator in Eq. (85) can be directly specified by means of the domain-specific language UFL [119], and discretized by the finite-element component DOLFIN [120,121]. Among the various tests [110], here we discuss in details three cases only.

Single vortex. In the simplest example, the domain $\overline{\Omega} = [0,1]^2$ is discretized by a uniform grid of 64×64 nodes. The vorticity field $u = \omega$ is approximated in the space of second-order Lagrange elements that are available in FEniCS [118]. The entropy density is quadratic, i.e. $s(\omega) = \omega^2/2$, hence the analytical solution of the variational principle is given in Eq. (87). The initial condition is

$$u_0(x) = u_G(x), \ \text{with}\, u_G \text{ defined in Eq. (62)}, \tag{89}$$

and with parameters $x_{0,1} = x_{0,2} = 1/2$, $w_1^2 = 0.01$, $w_2^2 = 0.07$, and $N = 1$. The numerical scheme provably preserves the Hamiltonian, which in this case is the kinetic energy of the fluid, and dissipates monotonically the entropy, independently of the magnitude of the time step.

Fig. 9 shows the initial condition, the final state, and the "scatter plot", which we use to identify functional relations between different fields, cf. the analysis in Section 4. Qualitatively, we see that the initial condition is quite far from an equilibrium of the Euler equations as the contours of the streaming function ϕ and those of vorticity ω are misaligned. Metriplectic relaxation with collision-like brackets yields a solution that appears to be an equilibrium, and from the scatter plot (right-hand-side panel in Fig. 9), one can see that the relaxed state is indeed an equilibrium characterized by the same linear relation of the exact solution (87). Therefore, the collision-like metric bracket appears to have completely relaxed the initial solution in the sense discussed in Section 3. From Fig. 10, one can verify energy conservation and entropy monotonic dissipation.

For this test case, the exact solution of the variational problem (13) has been computed analytically, Eq. (87), and we can evaluate the difference between the relaxed state of the metriplectic system and the solution of the variational problem. Fig. 11 shows that the relaxed state appears to be close to the expected solution of the variational principle.

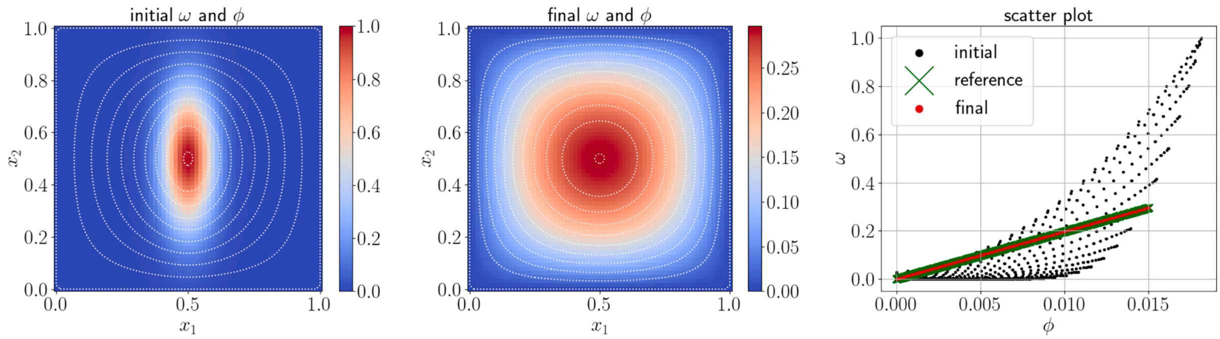

Fig. 9. Relaxation of an initial vortex with initial vorticity given in Eq. (89), according to the metriplectic system (85) with the state variable u being the fluid scalar vorticity ω. The initial condition is shown in the left-hand-side panel: the color map represents the vorticity field ω and the dashed (white) lines the contours of the corresponding streaming function ϕ, cf. Eq. (10). The relaxed state is displayed in the middle panel. The right-hand-side panel represents the functional relation between ω and ϕ for the initial condition (black dots), the final state (red bullets), and the expected linear relation $\omega = \lambda_{1,1}\phi$, cf. Eq. (87), plotted using the numerical solution for ϕ and the analytical value of $\lambda_{1,1}$ (green crosses).

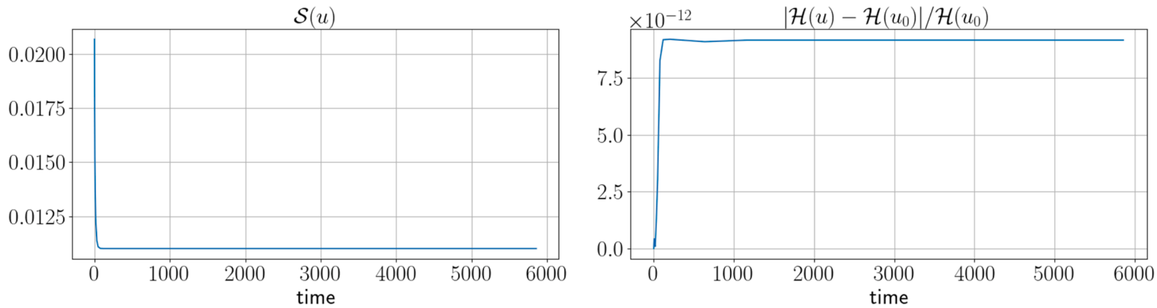

Fig. 10. Evolution of entropy (left-hand-side panel) and of the variation of the Hamiltonian relative to its initial value (right-hand-side panel), for the case in Fig. 9.

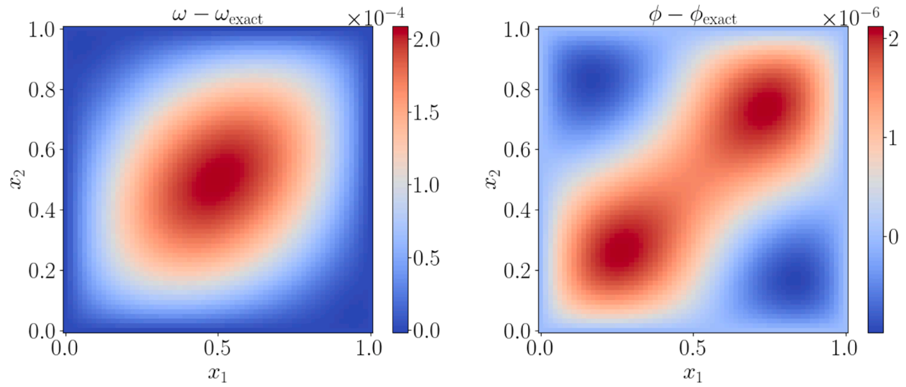

Fig. 11. Difference between the relaxed state and the exact analytical solution (87) of the variational principle (13) for the case in Fig. 9. The difference of the vorticity fields is shown in the left-hand-side panel, while the difference of the corresponding potentials is shown in the right-had-side panel. For the evaluation of the analytical solution (87), the initial energy $\mathcal{H}_0$ has been computed numerically.

Perturbed equilibrium. We repeat the experiment of Fig. 9 with an initial condition close to a critical point of entropy restricted to the constant-Hamiltonian surface. Specifically, the initial condition is

$$u_0(x) = \sin(6\pi x_1)\sin(4\pi x_2) + u_G(x), \text{ with } u_G \text{ defined in Eq. (62)}, \tag{90}$$

and with the same parameters as in Eq. (89) except for $N = 100$.

Fig. 12 shows the initial condition, the final state, and the usual scatter plot, which visualizes the relationship between ω and ϕ. The initial condition (Fig. 12, left-hand-side panel) is basically an eigenfunction of the Laplace operator corresponding to a large eigenvalue, the perturbation being hardly visible. This is confirmed by the scatter plot (Fig. 12 right-hand-side panel, black dots) in which the initial state is concentrated on a straight line, with a small spread due to the Gaussian perturbation. Therefore the initial

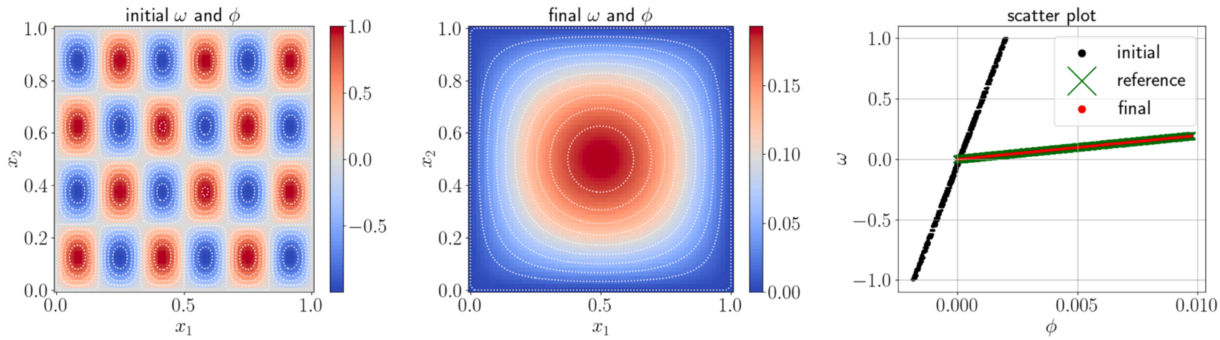

Fig. 12. The same as Fig. 9, but for the initial condition (90).

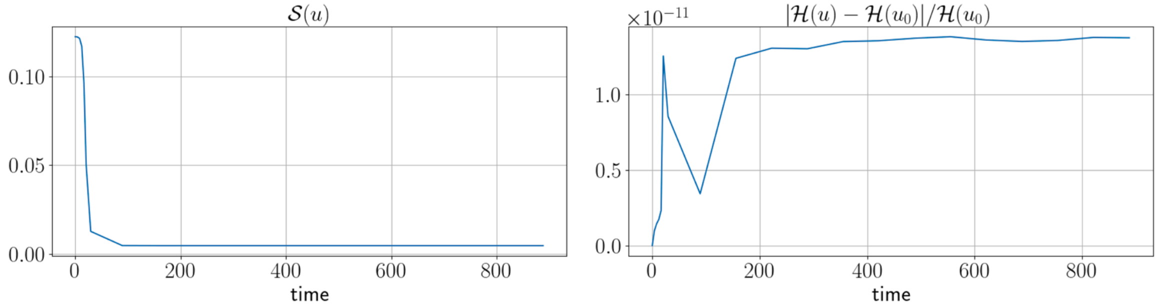

Fig. 13. Evolution of entropy (left-hand-side panel) and of the variation of the Hamiltonian relative to its initial value (right-hand-side panel), for the case in Fig. 12.

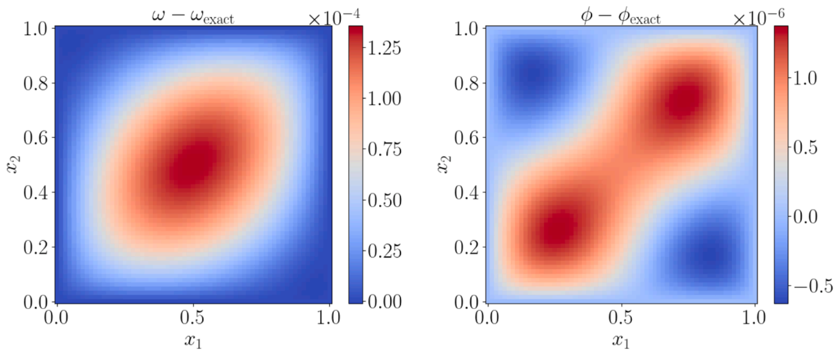

Fig. 14. The same as Fig. 11, but for the initial condition (90).

condition is close to an equilibrium. The relaxed state (Fig. 12, middle panel) is not exactly the same as in Fig. 9, since the initial value $\mathcal{H}_0$ of the Hamiltonian is different, but is it consistent with the exact solution (87) as shown in the scatter plot. Therefore, the initial condition has been completely relaxed to a solution of variational principle (13). In Fig. 13, one can see that the Hamiltonian function is preserved to machine accuracy, while entropy decays monotonically as expected. However, initially entropy appears to remain constant, due to the proximity of the initial condition to an equilibrium point. The difference between the relaxed state and the analytical solution of the variational principle (13) is shown in Fig. 14.

Gibbs entropy density. We consider now a more complicated entropy function, namely, the Gibbs entropy, which is given by the entropy density $s(\omega) = \omega \log \omega$. In this case the expected relaxed state is determined by, cf. Eq. (15),

$$1 + \log \omega = \lambda \phi \quad \Longleftrightarrow \quad \omega = e^{\lambda\phi - 1}, \tag{91}$$

where the eigenvalue λ could be numerically estimated by means of Eq. (88). The initial condition is the same as the one in Eq. (89) except for the amplitude, which here is increased to $1/N = 10$.

Fig. 15 shows the initial condition, the final state after the relaxation and the usual scatter plot. In this case the solution of the variational principle (green crosses) is computed using the numerically computed values of $\mathcal{M}(\omega)$ and $\mathcal{S}(\omega)$ in Eq. (88). Again we see

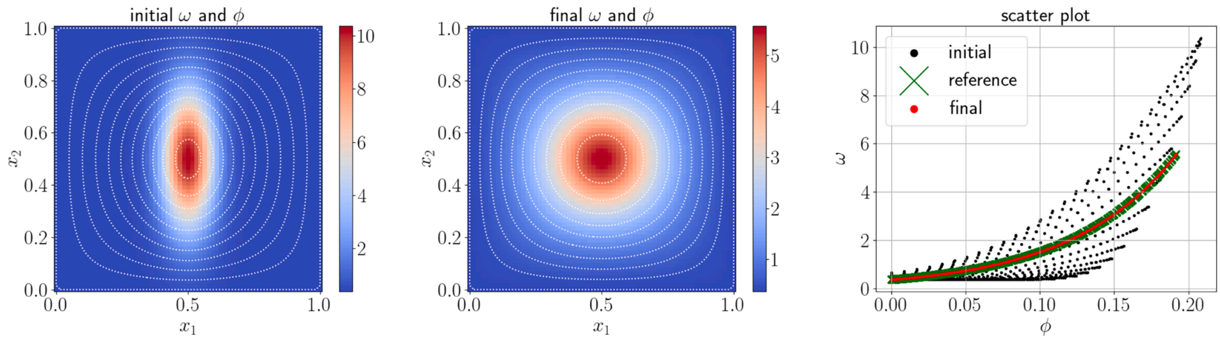

Fig. 15. The same as Fig. 9, but with the entropy density $s(\omega) = \omega \log \omega$. For the reference solution (green crosses) we used Eq. (91) with ϕ given by the numerical solution and λ computed from Eq. (88).

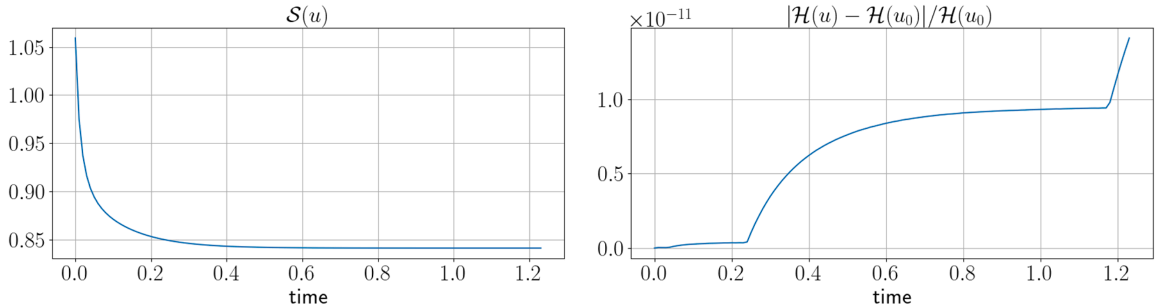

Fig. 16. Evolution of entropy (left-hand-side panel) and of the variation of the Hamiltonian relative to its initial value (right-hand-side panel), for the case in Fig. 15.

evidence of complete relaxation of the system toward the solution of the variational principle. Fig. 16 shows the expected behavior of entropy and Hamiltonian functions. It is worth noting that since the entropy is not quadratic in ω, the numerical scheme does not preserve the property of monotonic entropy dissipation, hence sufficiently small time steps must be used to ensure the evolution of the system is approximated with sufficient accuracy.

5.5. *Application to Grad-Shafranov equilibria*

As a second example, we construct a relaxation method for axisymmetric MHD equilibria, cf. Section 2.2.2. Essentially, this amounts to a different iterative method for the solution of the Grad-Shafranov equation. The metriplectic relaxation method ensures conservation of the poloidal magnetic energy and monotonic dissipation of an "ad hoc" entropy, but at a higher computational cost.

Specifically, we construct a relaxation method for the variational principle (20). On a bounded domain Ω, satisfying $\overline{\Omega} \subset \mathbb{R}_+ \times \mathbb{R}$ with coordinates $x = (x_1, x_2) = (r, z)$, cf. Section 2.2.2, the state variable is a scalar field $u(t)$ proportional to the toroidal component of the plasma current, $u(t, r, z) = (4\pi/c) r J_\varphi(t, r, z)$, and it is evolved in time according to Eq. (85) as in the case of the reduced Euler equations.

The entropy and Hamiltonian functions are chosen as in Eq. (21), with entropy density and measure μ on Ω given by

$$s(r, y) = \frac{1}{2} \frac{y^2}{Cr^2 + D}, \qquad d\mu(r, z) = \frac{1}{r} dr dz, \tag{92}$$

and the assumptions on the domain imply $r \geq r_0 > 0$ in $\overline{\Omega}$. Then, the first condition in Eq. (23) defines the toroidal current

$$\frac{4\pi}{c} J_\varphi = \lambda \Big(Cr + \frac{D}{r} \Big) \psi,$$

which is the current of the equilibrium found by Herrnegger [122] and Maschke [123], cf. also Mc Carthy [124, Section II.B]. Equivalently, since the state variable is $u = (4\pi/c) r J_\varphi$, the condition in Eq. (23) can be written as

$$\frac{u}{Cr^2 + D} = \lambda \psi. \tag{93}$$

In order to have a reference solution, we resort to the direct numerical solution of the Grad-Shafranov equation by using the classical iterative scheme [77, pp. 22–23, Eqs. (2.111) and (2.112)]. In the following experiments $C = 0.6$ and $D = 0.2$.

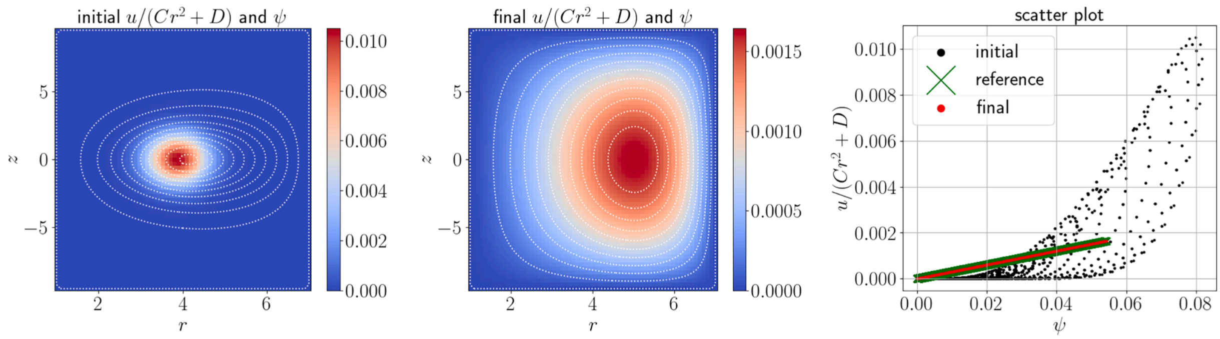

Fig. 17. Relaxation of an initial current $J_\varphi = (c/4\pi r)u_0$, with u_0 Gaussian, according to the evolution Eq. (85) applied to the Grad-Shafranov problem (20) with entropy (92) on a rectangular domain. The initial condition and the final state of the system are given in left-hand-side and middle panels, respectively, while the right-hand side panel shows the scatter plot, with the same color/symbol code as in Fig. 9. The color map represents the field $u/(Cr^2 + D)$, and the white contours are the flux function ψ, so that condition (93) is easily checked. Analogously the axes in the scatter plot refer to the values of the flux function ψ and the field $u/(Cr^2 + D)$. The reference solution (green crosses) is computed using relation (93), with ψ being given by the numerical solution and with the eigenvalue λ computed by a standard Grad-Shafranov solver. Here, $C = 0.6$ and $D = 0.2$.

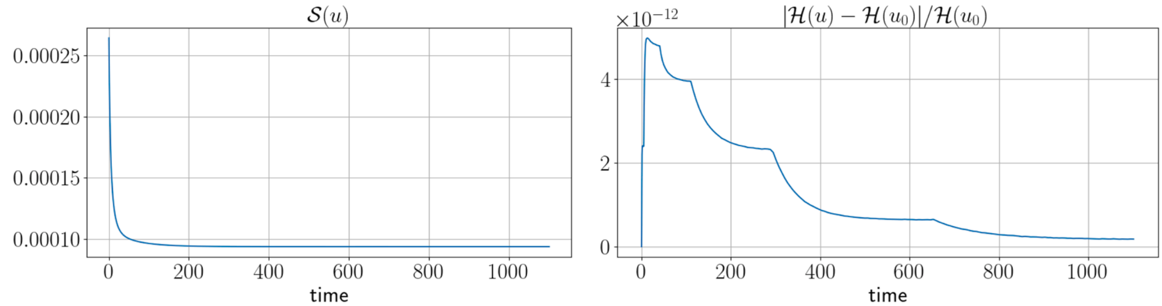

Fig. 18. Evolution of entropy (left-hand-side panel) and of the variation of the Hamiltonian relative to its initial value (right-hand-side panel), for the case in Fig. 17.

Rectangular domain. We start with a rectangular domain $\overline{\Omega} = [1, 7] \times [-9.5, +9.5]$, with coordinates $x = (x_1, x_2) = (r, z)$, discretized by a uniform grid of 64×64 nodes. The initial condition u_0 for the state variable is the same as in Eq. (89) with parameters $x_{0,1} = r_0 = 4$, $x_{0,2} = z_0 = 0$, $w_1^2 = 0.5$, $w_2^2 = 3.2$, and $N = 1$.

Fig. 17 shows the results of this numerical experiment. Instead of plotting the state variable u directly, the color plot represents the field $u/(Cr^2 + D)$, which should be proportional to ψ if the system reaches a state consistent with Eq. (93). Qualitatively the results are similar to those of Fig. 9 for the reduced Euler equations: the initial condition evolves toward an equilibrium state consistent with Eq. (93). The scatter plot in Fig. 17 shows that the functional relation between $u/(Cr^2 + D)$ and the potential ψ is linear. The reference solution (green cross) is obtained computing the field $u/(Cr^2 + D)$ from Eq. (93), with ψ being the numerical solution and with the eigenvalue $\lambda = 0.030302$ being obtained from the standard iterative solver of the Grad-Shafranov equation, which has also been implemented in FEniCS. Fig. 18 confirms the expected behavior of the entropy and Hamiltonian functions.

Mapped domain. At last, we give an example of the relaxation method for the Grad-Shafranov equation on a non-trivial mapped domain with a smooth boundary. The domain Ω is obtained by mapping the unit disk $\{z \in \mathbb{C} : |z| < 1\}$ with the map defined by

$$\begin{aligned} r &= a\left[b + \frac{1}{\varepsilon}\left(1 - \sqrt{1 + \varepsilon(\varepsilon + 2\, s\cos\theta)}\right)\right], \\ z &= c\,\frac{e\xi s\sin\theta}{2 - \sqrt{1 + \varepsilon(\varepsilon + 2\, s\cos\theta)}}, \end{aligned} \tag{94}$$

where $z = s\exp(i\theta)$ is a point in the unit disk, and the parameters are $e = 1.4$, $\varepsilon = 0.3$, $a = 4$, $b = 3$, and $c = 6.3$, with $\xi = 1/\sqrt{1 - \varepsilon^2/4}$. This map is a slightly modified version of the one used by Zoni and Güçlü [125, Eq. (3) and dummyTXdummy-[references therein]. The initial condition is the same as for the rectangular domain, except for $r_0 = 12$, $w_1^2 = 0.6$, and $w_2^2 = 6.0$.

The results for the mapped domain are shown in Fig. 19. The color map represents the field $u/(Cr^2 + D)$, as in the previous case. The relaxed state is again consistent with Eq. (93), with the eigenvalue $\lambda = 0.002599$ computed by a standard Grad-Shafranov solver. Fig. 20 shows the expected monotonic entropy dissipation and the conservation of the Hamiltonian to machine precision.

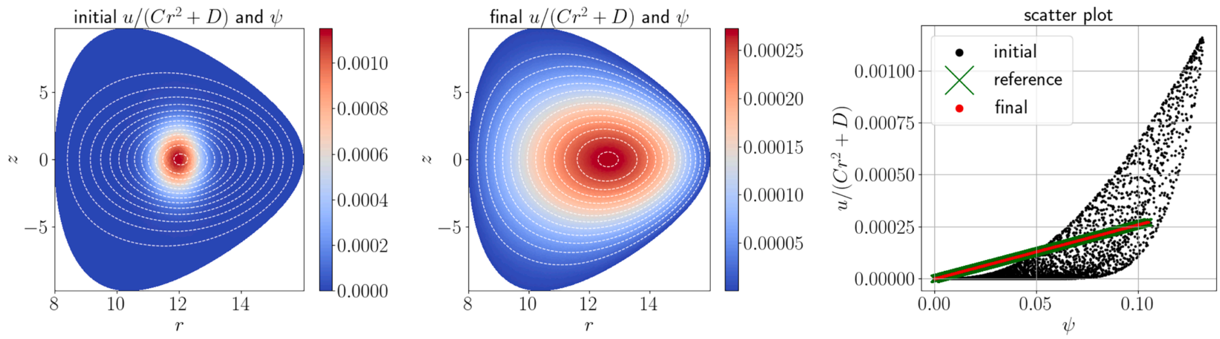

Fig. 19. The same as in Fig. 17, but on the domain obtained mapping the unit disk with Eq. (94).

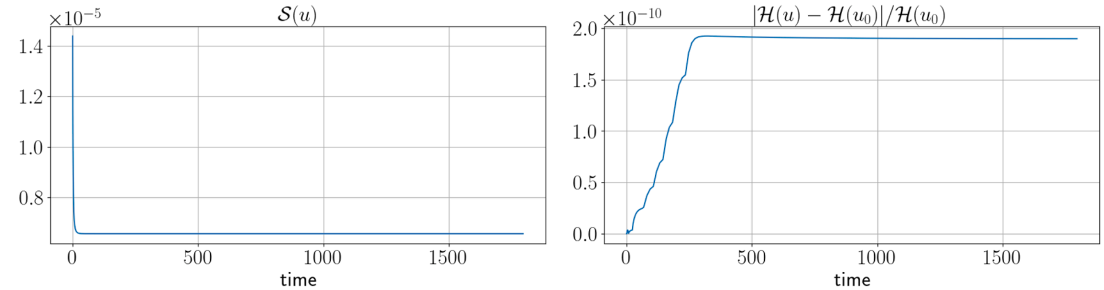

Fig. 20. Evolution of entropy (left-hand-side panel) and of the variation of the Hamiltonian relative to its initial value (right-hand-side panel), for the case in Fig. 19.

6. Diffusion-like metric brackets

A feature of the general collision-like bracket is that the generalized diffusion tensor and friction flux (81) are *nonlocal* functions of the unknown u. Therefore, their evaluation requires an integration over the whole domain Ω. This poses an issue of computational complexity even harder than that of the standard Landau collision operator, which is local in half of the variables (cf. Example 5). In order to mitigate the computational cost of a relaxation method based on these brackets, we have studied the special case of brackets (70) that corresponds to choosing $\mathcal{O} = \Omega$, i.e., we do not increase the size of the domain ($n = d$). We will refer to this case as a diffusion-like bracket, and reserve the name "collision-like" for the case $n > d$. As we shall see, in general, one cannot expect complete relaxation (in the sense defined in Section 3) from the diffusion-like brackets. The metric double bracket (57) is a special case of a diffusion-like bracket.

6.1. General construction of the brackets

With the same setup of Section 5.1, which is summarized in Fig. 8, let us consider the case $\mathcal{O} = \Omega$, $\nu = \mu$, but allow $\tilde{N} \neq N$, so that in general $\tilde{W} \neq W$.

Then bracket (70) reduces to

$$(\mathcal{F}, \mathcal{G}) := \int_\Omega P\frac{\delta \mathcal{F}(u)}{\delta u} \cdot \mathsf{T}(u) P\frac{\delta \mathcal{G}(u)}{\delta u} d\mu, \tag{95}$$

where $P : W \to \tilde{W}$ is a linear (possibly unbounded) operator, and $\mathsf{T}(u) \in \mathcal{B}(\tilde{W})$ is symmetric, positive semidefinite, bounded linear operator that satisfies (69). The evolution equation generated by the bracket (95) is formally the same as Eq. (72), but we shall see in the examples that the dual operator P' does not involve any integral operator.

Conditions (73) are of course still sufficient conditions for brackets of the form (95) to be minimally degenerate, but we shall see that condition (73b) is usually not satisfied in this case.

Example 8.

Brackets of the form (57), i.e., metriplectic double brackets acting on scalar fields, are special cases of diffusion-like brackets (95). In order to see this, let $N = 1$, thus $W = L^2(\Omega, \mu)$, $\Phi = C^\infty(\overline{\Omega})$, and

$$Pw = \nabla w.$$

Hence, $\tilde{N} = d$ and $\tilde{W} = L^2(\Omega, \mu; \mathbb{R}^d)$ is the space of L^2 vector fields over Ω. Given a function $J(x)$ of class $C^\infty(\overline{\Omega})$ taking values in the space of antisymmetric $d \times d$ matrices, we can define the antisymmetric bilinear operation

$$[w_1, w_2]_J = \nabla w_1(x) \cdot J(x) \nabla w_2(x).$$

As in Section 4.1, it is not necessary that $[\cdot,\cdot]_J$ satisfies the Jacobi identity. In addition let us consider a symmetric, positive-definite bi-linear form $\gamma : W \times W \to \mathbb{R}$, together with the associated linear bounded, symmetric positive definite operator $\Gamma : W \to W$, that is, cf. Appendix B,

$$\gamma(w_1, w_2) = \int_\Omega w_1(x) \cdot \Gamma w_2(x) d\mu(x),$$

for all $w_1, w_2 \in W$. In terms of J and γ, we define the kernel

$$\mathsf{T}(u) = X_h(u) \circ \Gamma \circ {}^t X_h(u),$$

where $X_h(u)$ and ${}^t X_h(u)$ are the operators of multiplication by the vector fields

$$X_h(u; x) = J(x)\nabla h(x), \quad {}^t X_h(u; x) = -\nabla h(x) \cdot J(x),$$

respectively, and $h = \delta\mathcal{H}(u)/\delta u \in \Phi = C^\infty(\overline{\Omega})$. If J is a Poisson tensor, i.e., $[\cdot,\cdot]_J$ satisfies the Jacobi identity, $X_h(u;\cdot)$ is the d-dimensional Hamiltonian vector field generated by the Hamiltonian function $h = \delta\mathcal{H}(u)/\delta u$. Since $X_h(u;\cdot) \in C^\infty(\overline{\Omega}; \mathbb{R}^d)$, $\mathsf{T}(u)$ maps $\tilde{W}$ into itself. The operator Γ is symmetric and positive definite, hence $\mathsf{T}(u)$ is symmetric and positive semidefinite, and we have $\nabla h \in \ker \mathsf{T}(u)$ as required by condition (69).

With the foregoing choices of P and $\mathsf{T}(u)$, Eq. (95) becomes

$$\begin{aligned}(\mathcal{F}, \mathcal{G}) &= \int_\Omega \Big(\nabla \frac{\delta\mathcal{F}(u)}{\delta u} \cdot J \nabla \frac{\delta\mathcal{H}(u)}{\delta u}\Big) \Gamma \Big(\nabla \frac{\delta\mathcal{G}(u)}{\delta u} \cdot J \nabla \frac{\delta\mathcal{H}(u)}{\delta u}\Big) d\mu \\ &= \gamma\Big(\Big[\frac{\delta\mathcal{F}(u)}{\delta u}, \frac{\delta\mathcal{H}(u)}{\delta u}\Big]_J, \Big[\frac{\delta\mathcal{G}(u)}{\delta u}, \frac{\delta\mathcal{H}(u)}{\delta u}\Big]_J\Big),\end{aligned}$$

which is Eq. (57). As discussed in Section 4.1, this bracket is in general not minimally degenerate. In fact, while condition (73a) amounts to the Poincaré inequality and holds true on a bounded domain Ω, condition (73b) fails since for any sufficiently regular function $f : \mathbb{R} \to \mathbb{R}$, any function of the form $\tilde{w} = \nabla f(h)$, with $h = \delta\mathcal{H}(u)\delta u$, belongs to $\ker \mathsf{T}(u) \cap \operatorname{ran}(P)$.

6.2. Diffusion-like brackets based on div–grad operators

We address the diffusion-like version of the bracket introduced in Section 5.2. For scalar fields ($N = 1$) on a bounded domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$, let $P = \nabla$ with domain $\Phi = H^1_0(\Omega)$. The kernel of the bracket is constructed from the matrix Q_d, cf. Eq. (76),

$$\mathsf{T}(u; x) = \kappa(u; x) Q_d\Big(\nabla \frac{\delta\mathcal{H}(u)}{\delta u}(x)\Big),$$

where $\kappa(u; x)$ is a positive scalar function, and $\mathsf{T}(u)$ is defined as the operator of multiplication by $\mathsf{T}(u;\cdot)$. Then, bracket (95) becomes [110]

$$(\mathcal{F}, \mathcal{G}) = \int_\Omega \kappa(u) \nabla \frac{\delta\mathcal{F}(u)}{\delta u} \cdot Q_d\Big(\nabla \frac{\delta\mathcal{H}(u)}{\delta u}\Big) \nabla \frac{\delta\mathcal{G}(u)}{\delta u} d\mu, \tag{96}$$

and the corresponding evolution equation is

$$\partial_t u = \widetilde{\operatorname{div}}_\mu \Big[\kappa(u) Q_d\Big(\nabla \frac{\delta\mathcal{H}(u)}{\delta u}\Big) \nabla \frac{\delta\mathcal{S}(u)}{\delta u}\Big] \quad \text{in } \Phi' = H^{-1}(\Omega).$$

This is a "local version" of Eq. (85) which justifies the name "diffusion-like" for this bracket. Condition (73b) fails in the same way as in Example 8.

It is worth noting that in two spatial dimensions, $d = 2$, one has

$$Q_2(\nabla h) = X_h \otimes X_h = X_h {}^t X_h, \qquad h = \delta\mathcal{H}(u)/\delta u,$$

where $X_h = J_c \nabla h = (-\partial_2 h, \partial_1 h)$ is the canonical Hamiltonian vector field generated by $h(x)$ and ${}^t X_h$ denotes its transpose. This means that the diffusion-like bracket (96) in a two dimensional domain amounts to the metric double bracket addressed in Example 8, with $\Gamma = I$ being the identity operator. For $d > 2$ the bracket (96) is however different from the metric double brackets in Example 8. In fact, if $X_h \neq 0$, the null space of the matrix Q_d is always one dimensional for any dimension d, while the null space of $X_h \otimes X_h$ is $d-1$ dimensional, hence the two matrices have the same null space only if $d = 2$.

Nonetheless, for $d \geq 3$, one can write the matrix $Q_d(\nabla h)$ in terms of suitable pairing of two antisymmetric operations by using the identity

$$\frac{1}{(d-2)!} \sum_{i_1,\dots,i_{d-2}} \epsilon_{i_1,\dots,i_{d-2},i,k} \epsilon_{i_1,\dots,i_{d-2},j,l} = \delta_{ij}\delta_{kl} - \delta_{il}\delta_{jk},$$

with $\epsilon_{i_1,\dots,i_d}$ being the completely antisymmetric symbol. We obtain

$$\begin{aligned}[Q_d(\nabla h)]_{ij} &= |\nabla h|^2 \delta_{ij} - \partial_i h \partial_j h = \sum_{kl} \big[\delta_{ij}\delta_{kl} - \delta_{il}\delta_{jk}\big] \partial_k h \partial_l h \\ &= \frac{1}{(d-2)!} \sum_{i_1,\dots,i_{d-2}} \sum_{k,l} \epsilon^{i_1,\dots,i_{d-2},i,k} \epsilon^{i_1,\dots,i_{d-2},j,l} \partial_k h \partial_l h,\end{aligned} \tag{97}$$

where $\partial_j h = \partial h/\partial x_j$, and thus, Eq. (96) can be written equivalently as

$$(\mathcal{F},\mathcal{G}) = \frac{1}{(d-2)!}\sum_{\alpha}\int_{\Omega}\kappa(u)\,\mathcal{E}_d^{\alpha}(\nabla f,\nabla h)\,\mathcal{E}_d^{\alpha}(\nabla g,\nabla h)\,d\mu,$$

where $\alpha = (i_1,\dots,i_{d-2})$ is a multi-index, $f = \delta\mathcal{F}(u)/\delta u$, $g = \delta\mathcal{G}(u)/\delta u$, and

$$\mathcal{E}_d^{\alpha}(\nabla\varphi,\nabla\psi) = \sum_{i,k}\epsilon^{i_1,\dots,i_{d-2},i,k}\partial_i\varphi\partial_k\psi, \qquad \alpha = (i_1,\dots,i_{d-2}).$$

In dimension $d = 3$, $\mathcal{E}_3$ defines a Lie bracket in $\mathbb{R}^3$. This is the standard Lie algebra structure on $\mathbb{R}^3$ given by the cross product arising in the case of rigid body rotation [3].

Yet another form of this bracket makes use of the Kulkarni-Nomizu (K-N) product and the metriplectic 4-bracket structure [35]. Using the first identity in (97), we can write

$$(\mathcal{F},\mathcal{G}) = \frac{1}{2}\sum_{i,j,k,l}\int_{\Omega}\kappa(u)(\delta \owedge \delta)_{ijkl}\Big[\partial_i\frac{\delta\mathcal{F}(u)}{\delta u}\Big]\Big[\partial_j\frac{\delta\mathcal{H}(u)}{\delta u}\Big]\Big[\partial_k\frac{\delta\mathcal{G}(u)}{\delta u}\Big]\Big[\partial_l\frac{\delta\mathcal{H}(u)}{\delta u}\Big]d\mu,$$

where $(\delta \owedge \delta)_{ijkl} = 2(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})$ is the K-N product of two identity tensors.

6.3. Diffusion-like brackets based on curl–curl operators

As a last example, we address the diffusion-like version of the $\mathrm{curl}-\mathrm{curl}$ brackets of Section 5.3. We consider a vector field over a bounded domain $\Omega \subset \mathbb{R}^3$ with Lebesgue measure $d\mu(x) = dx$, hence $d = N = 3$. We choose

$$Pw = \operatorname{curl} w,$$

so that $\tilde{N} = N = 3$ and $\tilde{W} = W$, with $\mathrm{dom}(P)$ given by

$$\Phi = \{w \in H(\operatorname{curl},\Omega)\cap H(\operatorname{div},\Omega):\ \operatorname{div} w = 0 \text{ in } \Omega,\ n\cdot w = 0 \text{ on } \partial\Omega\}.$$

This differs from the space Φ considered in Section 5.3 by the "opposite" choice of boundary conditions: the normal component is set to zero instead of the tangential component. The Poincaré inequality for the operator curl holds for this space as well [116, Corollary 3.51], so that condition (73a) holds true. (Here, we have the choice of the boundary condition since we do not need to satisfy the second identity in Eq. (82).)

As an example, let the kernel be once again constructed from Q_3, defined in Eq. (76),

$$\mathsf{T}(u;x) = \kappa(u;x)Q_3\Big(\operatorname{curl}\frac{\delta\mathcal{H}(u)}{\delta u}\Big),$$

where $\kappa(u;x)$ is a positive function. Even though the operator $P = \mathrm{curl}$ with domain $\mathrm{dom}(P) = \Phi$ satisfies a Poincaré inequality, in general, the kernel fails to satisfy condition (73b): a function $\tilde{w} \in \ker\mathsf{T}(u)\cap\mathrm{ran}(P) \subset \tilde{W} = L^2(\Omega;\mathbb{R}^3)$ must satisfy $\tilde{w} = \operatorname{curl} w$, with $w \in \Phi$ and $\tilde{w}(x) = \Lambda(x)b(x)$, $b = \operatorname{curl} h$, $h = \delta\mathcal{H}/\delta u$, and thus the pair (Λ, w) must solve

$$\begin{cases} \operatorname{curl} w = \Lambda b, & \text{in } \Omega,\\ \operatorname{div} w = 0, & \text{in } \Omega,\\ b\cdot\nabla\Lambda = 0, & \text{in } \Omega,\\ n\cdot w = 0, & \text{on } \partial\Omega. \end{cases}$$

As a special case let b be a nonlinear Beltrami field, i.e., a solution of (24) such that $\operatorname{curl} b = \mu b$ with $\mu(x)$ not a constant, then $\Lambda = \mu$ and $w = b$ is a solution that violates condition (73b).

With the foregoing choices, Eq. (95) amounts to

$$(\mathcal{F},\mathcal{G}) = \int_{\Omega}\kappa(u)\operatorname{curl}\frac{\delta\mathcal{F}(u)}{\delta u}\cdot Q_3\Big(\operatorname{curl}\frac{\delta\mathcal{H}(u)}{\delta u}\Big)\operatorname{curl}\frac{\delta\mathcal{G}(u)}{\delta u}dx, \tag{98}$$

and the corresponding evolution equation becomes

$$\partial_t u = -\widetilde{\operatorname{curl}}\Big[\kappa(u)Q_3\Big(\operatorname{curl}\frac{\delta\mathcal{H}(u)}{\delta u}\Big)\operatorname{curl}\frac{\delta\mathcal{S}(u)}{\delta u}\Big] \quad \text{in } \Phi' = H'(\operatorname{curl},\Omega).$$

This bracket can be written in terms of the antisymmetric bilinear operator

$$\mathcal{E}_3(X,Y) = [X,Y]_{\mathbb{R}^3} := X\times Y, \qquad X,Y\in\mathbb{R}^3,$$

which is the standard Lie bracket in $\mathbb{R}^3$. In fact Eq. (98) can be shown to be a special case of the following (cf. [3,64]):

$$(\mathcal{F},\mathcal{G}) = \int_{\Omega}\Big[\operatorname{curl}\frac{\delta\mathcal{F}(u)}{\delta u},\operatorname{curl}\frac{\delta\mathcal{H}(u)}{\delta u}\Big]_{\mathbb{R}^3}\Gamma\Big[\operatorname{curl}\frac{\delta\mathcal{G}(u)}{\delta u},\operatorname{curl}\frac{\delta\mathcal{H}(u)}{\delta u}\Big]_{\mathbb{R}^3}dx, \tag{99}$$

where $\Gamma(u) \in \mathcal{B}\big(L^2(\Omega;\mathbb{R}^3)\big)$ is a symmetric, positive definite operator; Eq. (98) is obtained for $\Gamma(u) = \kappa(u)$, the multiplication operator by the function $\kappa(u;\cdot)$. Eq. (99) is a metric double bracket of the form (57). Applied to magnetic fields this gives a generalization of the relaxation method of Chodura and Schlüter [30] with constant pressure, cf. also Moffatt [23]. An explicit example will be briefly reported in Section 6.4 below.

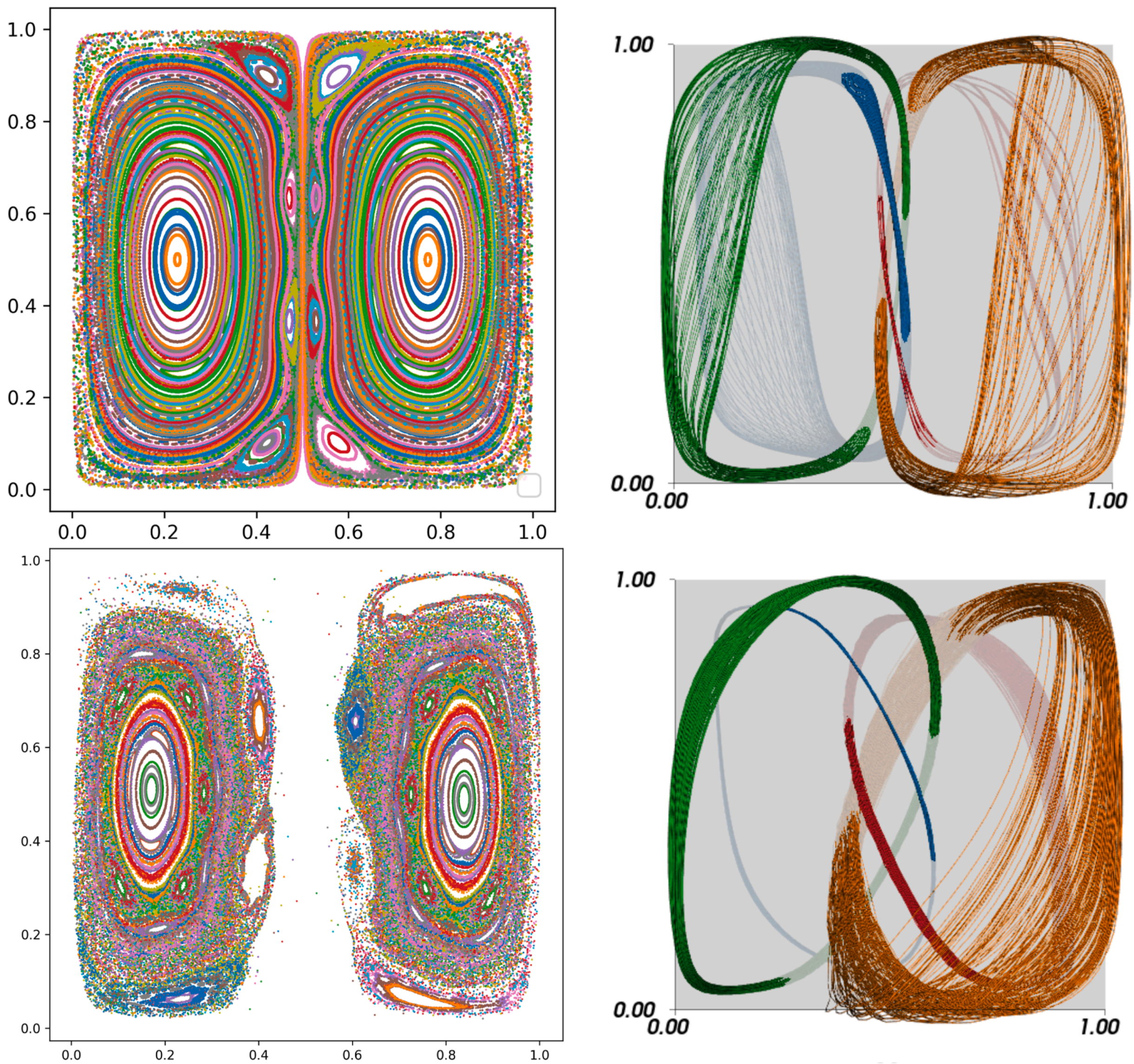

Fig. 21. Poincaré plot in the plane $x_1 - x_2$ and selected field lines of the magnetic field B, for the initial condition (top row) and the final state (bottom row), after the relaxation process. The initial condition is given in Eq. (104) with $m = n = 1$ and $a = 1$, while the evolution equation is the magneto-frictional method (101). The selected field lines correspond to the four large islands visible in the Poincaré plot around $x_1 = 1/2$. The Poincaré section is defined by $x_2 = 1/2$ and it is shown in light gray in the panels on the right-hand side. The initial points of the field lines are sampled differently for the initial and finals state of the field, hence they are not exactly the evolution of one another.

6.4. Application to nonlinear Beltrami fields

So far we have focused on examples of equilibrium problems for which complete relaxation of the solution is essential. We have shown that a metriplectic relaxation method for such problems should be based on metric brackets that are minimally degenerate (or specifically degenerate if more than one constraint is considered), cf. Section 3. Diffusion-like brackets do not appear to be appropriate for those problems.

For sake of completeness, we address an example of equilibrium problems that are characterized as minima of a function subject to topological constraints. This is the case of nonlinear Beltrami fields, for which the variational principle is given in Lagrangian representation, cf. Section 2.2.3 and Appendix C. Full three-dimensional MHD equilibria satisfy the same type of Lagrangian variational principle.

Because of their larger null space, metric double brackets of the form (99) allow us to obtain an evolution equation that preserves the necessary constraints. To this end, we identify the state variable $u(t)$ with a magnetic field $u(t, x) = B(t, x)$ on a simply connected,

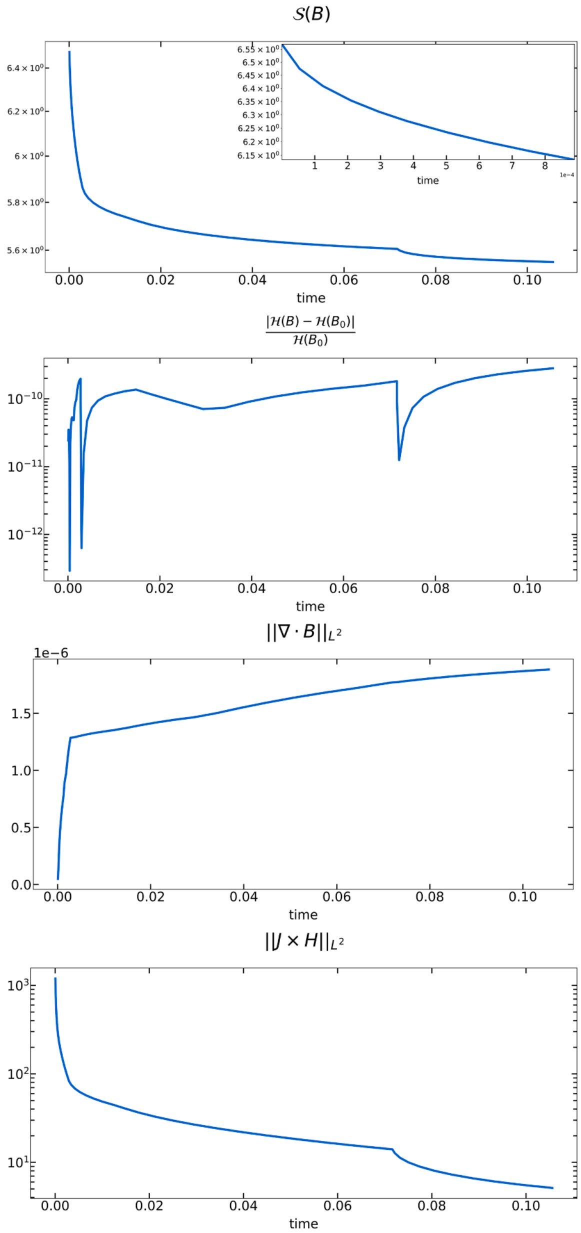

Fig. 22. From top to bottom, time evolution of the entropy (magnetic energy), the relative variation of the Hamiltonian (magnetic helicity), the L^2 norm of the divergence, and of the vector $j \times H$, where $j = \widetilde{\mathrm{curl}} B$ is the current density (computed with the weak curl operator), and H is the L^2-orthogonal projection of B onto $H_0(\mathrm{curl}, \Omega)$. The equilibrium condition for the considered numerical scheme reads $j \times H = 0$. Corners and jumps in the time traces corresponds to restarts with larger time steps.

bounded domain $\Omega \subset \mathbb{R}^3$. More specifically, we assume that $B(t) \in V \subset \Phi$, where Φ is the same space defined in Section 5.3. The evolution equation for $B(t)$ is given by Eq. (7a) and bracket (99), with $\Gamma = I$, the identity operator, for simplicity, and with entropy and Hamiltonian given in Eq. (27). Therefore, if an orbit of this metriplectic system completely relaxes, it would converge in time to a *linear* Beltrami field, cf. Section 2.2.3. In fact, this bracket has a much larger null space. The equilibrium points, given by $B \in \Phi$ such that $(\mathcal{S}, \mathcal{S})(B) = 0$, satisfy the Beltrami condition $(\mathrm{curl}\, B) \times B = 0$, in the weak formulation discussed in Section 2.2.3.

The resulting evolution equation amounts to

$$\partial_t B = \widetilde{\mathrm{curl}}\big[B \times \big(B \times \mathrm{curl}\, B\big)\big], \tag{100}$$

where we have accounted for the identity $\mathrm{curl}[\delta\mathcal{H}(B)/\delta B] = \mathrm{curl}\, A = B$. If B is sufficiently regular, we can replace $\widetilde{\mathrm{curl}}$ by curl and write

$$\begin{cases} \partial_t B - \mathrm{curl}\left[V \times B\right] = 0, & \text{in } \Omega, \\ V - (\mathrm{curl}\, B) \times B = 0, & \text{in } \Omega, \\ n \cdot B = 0, & \text{on } \partial\Omega\,, \end{cases} \tag{101}$$

which shows that the magnetic field B is advected by the flow of the effective "velocity" field V. Hence, so long as the solution remains smooth, the field lines of B are frozen into the flow (actually flux), i.e., they cannot change their topological properties. This is a much stronger constraint than just preservation of magnetic helicity $2\mathcal{H}(B)$. As anticipated, Eq. (101) is exactly the relaxation method of Chodura and Schlüter [30] with constant pressure. The method itself is therefore not new. In solar physics this relaxation method is known as the magneto-frictional method [126–129], and it has been applied to the computation of force-free magnetic fields in coronal active regions [31]. The bracket formalism, however, opens the way to possible generalizations by means of different choices of the kernel. This possibility will be explored in future work. Since this relaxation method is based on the MHD induction equation, smoothness of the solution may be lost in a finite time due to the formation of current sheets, as conjectured by Parker and discussed in Section 1.1. In this work we allow for weak solutions. In fact, Eq. (100) is reformulated with $B(t) \in H_0(\mathrm{div}, \Omega)$ only. More precisely, we search for $B \in C^1\big(([0,T]; H_0(\mathrm{div}, \Omega)\big)$ and auxiliary variables $E, j, H \in C\big([0,T]; H_0(\mathrm{curl}, \Omega)\big)$ satisfying

$$\begin{cases} \partial_t B + \mathrm{curl}\, E = 0, & \text{in } H_0(\mathrm{div}, \Omega), \\ (H, G)_{L^2} - (B, G)_{L^2} = 0, & \forall G \in H_0(\mathrm{curl}, \Omega), \\ (j, k)_{L^2} - (B, \mathrm{curl}\, k)_{L^2} = 0, & \forall k \in H_0(\mathrm{curl}, \Omega), \\ (E, F)_{L^2} - (H \times j, H \times F)_{L^2} = 0, & \forall F \in H_0(\mathrm{curl}, \Omega), \end{cases} \tag{102}$$

pointwise in time, with $F, G, k \in H_0(\mathrm{curl}, \Omega)$ being test functions. Faraday's equation is posed strongly as an identity in $H_0(\mathrm{div}, \Omega)$. As a result the condition $\mathrm{div}\, B = 0$ is preserved. One can also show directly that a solution of this system preserves magnetic helicity and dissipate magnetic energy, that is, the properties of the bracket hold for this reformulation. In particular, we observe that

$$\frac{1}{2}\frac{d}{dt}\int_\Omega |B|^2 dx = -\|j \times H\|_{L^2}^2,$$

and the equilibrium condition is $j \times H = 0$, which is the weak formulation of the Beltrami condition anticipated in Section 2.2.3. Here we present a single numerical experiment obtained by a structure-preserving numerical scheme [110], which we derived by adapting the finite-element exterior calculus (FEEC) scheme of Hu et al. [86] for incompressible MHD. The scheme provably preserves the Hamiltonian (magnetic helicity), the constraint $\mathrm{div}\, B = 0$, and the monotonic behavior of entropy (magnetic energy), but it does *not* preserve the topology of the field lines exactly. Similar work has been recently published by He et al. [130]. Previously, the magnetic relaxation problem has been dealt with by means of Lagrangian [131] and finite difference [132] methods. More recently, a different kind of Lagrangian numerical scheme has been proposed [133,134], which is based on the discretization of the domain in narrow flux tubes, each one being relaxed by a curve-shortening flow in a modified metric. This interesting scheme therefore preserves the topological properties of the field lines. Yet with the domain discretized by a finite set of lines, the reconstruction of the magnetic field at arbitrary points of the domain, needs to be addressed.

Before describing the considered test case, let us address the role of magnetic-helicity conservation. In a domain Ω where the Poincaré inequality for the curl operator holds true, magnetic helicity $H_m(B) = 2\mathcal{H}(B)$ provides a lower bound for magnetic energy. In fact, one has [22]

$$|H_m(B)| \le \|A\|_{L^2(\Omega)}\|B\|_{L^2(\Omega)} \le C_P\|B\|^2_{L^2(\Omega)}. \tag{103}$$

For an initial condition B_0 with $H_m(B_0) = 0$, it is possible that the solution of (101) with the chosen boundary conditions ($B \cdot n = 0$ on $\partial\Omega$) relaxes to a trivial field, i.e., $|B(t)| \to 0$ for $t \to +\infty$, even if the topology of the field lines is preserved. This is the case for the class of one-dimensional solutions of (101), which are obtained, for instance, by assuming

$$B(t,x) = \begin{pmatrix} 0 \\ 0 \\ b(x_1) \end{pmatrix} = \mathrm{curl}\begin{pmatrix} 0 \\ a(x_1) \\ 0 \end{pmatrix},$$

where $a'(x_1) = b(x_1)$, $x = (x_1, x_2, x_3)$ and the field is constant in (x_2, x_3). We have $A \cdot B = 0$ and thus $H_m(B) = 0$. Correspondingly, Eq. (101) reduce to

$$\partial_t b - (b^2 b')' = 0,$$

where a prime denotes spatial differentiation. This is a standard heat equation. The steady states are solution to $b^2 b' =$ constant, which gives $b(x_1) = (c_1 + c_2 x_1)^{1/3}$, with c_1, c_2 being integration constants. For instance homogeneous boundary conditions for b on an interval yields the unique solution $b(x_1) = 0$. Magnetic relaxation in one dimension has been recently considered by Yeates [135] and compared to the corresponding full MHD relaxation, thus exposing the limitations of the magneto-frictional method.

It is therefore meaningful to consider initial conditions with non-trivial magnetic helicity $\mathcal{H}(B) = \frac{1}{2}H_m(B) \neq 0$. We construct such an initial condition from the vector potential

$$\tilde{A}(x) := a\begin{pmatrix} (n/\sqrt{m^2+n^2})\sin(\pi m x_1)\cos(\pi n x_2) \\ -(m/\sqrt{m^2+n^2})\cos(\pi m x_1)\sin(\pi n x_2) \\ \sin(\pi m x_1)\sin(\pi n x_2) \end{pmatrix},$$

with $a \in \mathbb{R}$ and $m, n \in \mathbb{N}$ being parameters (we shall choose $m = n = 1$). For any choice of the parameters, $\tilde{A}$ is divergence-free and a linear Beltrami field, periodic is all variables; it is an eigenvalue of curl corresponding to the eigenvalue $\lambda_{m,n} = \pi(m^2+n^2)^{1/2}$. We localize this field in the unit cube $\overline{\Omega} = [0,1]^3$ by means of the cut-off function

$$\eta(x) := \chi(x_1)\chi(x_2)\chi(x_3), \qquad \chi(y) := y^2(1-y)^2, \quad y \in [0,1].$$

We have $\eta(x) = 0$ and $\nabla\eta(x) = 0$ for $x \in \partial\Omega$ since both χ and its derivative χ' vanish for $y = 0$ and $y = 1$. We construct the initial condition on the domain $\overline{\Omega} = [0,1]^3$ as

$$\begin{aligned} A_0 &:= \eta\tilde{A}, \\ B_0 &:= \operatorname{curl} A_0 = \nabla\eta \times \tilde{A} + \eta \operatorname{curl}\tilde{A}, \end{aligned} \tag{104}$$

and $A_0 \in H_0(\mathrm{curl}, \Omega)$, $B_0 \in H_0(\mathrm{div}, \Omega)$ with $\operatorname{div} B_0 = 0$. As for magnetic helicity,

$$H_m(B_0) = 2\mathcal{H}(B_0) = \int_\Omega A_0 \cdot B_0 dx = \lambda_{m,n}\|A_0\|^2_{L^2(\Omega)} > 0.$$

We can control the initial helicity by means of the parameters $a \in \mathbb{R}$ and $m, n \in \mathbb{N}$. The magnetic field (104) is represented in Fig. 21 (top row), by means of a Poincaré plot using the plane $x_2 = 1/2$ as a Poincaré section. From the plot (Fig. 21, top-left panel), one can identify a rather complex topology of the field lines, with, in particular, four large islands of period two that are rendered in three dimensions in Fig. 21, top-right panel, by tracing a few selected field lines for each island.

The time evolution of the initial condition (104) is obtained numerically by means of the FEEC scheme, which has been implemented in FEniCS as in the case of the tests reported in Section 5. Here, we use a relatively course resolution: the domain $\overline{\Omega} = [0,1]^3$ has been discretized by a uniform grid of 32^3 nodes. The time step is adapted, but limited to a maximum of 10^{-3}. The obtained final state is represented in Fig. 21 (bottom row), again by means of a Poincaré plot, Fig. 21, bottom-left panel. The four large islands appear to have been preserved by the relaxation process and are rendered in Fig. 21, bottom-right panel. Other large islands appear to have been preserved, but a closer analysis shows that the field line topology is not exactly preserved [110]. Indeed the numerical scheme preserves magnetic helicity only, and this alone does not completely guarantee the exact preservation of the field line topology.

When using Poincaré plots for the visualization of the field line topology, one should address the effect of the error of the projection onto the finite-element space used for the representation of the magnetic field. In this case, B is approximated in the space of linear Raviart-Thomas elements for computations, and the discrete approximation is further projected onto the space of linear Lagrange elements, for visualization purposes (precise definitions of these finite-element spaces can be found, e.g., in the FEniCS book [118]). Lagrange elements are nodal so that the degrees of freedom coincide with the value of the field at the grid nodes and can be directly interpolated. We have qualitatively checked the effect of all those operations on the results by comparing the Poincaré sections of the analytical field (104) with that of its projection onto the finite element space on the considered grid.

Fig. 22 shows the evolution in time of the main quantities of interest, namely, the entropy, the relative variation of the Hamiltonian with respect to its initial value, and the L^2-norms of $\operatorname{div} B$ and of the vector $j \times H$, where $j = \widetilde{\operatorname{curl}} B$ is (proportional to) the current density (computed weakly), while $H(t)$ is the L^2-orthogonal projection of $B(t) \in H_0(\mathrm{div}, \Omega)$ onto the space $H_0(\mathrm{curl}, \Omega)$. For sufficiently regular fields, we have $j \times H = (\operatorname{curl} B) \times B$, but in general $B(t) \in H_0(\mathrm{div}, \Omega)$ and $H(t) \in H_0(\mathrm{curl}, \Omega)$ are different and the numerical equilibrium condition is $j \times H = 0$. We recall that in this application the entropy and the Hamiltonian coincide with the magnetic energy and the magnetic helicity, respectively. From Fig. 22, we see that the qualitative properties of the relaxation method are preserved: the entropy decreases monotonically, the Hamiltonian is constant within a relative error of 10^{-10}, and $\operatorname{div} B = 0$ within an absolute error of 2×10^{-6} measured by the norm in $L^2(\Omega)$. We also see that $j \times H$ decreases, which indicates that the solution is approaching a configuration that satisfies the Beltrami condition $j \times H = 0$.

7. Summary and conclusions

We have considered the question of whether, given a metriplectic system with Hamiltonian function $\mathcal{H}$ and (dissipated) entropy $\mathcal{S}$, the orbit with initial condition u_0, in the long-time limit, converges to a minimum of $\mathcal{S}$ on the surface of constant Hamiltonian $\{u : \mathcal{H}(u) = \mathcal{H}(u_0)\}$. This question is interesting in itself, since many physical systems are metriplectic, but our work is mainly motivated by the idea of utilizing artificial metriplectic systems as relaxation methods for the computation of equilibria of fluids and plasmas.

We have shown that, in general, the answer is negative. For finite-dimensional metriplectic systems, we have given a sufficient condition, Proposition 3, under which an orbit relaxes to a constrained entropy minimum. These results are proven by means of a natural extension of the Lyapunov stability theorem for systems with constants of motion. One key assumption in Proposition 3 is that the metric bracket should be specifically degenerate with respect to a given finite set of constants of motion, or minimally degenerate if the Hamiltonian is the only constant of motion. Recall, a metric bracket is specifically degenerate if its null space is spanned by the gradients of a finite number of invariant functions $\mathcal{I}^{\alpha}$, the constants of motion, and minimally degenerate if its null space is spanned by the gradient of the Hamiltonian alone. The introduction of the concepts of specifically and minimally degenerate brackets is justified by Proposition 3 for finite dimensional systems, and generalized without proofs to the case of infinite-dimensional systems in Section 3.3. In addition, we have generalized the Polyak–Łojasiewicz condition for the exponential convergence of gradient flows, to the case of metriplectic system. The finite-dimensional results have been extended to the infinite-dimensional case without proof, and supported by a number of examples. In Section 4, we have studied quite in detail a specific equilibrium problem for the reduced Euler equation. We constructed two relaxation methods based upon two metriplectic systems: one is specifically degenerate and the other one is not. The results of our numerical experiments with these two relaxation mechanisms can be explained in terms of the theoretical results.

In the second part of the paper, we have proposed a class of metric brackets that have been put forward as a generalization of the Landau collision operator, which is included in the class as a special case, Section 5. For this reason we propose the term “collision-like” brackets. Checking if collision-like brackets are specifically degenerate reduces to checking two separate conditions, and this is usually simpler.

We demonstrate the use of such brackets as the basis for relaxation methods for various equilibrium problems for both the reduced Euler equations and axisymmetric MHD equilibria, the latter being equivalent to solving the Grad-Shafranov equation. These are well-known equilibrium problems, for which various methods of solution exist, and are used here only as a proof of concept. From a purely computational point of view, the direct solution of the Grad-Shafranov equation, in particular, is usually faster than the relaxation method constructed here, but the latter provides a recipe that can be adapted to more complicated equilibrium problems. Specifically, we have in mind equilibrium problems in kinetic theories, such as the Vlasov-Maxwell system, drift- and gyro-kinetic equations. There is also the possibility of generalizing the variational formulation for the Grad-Shafranov equation to non-monotonic equilibrium profiles, which have not been treated in the present work.

At last, we have discussed a simplified class of brackets, in which the nonlocal nature of collision-like brackets is removed, Section 6. These brackets lead to diffusion-like evolution equations and reduce to a number of known brackets in special cases. Without the nonlocality, characteristic of collision-like operators, these diffusion-like brackets are not minimally degenerate, but they can still be used to construct relaxation methods for equilibrium problems with stronger constraints. In the simplest case, we recover the known magnetic relaxation method of Chodura and Schlüter, and we give a numerical example based upon a structure-preserving numerical scheme obtained via finite-element exterior calculus (FEEC).

The results presented in this paper are far from complete. The proof of convergence for infinite-dimensional systems has not been addressed, and the formal argument for the minimal degeneracy of the collision-like bracket in Example 7 is incomplete as it relies on a technical condition being true.

Our results however can have consequences in designing relaxation methods for equilibrium problems. Specifically, for equilibrium problems that can be characterized by a variational principle of the form (1), that is finding a (local) minimum of entropy subject to the constraint of energy and possibly other quantities being constant, one should make sure that the relaxation method is based on specifically degenerate brackets, with the kernel generated by the gradients of exactly the same quantities defining the constraints. If this is not the case, examples show that the relaxation method may not find a solution of the considered problem. Equilibria of the reduced Euler equations, Grad-Shafranov equilibria, and linear Beltrami fields are examples of problems that belong to this category. Alternatively, one might search for equilibria that can be characterized by entropy minima subject to stronger constraints. This is the case, for instance, of nonlinear Beltrami fields, which can be characterized as minima of magnetic energy over the set of fields that are smooth deformations (push-forward) of a given initial configuration, cf. Section 2.2.3 and Appendix C. For this type of problem the metric bracket cannot be specifically degenerate, but must be designed to satisfy the needed constraints.

The actual implementation of the relaxation methods can be more subtle, due to the fact that, depending on the initial condition, the corresponding evolution equation might not admit a smooth solution, so that low-regularity solutions need to be considered. As an example in Section 6.4, we have discussed the case of a relaxation method for Beltrami fields, and the need for a weak formulation of the evolution equation and the corresponding equilibrium condition.

As for physical metriplectic systems, Example 5 in Section 5.1 shows how the techniques developed here could be used to study physically relevant metric brackets (in this example, Morrison’s bracket for the Landau collision operator). By checking if the bracket is specifically degenerate, one can gain some information on the long-time limit of the solution. We find that, in general, Morrison’s bracket is not specifically degenerate with respect to the three collision invariants, but this is only due to the fact that the Landau collision operator acts pointwise in space.

CRediT authorship contribution statement

C. Bressan: Writing – review & editing, Visualization, Software, Methodology, Investigation, Data curation, Conceptualization; **M. Kraus:** Writing – review & editing, Supervision, Software, Methodology, Investigation, Conceptualization; **O. Maj:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **P.J. Morrison:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Data availability

No data was used for the research described in the article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Bilinear forms and Leibniz identity

In this appendix we recall the basic definition of a Poisson structure and give a formal derivation of (5). The material is standard but not always easy to find in textbooks. Let $V \subseteq L^2(\mathbb{R}^d, \mu; \mathbb{R}^N)$ be a Banach space of squared-integrable functions over a domain $\Omega \subseteq \mathbb{R}^d$ with values in $\mathbb{R}^N$. We define functional derivatives of a function in $C^1(V)$ as an element of $L^2(\Omega, \mu; \mathbb{R}^N)$ according (2) with the pairing given by the L^2 scalar product,

$$\langle w, v\rangle = \int_\Omega w(x)\cdot v(x) d\mu(x),$$

for all $v \in V$ and $w \in L^2(\Omega, \mu; \mathbb{R}^N)$, where μ is a measure on Ω, e.g., the Lebesgue measure.

We consider in particular the consequences of the Leibniz identity on the structure of the bracket. Let $\alpha : C^\infty(V) \times C^\infty(V) \to C^\infty(V)$ be a generic bi-linear map satisfying the symmetry condition

$$\alpha(\mathcal{F}, \mathcal{G}) = \sigma\alpha(\mathcal{G}, \mathcal{F}), \quad \sigma^2 = 1, \tag{A.1}$$

and the Leibniz identity,

$$\alpha(\mathcal{F}, \mathcal{G}\mathcal{H}) = \alpha(\mathcal{F}, \mathcal{G})\mathcal{H} + \mathcal{G}\alpha(\mathcal{F}, \mathcal{H}). \tag{A.2}$$

Hence, α is either symmetric or antisymmetric, depending on whether $\sigma = 1$ or $\sigma = -1$, and a derivation in all arguments. Both Poisson and metric brackets, cf. Section 2.1, are special cases of such a bilinear form.

Given the restriction functions $\mathcal{R}_{x,j}(u) = u_j(x)$, which are defined for functions u that are at least continuous, let

$$\mathscr{A}_{ij}(u; x, x') := \alpha(\mathcal{R}_{x,i}, \mathcal{R}_{x',j})(u),$$

and we have

$$\mathscr{A}_{ij}(u; x, x') = \sigma\mathscr{A}_{ji}(u; x', x).$$

We claim that a bilinear form α satisfying (A.1) and (A.2) is such that

1) it vanishes on constants, that is,

$$\mathcal{G}(v) = a \in \mathbb{R} \text{ for all } v \implies \alpha(\mathcal{F}, \mathcal{G}) = 0 \text{ for all } \mathcal{F};$$

2) it has the representation, for any $\mathcal{F}, \mathcal{G}$,

$$\alpha(\mathcal{F}, \mathcal{G}) = \sum_{i,j}\int_\Omega\int_\Omega \frac{\delta\mathcal{F}(u)}{\delta u_i}(x)\mathscr{A}_{ij}(u; x, x')\frac{\delta\mathcal{F}(u)}{\delta u_j}(x')d\mu(x')d\mu(x).$$

In particular, 2) implies Eq. (5).

Claim 1) follows from the bilinearity of the form and the Leibniz property. If $\mathcal{G}(v) = a$ is a constant function, for any pair of functions $\mathcal{F}$ and $\mathcal{H}$, one has

$$\begin{aligned} a\alpha(\mathcal{F}, \mathcal{H}) &= \alpha(\mathcal{F}, a\mathcal{H}) = \alpha(\mathcal{F}, \mathcal{G}\mathcal{H}) \\ &= \alpha(\mathcal{F}, \mathcal{G})\mathcal{H} + a\alpha(\mathcal{F}, \mathcal{H}), \end{aligned}$$

from which one deduces $\alpha(\mathcal{F}, \mathcal{G}) = 0$ as claimed.

Claim 2) requires Taylor's formula: For any $u, u_0 \in V$

$$\mathcal{F}(u) = \mathcal{F}(u_0) + \theta_\mathcal{F}(u_0, u)(u - u_0),$$

where, with $v = u - u_0$,

$$\begin{aligned} \theta_\mathcal{F}(u_0, u)v &= \int_0^1 D\mathcal{F}\big((1-t)u_0 + tu\big)v\,dt \\ &= \int_0^1\int_\Omega \frac{\delta\mathcal{F}}{\delta u}\big((1-t)u_0 + tu\big)\cdot v\,d\mu(x)dt \\ &= \sum_i\int_0^1\int_\Omega \frac{\delta\mathcal{F}}{\delta u_i}\big((1-t)u_0 + tu\big)v_i\,d\mu(x)dt. \end{aligned}$$

For a fixed u_0, $\mathcal{F}(u_0)$ and $\mathcal{G}(u_0)$ are constants and using claim 1) we have

$$\begin{aligned} \alpha(\mathcal{F}, \mathcal{G}) &= \alpha\big(\mathcal{F} - \mathcal{F}(u_0), \mathcal{G} - \mathcal{G}(u_0)\big) \\ &= \alpha\big(\theta_\mathcal{F}(u_0, \cdot)(\cdot - u_0), \theta_\mathcal{G}(u_0, \cdot)(\cdot - u_0)\big) \end{aligned}$$

$$= \sum_{i,j} \int_0^1 \int_0^1 \int_\Omega \int_\Omega A_{ij}\, dt\, ds\, d\mu(x) d\mu(x'),$$

where we have *formally* exchanged the integrals and the bi-linear form and for brevity we have defined

$$A_{ij} = \alpha\Big(\frac{\delta\mathcal{F}((1-t)u_0 + tu)}{\delta u_i}(x)(u(x) - u_0(x))_i, \\ \frac{\delta\mathcal{G}((1-s)u_0 + su)}{\delta u_j}(x')(u(x') - u_0(x'))_j \Big).$$

In the latter expression the first argument of α is the product of the functions $u \mapsto \delta\mathcal{F}((1-t)u_0 + tu)/\delta u_i|_x$ and $u \mapsto (u - u_0)_j|_x$; analogously for the second argument. We can use Leibniz identity and evaluate at $u = u_0$ with the result that

$$A_{ij} = \frac{\delta\mathcal{F}(u_0)}{\delta u_i}(x)\alpha\big(\mathcal{R}_{x,i}, \mathcal{R}_{x',j}\big)(u_0)\frac{\delta\mathcal{G}(u_0)}{\delta u_j}(x').$$

Therefore,

$$\alpha(\mathcal{F}, \mathcal{G})(u_0) = \sum_{i,j} \int_\Omega \int_\Omega \frac{\delta\mathcal{F}(u_0)}{\delta u_i}(x)\mathscr{A}_{ij}(u_0, x, x')\frac{\delta\mathcal{G}(u_0)}{\delta u_j}(x')d\mu(x')d\mu(x),$$

and since the point $u_0 \in V$ is arbitrary this is claim 2). This argument however is purely formal: The restriction function $\mathcal{R}_{x,j}$ is defined only for functions that can be evaluated at a point x, e.g. continuous functions, thus excluding L^p functions for any p. We have assumed that functional derivative exists and in exchanging the integral with the form α one needs some continuity in order to pass to the limit after approximating the integrals by finite sums. We have also freely exchanged the integration order.

Appendix B. On continuous bilinear forms on Hilbert spaces

Let H be a Hilbert space over the fields of real numbers and with scalar product $(\cdot,\cdot)$ and induced norm $\|\cdot\|$. If $a : H \times H \to \mathbb{R}$ is a continuous, positive bilinear form, where continuity means that there exists $C > 0$ such that

$$0 \le a(u, v) \le C\|u\|\|v\|,$$

for all $u, v \in H$, then one can find a bounded, symmetric, positive-definite linear operator $A : H \to H$ such that

$$a(u, v) = (u, Av),$$

for all $u, v \in H$. In order to find A one observes that, for any u fixed, $\ell(v) = a(u, v)$ is a bounded linear function from $H \to \mathbb{R}$. The Riesz representation theorem [65, Theorem 8.12] yields a unique element $w \in H$ such that

$$\ell(v) = (w, v),$$

for all $v \in H$ and $\|w\| = \sup\{\ell(v) : v \in H,\ \|v\| = 1\} \le C\|u\|$. Since w is unique, one can set $w = Au$, and A is a bounded linear operator on H. Then

$$a(u, v) = \ell(v) = (Au, v).$$

Positivity and symmetry follow from the positivity and symmetry of a.

Appendix C. A Lagrangian variational principle for Beltrami fields

In this appendix, we give a self-contained overview of the variational principle for Beltrami fields. This is the constant-pressure version of the variational principle for full MHD equilibria obtained by Kendall [96], and formulated in a modern language.

While the special case of linear Beltrami fields obey Woltjer's principle of least magnetic energy at constant magnetic helicity, cf. Section 2.2.3, general Beltrami fields minimize energy under a much stronger constraint.

On a bounded (not necessarily simply connected) domain $\Omega \subset \mathbb{R}^3$, we fix a reference magnetic field $B_0 \in V$, where V is the space of vector fields $B \in [L^2(\Omega)]^3$ satisfying the conditions

$$\begin{aligned} \operatorname{div} B &= 0, && \text{in } \Omega, \\ n \cdot B &= 0, && \text{on } \partial\Omega. \end{aligned} \tag{C.1}$$

For any $\Phi : \Omega \to \Omega$ an element of the group $\mathrm{Diff}(\Omega)$ of diffeomorphisms of the domain Ω, we define

$$B = \Phi_* B_0 = \frac{D\Phi B_0}{\det D\Phi} \circ \Phi^{-1}, \tag{C.2}$$

where $D\Phi$ is the Jacobian matrix of Φ (defined by $(D\Phi)_{ij} = \partial_{x_j}\Phi_i$) and $\det D\Phi \neq 0$ is its determinant. Then B is the push-forward of the fields B_0 with the map Φ. A direct calculation show that

$$(\det D\Phi) \operatorname{div} B = \operatorname{div} B_0,$$

hence div $B_0 = 0$ imply div $B = 0$. Analogously one can show that the boundary condition $B_0 \cdot n = 0$ on $\partial\Omega$ is preserved by the diffeomorphism since if $x = \Phi(x_0)$ and $x_0 \in \partial\Omega$, then $x \in \partial\Omega$,

$$n(x) = \frac{{}^t D\Phi^{-1}(x) n(x_0)}{|{}^t D\Phi^{-1}(x) n(x_0)|},$$
$$n(x) \cdot B(x) = \frac{B_0(x_0) \cdot n(x_0)}{\det D\Phi \cdot |{}^t D\Phi^{-1}(x) n(x_0)|}.$$

(This can be proven by recalling that for a sufficiently regular domain, near a point $x_0 \in \partial\Omega$ there is a function f such that $f > 0$ in Ω and $f = 0$ on $\partial\Omega$; then $n(x_0) \propto \nabla f(x_0)$ and this transforms like a 1-form under Φ.) Therefore the push-forward formula (C.2) maps $B_0 \in V$ into $B \in V$.

Given $B_0 \in V$, we define the entropy functional on the group $\mathrm{Diff}(\Omega)$ as the magnetic energy stored in B, that is

$$\mathcal{S}(\Phi) = \mathcal{S}(\Phi; B_0) = \int_\Omega \frac{|B|^2}{8\pi} dx, \tag{C.3}$$

where $B = \Phi_* B_0$. The entropy depends parametrically on the initial field B_0.

We can now state the variational principle for (24). For any $B_0 \in V$ fixed, if Φ is a critical point of (C.3), then $B = \Phi_* B_0 \in V$ satisfies (24). More explicitly, this means that if

$$\frac{d}{d\varepsilon} \mathcal{S}(\Phi^\varepsilon)\Big|_{\varepsilon=0} = 0, \tag{C.4}$$

for any curve $\varepsilon \mapsto \Phi^\varepsilon \in \mathrm{Diff}(\Omega)$ such that $\Phi^\varepsilon|_{\varepsilon=0} = \Phi$, then $B = \Phi_* B_0$ is a Beltrami field obtained by mapping the given field B_0 by the action of the diffeomorphism Φ.

In order to prove the variational principle (C.4) let $B^\varepsilon(x) = \Phi^\varepsilon_* B_0(x)$, and introduce the displacement field

$$\xi^\varepsilon = \partial_\varepsilon \Phi^\varepsilon \circ (\Phi^\varepsilon)^{-1}. \tag{C.5}$$

The definition is equivalent to

$$\partial_\varepsilon \Phi^\varepsilon(x_0) = \xi^\varepsilon(x), \quad x = \Phi^\varepsilon(x_0).$$

Then, one obtains

$$\partial_\varepsilon B^\varepsilon = \mathrm{curl}(\xi^\varepsilon \times B^\varepsilon), \quad B^\varepsilon|_{\varepsilon=0} = B, \tag{C.6}$$

and we compute from (C.4),

$$\frac{d}{d\varepsilon} \mathcal{S}(\Phi^\varepsilon)\Big|_{\varepsilon=0} = -\int_\Omega \left[\frac{1}{4\pi}(\mathrm{curl}\, B) \times B\right] \cdot \xi dx + \int_\Omega n \cdot [(\xi \times B) \times B] d\sigma = 0,$$

where $\xi = \xi^\varepsilon|_{\varepsilon=0}$. The boundary term vanishes due to the identity $(\xi \times B) \times B = (B \cdot \xi)B - B^2 \xi$ and the boundary condition $B \cdot n = 0$, $\xi \cdot n = 0$; the latter follows from the fact that Φ preserves the boundary, hence $\xi|_{\partial\Omega}$ must be tangent to $\partial\Omega$. Since the derivative of $\mathcal{S}(\Phi^\varepsilon)$ has to vanish for any curve $\Phi^\varepsilon \in \mathrm{Diff}(\Omega)$ and thus for every ξ, we deduce that B satisfies (24).

We remark that, since B is the push-forward of a known field B_0, the field-line topology of B is the same as that of B_0. Magnetic helicity is also preserved.

This is a variant of the variational principle (C.4) at the basis of the relaxation method of Chodura and Schlüter [30,31], Moffatt [23], and of the SIESTA code [33].

References

[1] Morrison PJ. Bracket formulation for irreversible classical fields. Phys Lett A 1984;100(8):423–7. https://doi.org/10.1016/0375-9601(84)90635-2
[2] Morrison PJ. Some observations regarding brackets and dissipation. Tech. Rep. PAM–228; University of California at Berkeley; 1984. Available at: 15 Mar 2024 arXiv:2403.14698v1.
[3] Morrison PJ. A paradigm for joined Hamiltonian and dissipative systems. Phys D Nonlinear Phenom 1986;18:410–19. https://doi.org/10.1016/0167-2789(86)90209-5
[4] Morrison PJ. Hamiltonian description of the ideal fluid. Rev Mod Phys 1998;70:467–521. https://doi.org/10.1103/RevModPhys.70.467
[5] Freidberg JP. Ideal MHD. Cambridge University Press; 2014. https://doi.org/10.1017/CBO9780511795046
[6] Berk HL, Freidberg JP, Llobet X, Morrison PJ, Tataronis JA. Existence and calculation of sharp boundary magnetohydrodynamic equilibrium in three-dimensional toroidal geometry. Phys Fluids 1986;29(10):3281–90. https://doi.org/10.1063/1.865845.
[7] de la Llave R. A tutorial on KAM theory. American Mathematical Society. ISBN 9780821893746; 2001, pp. 175–292. https://doi.org/10.1090/pspum/069/1858536
[8] Bruno OP, Laurence P. Existence of three-dimensional toroidal MHD equilibria with nonconstant pressure. Commun Pure Appl Math 1996;49(7):717–64. https://doi.org/10.1002/(SICI)1097-0312(199607)49:7<717::AID-CPA3>3.0.CO;2-C
[9] Enciso A, Luque A, Peralta-Salas D. Mhd equilibria with nonconstant pressure in nondegenerate toroidal domains. J Eur Math Soc 2025;27(6):2251–91. https://doi.org/10.4171/JEMS/1410
[10] Bauer F, Betancourt O, Garabedian P. A computational method in plasma physics. Springer-Verlag; 1978. https://doi.org/10.1007/978-3-642-85470-5
[11] Hirshman SP, Whitson JC. Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria. Phys Fluids 1983;26(12):3553–68. https://doi.org/10.1063/1.864116
[12] Kruskal MD, Kulsrud RM. Equilibrium of a magnetically confined plasma in a toroid. Phys Fluids 1958;1(4):265–74. https://doi.org/10.1063/1.1705884
[13] Grad H. Toroidal containment of a plasma. Phys Fluids 1967;10(1):137–54. https://doi.org/10.1063/1.1761965
[14] Garabedian PR. Magnetohydrodynamic stability of fusion plasmas. Commun Pure Appl Math 1998;51(9–10):1019–33. https://doi.org/10.1002/(SICI)1097-0312(199809/10)51:9/10<1019::AID-CPA4>3.0.CO;2-G

[15] Garabedian PR. Three-dimensional analysis of tokamaks and stellarators. Proc Natl Acad Sci 2008;105(37):13716–19. https://doi.org/10.1073/pnas.0806354105
[16] Dudt DW, Kolemen E. Desc: a stellarator equilibrium solver. Phys Plasmas 2020;27(10):102513. https://doi.org/10.1063/5.0020743
[17] Hindenlang F, Plunk GG, Maj O. Computing MHD equilibria of stellarators with a flexible coordinate frame. Plasma Phys Control Fusion 2025;67(4):045002. https://doi.org/10.1088/1361-6587/adba11
[18] Imbert-Gérard L-M, Paul EJ, Wright AM. An introduction to stellarators: from magnetic fields to symmetries and optimization. Society for Industrial and Applied Mathematics; 2024. ISBN 9781611978223. https://doi.org/10.1137/1.9781611978223
[19] Grad H, Rubin H. Hydromagnetic equilibria and force-free fields. J Nucl Energy 1958;7(3–4):284–5. https://doi.org/10.1016/0891-3919(58)90139-6
[20] Reiman A, Greenside H. Calculation of three-dimensional MHD equilibria with islands and stochastic regions. Comput Phys Commun 1986;43(1):157–67. https://doi.org/10.1016/0010-4655(86)90059-7
[21] Moffatt HK. Magnetostatic equilibria and analogous Euler flows of arbitrarily complex topology. Part 1. Fundamentals. J Fluid Mech 1985;159:359–78. https://doi.org/10.1017/S0022112085003251
[22] Arnold VI, Khesin BA. Topological methods in hydrodynamics. Springer-Verlag; 1998.
[23] Moffatt HK. Some topological aspects of fluid dynamics. J Fluid Mech 2021;914:P1. https://doi.org/10.1017/jfm.2020.230
[24] Alfvén H. Existence of electromagnetic-hydrodynamic waves. Nature 1942;150:405–6. https://doi.org/10.1038/150405d0
[25] Davidson PA. An introduction to magnetohydrodynamics. Cambridge texts in applied mathematics; Cambridge University Press; 2001. ISBN 9780521794879.
[26] Parker EN. Topological dissipation and the small-scale fields in turbulent gases. Astrophys J 1972;174:499. https://doi.org/10.1086/151512
[27] Enciso A, Peralta-Salas D. Obstructions to topological relaxation for generic magnetic fields. Arch Ration Mech Anal 2025;249(1):6. https://doi.org/10.1007/s00205-024-02078-5
[28] Harafuji K, Hayashi T, Sato T. Computational study of three-dimensional magnetohydrodynamic equilibria in toroidal helical systems. J Comput Phys 1989;81(1):169–92. https://doi.org/10.1016/0021-9991(89)90069-7
[29] Suzuki Y, Nakajima N, Watanabe K, et al. Development and application of HINT2 to helical system plasmas. Nucl Fusion 2006;46(11):L19. https://doi.org/10.1088/0029-5515/46/11/L01
[30] Chodura R, Schlueter A. A 3D code for MHD equilibrium and stability. J Comput Phys 1981;41(1):68–88. https://doi.org/10.1016/0021-9991(81)90080-2
[31] Wiegelmann T, Sakurai T. Solar force-free magnetic fields. Living Rev Sol Phys 2012;9(1):5. https://doi.org/10.1007/s41116-020-00027-4
[32] Brenier Y, Duan X. An integrable example of gradient flow based on optimal transport of differential forms. Calc Var Partial Differ Equ 2018;57(5):125. https://doi.org/10.1007/s00526-018-1370-6
[33] Hirshman SP, Sanchez R, Cook CR. SIESTA: a scalable iterative equilibrium solver for toroidal applications. Phys Plasmas 2011;18(6). 062504. https://doi.org/10.1063/1.3597155
[34] Coquinot B, Morrison PJ. A general metriplectic framework with application to dissipative extended magnetohydrodynamics. J Plasma Phys 2020;86(3):835860302. https://doi.org/10.1017/s0022377820000392
[35] Morrison PJ, Updike MH. Inclusive curvature-like framework for describing dissipation: metriplectic 4-bracket dynamics. Phys Rev E 2024;109:045202.
[36] Zaidni A, Morrison PJ, Benjelloun S. Thermodynamically consistent Cahn-Hilliard-Navier-Stokes equations using the metriplectic dynamics formalism. Phys D 2024;468:134303. https://doi.org/10.1016/j.physd.2024.134303
[37] Sato N, Morrison PJ. A collision operator for describing dissipation in noncanonical phase space. Fund Plasma Phys 2024;10:100054.
[38] Sato N, Morrison PJ. Scattering theory in noncanonical phase space: a drift-kinetic collision operator for weakly collisional plasmas. Phys Plasmas 2025;32(10). https://doi.org/https://doi.org/10.1063/5.0289410
[39] Zaidni A, Morrison PJ. Metriplectic four-bracket algorithm for constructing thermodynamically consistent dynamical systems. Phys Rev E 2025;112(2):025101. https://doi.org/10.1103/r2lb-xkq6
[40] Hirsch MW, Smale S, Devaney RL. Differential equations, dynamical systems, and an introduction to chaos. Elsevier; 3rd ed.; 2013. https://doi.org/10.1016/C2009-0-61160-0
[41] Materassi M. Metriplectic algebra for dissipative fluids in Lagrangian formulation. Entropy 2015;17(3):1329–46. https://doi.org/10.3390/e17031329
[42] Materassi M, Tassi E. Metriplectic framework for dissipative magneto-hydrodynamics. Phys D Nonlinear Phenom 2012;241(6):729–34. https://doi.org/10.1016/j.physd.2011.12.013
[43] Kraus M, Hirvijoki E. Metriplectic integrators for the Landau collision operator. Phys Plasmas 2017;24(10):102311. https://doi.org/10.1063/1.4998610
[44] Barham W, Morrison PJ, Zaidni A. A thermodynamically consistent discretization of 1D thermal-fluid models using their metriplectic 4-bracket structure. Commun Nonlinear Sci Numer Simul 2025;145:108683. https://doi.org/10.2139/ssrn.5035052
[45] Vallis GK, Carnevale GG, Young WR. Extremal energy properties and construction of stable solutions of the Euler equations. J Fluid Mech 1989;207:133–52. https://doi.org/10.1017/S0022112089002533
[46] Carnevale GG, Vallis GK. Pseudo-advective relaxation to stable states of inviscid two-dimensional fluids. J Fluid Mech 1990;213:549–71. https://doi.org/10.1017/S0022112090002440
[47] Flierl GR, Morrison PJ. Hamiltonian Dirac simulated annealing: application to the calculation of vortex states. Phys D Nonlinear Phenom 2011;240(2):212–32. “Nonlinear Excursions” Symposium and Volume in Physica D to honor Louis N. Howard’s scientific career. https://doi.org/10.1016/j.physd.2010.08.011
[48] Chikasue Y, Furukawa M. Adjustment of vorticity fields with specified values of Casimir invariants as initial condition for simulated annealing of an incompressible, ideal neutral fluid and its MHD in two dimensions. J Fluid Mech 2015;774:443–59. https://doi.org/10.1017/jfm.2015.263
[49] Chikasue Y, Furukawa M. Simulated annealing applied to two-dimensional low-beta reduced magnetohydrodynamics. Phys Plasmas 2015;22(2):022511. https://doi.org/10.1063/1.4913234
[50] Furukawa M, Morrison PJ. Simulated annealing for three-dimensional low-beta reduced MHD equilibria in cylindrical geometry. Plasma Phys Control Fusion 2017;59(5):054001. https://doi.org/10.1088/1361-6587/aa5863
[51] Furukawa M, Watanabe T, Morrison PJ, et al. Calculation of large-aspect-ratio tokamak and toroidally-averaged stellarator equilibria of high-beta reduced magnetohydrodynamics via simulated annealing. Phys Plasmas 2018;25(8):082506. https://doi.org/10.1063/1.5038043
[52] Furukawa M, Morrison PJ. Simulated annealing of reduced magnetohydrodynamic systems. Rev Mod Plasma Phys 2025;9(1):15. https://doi.org/10.1007/s41614-025-00185-8
[53] Brockett RW. Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems. Linear Algebra Appl 1991;146:79–91. https://doi.org/10.1109/cdc.1988.194420
[54] Bloch AM, Brockett RW, Ratiu TS. Completely integrable gradient flows. Commun Math Phys 1992;147(1):57–74. https://doi.org/10.1007/bf02099528
[55] Bloch AM, Morrison PJ, Ratiu TS, et al. Flows in the normal and Kaehler metrics and triple bracket generated metriplectic systems. in: JohannA et al.(editor), Recent Trends in Dynamical Systems, Springer Proceedings in Mathematics & Statistics 2013; Vol. 35, pp.371–415.
[56] Hasegawa A. Self-organization processes in continuous media. Adv Phys 1985;34(1):1–42. https://doi.org/10.1080/00018738500101721
[57] Yoshida Z, Mahajan SM. Variational principles and self-organization in two-fluid plasmas. Phys Rev Lett 2002;88:095001. https://doi.org/10.1103/PhysRevLett.88.095001
[58] Woltjer L. On hydromagnetic equilibrium. Proc Natl Acad Sci 1958;44(9):833–41. https://doi.org/10.1073/pnas.44.9.833
[59] Taylor JB. Relaxation of toroidal plasma and generation of reverse magnetic fields. Phys Rev Lett 1974;33:1139–41. https://doi.org/10.1103/PhysRevLett.33.1139
[60] Taylor JB. Relaxation and magnetic reconnection in plasmas. Rev Mod Phys 1986;58:741–63. https://doi.org/10.1103/revmodphys.58.741
[61] Qin H, Liu W, Li H, et al. Woltjer-Taylor state without Taylor’s conjecture: plasma relaxation at all wavelengths. Phys Rev Lett 2012;109:235001. https://doi.org/10.1103/PhysRevLett.109.235001
[62] Adams MF, Hirvijoki E, Knepley MG, et al. Landau collision integral solver with adaptive mesh refinement on emerging architectures. SIAM J Sci Comput 2017;39(6):C452–C465. https://doi.org/10.1137/17M1118828

[63] Gay-Balmaz F, Holm DD. Selective decay by Casimir dissipation in inviscid fluids. Nonlinearity 2013;26(2):495. https://doi.org/10.1088/0951-7715/26/2/495
[64] Gay-Balmaz F, Holm DD. A geometric theory of selective decay with applications in MHD. Nonlinearity 2014;27(8):1747. https://doi.org/10.1088/0951-7715/27/8/1747
[65] Hunter JK, Nachtergaele B. Applied analysis. World Scientific Publishing Company; 2001.
[66] Marsden JE, Ratiu T, Abraham R. Manifolds, tensor analysis, and applications. Springer; 3rd ed.; 2001.
[67] Morrison PJ. Poisson brackets for fluids and plasmas. Cont Math 1982;88:13–46. https://doi.org/10.1063/1.33633
[68] Meyer KR. Symmetries and integrals in mechanics. In: Peixoto MM, editor. Dynamical systems. Academic Press; 1973, pp. 259–72. https://doi.org/10.1016/B978-0-12-550350-1.50025-4
[69] Marsden J, Weinstein A. Reduction of symplectic manifolds with symmetry. Rep Math Phys 1974;5(1):121–30. https://www.sciencedirect.com/science/article/pii/0034487774900214.
[70] Marsden JE, Ratiu TS. Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems. Springer New York; 1999. ISBN 9780387217925. https://doi.org/10.1007/978-0-387-21792-5
[71] Holm DD, Schmah T, Stoica C, Ellis DCP. Geometric mechanics and symmetry: from finite to infinite dimensions. Oxford texts in applied and engineering mathematics; Oxford University Press; 2009. ISBN 9780199212910.
[72] Morrison PJ, Eliezer S. Spontaneous symmetry breaking and neutral stability in the noncanonical Hamiltonian formalism. Phys Rev A 1986;33:4205–14.
[73] Yoshida Z, Tokieda T, Morrison PJ. Rattleback: a model of how geometric singularity induces dynamic chirality. Phys Lett A 2017;381:2772–7.
[74] Yoshida Z, Morrison PJ. Deformation of Lie-Poisson algebras and chirality. J Math Phys 2020;61:082901.
[75] Yoshida Z, Morrison PJ. Hierarchical structure of noncanonical Hamiltonian systems. Phys Scr 2016;91(2):024001. https://doi.org/10.1088/0031-8949/91/2/024001
[76] Taylor ME. Partial differential equations III: nonlinear equations; vol. 117 of Applied Mathematical Sciences. Springer New York; 2011. https://doi.org/10.1007/978-1-4419-7049-7
[77] Takeda T, Tokuda S. Computation of MHD equilibrium of tokamak plasma. J Comput Phys 1991;93(1):1–107. https://doi.org/10.1016/0021-9991(91)90074-u
[78] Taylor ME. Partial Differential Equations I: Basic Theory; vol. 115 of Applied Mathematical Sciences. Springer New York; 2011. https://doi.org/10.1007/978-1-4419-7055-8
[79] Grad H. Some new variational properties of hydromagnetic equilibria. Phys Fluids 1964;7(8):1283–92. https://doi.org/10.1063/1.1711373
[80] LoDestro LL, Pearlstein LD. On the Grad-Shafranov equation as an eigenvalue problem, with implications for q solvers. Phys Plasmas 1994;1(1):90–5. https://doi.org/10.1063/1.870464
[81] Pataki A, Cerfon AJ, Freidberg JP, et al. A fast, high-order solver for the Grad-Shafranov equation. J Comput Phys 2013;243:28–45. https://doi.org/10.1016/j.jcp.2013.02.045
[82] Boulmezaoud TZ, Amari T. On the existence of non-linear force-free fields in three-dimensional domains. Zeitschrift für angewandte Mathematik und Physik ZAMP 2000;51(6):942–67. https://doi.org/10.1007/pl00001531
[83] Amari T, Boulbe C, Boulmezaoud TZ. Computing Beltrami fields. SIAM J Sci Comput 2009;31(5):3217–54. https://doi.org/10.1137/070700942
[84] Enciso A, Peralta-Salas D. Beltrami fields with a nonconstant proportionality factor are rare. Arch Ration Mech Anal 2016;220(1):243–60. https://doi.org/10.1007/s00205-015-0931-5
[85] Hu K, Ma Y, Xu J. Stable finite element methods preserving $\nabla \cdot B = 0$ exactly for MHD models. Numer Math 2017;135(2):371–96. https://doi.org/10.1007/s00211-016-0803-4
[86] Hu K, Lee Y-J, Xu J. Helicity-conservative finite element discretization for incompressible MHD systems. J Comput Phys 2021;436:110284. https://doi.org/10.1016/j.jcp.2021.110284
[87] Yoshida Z, Giga Y. Remarks on spectra of operator rot. Math Z 1990;204(1):235–45. https://doi.org/10.1007/BF02570870
[88] Boulmezaoud T-Z, Maday Y, Amari T. On the linear force-free fields in bounded and unbounded three-dimensional domains. ESAIM: M2AN 1999;33(2):359–93. https://doi.org/10.1051/m2an:1999121
[89] Dewar RL, Hole MJ, McGann M, et al. Relaxed plasma equilibria and entropy-related plasma self-organization principles. Entropy 2008;10(4):621–34. https://doi.org/10.3390/e10040621
[90] Hudson SR, Dewar RL, Dennis G, et al. Computation of multi-region relaxed magnetohydrodynamic equilibria. Phys Plasmas 2012;19(11):112502. https://doi.org/10.1063/1.4765691
[91] Dixon AM, Berger MA, Browning PK, Priest ER. A generalization of the Woltjer minimum-energy principle. Astron Astrophys 1989;225:156–66.
[92] Laurence P, Avellaneda M. On Woltjer's variational principle for force-free fields. J Math Phys 1991;32(5):1240–53. https://doi.org/10.1063/1.529321
[93] Moffatt HK. The degree of knottedness of tangled vortex lines. J Fluid Mech 1969;null:117–29. https://doi.org/10.1017/s0022112069000991
[94] Arnold VI. The asymptotic Hopf invariant and its applications. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN 978-3-642-31031-7; 2014, p. 357–75. https://doi.org/10.1007/978-3-642-31031-7_32
[95] Vogel T. On the asymptotic linking number. Proc Am Math Soc 2003;131(7):2289–97. https://doi.org/10.1090/s0002-9939-02-06792-8
[96] Kendall PC. The variational formulation of the magneto-hydrostatic equations. Astrophys J 1960;131:681.
[97] Malhotra D, Cerfon A, Imbert-Gérard L-M, et al. Taylor states in stellarators: a fast high-order boundary integral solver. J Comput Phys 2019;397:108791. https://doi.org/10.1016/j.jcp.2019.06.067
[98] Henry D. Geometric theory of semilinear parabolic equations. Springer, Berlin; 1981.
[99] Temam R. Infinite-dimensional dynamical systems in mechanics and physics. Springer-Verlag New York; 1998.
[100] Wiggins S. Introduction to applied nonlinear dynamical systems and chaos. Springer New York, NY; 2003.
[101] Moretti V. Analytical mechanics. Springer Nature Switzerland; 2023. ISBN 9783031276125.
[102] La Salle J, Lefschetz S. Stability by Liapunov's direct method with applications. Academic Press; 1961.
[103] Beretta GP. A theorem on Lyapunov stability for dynamical systems and a conjecture on a property of entropy. J Math Phys 1986;27(1):305–8. https://doi.org/10.1063/1.527390
[104] Polyak BT. Gradient methods for the minimisation of functionals. USSR Comput Math Math Phys 1963;3(4):864–78. https://doi.org/10.1016/0041-5553(63)90382-3
[105] Łojasiewicz S. Sur les trajectoires du gradient d'une fonction analytique. In: Seminari di geometria 1982–1983. Universita di Bologna, Istituto di Geometria, Dipartamento di Matematica; 1984, p. 115–17.
[106] Palis J, Melo W. Geometric theory of dynamical systems. Springer New York; 1982. https://doi.org/10.1007/978-1-4612-5703-5
[107] Karimi H, Nutini J, Schmidt M. Linear convergence of gradient and proximal-gradient methods under the Polyak-Łojasiewicz condition. In: Machine learning and knowledge discovery in databases. Cham: Springer International Publishing; 2016, pp. 795–811.
[108] Rauch J. Hyperbolic partial differential equations and geometric optics. Graduate studies in mathematics; American Mathematical Society; 2012. ISBN 9780821872918.
[109] Denny D. Existence of a solution to a semilinear elliptic equation. AIMS Math 2016;1(3):208–11. https://doi.org/10.3934/Math.2016.3.208
[110] Bressan C. Metriplectic relaxation for calculating equilibria: theory and structure-preserving discretization. Ph.D. thesis; Technische Universität München; 2023. https://mediatum.ub.tum.de/1686142.
[111] Bressan C, Kraus M, Morrison PJ, et al. Relaxation to magnetohydrodynamics equilibria via collision brackets. J Phys Conf Ser 2018;1125:012002. https://doi.org/10.1088/1742-6596/1125/1/012002
[112] Evans LC. Partial differential equations. Graduate studies in mathematics; Providence (R.I.): American Mathematical Society; 1998. ISBN 0-8218-0772-2.
[113] Lenard A. On Bogoliubov's kinetic equation for a spatially homogeneous plasma. Ann Phys 1960;10(3):390–400.
[114] Villani C. Conservative forms of Boltzmann's collision operator : Landau revisited. ESAIM: Modélisation mathématique et analyse numérique 1999;33(1):209–27. https://doi.org/10.1051/m2an:1999112

[115] Villani C. Hypocoercive diffusion operators. Boll Unione Mat Ital 2007;10-B(2):257–75. https://doi.org/10.4171/022-3/25

[116] Monk P. Finite element methods for Maxwell's equations. Oxford Un. Press; 2003.

[117] Alnaes MS, Blechta J, Hake J, Johansson A, Kehlet B, Logg A, et al. The FEniCS project version 1.5. Arch Numer Softw 2015;3. https://doi.org/10.11588/ans.2015.100.20553

[118] Logg A, Mardal K-A, Wells GN, et al. Automated solution of differential equations by the finite element method. Springer; 2012. https://doi.org/10.1007/978-3-642-23099-8

[119] Alnaes MS, Logg A, Ølgaard KB, Rognes ME, Wells GN. Unified form language: a domain-specific language for weak formulations of partial differential equations. ACM Trans Math Softw 2014;40. https://doi.org/10.1145/2566630

[120] Logg A, Wells GN. DOLFIN: Automated finite element computing. ACM Trans Math Softw 2010;37. https://doi.org/10.1145/1731022.1731030

[121] Logg A, Wells GN, Hake J. DOLFIN: a C++/Python finite element library. In: Logg A, Mardal K-A, Wells GN, editors. Automated solution of differential equations by the finite element method; vol. 84 of Lecture Notes in Computational Science and Engineering; chap. 10. Springer; 2012.

[122] Herrnegger F. In: Proceedings of 5th European conference on controlled fusion and plasma physics, Grenoble. European Physical Society; 1972, p. 26.

[123] Maschke EK. Exact solutions of the MHD equilibrium equation for a toroidal plasma. Plasma Phys 1973;15(6):535. https://doi.org/10.1088/0032-1028/15/6/006

[124] Mc Carthy PJ. Analytical solutions to the Grad-Shafranov equation for tokamak equilibrium with dissimilar source functions. Phys Plasmas 1999;6(9):3554–60. https://doi.org/10.1063/1.873630

[125] Zoni E, Güçlü Y, et al. Solving hyperbolic-elliptic problems on singular mapped disk-like domains with the method of characteristics and spline finite elements. J Comput Phys 2019;398:108889. https://doi.org/10.1016/j.jcp.2019.108889

[126] Yang WH, Sturrock PA, Antiochos SK. Force-free magnetic fields: the magneto-frictional method. Astrophys J 1986;309:383–91.

[127] Klimchuk JA, Sturrock PA. Three-dimensional force-free magnetic fields and flare energy buildup. Astrophys J 1992;385:344–53.

[128] Valori G, Kliem B, Fuhrmann M. Magnetofrictional extrapolations of Low and Lou's force-free equilibria. Sol Phys 2007;245(2):263–85. https://doi.org/10.1007/s11207-007-9046-y

[129] Valori G, Kliem B, Török T, Titov VS. Testing magnetofrictional extrapolation with the Titov-Démoulin model of solar active regions. A&A 2010;519. https://doi.org/10.1051/0004-6361/201014416

[130] He M, Farrell PE, Hu K, Andrews BD. Topology-preserving discretization for the magneto-frictional equations arising in the Parker conjecture 2025;https://doi.org/10.48550/arXiv.2501.11654

[131] Candelaresi S, Pontin D, Hornig G. Mimetic methods for Lagrangian relaxation of magnetic fields. SIAM J Sci Comput 2014;36(6):B952–B968. https://doi.org/10.1137/140967404

[132] Guo Y, Xia C, Keppens R, Valori G. Magneto-frictional modeling of coronal nonlinear force-free fields. I. Testing with analytic solutions. Astrophys J 2016;828(2):82. https://doi.org/10.3847/0004-637X/828/2/82

[133] Padilla M, Gross O, Knöppel F, Chern A, Pinkall U, Schröder P. Filament based plasma. ACM Trans Graph 2022;41(4). https://doi.org/10.1145/3528223.3530102

[134] Gross O, Pinkall U, Schröder P. Plasma knots. Phys Lett A 2023;480:128986.

[135] Yeates AR. On the limitations of magneto-frictional relaxation. Geophys Astrophys Fluid Dyn 2022;116(4):305–20. https://doi.org/10.1080/03091929.2021.2021197