On the information content of the difference from hamiltonian evolution

Marius Buliga
Institute of Mathematics of the Romanian Academy

Abstract

A dissipative version of hamiltonian mechanics is proposed via a principle of minimal information content of the deviation from hamiltonian evolution. We show that we can cover viscosity, plasticity, damage and unilateral contact.

1 Introduction

In hamiltonian mechanics a physical system is described by a state vector $q \in X$ and a momentum vector $p \in Y$. The evolution in time of the system is governed by a hamiltonian function $H = H(q, p, t)$, via the equations:

$$\begin{cases} \dot{q} & = & \frac{\partial H}{\partial p}(q, p, t) \\ -\dot{p} & = & \frac{\partial H}{\partial q}(q, p, t) \end{cases} \tag{1}$$

where $\dot{q}$, $\dot{p}$ denote derivatives with respect to time. The evolution is reversible.

We propose a dissipative modification of these equations, which extends the work from [2], [3], [4].

Definition 1.1 *Given a hamiltonian H and a likelihood function π*

$$H : X \times Y \times \mathbb{R} \to \mathbb{R} \quad , \quad \pi : (X \times Y)^3 \to [0, 1]$$

a physical system described by the state vector $q \in X$ and the momentum vector $p \in Y$ evolves such that the vector $\eta = (\eta_q, \eta_p) \in X \times Y$ defined by:

$$\begin{cases} \dot{q} & = & \frac{\partial H}{\partial p}(q, p, t) + \eta_q \\ -\dot{p} & = & \frac{\partial H}{\partial q}(q, p, t) + \eta_p \end{cases} \tag{2}$$

called the difference (or gap) from hamiltonian evolution driven by H, maximizes the likelihood:

$$\pi(z, \dot{z}, \eta) = \max_{z'' \in X \times Y} \pi(z, \dot{z}, z'') \tag{3}$$

with the notations $z = (q, p) \in X \times Y$, $\dot{z} = (\dot{p}, \dot{q}) \in X \times Y$.

The likelihood function π cannot be taken arbitrarily, otherwise the proposed modification would be too vague. We propose the following constraints on the likelihood function.

First we define the information content function associated to the likelihood π as

$$I : (X \times Y)^3 \to [0, +\infty] \ , \ I(z, z', z") \ = \ -\ln \pi(z, z', z") \tag{4}$$

with the convention that $-\ln 0 = +\infty$.

Definition 1.2 *The likelihood π and the associated information content I satisfy:*

(a) *The information content function (4) is convex in each of the 2nd and 3rd variables and it has the needed degree of smoothness required (for example it is lower semi-continuous with respect to the relevant topologies on X and Y)*

(b) *for any $z, z' \in X \times Y$, the following maxima exist*

$$\max_{z" \in X \times Y} \pi(z, z', z") \, , \ \max_{z" \in X \times Y} \pi(z, z", z')$$

and they are either 0 or 1.

The equation (3) can be rephrased as: given the hamiltonian H and the information content function I, the physical system evolves such that at any moment it minimizes the information content of the gap from a hamiltonian evolution. Indeed, an evolution of the system is a curve $c_0 : [0, T] \to X \times Y$ with the property that it minimizes the information content gap functional:

$$G(c) = \int_0^T I \left(c(t), \dot c(t), \dot q(t) - \frac{\partial H}{\partial p}(c(t), t), -\dot p(t) - \frac{\partial H}{\partial q}(c(t), t) \right) dt \tag{5}$$

among all admissible evolution curves $c(t) = (q(t), p(t))$.

Pure Hamiltonian evolution. Let's pick the information content (4) to be:

$$I(z, z', z") \ = \ \chi_0(z") = \begin{cases} 0 & \text{if } z" = 0 \\ +\infty & \text{otherwise} \end{cases}$$

This corresponds to a likelihood function:

$$\pi(z, z', z") \ = \ \begin{cases} 1 & \text{if } z" = 0 \\ 0 & \text{otherwise} \end{cases}$$

The maximization of the likelihood (3) implies that the gap vector $\eta = 0$, therefore the evolution equations (2) reduce to the pure Hamiltonian evolution equations (1).

This example is trivial, we need a method to construct more interesting ones. One such method is based on the following observation, adapted from [6], section 2. We use the notations explained in section 3, in particular we use the duality $\langle\langle \cdot, \cdot \rangle\rangle$ and we suppose that we have on X, Y a topology compatible with it, so that the information content function I is lower semicontinuous (lsc).

2

Proposition 1.3 *The information content function (4) satisfies the conditions from Definition 1.2 if and only if the function*

$$b : (X \times Y)^3 \rightarrow \mathbb{R} \cup \{+\infty\} \ , \ b(z, z', z") \ = \ I(z, z', z") + \langle\langle z', z" \rangle\rangle \tag{6}$$

is a bipotential, i.e. it satisfies: for any $z \in X \times Y$,

(a) for any $z', z" \in X \times Y$ the functions $b(z, z', \cdot)$ and $b(z, \cdot, z")$ are convex (and lsc),

(b) for any $z', z" \in X \times Y$ we have the equivalences

$$z' \in \partial b(z, z', \cdot)(z") \iff z" \in \partial b(z, \cdot, z")(z') \iff I(z, z', z") = 0 \tag{7}$$

where "∂" denotes a subgradient, see section 3 for notations.

Bipotentials were introduced in [11] as a convex analysis notion which is well adapted for applications to non-associated constitutive laws. Bipotentials were used in soil mechanics, plasticity, damage or friction. For the theory of bipotentials see the review paper [5].

By concentrating our attention to the function b, instead of the information content I, we can build a host of examples. Indeed, for any lsc and convex function

$$\Phi : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$$

the associated function

$$b(z, z', z") \ = \ \Phi(z') + \Phi^*(z") \tag{8}$$

satisfies the conditions (a), (b) from Proposition 1.3 , where Φ^* is the polar, or Fenchel conjugate of Φ. Indeed, the mentioned condition (b) is just a reformulation of the Fenchel inequality [11].

Corollary 1.4 *For an information content function of the form*

$$I(z, z', z") \ = \ \Phi(z') + \Phi^*(z") - \langle\langle z', z" \rangle\rangle \tag{9}$$

the equation(3) is equivalent with the symplectic Brezis-Ekeland-Nayroles principle [3] definition 4.1, [4] definition 1.1:

$$\eta \in \partial \Phi (\dot{z}) \tag{10}$$

2   Examples

Viscosity, Rayleigh dissipation.   In particular, let's pick

$$\Phi(z') \ = \ \Phi(q', p') \ = \ \phi(q')$$

where $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is a convex, lsc function. A straightforward computation of Φ^* gives:

$$\Phi^*(z") \ = \ \Phi^*(q", p") \ = \ \sup \left\{ \langle\langle (q', p'), (q", p") \rangle\rangle - \phi(q') \mid q' \in X, p' \in Y \right\} =$$

3

$$= \sup \left\{ \langle q', p"\rangle + \langle q", p'\rangle - \phi(q') \mid q' \in X, p' \in Y \right\} =$$
$$= \chi_0(q") + \phi^*(p")$$

therefore the information content has the expression:

$$I(z, z', z") = \phi(q') + \phi^*(p") + \chi_0(q") - \langle q', p"\rangle$$

By the corollary 1.4 we obtain the equations:

$$\begin{cases} \dot{q} &=& \frac{\partial H}{\partial p}(q, p, t) \\ -\dot{p} &=& \frac{\partial H}{\partial q}(q, p, t) + \eta_p \\ \eta_p &\in& \partial \phi(\dot{q}) \end{cases} \tag{11}$$

This shows that ϕ is a Rayleigh dissipation potential.

Elasto-plasticity. Take a hamiltonian system with space state X and momentum space Y and supplement the state and the momentum with a new pair of state and momentum spaces:

$$(q, q_I) \in X \times X_I \,, \ (p, p_I) \in Y \times Y_I$$

Suppose further that the pair of spaces (X_I, Y_I) are in duality, so that we can define

$$\langle (q, q_I), (p, p_I) \rangle = \langle q, p \rangle + \langle q_I, p_I \rangle$$

which leads us to a duality product of $(X \times X_I) \times (Y \times Y_I)$ with itself:

$$\langle \langle (q', q_I', p', p_I'), (q", q_I", p", p_I") \rangle \rangle = \langle q', p"\rangle + \langle q", p'\rangle + \langle q_I', p_I"\rangle + \langle q_I", p_I'\rangle$$

In this setting, we take $N = (X \times X_I) \times (Y \times Y_I)$ and

$$\Phi\left(q', q_I', p', p_I'\right) = \phi\left(p_I'\right)$$

where $\phi : Y_I \to \mathbb{R} \cup \{+\infty\}$ is a convex lsc function.

There are two differences with respect to the viscosity example: there is a cartesian decomposition of the state space and momentum space, and the dissipation potential depends on (a component of) the momentum variable, while previously the dependence was on the state variable.

By a computation analoguous with the one from the previous example we obtain:

$$\Phi^*\left(q", q_I", p", p_I"\right) = \phi^*\left(q_I"\right) + \chi_0\left(p"\right) + \chi_0\left(p_I"\right) + \chi_0\left(q"\right)$$

Via the corollary 1.4 we obtain the equations:

$$\begin{cases} \dot{q} &=& \frac{\partial H}{\partial p}(q, p, q_I, p_I, t) \\ \dot{q}_I &=& \frac{\partial H}{\partial p_I}(q, p, q_I, p_I, t) + \eta_{q,I} \\ -\dot{p} &=& \frac{\partial H}{\partial q}(q, p, q_I, p_I, t) \\ -\dot{p}_I &=& \frac{\partial H}{\partial q_I}(q, p, q_I, p_I, t) \\ \eta_{q,I} &\in& \partial \phi(\dot{p}_I) \end{cases} \tag{12}$$

In particular, let's take $X = X_I$, $Y = Y_I$, Hilbert spaces, and a hamiltonian of the form:

$$H(q, p, q_I, p_I, t) = K(p) + E(q - q_I) - \langle q, f(t) \rangle$$

where $K(p)$ is the kinetic energy and $E(q - q_I)$ is the elastic energy. We denote the elastic force by

$$\sigma = \frac{\partial E}{\partial q}(q - q_I)$$

The 2nd and the 4th equations of the system (12) give:

$$\eta_{q,I} = \dot{q}_I \, , \ \dot{p}_I = \sigma$$

The system (12) reduces to the familiar equations:

$$
\begin{cases}
\dot{q} & = & \frac{\partial K}{\partial p}(p) \\
\dot{p} & = & f(t) - \sigma \\
\dot{q}_I & \in & \partial \phi(\sigma)
\end{cases}
\tag{13}
$$

In the next example we encounter a small extension of (9).

Damage.   We consider here a standard model of damage. In [2] we introduced a formulation of hamiltonian evolution with dissipation which is the ancestor of the one proposed in this article. Then we applied that formulation to a more involved damage model, where we used an Ambrosio-Tortorelli functional [1] in the hamiltonian of the model.

As in the elasto-plasticity example, we supplement the spaces X, Y with a new pair (q_I, p_I). Here $q_I = d \in \mathbb{R}$ is a damage variable and $p_I = r \in \mathbb{R}$ is a conjugated variable to the damage one.

Thefore we have a duality product over $N = (X \times \mathbb{R}) \times (Y \times \mathbb{R})$:

$$\langle\langle (q', d', p', r'), (q", d", p", r") \rangle\rangle = \langle q', p" \rangle + \langle q", p' \rangle + d'r" + d", r' \rangle$$

We choose a hamiltonian

$$H(q, p, d, r, t) = K(p) + k(r) + (1 - d)E(q) - \langle q, f(t) \rangle$$

where K is the kinetic energy, k is a term akin to the kinetic energy, but for the variable r and E is the elastic energy.

The difference from the previous examples is that the damage variable and the sign of $\dot{d}$ are constrained (i.e. damage cannot decrease):

$$d \in [0, 1] \quad , \quad \dot{d} \geq 0$$

We have to include these constraints in the expression of the information content:

$$I(q, d, p, r, \dot{q}, \dot{d}, \dot{p}, \dot{r}, \eta_q, \eta d, \eta_p, \eta_r) = \chi_{[0,1]}(d) +$$

$$+ \chi_{[0,+\infty)}(\dot{d}) + Y\dot{d} + \chi_{(-\infty,Y]}(\eta_r) +$$

$$+ \chi_0(\eta_p) + \chi_0(\eta_q) + \chi_0(\eta_d) - \langle\langle(\dot{q}, \dot{d}, \dot{p}, \dot{r}), (\eta_q, \eta_d, \eta_p, \eta_r)\rangle\rangle$$

The expression of the information content does not have the form (9), but instead it has the form

$$I(z, \dot{z}, \eta) = \quad constraints \; + \; \Phi(\dot{z}) + \Phi^*(\eta) - \langle\langle\dot{z}, \eta\rangle\rangle$$

with, in this case

$$\Phi(\dot{z}) = \phi(\dot{d}) = \chi_{[0,+\infty)}(\dot{d}) + Y\dot{d}$$

We obtain the equations:

$$\begin{cases} \dot{q} & = & \frac{\partial K}{\partial p}(p) \\ \dot{d} & = & \frac{\partial k}{\partial r}(r) \\ -\dot{p} & = & (1-d)\frac{\partial E}{\partial q}(q) - f(t) \\ -\dot{r} & = & -E(q) + \eta_r \end{cases} \tag{14}$$

and the inequations:

$$\begin{cases} d & \in & [0,1] \\ \dot{d} & \geq & 0 \\ \eta_r & \in & \partial\phi\left(\dot{d}\right) \end{cases} \tag{15}$$

In the next example we shall need a smoothness assumption. The alternative would be to renounce at the condition 1.2 (a) and to keep only the condition (b).

Unilateral contact. We take X, Y to be dual, finite dimensional Banach spaces and $M \subset X$ to be a closed, nonempty set of admissible states q. With the notations about Fréchet normals from [7] 1.4, we shall need that M is smooth enough, in the sense that the tangent cone at any $q \in M$ is the polar of the normal cone at x, denoted by $N(q \mid M)$.

There is no special form of the hamiltonian required. The information content has the form:

$$I(z, \dot{z}, \eta) = \chi_M(q) + \chi_{N(q|M)}(\eta_p) + \left(\chi_{N(q|M)}\right)^*(\dot{q}) - \langle\langle\dot{z}, \eta\rangle\rangle$$

The smoothness assumption on M is that

$$\left(\chi_{N(q|M)}\right)^* = \chi_{T(q|M)}$$

i.e. that the tangent cone is the polar of the normal cone. In general it is true that the normal cone is the polar of the tangent cone, but not the other way around. An alternative would be to not make any smoothness assumptions on M and instead to take an information content of the form

$$I(z, \dot{z}, \eta) = \chi_M(q) + \chi_{N(q|M)}(\eta_p) + \chi_{T(q|M)}(\dot{q}) - \langle\langle\dot{z}, \eta\rangle\rangle$$

Again, in general the tangent cone is not convex, therefore this form of information content contradicts the convexity condition 1.2 (a).

We obtain the equations:

$$\begin{cases} \dot{q} & = & \frac{\partial H}{\partial p}(q, p, t) \\ -\dot{p} & = & \frac{\partial H}{\partial q}(q, p, t) + \eta_p \end{cases} \tag{16}$$

6

and the unilateral contact conditions:

$$\left\{ \begin{array}{rcl} q & \in & M \\ \dot{q} & \in & T(q \mid M) \\ \eta_p & \in & N(q \mid M) \\ \langle \dot{q}, \eta_p \rangle & = & 0 \end{array} \right. \tag{17}$$

3 Notations and useful definitions

General notations.   The space X of states $q \in X$ and the space Y of momenta $p \in Y$ are real topological vector spaces in duality:

$$(q, p) \in X \times Y \mapsto \langle q, p \rangle \in \mathbb{R}$$

We suppose the usual: the duality is bilinear, continuous and for any linear and continuous functions $L : X \to \mathbb{R}$, $G : Y \to \mathbb{R}$ there exist $q \in X$, $p \in Y$ such that $L(\cdot) = \langle \cdot, p \rangle$ and $G(\cdot) = \langle q, \cdot \rangle$.

The space $N = X \times Y$ is in duality $\langle\langle \cdot, \cdot \rangle\rangle : N \times N \to \mathbb{R}$ with itself by:

$$\langle\langle (q_1, p_1), (q_2, p_2) \rangle\rangle = \langle q_1, p_2 \rangle + \langle q_2, p_1 \rangle$$

The space N is also symplectic, with the symplectic form defined by: for any $z_1 = (q_1, p_1)$, $z_2 = (q_2, p_2)$

$$\omega(z_1, z_2) = \langle\langle J z_1, z_2 \rangle\rangle = \langle q_1, p_2 \rangle - \langle q_2, p_1 \rangle$$

For any differentiable function $H : N \to \mathbb{R}$ the gradient of H at a point $z \in N$ is the element $DH(z) \in N$ with the property that

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(H \left(z + \varepsilon z' \right) - H(z) \right) = \langle\langle DH(z), z' \rangle\rangle$$

and the symplectic gradient of H is $XH(z) \in N$ is defined in a similar way by the equality

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(H \left(z + \varepsilon z' \right) - H(z) \right) = \omega(XH(z), z')$$

If we use the partial derivatives notation

$$DH(q, p) = \left(\frac{\partial}{\partial p} H(q, p), \frac{\partial}{\partial q} H(q, p) \right)$$

then

$$XH(q, p) = \left(\frac{\partial}{\partial p} H(q, p), -\frac{\partial}{\partial q} H(q, p) \right)$$

With the introduction of the linear conjugation

$$\bar{\cdot} : N \to N \ , \ \overline{(q, p)} =, (q, -p)$$

we have $XH(z) = \overline{DH(z)}$.

Convex analysis notations. These are the classical ones from Moreau [8]. We add $+\infty$ the field of reals $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. The addition operation is extended with $a + (+\infty) = +\infty$ for any $a \in \bar{\mathbb{R}}$. The multiplication with positive numbers is extended with: if $a > 0$ then $a(+\infty) = +\infty$.

For any function $\phi : X \to \bar{\mathbb{R}}$, it's domain is $dom\, \phi = \{x \in X : \phi(x) \in \mathbb{R}\}$.

The set of lower semicontinuous (lsc), convex functions defined on X, with non-empty domain is $\Gamma_0(X)$. The indicator function $\chi_A \in \Gamma_0(X)$ of a convex and closed set $A \subset X$ is

$$\chi_A(x) = \begin{cases} 0 & \text{if } x \in A \\ +\infty & \text{otherwise} \end{cases}$$

For any natural number $n \geq 1$, a non empty set $A \subset X \times Y$ is n-monotone monotone if for any collection $\{(x_k, y_k) \in A : k = 0, 1, ..., n\}$ we have the inequality:

$$\langle x_n - x_0, y_n \rangle + \sum_1^n \langle x_{k-1} - x_k, y_{k-1} \rangle \geq 0 \quad .$$

The set $A \subset X \times Y$ is maximally n-monotone if it is n-monotone and maximal with respect to the inclusion of sets.

The set A is cyclically monotone if it is n monotone for any natural number $n \geq 1$. It is cyclically maximal monotone if it cyclically monotone and maximal with respect to the inclusion of sets.

The subdifferential of a function $\phi : X \to \bar{\mathbb{R}}$ at a point $x \in X$ is the set:

$$\partial\phi(x) = \{u \in Y \mid \forall z \in X \ \langle z - x, u \rangle \leq \phi(z) - \phi(x)\} \quad .$$

The polar of a function $\phi : X \to \bar{\mathbb{R}}$ is $\phi^* : Y \to \bar{\mathbb{R}}$

$$\phi^*(y) = \sup\{\langle x, y \rangle - \phi(x) \mid x \in X\}$$

The polar is always convex and lsc.

Polars and subgradients are related by the Fenchel inequality. For any function $\phi : X \to \bar{\mathbb{R}}$ which is convex, lsc, we define

$$c(x, y) = \phi(x) + \phi^*(y) - \langle x, y \rangle$$

for any $x \in X$, $y \in Y$. The Fenchel inequality has two parts:

(a) $c(x, y) \geq 0$

(b) $c(x, y) = 0 \iff y \in \partial\phi(x) \iff x \in \partial\phi^*(y)$

References

[1] L. Ambrosio, V. Tortorelli, On the Approximation of Free Discontinuity Problems, Bollettino UMI 7, 6-B (1992), 105-123

[2] M. Buliga, Hamiltonian inclusions with convex dissipation with a view towards applications, *Mathematics and its Applications* **1**, 2 (2009), 228-251, arXiv:0810.1419.

[3] M. Buliga, G. de Saxcé, A symplectic Brezis-Ekeland-Nayroles principle, *Mathematics and Mechanics of Solids* **22**, 6, (2017)

[4] M. Buliga, A stochastic version and a Liouville theorem for hamiltonian inclusions with convex dissipation, arXiv:1807.10480

[5] M. Buliga, G. de Saxcé, C. Vallée, Bipotentials for non monotone multivalued operators: fundamental results and applications, Acta Applicandae Mathematicae, 110, 2(2010), 955-972

[6] M. Buliga, G. de Saxcé, C. Vallée, Blurred maximal cyclically monotone sets and bipotentials , Analysis and Applications 8 (2010), no. 4, 1-14

[7] A. Ya. Kruger, On Fréchet subdifferentials, Journal of Mathematical Sciences 16 (2003), no. 3, 3325-3358

[8] J.-J. Moreau, Fonctionnelles convexes (Rome: Istituto Poligrafico e zecca dello stato, 2003), p. 71, 97.

[9] R.T. Rockafellar, On the maximal monotonicity of subdifferential mappings, Pacific J. Math. 33, 1 (1970), 209-216

[10] A. Oueslati, A.D. Nguyen, G. de Saxcé (2017). A Symplectic Minimum Variational Principle for Dissipative Dynamical Systems. In: Nielsen F., Barbaresco F. (eds) Geometric Science of Information. GSI 2017. Lecture Notes in Computer Science, vol 10589. Springer, Cham

[11] G. de Saxcé, Une généralisation de linégalité de Fenchel et ses applications aux lois constitutives, C. R. Acad. Sci., Paris, Sér. II 314 (1992) 125-129