

## Variational aspects of analytical mechanics

**Gianluca Gorni**

Università di Udine  
Dipartimento di Matematica e Informatica  
via delle Scienze 208, 33100 Udine, Italy  
*E-mail address:* gianluca.gorni@uniud.it

**Gaetano Zampieri**

Università di Verona  
Dipartimento di Informatica  
strada Le Grazie 15, 37134 Verona, Italy  
*E-mail address:* gaetano.zampieri@univr.it

*Dedicated to Waldyr M. Oliva for his 80<sup>th</sup> birthday*

### 1. Introduction

In this paper we set out to revisit some basic variational aspects of Analytical Mechanics in one independent variable, in a way that would be most appealing to those who are being introduced to Mechanics while having some background in elementary Functional Analysis. A slightly longer version [7], written in Italian, is going to appear in a supplementary volume to the reedition of the classical 1923 Lectures in Rational Mechanics by Levi-Civita and Amaldi [9].

We are going to start with *Hamilton's principle*, that we will state for systems that admit a Lagrangian function, that is just assumed to have continuous first partial derivatives. This minimal regularity is by no means new, even though it is unusual in Mechanics textbooks. We think that it leads to a more general and elegant theory, that may be useful also outside of Mechanics.

We will then proceed to *Emmy Noether's theorem*, a result whose first formulation goes back to 1918 in the context of PDE, and is still today an active source of inspiration for physical theories. We will only treat the one independent variable, ODE case.

In our opinion the usual treatments of Noether's theorem overburden the statement, by requiring that the one-parameter family  $q_\sigma(t)$ , that extends the natural motion  $q_0(t)$ , to be of the form  $(\sigma, t, q, \dot{q})$ , or even that  $q_\sigma$  be a group with respect to the parameter  $\sigma$ , although the proof does not exploit such assumptions. Moreover, some authors spare the reader some work by curtailing the statement so much that neither the time change nor the term introduced by Bessel-Hagen (1921) are mentioned, and therefore cannot even deduce conservation of energy (as in the otherwise excellent book by Arnold [1]). Here we adopt the suggestions contained in an article by Lévy-Leblond [10], and thereby use Bessel-Hagen's term, that we call here BH-function, but not the time change. Strictly speaking our setting will not be the same as Noether's original one, but it is equivalent, as was observed by Boyer (1967). Our treatment is new in that it is divided into two distinct parts. The first part is a very easy, abstract and general proof that a certain quantity is a constant of motion. The second part is a strategy for finding first integrals, centered around what we call the total derivative condition (formula (18)), that unifies all the applications of the theorem that we will give. The interested reader may consult our specialized paper [8], that contains more historical details, and where we explain the interplay between the BH-function and time change, and how the constants of motion can well take an integral form.

In Section 5 we will show some *examples* to illustrate Noether's theorem and to highlight the novelties of our approach, in particular the center role of the BH-function and of the total derivative condition, the possibility of BH-functions of functional type, and the computation of functions that are constant along particular motions. Among the examples, our treatment of the Laplace-Runge-Lenz vector is different, and somewhat simpler, than the one usually found in the literature; it is basically taken from our paper [8] on Noether's theorem, but rewritten in a more familiar vector notation.

Next topic will be the local *least Hamiltonian action*, in Section 6. We will treat Lagrangians  $L(t, q, \dot{q})$  of general form with Legendre condition, define the second variation of the Hamiltonian action, and give a new detailed, elementary proof of the local theorem of least Hamiltonian action, using neighbourhoods in the  $C^1$  norm (as explained in Definition 2, page 265). As functional analysis tools we only require the most basic properties of the  $L^2$  norm. Our technique is inspired by the one used in Giovanni Gallavotti's book [4, n. 3.3], who however restricts to not fully general Lagrangians and refrains from introducing a topology in the space of motions. Nowhere else in the literature on this topic have we found an approach as elementary and direct as this one. Fox [3, sec. 2.4] makes his proof in the scalar case using a technique originated by Jacobi, with only an outline for the vector case [3, sec. 3.3], and he only deals with proving the local minimum along each direction. A version of the theorem can be proven for neighbourhoods

in the  $C^0$  norm, as done for example by Buttazzo, Giaquinta and Hildebrandt [2, theor. 1.33] and Troutman [13, sec. 9.8], but their techniques of fields of extremals are much more advanced.

In the last Section we deal with time-independent Lagrangians that are the difference between a homogeneous quadratic kinetic energy and a potential energy, what we call *Jacobi Lagrangians*, dwelling in particular on the *geodesics*. We will take the opportunity to work out two geodesic examples in the framework of Noether's theorem. Last, we will discuss the Jacobi metric, which, in a sense, reduces the Jacobi Lagrangians to the geodesics.

## 2. Notations

We will denote with  $\mathbb{R}^n$  the usual  $n$ -th dimensional Euclidean space, with  $x \cdot y$  the scalar product and with  $\|x\| = \sqrt{x \cdot x}$  the norm. In dimension 3 we will also use symbols like  $\vec{r}, \vec{v}$  for vectors and  $\times$  for the cross product. The symbol  $q$  will denote either a vector in  $\mathbb{R}^n$  or vector function of one variable, as appropriate to the context. The gradient of a scalar function  $q \mapsto f(q)$  of  $n$  variable will be either  $\partial_q f(q)$  or  $\nabla f(q)$ , and regarded as a vector in  $\mathbb{R}^n$ . To help the reader get used to our notations, here is Taylor's second-order formula for a scalar function  $f$  of two vector variables  $p, q \in \mathbb{R}^n$ :

$$\begin{aligned} f(p+h, q+k) &= f(p, q) + \partial_p f(p, q) \cdot h + \partial_q f(p, q) \cdot k + \\ &\quad + \frac{1}{2} h \cdot \partial_{p,p}^2 f(p, q) h + \frac{1}{2} k \cdot \partial_{q,q}^2 f(p, q) k + h \cdot \partial_{p,q}^2 f(p, q) k + \\ &\quad + o(\|p\|^2 + \|q\|^2). \end{aligned}$$

For a matrix  $A$ , the symbol  $\|A\|$  will be the operator norm, that is, the smallest real number such that  $\|Ax\| \leq \|A\| \cdot \|x\|$  for all vectors  $x$ .

When working with functionals on infinite-dimensional spaces, we will talk of Fréchet differential instead of gradient. Our notation will be  $F'(q)[h]$  to mean the differential of  $q \mapsto F(q)$  at the point  $q$  in the direction  $h$ , that is,  $F'(q)$  will be the linear operator that maps  $h$  to the directional derivative  $\lim_{\lambda \rightarrow 0} (F(q + \lambda h) - F(q))/\lambda$ . We will similarly treat the second differential  $F''(q)[h, k]$  as a bilinear form.

Given a compact interval  $I = [a, b] \subset \mathbb{R}$ , the set  $C^0(I, \mathbb{R}^n)$  of the continuous functions from  $I$  to  $\mathbb{R}^n$  is a Banach space with the norm

$$\|q\|_{C^0} := \sup_{t \in I} \|q(t)\|_{\mathbb{R}^n}. \quad (1)$$

Also  $C^1(I, \mathbb{R}^n)$  is a Banach space with the norm

$$\|q\|_{C^1} := \|q\|_{C^0} + \|\dot{q}\|_{C^0}, \quad (2)$$

and we will use its closed linear subspace  $C_0^1(I, \mathbb{R}^n)$  of the functions  $q \in C^1(I, \mathbb{R}^n)$  that vanish at the endpoints:  $q(a) = q(b) = 0$ .

### 3. Hamilton's principle

In this and in the next Sections we will try to make statements more readable by assuming a Lagrangian  $L$  defined unrestrictedly on  $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ . The results hold true, with only marginal retouching, also when the domain of  $L$  is a nonempty open subset of  $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ .

Given a  $C^1$  Lagrangian function  $L: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ , let us consider the following *action functional*

$$S_I(q) := \int_I L(t, q(t), \dot{q}(t)) dt, \quad (3)$$

defined on the functions  $q \in C^1(I, \mathbb{R}^n)$ . Sometimes it is called Hamiltonian action, to distinguish it from different action functionals that are found in the literature [4, concluding comments to Chapter 3]. This functional  $S_I$  is Fréchet differentiable and its differential is

$$S'_I(q)[h] = \int_I \left( \partial_q L(t, q(t), \dot{q}(t)) \cdot h(t) + \partial_{\dot{q}} L(t, q(t), \dot{q}(t)) \cdot \dot{h}(t) \right) dt \quad (4)$$

for all  $h \in C^1(I, \mathbb{R}^n)$ . If we take differentiability for granted, the formula for  $S'_I(q)[h]$  can be easily obtained from the directional derivatives

$$\frac{\partial}{\partial h} S_I(q) = \frac{d}{d\lambda} S_I(q + \lambda h) \Big|_{\lambda=0}.$$

by derivation under integral sign.

Here is what is meant by stationary action and Hamilton's principle:

**Definition 1** (Fixed-endpoints stationarity). *Let  $L: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  be a  $C^1$  Lagrangian function. We will say that the action  $S_I$  is fixed-endpoints stationary at  $q \in C^1(I, \mathbb{R}^n)$ , or that  $q$  is fixed-endpoints stationary for  $S_I$ , if  $S'_I(q)[h] = 0$  for all  $h \in C_0^1(I, \mathbb{R}^n)$ . In the sequel, "stationary" will always be an abbreviation for fixed-endpoint stationary.*

**Theorem 1** (Hamilton's Principle). *Let  $L: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  be a  $C^1$  Lagrangian function. The motion  $q \in C^1(I, \mathbb{R}^n)$  is stationary for the action  $S_I$  if and only if it satisfies Lagrange equation*

$$\frac{d}{dt} \partial_{\dot{q}} L(t, q(t), \dot{q}(t)) - \partial_q L(t, q(t), \dot{q}(t)) = 0 \quad \forall t \in I; \quad (5)$$

or, more precisely, if and only the mapping  $t \mapsto \partial_{\dot{q}} L(t, q(t), \dot{q}(t))$  is  $C^1$  and equation (5) holds. In this case we will also say that  $q$  is a natural motion (for the given Lagrangian).

In the Calculus of Variations, equation (5) is called Euler, or Euler-Lagrange equation. We omit the proof, because it is standard in textbooks, even though it usually assumes more smoothness than we do here. The minimal regularity version that we show in our parent work [7] is taken from Italian references.

A first consequence of Hamilton's principle is that stationarity is a local property with respect to time, in the sense that  $q$  is stationary for  $S_I$  if and only if all  $t_0 \in I$  have a connected neighbourhood  $J$  (in the relative topology of  $I$ ) such that the restriction of  $q$  to  $J$  is stationary for  $S_J$ .

When the  $C^1$  motion  $q: I = [a, b] \rightarrow \mathbb{R}^n$  is natural, the action differential can be rewritten in a simpler, integral-free form:

$$S'_I(q)[h] = \partial_{\dot{q}}L(b, q(b), \dot{q}(b)) \cdot h(b) - \partial_{\dot{q}}L(a, q(a), \dot{q}(a)) \cdot h(a) \quad (6)$$

for all  $h \in C^1(I, \mathbb{R}^n)$ , not necessarily vanishing at the endpoints. In fact, integrating formula (4) by parts we get

$$\begin{aligned} S'_I(q)[h] &= [\partial_{\dot{q}}L(t, q(t), \dot{q}(t)) \cdot h(t)]_{t=a}^{t=b} + \\ &+ \int_I \left( \partial_q L(t, q(t), \dot{q}(t)) - \frac{d}{dt} \partial_{\dot{q}}L(t, q(t), \dot{q}(t)) \right) \cdot h(t) dt \end{aligned} \quad (7)$$

and the last integral vanishes because of Lagrange equation.

Among the natural motions  $q(t)$ , there does not necessarily exist one that has preassigned values at the endpoints of  $I$ , and, if it exists, it may not be unique. Think for example at the harmonic oscillator, whose Lagrangian is  $L(t, q, \dot{q}) = (\dot{q}^2 - q^2)/2$ . Since all solutions to Lagrange equation  $\ddot{q} + q = 0$  are  $2\pi$ -periodic, we can assign values in 0 and in  $2\pi$  only if they coincide. Moreover, there are infinitely many solutions such that  $q(0) = 0 = q(\pi)$ .

Also, it is readily verified that two Lagrangians  $L, \tilde{L}$  yields the same Lagrange equations whenever they differ by the total time derivative of some smooth function  $f(t, q)$

$$\tilde{L}(t, q, \dot{q}) = L(t, q, \dot{q}) + \partial_t f(t, q) + \partial_q f(t, q) \cdot \dot{q}. \quad (8)$$

Replacing  $L$  with such a  $\tilde{L}$  in the calculations is called "*gauge*" transform.

#### 4. Noether's theorem

In this section we ask the reader to differentiate between the concepts of "constant of motion" and of "first integral", that are usually interchangeable. By *first integral* we mean a function of the finite set of variables  $(t, q, \dot{q})$  which is constant whenever it is evaluated along a natural motion  $q(t)$ . Instead, by *constant of motion* we mean a quantity which is constant in  $t$ , but which may depend on the motion  $q(t)$  not necessarily

through the current position and velocity, but possibly also through other values at other times. In the instance that we have in mind, the constant of motion will generally involve integrals of functions of  $t, q(t), \dot{q}(t)$ . Also, we will not require that the quantity be constant for all natural motions, but possibly for *single* selected motions only.

The version of Noether's theorem that we present here, in its full generality, will provide constant of motions, in the sense explained above. We think that the proof will be much simpler and more natural this way.

**Theorem 2** (Noether). *Let  $L: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  be a  $C^1$  Lagrangian,  $I_0$  be an interval, let  $(\sigma, t) \mapsto q_\sigma(t)$  be a  $C^2$  map from  $\mathbb{R} \times I_0$  to  $\mathbb{R}^n$ , and  $(\sigma, t) \mapsto G(\sigma, t)$  be a  $C^2$  function (that we will call Bessel-Hagen function, or simply BH-function) from  $\mathbb{R} \times I_0$  to  $\mathbb{R}$ , such that  $t \mapsto q_0(t)$  be a natural motion, and, finally, that the following "infinitesimal invariance up to BH-function" condition holds:*

$$\left. \frac{\partial}{\partial \sigma} \left( L(t, q_\sigma(t), \dot{q}_\sigma(t)) + \partial_t G(\sigma, t) \right) \right|_{\sigma=0} = 0 \quad \forall t \in I_0. \quad (9)$$

Then the function

$$t \mapsto \partial_{\dot{q}} L(t, q_0(t), \dot{q}_0(t)) \cdot \partial_\sigma q_\sigma(t)|_{\sigma=0} + \partial_\sigma G(\sigma, t)|_{\sigma=0} \quad (10)$$

is constant on  $I_0$ .

*Proof.* Let  $I = [a, b] \subseteq I_0$  and define

$$f(\sigma) := S_I(q_\sigma) + G(\sigma, b) - G(\sigma, a). \quad (11)$$

The mapping  $\sigma \mapsto q_\sigma$  has  $C^1$  Fréchet regularity if seen from  $\mathbb{R}$  with values in  $C^1(I, \mathbb{R}^n)$ . The mapping  $\sigma \mapsto S_I(q_\sigma)$  is the composition of two Fréchet differentiable mappings, so that it is itself Fréchet differentiable, and its derivative at  $\sigma = 0$  can be written using the special form (6) of the differential that holds for natural motions:

$$\begin{aligned} f'(\sigma) &= S'_I(q_\sigma)[\partial_\sigma q_\sigma] + \partial_\sigma G(\sigma, b) - \partial_\sigma G(\sigma, a), \\ f'(0) &= \partial_{\dot{q}} L(b, q_0(b), \dot{q}_0(b)) \cdot \partial_\sigma q_\sigma(b)|_{\sigma=0} + \partial_\sigma G(\sigma, b)|_{\sigma=0} - \\ &\quad - \partial_{\dot{q}} L(a, q_0(a), \dot{q}_0(a)) \cdot \partial_\sigma q_\sigma(a)|_{\sigma=0} - \partial_\sigma G(\sigma, a)|_{\sigma=0}. \end{aligned} \quad (12)$$

On the other hand, bringing the derivative under the integral sign, we have that

$$\frac{\partial}{\partial \sigma} S_I(q_\sigma) = \int_I \frac{\partial}{\partial \sigma} L(t, q_\sigma(t), \dot{q}_\sigma(t)) dt.$$

As for the BH-function, we can write

$$\frac{\partial}{\partial \sigma} (G(\sigma, b) - G(\sigma, a)) = \frac{\partial}{\partial \sigma} \int_a^b \partial_t G(\sigma, t) dt = \int_a^b \partial_\sigma \partial_t G(\sigma, t) dt.$$

Hence, by using the assumption (9),

$$\begin{aligned} f'(0) &= \frac{\partial}{\partial \sigma} (S_I(q_\sigma) + G(\sigma, b) - G(\sigma, a)) \Big|_{\sigma=0} = \\ &= \int_I \frac{\partial}{\partial \sigma} (L(t, q_\sigma(t), \dot{q}_\sigma(t)) + \partial_t G(\sigma, t)) dt \Big|_{\sigma=0} = \int_I 0 dt = 0. \end{aligned}$$

By combining this result with (12) we obtain that the function (10) has the same value at  $t = a$  and  $t = b$ . Since  $a, b$  are arbitrary in  $I_0$ , the function (10) is actually constant on  $I_0$ .  $\square$

We have made the statement simple by assuming that the function  $(\sigma, t) \mapsto q_\sigma(t)$  be defined on  $\mathbb{R} \times I_0$ . Alternatively, we may require it to be defined on an open subset of  $\mathbb{R} \times \mathbb{R}$  that contains the vertical segment  $\{0\} \times I_0$ .

Of course, the infinitesimal invariance condition (9) is surely implied by the stricter “finite invariance” up to a BH-function:

$$L(t, q_\sigma(t), \dot{q}_\sigma(t)) + \partial_t G(\sigma, t) \text{ does not depend on } \sigma. \quad (13)$$

In integral terms, what is invariant is the function  $f$  of formula (11), which is defined as the sum of the action and the increment of the BH-function. In the Calculus of Variations textbooks the analogous notion, in the context of more independent variables, is called invariance up to a divergence.

We have chosen the letter  $G$  for the BH-function because it is connected with the concept of “gauge-variance” described by Levy-Leblond [10].

In our statement of Noether’s theorem we have given equal prominence to the family  $q_\sigma(t)$  and to the term  $G$ . Actually, what is important is  $q_\sigma$ , because a  $G$  that makes the theorem work is easily found as a consequence, for instance

$$G(\sigma, t) := - \int_{t_0}^t L(s, q_\sigma(s), \dot{q}_\sigma(s)) ds, \quad (14)$$

for an arbitrary constant  $t_0 \in I_0$ , or, alternatively, the following, which is linear with respect to  $\sigma$ :

$$G(\sigma, t) := -\sigma \int_{t_0}^t \frac{\partial}{\partial \sigma} (L(s, q_\sigma(s), \dot{q}_\sigma(s))) \Big|_{\sigma=0} ds; \quad (15)$$

other choices are conceivable, provided that they do not change what is essential here, that is, the partial derivative  $\partial_{\sigma,t}^2 G$  for  $\sigma = 0$ . The non-linear choice (14) yields a constant function  $f$ , i.e., what we called finite invariance.

These  $G$  functions, if replaced in formula (10) all give the following constant of motion:

$$t \mapsto \partial_{\dot{q}} L(t, q_0(t), \dot{q}_0(t)) \cdot \partial_{\sigma} q_{\sigma}(t)|_{\sigma=0} - \int_{t_0}^t \frac{\partial}{\partial \sigma} \left( L(s, q_{\sigma}(s), \dot{q}_{\sigma}(s)) \right) \Big|_{\sigma=0} ds, \quad (16)$$

which involves an integral along the motion, and hence it is not a first integral in the usual sense. If this were not enough, also the term  $\partial_{\sigma} q_{\sigma}(t)|_{\sigma=0}$  does not clearly appear to be a function of  $(t, q, \dot{q})$ . In stating our version of Noether's theorem, we may have even chosen to bypass any mention of  $G$  and exhibit directly the constant of motion (16). We prefer to give  $G$  a role of its own, however, also because in many examples it lends itself to useful mechanical interpretations.

Noether's theorem, as presented above, would be of modest interest, if it did not happen that there exist rare and precious choices of the Lagrangian  $L$  and of the family  $q_{\sigma}$  for which  $q_0$  can be a generic natural motion, and moreover we can write

$$\partial_{\sigma} q_{\sigma}(t)|_{\sigma=0} = \varphi(t, q_0(t), \dot{q}_0(t)), \quad (17)$$

$$\frac{\partial}{\partial \sigma} \left( L(t, q_{\sigma}(t), \dot{q}_{\sigma}(t)) \right) \Big|_{\sigma=0} = \frac{d}{dt} \left( \psi(t, q_0(t), \dot{q}_0(t)) \right), \quad (18)$$

for suitable functions  $\varphi, \psi$  of the variables  $(t, q, \dot{q})$ , if necessary by exploiting the hypothesis that  $q_0(t)$  solves Lagrange equation. Formula 18 will be called *total derivative condition*. In such cases we can take as  $G$  the function

$$G(\sigma, t) := -\sigma \cdot \psi(t, q_0(t), \dot{q}_0(t)). \quad (19)$$

Then the constant of motion simplifies to a proper *first integral* in the usual sense: a function of finite number of variables  $(t, q, \dot{q})$

$$\partial_{\dot{q}} L(t, q, \dot{q}) \cdot \varphi(t, q, \dot{q}) - \psi(t, q, \dot{q}), \quad (20)$$

which will be constant whenever evaluated along a natural motion  $q_0(t)$ .

All applications that we are going to see of Noether's theorem will follow the same unified steps: write the family  $q_{\sigma}$ , find a  $\psi$  that satisfies the total derivative condition (18), write the corresponding  $G$  from (19), and finally see what the first integral (20) looks like in that special situation.

## 5. Applications of Noether's theorem to first integrals

In this Section, and later concerning geodesics, we review a choice of the most notable systems where Noether's theorem can be applied, each detailed with its Lagrangian  $L$ , the family  $q_{\sigma}$ , the BH-function  $G$  and the first integral.

5.1. **Energy.** We start with the important special case when  $L$  is independent of time, that is,  $L(t, q, \dot{q}) = \mathcal{L}(q, \dot{q})$ . Then any regular vector function  $q(t)$  (natural motion or not) can be embedded into a time-shift family

$$q_\sigma(t) = q(t + \sigma), \tag{21}$$

defined for the couples  $(\sigma, t) \in \mathbb{R} \times \mathbb{R}$  such that  $t + \sigma \in I_0$ . We can compute the two relations

$$\partial_\sigma q_\sigma(t)|_{\sigma=0} = \dot{q}(t), \tag{22}$$

$$\left. \frac{\partial}{\partial \sigma} L(t, q_\sigma(t), \dot{q}_\sigma(t)) \right|_{\sigma=0} = \frac{d}{dt} \mathcal{L}(q(t), \dot{q}(t)). \tag{23}$$

that fit into the pattern of equations (17) and (18). If we set

$$G(\sigma, t) = -\sigma \mathcal{L}(q(t), \dot{q}(t)) \tag{24}$$

we see that the infinitesimal invariance (9) holds. Then the natural motions are granted by Noether's theorem the *conservation of energy*, that is, the first integral

$$\partial_{\dot{q}} \mathcal{L}(q, \dot{q}) \cdot \dot{q} - \mathcal{L}(q, \dot{q}). \tag{25}$$

This way of deriving the conservation of energy is not the most common in the literature, but it was promoted by Lévy-Leblond [10], and we share his view.

We can obtain the same first integral with a BH-function that gives finite invariance, as with the one of formula (15) or with the following variant

$$\tilde{G}(\sigma, t) = \int_t^{t+\sigma} \mathcal{L}(q(\xi), \dot{q}(\xi)) d\xi \tag{26}$$

This  $G$  is nonlinear in  $\sigma$  and is not a function of  $(\sigma, t, q(t), \dot{q}(t))$ . It leads to finite invariance, because the quantity

$$\begin{aligned} L(t, q_\sigma(t), \dot{q}_\sigma(t)) + \partial_t \tilde{G}(\sigma, t) &= \\ &= \mathcal{L}(q(t + \sigma), \dot{q}(t + \sigma)) - \left( \mathcal{L}(q(t + \sigma), \dot{q}(t + \sigma)) - \mathcal{L}(q(t), \dot{q}(t)) \right) = \\ &= \mathcal{L}(q(t), \dot{q}(t)) \end{aligned}$$

is independent of  $\sigma$ . On the other hand, with such an integral BH-function we can reasonably expect a mere constant of motion, and not generally a bona fide first integral.

5.2. **Momentum.** Suppose that  $L(t, q, \dot{q})$  be invariant in the direction of the vector  $u \in \mathbb{R}^n$ :

$$L(t, q + \sigma u, \dot{q}) \equiv L(t, q, \dot{q}) \quad \forall t, \sigma \in \mathbb{R}, \quad q, \dot{q} \in \mathbb{R}^n. \tag{27}$$

Given any smooth  $q(t)$ , consider the space-shift family

$$q_\sigma(t) := q(t) + \sigma u, \quad (28)$$

It is clear that the finite invariance condition (13) holds simply with  $G \equiv 0$ . Noether's theorem then ensures that the component of the momentum in the  $u$  direction

$$\partial_{\dot{q}}L(t, q(t), \dot{q}(t)) \cdot u \quad (29)$$

is a constant of motion. It is also evidently a first integral.

**5.3. Free fall.** Let us see how the simple problem of the free fall can fit into Noether's theorem and gauge transforms. Consider the free fall of a point weight  $q = (x, y, z)$  in  $\mathbb{R}^3$ , with the positive  $z$  axis pointing downward, and with no air resistance. The Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2}\|\dot{q}\|^2 + (0, 0, g) \cdot q = \frac{1}{2}\|\dot{q}\|^2 + gz. \quad (30)$$

The Lagrange equation is  $\ddot{q} = (0, 0, g)$ .

We observe first that the Lagrangian is time-independent, so that the energy

$$\partial_{\dot{q}}L(q, \dot{q}) \cdot \dot{q} - L(q, \dot{q}) = \dot{q} \cdot \dot{q} - L(q, \dot{q}) = \frac{1}{2}\|\dot{q}\|^2 - gz. \quad (31)$$

will be preserved along the free falls. Another straightforward invariance is with respect to the horizontal translations: taken a horizontal vector  $u = (u_1, u_2, 0)$  we can define the family (28) and obtain the conservation of

$$\partial_{\dot{q}}L(q, \dot{q}) \cdot u = \dot{q} \cdot u = u_1\dot{x} + u_2\dot{y}.$$

Since this happens for any  $u_1, u_2 \in \mathbb{R}$ , we deduce that the horizontal components  $\dot{x}, \dot{y}$  of the velocity are constant.

A third invariance is less obvious. Consider the family of the vertical translations

$$q_\sigma(t) := q(t) + (0, 0, \sigma). \quad (32)$$

The Lagrangian itself is not invariant, but still we can successfully try the total derivative condition (18)

$$\frac{\partial}{\partial \sigma}L(q_\sigma(t), \dot{q}_\sigma(t)) = g = \partial_{\sigma, t}^2(\sigma gt),$$

which suggests the term

$$G(\sigma, t) := -\sigma gt.$$

The consequence is that along any free fall the following time dependent-first quantity is conserved:

$$\partial_{\dot{q}}L(q(t), \dot{q}(t)) \cdot \partial_\sigma q_\sigma(t)|_{\sigma=0} + \partial_\sigma G(\sigma, t)|_{\sigma=0} = \dot{z} - gt, \quad (33)$$

which we can combine with  $\dot{x}, \dot{y}$ , to obtain a vector first integral

$$\dot{q} - (0, 0, gt). \tag{34}$$

The free fall can be described also with a different Lagrangian:

$$\tilde{L}(t, q, \dot{q}) = L(q, \dot{q}) + \partial_t(-gtz) + \partial_q(-gtz) \cdot \dot{q} = \frac{1}{2} \|\dot{q}\|^2 - gtz. \tag{35}$$

As we remarked at the end of Section 3, this is a gauge transform, that does not change the Lagrange equation. The Lagrangian  $\tilde{L}$  has one advantage: it is independent of  $q$ . This means that there is finite invariance under the space-shift family  $q_\sigma(t) := q(t) + \sigma u$  for any direction  $u \in \mathbb{R}^3$ , with  $G \equiv 0$ . The resulting first integral is  $\partial_{\dot{q}} \tilde{L} \cdot u = (\dot{x}, \dot{y}, \dot{z} - gt) \cdot u$ , which leads again to (34).

The new Lagrangian  $\tilde{L}$  has a drawback, of course: it is time-dependent, so that the conservation of energy is not automatic any more. Let us take the same time-shift family  $q_\sigma(t) = q(t + \sigma)$  as in (21), and rework the calculation (23) for  $\tilde{L}$ :

$$\begin{aligned} \frac{\partial}{\partial \sigma} \tilde{L}(t, q(t + \sigma), \dot{q}(t + \sigma)) \Big|_{\sigma=0} &= \frac{d}{dt} \tilde{L}(t, q(t), \dot{q}(t)) - (\partial_t \tilde{L})(t, q(t), \dot{q}(t)) = \\ &= \frac{d}{dt} \tilde{L}(t, q(t), \dot{q}(t)) + gz(t) = \frac{d}{dt} \left( \tilde{L}(t, q(t), \dot{q}(t)) + gz(t) \right). \end{aligned}$$

This suggests the new term

$$\tilde{G}(\sigma, t) := -\sigma \left( \tilde{L}(t, q(t), \dot{q}(t)) + gz(t) \right), \tag{36}$$

which gives infinitesimal invariance

$$\frac{\partial}{\partial \sigma} \left( \tilde{L}(t, q_\sigma(t), \dot{q}_\sigma(t)) + \partial_t \tilde{G}(\sigma, t) \right) \Big|_{\sigma=0} \equiv 0.$$

With  $\tilde{G}$  we recover the conservation of energy:

$$\begin{aligned} \partial_{\dot{q}} \tilde{L}(t, q(t), \dot{q}(t)) \cdot \partial_\sigma q_\sigma(t) \Big|_{\sigma=0} + \partial_\sigma \tilde{G}(\sigma, t) \Big|_{\sigma=0} &= \\ &= \dot{q} \cdot \dot{q} - gtz - \tilde{L} - gz = \frac{1}{2} \|\dot{q}\|^2 - gz. \end{aligned}$$

**5.4. Angular Momentum.** Let us consider the Lagrangian of a point mass in  $\mathbb{R}^3$  that is subject to a central field force, possibly time-dependent:

$$L(t, \vec{r}, \vec{v}) := \frac{1}{2} m \|\vec{v}\|^2 - U(t, \|\vec{r}\|). \tag{37}$$

We will need some preliminaries on space rotations. Given a vector  $\vec{u} = (u_1, u_2, u_3)$ , let  $A_{\vec{u}}$  be the skew-symmetric matrix defined by the action

$A_{\vec{u}} \vec{r} = \vec{u} \times \vec{r}$ , where  $\times$  is the cross product. In components

$$A_{\vec{u}} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}. \quad (38)$$

Then the matrix exponential  $e^{\sigma A_{\vec{u}}}$  is a rotation matrix: it is orthogonal because

$$(e^{\sigma A_{\vec{u}}})^T = e^{\sigma A_{\vec{u}}^T} = e^{-\sigma A_{\vec{u}}} = (e^{\sigma A_{\vec{u}}})^{-1};$$

and it has positive determinant, since  $(\det e^{\sigma A_{\vec{u}}})^2 \equiv 1$  and at  $\sigma = 0$  the continuous function  $\sigma \mapsto \det e^{\sigma A_{\vec{u}}}$  has value 1. For example, if we choose  $\vec{u} = \vec{e}_3 = (0, 0, 1)$  we get

$$\begin{aligned} e^{\sigma A_{\vec{u}}} &= \sum_{k=0}^{\infty} \frac{\sigma^k A_{\vec{u}}^k}{k!} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \sigma \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \frac{\sigma^2}{2!} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \\ &\quad - \frac{\sigma^3}{3!} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{\sigma^4}{4!} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \dots = \begin{pmatrix} \cos \sigma & -\sin \sigma & 0 \\ \sin \sigma & \cos \sigma & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

which is a rotation of angle  $\sigma$  around the axis  $\vec{e}_3$ .

Given now a generic smooth  $\vec{r}(t)$  and a direction vector  $\vec{u}$ , define the space rotation family

$$\vec{r}_{\sigma}(t) = e^{\sigma A_{\vec{u}}} \vec{r}(t). \quad (39)$$

There is finite invariance with null  $G$ :

$$\frac{\partial}{\partial \sigma} L(t, \vec{r}_{\sigma}(t), \dot{\vec{r}}_{\sigma}(t)) = 0$$

since

$$\|\vec{r}_{\sigma}(t)\| = \|\vec{r}(t)\|, \quad \|\dot{\vec{r}}_{\sigma}(t)\| = \|\dot{\vec{r}}(t)\|.$$

We can apply Noether's theorem and obtain the following first integral along all solutions to Lagrange equation:

$$\begin{aligned} m \dot{\vec{r}}(t) \cdot \frac{\partial}{\partial \sigma} e^{\sigma A_{\vec{u}}} \vec{r}(t) \Big|_{\sigma=0} &= m \dot{\vec{r}}(t) \cdot A_{\vec{u}} e^{\sigma A_{\vec{u}}} \vec{r}(t) \Big|_{\sigma=0} = \\ &= m \dot{\vec{r}}(t) \cdot A_{\vec{u}} \vec{r}(t) = m \dot{\vec{r}}(t) \cdot \vec{u} \times \vec{r}(t) = \vec{u} \cdot \vec{r}(t) \times m \dot{\vec{r}}(t). \end{aligned}$$

Even more, since  $\vec{u}$  is arbitrary, we deduce that the vector-valued *angular momentum*

$$\vec{r} \times m \dot{\vec{v}}. \quad (40)$$

is a first integral.

5.5. **Kepler’s problem.** The Lagrangian of Kepler’s problem is

$$L(t, \vec{r}, \vec{v}) = \frac{1}{2}m\|\vec{v}\|^2 + \frac{k}{r}, \quad \vec{r} \in \mathbb{R}^3 \setminus \{0\}, \tag{41}$$

where  $r = \|\vec{r}\|$ ,  $m, k > 0$ , with its associated Lagrange equation

$$m\ddot{\vec{r}} = -\frac{k}{r^3} \vec{r}. \tag{42}$$

Since  $L$  is obviously both time-independent and central, we can apply the results above and get two first integrals: energy and angular momentum

$$\frac{1}{2}m\|\vec{v}\|^2 - \frac{k}{r}, \quad \vec{r} \times m\vec{v}. \tag{43}$$

Kepler’s problem exhibits one further vector first integral, variously known as Runge-Lenz or Laplace-Runge-Lenz vector

$$m\vec{v} \times (\vec{r} \times m\vec{v}) - \frac{mk}{r} \vec{r} \tag{44}$$

that we will deduce below, and that makes Kepler’s system superintegrable. Goldstein [6] reports that this vector was discovered independently by various authors, starting from Jakob Hermann and Johann Bernoulli (1710) and then Laplace (1799), and that it was crucial in deriving the emission spectrum of the hydrogen atom. For curved noncircular Kepler orbits, it can be easily shown that the Laplace-Runge-Lenz vector points in the direction of the point of minimum  $r$  along the orbit, what is usually called the perihelion. For more about the mechanical meaning the reader can refer for example to the textbook by Goldstein, Poole and Safko [5, sec. 3.9].

5.5.1. *Laplace-Runge-Lenz vector.* We will prove the constance of the vector (44) using Noether’s theorem. This will require a nontrivial choice for the family  $\vec{r}_\sigma(t)$ . Let us start with an arbitrary smooth  $\vec{r}(t) = \vec{r}_0(t)$ , that does not go through the origin, and a vector  $\vec{u} \in \mathbb{R}^3$ . We want to define  $\vec{r}_\sigma(t)$  so that

$$\vec{u} \cdot m\vec{v} \times (\vec{r} \times m\vec{v}) = \partial_{\vec{v}}L(t, \vec{r}_0(t), \dot{\vec{r}}_0(t)) \cdot \partial_\sigma \vec{r}_\sigma(t)|_{\sigma=0}. \tag{45}$$

A suitable choice is

$$\vec{r}_\sigma(t) = \vec{r}(t) + m(\vec{r}(t) \times \vec{r}(t + \sigma)) \times \vec{u}, \tag{46}$$

as can be seen by noticing that

$$\partial_\sigma \vec{r}_\sigma(t)|_{\sigma=0} = (\vec{r}(t) \times m\dot{\vec{r}}(t)) \times \vec{u}. \tag{47}$$

Let us compute the left-hand side of equation (18):

$$\frac{\partial}{\partial \sigma} L(t, \vec{r}_\sigma(t), \dot{\vec{r}}_\sigma(t)) \Big|_{\sigma=0} = \left( m\dot{\vec{r}}_\sigma(t) \cdot \partial_\sigma \dot{\vec{r}}_\sigma(t) - \frac{k}{\|\vec{r}_\sigma(t)\|^3} \vec{r}_\sigma(t) \cdot \partial_\sigma \vec{r}_\sigma(t) \right) \Big|_{\sigma=0}.$$

If we now add the assumption that  $\vec{r}(t) = \vec{r}_0(t)$  is a Kepler motion, that is, a solution to Lagrange equation (42), we can check that

$$\partial_\sigma \dot{\vec{r}}_\sigma(t) \Big|_{\sigma=0} = 0.$$

Replacing this into (47) we can further compute and find a total derivative:

$$\begin{aligned} \frac{\partial}{\partial \sigma} L(t, \vec{r}_\sigma(t), \dot{\vec{r}}_\sigma(t)) \Big|_{\sigma=0} &= -\frac{k}{r^3} \vec{r} \cdot (\vec{r} \times m\dot{\vec{r}}) \times \vec{u} = -\frac{k}{r^3} \vec{r} \times (\vec{r} \times m\dot{\vec{r}}) \cdot \vec{u} = \\ &= -\frac{k}{r^3} \vec{u} \cdot (\vec{r}(\vec{r} \cdot m\dot{\vec{r}}) - m\dot{\vec{r}}r^2) = \frac{d}{dt} \left( \frac{mk}{r(t)} \vec{r}(t) \cdot \vec{u} \right). \end{aligned}$$

There is then infinitesimal invariance with the following BH-function

$$G(\sigma, t) = -\sigma \frac{mk}{r(t)} \vec{r}(t) \cdot \vec{u}. \quad (48)$$

The associated conserved quantity is found to be the scalar product of  $\vec{u}$  with the Laplace-Runge-Lenz vector.

Notice that here we needed to use Lagrange equation explicitly to prove the infinitesimal invariance. In the previous examples Lagrange equation were only involved implicitly when finally invoking Noether's theorem, which states that the first integral is indeed constant along all natural motions.

5.5.2. *Constants along single motions.* Let us show how our formulation of Noether's theorem 2 can yield a function that is constant along a single natural motion, but not along others. Fix a radius  $R > 0$  and consider the following family of uniform circular motions around the origin in  $\mathbb{R}^3$  with a time-shift:

$$\vec{r}_\sigma(t) := e^{\omega\sigma} R(\cos(\omega(\sigma+t)), \sin(\omega(\sigma+t)), 0), \quad \text{where } \omega = \sqrt{k/(mR^3)}. \quad (49)$$

Within this family, the only Kepler motion is the one for  $\sigma = 0$ . However, there is infinitesimal invariance of the Kepler Lagrangian with  $G \equiv 0$ :

$$\frac{\partial}{\partial \sigma} L(t, \vec{r}_\sigma(t), \dot{\vec{r}}_\sigma(t)) \Big|_{\sigma=0} = \frac{\partial}{\partial \sigma} \left( \frac{mR^2\omega^2}{2} e^{2\omega\sigma} + \frac{k}{R e^{\omega\sigma}} \right) \Big|_{\sigma=0} = 0.$$

Applying Noether's theorem we get that twice the kinetic energy  $t \mapsto m \cdot \|\dot{\vec{r}}_0(t)\|^2$  is constant along that single natural motion  $r_0(t)$ . Of course, it is not constant along most Kepler motions.

**5.6. A separated degree of freedom.** In a recent work [14] the author introduced the following Lagrangian, with its associated Lagrange equations:

$$L(x, y, \dot{x}, \dot{y}) = \dot{x}\dot{y} - g(x)y, \quad (50)$$

$$\ddot{x} = -g(x), \quad \ddot{y} = -g'(x)y, \quad (51)$$

where  $g \in C^1$  in a neighbourhood of 0 in  $\mathbb{R}$ ,  $g(0) = 0$ ,  $g'(0) > 0$ . The dynamic is interesting because the origin is weakly unstable for almost all choices of the function  $g$ , whilst for some exceptional  $g$  there is superintegrability with isochrony. For all  $g$  the system has the following two first integrals:

$$H(x, y, \dot{x}, \dot{y}) = \dot{y}\dot{x} + g(x)y, \quad K(x, \dot{x}) = \frac{\dot{x}^2}{2} + V(x), \quad (52)$$

where

$$V(x) := \int_0^x g(s) ds.$$

Let us see how  $H$  and  $K$  can be obtained from Noether's theorem. What is obvious at once is that the Lagrangian is time-independent, so that there is invariance under the time-shift family  $q_\sigma(t) = (x(t + \sigma), y(t + \sigma))$ , leading to the conservation of energy, which is  $H$ .

As for  $K$ , we can use the family

$$q_\sigma(t) = (x(t), y(t) + \sigma\dot{x}(t)).$$

The Lagrangian along this family is

$$L(q_\sigma, \dot{q}_\sigma) = (\dot{y} - \sigma g(x))\dot{x} - (y + \sigma\dot{x})g(x).$$

Let us try the total derivative condition (18), using explicitly the Lagrange equations (51)

$$\frac{\partial}{\partial \sigma} L(q_\sigma(t), \dot{q}_\sigma(t)) \Big|_{\sigma=0} = -2g(x)\dot{x} = -2\frac{d}{dt}V(x(t)).$$

Indeed there is infinitesimal invariance up to the BH-function

$$G(\sigma, t) = 2\sigma V(x(t)), \quad (53)$$

whence the conservation of  $K$ :

$$\begin{aligned} \partial_{\dot{q}} L(q(t), \dot{q}(t)) \cdot \partial_\sigma q_\sigma(t) \Big|_{\sigma=0} + \partial_\sigma G(\sigma, t) \Big|_{\sigma=0} &= \\ &= (\dot{y}, \dot{x}) \cdot (0, \dot{x}) + 2V(x) = 2K(x(t), \dot{x}(t)). \end{aligned}$$

In another work of ours [8] we show how the total derivative condition can be effectively used to compute families of functions  $g$  for which there exists a third independent first integral, i.e., there is superintegrability.

## 6. Local theorem of least Hamiltonian action

From this Section onward we will confine our attention to the Lagrangians  $L(t, q, \dot{q})$  for which the Hessian matrix with respect to  $\dot{q}$  is positive definite:

$$\partial_{\dot{q}, \dot{q}}^2 L(t, q, \dot{q}) > 0, \quad (54)$$

which is called Legendre condition. The most classical case of this situation is when  $L = T - V$ , where the kinetic energy  $T$  is a positive definite quadratic form with respect to  $\dot{q}$ , and the generalized potential energy  $V$  has an affine dependency on  $\dot{q}$ .

If  $L$  is of class  $C^2$  with respect to all variables and the strict Legendre condition holds, in particular the Hessian matrix  $\partial_{\dot{q}, \dot{q}}^2 L$  is invertible, and we are guaranteed that every natural motion is not just of class  $C^1$  (as was minimally assumed in Section 3) but actually of class  $C^2$  with respect to time. In fact, the Lagrange equation gives

$$\partial_{\dot{q}} L(t, q(t), \dot{q}(t)) = \partial_{\dot{q}} L(a, q(a), \dot{q}(a)) + \int_a^t \partial_q L(\xi, q(\xi), \dot{q}(\xi)) d\xi. \quad (55)$$

Since the Hessian is invertible, by the implicit function theorem we can locally solve the equation

$$\partial_{\dot{q}} L(t, q, \dot{q}) = p \quad (56)$$

and express  $\dot{q}$  as a  $C^1$  function of  $(t, q, p)$ . If we replace  $p$  with the right-hand side of (55), and  $q, \dot{q}$  with  $q(t), \dot{q}(t)$ , we see that the function  $t \mapsto \dot{q}(t)$  is of class  $C^1$ . We can then rewrite the Lagrange equations in a regular form that is linear with respect to  $\ddot{q}$ :

$$\partial_{\dot{q}, \dot{q}}^2 L \ddot{q} + \partial_{\dot{q}, q}^2 L \dot{q} + \partial_t \partial_{\dot{q}} L - \partial_q L = 0, \quad (57)$$

where

$$(\partial_{\dot{q}, q}^2 L \dot{q})_i = \sum_j \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j} \dot{q}_j.$$

If we multiply equation (57) by the inverse of the Hessian matrix  $\partial_{\dot{q}, \dot{q}}^2 L$  we further reduce the Lagrange equation into normal form.

Some special cases of what we have called Hamilton's principle are referred to as "principle of least action" by some authors, as for example Arnold [1]. We avoided using such expression, because it is meaningless for such general Lagrangians as the ones in Section 3, and should be replaced by principle of stationary action. The main result of this Section is to prove that, when Legendre conditions holds, then we are partially justified in talking of least action, in the sense that the solutions of the Lagrange equation actually minimize the action (among small enough fixed-endpoints variations), provided that the integration interval is small enough too. On

arbitrary integration intervals the minimization is not guaranteed, as we are going to see soon with a simple example.

Let us be more precise about what we mean by minimizing the action:

**Definition 2.** We will say that the Hamiltonian action  $S_I$  has a fixed-endpoints local minimum for the  $C^1$  topology at the motion  $q \in C^1(I, \mathbb{R}^n)$  if there exists  $\delta > 0$  such that  $S_I(q) \leq S_I(q + h)$  for all  $h \in C_0^1$  such that  $\|h\|_{C^1} < \delta$ . Later we will abbreviate and simply say that  $q$  minimizes the action. The minimum will be called strict if  $S_I(q) < S_I(q + h)$  when  $0 < \|h\|_{C^1} < \delta$ .

In the Calculus of Variations the minima that we have just defined are often called *weak minima*, and contrasted with *strong minima*, for which the definition is the same, except that the condition  $\|h\|_{C^1} < \delta$  is replaced with  $\|h\|_{C^0} < \delta$ . The adjectives “weak” and “strong” are due to the fact that the neighbourhoods in the  $C^1$  topology are smaller than in the  $C^0$  topology. We will not mention or use strong minima again, but the reader that has already met them should take some care not to confuse strong minima with we call here strict minima.

It is obvious that minimum implies stationary. Contrary to what Legendre (1786) thought, as mentioned for example by Troutman [13, sec. 9.8, p. 320], the reverse is false even if we assume that  $L \in C^2$  and the strict Legendre condition. As a simple example consider the harmonic oscillator

$$L(t, q, \dot{q}) = \frac{1}{2}(\dot{q}^2 - q^2), \quad \ddot{q} + q = 0, \quad (58)$$

its solution  $q(t) = \sin t$  on the interval  $I = [0, 3\pi/2]$ , and the variation  $h(t) = t(t - 3\pi/2)$ , that vanishes at the endpoints of  $I$ . The action on  $q + \lambda h$  is elementary to compute:

$$\begin{aligned} S_I(q + \lambda h) &= \int_0^{3\pi/2} L(t, q(t) + \lambda h(t), \dot{q}(t) + \lambda \dot{h}(t)) dt = \\ &= -\lambda^2 \cdot \frac{9\pi^3}{640} (9\pi^2 - 40). \end{aligned}$$

Since  $9\pi^2 > 40$ , the action  $S_I$  has a strict *maximum* at the motion  $q$  if restricted to the straight line going through  $q$  in the direction  $h$ . In particular, it cannot possibly be a minimum according to the Definition 2. It can be proved that, if we restrict the interval to  $J = [0, b]$  with  $0 < b < \pi$ , the Hamiltonian action  $S_J$  has indeed a strict minimum at  $q$ .

There is a theory, called of the “conjugate points”, that explains very elegantly the counterexample we have just seen. We will not touch further on this topic, except by observing that all the solutions of  $\ddot{q} + q = 0$  such that  $q(0) = 0$  will also vanish at  $\pi$ .

Let us finally come to main result that we announced:

**Theorem 3** (Local least Hamiltonian action). *Suppose that  $L$  is  $C^2$  on an open set  $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ ,  $I_0$  is a compact interval,  $q: I_0 \rightarrow \mathbb{R}^n$  is  $C^1$  and such that  $(t, q(t), \dot{q}(t)) \in \Omega$  for all  $t \in I_0$ . Suppose that the symmetric matrix  $\partial_{\dot{q}, \dot{q}}^2 L$  is positive definite on  $\Omega$ . Suppose finally that  $q: I_0 \rightarrow \mathbb{R}^n$  is stationary for  $S_{I_0}$ . Then there exists  $a > 0$  such that, for any subinterval  $I$  shorter than  $a$ , the restriction of  $q$  to  $I$  is a strict minimum for  $S_I$ .*

Let us introduce the following quadratic form on  $C_0^1(I, \mathbb{R}^n)$ , called the *second variation* of  $S_I$

$$S_I''(q)[h, h] = \int_I \left( h(t) \cdot \partial_{q, q}^2 L h(t) + 2\dot{h}(t) \cdot \partial_{q, q}^2 L h(t) + \dot{h}(t) \cdot \partial_{\dot{q}, \dot{q}}^2 L \dot{h}(t) \right) dt, \quad (59)$$

where the matrices  $\partial_{q, q}^2 L$ ,  $\partial_{\dot{q}, \dot{q}}^2 L$ ,  $\partial_{q, \dot{q}}^2 L$  are evaluated at  $(t, q(t), \dot{q}(t))$ . The expression for the second variation is obtained by taking the second derivative of the scalar function  $\lambda \mapsto S_I(q + \lambda)$  at  $\lambda = 0$ , using the theorems on derivation under the integral sign, and using the symmetry  $x \cdot \partial_{q, \dot{q}}^2 L y = y \cdot \partial_{\dot{q}, q}^2 L x$ .

The proof of the theorem will need the next three lemmas.

**Lemma 1.** *If  $h \in C_0^1(I)$ , then*

$$\|h\|_{L^2(I)} \leq \frac{|I|}{\sqrt{2}} \|\dot{h}\|_{L^2(I)}.$$

*Proof.* Let  $I = [a, b]$ . Since  $h(a) = 0$  we can write  $h(t) = h(a) + \int_a^t \dot{h} = \int_a^t \dot{h}$ , whence  $\|h(t)\| \leq \int_a^t \|\dot{h}(s)\| ds$ , and, using Schwarz inequality,

$$\|h(t)\| \leq \sqrt{t-a} \left( \int_a^t \|\dot{h}(s)\|^2 ds \right)^{1/2} \leq \sqrt{t-a} \|\dot{h}\|_{L^2(I)},$$

so that

$$\|h\|_{L^2(I)}^2 = \int_a^b \|h(t)\|^2 dt \leq \|\dot{h}\|_{L^2(I)}^2 \int_a^b (t-a) dt = \frac{(b-a)^2}{2} \|\dot{h}\|_{L^2(I)}^2,$$

whence the result by taking square roots.  $\square$

**Lemma 2** (Remainder for 2nd order Taylor). *Suppose that  $L$  is  $C^2$  on an open set  $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ ,  $I_0$  is a compact interval,  $q: I_0 \rightarrow \mathbb{R}^n$  is  $C^1$  and such that  $(t, q(t), \dot{q}(t)) \in \Omega$  for all  $t \in I_0$ . Let  $\varepsilon > 0$ . Then there exists*

$\delta > 0$  such that for all compact subintervals  $I \subseteq I_0$  and for all  $h \in C_0^1(I)$ , if  $\|h\|_{C^1(I)} < \delta$  then

$$\left| S_I(q+h) - S_I(q) - A'_I(q)[h] - A''_I(q)[h, h] \right| \leq \varepsilon \left( \frac{|I_0|^2}{2} + 1 \right) \|\dot{h}\|_{L^2(I)}^2.$$

*Proof.*  $\Omega$  is open in  $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$  and it contains the image of the compact set  $I_0$  through the continuous mapping  $t \mapsto (t, q(t), \dot{q}(t))$ . By (uniform) continuity of the second derivatives of  $L$ , there exists  $\delta > 0$  such that if  $t \in I_0$ ,  $x, y \in \mathbb{R}^n$  and  $\|x\| + \|y\| < \delta$ , then  $(t, q(t) + x, \dot{q}(t) + y) \in \Omega$  and

$$\begin{aligned} \left\| \partial_{\dot{q}, \dot{q}}^2 L(t, q(t) + x, \dot{q}(t) + y) - \partial_{\dot{q}, \dot{q}}^2 L(t, q(t), \dot{q}(t)) \right\| &< \varepsilon, \\ \left\| \partial_{\dot{q}, q}^2 L(t, q(t) + x, \dot{q}(t) + y) - \partial_{\dot{q}, q}^2 L(t, q(t), \dot{q}(t)) \right\| &< \varepsilon, \\ \left\| \partial_{q, q}^2 L(t, q(t) + x, \dot{q}(t) + y) - \partial_{q, q}^2 L(t, q(t), \dot{q}(t)) \right\| &< \varepsilon. \end{aligned}$$

Let  $I$  be a compact subinterval of  $I_0$ , and  $h \in C_0^1(I)$  such that  $\|h\|_{C^1(I)} < \delta$ . The mapping  $\lambda \mapsto S_I = (q + \lambda h)$  turns out to be  $C^2$  in  $[0, 1]$  because we can take the derivative under the integral sign, and Taylor's formula applies:

$$\begin{aligned} S_I(q+h) &= S_I(q) + A'_I(q)[h] + \frac{1}{2} A''_I(q + \xi h)[h, h] = \\ &= S_I(q) + A'_I(q)[h] + \frac{1}{2} A''_I(q)[h, h] + \\ &\quad + \frac{1}{2} \left( A''_I(q + \xi h)[h, h] - A''_I(q)[h, h] \right) \end{aligned} \tag{60}$$

for some constant  $\xi \in [0, 1]$ . We can estimate the remainder as

$$\begin{aligned} \frac{1}{2} \left| A''_I(q + \xi h)[h, h] - A''_I(q)[h, h] \right| &\leq \\ &\leq \frac{1}{2} \varepsilon \|\dot{h}\|_{L^2(I)}^2 + \varepsilon \int_I \|h(t)\| \cdot \|\dot{h}(t)\| dt + \frac{1}{2} \varepsilon \|h\|_{L^2(I)}^2 \leq \\ &\leq \frac{1}{2} \varepsilon \|\dot{h}\|_{L^2(I)}^2 + \varepsilon \|h\|_{L^2(I)} \cdot \|\dot{h}\|_{L^2(I)} + \frac{1}{2} \varepsilon \|h\|_{L^2(I)}^2 = \\ &= \frac{\varepsilon}{2} \left( \|h\|_{L^2(I)} + \|\dot{h}\|_{L^2(I)} \right)^2 \leq \varepsilon \left( \|h\|_{L^2(I)}^2 + \|\dot{h}\|_{L^2(I)}^2 \right) \leq \\ &\leq \varepsilon \left( \frac{|I|^2}{2} + 1 \right) \|\dot{h}\|_{L^2(I)}^2. \end{aligned}$$

□

**Lemma 3** (Estimate of the 2nd variation). *Suppose that  $L$  is  $C^2$  on an open set  $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ ,  $I_0$  is a compact interval,  $q: I_0 \rightarrow \mathbb{R}^n$  is  $C^1$*

and such that  $(t, q(t), \dot{q}(t)) \in \Omega$  for all  $t \in I_0$ . Suppose that the symmetric matrix  $\partial_{\dot{q}, \dot{q}}^2 L$  is positive definite on  $\Omega$ . Then there exist constants  $\alpha, \beta > 0$  such that for all compact subintervals  $I \subseteq I_0$  and for all  $h \in C_0^1(I)$

$$S_I''(q)[h, h] \geq \left( \alpha - \beta \left( \frac{|I|}{\sqrt{2}} + \frac{|I|^2}{2} \right) \right) \|\dot{h}\|_{L^2(I)}^2$$

*Proof.* Define the following matrix-valued functions:

$$\begin{aligned} \varphi_1(t) &= \partial_{\dot{q}, \dot{q}}^2 L(t, q(t), \dot{q}(t)), & \varphi_2(t) &= \partial_{\dot{q}, \dot{q}}^2 L(t, q(t), \dot{q}(t)), \\ \varphi_3(t) &= \partial_{\dot{q}, q}^2 L(t, q(t), \dot{q}(t)). \end{aligned}$$

By compactness and continuity, there exist  $\alpha, \beta > 0$  be such that

$$\begin{aligned} y \cdot \varphi_1(t)y &\geq \alpha \|y\|^2, \\ |y \cdot \varphi_2(t)z| &\leq \beta \|y\| \cdot \|z\|, \quad |y \cdot \varphi_3(t)y| \leq \beta \|y\|^2 \end{aligned}$$

for all  $t \in I_0$ ,  $y, z \in \mathbb{R}^n$ . Then for all compact subintervals  $I \subseteq I_0$  and for all  $h \in C_0^1(I)$

$$\begin{aligned} S_I''(q)[h, h] &= \int_I \left( \dot{h}(t) \cdot \varphi_1(t)\dot{h}(t) + \right. \\ &\quad \left. + 2\dot{h}(t) \cdot \varphi_2(t)h(t) + h(t) \cdot \varphi_3(t)h(t) \right) dt \geq \\ &\geq \alpha \|\dot{h}\|_{L^2(I)}^2 - \beta (\|\dot{h}\|_{L^2(I)} \|h\|_{L^2(I)} + \|h\|_{L^2(I)}^2) \geq \\ &\geq \alpha \|\dot{h}\|_{L^2(I)}^2 - \beta \left( \frac{|I|}{\sqrt{2}} + \frac{|I|^2}{2} \right) \|\dot{h}\|_{L^2(I)}^2. \end{aligned}$$

□

*Proof.* (Of the local least action theorem) Let  $\alpha, \beta > 0$  be as in Lemma 3. Choose  $c > 0$  small enough so that

$$\frac{\alpha}{2} - \beta \left( \frac{a}{\sqrt{2}} + \frac{a^2}{2} \right) > 0.$$

Let  $\varepsilon > 0$  be such that

$$\varepsilon \left( \frac{|I_0|^2}{2} + 1 \right) \leq \frac{\alpha}{2}.$$

Let  $\delta > 0$  be associated to  $\varepsilon$  by Lemma 2. Let  $I$  be a subinterval of  $I_0$  with length  $\leq c$ , and  $h \in C_0^1(I)$  such that  $0 < \|h\|_{C^1(I)} < \delta$ . Then again using Taylor's formula (60)

$$S_I(q+h) = S_I(q) + A_I'(q)[h] + \frac{1}{2} A_I''(q + \xi h)[h, h] =$$

$$\begin{aligned}
&= S_I(q) + \frac{1}{2}A_I''(q)[h, h] + \\
&\quad + \frac{1}{2}\left(A_I''(q + \xi h)[h, h] - A_I''(q)[h, h]\right) \geq \\
&\geq S_I(q) + \left(\alpha - \beta\left(\frac{|I|}{\sqrt{2}} + \frac{|I|^2}{2}\right)\right)\|\dot{h}\|_{L^2(I)}^2 - \\
&\quad - \varepsilon\left(\frac{|I_0|^2}{2} + 1\right)\|\dot{h}\|_{L^2(I)}^2 \geq \\
&\geq S_I(q) + \left(\frac{\alpha}{2} - \beta\left(\frac{|I|}{\sqrt{2}} + \frac{|I|^2}{2}\right)\right)\|\dot{h}\|_{L^2(I)}^2 \geq \\
&\geq S_I(q) + \left(\frac{\alpha}{2} - \beta\left(\frac{c}{\sqrt{2}} + \frac{c^2}{2}\right)\right)\|\dot{h}\|_{L^2(I)}^2 > \\
&> S_I(q).
\end{aligned}$$

We conclude that the restriction of  $q$  to  $I$  is a weak strong minimum for  $S_I$ .  $\square$

## 7. Geodesics and Jacobi Lagrangians

Let us consider a conservative Lagrangian of the form

$$L(q, \dot{q}) = \frac{1}{2}\dot{q} \cdot G(q)\dot{q} - U(q), \quad (61)$$

where the *potential energy*  $U$  is of class  $C^1$  on an open set  $D \subseteq \mathbb{R}^n$  and  $G \in C^1$  is a matrix-valued function on  $D$ . The term  $\dot{q} \cdot G(q)\dot{q}/2$  is the *kinetic energy*. Each matrix  $G(q)$  is assumed symmetric and positive definite:  $G(q) = G(q)^T > 0$ . This is the Lagrangian of a system with smooth constraints and subject to forces that are time-independent gradients. What we are going to say can be easily generalized to Riemannian manifolds, where  $G(q)$  is the matrix of the components of the metric tensor in a local chart. In the sequel we will call *metric* the matrix field  $G$ .

Although within our assumptions the Lagrangian (61) is only  $C^1$  with respect to  $q$ , the reasoning that we made at the beginning of Section 6 can be repeated to show that here too the natural motions are  $C^2$  and the Lagrange equation can be rewritten into the following normal form

$$\ddot{q} = -\Gamma(q)[\dot{q}, \dot{q}] - G(q)^{-1}\nabla U(q), \quad (62)$$

where  $\nabla$  is the gradient operator, and for each  $q$  the mapping  $\Gamma(q): \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $(u, v) \mapsto \Gamma(q)[u, v]$ , is a symmetric bilinear form whose components are called *Christoffel symbols*

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{l=1}^n G_{kl}^{-1} (\partial_i G_{jl} + \partial_j G_{il} - \partial_l G_{ij}), \quad (63)$$

$$(\Gamma(q)[u, v])_k = \sum_{i,j=1}^n \Gamma(q)_{ij}^k u_i v_j. \quad (64)$$

The Lagrangian is time-independent, bringing along the *energy* first integral

$$E(q, \dot{q}) = \partial_{\dot{q}} L(q, \dot{q}) \cdot \dot{q} - L(q, \dot{q}) = \frac{1}{2} \dot{q} \cdot G(q) \dot{q} + U(q), \quad (65)$$

which is the sum of kinetic and potential energy.

Specially notable is the case when  $U(q) \equiv 0$ , that is, when the Lagrangian coincides with the kinetic energy. The solutions to the Lagrange equation (62) are then called *geodesics* for the metric  $G$ . Otherwise said, the geodesics are for which the Hamiltonian action functional

$$S_I(q) = \frac{1}{2} \int_I \dot{q}(t) \cdot G(q(t)) \dot{q}(t) dt, \quad (66)$$

is stationary, or, even more, the motions that actually minimize it when  $I$  is small enough, as we saw in Section 6. In Riemann geometry the *geodesic equation*  $\ddot{q} = -\Gamma(q)[\dot{q}, \dot{q}]$  represents the Levi-Civita parallel transport of the tangent vector.

In the sequel we will concentrate on geodesics, until the last section on the Jacobi metric, when we will see how the natural motions for a Lagrangian of the form (61) are closely tied to suitable geodesics. This is the reason why we propose the following terminology.

**Definition 3.** *The Lagrangians of the form  $L(q, \dot{q}) = \frac{1}{2} \dot{q} \cdot G(q) \dot{q} - U(q)$ , where the  $G(q)$  are symmetric, positive definite matrices, will be called *Jacobi Lagrangians*.*

These are the same as what some authors call “natural” Lagrangians, whilst others use the word “natural” to refer to a wider set, that includes for example Foucault’s pendulum.

**7.1. Geodesics and stationary arc length.** In the  $G$  metric, the length of a  $C^1$  curve  $q: [a, b] \rightarrow D \subseteq \mathbb{R}^n$  is given by

$$\ell_{[a,b]}(q) = \int_a^b \sqrt{\dot{q}(t) \cdot G(q(t)) \dot{q}(t)} dt. \quad (67)$$

If the radicand  $\dot{q}(t) \cdot G(q(t))\dot{q}(t)$  happens to be identically 1 for all  $t \in [a, b]$  then  $t$  will be called an *arc length* parameter. More generally, if the radicand is a positive constant we will say that the curve is parameterized proportionally to the arc length.

We are going to study how the length functional is related to the geodesics, whose Lagrangian and differential equation are

$$L(q, \dot{q}) = \frac{1}{2} \dot{q} \cdot G(q) \dot{q}, \quad \ddot{q} = -\Gamma(q)[\dot{q}, \dot{q}]. \quad (68)$$

The energy first integral (65) in this situation reduces to the kinetic energy, which means that the (nonconstant) geodesics are always parameterized proportionally to the arc length.

Since the right-hand side of the geodesic equation  $\ddot{q} = -\Gamma(q)[\dot{q}, \dot{q}]$  is homogeneous with respect to  $\dot{q}$ , whenever  $q(t)$  is a solution and  $t_0, c \in \mathbb{R}$ , also  $t \mapsto q(t_0 + ct)$  is a geodesic motion. Reparameterizations other than the affine ones are not admitted, because the geodesic equation has unique solution once  $q, \dot{q}$  are known at a given time.

We claim that the geodesics are also solutions to the variational problem associated to the length functional (67). That is, they make stationary the Hamiltonian action associated to the arc length Lagrangian

$$\mathcal{L}(q, \dot{q}) = \sqrt{\dot{q} \cdot G(q) \dot{q}} = \sqrt{2L(q, \dot{q})}, \quad (69)$$

in addition, of course, to the action associated to the kinetic energy Lagrangian  $L$ . In fact, let  $q(t)$  be a geodesic motion. Thanks to the fact that  $L$  is constant with respect to  $t$  along  $q$ , we can write

$$\frac{d}{dt} \partial_{\dot{q}} \mathcal{L} - \partial_q \mathcal{L} = \frac{d}{dt} \left( \frac{\partial_{\dot{q}} L}{\sqrt{2L}} \right) - \frac{\partial_q L}{\sqrt{2L}} = \quad (70)$$

$$= \frac{1}{\sqrt{2L}} \left( \frac{d}{dt} \partial_{\dot{q}} L - \partial_q L \right), \quad (71)$$

and the last member (71) vanishes identically because  $q(t)$  is a geodesic. We conclude that  $q$  solves also the Lagrange equation associated to  $\mathcal{L}$ .

The reverse is not true: the curves that make the length functional  $\ell$  are many more than the geodesics. Notice in fact that  $\mathcal{L}$  is positive homogeneous of degree 1 with respect to  $\dot{q}$ , which implies that *any* reparameterization of a stationary curve will be stationary too. In details, let  $\tau: J \rightarrow I = [a, b]$  be a  $C^1$  increasing diffeomorphism, and  $q: I \rightarrow D$  be any  $C^1$  function. Then

$$\ell_I(q) = \int_I \mathcal{L}(q(t), \dot{q}(t)) dt = \int_J \mathcal{L}(q(t), \dot{q}(t)) \Big|_{t=\tau(\xi)} \tau'(\xi) d\xi = \ell_J(q \circ \tau).$$

If  $q$  is stationary for  $\ell_I$ , then for all  $h \in C_0^1(J, \mathbb{R}^n)$

$$\frac{d}{d\lambda} \ell_J(q \circ \tau + \lambda h) = \frac{d}{d\lambda} \ell_J(q \circ \tau + \lambda h \circ \tau^{-1} \circ \tau) = \frac{d}{d\lambda} \ell_I(q + \lambda h \circ \tau^{-1}) = 0$$

so that  $q \circ \tau$  is stationary too.

Take now a  $q$  that is stationary for the length functional, and for which  $\dot{q}$  never vanishes. We say that  $q$  is a reparameterization of a geodesic. Consider in fact the parameter, proportional to the arc length,

$$\sigma(t) = \frac{1}{c} \int_a^t \mathcal{L}(q(\xi), \dot{q}(\xi)) d\xi$$

with  $c \neq 0$ , and consider the reparameterization  $\tilde{q} = q \circ \sigma^{-1}$ . We can write

$$\dot{\tilde{q}}(\sigma(t)) \frac{\mathcal{L}(q(t), \dot{q}(t))}{c} = \dot{q}(t), \quad L(\tilde{q}(\sigma(t)), \dot{\tilde{q}}(\sigma(t))) = c^2/2.$$

Hence the mapping  $s \mapsto L(\tilde{q}(s), \dot{\tilde{q}}(s))$  is constant, and the computation (70–71), read backward with  $s, \tilde{q}, \dot{\tilde{q}}$  replacing  $t, q, \dot{q}$ , shows that  $\tilde{q}$  is a geodesic. If we take  $c = \ell_I(q)$  then  $\tilde{q}$  is defined on  $[0, 1]$ . Another interesting choice is  $c = \mathcal{L}(q(a), \dot{q}(a))$ , because it gives  $\dot{\sigma}(a) = 1$  and  $\dot{\tilde{q}}(0) = \dot{q}(a)$ .

If the reparameterizations do not change the stationarity for the length functional, they will also leave invariant the property of being solutions to the associated Lagrange equation. A consequence is that the Cauchy problems for that differential equation do not have unique solution. In fact, if we start with a solution and we apply an arbitrary change of parameter  $\tau$  for which  $\tau(t_0) = t_0$  and  $\tau'(t_0) = 1$ , the result is a different solution, although the initial data at the time  $t_0$  are the same.

Increasing the regularity of the mapping  $q \mapsto G(q)$  from  $C^1$  to  $C^2$  would not help recover uniqueness, because the obstacle lies in the fact that the Lagrange equation for  $\ell_I$  cannot be written in normal form. In fact, starting from the fact that  $\lambda \mathcal{L}(q, \dot{q}) = \mathcal{L}(q, \lambda \dot{q})$  for all  $\lambda > 0$ , taking the derivative with respect to  $\lambda$  at  $\lambda = 1$  we obtain  $\mathcal{L}(q, \dot{q}) = \partial_{\dot{q}} \mathcal{L}(q, \dot{q}) \cdot \dot{q}$ , which implies that

$$\partial_{\dot{q}} \mathcal{L}(q, \dot{q}) = \partial_{\dot{q}} \mathcal{L}(q, \dot{q}) + \partial_{\dot{q}, \dot{q}}^2 \mathcal{L}(q, \dot{q}) \dot{q},$$

so that the Hessian matrix  $\partial_{\dot{q}, \dot{q}}^2 \mathcal{L}(q, \dot{q})$  cannot be invertible.

**7.2. Geodesics and Noether's theorem.** We are going to study two geodesic problems in the framework of first integrals and Noether's theorem.

**7.2.1. Logarithmic geodesics.** In  $\mathbb{R}^2$  define the metric  $G(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} e^{2x}$  with its associate Lagrangian

$$L(q, \dot{q}) = \frac{1}{2} e^{2q_1} (\dot{q}_1^2 + \dot{q}_2^2). \quad (72)$$

This system has two obvious symmetries: invariance by time translations and by vertical translations. Noether’s theorem supplies the two associated first integrals:  $L(q, \dot{q}), e^{2q_1} \dot{q}_2$ .

To find hidden symmetries it is convenient to identify  $\mathbb{R}^2$  with the complex plane  $\mathbb{C}$  and  $q(t)$  with the complex-valued curve  $z(t) = q_1(t) + iq_2(t)$ . The Lagrangian along  $z(t)$  takes the form

$$L(z(t), \dot{z}(t)) = \frac{1}{2} |\dot{z}(t)e^{z(t)}|^2 = \frac{1}{2} \left| \frac{d}{dt} e^{z(t)} \right|^2.$$

For a given  $z(t)$  defined in a neighbourhood of  $t_0$ , there is invariance of the Lagrangian under the following three-parameter family of curves:

$$z_\sigma(t) = z_{\sigma_1, \sigma_2, \sigma_3}(t) = \log \left( \sigma_1 + i\sigma_2 + e^{i\sigma_3} e^{z(t)} \right), \tag{73}$$

where we choose a branch of the complex logarithm that is regular in a neighbourhood of  $e^{z(t_0)}$ . In fact  $z_0(t) = z(t)$  for all  $t$  close enough to  $t_0$ , and

$$\frac{d}{dt} e^{z_\sigma(t)} = \frac{d}{dt} (\sigma_1 + i\sigma_2 + e^{i\sigma_3} e^{z(t)}) = e^{i\sigma_3} \frac{d}{dt} e^{z(t)},$$

so that

$$L(z_\sigma(t), \dot{z}_\sigma(t)) = \frac{1}{2} \left| \frac{d}{dt} e^{z_\sigma(t)} \right|^2 = \frac{1}{2} \left| e^{i\sigma_3} \frac{d}{dt} e^{z(t)} \right|^2 = L(z(t), \dot{z}(t)).$$

To find the associate first integrals, let us start with  $\sigma_1$ :

$$\frac{\partial}{\partial \sigma_1} z_{\sigma_1, 0, 0}(t) \Big|_{\sigma_1=0} = \frac{1}{\sigma_1 + e^{z(t)}} \Big|_{\sigma_1=0} = e^{-z(t)} = e^{-q_1(t)} (\cos q_2(t) - i \sin q_2(t)).$$

Hence, along the geodesics the following quantity is conserved:

$$\begin{aligned} \partial_{\dot{q}} L(q, \dot{q}) \cdot \partial_{\sigma_1} q_{\sigma_1, 0, 0}(t) \Big|_{\sigma_1=0} &= e^{2q_1} (\dot{q}_1, \dot{q}_2) \cdot e^{-q_1} (\cos q_2, -\sin q_2) = \\ &= e^{q_1} (\dot{q}_1 \cos q_2 - \dot{q}_2 \sin q_2) = \Re(e^z \dot{z}) = \Re \left( \frac{d}{dt} e^{z(t)} \right). \end{aligned}$$

With similar calculations with  $\sigma_2$  we arrive at the conservation of  $\Im \frac{d}{dt} e^{z(t)}$ . We deduce that  $\frac{d}{dt} e^{z(t)}$  is a complex-valued first integral. From this we can deduce an explicit formula for all geodesics:

$$z(t) = z(0) + \log(1 + t\dot{z}(0)), \tag{74}$$

where we take the principal branch of the complex logarithm. The geodesics are global in time, except when  $\dot{z}(0)$  is a nonzero real number.

We leave it as an exercise to find what first integral is associated to the parameter  $\sigma_3$ , and what relations (if any) hold among all these constants of motion.

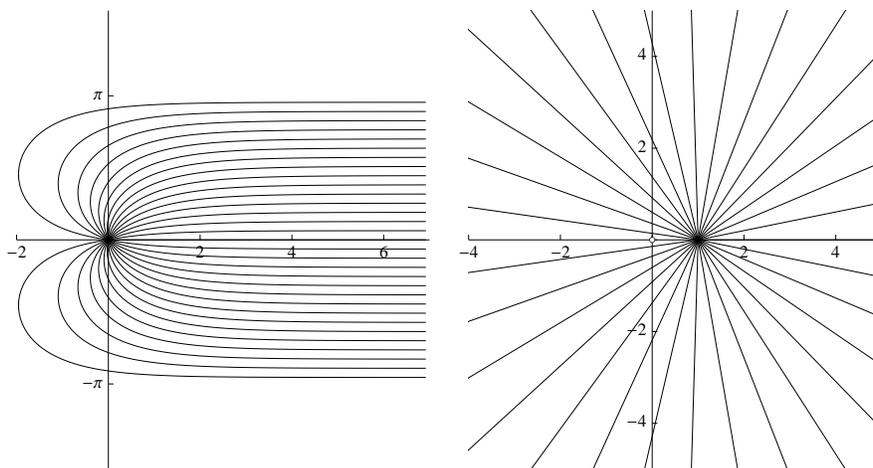


FIGURE 1. To the left, geodesics through the origin for the Lagrangian (72), which are the complex logarithm of the straight lines through  $(1, 0)$  (right).

An interesting feature of this example is the following: since the principal complex logarithm has imaginary part between  $-\pi$  and  $\pi$ , there are no geodesics that connect two points in the plane whose vertical distance is larger than  $2\pi$ . For example, as illustrated in the left half of Figure 1, no geodesic that goes through the origin ever exits the stripe  $-\pi < \Im z < \pi$ .

The Lagrange equations for the Lagrangian (72) are

$$e^{2q_1}(\ddot{q}_1 + \dot{q}_1^2 - \dot{q}_2^2) = 0, \quad e^{2q_1}(\ddot{q}_2 + 2\dot{q}_1\dot{q}_2) = 0, \quad (75)$$

and they are equivalent to  $\ddot{z} + \dot{z}^2 = 0$ , a differential equation that is readily solved. However, we wanted to show a Noetherian point of view, at the expense of speed.

**7.2.2. Poincaré's half-plane.** Our second example is Poincaré's half-plane, a classical model for Lobachevsky's hyperbolic geometry, that historically came after equivalent models by Klein and, earlier, by Beltrami. It is the half-plane  $D = \{(x, y) \in \mathbb{R}^2 : y > 0\}$  endowed with the metric defined by  $G(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} / y^2$ , so that the Lagrangian is

$$L(q, \dot{q}) = \frac{\dot{q}_1^2 + \dot{q}_2^2}{2q_2^2}. \quad (76)$$

This  $L$  has three obvious symmetries: it is time-independent, it is invariant by dilations  $q \mapsto e^\sigma q$ , and it is also invariant by horizontal translations.

The associated first integrals by Noether's theorem are

$$\frac{\dot{q}_1^2 + \dot{q}_2^2}{2q_2^2} = a, \quad \frac{q_1\dot{q}_1 + q_2\dot{q}_2}{q_2^2} = b, \quad \frac{\dot{q}_1}{q_2^2} = c, \quad (77)$$

with  $a > 0$  for nonconstant geodesics. When  $c = 0$  we see that the trajectory is a vertical straight line. For  $c \neq 0$ , with the following algebraic manipulation we can eliminate  $\dot{q}_1, \dot{q}_2$  from these equations. Solving the third equation for  $\dot{q}_1$  and replacing into the others, we have

$$\frac{c^2 q_2^4 + \dot{q}_2^2}{2q_2^2} = a, \quad cq_1 + \frac{\dot{q}_2}{q_2} = b.$$

Solving for  $\dot{q}_2$  in this last equation, replacing into the preceding one and simplifying, we obtain

$$(q_1 - b/c)^2 + q_2^2 = R^2, \quad \text{dove } R := \sqrt{2a/|c|},$$

which will be recognized as the equation of a circle with its center on the horizontal axis.

On Poincaré's half-plane every couple of points is connected by a geodesic. Figure 2 shows that given a geodesic  $r$  and a point  $P$  outside, there exist infinitely many geodesics through  $P$  that does not cross  $r$ . It is a property that replaces Euclid's parallel postulate.

Having found the geodesic trajectories, we can calculate explicitly their parameterization. When  $c = 0$  the motion is a straight vertical line with exponential parameterization  $q_2(t) = q_2(0)e^{bt}$ . For the semicircles, introducing the polar angle  $\theta(t)$ , we can write

$$q_1(t) = b/c + R \cos \theta(t), \quad q_2(t) = R \sin \theta(t). \quad (78)$$

From  $\dot{q}_1 = cq_2^2$  we get that  $-R\dot{\theta} \sin \theta = cR^2 \sin^2 \theta$  and  $\dot{\theta} = -cR \sin \theta$ , which is readily solved, giving

$$\theta(t) = 2 \arctan e^{cR(t_0-t)}, \quad (79)$$

where  $t_0$  is an integration constant, whose meaning is the instant when  $\theta(t_0) = \pi/2$ , and is related to the constants of motion (77) by  $q_1(t_0) = b/c$ .

**7.3. Jacobi metric.** It is very interesting to interpret some aspects of a mechanical problem within a geodesic framework. This is possible for what we called Jacobi Lagrangians, that is, for the Lagrangians of the form

$$L(q, \dot{q}) := \frac{1}{2} \dot{q} \cdot G(q) \dot{q} - U(q), \quad q \in D = D^\circ \subseteq \mathbb{R}^n, \quad \dot{q} \in \mathbb{R}^n. \quad (80)$$

As we noted, along the solutions to the Lagrange equation the energy is preserved

$$\frac{1}{2} \dot{q} \cdot G(q) \dot{q} + U(q) \equiv \mathcal{E}. \quad (81)$$

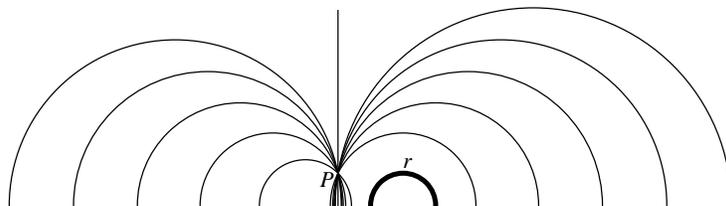


FIGURE 2. Geodesics in Poincaré's half-plane.

Let us concentrate our attention to the class of solutions on the same given energy level  $\mathcal{E} \in \mathbb{R}$ . Then further restrict to the (parts of) solutions that live in the open set  $D_{\mathcal{E}} := \{x \in D \mid U(x) < \mathcal{E}\}$ .

As an example, for the harmonic oscillator with Lagrangian  $L(q, \dot{q}) = (\dot{q}^2 - q^2)/2$ , and for the energy level  $\mathcal{E} = 1/2$ , the set  $D_{\mathcal{E}}$  is the space interval  $(-1, 1)$ , and the motion  $q(t) = \cos t$  will have to be restricted to the time intervals  $(m\pi, m\pi + \pi)$ , with  $m \in \mathbb{Z}$ .

To the Lagrangian (80) and to the energy level  $\mathcal{E}$  we associate the *Jacobi metric* with parameter  $\mathcal{E}$ , which is the matrix field

$$x \mapsto (\mathcal{E} - U(x))G(x) \quad (82)$$

on the open set  $D_{\mathcal{E}}$ . The new Lagrangian associated to the Jacobi metric is

$$\mathcal{J}_{\mathcal{E}}(q, \dot{q}) := \frac{1}{2}(\mathcal{E} - U(q))\dot{q} \cdot G(q)\dot{q}, \quad (q, \dot{q}) \in D_{\mathcal{E}} \times \mathbb{R}^n. \quad (83)$$

The geodesics for the Jacobi metric are the solution of the differential equation

$$\frac{d}{dt} \partial_{\dot{q}} \mathcal{J}_{\mathcal{E}}(q(t), \dot{q}(t)) - \partial_q \mathcal{J}_{\mathcal{E}}(q(t), \dot{q}(t)) = 0.$$

The final result that we are going to present is that the natural motions for the Lagrangian  $L$  on the energy level  $\mathcal{E}$  that live in  $D_{\mathcal{E}}$  coincide, up to reparameterizations, with the geodesics of  $\mathcal{J}_{\mathcal{E}}$ . We stress that this gives information about the trajectories of the natural motions, and not on how they are parameterized.

**Theorem 4.** *Let  $D$  be a nonempty open set in  $\mathbb{R}^n$ , let  $\mathcal{E} \in \mathbb{R}$  be a constant, and let  $U: D \rightarrow \mathbb{R}$  be a  $C^1$  function such that  $U(q) < \mathcal{E}$  for all  $q \in D$ . Let  $G(q)$  be a symmetric, positive definite matrix, depending in a  $C^1$  way on  $q \in D$ . Consider the two Lagrangians:*

$$L(q, \dot{q}) := \frac{1}{2}\dot{q} \cdot G(q)\dot{q} - U(q), \quad \mathcal{J}_{\mathcal{E}}(q, \dot{q}) := \frac{1}{2}(\mathcal{E} - U(q))\dot{q} \cdot G(q)\dot{q}. \quad (84)$$

Then a suitable reparameterization will turn a motion in  $D$  that is natural for  $L$  and has energy  $\mathcal{E}$ , into a nonconstant motion that is natural for  $\mathcal{J}_{\mathcal{E}}$ , and vice versa.

*Proof.* In the proof we will use the arc length Lagrangian associated to the Jacobi metric

$$\mathcal{L}_{\mathcal{E}}(q, \dot{q}) := \sqrt{2\mathcal{J}_{\mathcal{E}}(q, \dot{q})} = \sqrt{(\mathcal{E} - U(q))\dot{q} \cdot G(q)\dot{q}}, \quad (q, \dot{q}) \in D \times \mathbb{R}^n,$$

and the relations between  $\mathcal{J}_{\mathcal{E}}$  and  $\mathcal{L}_{\mathcal{E}}$  that were discussed in Subsection 7.1.

Let  $q(t)$  be a  $C^1$  motion in  $D$  that is stationary for the Hamiltonian action defined by  $L$ , and has energy  $\mathcal{E}$ . Because of conservation of energy (81), along the motion we can write

$$\begin{aligned} \mathcal{L}_{\mathcal{E}}(q(t), \dot{q}(t)) &= \sqrt{(\mathcal{E} - U(q))\dot{q} \cdot G(q)\dot{q}} = \sqrt{\frac{1}{2}(\dot{q} \cdot G(q)\dot{q})^2} = \\ &= \frac{1}{\sqrt{2}}\dot{q} \cdot G(q)\dot{q} = \sqrt{2}(\mathcal{E} - U(q)) > 0, \end{aligned}$$

since  $U < \mathcal{E}$  in  $D$ . Using these relations, again along the motion  $q(t)$  we have that

$$\begin{aligned} \frac{d}{dt}\partial_{\dot{q}}\mathcal{L}_{\mathcal{E}} - \partial_q\mathcal{L}_{\mathcal{E}} &= \frac{d}{dt}\left(\frac{\partial_{\dot{q}}\mathcal{J}_{\mathcal{E}}}{\mathcal{L}_{\mathcal{E}}}\right) - \frac{\partial_q\mathcal{J}_{\mathcal{E}}}{\mathcal{L}_{\mathcal{E}}} = \tag{85} \\ &= \frac{d}{dt}\left(\frac{(\mathcal{E} - U(q))G(q)\dot{q}}{\mathcal{L}_{\mathcal{E}}}\right) - \\ &\quad - \frac{(\mathcal{E} - U(q))\partial_q(\dot{q} \cdot G(q)\dot{q})}{\mathcal{L}_{\mathcal{E}}} + \\ &\quad + \frac{\frac{1}{2}(\dot{q} \cdot G(q)\dot{q})\nabla U(q)}{\mathcal{L}_{\mathcal{E}}} = \\ &= \frac{1}{\sqrt{2}}\left(\frac{d}{dt}(G(q)\dot{q}) - \partial_q(\dot{q} \cdot G(q)\dot{q}) + \nabla U(q)\right) = \\ &= \frac{1}{\sqrt{2}}\left(\frac{d}{dt}\partial_{\dot{q}}L - \partial_qL\right). \tag{86} \end{aligned}$$

This last member (86) vanishes identically because of the Lagrange equation for  $L$ . Hence the motion  $q(t)$  is a solution to the Lagrange equation for  $\mathcal{L}_{\mathcal{E}}$  too, and it stationary for the Hamiltonian action associated with  $\mathcal{L}_{\mathcal{E}}$ . Moreover,  $\dot{q}$  never vanishes, because the kinetic energy is  $\mathcal{E} - U > 0$  on  $D$ . For a general fact that we saw in subsection 7.1, there exists a reparameterization of  $q(t)$  that is a geodesic for the Jacobi metric.

Conversely, let  $\tilde{q}: I \rightarrow D$  be a nonconstant geodesic for  $\mathcal{J}_{\mathcal{E}}$ . We know that any of its reparameterization will be stationary for the arc length

Lagrangian  $\mathcal{L}_{\mathcal{E}}$ , that is, it will make the left-hand side of (85) vanish identically. We claim that there is a particular reparameterization that has the energy constantly equal to  $\mathcal{E}$ , that is, it satisfies condition (81). In fact, write the reparameterization as  $q(t) = \tilde{q}(\tau(t))$  and let us impose the energy level:

$$\begin{aligned}\mathcal{E} &\equiv \frac{1}{2}\dot{q} \cdot G(q)\dot{q} + U(q) = \\ &= \frac{1}{2}\dot{\tau}(t)^2 \dot{\tilde{q}}(\tau(t)) \cdot G(\tilde{q}(\tau(t)))\dot{\tilde{q}}(\tau(t)) + U(\tilde{q}(\tau(t))).\end{aligned}$$

We can solve the equation for  $\dot{\tau}(t)$  and obtain

$$\dot{\tau}(t) = \pm \sqrt{2 \frac{\mathcal{E} - U(\tilde{q}(\tau(t)))}{\dot{\tilde{q}}(\tau(t)) \cdot G(\tilde{q}(\tau(t)))\dot{\tilde{q}}(\tau(t))}}$$

which is an autonomous scalar differential equation of the first order in the unknown function  $\tau(t)$ , choosing the sign  $+$  or  $-$  according to our preference for an increasing or a decreasing change of parameter. The variables in the differential equation can be separated and therefore there exists a solution in the large  $\tau: J \rightarrow I$ ,  $\tau(J) = I$  (unique up to time translations). We end up with a  $q(t)$  that makes the term (85) vanish identically and that satisfies (81). We can then follow the calculations backward from (86) to (85) and reach the conclusion that  $q(t)$  solves the Lagrange equation for  $L$ , as desired.  $\square$

The reader can find alternative, geometric approaches in the books by Oliva [12, sec. 5.2] and Marsden and Ratiu [11, sec. 7.7].

## References

- [1] Vladimir I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed. Springer, 1989.
- [2] Giuseppe Buttazzo, Mariano Giaquinta, Stefan Hildebrandt, *One-dimensional variational problems, an introduction*. Oxford Lecture Series in Mathematics and its Applications **15**, Clarendon Press, Oxford, 1998.
- [3] Charles Fox, *An introduction to the calculus of variations*. (Reprint of the Original publ. by Oxford University Press, London, 1950, 1963 corr. printing). Dover, 1987.
- [4] Giovanni Gallavotti, *The elements of mechanics*, Springer 1983. The second, 2007 edition is available from the author's home page at: <http://ipparco.roma1.infn.it>
- [5] Herbert Goldstein, Charles Poole, John Safko, *Classical mechanics*, 3rd ed. Addison Wesley, 2000.
- [6] Herbert Goldstein, *Prehistory of the "Runge-Lenz" vector*. *American Journal of Physics* **43** (1975) 737-738. *More on the prehistory of the Laplace or Runge-Lenz vector*. *American Journal of Physics* **44** (1976) 1123-1124.

- [7] Gianluca Gorni, Gaetano Zampieri, *Aspetti variazionali della meccanica, un approccio analitico*. To appear (2012) in “Complementi alle Lezioni”, a complementary vol. 3 to [9].
- [8] Gianluca Gorni, Gaetano Zampieri, *Revisiting Noether’s theorem on constants of motion*. In preparation.
- [9] Tullio Levi-Civita, Ugo Amaldi, *Lezioni di Meccanica Razionale*, Vol. 1 and Vol. 2, 1923. Republication Edizioni CompoMat, 2012.
- [10] Jean-Marc Lévy-Leblond, *Conservation laws for gauge-variant Lagrangians in classical mechanics*, Am. J. Phys. **39** (1971) 502–506.
- [11] Jerrold E. Marsden, Tudor S. Ratiu, *Introduction to mechanics and symmetry. A basic exposition of classical mechanical systems*. 2nd ed. Springer, 1999.
- [12] Waldyr M. Oliva, *Geometric mechanics*. Lecture Notes in Mathematics **1798**. Springer, 2002.
- [13] John L. Troutman, *Variational calculus and optimal control. Optimization with elementary convexity*. 2nd ed. Springer, 1996.
- [14] Gaetano Zampieri, *Completely integrable Hamiltonian systems with weak Lyapunov instability or isochrony*, Commun. Math. Phys. **303** (2011) 73–87.