ELSEVIER

J. Math. Pures Appl. 81 (2002) 1011–1052

JOURNAL DE MATHÉMATIQUES PURES ET APPLIQUÉES

Global variational methods on smooth nonholonomic constraints

Roberto Giambò *, Fabio Giannoni

Dipartimento di Matematica e Fisica, Università di Camerino, Italy

Accepted 16 August 2002

Abstract

We develop a variational theory for critical points of integral functionals in a space of curves on a manifold $\mathcal{M}$, between a fixed point and a one-dimensional submanifold of $\mathcal{M}$, and satisfying a nonholonomic constraint equation $\phi = 0$, where ϕ is a C^2 function defined on $T\mathcal{M} \times \mathbb{R}$.

We obtain existence, regularity and multiplicity results, writing the integro-differential equations satisfied by critical points. Moreover, we present some results concerning a sort of *exponential map* relative to the integro-differential equations and some examples.

Résumé

Nous développons une théorie variationnelle pour les points critiques des opérateurs intégraux dans un espace des courbes sur une variété $\mathcal{M}$, entre un point fixe et une sous-variété 1-dimensionnelle de $\mathcal{M}$, satisfaisant une équation de liaison non-holonome $\phi = 0$, où ϕ est une fonction C^2 définie sur $T\mathcal{M} \times \mathbb{R}$.

Nous obtenons des résultats d'existence, de régularité et de multiplicité, écrivant les équations integro-différentielle satisfaites par les points critiques. Nous présentons ainsi quelques résultats au sujet d'une sorte de *application exponentielle* relativement aux équations, et quelques exemples.

Keywords: Nonholonomic mechanics; Variational problems

Mots-clés : Méchanique non-holonome ; Problèmes variationelles

* Corresponding author.
E-mail addresses: roberto.giambo@unicam.it (R. Giambò), fabio.giannoni@unicam.it (F. Giannoni).
URLs: http://www2.unicam.it/~giambo, http://www2.unicam.it/~giannoni.

0021-7824/02/$ – see front matter
PII: S0021-7824(02)01269-2

1. Introduction

Many examples in applied mathematics lead to the study of variational problems with nonholonomic constraints, that is where the constraints are not only imposed on the configurations but also on the velocities, and arise as a submanifold of the *velocity phase space* (also called *state space*) (see [2]).

A first case is given by sub-Riemannian geodesics, where one searches for curves on a manifold $\mathcal{M}$ locally minimizing distance and such that their velocity is in a given subspace of the tangent space, for instance is orthogonal to a given vector field Y. Regularity of sub-Riemannian geodesics between two given points is still an open problem. The situation changes if we let the end point free to move on an integral curve of Y. Indeed, in this case a variational theory, completely analogous to the classical Riemannian geodesics one, can be developed (see [7], and Section 6.1). Another example, concerning Lorentzian geometry, is shown in Section 6.2. We refer to [3] for the main definitions and properties in Riemannian and Lorentzian geometry.

The aim of this work is to develop a variational theory for problems of this kind. We will always deal with functionals defined on a space of curves with values in a differentiable manifold, say $\mathcal{M}$, of the form $\mathcal{L}(z) = \int_0^1 L(\dot{z}(t), z(t), t)\,\mathrm{d}t$, (see (11)), where the *Lagrangian* function L is defined on the tangent space $T\mathcal{M}$ of the manifold, and is possibly depending also on time. For sake of simplicity, we have focused our attention on one-co-dimensional smooth constraints, that is when the constraint itself is described by a single equation $\phi = 0$, where ϕ is a smooth function on $T\mathcal{M}$ (possibly depending also on time).

The theory is described in Sections 2 to 4. We have tried to formulate hypotheses as general as possible on the Lagrangian function and on the constraint equation, in order to cover several situations, for instance the examples shown in Section 6. The result of existence and regularity of minimizers for the constrained functional is stated in Theorem 3.1, along with the Euler–Lagrange integro-differential equation solved by critical points. Since the constraint is not closed with respect to the weak convergence, we needed the well known *Palais–Smale condition* (see Definition 3.2), in order to pass from weak to strong convergence of a minimizing sequence. The proof that Palais–Smale condition is verified in our framework is quite delicate and is given in Proposition 3.5. Thanks to this condition, we can also obtain multiplicity results using the classical theory of Ljusternik–Schnirelman (see Theorem 3.14). Moreover, a local theory is developed, building an *exponential map* as is usually done in classical theory of ordinary differential equations (see Section 4).

In Sections 2–4 the Lagrangian function is assumed to have a growth in the velocity w given by $|w|^p$, where $p > 1$. In Section 5 a brief description of the case when $p = 1$ is given.

A short Appendix about a geometric description of the framework used ends the work.

There are some other examples strictly linked to our variational theory. First of all the relativistic brachistochrones with respect to the travel time (to arrive as *young* as possible – see Ref. [10]). The problem is reduced to the search of sub-Riemannian geodesics between a point and a curve, therefore it is straightly covered by our theory.

About the relativistic brachistochrones curves with respect to the arrival time (to arrive as *early* as possible – see Ref. [12]) the situation is different because the functional to study is a length functional plus another functional (the arrival time) which is invariant under reparameterizations, but does not have a Lagrangian density (in particular it is not an integral functional).

In [9] we have a similar situation, with an arrival time functional without Lagrangian density. Here we have not discussed this kind of situation, such as the problem in [11], where the constraint is described by $\phi(w,z,t)=c_z\in\mathbb{R}$ (and not by $\phi(w,z,t)=0$).

All these cases will be considered in the future. Our hope is to give a general theory, including all the examples here briefly sketched.

2. Framework setup and assumptions

Throughout the paper $\mathcal{M}$ will be an n-dimensional manifold, $n>1$, that we suppose C^∞, Hausdorff, and second-countable. We will denote by $T\mathcal{M}\to\mathcal{M}$ the tangent bundle of $\mathcal{M}$, and by $T_z\mathcal{M}$ the tangent space at a point $z\in\mathcal{M}$.

Coordinate systems on $\mathcal{M}$ and $T\mathcal{M}$ will be considered, whose notation will be:

$$z=\left(z^1,\dots,z^n\right)$$

for $\mathcal{M}$,

$$(w,z)=\left(w^1,\dots,w^n,z^1,\dots,z^n\right)$$

for $T\mathcal{M}$, and t for $\mathbb{R}$.

Let us consider a C^2 real *Lagrangian* function L defined on $T\mathcal{M}\times\mathbb{R}$, and a C^1 *constraint* function ϕ,

$$L:T\mathcal{M}\times\mathbb{R}\to\mathbb{R},\qquad \phi:T\mathcal{M}\times\mathbb{R}\to\mathbb{R}.$$

We assume that

$$L(w,z,t)\geqslant 0,\tag{1}$$

$$\phi(0,z,t)=0,\quad \forall(z,t)\in\mathcal{M}\times\mathbb{R},\tag{2}$$

and

$$\frac{\partial L}{\partial w}(0,z,t)=0,\quad \forall(z,t)\in\mathcal{M}\times\mathbb{R}.\tag{3}$$

It will be as well convenient to introduce the space:

$$\mathcal{S}_{(w,z,t)}=\ker\frac{\partial\phi}{\partial w}(w,z,t)=\left\{\xi\in T_z\mathcal{M}\;\middle|\;\frac{\partial\phi}{\partial w}(w,z,t)[\xi]=0\right\}$$

for every $z\in\mathcal{M}$, $w\in T_z\mathcal{M}$, and $t\in\mathbb{R}$. We will require that ϕ is an *admissible* constraint (see [16]). This amounts to say that

$$\dim\mathcal{S}_{(w,z,t)}=n-1,\quad \forall(w,z,t)\in T\mathcal{M}\times\mathbb{R}.\tag{4}$$

In order to define the space of curves we will search for critical points in, we must fix a point $Q \in \mathcal{M}$ and a curve $\gamma : \mathbb{R} \to \mathcal{M}$ transversal to $\mathcal{S}$, i.e.,

$$\frac{\partial\phi}{\partial w}\big(w, \gamma(t), t\big)\big[\dot{\gamma}(t)\big] \neq 0, \quad \forall w \in T_{\gamma(t)}\mathcal{M}.$$

For this aim we will suppose the existence of a smooth vector field $Y : \mathcal{M} \to T\mathcal{M}$, not null everywhere, such that

$$\frac{\partial\phi}{\partial w}(w, z, t)\big[Y(z)\big] = 1, \quad \forall (w, z, t) \in \phi^{-1}(0). \tag{5}$$

Hereafter we will actually suppose that γ is an integral curve of Y.

The Lagrangian function L does not need to be *regular*—in the sense of a mechanical system—anyway we will suppose that, $\forall (w, z, t) \in \phi^{-1}(0) \subset T\mathcal{M} \times \mathbb{R}$,

$$G(w, z, t) = \frac{\partial^2 L}{\partial w^2}(w, z, t) : T_z\mathcal{M} \times T_z\mathcal{M} \to \mathbb{R}$$

is a bilinear application on $T_z\mathcal{M}$ such that

$$G(w, z, t)[\xi, \xi] > 0 \quad \text{and} \quad G(w, z, t)\big[\xi, Y(z)\big] = 0, \quad \forall \xi \in \mathcal{S}_{(w,z,t)} \setminus \{0\}. \tag{6}$$

Remark 2.1. The assumptions made so far, and Eq. (7) that we will state in a while, are actually thought in terms of local coordinates, because of the derivatives with respect to w. It can be shown that they can be stated in terms of intrinsic objects. See Appendix A for further details.

We will make a similar assumption on the second derivatives of the constraint equation ϕ, as specified in the following:

Assumption 2.2. We will require that, $\forall (w, z, t) \in \phi^{-1}(0) \subset T\mathcal{M} \times \mathbb{R}$,

$$\big(\mathrm{d}L(w, z, t)\big[Y(z)\big]\big) \cdot \frac{\partial^2\phi}{\partial w^2}(w, z, t)[\xi, \xi] \leqslant 0, \quad \forall \xi \in \mathcal{S}_{(w,z,t)}, \tag{7}$$

where

$$\mathrm{d}L(w, z, t)\big[Y(z)\big] = \frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial w}[Y']$$

is intrinsically defined in Appendix A.

Example 2.3. Let us consider the case of sub-Riemannian geodesics (see Section 6.1). Let us take the energy functional (87) of Section 6.1, so that the Lagrangian L is given by:

$$L(w, z) = \langle w, w\rangle.$$

Take into account the vector field Y defined in (84) of Section 6.1. Then we have:

$$\mathrm{d}L(w,z,t)\big[Y(z)\big] = \langle \nabla_w Y, w\rangle,$$

where ∇ is the covariant derivative with respect to $\langle\cdot,\cdot\rangle$.

We also remark that, in this case, the right-hand side in last equation above have not a specific sign, but Assumption 2.2 is satisfied, since the constraint equation is linear. Observe that the other assumptions made before also hold in this particular case, as it can be easily seen.

In addition we will suppose that there exists a number $p > 1$, and some functions α_i, δ_i of class C^0, defined on $\mathcal{M}\times\mathbb{R}$ and strictly positive such that

$$\forall (w,z,t) \in \phi^{-1}(0) \subset T\mathcal{M}\times\mathbb{R},$$

it is:

$$\frac{\partial L}{\partial w}(w,z,t)[w] \geqslant \alpha_1(z,t)|w|^p - \delta_1(z,t), \tag{8a}$$

$$L(w,z,t) \geqslant \alpha_2(z,t)|w|^p - \delta_2(z,t), \tag{8b}$$

where $|\cdot|$ is a (given) norm on $T_z\mathcal{M}$ depending continuously on $z\in\mathcal{M}$. Note that (8b) comes from (8a) in some particular cases (see Remark 2.5 below). We finally require the following asymptotical estimates on the derivatives of L and ϕ:

$$\left|\frac{\partial L}{\partial w}(w,z,t)\right| \leqslant a_1(z,t)|w|^{p-1} + b_1(z,t), \tag{9a}$$

$$\left|\frac{\partial L}{\partial z}(w,z,t)\right| \leqslant a_2(z,t)|w|^{p} + b_2(z,t), \tag{9b}$$

and

$$\left|\frac{\partial \phi}{\partial w}(w,z,t)\right| \leqslant b_3(z,t), \tag{10a}$$

$$\left|\frac{\partial \phi}{\partial z}(w,z,t)\right| \leqslant a_4(z,t)|w| + b_4(z,t), \tag{10b}$$

$\forall (w,z,t)\in T\mathcal{M}\times\mathbb{R}$, for some C^0 functions a_i, b_i defined on $\mathcal{M}\times\mathbb{R}$.

Remark 2.4. It can be proved that assumptions (8a)–(10b) are independent from the norm used. Moreover, these hypotheses, as the previous ones, are satisfied by the examples given in Section 6.

Remark 2.5. In case ϕ is a homogeneous function of degree 1 in the variable w, conditions (8a) and (8b) can be straightly derived from the following condition on $G(w,z,t)$, $\forall (w,z,t) \in \phi^{-1}(0)$:

$$G(w,z,t)[w,w] \geqslant \alpha(z,t)|w|^p,$$

for some strictly positive function $\alpha(z,t)$. Indeed, we observe that, in this particular case, if $\phi(w,z,t) = 0$,

$$\phi(\sigma \cdot w, z, t) = \sigma \cdot \phi(w,z,t) = 0,$$

and

$$\frac{\partial \phi}{\partial w}(w,z,t)[\sigma \cdot w] = \sigma \cdot \frac{\partial \phi}{\partial w}(w,z,t)[w] = \sigma \cdot \phi(w,z,t) = 0,$$

where $\sigma \in \mathbb{R}$. Therefore, if $(w,z,t) \in \phi^{-1}(0)$ then $\{\sigma w \colon \sigma \in \mathbb{R}\} \subset \mathcal{S}_{(w,z,t)}$.

From this we have, also using (3),

$$\begin{aligned}
\frac{\partial L}{\partial w}(w,z,t)[w] &= \int_0^1 \frac{\partial^2 L}{\partial w^2}(\sigma w, z, t)[w,w]\,\mathrm{d}\sigma \\
&= \int_0^1 \frac{1}{\sigma^2}\frac{\partial^2 L}{\partial w^2}(\sigma w, z, t)[\sigma w, \sigma w]\,\mathrm{d}\sigma \\
&\geqslant \alpha(z,t)|w|^p \int_0^1 \sigma^{p-1}\,\mathrm{d}\sigma = \frac{\alpha(z,t)}{p-1}|w|^p
\end{aligned}$$

and analogously (8b) can be derived. Note that this method cannot be used in general because we require $G(w,z,t)[w,w] \geqslant \alpha(z,t)|w|^p$ only on the constraint, and not in the whole $T\mathcal{M} \times \mathbb{R}$.

As pointed out in the Introduction, we will deal with functionals defined on a space of curves with values on $\mathcal{M}$. The Sobolev space $H^{1,p}([0,1],\mathcal{M})$ naturally arises as the main workspace. It can be defined as the set of all absolutely continuous curves $z:[0,1] \to \mathcal{M}$ such that, for every local chart (V,φ) on $\mathcal{M}$ and for every closed sub-interval $[a,b] \subset [0,1]$ such that $z([a,b]) \subset V$, $\varphi \circ z \in H^{1,p}([a,b],\mathbb{R}^n)$. The space $H^{1,p}([0,1],\mathcal{M})$ has an infinite dimensional manifold structure (see [15]), modeled on $H^{1,p}([0,1],\mathbb{R}^n)$. We stress the point that, although this definition is given in terms of local charts, $H^{1,p}([0,1],\mathcal{M})$ is independent from the chosen coordinate system.

In view of this, a functional $\mathcal{L}$ is then induced by L on $H^{1,p}([0,1],\mathcal{M})$ in a natural way:

$$\mathcal{L}(z) = \int_0^1 L\big(\dot z(t), z(t), t\big)\,\mathrm{d}t. \tag{11}$$

We can now define the space of curves $\Omega_{Q,\gamma}$,

$$\Omega_{Q,\gamma} = \big\{z \in H^{1,p}\big([0,1],\mathcal{M}\big) \bigm| z(0) = Q,\ z(1) \in \gamma(\mathbb{R})\big\}, \tag{12}$$

and its subspace $\Omega_{Q,\gamma}(\phi)$,

$$\Omega_{Q,\gamma}(\phi) = \big\{z \in \Omega_{Q,\gamma} \bigm| \phi\big(\dot z(t), z(t), t\big) \equiv 0 \text{ a.e. in } [0,1]\big\}, \tag{13}$$

that it is supposed to be non empty.

We require a *pseudo-coercivity* condition of $\mathcal{L}$ with respect to ϕ in the following way:

Assumption 2.6. $\forall c \in \mathbb{R}$ there exists a compact set $\mathcal{K}(c) \subset \mathcal{M}$ such that, $\forall z \in \Omega_{Q,\gamma}(\phi)$,

$$\mathcal{L}(z) \leqslant c \quad \Longrightarrow \quad z\big([0,1]\big) \subset \mathcal{K}(c). \tag{14}$$

Remark 2.7. The pseudo-coercivity assumption – satisfied by the examples of Section 6 – is an intrinsic way for requiring completeness. It is satisfied, for instance, if α_2 in (8b) is bounded away from zero, δ_2 is bounded and $(M, |\cdot|)$ is complete. Indeed, if there exists a constant $\tilde\alpha > 0$ with $\alpha_2(z,t) > \tilde\alpha$ and $\delta_2 \leqslant \tilde\alpha$, it is, in local coordinates, for each $z \in \Omega_{Q,\gamma}(\phi)$ such that $\mathcal{L}(z) \leqslant c$,

$$d\big(z(t), Q\big) \leqslant \int_0^t \big|\dot z(s)\big|\,\mathrm{d}s \leqslant \int_0^1 \big|\dot z(s)\big|\,\mathrm{d}s,$$

where $d(P,Q)$ is the distance in $\mathcal{M}$ defined by the $\inf \int |\dot z|$ calculated among all the paths defined in $[0,1]$ with values in $\mathcal{M}$ linking two points P, Q of $\mathcal{M}$.

Using (8b) we have:

$$\begin{aligned}\big(d(z(t),Q)\big)^p &\leqslant \left(\int_0^1 \big|\dot z(s)\big|\,\mathrm{d}s\right)^p \leqslant \int_0^1 \big|\dot z(s)\big|^p\,\mathrm{d}s \leqslant \int_0^1 \frac{L(\dot z(s), z(s), s) + \delta_2(z(s), s)}{\alpha_2(z(s), s)}\,\mathrm{d}s \\ &\leqslant \frac{1}{\tilde\alpha}\left(\int_0^1 L\,\mathrm{d}s + \tilde\delta\right) \leqslant \frac{c + \tilde\delta}{\tilde\alpha},\end{aligned}$$

and if $(M, |\cdot|)$ is complete the closed balls are compact.

It is well known that $\Omega_{Q,\gamma}$ is a Banach manifold (see [15]); its tangent space $T_z\Omega_{Q,\gamma}$, $\forall z \in \Omega_{Q,\gamma}$, can be defined as follows:

$$T_z\Omega_{Q,\gamma} = \big\{\xi \in H^{1,p}\big([0,1], T\mathcal{M}\big) \bigm| \xi \text{ vector field along } z, \\ \xi(0) = 0,\ \xi(1) \parallel \dot\gamma(t_z)\big\}, \tag{15}$$

where $t_z \in \mathbb{R}$ is the real value mapped by γ into the end point of z, that is $z(1) = \gamma(t_z)$. Moreover $T\Omega_{Q,\gamma}$ is endowed with the Finslerian structure induced by the $H^{1,p}$ norm

$$\|\xi\| = \left(\int_0^1 \big|\dot\xi(t)\big|^p \,\mathrm{d}t\right)^{1/p}. \tag{16}$$

Under the above hypothesis, we will prove results on existence, regularity and multiplicity of critical points for the functional $\mathcal{L}$ in the set $\Omega_{Q,\gamma}(\phi)$.

The set $\Omega_{Q,\gamma}(\phi)$ is a Banach submanifold of $\Omega_{Q,\gamma}$, as shown in the following:

Proposition 2.8. *$\Omega_{Q,\gamma}(\phi)$ is a C^1 Banach submanifold of $\Omega_{Q,\gamma}$, and its tangent space, $\forall z \in \Omega_{Q,\gamma}(\phi)$, is given by*

$$T_z\Omega_{Q,\gamma}(\phi) = \big\{\xi \in T_z\Omega_{Q,\gamma} \bigm| \mathrm{d}\phi(z)[\xi] \equiv 0 \ a.e.\big\}. \tag{17}$$

Remark 2.9. Here $\mathrm{d}\phi(z)[\xi]$ denotes the Gateaux derivative of ϕ along the direction ξ; in local coordinates it reads:

$$\mathrm{d}\phi(z)[\xi] = \frac{\partial\phi}{\partial z}\big(\dot z(t), z(t), t\big)[\xi(t)] + \frac{\partial\phi}{\partial w}\big(\dot z(t), z(t), t\big)\big[\dot\xi(t)\big].$$

Let us observe that $\mathrm{d}\phi(z)[\xi]$ makes sense, and is in $L^p([0,1],\mathbb{R})$, since ξ, z are continue, $\dot\xi$, $\dot z$ are in L^p, and (10a), (10b) holds.

Proof. Let us consider the application:

$$\begin{aligned} &F : \Omega_{Q,\gamma} \to L^p\big([0,1], \mathcal{M}\big), \\ &F(z)(t) = \phi\big(\dot z(t), z(t), t\big). \end{aligned} \tag{18}$$

Since $\Omega_{Q,\gamma}(\phi) = F^{-1}(0)$, we must show [13] that

(1) $F \in C^1$, $\forall z \in \Omega_{Q,\gamma}$,
(2) $\mathrm{d}F(z)$ is surjective,
(3) its kernel splits.

(1) We first actually show that F is Gateaux differentiable: considering a local coordinate system, let us fix $z \in \Omega_{Q,\gamma}$ and $\xi \in T_z\Omega_{Q,\gamma}$, and prove that

$$\lim_{\sigma\to 0}\left\|\frac{\phi(\dot z+\sigma\dot\xi, z+\sigma\xi, t)-\phi(\dot z, z, t)}{\sigma}-\frac{\partial\phi}{\partial z}(\dot z, z, t)[\xi]-\frac{\partial\phi}{\partial w}(\dot z, z, t)[\dot\xi]\right\|_{L^p}=0. \tag{19}$$

Indeed there exists, $\forall t\in[0,1]$, two values $h,k\in[0,1]$ such that

$$\begin{aligned}
&\phi(\dot z+\sigma\dot\xi, z+\sigma\xi, t)-\phi(\dot z, z, t)\\
&\quad=\phi(\dot z+\sigma\dot\xi, z+\sigma\xi, t)+\phi(\dot z+\sigma\dot\xi, z, t)-\phi(\dot z+\sigma\dot\xi, z, t)-\phi(\dot z, z, t)\\
&\quad=\sigma\frac{\partial\phi}{\partial z}(\dot z+\sigma\dot\xi, z+k\sigma\xi, t)[\xi]+\sigma\frac{\partial\phi}{\partial w}(\dot z+h\sigma\dot\xi, z, t)[\dot\xi],
\end{aligned}$$

and then the argument of the limit in (19) becomes:

$$\begin{aligned}
&\left\|\left[\frac{\partial\phi}{\partial w}(\dot z+h\sigma\dot\xi, z, t)-\frac{\partial\phi}{\partial w}(\dot z, z, t)\right][\dot\xi]\right.\\
&\qquad\left.+\left[\frac{\partial\phi}{\partial z}(\dot z+\sigma\dot\xi, z+k\sigma\xi, t)-\frac{\partial\phi}{\partial z}(\dot z, z, t)\right][\xi]\right\|_{L^p}.
\end{aligned} \tag{20}$$

Since $\dot z$ and $\dot\xi$ are fixed and in L^p, and ϕ is C^1, (19) holds by Lebesgue dominate convergence theorem. Moreover, we must now show that the application

$$z\mapsto \mathrm{d}F(z)[\,\cdot\,]=\frac{\partial\phi}{\partial z}\big(\dot z(t), z(t), t\big)\big[(\cdot)\big]+\frac{\partial\phi}{\partial w}\big(\dot z(t), z(t), t\big)\big[(\cdot)'\big] \tag{21}$$

is a continuous map. For this aim it is sufficient to consider a strongly converging sequence $\{z_n\}$ in $\Omega_{Q,\gamma}$ to a curve $z\in\Omega_{Q,\gamma}$.

Arguing in a similar way as before, using (10a) and (10b), one obtain that

$$\lim_{n\to\infty}\sup_{\|\xi\|_{H^{1,p}}\leqslant 1}\big\|\big(\mathrm{d}F(z_n)-\mathrm{d}F(z)\big)[\xi]\big\|_{L^p}=0. \tag{22}$$

(2) Let us consider $z\in F^{-1}(0)$, and let $h\in L^p([0,1],\mathcal{M})$; to prove surjectivity we look for $\psi_h\in H^{1,p}([0,1],\mathbb{R})$ such that $h=\mathrm{d}F(z)[\psi_h Y(z)]$, and $\psi_h(0)=0$. Since, in local coordinates,

$$\begin{aligned}
\mathrm{d}F(z)\big[\psi_h Y(z)\big] = {}&\frac{\partial\phi}{\partial z}\big(\dot z(t), z(t), t\big)\big[\psi_h Y(z)\big]\\
&+\frac{\partial\phi}{\partial w}\big(\dot z(t), z(t), t\big)\big[\psi_h Y'(z)+\dot\psi_h Y(z)\big],
\end{aligned}$$

using (5) we obtain the following ODE for ψ_h:

$$\begin{cases} h=\dot\psi_h+\psi_h\,\mathrm{d}F(z)\big[Y(z)\big],\\ \psi_h(0)=0,\end{cases} \tag{23}$$

where $\mathrm{d}F(z)$ is given by (21). The explicit solution of the ODE is:

$$\psi_h(t) = \int_0^t h(t)\,\mathrm{e}^{\int_t^s \mathrm{d}F(z(r))[Y(z(r))]\,\mathrm{d}r}\,\mathrm{d}s. \tag{24}$$

It is easily seen that ψ_h is $H^{1,p}$, and $\eta(t) = \psi_h(t)Y(z(t)) \in T_z\Omega_{Q,\gamma}$.

(3) We argue in a similar way as point (2) to prove that $\ker \mathrm{d}F(z)$ splits $\forall z \in F^{-1}(0)$: fixed $V \in T_z\Omega_{Q,\gamma}$, we look for $\psi_V \in H^{1,p}([0,1],\mathbb{R})$ such that

$$\begin{cases} 0 = \mathrm{d}F(z)[V - \psi_V Y] = \mathrm{d}F[V] - \psi_V\,\mathrm{d}F\big[Y(z)\big] - \dot{\psi}_V, \\ \psi_V(0) = 0, \end{cases} \tag{25}$$

and then V can be written as the sum of two components

$$V = (V - \psi_V Y) + \psi_V Y,$$

such that $(V - \psi_V Y) \in \ker \mathrm{d}F(z)$. Here

$$\psi_V(t) = \int_0^t \mathrm{d}F\big(z(s)\big)\big[V(s)\big]\,\mathrm{e}^{\int_t^s \mathrm{d}F(z(r))[Y(z(r))]\,\mathrm{d}r}\,\mathrm{d}s. \qquad \square \tag{26}$$

Thanks to the above proposition, we can give the following:

Definition 2.10. $z \in \Omega_{Q,\gamma}(\phi)$ is a *critical point* for the functional $\mathcal{L}$ on $\Omega_{Q,\gamma}(\phi)$ if

$$\mathrm{d}\mathcal{L}(z)[\xi] = 0 \tag{27}$$

for every *admissible variation* ξ, that is $\xi \in T_z\Omega_{Q,\gamma}(\phi)$.

Remark 2.11. When studying a dynamical system described by a Lagrangian function and a set of constraints, two different approaches can be followed, depending on the choice made of the admissible variation.

Without getting into details—that can anyway be found, for instance, in [2]—the classical approach takes into account the Principle of Virtual Work, to write a system of equation in which the reaction force of the constraints does not occur. In this setup the *virtual displacements*, namely the admissible variations, are given by the space $\mathcal{S}$ defined at the beginning of this section.

The second way to handle the problem, that is the one we use, as we can understand from Definition 2.10, is a Lagrangian counterpart of Hamilton's principle. In short, it takes the tangent vectors of the constraint manifold as virtual displacements. In mechanics this approach is referred to as dynamics of *variational axiomatic kind* (vakonomic dynamics).

This two methods leads to different Euler–Lagrange equation in case of nonholonomic systems.

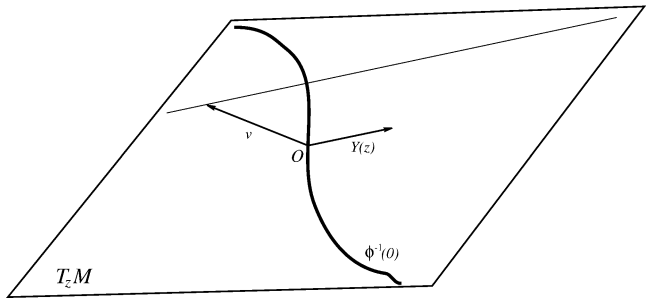

Fig. 1. The tangent space $T_z\mathcal{M}$ at a point $z \in \mathcal{M}$ (Lemma 2.13).

Remark 2.12. In this framework a crucial role is played by the vector field Y, that someway links the Lagrangian function and the constraint equation, in the sense described by Eqs. (5)–(7), and by the following result, that we will prove in Proposition 2.15:

$$\frac{\partial L}{\partial w}(w,z,t)\big[Y(z)\big]=0,\quad \forall(w,z,t)\in\phi^{-1}(0)\subset T\mathcal{M}\times\mathbb{R}.$$

In order to prove the above relation, the following Lemma is needed:

Lemma 2.13. *For each* (v,z,t) *in* $T\mathcal{M}\times\mathbb{R}$ *there is a unique real number* $\mu(v,z,t)$ *such that* (*see Fig.* 1)

$$\phi(v+\mu(v,z,t)Y(z),z,t)=0.$$

Proof. Fixed (v,z,t), it suffices to consider the function in $\mathbb{R}^2$:

$$H(\tau,\mu)=\phi\big(\tau v+\mu Y(z),z,t\big).$$

Since $H(0,0)=0$ by (2), and $\partial H/\partial\mu(0,0)=\partial\phi/\partial w(0,z,t)[Y(z)]=1$ by (5), from the Implicit Function Theorem there exists $\theta_0>0$ such that $\mu=\mu(\tau)$ (for $\tau\in(-\theta_0,\theta_0)$) is the unique solution of $H(\tau,\mu)=0$. But

$$\frac{\mathrm{d}\mu}{\mathrm{d}\tau}=-\frac{\partial\phi}{\partial w}\big(\tau v+\mu Y(z)\big)[v]$$

and $\frac{\partial\phi}{\partial w}(w,z,t)$ is bounded, fixed z and t, by (10b). Then there exists

$$\mu_0=\lim_{\tau\to\theta_0}\mu(\tau)<+\infty,$$

and by continuity of ϕ, $H(\theta_0,\mu_0)=0$. Then μ can be extended to the whole interval $[0,1]$, and $\mu(1)$ is the required value $\mu(v,z,t)$. Let us now show that it is unique. Let $\mu_1\neq\mu_2$ such that $\phi(v-\mu_iY,z,t)=0$, $i=1,2$. Then we can consider $H(\tau,\mu)$ as before, and in this case we have $H(1, \mu_i(1)) = 0$, $i = 1, 2$. This means that there exists two functions $\mu_i(\tau)$ solutions of the ODE:

$$\begin{cases} \mu_i'(\tau) = -\frac{\partial\phi}{\partial w}\big(\tau v + \mu_i(\tau)Y(z), z, t\big)[v], \\ \mu_i(1) = \mu_i. \end{cases}$$

These two functions can be extended until $\tau = 0$, where they coincide since

$$\phi(\mu Y, z, t) = 0 \quad \Longleftrightarrow \quad \mu = 0;$$

therefore they coincide also for $\tau = 1$, that implies $\mu_1 = \mu_2$.

By smoothness of ϕ it also follows that μ is smooth. □

Remark 2.14. In particular the argument in the proof the above lemma implies that $\{w \in T_z\mathcal{M} \mid \phi(w, z, t) = 0\}$ is contractible for each (z, t).

Proposition 2.15. *For each* $(w, z, t) \in \phi^{-1}(0) \subset T\mathcal{M} \times \mathbb{R}$, *it is*

$$\frac{\partial L}{\partial w}(w, z, t)\big[Y(z)\big] = 0. \tag{28}$$

Proof. Let us fix $(w, z, t) \in \phi^{-1}(0)$; then we can write, using (3),

$$\frac{\partial L}{\partial w}(w, z, t)\big[Y(z)\big] = \int_0^1 \frac{\partial^2 L}{\partial w^2}\big(\sigma\, w + \mu(\sigma)Y(z), z, t\big)\big[w + \dot{\mu}(\sigma)Y(z), Y(z)\big]\,\mathrm{d}\sigma, \tag{29}$$

where $\sigma \in [0, 1]$, and $\mu(\sigma)$ is chosen such that $\phi(\sigma\, w + \mu(\sigma)Y(z), z, t) = 0$, by the above lemma. Then, if we denote

$$K(\sigma) = \phi\big(\sigma\, w + \mu(\sigma)Y(z), z, t\big),$$

it is

$$0 = K'(\sigma) = \frac{\partial\phi}{\partial w}\big(\sigma\, w + \mu(\sigma)Y(z), z, t\big)\big[w + \dot{\mu}(\sigma)Y(z)\big],$$

that is, $w + \dot{\mu}(\sigma)Y(z) \in \mathcal{S}_{(\sigma\, w+\mu(\sigma)Y(z),z,t)}$. Using (6) we see that the function into the integral of (29) in null, proving (28). □

3. Existence, regularity and multiplicity of critical points

We are now ready to start the proof of one of the main result of the section, that is the following:

Theorem 3.1. *Under the hypotheses made in Section* 2 (*see* (1)–(10b), *and pseudo-coercivity Assumption* 2.6), *there exists a minimizer z for the functional $\mathcal{L}$ in the set $\Omega_{Q,\gamma}(\phi)$.*

Moreover, if z is a critical point of $\mathcal{L}$ in $\Omega_{Q,\gamma}(\phi)$, then $z \in C^2([0,1],\mathcal{M})$, and satisfies the following Euler–Lagrange equation:

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial w} - \lambda\frac{\partial\phi}{\partial w}\right) - \left(\frac{\partial L}{\partial z} - \lambda\frac{\partial\phi}{\partial z}\right)\right]\big(\dot z(t), z(t), t\big) = 0, \tag{30}$$

where $\lambda(t) \in C^1([0,1],\mathbb{R})$ is the following Lagrange multiplier:

$$\lambda(t) = \int_t^1 \left(\frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial w}[Y']\right) \mathrm{e}^{\int_s^t (\frac{\partial\phi}{\partial z}[Y] + \frac{\partial\phi}{\partial w}[Y'])\,\mathrm{d}r}\,\mathrm{d}s. \tag{31}$$

We remind that

$$\mathrm{d}L[Y] \equiv \left(\frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial w}[Y']\right) \quad \text{and} \quad \mathrm{d}\phi[Y] \equiv \left(\frac{\partial\phi}{\partial z}[Y] + \frac{\partial\phi}{\partial w}[Y']\right)$$

are intrinsically defined (see Appendix A).

Let us start proving the existence of a minimizer for the system. For this aim we will show that the functional satisfy a good compactness property: the *Palais–Smale condition*.

Definition 3.2. Given a C^1 functional $F: X \to \mathbb{R}$ on a Banach manifold (X,h), then F satisfies the *Palais–Smale condition* at level $c \in \mathbb{R}$ if every sequence $\{z_n\}_{n\in\mathbb{N}} \subset X$ such that

$$\lim_{n\to\infty} F(z_n) = c, \qquad \lim_{n\to\infty} \big\|\mathrm{d}F(z_n)\big\| = 0, \tag{32}$$

(where $\|\cdot\|$ denotes the norm in the dual space of $T_{x_n}X$), has a subsequence converging in X. The sequence $\{z_n\}$ is called a *Palais–Smale sequence in X for E at level c*.

Lemma 3.3. *Let $\{z_n\}$ be a sequence in $\Omega_{Q,\gamma}(\phi)$ such that there exists $c \in \mathbb{R}$ with $\mathcal{L}(z_n) \leqslant c$ for each $n \in \mathbb{N}$. Then $\{z_n\}$ is uniformly bounded in $H^{1,p}$, and included in a compact subset of $\mathcal{M}$.*

Proof. From pseudo-coercivity it follows that there exists a compact set $\mathcal{K} \subset \mathcal{M}$ such that $z_n([0,1]) \subset \mathcal{K}$, $\forall n$. The result immediately follows from (8b), recalling that α_2 in this equation is strictly positive. □

Lemma 3.4. *Let $\{z_n\}$ be a Cauchy sequence, with respect to the Finslerian structure* (16), *in*

$$\mathcal{L}^c = \big\{z \in \Omega_{Q,\gamma}(\phi) \,\big|\, \mathcal{L}(z) \leqslant c\big\}.$$

Then z_n converges in $H^{1,p}$ to a curve $z \in \Omega_{Q,\gamma}(\phi)$.

Proof. Let z_n be a Cauchy sequence with respect to (16) and such that $\mathcal{L}(z) \leqslant c$. By Lemma 3.3 there exists a subsequence z_{n_k} uniformly converging to $z \in \Omega_{Q,\gamma}$.

We can consider a finite number of local charts $(\mathcal{U}_i, \psi_i)$ that covers $z([0,1])$. Fixed i, let $[\alpha_i, \beta_i] \subset [0,1]$ such that $z_{n_k} \mid [\alpha_i, \beta_i] \subset \mathcal{U}_i$, for each $n \in \mathbb{N}$, so we can suppose that $\{z_{n_k} \mid [\alpha_i, \beta_i]\}$ is a Cauchy sequence in the complete space $H^{1,p}([\alpha_i, \beta_i], \mathbb{R}^n)$, and then it converges in $H^{1,p}$ to a function y. The curve y satisfies the constraints because $\dot{z}_{n_k}$ has a pointwise converging subsequence. On every local chart the convergence is also uniform, and then $y = z$. Then z_{n_k} converges to z in $H^{1,p}$ and $z \in \Omega_{Q,\gamma}(\phi)$. Since z_n is a Cauchy sequence, it follows that z_n converges to z in $H^{1,p}$. □

Proposition 3.5. *The functional $\mathcal{L} : \Omega_{Q,\gamma}(\phi) \to \mathbb{R}$ satisfies the Palais–Smale condition at every level $c \in \mathbb{R}$.*

Proof. Let $\{z_k\} \in \Omega_{Q,\gamma}(\phi)$ be a Palais–Smale sequence at level c for some $c > \inf_{\Omega_{Q,\gamma}(\phi)} \mathcal{L}$.

We observe that, from pseudo-coercivity, we can suppose, up to subsequences, that z_k is uniformly convergent to $z \in H^{1,p}([0,1], \mathcal{M})$. Moreover $\dot{z}_k$ is weakly convergent to $\dot{z}$ in L^p. We have just to prove, using that z_k is a Palais–Smale sequence, that $\dot{z}_k \to \dot{z}$ strongly in L^p.

Let us consider $V \in T_{z_k}\Omega_{Q,\gamma}$; we know from Proposition 2.8 that every admissible variation can be expressed in the form $(V - \psi_V Y)$, where ψ_V is given by (26):

$$\psi_V(t) = \int_0^t \left(\frac{\partial \phi}{\partial z}(\dot{z}_k, z_k, t)\big[V(s)\big] + \frac{\partial \phi}{\partial w}(\dot{z}_k, z_k, t)\big[V'(s)\big] \right) \times \mathrm{e}^{\int_t^s (\frac{\partial \phi}{\partial z}(\dot{z}_k, z_k, t)[Y(s)] + \frac{\partial \phi}{\partial w}(\dot{z}_k, z_k, t)[Y'(s)])\,\mathrm{d}r}\,\mathrm{d}s.$$

Let us now work on the quantity $\mathrm{d}\mathcal{L}(z_k)[V - \psi_V Y]$:

$$\begin{aligned}
\mathrm{d}\mathcal{L}(z_k)[V - \psi_V Y] &= \int_0^1 \frac{\partial L}{\partial z}[V] + \frac{\partial L}{\partial w}[V'] - \psi \frac{\partial L}{\partial z}[Y] - \dot{\psi} \frac{\partial L}{\partial w}[Y] - \psi \frac{\partial L}{\partial w}[Y']\,\mathrm{d}t \\
&= \int_0^1 \frac{\partial L}{\partial z}[V] + \frac{\partial L}{\partial w}[V'] - \psi \left(\frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial w}[Y'] \right) \mathrm{d}t. \qquad (33)
\end{aligned}$$

Note that here we have dropped the subscript $_V$ from ψ to lighten the notation, such as the argument $(\dot{z}_k(t), z_k(t), t)$

Substituting (26) in the last addendum of (33), and applying Fubini's theorem, we get:

$$-\int_0^1 \psi \left(\frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial \dot{z}}[Y'] \right) \mathrm{d}t$$

$$= -\int_0^1 \left(\frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial w}[Y']\right) \int_0^t \left(\frac{\partial \phi}{\partial z}[V] + \frac{\partial \phi}{\partial w}[V']\right) \mathrm{e}^{\int_t^s (\frac{\partial \phi}{\partial z}[Y] + \frac{\partial \phi}{\partial w}[Y'])\,\mathrm{d}r}\, \mathrm{d}s\, \mathrm{d}t$$

$$= -\int_0^1 \left(\frac{\partial \phi}{\partial z}[V] + \frac{\partial \phi}{\partial w}[V']\right) \int_s^1 \left(\frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial w}[Y']\right) \mathrm{e}^{\int_t^s (\frac{\partial \phi}{\partial z}[Y] + \frac{\partial \phi}{\partial w}[Y'])\,\mathrm{d}r}\, \mathrm{d}t\, \mathrm{d}s$$

$$= -\int_0^1 \lambda_k(t) \left(\frac{\partial \phi}{\partial z}[V] + \frac{\partial \phi}{\partial w}[V']\right) \mathrm{d}t, \tag{34}$$

where

$$\lambda_k(t) = \int_t^1 \left(\frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial w}[Y']\right) \mathrm{e}^{\int_s^t (\frac{\partial \phi}{\partial z}[V] + \frac{\partial \phi}{\partial w}[V'])\,\mathrm{d}r}\, \mathrm{d}s. \tag{35}$$

Using (34) and (28) into (33) we obtain:

$$\mathrm{d}\mathcal{L}(z_k)[V - \psi_V Y] = \int_0^1 \left(\frac{\partial L}{\partial z} - \lambda_k \frac{\partial \phi}{\partial z}\right)[V] + \left(\frac{\partial L}{\partial w} - \lambda_k \frac{\partial \phi}{\partial w}\right)[V']\, \mathrm{d}t. \tag{36}$$

The terms $\partial L/\partial z$ and $\partial \phi/\partial z$, such as V', are not intrinsically meaningful. Nevertheless, $z([0,1])$ can be covered with a finite number of local charts, where we can consider each term in (36) separately. Then, since z_k uniformly converges to z, for sake of simplicity we make our computations assuming we are in a single chart.

Recall now that z_k is a Palais–Smale sequence, and let p^* denote the conjugate exponent of p $(1/p + 1/p^* = 1)$. Therefore there exists a sequence $\{a_k\}$ converging to 0 in L^{p^*} such that, for any V satisfying $V(0) = V(1) = 0$,

$$\int_0^1 \left(\frac{\partial L}{\partial z} - \lambda_k \frac{\partial \phi}{\partial z}\right)[V] + \left(\frac{\partial L}{\partial w} - \lambda_k \frac{\partial \phi}{\partial w}\right)[V']\, \mathrm{d}t = \int_0^1 a_k\big[(V - \psi_V Y)'\big]\, \mathrm{d}t. \tag{37}$$

Integrating by parts the first term in the left-hand side of (36) we get:

$$\mathrm{d}\mathcal{L}(z_k)[V - \psi_V Y] = \int_0^1 \left[\left(\frac{\partial L}{\partial w} - \lambda_k \frac{\partial \phi}{\partial w}\right) - \int_0^t \left(\frac{\partial L}{\partial z} - \lambda_k \frac{\partial \phi}{\partial z}\right) \mathrm{d}s\right][V']\, \mathrm{d}t. \tag{38}$$

Note that there is not boundary contribute from the integration, since $V(0) = V(1) = 0$.

Let us now take into account the right-hand side of (37). Using again (26) we have, for any $V \in T_{z_k}\Omega_{Q,\gamma}$,

$$\int_0^1 a_k\big[(V-\psi Y)'\big]\,\mathrm{d}t = \int_0^1 a_k[V'] - \dot\psi a_k[Y] - \psi a_k[Y']\,\mathrm{d}t. \tag{39}$$

With some tedious calculus, using again Fubini's theorem, we get:

$$-\int_0^1 \dot\psi a_k[Y]\,\mathrm{d}t = \int_0^1 g_k(t)\left(\frac{\partial\phi}{\partial z}[V] + \frac{\partial\phi}{\partial w}[V']\right)\mathrm{d}t, \tag{40a}$$

where

$$g_k(t) = \int_t^1 a_k[Y]\left(\frac{\partial\phi}{\partial z}[Y] + \frac{\partial\phi}{\partial w}[Y']\right)\mathrm{e}^{\int_s^t(\frac{\partial\phi}{\partial z}[Y]+\frac{\partial\phi}{\partial w}[Y'])\,\mathrm{d}r}\,\mathrm{d}s - a_k[Y], \tag{40b}$$

and

$$-\int_0^1 \psi a_k[Y']\,\mathrm{d}t = \int_0^1 h_k(t)\left(\frac{\partial\phi}{\partial z}[V] + \frac{\partial\phi}{\partial w}[V']\right)\mathrm{d}t, \tag{41a}$$

where

$$h_k(t) = -\int_t^1 a_k[Y']\,\mathrm{e}^{\int_s^t(\frac{\partial\phi}{\partial z}[Y]+\frac{\partial\phi}{\partial w}[Y'])\,\mathrm{d}r}\,\mathrm{d}s. \tag{41b}$$

Substituting (40a)–(41b) in (39), and considering as before all the variations V such that $V(0) = V(1) = 0$, we have:

$$\int_0^1 a_k\big[(V-\psi Y)'\big]\,\mathrm{d}t = \int_0^1 \left\{a_k + (g_k+h_k)\frac{\partial\phi}{\partial w} - \left(\int_0^t (g_k+h_k)\frac{\partial\phi}{\partial z}\,\mathrm{d}s\right)\right\}[V']\,\mathrm{d}t. \tag{42}$$

Since $\{a_k\}$ converges to 0 in L^{p^*}, combining together (38) and (42) into (37), and using (9a)–(10b), there exists a sequence $\{b_k\}$ in L^{p^*} such that $\|b_k\|_{L^{p^*}} \xrightarrow{k\to\infty} 0$ and

$$\left(\frac{\partial L}{\partial w} - \lambda_k\frac{\partial\phi}{\partial w}\right) - \int_0^t\left(\frac{\partial L}{\partial z} - \lambda_k\frac{\partial\phi}{\partial z}\right)\mathrm{d}s - b_k = c_k, \quad \text{in } [0,1], \tag{43}$$

where $c_k' = 0$.

Since $\|\dot{z}_k\|_p$ is equibounded too, by (9a)–(10b), integrating (43) we obtain also c_k equibounded. Let us now observe that, by (9a)–(10b), $\lambda_k(t)$ and

$$\int_0^t \left(\frac{\partial L}{\partial z} - \lambda_k \frac{\partial \phi}{\partial z}\right) \mathrm{d}s$$

are equibounded in $H^{1,1}$. Therefore

$$\frac{\partial L}{\partial w}(\dot{z}_k, z_k, t) - \lambda_k \frac{\partial \phi}{\partial w}(\dot{z}_k, z_k, t) = \chi_k, \tag{44}$$

for some χ_k converging in L^{p^*} (up to subsequences, because $H^{1,1}$ is compactly embedded in L^{p^*} (see [4]), and b_n goes to zero in L^{p^*}.

Let us apply both members of this equation to $\dot{z}_k$, and exploit (8a) to find the existence of a constant C such that

$$|\dot{z}_k|^p \leqslant C\big(1 + |\chi_k|^{p^*}\big),$$

and the right-hand side above converges in L^1.

Then we have reduced ourselves to find a pointwise convergence for $\{\dot{z}_k\}$, up to subsequences, in order to apply Lebesgue Theorem, since it is, for some constant D,

$$|\dot{z}_n - \dot{z}|^p \leqslant D\big(1 + |\dot{z}|^p + |\chi_k|^{p^*}\big), \tag{45}$$

and the right-hand side above converges in L^1.

For this aim (that is, to find pointwise convergence), we will use the Implicit Function Theorem and Caccioppoli Global Inversion Theorem (see [5]), starting from Eq. (44).

Let us choose a local coordinate system where

$$Y = \frac{\partial}{\partial z^n}.$$

From (5), in this system the constraint can be written as

$$\phi(w, z, t) = w^n - g\big(w^1, \ldots, w^{n-1}, z, t\big), \tag{46}$$

with $g \in C^2$. Using this equation and (28) in (44) we have, for each $k \in \mathbb{N}$,

$$\begin{aligned}
&\frac{\partial L}{\partial w^i}\big(\dot{z}_k^1, \ldots, \dot{z}_k^{n-1}, g\big(\dot{z}_k^1, \ldots, \dot{z}_k^{n-1}\big), z_k, t\big) \\
&\quad - \lambda_k(t) \frac{\partial \phi}{\partial w^i}\big(\dot{z}_k^1, \ldots, \dot{z}_k^{n-1}, g\big(\dot{z}_k^1, \ldots, \dot{z}_k^{n-1}\big), z_k, t\big) \\
&\quad - (\chi_k)_i(t) = 0, \quad i = 1, \ldots, n-1,
\end{aligned} \tag{47}$$

and

$$\lambda_k(t) = -(\chi_k)_n(t).$$

Let us denote by y_k the point in $\mathbb{R}^{n-1}$:

$$y_k = \left(\dot{z}_k^1, \dots, \dot{z}_k^{n-1}\right).$$

We claim that we can apply Implicit Function Theorem and Caccioppoli Theorem to (47), in order to prove that y_k is pointwise convergent. Note that, up to subsequences, λ_k and χ_k are convergent almost everywhere to λ and χ, respectively.

Let $t \in [0,1]$ be fixed and such that $\lambda_k(t) \to \lambda(t)$ and $\chi_k(t) \to \chi(t)$. We want to prove that $\dot{z}_k(t)$ converges. To this aim, we consider the application Λ, locally defined as:

$$\begin{aligned} &\Lambda\left(w^1, \dots, w^{n-1}\right) \\ &\quad = \frac{\partial L}{\partial w^i}\left(w^1, \dots, w^{n-1}, g\left(w^1, \dots, w^{n-1}, z, t\right), z, t\right) \\ &\qquad - \lambda(t)\frac{\partial \phi}{\partial w^i}\left(w^1, \dots, w^{n-1}, g\left(w^1, \dots, w^{n-1}, z, t\right), z, t\right) - \chi_i(t) \\ &\quad = 0, \quad i = 1, \dots, n-1, \end{aligned} \tag{48}$$

Deriving (48) with respect to w^j we get, for $i, j = 1, \dots, n-1$,

$$\begin{aligned} &\Lambda_{ij}\left(y_k, g(y_k, z_k, t), z_k, t\right) \\ &\quad = \left[\left(\frac{\partial^2 L}{\partial w^i \partial w^j} + \frac{\partial^2 L}{\partial w^i \partial w^n}\frac{\partial g}{\partial w^j}\right) - \lambda(t)\left(\frac{\partial^2 \phi}{\partial w^i \partial w^j} + \frac{\partial^2 \phi}{\partial w^i \partial w^n}\frac{\partial g}{\partial w^j}\right)\right] \\ &\qquad \times \left(y_k, g(y_k, z_k, t), z_k, t\right), \end{aligned} \tag{49}$$

and we must show that[1]

$$\Lambda_{ij}\xi^i\xi^j \neq 0.$$

Now we use hypothesis (6). A vector $(\xi^1, \dots, \xi^n) \in \mathcal{S}_{(y_k, g(y_k, z_k, t), z_k, t)}$ is such that, in this local system,

$$\xi^n = \frac{\partial g}{\partial w^i}(y_k, z_k, t)\xi^i.$$

Writing (6) in coordinates we have (dropping the argument $(y_k, g(y_k, z_k, t), z_k, t)$):

[1] We use Einstein's indices convention: here, and in the following, the repeated indices i, j run from 1 to $n-1$.

$$
\begin{aligned}
0 &< \frac{\partial^2 L}{\partial w^i \partial w^j}\xi^i\xi^j + \frac{\partial^2 L}{\partial w^i \partial w^n}\xi^i\frac{\partial g}{\partial w^j}\xi^j + \frac{\partial^2 L}{\partial w^n \partial w^j}\frac{\partial g}{\partial w^i}\xi^i\xi^j + \frac{\partial^2 L}{\partial (w^n)^2}\frac{\partial g}{\partial w^i}\frac{\partial g}{\partial w^j}\xi^i\xi^j \\
&= \left(\frac{\partial^2 L}{\partial w^i \partial w^j} + \frac{\partial^2 L}{\partial w^i \partial w^n}\frac{\partial g}{\partial w^j}\right)\xi^i\xi^j + \left(\frac{\partial^2 L}{\partial w^n \partial w^j} + \frac{\partial^2 L}{\partial (w^n)^2}\frac{\partial g}{\partial w^j}\right)\frac{\partial g}{\partial w^i}\xi^i\xi^j \\
&= \left(\frac{\partial^2 L}{\partial w^i \partial w^j} + \frac{\partial^2 L}{\partial w^i \partial w^n}\frac{\partial g}{\partial w^j}\right)\xi^i\xi^j. \qquad (50)
\end{aligned}
$$

Last equality above follows from (28), that in local coordinates reads:

$$
\frac{\partial L}{\partial w^n}\big(y_k, g(y_k, z_k, z, t), z, t\big) = 0,
$$

and deriving this with respect to w^j we have:

$$
\left(\frac{\partial^2 L}{\partial w^n \partial w^j} + \frac{\partial^2 L}{\partial (w^n)^2}\frac{\partial g}{\partial w^j}\right)\big(y_k, g(y_k, z_k, z, t), z, t\big) = 0,
$$

from which we obtain last inequality in (50).

Moreover we have, for each $\xi \in \mathcal{S}_{(y_k, g(y_k, z_k, t), z_k, t)}$,

$$
\begin{aligned}
&-\lambda_k(t)\frac{\partial^2 \phi}{\partial w^2}\big(y_k, g(y_k, z_k, t), z, t\big)[\xi, \xi] \\
&\quad = -\lambda_k(t)\left(\frac{\partial^2 \phi}{\partial w^i \partial w^j} + \frac{\partial^2 \phi}{\partial w^i \partial w^n}\frac{\partial g}{\partial w^j}\right)\big(y_k, g(y_k, z_k, t), z_k, t\big)\xi^i\xi^j. \qquad (51)
\end{aligned}
$$

We now observe that, from (7), this quantity is always $\geqslant 0$ (indeed, if (7) holds, from (35) we have $\lambda_k(t) \leqslant 0$).

Using this fact and (50) in (49) we find

$$
\Lambda_{ij}\xi^i\xi^j > 0 \quad \text{if } \xi \neq 0.
$$

Now, multiplying both terms in (44) by $\dot{z}_k$ and using (8a) we obtain that $\dot{z}_k(t)$ is bounded. Let w be a limit point. We have the following situation: fixed

$$
\big(z(t), t, \chi(t), \lambda(t)\big),
$$

there exists w such that $\phi(w, z, t) = 0$ and $\Lambda_{(z(t), t, \chi(t), \lambda(t))}(w) = 0$, where Λ is defined in (48). Taking into account $\Lambda(w)[w]$ and using (8a) we have that

$$
\Lambda : \big\{\phi^{-1}(0) \cap T_{z(t)}\mathcal{M}\big\} \to \Lambda\big(\big\{\phi^{-1}(0) \cap T_{z(t)}\mathcal{M}\big\}\big)
$$

is a proper function between two metric spaces.

Considering Remark 2.14 and $\Lambda(\sigma w + \mu(\sigma)Y(z))$ we see that

$$
\Lambda\big(\big\{\phi^{-1}(0) \cap T_{z(t)}\mathcal{M}\big\}\big)
$$

is contractible to $\Lambda_{(z(t),t,\chi(t),\lambda(t))}(0)$. Moreover $\phi^{-1}(0) \cap T_{z(t)}\mathcal{M}$ is arcwise connected, again by Remark 2.14. Hence, using Caccioppoli Theorem, we see that $\Lambda(w) = 0$ has a unique solution, which is the limit of $\dot{z}_k(t)$.

Repeating this argument for almost every $t \in [0, 1]$, we have pointwise convergence a.e. of $\dot{z}_k$, and we can apply Lebesgue Theorem, finding that, by (45), z_k has a converging subsequence in $H^{1,p}$. □

Remark 3.6. Proposition 3.5 immediately gives the existence of the minimizer for $\mathcal{L}$ in $\Omega_{Q,\gamma}(\phi)$. Indeed the following well-known theorem applies to our case:

Theorem 3.7 [14]. *Let $F : X \to \mathbb{R}$ be a C^1 functional on a C^1 Banach manifold with the following properties*:

(1) *there exists $c \geqslant \inf_X F$ such that $F^c = \{x \in X\colon\ F(x) \leqslant c\}$ is complete,*
(2) $\inf_X F > -\infty$, *and*
(3) *F satisfies the Palais–Smale condition at* $\inf_X F$.

There F admits a minimizer in X.

The pseudo-coercivity assumption of $\mathcal{L}$ on $\Omega_{Q,\gamma}(\phi)$ gives hypothesis (1) of the theorem, and moreover $\mathcal{L}$ is bounded from below by hypotheses (1). Therefore the Palais–Smale condition proved in Proposition 3.5 allows us to obtain the existence result stated in Theorem 3.1.

We stress the point that the constraint in our problem is in general not closed with respect to the weak convergence. This is the reason why we had to introduce the Palais–Smale condition and use Theorem 3.7 to prove existence of minimizers. Clearly, they are critical points of $\mathcal{L}$ in $\Omega_{Q,\gamma}(\phi)$.

Regularity of solutions is stated by the following:

Proposition 3.8. *The critical points $z \in \Omega_{Q,\gamma}(\phi)$ of $\mathcal{L}$ belong to $C^2([0,1], \mathcal{M})$. Moreover, there exists $\lambda \in C^1([0,1], \mathbb{R})$ such that the couple (z, λ) solves Eqs.* (30)–(31).

Proof. We will use a bootstrap argument to prove the first part of the assertion. Let $z(t) \in \Omega_{Q,\gamma}(\phi)$ a critical point of $\mathcal{L}$. Arguing as in the proof of Proposition 3.5 there exists a function $\lambda(t) \in H^{1,1}$ such that

$$\lambda(t) = \int_t^1 \left(\frac{\partial L}{\partial z}[Y] + \frac{\partial L}{\partial w}[Y']\right) \mathrm{e}^{\int_s^t (\frac{\partial \phi}{\partial z}[Y] + \frac{\partial \phi}{\partial w}[Y'])\,\mathrm{d}r}\,\mathrm{d}s, \tag{52}$$

and

$$\mathrm{d}\mathcal{L}(z)[V - \psi_V Y] = \int_0^1 \left(\frac{\partial L}{\partial z} - \lambda\frac{\partial \phi}{\partial z}\right)[V] + \left(\frac{\partial L}{\partial w} - \lambda\frac{\partial \phi}{\partial w}\right)[V']\,\mathrm{d}t, \tag{53}$$

$\forall V \in T_z\Omega_{Q,\gamma},\ V(0) = V(1) = 0$. Then

$$\frac{\partial L}{\partial w}(\dot z, z, t) - \lambda(t)\,\frac{\partial \phi}{\partial w}(\dot z, z, t) = \theta(t), \tag{54}$$

where $\theta \in H^{1,1}([0,1], \mathcal{M})$ is

$$\theta(t) = c + \int_0^t \left(\frac{\partial L}{\partial z} - \lambda \frac{\partial \phi}{\partial z}\right)[V]\,\mathrm{d}s, \tag{55}$$

and $c' = 0$.

Since θ and λ are now continue, we can use the Implicit Function Theorem exactly as done in the proof of Proposition 3.5, finding locally a C^1 function h such that

$$\dot z(t) = h\big(z(t), t, \chi(t), \lambda(t)\big),$$

and since z, θ and λ are continue, so is $\dot z(t)$. This implies (see (52) and (55)) that also λ is C^1, and θ is C^1. Thus, the above equality gives $\dot z$ of class C^1.

It is now possible to integrate by parts expression (53) in the “right” direction, obtaining the Euler–Lagrange equation (30).

Corollary 3.9. *Let us suppose L and ϕ does not depend on time, that is $L = L(w, z)$ and $\phi = \phi(w, z)$. Then, given a critical point z for $\mathcal{L}$ in $\Omega_{Q,\gamma}(\phi)$ and λ as in Proposition* 3.8, *the following quantity is constant along $z(t)$*:

$$E = \frac{\partial L}{\partial \dot z}\dot z(t) - L - \lambda(t)\frac{\partial \phi}{\partial \dot z}\dot z(t). \tag{56}$$

Proof. Let us rewrite Euler–Lagrange equation (30):

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial w} - \lambda \frac{\partial \phi}{\partial w}\right) - \left(\frac{\partial L}{\partial z} - \lambda \frac{\partial \phi}{\partial z}\right)\right]\big(\dot z(t), z(t), t\big) = 0.$$

Applying this to $\dot z$ we get:

$$\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial w}\right)[\dot z] - \frac{\partial L}{\partial z}[\dot z] - \frac{\mathrm{d}}{\mathrm{d}t}\left(\lambda \frac{\partial \phi}{\partial w}\right)[\dot z] + \lambda \frac{\partial \phi}{\partial z}[\dot z] \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial w}[\dot z]\right) - \left(\frac{\partial L}{\partial z}[\dot z] + \frac{\partial L}{\partial w}[\ddot z]\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\lambda \frac{\partial \phi}{\partial w}[\dot z]\right) + \lambda\left(\frac{\partial \phi}{\partial z}[\dot z] + \frac{\partial \phi}{\partial w}[\ddot z]\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial w}[\dot z] - L - \lambda \frac{\partial \phi}{\partial w}[\dot z] + \lambda \phi\right) + \frac{\partial L}{\partial t} - \lambda(t)\frac{\partial \phi}{\partial t}.
\end{aligned}$$

The result follows using time-independence of L and ϕ. □

Remark 3.10. Let us consider the case of a function $z:[0,1]\in\mathcal{M}$, of class C^2, that solves Euler–Lagrange equation (30), where $\lambda(t)$ is given by (31). We ask whether this conditions are sufficient to guarantee that $(\dot z(t),z(t),t)$ is in the constraint $\phi=0$ or not.

We first observe that, from (31), λ solves the following ODE:

$$\dot\lambda(t)-\left[\left(\frac{\partial\phi}{\partial z}[V]+\frac{\partial\phi}{\partial w}[V']\right)[Y]\right]\lambda(t)+\left(\frac{\partial L}{\partial z}[Y]+\frac{\partial L}{\partial w}[Y']\right)[Y]=0, \tag{57}$$

with initial condition $\lambda(0)=0$.

Applying (30) to Y we find:

$$\begin{aligned}&\frac{\partial\phi}{\partial w}(\dot z,z,t)\big[Y(z)\big]\dot\lambda+\lambda\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\phi}{\partial w}-\frac{\partial\phi}{\partial z}\right)(\dot z,z,t)\big[Y(z)\big]\\&\quad-\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial w}-\frac{\partial L}{\partial z}\right)(\dot z,z,t)\big[Y(z)\big]=0.\end{aligned} \tag{58}$$

But we have:

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\phi}{\partial w}-\frac{\partial\phi}{\partial z}\right)(\dot z,z,t)\big[Y(z)\big]=-\left(\frac{\partial\phi}{\partial z}[V]+\frac{\partial\phi}{\partial w}[V']\right)[Y]+\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\phi}{\partial w}[Y]\right),$$

and analogously

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial w}-\frac{\partial L}{\partial z}\right)(\dot z,z,t)\big[Y(z)\big]=-\left(\frac{\partial L}{\partial z}[V]+\frac{\partial L}{\partial w}[V']\right)[Y]+\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial w}[Y]\right).$$

Let us now suppose that

$$\frac{\partial\phi}{\partial w}[Y]\equiv 1$$

on *all* $T\mathcal{M}\times\mathbb{R}$. Using this fact and the above expressions together with (57) we find:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial w}[Y]=0, \tag{59}$$

that is $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial w}[Y]=\text{const}$. Then if the equation

$$\frac{\partial L}{\partial w}(w,z,t)\big[Y(z)\big]=0$$

describes *exactly* the constraint, that is (28) it is satisfied if and only if $\phi(w,z,t)=0$, we conclude that if the initial data $(\dot z(0),z(0),0)$ is in the constraint, then the curve z is in the constraint for all $t\in[0,1]$. This is, for example, the case of the sub-Riemannian geodesics (see (90) of Section 6.1).

Also multiplicity results can be obtained. We first recall the following:

Definition 3.11. The *Ljusternik–Schnirelman category* $\mathrm{cat}(X)$ of a topological space X is the possibly infinite minimal number of closed contractible subsets of X that form a covering of X.

We can apply the classical Ljusternik–Schnirelman theory (see [14]) to obtain a multiplicity result for critical points between Q and γ given in terms of the Ljusternik–Schnirelman category of the Banach manifold $\Omega_{Q,\gamma}(\phi)$:

Theorem 3.12. *Under the assumptions of Theorem* 3.1, *there are at least*

$$\mathrm{cat}\big(\Omega_{Q,\gamma}(\phi)\big)$$

critical points between Q and γ. Moreover, if $\mathrm{cat}(\Omega_{Q,\gamma}(\phi))$ *is infinite, then there exists a sequence* $\{z_n\}_{n\in\mathbb{N}}$ *of critical points between Q and γ such that*

$$\lim_{n\to\infty}\mathcal{L}(z_n)=+\infty.$$

Note that $\sup_{\Omega_{Q,\gamma}(\phi)}\mathcal{L}=+\infty$.

Now we want to show some results that allows us to calculate $\mathrm{cat}(\Omega_{Q,\gamma}(\phi))$, supposing that Y does not allow closed integral curves. Let us introduce the spaces:

$$C^{\alpha}_{\Omega_{Q,\gamma}}=\Omega_{Q,\gamma}\cap C^{\alpha}\big([0,1],\mathcal{M}\big), \tag{60}$$

and

$$C^{\alpha}_{\Omega_{Q,\gamma}(\phi)}=\Omega_{Q,\gamma}(\phi)\cap C^{\alpha}\big([0,1],\mathcal{M}\big). \tag{61}$$

Let $z\in C^{1}_{\Omega_{Q,\gamma}}$ and $\psi:\mathcal{M}\times\mathbb{R}\to\mathbb{R}$ be the flow of the vector field Y ($\psi(z,t)=\gamma_x(t)$, where γ_x is the integral curve of Y such that $\gamma_x(0)=x$). We define:

$$z_h(t)=\psi\big(z(t),h\cdot\eta(t)\big),$$

where $\eta:[0,1]\to\mathbb{R}$ must be opportunely chosen. Then it must be $\eta(0)=0$, in such a way that $z_1(0)=Q$. Moreover the definition of flow assures $z_1(1)=\psi(z(1),\eta(1))\in\gamma(\mathbb{R})$. Then we would like to have, $\forall t\in[0,1]$,

$$\begin{cases}\phi\Big(d_z\psi\big(z(t),\eta(t)\big)\big[\dot z(t)\big]+Y\Big(\psi\big(z(t),\eta(t)\big)\Big)\dot\eta(t),\,\psi\big(z(t),\eta(t)\big),t\Big)=0,\\ \eta(0)=0.\end{cases} \tag{62}$$

Using the initial condition we have:

$$\left.\frac{\partial\phi}{\partial\dot\eta}\right|_{\eta=0}=\frac{\partial\phi}{\partial\dot z}\big(\dot z(0),Q,0\big)\big[Y(x)\big]=1,$$

from (5). Then there exists a unique $\Theta(\eta,t)$ defined in $[0,\eta_0)\times[0,t_0)$, such that $\Theta(0,0)=\mu(\dot z(0),Q,0)$ ($\mu(\dot z(0),Q,0)$ is defined in Lemma 2.13), and (62) is equivalent to

$$\dot{\eta}(t) = \Theta(\eta, t), \qquad \eta(0) = 0.$$

Let us suppose that we can extend this solution to the whole interval $[0, 1]$; we can now observe that, if $z \in C^1_{\Omega_{Q,\gamma}(\phi)}$, $\eta(t) \equiv 0$ is the unique solution of (62). Then $z_h(t) = z(t)$ in $[0, 1]$. This assures that $C^1_{\Omega_{Q,\gamma}(\phi)}$ is a strong retract of $C^1_{\Omega_{Q,\gamma}}$, and then the following proposition is proved:

Proposition 3.13. *If Eq.* (62) *can be solved in* $[0, 1]$ *for every* $z \in C^1_{\Omega_{Q,\gamma}}$, *then* $C^1_{\Omega_{Q,\gamma}}$ *and* $C^1_{\Omega_{Q,\gamma}(\phi)}$ *are homotopically equivalent.*

Using convolution operators it can be seen that $\Omega_{Q,\gamma}(\phi)$ and $C^1_{\Omega_{Q,\gamma}(\phi)}$ are homotopically equivalent. Moreover, since $\mathrm{cat}(\,\cdot\,)$ is invariant under homotopic equivalences, we get $\mathrm{cat}(C^1_{\Omega_{Q,\gamma}}) = \mathrm{cat}(C^1_{\Omega_{Q,\gamma}(\phi)})$, and using again convolution operators it can be seen that

$$\mathrm{cat}\big(C^1_{\Omega_{Q,\gamma}}\big) = \mathrm{cat}\big(C^0_{\Omega_{Q,\gamma}}\big).$$

Then, from a well know result by Fadell and Husseini [6], and from regularity of critical points stated by Theorem 3.8, we can state the following:

Theorem 3.14. *Under the above hypotheses* (*see in particular Proposition* 3.13), *if* $\mathcal{M}$ *is not contractible, there exist infinite critical points* z_n *between* Q *and* γ, *such that* $\lim_{n\to\infty} \mathcal{L}(z_n) = +\infty$.

As a particular situation, we will consider the case when $\mathcal{M}$ admits a global splitting $\mathcal{M} = \mathcal{M}_0 \times \mathbb{R}$. We put on $\mathcal{M}$ the coordinates

$$\big(z^1, \ldots, z^{n-1}, \theta\big),$$

and suppose that the vector field Y of the hypotheses is

$$Y = \frac{\partial}{\partial\theta}.$$

Let us denote $Q = (x_0, \theta_0)$ and let $\gamma = (x_1, \gamma(t))$ be and integral curve of Y, $x_0 \neq x_1$. Let us also fix a curve $z(t)$ in the space

$$C^1_0 = \big\{z \in \big([0, 1], \mathcal{M}_0\big) \bigm| z(0) = x_0\ z(1) = x_1\big\}.$$

Then, arguing as before, or some function Θ_z the constraint equation

$$\phi\big(\dot{z}(t), \dot{\theta}, z(t), \theta, t\big) = 0$$

is equivalent to the ODE:

$$\dot{\theta}(t) = \Theta_z(\theta, t), \quad \theta(0) = \theta_0,$$

Thus, the curve $(z(t), \theta(t))$ just found is in the space $C^1_{\Omega_{Q,\gamma}(\phi)}$ (61). This process is obviously reversible, giving a homomorphism between C^1 curves in $\mathcal{M}$ between point and line satisfying the constraint and C^1 curves in $\mathcal{M}_0$ between two points.

4. Local theory

In this section we will prove *local* uniqueness of minimizers for $\mathcal{L}$ in $\Omega_{Q,\gamma}(\phi)$, through the study of the flow induced by the Euler–Lagrange equations (30).

For this aim we will make the assumptions

$$\frac{\partial L}{\partial z}(0,0,0) = \frac{\partial^2 L}{\partial z \partial z}(0,0,0) = \frac{\partial^2 L}{\partial w \partial z}(0,0,0) = 0. \tag{63}$$

Remark 4.1. The above assumption is satisfied by the particular cases exposed in Section 6.1.

We will begin proving the following theorem:

Proposition 4.2. *Let $Q \in \mathcal{M}$ and $v_0 \in T_Q\mathcal{M}$ such that $|v_0|$ is sufficiently small. Then there exists a unique solution $z(t)$ of the integro-differential* (30)–(31)*, defined in the interval* $[0,1]$*, such that $\phi(\dot z(t), z(t), t) = 0$, satisfying the initial conditions $z(0) = Q,\ \dot z(0) = v_0$.*

Proof. We will work in a local coordinate system of $0 \in \mathbb{R}^n$ such that

$$Y = \frac{\partial}{\partial z^n}. \tag{64}$$

From (5) and the Implicit Function Theorem there exists a neighborhood $\mathcal{U}^n \times \mathcal{V}^n \times (-\varepsilon, +\varepsilon)$ of $(0,0,0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ such that

$$\phi = w^n - g\big(w^1, \ldots, w^{n-1}, z, t\big), \tag{65}$$

where g is a C^1 function, uniquely determined. Let us now write Euler–Lagrange equations: in local coordinates we have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial w^i} - \lambda \frac{\partial \phi}{\partial w^i}\right) - \left(\frac{\partial L}{\partial z^i} - \lambda \frac{\partial \phi}{\partial z^i}\right) = 0, \quad \forall i = 1, \ldots, n, \tag{66}$$

where $\lambda(t)$ is—using (64) together with (5) and (28)—

$$\lambda(t) = \int_t^1 \frac{\partial L}{\partial z^n}\big(\dot z(s), z(s), s\big) \mathrm{e}^{-\int_s^t \frac{\partial g}{\partial z^n}(\dot z(r), z(r), r)\,\mathrm{d}r}\,\mathrm{d}s. \tag{67}$$

We observe that, integrating (66) in $[t,1]$ for $i=n$ we obtain (67). Then, if $\mathcal{U}^{n-1}$ denotes the projection onto $\mathbb{R}^{n-1}$ of $\mathcal{U}^n$, we define the application:

$$\Phi : \bigl\{z \in C^0\bigl([0,1], \mathcal{U}^{n-1} \times \mathcal{V}^n\bigr) \bigm| z(0)=0\bigr\} \times \mathcal{U}^{n-1} \to C^0\bigl([0,1], \mathbb{R}^{n-1} \times \mathbb{R}^n\bigr),$$

$$\Phi\bigl(w^1,\dots,w^{n-1},z,v\bigr)(t) = \begin{pmatrix} \Phi_1(w,z,v)(t) \\ \dots \\ \Phi_{n-1}(w,z,v)(t) \\ z^1 - \int_0^t w^1\,\mathrm{d}s \\ \dots \\ z^n - \int_0^t w^n\,\mathrm{d}s \end{pmatrix}, \tag{68}$$

where, for $i=1,\dots,n-1$,

$$\begin{aligned}
&\Phi_i\bigl(w^1,\dots,w^{n-1},z,v\bigr)(t) \\
&\quad = \left(\frac{\partial L}{\partial w^i} - \lambda \frac{\partial \phi}{\partial w^i}\right)\bigl(w(t),z(t),t\bigr) - \int_0^t \left(\frac{\partial L}{\partial z^i} - \lambda \frac{\partial \phi}{\partial z^i}\right)\bigl(w(s),z(s),s\bigr)\,\mathrm{d}s \\
&\qquad - \left(\frac{\partial L}{\partial w^i}(v,0,0) - \lambda(0)\frac{\partial \phi}{\partial w^i}(v,0,0)\right),
\end{aligned} \tag{69}$$

with the functional dependencies

$$w^n = g\bigl(w^1,\dots,w^{n-1},z,t\bigr), \tag{70}$$

$$\lambda(w,z)(t) = \int_t^1 \frac{\partial L}{\partial z^n}\bigl(w(s),z(s),s\bigr) \mathrm{e}^{-\int_s^t \frac{\partial g}{\partial z^n}(w(r),z(r),r)\,\mathrm{d}r}\,\mathrm{d}s. \tag{71}$$

We also set, for any $v=(v^1,\dots,v^{n-1})$,

$$\bar{v} = \bigl(v_1,\dots,v_{n-1}, g(v_1,\dots,v_{n-1},0,0)\bigr).$$

Moreover, if $\Phi(w^1(t),\dots,w^{n-1}(t),z(t),v_0)=0$ we can set:

$$w^n(t) = g\bigl(w^1(t),\dots,w^{n-1}(t),z(t),t\bigr)$$

and then, using (67)–(70), $w(t)=\dot{z}(t)$, $\phi(w,z,t)=0$, and (w,z) solves Euler–Lagrange equations with $z(0)=0$, $w(0)=\overline{v_0}=(v_0,g(v_0,0,0))$. Here we stress the fact that $w(0)=\overline{v_0}$ comes from the local inversion showed in the proof of Theorem 3.5 (see pp. 12–13).

To prove the theorem we want to show that

$$\frac{\partial \Phi}{\partial(w,z)}(0,0,0) : C^0\bigl([0,1],\mathbb{R}^{n-1}\times\mathbb{R}^n\bigr) \to C^0\bigl([0,1],\mathbb{R}^{n-1}\times\mathbb{R}^n\bigr)$$

is invertible, so that there exists a C^1 map

$$v \mapsto (w_v, z_v),$$

unique for $|v|$ sufficiently small, with $\Phi(w_v, z_v, v) = 0$. With the notation:

$$G(w,z,t) = G_{ij}(w,z,t) = \frac{\partial^2 L}{\partial w^i \partial w^j}(w,z,t), \quad \forall i, j = 1, \ldots, n-1,$$

and using (63)–(71) it can be seen that, given $[\omega, \xi] \in \mathbb{R}^{n-1} \times \mathbb{R}^n$, we have:

$$\frac{\partial \Phi}{\partial (w,z)}(0,0,0)[\omega,\xi] = \left(G(0,0,0) \cdot \omega, \xi - \int_0^t \omega \, \mathrm{d}s \right), \tag{72}$$

which is an invertible map, with inverse given by:

$$\left[\frac{\partial \Phi}{\partial (w,z)}(0,0,0) \right]^{-1} [\omega, \xi] = \left(G^{-1}(0,0,0) \cdot \omega, \xi - \int_0^t G^{-1}(0,0,0) \cdot \omega \, \mathrm{d}s \right). \tag{73}$$

The result is proved. $\square$

Using Φ of Proposition 4.2 $\forall Q \in \mathcal{M}$ we define an *exponential map* as follows. Let us denote by Q a point of $\mathcal{M}$ with coordinates (z), and set

$$\mathcal{T}_Q = \left\{ v \in T_Q\mathcal{M} \;\middle|\; \phi(v, z, 0) = 0 \right\}.$$

We can think the map $v \mapsto (w_v, z_v)$ as defined in $\mathcal{U}_Q = \mathcal{T}_Q \cap \mathcal{V}_Q$, where $\mathcal{V}_Q$ is an open subset of 0 in $T_Q\mathcal{M}$. Then we define $\mathfrak{exp}_Q(v)$ as the end-point of the (unique) solution of Euler–Lagrange equations (66), that is

$$\mathfrak{exp}_Q(v) = z_v(1). \tag{74}$$

From the Implicit Function Theorem, the differential at $v = 0$ of $v \to (w_v, z_v)$ is given by:

$$\left[\frac{\partial \Phi}{\partial (w,z)}(0,0,0) \right]^{-1} \circ \left(\frac{\partial \Phi}{\partial v}(0,0,0)[w] \right) = (w, w \cdot t), \quad \forall w \in \mathbb{R}^{n-1}, \tag{75}$$

where Φ is given by (69).

Since the derivative of the exponential map at $v = 0$ in the direction w is given by the evaluation at $t = 1$ of the second component of (75), it is $\mathrm{d}\,\mathfrak{exp}_Q(0)[w] = w$, that is $\mathrm{d}\,\mathfrak{exp}_Q(0)$ is the identity map on $T_Q\mathcal{U}_Q \subset T_Q\mathcal{M}$, then $\mathfrak{exp}_Q$ is a local diffeomorphism. Then the integral curves of Y, that in local coordinates are given by:

$$\gamma(t) = \left(\gamma^1, \dots \gamma^{n-1}, \gamma_k(t)\right),$$

are transversal to the image Σ_Q of $\mathcal{U}_Q$ through the exponential map, and intercept it exactly once if $\gamma(t)$ is sufficiently close to Q. By the local uniqueness of the flow Φ, we have here showed the following:

Theorem 4.3. *For all $Q \in \mathcal{M}$, if γ is sufficiently close to Q then there is a unique minimal curve z for $\mathcal{L}$ in $\Omega_{Q,\gamma}(\phi)$.*

5. The case $p = 1$

We dealt so far with problems where the Lagrangian function $L(w, z, t)$ was estimated by a power p of $|w|$, with $p > 1$. This is the case of most mechanical systems, where L is a quadratic function in the velocities ($p = 2$), and this is the reason why it has been so widely treated.

The sub-Riemannian geodesics case (see [7]) suggests a common technique to deal with problems where the Lagrangian is estimated by $|w|$, that is well-working with functionals where both L and ϕ are homogeneous of degree 1. It is an open problem whether suitable techniques can be exploited where the homogeneity condition is lost.

Let us consider a Lagrangian $L(w, z)$ on a Banach manifold $\mathcal{M}$ such that

$$L(\mu w, z) = \mu L(w, z), \quad \forall (w, z) \in T\mathcal{M},\ \forall \mu \in \mathbb{R}, \tag{76}$$

and we want to study critical points of

$$\mathcal{L}(z) = \int_0^1 L\left(\dot z(t), z(t)\right) \mathrm{d}t \tag{77}$$

parameterized in such a way that

$$L\left(\dot z(t), z(t)\right) = \text{const.},$$

in the set:

$$\begin{aligned} \mathcal{C}^1_{Q,\gamma}(\phi) = \big\{ z \in \mathcal{C}^1\left([0,1], \mathcal{M}\right) \bigm| z(0) = Q,\ z(1) \in \gamma(\mathbb{R}), \\ \phi\left(\dot z(t), z(t)\right) \equiv 0 \text{ a.e.} \big\}, \end{aligned} \tag{78}$$

where γ is an integral curve of the usual vector field Y on $\mathcal{M}$, $Q \notin \gamma(\mathbb{R})$, and the constraint equation ϕ is a linear function in w:

$$\phi(w, z) = D(z) \cdot w, \tag{79}$$

with $D(z)$ linear operator in $T_z\mathcal{M}$. We will first study critical points for the functional

$$\mathcal{E}(z) = \int_0^1 \big(L\big(\dot{z}(t), z(t)\big)\big)^2 \,\mathrm{d}t, \tag{80}$$

in the space $\Omega_{Q,\gamma}(\phi)$,

$$\Omega_{Q,\gamma}(\phi) = \big\{z \in H^{1,2}\big([0,1], \mathcal{M}\big) \bigm| z(0) = Q,\ z(1) \in \gamma(\mathbb{R}),\ \phi\big(\dot{z}(t), z(t)\big) \equiv 0 \text{ a.e.}\big\}.$$

Remark 5.1. The function

$$E(w, z) = \big(L(w, z)\big)^2$$

is homogeneous in w of degree 2. This implies

$$E(w, z) = |w|^2 E\left(\frac{w}{|w|}, z\right),$$

and since

$$S^1 \cap T_z\mathcal{M} = \big\{w \in T_z\mathcal{M} \bigm| |w| = 1\big\}$$

is compact, $\forall z \in \mathcal{M}$, we can obtain the estimates on E and its derivatives as in (8a)–(9b).

Similar estimates on ϕ are straightforwardly found since ϕ is linear.

As usual, we require pseudo-coercivity of $\mathcal{E}$ with respect to the constraint. This happens for instance, recalling Remark 2.7, if

$$E(w, z)|_{S^1 \cap T_z\mathcal{M}}$$

is bounded away from 0, for each $z \in \mathcal{M}$, and $\mathcal{M}$ is complete.

In order to relate the two functionals, we need the hypotheses made in Section 2 regarding the vector field Y.

Therefore, using theory from Sections 2–3, we have existence and regularity for critical points of $\mathcal{L}$ in $\Omega_{Q,\gamma}(\phi)$. Since hypotheses (63) of Section 4 are satisfied, as it can be easily seen using homogeneity, then local uniqueness of critical points is obtained.

Moreover, using Corollary 3.9, we have the following first integral for solution of the problem:

$$\frac{\partial E}{\partial \dot{z}}\dot{z} - E - \lambda \frac{\partial \phi}{\partial \dot{z}}\dot{z} = E - \lambda\phi = E, \tag{81}$$

that is

$$L(\dot z, z) = \text{const.}$$

if z is a critical point for the "energy" functional.

Minimizers of $\mathcal{L}$ and $\mathcal{E}$ are related by the following

Proposition 5.2. *If a curve z is a minimizer for $\mathcal{E}$ in $\Omega_{Q,\gamma}(\phi)$, then it is a minimizer for $\mathcal{L}$ in $\mathcal{C}^1_{Q,\gamma}(\phi)$ parameterized by $L(w, z) = \text{const.}$*

Remark 5.3. The converse is straightforwardly true, recalling a standard argument using Hölder's inequality.

Proof of Proposition 5.2. Let $z \in \Omega_{Q,\gamma}(\phi)$ be a minimizer for $\mathcal{E}$. We know z is C^2 and $L(z(t), z(t))$ is constant. We want to show that z is a minimizer for $\mathcal{L}$ in $\mathcal{C}^1_{Q,\gamma}(\phi)$.

Indeed, let us suppose there exists $y \in \mathcal{C}^1_{Q,\gamma}(\phi)$, with $\mathcal{L}(y) < \mathcal{L}(z)$.

Thus we can reparameterize y obtaining a curve x:

$$x(\rho) = y(\sigma), \quad \rho(\sigma) = \frac{1}{\mathcal{L}(y)} \int_0^\sigma L\big(\dot y(s), y(s)\big)\,\mathrm{d}s.$$

Then $x \in \Omega_{Q,\gamma}(\phi)$, and

$$\mathcal{E}(x) = \int_0^1 L\big(\dot x(\rho), x(\rho)\big)\,\mathrm{d}\rho = \mathcal{L}(y) \int_0^1 \frac{(L(\dot y(\sigma), y(\sigma)))^2}{L(\dot y(\sigma), y(\sigma))}\,\mathrm{d}\sigma = \big[\mathcal{L}(y)\big]^2. \tag{82}$$

Therefore

$$\mathcal{E}(x) = \big[\mathcal{L}(y)\big]^2 < \big[\mathcal{L}(z)\big]^2 = \mathcal{E}(z), \tag{83}$$

obtaining a contradiction. □

Thus we come to the following:

Theorem 5.4. *Under the assumptions made, the functional $\mathcal{L}$ attains its minimum z in $\mathcal{C}^1_{Q,\gamma}(\phi)$, that is of class C^2 and satisfies Euler–Lagrange equations* (30)*, with L replaced by $E(w, z) = (L(w, z))^2$. Moreover z can be parameterized in such a way that $L(\dot z, z)$ is a constant.*

Since local uniqueness results holds for $\mathcal{E}$, Proposition 5.2 yields the following:

Corollary 5.5. *Under the assumptions made, if γ is sufficiently close to Q there exists a unique minimizer for $\mathcal{L}$ in $\mathcal{C}^1_{Q,\gamma}(\phi)$ with $L(\dot z, z)$ constant.*

Also multiplicity results can be found, using the same techniques as in Section 3.

6. Examples

6.1. Example 1. Sub-Riemannian geodesics

A sub-Riemannian manifold consists of a triple $(\mathcal{M}, \Delta, g)$, where $\mathcal{M}$ is a smooth manifold, $\Delta \subset T\mathcal{M}$ is a smooth distribution in $\mathcal{M}$ and g is a positive definite metric tensor on Δ. We recall that a *distribution* $\mathcal{D}$ of dimension d on $\mathcal{M}$ is a differentiable map that associates to every point P of $\mathcal{M}$ a d-dimensional subspace of $T_P\mathcal{M}$.

We are interested in what are usually called *normal geodesics*, i.e. those curves z such that $\dot z \in \Delta$—called *horizontal* curves—and that "locally" minimize their length—that is, their restriction to a sufficiently small interval $[a, b] \subset [0, 1]$ are horizontal curves of minimal length between $z(a)$ and $z(b)$ (see Appendix B in [7] for further details).

The main obstruction for the approach to the problem is that, usually, the set of horizontal curves between two fixed point does not have a differential structure in general. This obstruction is overcome if we let the end point free to move on a submanifold which is transversal to Δ.

This suggests the following setup: take a complete Riemannian manifold $(\mathcal{M},\ g = \langle\cdot,\cdot\rangle)$, and a never vanishing vector field Y on $\mathcal{M}$, that we will suppose without loss of generality to be normalized:

$$\langle Y, Y\rangle = 1. \tag{84}$$

For sake of simplicity we consider the case of a codimension 1 distribution, precisely $\Delta = Y^\perp$, the orthogonal distribution to Y, fix a point $Q \in \mathcal{M}$, and consider a maximal integral curve $\gamma : \mathbb{R} \to \mathcal{M}$, $Q \notin \gamma(\mathbb{R})$, letting the end point free to move on it. Sub-Riemannian length minimizers are related to critical points of the functional,

$$\mathcal{L}(z) = \int_0^1 \sqrt{\bigl\langle \dot z(t), \dot z(t)\bigr\rangle}\,\mathrm{d}t, \tag{85}$$

in the set

$$\mathcal{C}^1_{Q,\gamma}(\Delta) = \bigl\{z \in C^1\bigl([0,1], \mathcal{M}\bigr) \bigm| z(0) = Q,\ z(1) \in \gamma(\mathbb{R}),\ \langle \dot z, Y\rangle = 0\bigr\}, \tag{86}$$

satisfying $|\dot z(t)| = \text{const}$. The presence of the square root in the expression of $\mathcal{L}$ may result in some technical difficulties that are overcome – as for the Riemannian geodesics – taking into account the *energy* functional

$$E(z) = \int_0^1 \bigl\langle \dot z(t), \dot z(t)\bigr\rangle\,\mathrm{d}t \tag{87}$$

in the space:

$$\Omega_{Q,\gamma}(\Delta) = \big\{z \in H^{1,2}\big([0,1],\mathcal{M}\big) \bigm| z(0) = Q,\ z(1) \in \gamma(\mathbb{R}),\ \langle \dot z, Y\rangle = 0 \text{ a.e.}\big\}. \tag{88}$$

Then results from Section 5 can be exploited to link critical points of $\mathcal{L}$ and E.

The constraint equation ϕ:

$$\phi(w,z) = \big\langle w, Y(z)\big\rangle$$

is linear in w, and the constraint is admissible (in the sense of (4)), since

$$\mathcal{S}_{(w,z)} = \Delta(z) = Y^{\perp}(z), \quad \forall (w,z) \in T\mathcal{M}, \tag{89}$$

and therefore $\dim \mathcal{S}_{(w,z)} = n-1$. Condition (5) is equivalent to assumption (84) on Y, since

$$\frac{\partial L}{\partial w} = \langle w, \cdot\rangle. \tag{90}$$

Moreover, Eq. (6) is straightly verified. Indeed $\partial^2 L/\partial w^2$ is now a real metric, and Y is orthogonal to the vectors in the constraint, that in this case are the same vectors of $\mathcal{S}$ (recall that $\mathcal{S} = \ker \partial\phi/\partial w$).

Also Eq. (7) holds, since ϕ is a linear constraint. It is also straightforward to verify the estimates on L and ϕ (8a)–(10b) for $p = 2$ and, using the fact that $(\mathcal{M}, \langle\cdot,\cdot\rangle)$ is complete, also pseudo-coercivity assumption is easily checked to hold.

The theory exposed ensures existence and regularity for critical points of E. Then, using results from Section 5, we can pass to the length functional and obtain the same results for it.

Defined the transpose of the covariant derivative $(\nabla Y)^*$ as the $(1,1)$-type tensor field on $\mathcal{M}$ such that $\forall x \in \mathcal{M},\ \forall v_1, v_2 \in T_x\mathcal{M}$,

$$\big\langle (\nabla Y)^*[v_1], v_2\big\rangle = \big\langle \nabla_{v_2} Y(x), v_1\big\rangle, \tag{91}$$

we have that a normal geodesic z is a curve of class C^2, parameterized with $|\dot z| = \text{const.}$, that satisfies the equation:

$$\nabla_{\dot z}\dot z - \nabla_{\dot z}(\lambda_z \cdot Y) + \lambda_z \cdot (\nabla Y)^*[\dot z] = 0, \tag{92}$$

where

$$\lambda_z(t) = \mathrm{e}^{\int_0^t \langle \dot z, \nabla_Y Y\rangle\,\mathrm{d}s} \cdot \left[\int_t^1 \langle \dot z, \nabla_{\dot z} Y\rangle \mathrm{e}^{-\int_0^s \langle \dot z, \nabla_Y Y\rangle\,\mathrm{d}r}\,\mathrm{d}s\right]. \tag{93}$$

About multiplicity, we first observe that Eq. (62) now takes the form:

$$\begin{cases} \langle d_z\psi(z(t),\eta(t))[\dot z(t)], Y(\psi(z(t),\eta(t)))\rangle + \dot\eta(t) = 0, \\ \eta(0) = 0, \end{cases} \tag{94}$$

that can be solved in [0, 1] for any $z \in \mathcal{C}^1_{Q,\gamma}(\Delta)$. Using Ljusternik–Schnirelman theory we get a multiplicity result for sub-Riemannian geodesics between Q and γ:

Theorem 6.1. *There are at least* $\operatorname{cat}(\mathcal{C}^1_{Q,\gamma}(\Delta))$ *normal geodesics between* P *and* γ. *Moreover, if* $\operatorname{cat}(\mathcal{C}^1_{Q,\gamma}(\Delta))$ *is infinite, then there exists a sequence* $\{z_n\}_{n\in\mathbb{N}}$ *of normal geodesics between* Q *and* γ *such that*:

$$\lim_{n\to\infty} E(z_n) = +\infty.$$

Note that L is quadratic in the velocities w, therefore (63) is verified, and local uniqueness of critical points holds when Q and γ are sufficiently close.

Example 6.2. As a dynamical interpretation of sub-Riemannian geodesics, let us consider the free motion of a solid body $\mathcal{B}$ that slides on a horizontal plane π. All but one contact points between the body $\mathcal{B}$ and π are free to slide in all directions, whereas the last contact point P is realized by a knife edge, and such that $\mathcal{B}$ can move on π along the knife edge.

We consider the special case when the projection of the mass center of $\mathcal{B}$ on the plane coincides with the contact point P.

If (x, y) are the coordinates of the projection of the mass center on the plane ($\pi = \mathbb{R}^2$), and θ is the angle between the plane of the knife edge and a fixed axes (say $0x$), the Lagrangian is given by (the body is assumed to have unit mass, and k is a constant – namely the radius of gyration)

$$L = \frac{1}{2}(\dot x^2 + \dot y^2 + k^2\dot\theta^2).$$

The condition on motion of the contact point P can be prescribed by the equation:

$$\mathrm{d}y = \tan\theta\,\mathrm{d}x,$$

where $\mathrm{d}x$ and $\mathrm{d}y$ are infinitesimal displacements respectively along the directions x and y. Therefore the equation of constraint comes from the above relation:

$$\phi = \dot x\sin\theta - \dot y\cos\theta = 0.$$

This can be viewed as follows: let $\mathcal{M} = \mathbb{R}^2 \times S^1$, and fix a point of $\mathcal{M}$ with coordinates (x, y, θ). Given $(\xi, \eta, \psi) \in T_{(x,y,\theta)}\mathcal{M}$, define the metric – induced by L,

$$\langle(\xi, \eta, \psi), (\xi, \eta, \psi)\rangle = \xi^2 + \eta^2 + k\psi^2.$$

Also let Y be the vector field:

$$Y = \sin\theta \frac{\partial}{\partial x} - \cos\phi \frac{\partial}{\partial y}.$$

Then the constraint equation $\phi = 0$ can be written as

$$\langle Y(z), \dot z \rangle = 0,$$

for each $z \in \mathcal{M}$, and the Lagrangian L is

$$L = \frac{1}{2}\langle \dot z, \dot z\rangle.$$

Since, fixed $\theta_0 \in S^1$, an integral curve γ of Y is given by

$$\gamma(s) = \big((\sin\theta_0)\, s, -(\cos\theta_0)\, s, \theta_0\big),$$

this means that we can study the motion of the wheel from a configuration (x_0, y_0, θ_0) given, to the set of configurations described by $\gamma(s)$.

6.2. *Example 2. Stably-causal Lorentzian manifold*

Let $(M, \langle\cdot,\cdot\rangle)$ be a Lorentzian manifold, endowed with a smooth *absolute time* function $T(z): \mathcal{M} \to \mathbb{R}$, such that $\langle \nabla T(z), \nabla T(z)\rangle = -1$. This is a particular case of *stably-causal* Lorentzian manifold [8]. Light-rays between a point $Q \in \mathcal{M}$ and an observer, i.e. an integral curve γ of ∇T, are related to critical points of

$$\mathcal{L} = \int_0^1 \langle \dot z(s), \dot z(s)\rangle_{(P)}\, \mathrm{d}s \tag{95}$$

in the space

$$\begin{aligned} \Omega^+_{Q,\gamma} = \big\{ z \in H^{1,2}\big([0,1], \mathcal{M}\big) \bigm| z(0) = Q,\ z(1) \in \gamma(\mathbb{R}), \\ \langle \nabla T(z), \dot z\rangle - \sqrt{\langle \dot z, \dot z\rangle_{(P)}} = 0 \text{ a.e.}\big\}, \end{aligned} \tag{96}$$

where $\langle\cdot,\cdot\rangle_{(P)}$ is the pseudo-Riemannian structure given by:

$$\langle \xi, \xi\rangle_{(P)} = \langle \xi, \xi\rangle + \langle \nabla T(z), \xi\rangle^2, \quad \xi \in T_z\mathcal{M}.$$

The constraint equation contained in (96) is not smooth at $\dot z = 0$. This problem can be approximated studying critical points of the functional (95) among all the $H^{1,2}$ curves between Q and $\gamma(\mathbb{R})$ satisfying the constraint $\phi_\varepsilon(\dot z, z) \equiv 0$ a.e., where

$$\phi_\varepsilon(w,z) \equiv \left\langle \nabla T(z), w \right\rangle + \varepsilon - \sqrt{\langle w, w\rangle_{(P)} + \varepsilon^2}. \tag{97}$$

This is a C^2 constraint, not linear in w. All the assumptions made are verified, as can be checked by some calculations. The only additional requirement

$$\left\langle H^T(z)[w], w \right\rangle \leqslant 0, \quad \forall (w,z) \in \phi^{-1}(0)$$

must be imposed to satisfy (7). Here $H^T(z)$ denotes the Hessian of the function T at a point $z \in \mathcal{M}$. Note that it can always be possible to reduce to this case, as shown in [8].

Appendix A. Geometry of Lagrangian systems

The aim of this Appendix is to give an intrinsic description of the objects used throughout the theory exposed in Sections 2–5, without dropping the coordinate representation we have so far used.

First, we need to recall some basic notion about the tangent bundle[2] of a manifold $\mathcal{M}$. We denote by:

$$\pi_{\mathcal{M}} : T\mathcal{M} \to \mathcal{M} \tag{A.1}$$

the tangent bundle of $\mathcal{M}$. Let $(z^i) = (z^1, \ldots, z^n)$ be a coordinate system on $\mathcal{M}$. Then a coordinate system[3] (z^i, w^i), $i = 1, \ldots, n$, is naturally induced on $T\mathcal{M}$; this system is *adapted* to the fibration $\pi_{\mathcal{M}}$, namely the expression of π_M in coordinates simply reads

$$\left(w^i, z^i\right) \xrightarrow{\pi_M} \left(z^i\right).$$

Fixed a point Q in this coordinate chart of $\mathcal{M}$, we denote by:

$$\left\{ \frac{\partial}{\partial z^i} \right\}_{i=1,\ldots,n} \tag{A.2}$$

the basis of $T_Q\mathcal{M}$ induced by the coordinate system (z^i).

Example A.1. Let $Y : \mathcal{M} \to T\mathcal{M}$ be a vector field on $\mathcal{M}$. We can express Y in coordinates using either the notation:

$$Y = Y^i(z) \frac{\partial}{\partial z^i}, \tag{A.3a}$$

[2] For further details about the tangent bundle and the tangent map, that we use later, see [1].

[3] From now we will use the notation (z^i) to mean the n-tuple $(z^1, \ldots, z^n)$. We will, moreover, use Einstein's repeated indices convention, as done in Section 3, p. 12.

or

$$\left(z^i\right) \xrightarrow{Y} \left(Y^i(z), z^i\right). \tag{A.3b}$$

Remark A.2. If $(\bar{z}^i)$ is another coordinate system on $\mathcal{M}$, and $\bar{z}^i = \bar{z}^i(z)$ are the transition functions, these are the expression for the change of basis of $T_Q\mathcal{M}$:

$$\frac{\partial}{\partial z^i} = \frac{\partial \bar{z}^j}{\partial z^i}(z)\frac{\partial}{\partial \bar{z}^j}. \tag{A.4a}$$

Analogously, if $(\overline{w}^i)$ are the induced coordinates on $T\mathcal{M}$ by $(\bar{z}^i)$, we have:

$$\overline{w}^i = \frac{\partial \bar{z}^i}{\partial z^j}(z)w^j. \tag{A.4b}$$

Given a manifold, we can always build its tangent bundle. So we can do for $T\mathcal{M}$. Its tangent bundle is denoted by:

$$\pi_{T\mathcal{M}} : TT\mathcal{M} \to T\mathcal{M}. \tag{A.5}$$

If (u^i, y^i, w^i, z^i) denotes the coordinate system on $TT\mathcal{M}$ induced by (w^i, z^i) of $T\mathcal{M}$, the projection (A.5) simply reads

$$\left(u^i, y^i, w^i, z^i\right) \xrightarrow{\pi_{T\mathcal{M}}} \left(w^i, z^i\right). \tag{A.6}$$

Fixed an element $(w, z) \in T\mathcal{M}$, the basis on $T_{(w,z)}T\mathcal{M}$ induced by the system (w^i, z^i) is denoted by

$$\left\{\frac{\partial}{\partial z^i}, \frac{\partial}{\partial w^i}\right\}_{i=1,\dots,n}. \tag{A.7}$$

Let us now take into account the tangent map of (A.1),

$$T\pi_{\mathcal{M}} : TT\mathcal{M} \to T\mathcal{M}. \tag{A.8}$$

Its expression, in terms of the basis (A.7), reads

$$u^i\frac{\partial}{\partial w^i} + y^i\frac{\partial}{\partial z^i} \xrightarrow{T\pi_{\mathcal{M}}} y^i\frac{\partial}{\partial z^i}. \tag{A.9}$$

We define the *vertical subbundle* of $TT\mathcal{M}$ to be the kernel of (A.8),

$$VTM = \ker T\pi_M. \tag{A.10}$$

This means the following: for each $(w, z) \in T\mathcal{M}$, we consider the restriction of $T\pi_M$ (A.8) to $T_{(w,z)}T\mathcal{M}$. This is a linear application from the vector space $T_{(w,z)}T\mathcal{M}$ to the vector space $T_z\mathcal{M}$. Its kernel, by definition (A.8), it's just $V_{(w,z)}T\mathcal{M}$, and

$$VTM = \bigcup_{(w,z)\in T\mathcal{M}} V_{(w,z)}T\mathcal{M}$$

is shown to possess a bundle structure over $T\mathcal{M}$.

As it can be seen from (A.9),

$$\left\{\frac{\partial}{\partial w^i}\right\}_{i=1,\dots,n} \tag{A.11}$$

forms a local basis for $V_{(w,z)}T\mathcal{M}$.

An injection v

$$VTM \overset{v}{\hookrightarrow} TTM \tag{A.12}$$

is naturally induced by (A.10). This injection does not depend on the choice of the coordinates, however, given the usual system (w^i, z^i) on $T\mathcal{M}$, the induced system on VTM is (u^i, w^i, z^i). Then (A.12) reads

$$\left(u^i, w^i, z^i\right) \overset{v}{\hookrightarrow} \left(u^i, 0, w^i, z^i\right). \tag{A.13}$$

Comparing (A.2) and (A.11), we get that, for every $(w, z) \in T\mathcal{M}$, a canonical isomorphism $\mathcal{I}_{(w,z)}$ between $V_{(w,z)}T\mathcal{M}$ and $T_z\mathcal{M}$ can be defined:

$$\frac{\partial}{\partial z^i} \in T_z\mathcal{M} \xrightarrow[\cong]{\mathcal{I}_{(w,z)}} \frac{\partial}{\partial w^i} \in V_{(w,z)}T\mathcal{M}. \tag{A.14}$$

Remark A.3. It can be seen that this isomorphism does not depend on the choice of the coordinate system. Indeed, if $\bar{z}^i = \bar{z}^i(z)$ is another coordinate system on $\mathcal{M}$, (A.4a) gives the change of basis in $T_z\mathcal{M}$, and analogously we can find the expression for the basis component (A.7) $\partial/\partial z^i$ and $\partial/\partial w^i$ in the new coordinates:

$$\frac{\partial}{\partial z^i} = \frac{\partial \bar{z}^j}{\partial z^i}(z)\frac{\partial}{\partial \bar{z}^j} + \frac{\partial \overline{w}^j}{\partial z^i}(w, z)\,\frac{\partial}{\partial \overline{w}^j}, \tag{A.15a}$$

$$\frac{\partial}{\partial w^i} = \frac{\partial \overline{w}^j}{\partial w^i}(w, z)\,\frac{\partial}{\partial \overline{w}^j}. \tag{A.15b}$$

But recalling (A.4b) we have:

$$\frac{\partial \overline{w}^j}{\partial w^i} = \frac{\partial \bar{z}^j}{\partial z^i},$$

so that (A.15b) becomes:

$$\frac{\partial}{\partial w^i} = \frac{\partial \bar{z}^j}{\partial z^i}(z)\,\frac{\partial}{\partial \overline{w}^j}. \tag{A.16}$$

Comparing (A.4a) and (A.16) we have that $\mathcal{I}$ in (A.14) is independent from the coordinates used.

Remark A.4. Relation (A.15b) also says that the vertical vector fields—linear combinations of $\partial/\partial w^i$—are intrinsically defined. We cannot say the same for linear combinations of $\partial/\partial z^i$, because, under coordinate changes, a vertical term arises in (A.15a).

Nevertheless, there exists the injection $v : VT\mathcal{M} \hookrightarrow TT\mathcal{M}$ (A.12), such as its dual counterpart, the projection

$$T^*T\mathcal{M} \xrightarrow{v^*} V^*T\mathcal{M} \tag{A.17}$$

from the cotangent bundle to the vertical bundle of $T\mathcal{M}$. This projection will be used later.

Remark A.5. We have so far shown how $TT\mathcal{M}$ can be projected on $T\mathcal{M}$ in two ways, namely using either $\pi_{T\mathcal{M}}$ (A.6) or $T\pi_{\mathcal{M}}$ (A.8). There exists a canonical involution ι on $TT\mathcal{M}$ – namely a map on $TT\mathcal{M}$ such that $\iota^2 = \mathrm{id}_{TT\mathcal{M}}$ –

$$\iota : TT\mathcal{M} \to TT\mathcal{M} \tag{A.18}$$

such that the following diagram

$$\begin{tikzcd} TT\mathcal{M} \arrow[r, "\iota"] \arrow[d, "\pi_{T\mathcal{M}}"'] & TT\mathcal{M} \arrow[d, "T\pi_{\mathcal{M}}"] \\ T\mathcal{M} \arrow[r, "\mathrm{id}_{T\mathcal{M}}"] & T\mathcal{M} \end{tikzcd}$$

is commutative ($\mathrm{id}_{T\mathcal{M}}$ is the identity map on $T\mathcal{M}$). Its coordinate expression reads

$$\iota : \left(u^i, y^i, w^i, z^i\right) \to \left(u^i, w^i, y^i, z^i\right). \tag{A.19}$$

Once we have set up the framework, to better understand how the objects can be intrinsically defined, we will begin from *autonomous* system, (that is, time-independent), and then we will extend to the case when time enters in the expression of either the Lagrangian or the constraint equation.

Then, let us take into account a general C^2 real function defined in $T\mathcal{M}$,

$$f : T\mathcal{M} \to \mathbb{R}.$$

Its differential $\mathrm{d}f$ is a map

$$\mathrm{d}f : T\mathcal{M} \to T^*T\mathcal{M}, \tag{A.20}$$

such that, $\forall z \in \mathcal{M}$ and $w \in T_z\mathcal{M}$, $\mathrm{d}f(w,z)$ is a linear function on $T_{(w,z)}T\mathcal{M}$.

We consider a vector field $W:\mathcal{M}\to T\mathcal{M}$ on $\mathcal{M}$. Its tangent map TW is an application

$$TW:T\mathcal{M}\to TT\mathcal{M} \tag{A.21}$$

that in coordinates reads:

$$\big(w^i,z^i\big)\xrightarrow{TW}\left(\frac{\partial W^i}{\partial z^j}(z)w^j,\,w^i,\,W^i(z),\,z^i\right),$$

and then it is not a vector field on $T\mathcal{M}$, since $\pi_{T\mathcal{M}}\circ TW(z,w)$ is $W(z)$, and not the identity map on $T\mathcal{M}$. But, applying the canonical involution $\imath$ (A.18) to TW we obtain a true vector field on $T\mathcal{M}$, whose coordinate expression reads:

$$\big(w^i,z^i\big)\xrightarrow{\imath\circ TW}\left(\frac{\partial W^i}{\partial z^j}(z)w^j,\,W^i(z),\,w^i,\,z^i\right), \tag{A.22a}$$

and now $\pi_{T\mathcal{M}}\circ(\imath\circ TW(z,w))$ is the identity map on $T\mathcal{M}$. Note that we can also write:

$$\imath\circ TW(w,z)=W^i(z)\frac{\partial}{\partial z^i}+\left(\frac{\partial W^i}{\partial z^j}(z)w^j\right)\frac{\partial}{\partial w^i}. \tag{A.22b}$$

We define:

$$\begin{aligned}&\mathrm{d}f[W]:T\mathcal{M}\to\mathbb{R},\\&\mathrm{d}f[W](w,z)=\mathrm{d}f(w,z)\big[\imath\circ TW(w,z)\big].\end{aligned} \tag{A.23}$$

Its coordinate expression is:

$$\mathrm{d}f[W](w,z)=\frac{\partial f}{\partial z^i}(z)W^i(z)+\frac{\partial f}{\partial w^i}(z)\left(\frac{\partial W^i}{\partial z^j}(z)w^j\right). \tag{A.24}$$

We can also define a fiber derivative of f in the following way. From (A.20), we can project $\mathrm{d}f$ on $V^*T\mathcal{M}$, taking into account (A.17), obtaining

$$v^*\mathrm{d}f:T\mathcal{M}\to V^*T\mathcal{M}. \tag{A.25}$$

Moreover, given a vector field $W:\mathcal{M}\to T\mathcal{M}$, and using the canonical isomorphism $\mathcal{I}$ (A.14) (dropping the subscript $_{(w,z)}$ to lighten the notation), we define:

$$\begin{aligned}&\frac{\partial f}{\partial w}[W]:T\mathcal{M}\to\mathbb{R},\\&\frac{\partial f}{\partial w}[W](w,z)=v^*\mathrm{d}f(w,z)\big[\mathcal{I}\big(W(z)\big)\big].\end{aligned} \tag{A.26}$$

Its coordinate expression simply reads:

$$\frac{\partial f}{\partial w}[W](w,z)=\frac{\partial f}{\partial w^i}(z)W^i(z). \tag{A.27}$$

It is clear that we can define in the same way higher order fiber derivatives. For instance, given two vector fields W_1, W_2 on $\mathcal{M}$, we can consider the second-order fiber derivative $\frac{\partial^2 f}{\partial w^2}[W_1, W_2]$ whose coordinate expression is:

$$\frac{\partial^2 f}{\partial w^2}[W_1,W_2](w,z)=\frac{\partial^2 f}{\partial w^i \partial w^j}(w,z)W_1^i(z)W_2^j(z).$$

All these objects can be naturally extended when we deal functions f defined on $T\mathcal{M}\times\mathbb{R}$ (as it happens, for instance, for *non-autonomous* systems), and with time-dependent vector fields $V: T\mathcal{M}\times\mathbb{R}\to T\mathcal{M}$.

Taking into account the natural projections:

$$p_{T\mathcal{M}}: T\mathcal{M}\times\mathbb{R}\to T\mathcal{M}, \tag{A.28}$$

$$T^*p_{T\mathcal{M}}: T^*(T\mathcal{M}\times\mathbb{R})\to T^*T\mathcal{M}, \tag{A.29}$$

and the canonical injection

$$\jmath: T\mathcal{M}\times\mathbb{R}\to T(\mathcal{M}\times\mathbb{R}) \tag{A.30}$$

whose coordinate expression is:

$$\jmath:\left(z^i\frac{\partial}{\partial z^i},t\right)\to\frac{\partial}{\partial t}+z^i\frac{\partial}{\partial z^i}, \tag{A.31}$$

we define:

$$\begin{aligned}&\mathrm{d}f[V]: T\mathcal{M}\times\mathbb{R}\to\mathbb{R},\\&\mathrm{d}f[V](w,z,t)=(T^*p_{T\mathcal{M}}\circ\mathrm{d}f)(w,z,t)\big[\imath\circ TV\circ\jmath(w,z,t)\big],\end{aligned} \tag{A.32}$$

and

$$\begin{aligned}&\frac{\partial f}{\partial w}[V]: T\mathcal{M}\times\mathbb{R}\to\mathbb{R},\\&\frac{\partial f}{\partial w}[V](w,z,t)=v^*(T^*p_{T\mathcal{M}}\circ\mathrm{d}f)(w,z,t)\big[\mathcal{I}\big(V(z,t)\big)\big].\end{aligned} \tag{A.33}$$

Their coordinate expression respectively reads:

$$\mathrm{d}f[V](w,z,t)=\frac{\partial f}{\partial z^i}(z)V^i(z,t)+\frac{\partial f}{\partial w^i}(z)\left(\frac{\partial V^i}{\partial t}(z,t)+\frac{\partial V^i}{\partial z^j}(z,t)w^j\right), \tag{A.34}$$

and

$$\frac{\partial f}{\partial w}[V](w,z,t)=\frac{\partial f}{\partial w^i}(z)\,V^i(z,t). \tag{A.35}$$

Example A.6. Let $V=\lambda(t)\,Y(z)$, where $\lambda:\mathbb{R}\to\mathbb{R}$ and Y is a vector field on $\mathcal{M}$. Then

$$\begin{aligned}
&\mathrm{d}f[V](w,z,t)\\
&\quad=\lambda(t)\,\frac{\partial f}{\partial z^i}(w,z,t)Y^i(z)+\frac{\partial f}{\partial w^i}(w,z,t)\left(\dot\lambda(t)\,Y^i(z)+\lambda(t)\,\frac{\partial Y^i}{\partial z^j}(z)\,w^j\right)\\
&\quad=\lambda(t)\,\mathrm{d}f[Y](w,z,t)+\dot\lambda(t)\,\frac{\partial f}{\partial w}[Y](w,z,t).
\end{aligned} \tag{A.36}$$

Example A.7. Let $z:[0,1]\to\mathcal{M}$ be a curve on $\mathcal{M}$, and

$$c:[0,1]\subset\mathbb{R}\to T\mathcal{M},\qquad c(t)=\big(\xi^i(t),z^i(t)\big),$$

a vector field on $\mathcal{M}$ along $z(t)$. Its *lift* to the tangent space of $T\mathcal{M}$ is given by

$$\dot c:[0,1]\to TT\mathcal{M},\qquad c(t)=\big(\dot\xi^i(t),\dot z^i(t),\xi^i(t),z^i(t)\big).$$

Applying the involution $\imath$ (A.18) to $\dot c$ we obtain a vector field on $T\mathcal{M}$ over $(\dot z(t),z(t))$, whose coordinate expression is:

$$\imath\circ\dot c=\xi^i(t)\left(\frac{\partial}{\partial z^i}\circ(\dot z,z)\right)+\dot\xi^i(t)\left(\frac{\partial}{\partial w^i}\circ(\dot z,z)\right). \tag{A.37}$$

Thus, given a function f on $T\mathcal{M}$ or on $T\mathcal{M}\times\mathbb{R}$, we can apply $\mathrm{d}f$ to $\imath\circ\dot c$, exactly as done for W in (A.24) and for V in (A.32). Analogously can be done taking into account $\partial f/\partial w$ instead of $\mathrm{d}f$.

References

[1] R. Abraham, J.E. Marsden, Foundation of Mechanics, Benjamin, London, 1978.
[2] V.I. Arnol'd (Ed.), Dynamical System, III, Encyclopedia of Mathematical Sciences, Springer-Verlag, New York, 1988.
[3] J.K. Beem, P.E. Ehrlich, K.L. Easly, Global Lorentzian Geometry, Dekker, New York, 1996.
[4] H. Brezis, Analyse fonctionnelle, Masson Editeur, Paris, 1983.
[5] R. Caccioppoli, Un principio di inversione per le corrispondenze funzionali e sue applicazioni alle equazioni alle derivate parziali, Atti Accad. Naz. Lincei 16 (1932).
[6] E. Fadell, S. Husseini, Category of loop spaces of open subsets in Euclidean spaces, Nonlinear Anal. 17 (1991).
[7] R. Giambò, F. Giannoni, P. Piccione, Existence, multiplicity and regularity for sub-Riemannian geodesics by variational methods, SIAM J. Control Optim. 40 (2002).
[8] F. Giannoni, A. Masiello, P. Piccione, A variational theory for light rays in stably causal Lorentzian manifolds: Regularity and multiplicity results, Comm. Math. Phys. 187 (1997).

[9] F. Giannoni, A. Masiello, P. Piccione, A Morse theory for massive particles and photons in general relativity, J. Geom. Phys. 35 (2000).
[10] F. Giannoni, P. Piccione, An existence theory for relativistic brachistochrones in stationary space-times, J. Math. Phys. 39 (1998).
[11] F. Giannoni, P. Piccione, An intrinsic approach to the geodesical connectedness of stationary Lorentzian manifolds, Comm. Anal. Geom. 7 (1999).
[12] F. Giannoni, P. Piccione, The arrival time brachistochrones in a general relativistic spacetime, J. Geom. Anal. 12 (2002).
[13] S. Lang, Differential and Riemannian Manifolds, Springer-Verlag, New York, 1995.
[14] J. Mawhin, M. Willem, Critical Point Theory and Hamiltonian Systems, Springer-Verlag, New York, 1989.
[15] R.S. Palais, Foundations of Global Nonlinear Analysis, Benjamin, New York, 1968.
[16] A.M. Vershik, in: Lecture Notes in Math., Vol. 1108, Springer-Verlag, New York, 1986.