A coordinate-free theory of virtual holonomic constraints

JOURNAL OF GEOMETRIC MECHANICS doi:10.3934/jgm.2018018

Volume **10**, Number **4**, December **2018** pp. **X–XX**

A COORDINATE-FREE THEORY OF VIRTUAL HOLONOMIC CONSTRAINTS

Luca Consolini and Alessandro Costalunga

Dipartimento di Ingegneria dell'Informazione, Università di Parma
Parco Area delle Scienze 181/a, 43124 Parma, Italy

Manfredi Maggiore*

Department of Electrical and Computer Engineering, University of Toronto
10 King's College Road, Toronto, Ontario, M5S 3G4, Canada

(Communicated by Witold Respondek)

Abstract. This paper presents a coordinate-free formulation of virtual holonomic constraints for underactuated Lagrangian control systems on Riemannian manifolds. It is shown that when a virtual constraint enjoys a regularity property, the constrained dynamics are described by an affine connection dynamical system. The affine connection of the constrained system has an elegant relationship to the Riemannian connection of the original Lagrangian control system. Necessary and sufficient conditions are given for the constrained dynamics to be Lagrangian. A key condition is that the affine connection of the constrained dynamics be metrizable. Basic results on metrizability of affine connections are first reviewed, then employed in three examples in order of increasing complexity. The last example is a double pendulum on a cart with two different actuator configurations. For this control system, a virtual constraint is employed which confines the second pendulum to within the upper half-plane.

1. **Introduction.** A virtual holonomic constraint (VHC) for a Lagrangian control system is a collection of relations among the configuration variables of the system that can be made invariant via feedback control. The precise meaning of this terminology is clarified in what follows, but the key idea is to emulate via feedback control the presence of a holonomic constraint in the Lagrangian control system. By appropriate design of the VHC, the constrained system may display useful properties.

The notion of VHC can be traced back to early twentieth century work of P. Appell in [2] and H. Beghin in [5], but it has emerged prominently in the last fifteen years as a tool for control of biped robots (see, e.g., [24, 27, 38, 37, 8]), and as an approach to motion planning for general robotic systems (e.g., [31, 32, 33, 13]). In VHC-based robot control, the motion one wants to induce is represented implicitly in terms of constraints on the robot's configuration variables, and the control loop is designed to asymptotically stabilize a subset of the state space, the

2010 *Mathematics Subject Classification.* 53B05, 93C10, 37Jxx.

Key words and phrases. Mechanical control systems, virtual holonomic constraints, inverse Lagrangian problem, geometric mechanics.

Some of the ideas of this paper appeared in preliminary form in [11].

* Corresponding author.

so-called constraint manifold. This control philosophy stands in contrast to the standard technique of parametrizing a desired motion by time, and then stabilizing the resulting reference signals. The VHC control paradigm has proved particularly effective in inducing complex behaviours in underactuated robots, and gives rise to a feedback loop that is intrinsically robust because it is not driven by any exogenous signal.

For biped robots, Grizzle and collaborators (see, e.g., [37]) defined VHCs in terms of invariance of a submanifold of the state space. The paper [21] generalized Grizzle's notion of VHC to mechanical control systems whose generalized coordinates are linear displacements or angles (i.e., systems whose configuration manifold is a generalized cylinder) and whose degree of underactuation is equal to one. For this class of systems, the authors of [22, 23] showed that, generically, the constrained dynamics in the presence of a VHC do *not* possess a Lagrangian structure. They then gave necessary and sufficient conditions for a Lagrangian structure to exist. The theory of [21, 23] does not handle mechanical systems whose configuration space is not a generalized cylinder, or whose degree of underactuation is greater than one. To illustrate, the configuration manifold of a rigid body is $\mathsf{SE}(3)$, a manifold that cannot be handled by the theory in [21, 23]. Similarly, a double pendulum on a cart has degree of underactuation two, which again is not contemplated in the theory of [21, 23].

Main contributions. This paper generalizes the theory in [21, 23] by presenting a coordinate-free formulation of VHCs for arbitrary configuration manifolds and arbitrary degrees of underactuation. We give a new geometric definition of VHC, and define a regularity property of VHCs in terms of transversality of two subbundles. We show that a regular VHC induces on the constraint manifold an affine connection, the so-called *induced connection.* In the absence of a potential function, orbits of the constrained dynamics are geodesics of this induced connection. We give an explicit characterization of the constrained dynamics in coordinates with formulas for the Christoffel symbols of the induced connection. We show that the problem of determining whether or not the constrained dynamics are Lagrangian amounts in great part to determining whether or not the induced connection is metrizable, i.e., it is Riemannian for a suitable metric. Leveraging this insight, and using existing results from the theory of affine connections, we give conditions for the existence of a Lagrangian structure for the constrained dynamics arising from a regular VHC. These conditions are applicable to Lagrangian control systems with arbitrary degree of underactuation. In the special case when the subbundle associated with the control accelerations is orthogonal to the constraint manifold, the constrained dynamics are always Lagrangian, and we show that they coincide with the dynamics one would obtain in the presence of an ideal holonomic constraint. Thus, the classics mechanics notion of holonomic constraint is a special case of our theory. For systems with underactuation degree one, our results provide an elegant geometric insight for the results in [23].

The focus of this paper is on the case when all control inputs are used to enforce the VHC, so that the constrained dynamics are unforced. In the more general case when the constrained dynamics are forced, the question of existence of a Lagrangian structure for the constrained dynamics turns into the more general question of feedback equivalence of the constrained dynamics to a Lagrangian control system. A local version of the latter question has been investigated in [29] for general control systems on smooth manifolds near zero velocity points.

Organization of the paper. In Section 2 we review concepts of Riemannian geometry, and the definition from [7] of a Lagrangian (control) system on a Riemannian manifold. Section 3 reviews the definition of VHC from [21] and the Lagrangian properties of the constrained dynamics from [23], valid for the case of systems with degree of underactuation one. Section 4 formulates a new coordinate-free theory of VHCs, characterizing the regularity of VHCs in terms of transversality of the constraint manifold and the distribution induced by control forces. It is shown that a VHC induces an affine connection on the constraint manifold, and this connection is then used to characterize the constrained dynamics. In Section 5 we give necessary and sufficient conditions under which the constrained dynamics are Lagrangian. We also treat the special case when the distribution induced by control forces is orthogonal to the constraint manifold. In Section 6 we give a tutorial overview of holonomy groups and results on metrizability of affine connections, treating the special cases of flat connections, of simply connected constraint manifolds, and one and two-dimensional constraint manifolds. Here we show that the results of Section 3 are a special case of the general theory of this paper. In Section 7, we present three examples illustrating the theory, in order of increasing complexity. The last example is a double pendulum on a cart with a VHC that constrains the angle of the second pendulum to be a function of the angle of the first pendulum, in such a way that the second pendulum is always confined to the upper half plane.

2. Preliminaries.

In this section we present the notation used in this paper, review notions of Riemannian geometry, and review the definition of a Lagrangian (control) system on a Riemannian manifold. All results are found in [19, 12, 6, 7].

Smooth manifolds. If $\mathcal{M}$ is a smooth manifold, we denote by $C^\infty(\mathcal{M})$ the ring of smooth real-valued functions on $\mathcal{M}$, by $\mathfrak{X}(\mathcal{M})$ the set of smooth vector fields on $\mathcal{M}$, and by $\Omega(\mathcal{M})$ the set of smooth one-forms on $\mathcal{M}$. The tangent space to $\mathcal{M}$ at $p \in \mathcal{M}$ is denoted by $T_p\mathcal{M}$, and it dual, the cotangent space, is denoted by $T_p^\star\mathcal{M}$. We denote by $T\mathcal{M}$ and $T^\star\mathcal{M}$ the tangent and cotangent bundles of $\mathcal{M}$, and by $\pi : T\mathcal{M} \to \mathcal{M}$ the natural projection on $T\mathcal{M}$. An element of $T\mathcal{M}$ will be denoted by v_q, with the understanding that $v_q \in T_q\mathcal{M}$. If $\mathcal{N}$ is a submanifold of $\mathcal{M}$, $T\mathcal{M}|_\mathcal{N}$ denotes the restriction of $T\mathcal{M}$ to $\mathcal{N}$, defined as $T\mathcal{M}|_\mathcal{N} = \bigcup_{p\in\mathcal{N}} T_p\mathcal{M}$.

If $(U, (x^1, \dots, x^n))$ is a coordinate chart of $\mathcal{M}$, for each $p \in U$ the basis for $T_p\mathcal{M}$ induced by the chart is denoted by $\partial/\partial x^1|_p, \dots, \partial/\partial x^n|_p$. The vector fields $\{\partial/\partial x^1, \dots, \partial/\partial x^n\}$ form a *local frame* for $T\mathcal{M}$. If $F : \mathcal{M} \to \mathcal{N}$ is a smooth function between smooth manifolds and $p \in \mathcal{M}$, we let $F_p := F(p)$, and we denote by $dF_p : T_p\mathcal{M} \to T_{F(p)}\mathcal{N}$ the differential of F at p. If $F : U \to V$ is a smooth function and $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ are open sets, we denote by $\partial_{x^i}F$ the partial derivative F with respect to its i-th argument. The notation $\partial^2_{x^i x^j}F$ indicates second-order partial differentiation with respect to the i-th and j-th argument. More generally, if $U \subset \mathbb{R}^n$ is an open set and $F : U \to F(U) \subset M$ is smooth, then we denote $\partial_{x^i}F := dF_x(\partial/\partial x^i)$, where $\partial/\partial x^i$ denotes the i-th natural basis vector of $T_x\mathbb{R}^n$. In the special case of a function of one variable in $\mathbb{R}$ or $\mathbb{S}^1$, we let $F'(x) := \partial_x F$ and $F''(x) := \partial_x^2 F$.

A smooth function $h : \mathcal{M} \to \mathbb{R}^k$ is a *submersion* if $\operatorname{rank} dh_p = k$ for all $p \in \mathcal{M}$. If $\operatorname{rank} dh_p = k$ for all $p \in h^{-1}(0)$, then we say that 0 is a *regular value of* h. If $f \in C^\infty(\mathcal{M})$ and $X \in \mathfrak{X}(\mathcal{M})$, the *Lie derivative* of f along X is the smooth function $X(f) \in C^\infty(\mathcal{M})$ defined as $p \mapsto X(f)(p) := df_p(X(p))$. If $X, Y \in \mathfrak{X}(\mathcal{M})$, $[X, Y] \in \mathfrak{X}(\mathcal{M})$ denotes the *Lie Bracket* of X and Y.

Riemannian manifolds and connections. A *Riemannian manifold* is a pair $(\mathcal{M}, g)$, where $\mathcal{M}$ is a smooth manifold, and $g : T\mathcal{M} \times T\mathcal{M} \to \mathbb{R}$, the *Riemannian metric*, is a smooth function such that, for each $p \in \mathcal{M}$, g_p is a bilinear form $T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$ which is symmetric and positive definite, i.e., for each $v_p, w_p \in T_p\mathcal{M}$, $g_p(v_p, w_p) = g_p(w_p, v_p)$, and the function $v_p \mapsto g_p(v_p, v_p)$ is positive definite. Thus, g_p is an inner product on $T_p\mathcal{M}$ which varies smoothly with p. In the language of tensors, g is a type $(0, 2)$ symmetric and positive definite tensor field on M. A Riemannian metric induces two maps. The *flat* map is the function $T\mathcal{M} \to T^\star\mathcal{M}$, $X \mapsto X^\flat$, defined as $X^\flat(Y) = g(X, Y)$ for all $Y \in T\mathcal{M}$. The *sharp* map is the function $T^\star\mathcal{M} \to T\mathcal{M}$, $\omega \mapsto \omega^\sharp$, defined uniquely through the identity $\omega(X) = g(\omega^\sharp, X)$ for all $X \in T\mathcal{M}$. Given a function $f \in C^\infty(\mathcal{M})$, $\operatorname{grad} f : \mathcal{M} \to T\mathcal{M}$ is the smooth vector field defined as $\operatorname{grad} f := df^\sharp$.

An *affine connection* on $\mathcal{M}$ is a smooth function $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$, $(X, Y) \mapsto \nabla_X Y$ satisfying the following properties:

$$\begin{aligned} &\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z, \\ &\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z, \\ &\nabla_X (fY) = f\nabla_X Y + X(f)Y, \end{aligned} \tag{1}$$

for any $f, g \in C^\infty(\mathcal{M})$ and $X, Y, Z \in \mathfrak{X}(\mathcal{M})$. The vector field $\nabla_X Y$ is called the *covariant derivative* of Y in the direction of X. The covariant derivative of vector fields induces a covariant derivative of tensor fields, also denoted ∇, enjoying the properties listed in [19, Lemma 4.6]. Among them, we mention the following. If F is a tensor field on $\mathcal{M}$ of type $(0, s)$, and $X \in \mathfrak{X}(\mathcal{M})$, then $\nabla_X F$ is a type $(0, s)$ tensor field satisfying the following identity

$$(\nabla_X F)(Y_1, \ldots, Y_s) = X(F(Y_1, \ldots Y_s)) - \sum_{j=1}^{s} F(Y_1, \ldots, \nabla_X Y_j, \ldots, Y_s), \tag{2}$$

for all $Y_1, \ldots, Y_s \in \mathfrak{X}(\mathcal{M})$. The *total covariant derivative* of a type $(0, s)$ tensor field F is the type $(0, s + 1)$ tensor field ∇F given by

$$\nabla F(X, Y_1, \ldots, Y_s) = \nabla_X F(Y_1, \ldots, Y_s), \quad \text{for all } X, Y_1, \ldots, Y_s \in \mathfrak{X}(\mathcal{M}). \tag{3}$$

The connection ∇ is *symmetric* (or torsionless) if

$$(\forall X, Y \in \mathfrak{X}(\mathcal{M}))\ \nabla_X Y - \nabla_Y X = [X, Y].$$

The connection ∇ is *compatible with* g if

$$(\forall X, Y, Z \in \mathfrak{X}(\mathcal{M}))\ X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \tag{4}$$

or, in terms of the total covariant derivative,

$$\nabla g = 0. \tag{5}$$

The Fundamental Lemma of Riemannian Geometry (e.g., [19]) states that there is a unique affine connection ∇ on a Riemannian manifold $(\mathcal{M}, g)$ with the property of being symmetric and compatible with g. This connection is called the *Riemannian connection* or the *Levi-Civita connection* of g.

The covariant derivative $\nabla_X Y$ may be viewed as a differentiation of the vector field Y along X. If $\gamma(t)$ is a smooth curve on $\mathcal{M}$ and $Y \in \mathfrak{X}(\mathcal{M})$, the restriction of Y to $\gamma(t)$, $V(t) := Y(\gamma(t))$, is a vector field along γ. The derivative $D_t V := \nabla_{\dot\gamma} Y$ is called the *covariant derivative of* V *along* γ. Although the definition just given relies on expressing V as the restriction to γ of a vector field Y on $\mathcal{M}$, $D_t V$ does not depend on the values of Y outside of $\gamma(t)$, in that *any* smooth extension of V outside of γ gives the same value of D_tV. The geometric intuition of the notion of covariant derivative is as follows (see, e.g., [6, Chapter VII, Section 2]). If $\mathcal{M}$ is an embedded submanifold of $\mathbb{R}^n$ with Riemannian metric induced from an inner product on $\mathbb{R}^n$, D_tV is the orthogonal projection of the time derivative of $Y(\gamma(t))$ onto the tangent space $T_{\gamma(t)}\mathcal{M}$. Thus, roughly speaking, D_tV measures how much the vector field $V(t)$ turns as seen from the point of view of $\mathcal{M}$. In essence, covariant derivatives embody the notion of acceleration of a curve. More precisely, the *acceleration* of a curve γ on $\mathcal{M}$ is the vector field $D_t\dot{\gamma}$ along γ, and γ is called a *geodesic* of ∇ if its acceleration is zero, i.e., $\nabla_{\dot{\gamma}}\dot{\gamma}(t) = D_t\dot{\gamma}(t) \equiv 0$. We remark that this definition of geodesic curve does not require ∇ to be a Riemannian connection.

Coordinate representation of the covariant derivative. In coordinates, covariant derivatives associated with a Riemannian connection take on a familiar form, which we now review. Consider a coordinate chart $(U, (x^1, \dots, x^n))$ on $\mathcal{M}$ and the associated local frame $\{\partial/\partial x^1, \dots, \partial/\partial x^n\}$ for $T\mathcal{M}$. Let $X_i := \partial/\partial x^i$. Given an affine connection ∇ on $\mathcal{M}$, not necessarily Riemannian, the *Christoffel symbols* of ∇ associated with the local frame $\{X_1, \dots, X_n\}$ are the n^3 functions Γ^k_{ij} in $C^\infty(U)$ that are coefficients of the expansion of $\nabla_{X_i}X_j$ in the local frame $\{X_1, \dots, X_n\}$, i.e.,

$$\nabla_{X_i}X_j = \sum_{k=1}^{n} \Gamma^k_{ij} X_k.$$

One can show that if ∇ is symmetric, then $\Gamma^k_{ij} = \Gamma^k_{ji}$. Whether or not ∇ is symmetric, if $Y, Z \in \mathfrak{X}(\mathcal{M})$, expanding $Y = \sum_i y_i X_i$, $Z = \sum_i z_i X_i$, with $y_i, z_i \in C^\infty(U)$, the covariant derivative $\nabla_X Y$ can be computed through the formula

$$\nabla_Y Z = \sum_k \left(Y(z_k) + \sum_{i,j} \Gamma^k_{ij} y_i z_j \right) X_k, \tag{6}$$

where $Y(z_k)$ is the Lie derivative of z_k along Y. The acceleration of a smooth curve $\gamma : I \to \mathcal{M}$, $I \subset \mathbb{R}$, can be computed as follows. Letting $\gamma^i(t) := x^i(\gamma(t))$ denote the i-th component of the coordinate representation of γ, we have

$$\nabla_{\dot{\gamma}}\dot{\gamma} = \sum_k \left(\ddot{\gamma}^k + \sum_{i,j} \Gamma^k_{ij} \dot{\gamma}^i \dot{\gamma}^j \right) X_k. \tag{7}$$

We see that, in local coordinates, geodesics are solutions of the system of second-order differential equations

$$\ddot{\gamma}^k = -\sum_{i,j} \Gamma^k_{ij} \dot{\gamma}^i \dot{\gamma}^j, \ k = 1, \dots, n.$$

If ∇ is Riemannian, the Christoffel symbols may be computed using a matrix representation of the metric g. Using again the local frame $\{X_1, \dots, X_n\}$, let $g_{ij}(p) := g(\partial/\partial x^i|_p, \partial/\partial x^j|_p)$, and let g^{kl} be (k,l)-th element of the inverse of the matrix (g_{ij}). Then,

$$\Gamma^k_{ij} = \frac{1}{2} \sum_l g^{kl} \big(X_i(g_{jl}) + X_j(g_{il}) - X_l(g_{ij}) \big). \tag{8}$$

Lagrangian control systems on manifolds. Having reviewed basic notions of Riemannian geometry, we are ready to present the class of mechanical systems considered in this paper. The definitions below are adapted from [7].

Definition 2.1 (Lagrangian system)**.** A *Lagrangian system* is a triple $(\mathcal{Q}, g, P)$, where $(\mathcal{Q}, g)$ is an n-dimensional Riemannian manifold called the *configuration manifold*, and $P : \mathcal{Q} \to \mathbb{R}$ is a smooth function called the *potential function*. The triple $(\mathcal{Q}, g, P)$ is also called a *Lagrangian structure*. A smooth curve $q : I \to \mathcal{Q}$, where I is an open interval in $\mathbb{R}$, is a *base integral curve* of the Lagrangian system if

$$\nabla_{\dot{q}(t)}\dot{q}(t) = -\operatorname{grad} P(q(t)), \tag{9}$$

for all $t \in I$. △

For each $q_0 \in \mathcal{Q}$ and each $v_{q_0} \in T_{q_0}\mathcal{Q}$, there exists a unique maximal base integral curve $q(t)$ such that $q(0) = q_0$ and $\dot{q}(0) = v_{q_0}$, where maximality is defined in the same way as for integral curves of vector fields. We will call $q(t)$ the *maximal base integral curve of* (9) *with initial condition* (q_0, v_{q_0}).

Base integral curves have the property of being extremizers of the action functional $\int_I L(q(t), \dot{q}(t))dt$, $I \subset \mathbb{R}$, where $L : T\mathcal{Q} \to \mathbb{R}$ is the Lagrangian function defined as

$$L(q, \dot{q}) := \frac{1}{2} g_q(\dot{q}, \dot{q}) - P(q). \tag{10}$$

In Lagrangian systems, controls appear by way of forces. On Riemannian manifolds, forces are modelled as one-forms because, under coordinate changes, they transform like the components of one-forms (see [7]). Thus, a force on $\mathcal{Q}$ is a one-form $F \in \Omega(\mathcal{Q})$. The corresponding vector field $F^\sharp \in \mathfrak{X}(\mathcal{Q})$ can be thought of as the portion of the acceleration due to F. With this in mind, we proceed to the definition of a Lagrangian control system.

Definition 2.2 (Lagrangian control system)**.** A *Lagrangian control system* is a quadruple $(\mathcal{Q}, g, P, \mathsf{F})$, where $(\mathcal{Q}, g, P)$ is a Lagrangian system and $\mathsf{F} = \{F^1, \ldots, F^m\}$, $F^i \in \Omega(\mathcal{Q})$, are called the *control forces*. A curve $q : I \to \mathcal{Q}$, where I is an open interval in $\mathbb{R}$, is a *base integral curve* of the Lagrangian control system if there exist m smooth functions $\tau_i : I \to \mathbb{R}$ such that

$$\nabla_{\dot{q}(t)}\dot{q}(t) = -\operatorname{grad} P(q(t)) + \sum_{i=1}^{m} (F^i)^\sharp_{q(t)} \tau_i(t), \tag{11}$$

for all $t \in I$. △

Let $\tau^\star = (\tau_1^\star, \ldots, \tau_m^\star) : T\mathcal{Q} \to \mathbb{R}^m$ be a smooth map. Then, for each $q_0 \in \mathcal{Q}$ and each $v_{q_0} \in T_{q_0}\mathcal{Q}$, there exists a unique maximal solution $q : I \to \mathcal{Q}$ of (11) with $\tau_i(t) = \tau_i^\star(q(t), \dot{q}(t))$, $i = 1, \ldots, m$. We call it the *maximal base integral curve of* (11) *with feedback* $\tau = \tau^\star(q, \dot{q})$ *and initial condition* (q_0, v_{q_0}).

It is shown in [7] that the equations of motion of the Lagrangian system in (9) can be equivalently expressed as a smooth vector field on $T\mathcal{Q}$. Similarly, the equations of motion of the Lagrangian control system in (11) have an equivalent representation as a smooth control-affine system with state space $T\mathcal{Q}$. In order to define such control-affine system, we need the notion of *vertical lift* at a point $v_q \in T\mathcal{Q}$ ([7]), the linear map $\operatorname{vlft}_{v_q} : T_q\mathcal{Q} \to T_{v_q}T\mathcal{Q}$ defined by

$$\operatorname{vlft}_{v_q}(X_q) = \frac{d}{dt}\Big|_{t=0} (v_q + tX_q). \tag{12}$$

The vertical lift of a vector field $X \in \mathfrak{X}(\mathcal{Q})$ is the vector field $\operatorname{vlft}(X)$ on $T\mathcal{Q}$ defined by $\operatorname{vlft}(X)(v_q) = \operatorname{vlft}_{v_q}(X(q))$. The vertical lift of a distribution $\mathcal{D}$ on $\mathcal{Q}$ is the distribution $\operatorname{vlft}(\mathcal{D})$ on $T\mathcal{Q}$ defined by $\operatorname{vlft}(\mathcal{D})(v_q) = \operatorname{vlft}_{v_q}(\mathcal{D}(q))$.

The control-affine system on $T\mathcal{Q}$ associated with the Lagrangian control system (11) is given by

$$\dot{X} = S(X) - \mathrm{vlft}(\mathrm{grad}\, P)(X) + \sum_{i=1}^{m} \tau_i \, \mathrm{vlft}\left((F^i)^\sharp\right)(X). \tag{13}$$

In the above, the vector field $S \in \mathfrak{X}(TQ)$ is the *geodesic spray* associated with the metric g, and it has the property that the integral curves of S project to geodesics of the metric g on $\mathcal{Q}$ via the canonical projection $\pi : T\mathcal{Q} \to \mathcal{Q}$. The control system (13) is equivalent to (11) in the following sense. All maximal base integral curves of (11) are projections under $\pi : T\mathcal{Q} \to \mathcal{Q}$ of maximal integral curves of (13). Vice versa, the projection of any maximal integral curve of (13) is a maximal base integral curve of (11).

In coordinates, equations (9) and (11) take on the familiar form of the Euler-Lagrange equations (e.g., [3, Chapter 3]). More precisely, let $q : I \to \mathcal{Q}$ be a base integral curve of a Lagrangian system $(\mathcal{Q}, g, P)$. Given a chart (U, ϕ) for $\mathcal{Q}$, let $x = \phi \circ q$ be the coordinate representation of q, and let $\hat{L}(x, \dot{x}) := L(\phi^{-1}(x), (d\phi^{-1})_x \dot{x})$, where L is the Lagrangian function defined in (10). Then, $x = (x^1, \dots, x^n) : I \to \mathbb{R}^n$ satisfies the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial \hat{L}}{\partial \dot{x}^i} - \frac{\partial \hat{L}}{\partial x^i} = 0, \; i = 1, \dots, n.$$

Vice versa, if $x : I \to \mathbb{R}^n$ satisfies the above Euler-Lagrange equations, then $q : I \to \mathcal{Q}$, $q := \phi^{-1} \circ x$ is a base integral curve of $(\mathcal{Q}, g, P)$. An analogous property holds for Lagrangian control systems $(\mathcal{Q}, g, P, \mathsf{F})$, where now $q : I \to \mathcal{Q}$ is a base integral curve of $(\mathcal{Q}, g, P, \{F^i\}_{i=1,\dots,m})$ if, and only if, $x = \phi \circ q$ is a solution of the forced Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial \hat{L}}{\partial \dot{x}^j} - \frac{\partial \hat{L}}{\partial x^j} = \sum_i B_{ij}(x)\tau_i, \; j = 1, \dots, n. \tag{14}$$

In the above $B_{ij}(x) := F^j(\partial/\partial x^i|_{\phi^{-1}(x)})$ is the i-th coefficient of the expansion of F^j in the local frame $\{\partial/\partial x^1, \dots, \partial/\partial x^n\}$ induced by the chart. Let $D(x)$ be the matrix with components $D_{ij}(x) = g_{ij}(\phi^{-1}(x))$, and set $C(x, \dot{x}) = D(x)\tilde{C}(x, \dot{x})$, where $\tilde{C}_{kj}(x, \dot{x}) = \sum_i \Gamma^k_{ij}(\phi^{-1}(x))\dot{x}_i$. Let $\hat{P} := P \circ \phi^{-1}$, and let $B(x)$ be the matrix with elements $B_{ij}(x)$. Then the Euler-Lagrange equations (14) take on the familiar form

$$D(x)\ddot{x} + C(x, \dot{x})\dot{x} + \nabla_x \hat{P}(x) = B(x)\tau, \tag{15}$$

where τ is the vector whose components are the control inputs τ_i in (14).

3. **Review of virtual holonomic constraints.** The configuration manifold $\mathcal{Q}$ of a robot whose joints are revolute or prismatic is a generalized cylinder. In other words, an element of $\mathcal{Q}$ may be represented as an n-tuple $(q_1, \dots, q_n)$, where each q_i is either in $\mathbb{R}$ if the i-th joint is prismatic, or in $\mathbb{S}^1$ if the i-th joint is revolute. In this case, the Lagrangian control system (11) admits a global coordinate representation of the form (15):

$$D(q)\ddot{q} + C(q, \dot{q})\dot{q} + \nabla_q \hat{P}(q) = B(q)\tau. \tag{16}$$

In this section, we review basic facts concerning VHCs for the class of mechanical control systems (16). The rest of this paper will be devoted to the generalization of these results to the coordinate-free setting. We assume, throughout, that the matrix B has full-rank m.

Definition 3.1 (Virtual holonomic constraint in coordinates, [21])**.** A *virtual holonomic constraint of order* k for system (16) is a relation $h(q) = 0$, where $h : \mathcal{Q} \to \mathbb{R}^k$ is a smooth function such that 0 is a regular value of h and the set

$$\Gamma = \{(q, \dot{q}) \in T\mathcal{Q} : h(q) = 0,\ dh_q \dot{q} = 0\} \tag{17}$$

is controlled invariant for (16). The set Γ is called the *constraint manifold* associated with the VHC $h(q) = 0$. △

Requiring Γ to be controlled invariant means requiring that there exists a smooth feedback $\tau = \tau^\star(q, \dot{q})$ rendering Γ positively invariant for the closed-loop system (see, e.g., [25, Definition 11.1]). If we let $\mathcal{C} := h^{-1}(0)$, then the hypothesis that $\operatorname{rank}(dh_q) = k$ for all $q \in h^{-1}(0)$ implies that $\mathcal{C}$ is a closed embedded submanifold of $\mathcal{Q}$. Moreover, the constraint manifold Γ in (17) is the tangent bundle of $\mathcal{C}$, $\Gamma = T\mathcal{C}$.

In the context of nonlinear control, the constraint manifold associated with a VHC $h(q) = 0$ is the zero dynamics manifold of system (16) with output $e = h(q)$ (see [15]). A special case of interest is when this output function yields a well-defined relative degree, as in the next definition.

Definition 3.2 (Regular VHC in coordinates, [21])**.** A relation $h(q) = 0$ is a *regular VHC of order* k for system (16) if $h : \mathcal{Q} \to \mathbb{R}^k$ is smooth, and system (16) with output $e = h(q)$ has vector relative degree[1] $\{2, \dots, 2\}$ everywhere on the constraint manifold Γ in (17), i.e., $\operatorname{rank}(dh_q D^{-1}(q) B(q)) = k$ for all $q \in h^{-1}(0)$. △

Since the matrix $dh_q D^{-1}(q) B(q)$ has dimension $k \times m$, for a regular VHC it must hold that the number of constraints, k, be less than or equal to the number of controls, m. The next result provides a geometric interpretation of the regularity condition.

Proposition 3.3 ([16])**.** *A relation $h : \mathcal{Q} \to \mathbb{R}^k$ is a regular VHC for system* (16) *if, and only if, letting $\mathcal{C} = h^{-1}(0)$,*

$$(\forall q \in \mathcal{C})\ T_q\mathcal{C} + \operatorname{Im}(D^{-1}(q)B(q)) = T_q\mathcal{Q}. \tag{18}$$

In light of the proposition above, the VHC $h(q) = 0$ is regular if, for each $q \in \mathcal{C}$, the mechanical system can produce control accelerations $D^{-1}B\tau$ in any direction transversal to $T_q\mathcal{C}$. In the special case $k = m$, when the number of constrains is equal to the number of controls, the subspace sum in (18) becomes direct.

Regular VHCs are important in two respects. First, since the output $e = h(q)$ yields a well-defined vector relative degree on the constraint manifold Γ, one may use input-output linearization to asymptotically stabilize Γ. For this, some technical assumptions on h and its differential are required, see [21]. Second, when $k = m$, there is a unique smooth feedback $\tau^\star : \Gamma \to \mathbb{R}^m$ rendering Γ invariant, resulting in constrained dynamics on Γ described by an autonomous differential equation. This is stated in the next proposition for the case[2] $m = n - 1$. In what follows, we let $B^\perp : \mathcal{Q} \to \mathbb{R}^{1\times n}\backslash\{0\}$ be a smooth left annihilator of B of rank one everywhere on $h^{-1}(0)$.

Proposition 3.4 ([21])**.** *Let $m = n - 1$, and let $h(q) = 0$ be a regular VHC of order $n - 1$ for system* (16)*. Then there exists a unique smooth feedback $\tau^\star : \Gamma \to$*

[1] This means [15] that the control input τ appears nonsingularly in the second time derivative, $\ddot{e}$, of e along solutions of (16).

[2] The more general case $k \leq m \leq n$ is addressed in [16].

$\mathbb{R}^m$ rendering Γ in (17) invariant, and the resulting closed-loop dynamics on Γ are given as follows. Let $\varphi : \Theta \to \mathcal{Q}$ be a regular parametrization of the curve $\mathcal{C} = h^{-1}(0)$, where $\Theta = \mathbb{S}^1$ if $\mathcal{C}$ is a Jordan curve and $\Theta = \mathbb{R}$ otherwise, and let $(q, \dot{q}) = (\varphi(s), \varphi'(s)\dot{s})$. The closed-loop dynamics on Γ are given by

$$\ddot{s} = \Psi_1(s) + \Psi_2(s)\dot{s}^2, \tag{19}$$

where $(s, \dot{s}) \in \Theta \times \mathbb{R}$ and

$$\begin{aligned}\Psi_1(s) &= -\frac{B^\perp \nabla_q \hat{P}}{B^\perp D\varphi'}\bigg|_{q=\varphi(s)},\\ \Psi_2(s) &= -\frac{B^\perp D\varphi'' + \sum_{i=1}^n (B^\perp D)_i \varphi'^\top \Gamma^i \varphi'}{B^\perp D\varphi'}\bigg|_{q=\varphi(s)},\end{aligned} \tag{20}$$

where $(\Gamma^i)_{jk} = \Gamma^i_{jk}$ is the matrix containing the Christoffel symbols of the metric $g_q(v_q, w_q) = v_q^\top D(q) w_q$.

The dynamics in (19) are called the *constrained (or reduced) dynamics* associated with the VHC $h(q) = 0$. The next result, taken from [23], characterizes when the constrained dynamics (19) possess a Lagrangian structure.

Theorem 3.5 ([23])**.** *Define a map $\pi : \mathbb{R} \to \Theta$ as*

$$\pi(x) = \begin{cases} x & \text{if } \Theta = \mathbb{R} \\ x \bmod 2\pi & \text{if } \Theta = \mathbb{S}^1. \end{cases}$$

Define smooth functions $\hat{M}_\mathcal{C}, \hat{P}_\mathcal{C} : \mathbb{R} \to \mathbb{R}$ as,

$$\begin{aligned}\hat{M}_\mathcal{C}(x) &:= \exp\left(-2\int_0^x \Psi_2 \circ \pi(z) dz\right),\\ \hat{P}_\mathcal{C}(x) &:= -\int_0^x \big(\Psi_1 \circ \pi(z)\big)\hat{M}_\mathcal{C}(z) dz,\end{aligned}$$

and (generally multi-valued) functions $M_\mathcal{C}, P_\mathcal{C} : \Theta \rightrightarrows \mathbb{R}$ as

$$M_\mathcal{C} := \hat{M}_\mathcal{C} \circ \pi^{-1}, \ P_\mathcal{C} := \hat{P}_\mathcal{C} \circ \pi^{-1}.$$

Let

$$L(s, \dot{s}) = \frac{1}{2} M_\mathcal{C}(s)\dot{s}^2 - P_\mathcal{C}(s). \tag{21}$$

Then the following statements are true.

(a) If $\Theta = \mathbb{R}$, then (19) is a Lagrangian system with Lagrangian L in (21).

(b) If $\Theta = \mathbb{S}^1$, then (19) is Lagrangian if, and only if, $\hat{M}_\mathcal{C}$ and $\hat{P}_\mathcal{C}$ are 2π-periodic, in which case $M_\mathcal{C}, P_\mathcal{C}$ in (21) are single-valued and smooth, and L in (21) is the Lagrangian function of (19).

As will become apparent in the development that follows, the results reviewed in this section are a special case of the general theory developed in this paper.

4. **Coordinate-free formulation of virtual holonomic constraints.** In this section we reformulate and generalize the theory of Section 3 in a coordinate-free context. We consider throughout a Lagrangian control system $(\mathcal{Q}, g, P, \mathsf{F})$ with equations of motion

$$\nabla_{\dot{q}}\dot{q} = -\operatorname{grad} P(q) + \sum_{i=1}^m (F^i)_q^\sharp \tau_i. \tag{22}$$

We assume that the one-forms $\mathsf{F} = \{F^1, \ldots, F^m\}$ are independent, and define the *acceleration distribution*

$$\mathcal{D}_A = \operatorname{span}\{(F^1)^\sharp, \ldots, (F^m)^\sharp\}. \tag{23}$$

We recall that base integral curves of (22) are projections onto $\mathcal{Q}$ of integral curves of the control affine system

$$\dot{X} = S(X) - \operatorname{vlft}(\operatorname{grad} P)(X) + \sum_{i=1}^{m} \tau_i \operatorname{vlft}\big((F^i)^\sharp\big)(X) \tag{24}$$

via the canonical projection map $\pi : T\mathcal{Q} \to \mathcal{Q}$. Note that

$$\operatorname{vlft}(\mathcal{D}_A) = \operatorname{span}\big\{\operatorname{vlft}((F^1)^\sharp), \ldots, \operatorname{vlft}((F^m)^\sharp)\big\}.$$

4.1. **VHC definitions and relationships.**

Definition 4.1 (Controlled invariant submanifold)**.** Let $\mathcal{C}$ be a closed embedded submanifold of $\mathcal{Q}$. The tangent bundle $T\mathcal{C}$ is *controlled invariant* for (22) if there exists a smooth feedback $\tau^\star = (\tau_1^\star, \ldots, \tau_m^\star) : T\mathcal{C} \to \mathbb{R}^m$ such that for each $q_0 \in \mathcal{C}$ and each $v_{q_0} \in T_{q_0}\mathcal{C}$, the maximal base integral curve $q : I \to \mathcal{Q}$ of (11) with feedback $\tau = \tau^\star(q, \dot{q})$ and initial condition (q_0, v_{q_0}) satisfies $q(t) \in \mathcal{C}$ for all $t \in I$. $\triangle$

In reference to the control-affine system (24), the definition above can be rephrased as the requirement that there exists a smooth feedback rendering $T\mathcal{C}$ an invariant set for the closed-loop system, which is the standard concept of controlled invariance of submanifolds used in control theory (see, e.g., [15]).

Definition 4.2 (Virtual holonomic constraint)**.** A *virtual holonomic constraint of order k* for the Lagrangian control system (11) is a closed embedded submanifold $\mathcal{C}$ of $\mathcal{Q}$ of codimension k such that $T\mathcal{C}$ is controlled invariant for (11). The set $T\mathcal{C}$ is called the *constraint manifold.* $\triangle$

Definition 4.3 (Regular VHC of order m)**.** A closed embedded submanifold $\mathcal{C}$ of $\mathcal{Q}$ is a *regular VHC of order m* for the Lagrangian control system (11) if $\mathcal{C}$ has codimension m and

$$(\forall q \in \mathcal{C})\ T_q\mathcal{C} \oplus \mathcal{D}_A(q) = T_q\mathcal{Q}, \tag{25}$$

where $\mathcal{D}_A$ is the acceleration distribution defined in (23). $\triangle$

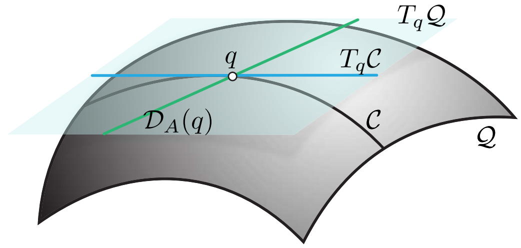

FIGURE 1. Transversality condition in the definition of regular VHC.

The transversality condition (25), illustrated in Figure 1, generalizes (18) in the case when the number of constraints, k, is equal to the number of controls, m.

A regular VHC is a VHC in the sense of Definition 4.2, as the next result shows.

Proposition 4.4. *If a closed embedded submanifold $\mathcal{C}$ of $\mathcal{Q}$ is a regular VHC of order m for system (22), then $\mathcal{C}$ is also a VHC in the sense of Definition 4.2, and the smooth feedback $\tau^\star : T\mathcal{C} \to \mathbb{R}^m$ rendering $T\mathcal{C}$ invariant is unique.*

We need the following lemma, whose proof is in the appendix.

Lemma 4.5. *Consider the Lagrangian control system (22) and its associated control-affine system (24). If $\mathcal{C}$ is a regular VHC, then for each $X_q \in T\mathcal{C}$,*

$$S(X_q) - \mathrm{vlft}(\mathrm{grad}\, P)(X_q) \in T_{X_q}T\mathcal{C} \oplus \mathrm{vlft}(\mathcal{D}_A)(X_q). \tag{26}$$

Proof of Proposition 4.4. Suppose $\mathcal{C}$ is a regular VHC for the Lagrangian control system (22). We need to show that the closed embedded submanifold $T\mathcal{C} \subset T\mathcal{Q}$ is controlled invariant for the control-affine system (24), and the smooth feedback rendering it invariant is unique. By Lemma 4.5, for each $X_q \in T\mathcal{C}$ we have that

$$S(X_q) - \mathrm{vlft}(\mathrm{grad}\, P)(X_q) \in T_{X_q}T\mathcal{C} \oplus \mathrm{span}\{\mathrm{vlft}\left((F^i)^\sharp\right), i \in \{1,\ldots,m\}\}, \tag{27}$$

from which it follows that there is a unique vector $\tau^\star(X_q) = (\tau_1^\star(X_q), \ldots, \tau_m^\star(X_q))$ such that

$$S(X_q) - \mathrm{vlft}(\mathrm{grad}\, P)(X_q) + \sum_{i=1}^m \tau_i^\star(X_q)\, \mathrm{vlft}\left((F^i)^\sharp\right) \in T_{X_q}T\mathcal{C}. \tag{28}$$

The map $T\mathcal{C} \to \mathbb{R}^m$, $X_q \to \tau^\star(X_q)$ is smooth because in any set of local coordinates $\hat{X}$ on $T\mathcal{C}$, the requirement (28) can be expressed as a matrix equation of the form $A(\hat{X})\tau = b(\hat{X})$, where, by (27), $\mathrm{rank}\, A = m$ and $b(\hat{X}) \in \mathrm{Im}\, A(\hat{X})$. The unique solution $\tau(\hat{X})$ of this equation is $\tau(\hat{X}) = \left(A(\hat{X})^\top A(\hat{X})\right)^{-1} A(\hat{X})^\top b(\hat{X})$, which is a smooth function. In conclusion, there exists a unique smooth feedback $\tau^\star : T\mathcal{C} \to \mathbb{R}^m$ such that the closed-loop vector field given by (13) with $\tau_i = \tau_i^\star(X)$ is tangent to $T\mathcal{C}$. By [9, Theorem 2.1], $T\mathcal{C}$ is an invariant set for the closed-loop vector field, and thus $\mathcal{C}$ is a VHC in the sense of Definition 4.2. ☐

Definition 4.3 of regular VHCs is a coordinate-free generalization of Definition 3.2 in the following sense. When $\mathcal{C}$ is globally described by the zero level set of a smooth submersion $h : \mathcal{Q} \to \mathbb{R}^m$, then $T\mathcal{C} = \{v_q \in T\mathcal{Q} : h(q) = 0,\ dh_q v_q = 0\}$, and system (22) with output function $e = h(q)$ has vector relative degree $\{2, \ldots, 2\}$. Indeed, differentiating each component of the output, $e_i = h_i(q)$, twice along the base integral curves of (22) one can show that

$$\ddot{e}_i = g(\nabla_{\dot{q}}\, \mathrm{grad}\, h_i, \dot{q}) - g(\mathrm{grad}\, h_i, \mathrm{grad}\, P) + \sum_{j=1}^m b_{ij}(q)\tau_j,\ \ i = 1, \ldots, k, \tag{29}$$

where $b_{ij} = g(\mathrm{grad}\, h_i, (F^j)^\sharp)$. The transversality condition (25) in Definition 4.3 implies that the $m \times m$ matrix with component b_{ij} is invertible on $\mathcal{C}$, and thus system (22) with output function $e = h(q)$ has vector relative degree $\{2, \ldots, 2\}$.

Just like in Section 3, the expression (29) suggests a way to asymptotically stabilize[3] the constraint manifold $T\mathcal{C}$ using an input-output linearizing feedback

$$\tau_i^\star = \sum_j b^{ij}(q) \left[-g(\nabla_{\dot{q}}\, \mathrm{grad}\, h_j, \dot{q}) + g(\mathrm{grad}\, h_j, \mathrm{grad}\, P) - K_{p,j} h_j - K_{d,j}(dh_j)_q \dot{q}\right],$$

where b^{ij} is the (i,j)-th element of the inverse of the matrix (b_{ij}), and $K_{p,j}$, $K_{d,j}$ are positive design parameters.

[3]Provided that certain technical assumptions hold for h and dh (see [21]).

The restriction of $\tau^\star$ above to $T\mathcal{C}$ is the unique feedback rendering $T\mathcal{C}$ invariant predicted by Proposition 4.4, and it is given by

$$\tau^\star|_{T\mathcal{C}} = \sum_j b^{ij}(q)\left[-g(\nabla_{\dot q}\operatorname{grad} h_j, \dot q) + g(\operatorname{grad} h_j, \operatorname{grad} P)\right].$$

For base integral curves $q(t)$ in $\mathcal{C}$ it holds that $\dot q \in T_q\mathcal{C}$. Moreover, since $\mathcal{C} = h^{-1}(0)$, $\operatorname{grad} h_i(q) \in T_q\mathcal{C}^\perp$. The Weingarten equation [19] then gives $g(\nabla_{\dot q}\operatorname{grad} h_j, \dot q) = -g(\operatorname{grad} h_j, \mathrm{I\!I}(\dot q, \dot q))$, where $\mathrm{I\!I}$ is the second fundamental form of $\mathcal{C}$. Thus the unique feedback rendering $T\mathcal{C}$ invariant is

$$\tau^\star|_{T\mathcal{C}} = \sum_j b^{ij}(q) g(\operatorname{grad} h_j, \mathrm{I\!I}(\dot q, \dot q) + \operatorname{grad} P).$$

4.2. **Constrained dynamics.** Our next objective is to characterize the constrained dynamics on $T\mathcal{C}$, by which we mean the closed-loop dynamics resulting from the application of the unique smooth feedback $\tau^\star : T\mathcal{C} \to \mathbb{R}^m$ rendering $T\mathcal{C}$ invariant. We would like a coordinate-free generalization of Proposition 3.4 valid for any m, not just $m = n-1$. As we now show, such dynamics are described by a special affine connection on $\mathcal{C}$ induced by the VHC. This so-called *induced connection* was originally developed in the context of affine differential geometry (see [26, Chapter 2]). We adopt it in the context of regular VHCs.

Before giving a formal definition of the induced connection, we present the basic idea behind it. If $\mathcal{C}$ is a regular VHC, the transversality condition (25) in Definition 4.3 states that, for each $q \in \mathcal{C}$, the tangent space $T_q\mathcal{Q}$ is the direct sum of $T_q\mathcal{C}$ and $\mathcal{D}_A(q)$. We may then define the projection $\sigma_q : T_q\mathcal{Q} \to T_q\mathcal{C}$ of the vector space $T_q\mathcal{Q}$ onto $T_q\mathcal{C}$ along the subspace $\mathcal{D}_A(q)$. The map σ_q is uniquely determined by the following properties:

(i) $\sigma_q^2 = \sigma_q$,
(ii) $\operatorname{Im}\sigma_q = T_q\mathcal{C}$,
(iii) $\operatorname{Ker}\sigma_q = \mathcal{D}_A(q)$.

Now consider the vector bundle map $\sigma : T\mathcal{Q}|_{\mathcal{C}} \to T\mathcal{C},\ w_q \mapsto \sigma_q(w_q)$, illustrated in Figure 2.

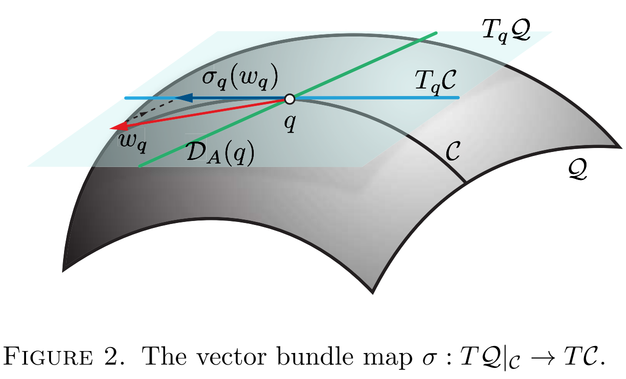

FIGURE 2. The vector bundle map $\sigma : T\mathcal{Q}|_{\mathcal{C}} \to T\mathcal{C}$.

Since the acceleration distribution $\mathcal{D}_A(q)$ is smooth, so is σ. Using σ, we define a new connection on $\mathcal{C}$ as follows. Given two vector fields $X, Y \in \mathfrak{X}(\mathcal{C})$, $\nabla_X Y$

is generally not a vector field on $\mathcal{C}$, but its projection $\sigma(\nabla_X Y)$ is, and the next theorem shows that this operation identifies an affine connection on $\mathcal{C}$.

Before presenting the theorem, we need to justify the notation $\nabla_X Y$ for vector fields $X, Y \in \mathfrak{X}(\mathcal{C})$, since the affine connection ∇ accepts vector fields on $\mathcal{Q}$. Consider arbitrary smooth extensions[4] $\tilde{X}, \tilde{Y}$ of X, Y on a neighbourhood of $\mathcal{C}$ in $\mathcal{Q}$ such that $\tilde{X}|_{\mathcal{C}} = X$ and $\tilde{Y}|_{\mathcal{C}} = Y$. Given any $p \in \mathcal{C}$, by [19, Exercise 4.7, p.58], the value of $\nabla_{\tilde{X}}\tilde{Y}(p)$ depends only on $\tilde{X}_p$ (and thus X_p) and the value of $\tilde{Y}$ along any smooth curve $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{Q}$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = X_p$. Since $X_p \in T_p\mathcal{C}$, we may pick a curve γ contained in $\mathcal{C}$, so that the value of $\tilde{Y}$ along γ coincides with that of Y. Therefore, on $\mathcal{C}$ the function $\nabla_{\tilde{X}}\tilde{Y}$ is uniquely determined by X, Y. These considerations justify the slight abuse of notation $\nabla_X Y$ for vector fields $X, Y \in \mathfrak{X}(\mathcal{C})$.

Theorem 4.6 ([26])**.** *Let $\mathcal{C}$ be a regular VHC of order m for the Lagrangian control system* (22)*, and define the map $\overset{\mathcal{C}}{\nabla} : \mathfrak{X}(\mathcal{C}) \times \mathfrak{X}(\mathcal{C}) \to \mathfrak{X}(\mathcal{C})$ as*

$$\overset{\mathcal{C}}{\nabla}_X Y := \sigma\big(\nabla_X Y\big), \tag{30}$$

where ∇ is the Riemannian connection of $(\mathcal{Q}, g)$. The map $\overset{\mathcal{C}}{\nabla}$ is a symmetric affine connection on $\mathcal{C}$.

The above result is mentioned in [26, Chapter 2, p. 28]. The straightforward proof is omitted.

We call $\overset{\mathcal{C}}{\nabla}$ the *induced connection*, or the connection induced by the regular VHC $\mathcal{C}$. While $\overset{\mathcal{C}}{\nabla}$ is symmetric, it is generally not a Riemannian connection with respect to the induced Riemannian metric on $\mathcal{C}$. This fact is discussed in the next section. Now the main result of this section. In what follows, let $\iota : \mathcal{C} \to \mathcal{Q}$ denote the inclusion map.

Theorem 4.7. *If $\mathcal{C}$ is a regular VHC of order m for the Lagrangian control system* (22)*, then the constrained dynamics on $T\mathcal{C}$ are described by the equation of motion*

$$\overset{\mathcal{C}}{\nabla}_{\dot{q}}\dot{q} = -\sigma_q(\operatorname{grad} P(q)), \tag{31}$$

in the following sense. If $q : I \to \mathcal{Q}$ is a maximal base integral curve of (22) *such that $q(I) \subset \mathcal{C}$, then $q : I \to \mathcal{C}$ is a maximal base integral curve of system* (31)*. Vice versa, if $q : I \to \mathcal{C}$ is a maximal base integral curve of* (31)*, then $\iota \circ q$ is a maximal base integral curve of* (22)*.*

Proof. Let $q : I \to \mathcal{Q}$ be a base integral curve of (22) such that $q(I) \subset \mathcal{C}$. Then

$$\nabla_{\dot{q}}\dot{q} + \operatorname{grad} P(q) - \sum_{i=1}^{m} (F^i)^\sharp \tau_i^\star(q) = 0,$$

where $\tau_i^\star(q)$ is the i-th component of the unique feedback $\tau^\star : T\mathcal{C} \to \mathbb{R}^m$ rendering $T\mathcal{C}$ invariant (see Proposition 4.4). Using σ_q to project both sides of the above identity onto $T_q\mathcal{C}$ and using the fact that $\sigma_q((F^i)^\sharp) = 0$, we get

$$\sigma_q(\nabla_{\dot{q}}\dot{q} + \operatorname{grad} P(q)) = 0.$$

By Theorem 4.6 we get

$$\overset{\mathcal{C}}{\nabla}_{\dot{q}}\dot{q} + \sigma_q(\operatorname{grad} P(q)) = 0,$$

[4] These exist by [20, Problem 8.15].

which proves that $q : I \to \mathcal{C}$ is a base integral curve of (31).

Now let $q : I \to \mathcal{C}$ be a base integral curve of (31). Then,

$$\overset{\mathcal{C}}{\nabla}_{\dot q}\dot q + \sigma_q(\operatorname{grad} P(q)) = 0,$$

or

$$\sigma_q(\nabla_{\dot q}\dot q + \operatorname{grad} P(q)) = 0.$$

Since $\operatorname{Ker}\sigma_q = \mathcal{D}_A(q)$, we have

$$\nabla_{\dot q}\dot q + \operatorname{grad} P(q) \in \mathcal{D}_A(q),$$

from which it follows that, for each $t \in I$, there exists $\bar\tau(q(t)) = (\bar\tau_1(q(t)), \ldots, \bar\tau_m(q(t))) \in \mathbb{R}^m$ such that

$$\nabla_{\dot q(t)}\dot q(t) + \operatorname{grad} P(q(t)) = \sum_{i=1}^{m} (F^i)^\sharp(q(t))\tau_i(q(t)).$$

By the uniqueness of the feedback $\tau^\star$ rendering $T\mathcal{C}$ invariant (see Proposition 4.4), it must hold that $\bar\tau = \tau^\star$, a smooth feedback. This proves that $\iota(q)$ is an integral curve of (22).

We now prove maximality. Suppose, by way of contradiction, that $q : I \to \mathcal{Q}$ a maximal base integral curve of (22) such that $q(I) \subset \mathcal{C}$ but the corresponding base integral curve $q : I \to \mathcal{C}$ of (31) is not maximal. Let $\tilde q : \tilde I \to \mathcal{C}$, $\tilde I \supset I$, be the unique maximal base integral curve of (31) such that $\tilde q|_I = q$. Then $\iota(\tilde q)$ is a base integral curve of (22) with a larger interval of existence than q, which contradicts the maximality of q. In an analogous way one shows that if $q : I \to \mathcal{C}$ is a maximal base integral curve of (31) then $\iota(q)$ is maximal for (22). □

4.3. **Constrained dynamics in coordinates.** We now characterize the constrained dynamics in coordinates. Pick a coordinate chart for $\mathcal{Q}$, (U, ϕ), with $\phi : U \to \hat U \subset \mathbb{R}^n$ and $\mathcal{C} \cap U \neq \emptyset$. Letting $x = \phi(q) = (x^1(q), \ldots, x^n(q))$, the equations of motion (22) in x coordinates read as (cf. (15)),

$$D(x)\ddot x + C(x, \dot x)\dot x + \nabla_x \hat P(x) = B(x)\tau.$$

Tangent space of $\mathcal{C}$. The chart domain U can always be chosen small enough that the local representation of the constraint manifold, $\hat{\mathcal{C}} = \phi(\mathcal{C} \cap U)$, is the image of a diffeomorphism $\varphi : W \subset \mathbb{R}^{n-m} \to \hat{\mathcal{C}}$, $s = (s^1, \ldots, s^{n-m}) \mapsto \varphi(s)$. Using this parametrization, we have $T_{\varphi(s)}\hat{\mathcal{C}} = \operatorname{Im}(d\varphi_s)$. Thus, letting

$$V^i(x) := d\varphi_{\varphi^{-1}(x)}(\partial/\partial s^i) = \partial_{s^i}\varphi(\varphi^{-1}(x)),\ i = 1, \ldots, n-m,$$

we have

$$T\hat{\mathcal{C}} = \operatorname{span}\{V^1, \ldots, V^{n-m}\}.$$

This construction is depicted in Figure 3.

Projection map. In coordinates, we have

$$d\phi_{\phi^{-1}(x)}\big(\mathcal{D}_A(\phi^{-1}(x))\big) = \operatorname{Im} D^{-1}(x)B(x).$$

Letting $B^\perp$ be a full-rank left-annihilator of B, the coordinate representation of the projection map σ is the map $\hat\sigma : T\hat U \to T\hat{\mathcal{C}}$ defined as

$$\hat\sigma_x(v_x) = d\varphi_s\big((B^\perp D d\varphi_s)^{-1}B^\perp D\big)\big|_{x=\varphi(s)}(v_x). \tag{32}$$

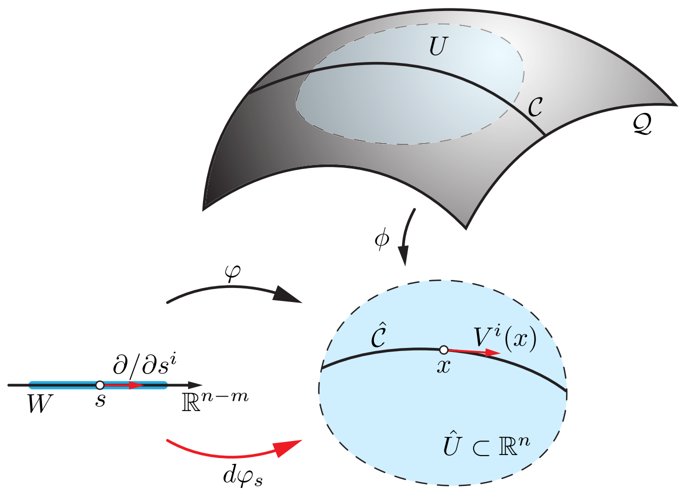

FIGURE 3. Coordinate systems used in Section 4.3.

Indeed, one can readily verify that $\hat{\sigma}_x^2 = \hat{\sigma}_x$, $\mathrm{Im}(\sigma_x) = T_x\hat{\mathcal{C}}$, and $\mathrm{Ker}\, \sigma_x = \mathrm{Im}(D^{-1}(x)$ $B(x))$. These properties imply that $\hat{\sigma}_x$ is the projection onto $T_x\hat{\mathcal{C}}$ along the subspace $\mathrm{Im}(D^{-1}(x)B(x))$, as required.

Induced connection $\overset{\mathcal{C}}{\nabla}$. The coordinate chart (U, ϕ) induces Christoffel symbols Γ_{ij}^k, $i, j, k \in \{1, \dots, n\}$, of the Riemannian connection ∇. We now derive the Christoffel symbols of the induced connection, defined through the identity

$$\hat{\sigma}(\nabla_{V^i} V^j) = \overset{\mathcal{C}}{\nabla}_{V^i} V^j = \sum_{k=1}^{n-m} \overset{\mathcal{C}}{\Gamma}{}_{ij}^k V^k,$$

where $\overset{\mathcal{C}}{\Gamma}{}_{ij}^k$ are the symbols we are looking for. Using the definition of V^i, V^j, identity (6), and the expression for $\hat{\sigma}$ in (32), one gets

$$\overset{\mathcal{C}}{\Gamma}{}_{ij}^k = \sum_{a=1}^{n} \partial^2_{s^i s^j} \varphi^a + [(B^\perp D d\varphi_s)^{-1} B^\perp D]_{ka} \sum_{b,c=1}^{n} \Gamma_{bc}^a (\partial_{s^i} \varphi^b)(\partial_{s^j} \varphi^c),$$

where φ^a denotes the a-th component of φ. Letting $\Gamma^a(x)$ be the matrix with components $(\Gamma^a)_{bc} = \Gamma_{bc}^a$, we may rewrite the Christoffel symbols of the induced connection in the more economical form

$$\overset{\mathcal{C}}{\Gamma}{}_{ij}^k = \sum_{a=1}^{n} [(B^\perp D d\varphi_s)^{-1} B^\perp D]_{ka} \left(\partial^2_{s^i s^j} \varphi^a + (\partial_{s^i} \varphi)^\top \Gamma^a (\partial_{s^j} \varphi) \right) \Big|_{x=\varphi(s)}, \tag{33}$$

$i, j, k \in \{1, \dots, n-m\}$.

Constrained dynamics. The coordinate representation of $\operatorname{grad} P(q)$ is

$$\left[D^{-1}(x) \nabla_x \hat{P}(x) \right]_{x=\varphi(s)}.$$

Using (32), the projection $\sigma_q(\operatorname{grad} P(q))$ in s-coordinates reads as

$$\left[(B^\perp D d\varphi_s)^{-1} B^\perp \nabla_x \hat{P} \right]_{x=\varphi(s)}.$$

Letting e_k denote the k-th natural basis vector of $\mathbb{R}^{n-m}$, the coordinate representation of the constrained dynamics (31) is

$$\ddot{s}^k = -\sum_{ij} \overset{\mathcal{C}}{\Gamma}{}^k_{ij}(s)\dot{s}^i\dot{s}^j - e_k^\top (B^\perp D d\varphi_s)^{-1} B^\perp \nabla_x \hat{P}\Big|_{x=\varphi(s)}, \; k=1,\dots,n-m. \tag{34}$$

In the special case when $\mathcal{C}$ is diffeomorphic to a generalized cylinder, one may pick φ to be a global diffeomorphism $(\mathbb{S}^1)^k \times (\mathbb{R})^{n-m-k} \to \mathcal{C}$, in which case the ODEs in (34) constitute a global representation of the constrained dynamics. In particular, for systems with degree of underactuation one, i.e., when $n-m=1$, (34) is always valid globally, and it reduces to

$$\begin{aligned}\ddot{s} &= -\overset{\mathcal{C}}{\Gamma}{}^1_{11}(s)\dot{t}h^2 - \sigma_s(\operatorname{grad} P(s))\\ &= -\frac{\sum_a (B^\perp D)_{1a}((\varphi^a)'' + \varphi'^\top \Gamma^a \varphi)}{B^\perp D\varphi'}\Big|_{x=\varphi(s)}\dot{s}^2 - \frac{B^\perp \nabla_x \hat{P}}{B^\perp D\varphi'}\Big|_{x=\varphi(s)},\end{aligned} \tag{35}$$

The above is precisely the form of the constrained dynamics in (19)-(20).

4.4. **Examples of computation of constrained dynamics.** We present two examples illustrating the formulas in Section 4.3.

Example 1. Consider the unit point-mass particle on the plane with inertial coordinates $q = [q_1 \ q_2]^\top \in \mathbb{R}^2$ depicted in Figure 4.

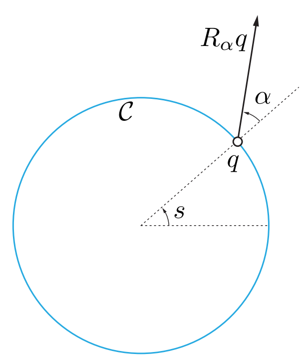

FIGURE 4. The set $\mathcal{C}$ in Example 1 and its parametrization.

The particle is actuated by a force $(R_\alpha\, q)\tau$, where $\tau \in \mathbb{R}$ is the control input, α is a fixed parameter, and $R_\alpha \in \mathsf{SO}(2)$ is the matrix operating a counterclockwise rotation by angle α. The equations of motion are

$$\ddot{q} = (R_\alpha\, q)\tau. \tag{36}$$

This is a Lagrangian system $(\mathcal{Q}, g, 0, F)$, with $\mathcal{Q} = \mathbb{R}^2$, g the Euclidean inner product, and $F(q) = (q_1 \cos\alpha - q_2 \sin\alpha)dq_1 + (q_1 \sin\alpha + q_2 \cos\alpha)dq_2$. Let $\mathcal{C}$ be the unit circle centred at the origin. If $\alpha \in (-\pi/2, \pi/2)$, then $\mathcal{C}$ is a regular VHC since, for all $q \in \mathcal{C}$, the vector $R_\alpha\, q$ is transversal to $\mathcal{C}$:

$$T_q\mathcal{C} + \operatorname{span} F^\sharp(q) = \operatorname{span}\begin{bmatrix} -q_2 \\ q_1 \end{bmatrix} + \operatorname{span}\{R_\alpha\, q\} = T_q\mathbb{R}^2.$$

The output function $e = q^\top q - 1$ yields vector relative degree $\{2,2\}$ everywhere on $\mathcal{C}$, and the feedback

$$\tau^\star(q,\dot q) = \frac{1}{2q^\top R_\alpha q}(-2\dot q^\top \dot q - K_p e - K_d \dot e),\ K_p, K_d > 0,$$

asymptotically stabilizes the constraint manifold $T\mathcal{C}$ and renders it invariant. The map $\varphi : \mathbb{S}^1 \to \mathbb{R}^2$, $\varphi(s) = [\cos(s)\ \sin(s)]^\top$, is a parametrization of $\mathcal{C}$. The Christoffel symbols Γ^k_{ij} of g are all zero. Letting $B^\perp(q) := q^\top R_{\alpha+\pi/2}$ and using (33), the Christoffel symbol of the induced connection is $\overset{\mathcal{C}}{\Gamma}{}^1_{11} = \tan\alpha$, so by (35) the constrained dynamics are given by

$$\ddot s = -(\tan\alpha)\dot s^2.$$

One can also derive the constrained dynamics by multiplying both sides of (36) on the left by $B^\perp$, and substituting $q = \varphi(s)$, $\ddot q = \varphi'(s)\ddot s + \varphi''(s)\dot s^2$ in the resulting expression. The ODE one gets this way is the same as above. △

Example 2. Consider now a unit point-mass in $\mathbb{R}^3$ with inertial coordinates $q = [q_1\ q_2\ q_3]^\top \in \mathbb{R}^3$, actuated by a control force $(\mathrm{diag}(1,1,2)q)\tau$, where $\tau \in \mathbb{R}$ is the control input:

$$\ddot q = B(q)\tau,$$

where $B(q) = \mathrm{diag}(1,1,2)q$. This is a Lagrangian system $(\mathcal{Q}, g, 0, F)$, where $\mathcal{Q} = \mathbb{R}^3$, g is the Euclidean inner product, and $F(q) = q_1 dq_1 + q_2 dq_2 + 2q_3 dq_3$. In this example, $\mathcal{D}_A(q) = \mathrm{span}\{F^\sharp(q)\} = \mathrm{Im}\, B(q)$. Let $\mathcal{C}$ be the unit sphere centred at the origin, $\mathcal{C} = \{q \in \mathbb{R}^3 : q^\top q = 1\}$. The set $\mathcal{C}$ is illustrated in Figure 5. For each $q \in \mathcal{C}$, $T_q\mathcal{C}$

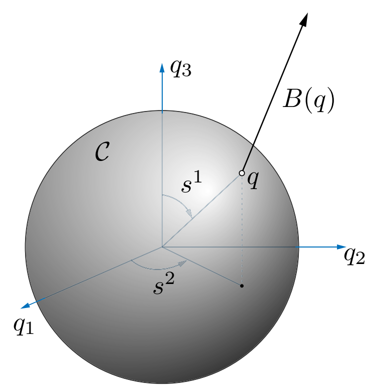

FIGURE 5. The VHC $\mathcal{C}$ in Example 2 and its parametrization.

is the orthogonal complement of $\mathrm{span}\{q\}$. Since $g(q, F^\sharp(q)) = q^\top \mathrm{diag}(1,1,2)q > 0$, the control force is transversal everywhere to the sphere, and therefore, for any $q \in \mathcal{C}$,

$$T_q\mathcal{C} \oplus \mathcal{D}_A(q) = T_q\mathbb{R}^3.$$

Thus $\mathcal{C}$ is a regular VHC. For a parametrization of $\mathcal{C}$, we use spherical coordinates:

$$\varphi(s^1, s^2) = \begin{bmatrix} \sin(s^1)\cos(s^2) \\ \sin(s^1)\sin(s^2) \\ \cos(s^1) \end{bmatrix}. \tag{37}$$

Letting $W = (0,\pi) \times (-\pi,\pi)$ and $\hat{\mathcal{C}} = \mathbb{S}^2/\{N,P\}$, where N and P are the north and south poles of $\mathcal{C}$, the map $\varphi : W \to \hat{C}$ is a diffeomorphism. To compute the Christoffel symbols of the induced connection on $\mathcal{C}$, we define a left-annihilator of $B(q) = \operatorname{diag}(1,1,2)q$:

$$B^{\perp}(q) = \operatorname{Im}\begin{bmatrix} -q_2 & q_1 & 0 \\ -q_1 q_3 & -q_2 q_3 & (q_1^2+q_2^2)/2 \end{bmatrix}.$$

For all $q \in \hat{C}$, $\operatorname{rank} B^{\perp}(q) = 2$ and $B^{\perp}B = 0$, as required. Using $B^{\perp}$ above, φ in (37), $D = I_3$, and $\Gamma_{ij}^k = 0$, we get $\overset{\mathcal{C}}{\Gamma}{}_{ij}^k$ from (33) as

$$\begin{aligned}
&\overset{\mathcal{C}}{\Gamma}{}_{11}^1 = \frac{-\sin(2s^1)}{2(\cos^2(s^1)+1)}, && \overset{\mathcal{C}}{\Gamma}{}_{22}^1 = \frac{-\sin(2s^1)}{\cos^2(s^1)+1}, && \overset{\mathcal{C}}{\Gamma}{}_{12}^1 = \overset{\mathcal{C}}{\Gamma}{}_{21}^1 = 0,\\
&\overset{\mathcal{C}}{\Gamma}{}_{11}^2 = 0, && \overset{\mathcal{C}}{\Gamma}{}_{22}^2 = 0, && \overset{\mathcal{C}}{\Gamma}{}_{12}^2 = \overset{\mathcal{C}}{\Gamma}{}_{21}^2 = \cot(s^1).
\end{aligned}$$

Therefore, the coordinate representation of the constrained dynamics on $T\mathcal{C}$ is given by the ODEs

$$\begin{aligned}
\ddot{s}^1 &= \frac{\sin(2s^1)}{2(\cos^2(s^1)+1)}(\dot{s}^1)^2 + \frac{\sin(2s^1)}{\cos^2(s^1)+1}(\dot{s}^2)^2 \\
\ddot{s}^2 &= -2\cot(s^1)\dot{s}^1\dot{s}^2.
\end{aligned} \tag{38}$$

In Section 7 we will investigate the Lagrangian structure of (38). △

5. **Existence of a Lagrangian structure for the constrained dynamics.**

In this section we investigate this question: given the Lagrangian control system (22) and a regular VHC $\mathcal{C}$ of order m, determine whether there exists a Riemannian metric $g_{\mathcal{C}}$ on $\mathcal{C}$ and a smooth potential function $P_{\mathcal{C}} : \mathcal{C} \to \mathbb{R}$ such that the constrained dynamics (31) are generated by the Lagrangian structure $(\mathcal{C}, g_{\mathcal{C}}, P_{\mathcal{C}})$. If this is the case, we say that *the constrained dynamics are Lagrangian.* The solution in the special case $m = n-1$ was reviewed in Theorem 3.5. Here we investigate the problem from a more general perspective.

5.1. **A general result.**

Theorem 5.1. *If $\mathcal{C}$ is a regular VHC of order m for the Lagrangian control system (22), then the constrained dynamics (31) are Lagrangian if and only if the following two conditions hold:*

(i) The induced connection $\overset{\mathcal{C}}{\nabla}$ is metrizable, i.e., there exists a Riemannian metric $g_{\mathcal{C}}$ on $\mathcal{C}$ such that $\overset{\mathcal{C}}{\nabla}$ is the Riemannian connection associated with $g_{\mathcal{C}}$.

(ii) There exists a smooth function $P_{\mathcal{C}} : \mathcal{C} \to \mathbb{R}$ such that

$$\sigma(\operatorname{grad} P) = \operatorname{grad}_{\mathcal{C}} P_{\mathcal{C}},$$

where $\operatorname{grad}_C P_{\mathcal{C}} \in \mathfrak{X}(\mathcal{C})$ is the gradient vector field of $P_{\mathcal{C}}$ induced by the metric $g_{\mathcal{C}}$, i.e., defined by the identity $dP_{\mathcal{C}}(v_q) = g_{\mathcal{C}}(\operatorname{grad}_{\mathcal{C}} P_{\mathcal{C}}, v_q)$ for all $v_q \in T\mathcal{C}$.

Moreover, if (i) and (ii) hold, the Lagrangian structure of the constrained dynamics is $(\mathcal{C}, g_{\mathcal{C}}, P_{\mathcal{C}})$.

Proof. ($\Longleftarrow$) If conditions (i) and (ii) hold, then it follows directly from the definition that the constrained dynamics (31) are generated by the Lagrangian system $(\mathcal{C}, g_{\mathcal{C}}, P_{\mathcal{C}})$.

($\Longleftarrow$) Suppose the constrained dynamics (31) are generated by a Lagrangian system $(\mathcal{C}, g_{\mathcal{C}}, P_{\mathcal{C}})$. Let $\bar{\nabla}$ be the Riemannian connection associated with $g_{\mathcal{C}}$. By Theorem 4.7, a curve $q : I \to \mathcal{C}$ satisfies

$$\overset{\mathcal{C}}{\nabla}_{\dot{q}}\dot{q} + \sigma_q(\operatorname{grad} P(q)) = 0 \tag{39}$$

if and only if it satisfies

$$\bar{\nabla}_{\dot{q}}\dot{q} + \operatorname{grad}_{\mathcal{C}} P_{\mathcal{C}}(q) = 0. \tag{40}$$

For any $q_0 \in \mathcal{C}$, let $q : I \to \mathcal{C}$ be the maximal integral curve of the constrained dynamics (31) with initial condition $(q_0, 0)$. If $\{X_1, \ldots, X_{n-m}\}$ is any local frame for $T\mathcal{C}$ defined in a neighbourhood of q_0, identity (7) and the fact that $\dot{q}|_{t=0} = 0$ imply that

$$\overset{\mathcal{C}}{\nabla}_{\dot{q}}\dot{q}\Big|_{t=0} = \bar{\nabla}_{\dot{q}}\dot{q}\Big|_{t=0}.$$

Since (39) and (40) hold, we deduce that

$$\sigma_{q_0}(\operatorname{grad} P(q_0)) = \operatorname{grad}_{\mathcal{C}}(P(q_0)),$$

proving that $P_{\mathcal{C}}$ satisfies condition (ii).

Next, subtracting (39) from (40) we get

$$\overset{\mathcal{C}}{\nabla}_{\dot{q}}\dot{q} = \bar{\nabla}_{\dot{q}}\dot{q}.$$

Since symmetric connections having the same geodesics are equal (see, instance, [28, Theorem 2.101]), the above implies that $\overset{\mathcal{C}}{\nabla} = \bar{\nabla}$. Hence, $\overset{\mathcal{C}}{\nabla}$ is metrizable, which proves condition (i). □

5.2. **Case of orthogonal control accelerations.** Referring to the regularity condition (25), when the acceleration distribution $\mathcal{D}_A$ is fibrewise orthogonal to $T\mathcal{C}$ (see Figure 6), the feedback rendering $T\mathcal{C}$ invariant produces a control force that does no work on base integral curves contained in $\mathcal{C}$. In this setting, the control force

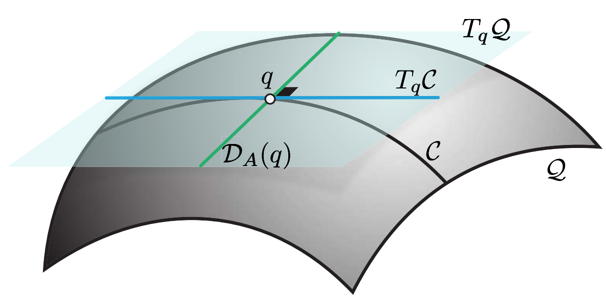

FIGURE 6. Illustration of the case when the control accelerations are orthogonal to $\mathcal{C}$.

is identical to the constraint force that would arise if $\mathcal{C}$ were a holonomic constraint[5]. Just like in classical mechanics, one expects the constrained dynamics to

[5]In classical mechanics, holonomic constraints such that the constraint force does no work along constrained solutions are called *ideal*.

be Lagrangian, with Lagrangian structure given by the restriction of the original Lagrangian structure to $\mathcal{C}$. The next proposition makes this intuition precise. Recall the inclusion map $\iota : \mathcal{C} \to \mathcal{Q}$. The metric $g : T\mathcal{Q} \times T\mathcal{Q} \to \mathbb{R}$ on $\mathcal{Q}$ gives rise to a metric on $\mathcal{C}$ via the pullback

$$\iota^* g(v_q, w_q) = g(d\iota_q(v_q), d\iota_q(w_q)) \text{ for all } v_q, w_q \in T_q\mathcal{C}.$$

The metric $\iota^* g$ is called the *induced metric on* $\mathcal{C}$.

Proposition 5.2. *If $\mathcal{C}$ is a regular VHC of order m for the Lagrangian control system (22) such that*

$$(\forall q \in \mathcal{C}) \; T_q\mathcal{C} \overset{\perp}{\oplus} \mathcal{D}_A(q) = T_q\mathcal{Q}, \tag{41}$$

with orthogonality holding with respect to the metric g, then the constrained dynamics (31) are Lagrangian with Lagrangian structure $(\mathcal{C}, g_\mathcal{C}, P_\mathcal{C})$, where $g_\mathcal{C} = \iota^* g$ and $P_\mathcal{C} = P|_\mathcal{C} = P \circ \iota$.

Proof. Since ∇ is a Riemannian connection, it satisfies

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \tag{42}$$

for all $X, Y, Z \in \mathfrak{X}(\mathcal{Q})$, and therefore also for all $X, Y, Z \in \mathfrak{X}(\mathcal{C})$. Let $X, Y, Z \in \mathfrak{X}(\mathcal{C})$ be arbitrary. In light of the regularity condition (18), we have

$$\nabla_X Y = \sigma(\nabla_X Y) + N_X Y = \overset{\mathcal{C}}{\nabla}_X Y + N_X Y,$$

where $N_X Y$ is a vector field in the control distribution $\operatorname{span}\{(F^i)^\sharp, i = 1, \ldots, m\}$. By the orthogonality hypothesis, we have $g(N_X Y, Z) = 0$ for all $Z \in \mathfrak{X}(\mathcal{C})$, implying that

$$g(\nabla_X Y, Z) = g(\overset{\mathcal{C}}{\nabla}_X Y, Z) = g_\mathcal{C}(\overset{\mathcal{C}}{\nabla}_X Y, Z). \tag{43}$$

The second identity in (43) is due to the fact that $\overset{\mathcal{C}}{\nabla}_X Y$ and Z are vector fields on $\mathcal{C}$. Analogously to (43), we have

$$g(Y, \nabla_X Z) = g_\mathcal{C}(Y, \overset{\mathcal{C}}{\nabla}_X Z). \tag{44}$$

Substituting (43) and (44) into (42) and using the fact that $g(Y, Z) = g_\mathcal{C}(Y, Z)$, we get

$$X(g_\mathcal{C}(Y, Z)) = g_\mathcal{C}(\overset{\mathcal{C}}{\nabla}_X Y, Z) + g_\mathcal{C}(Y, \overset{\mathcal{C}}{\nabla}_X Z),$$

implying that $\overset{\mathcal{C}}{\nabla}$ is compatible with $g_\mathcal{C}$. Since, by Theorem 4.6, $\overset{\mathcal{C}}{\nabla}$ is symmetric, $\overset{\mathcal{C}}{\nabla}$ is Riemannian with respect $g_\mathcal{C}$, proving that condition (i) of Theorem 5.1 holds.

Next, we need to show that, for each $q \in \mathcal{C}$, $\sigma_q(\operatorname{grad} P(q)) = \operatorname{grad}_\mathcal{C} P|_\mathcal{C}$, where $\operatorname{grad}_\mathcal{C} P|_\mathcal{C}$ is the gradient vector field of $P|_\mathcal{C}$ induced by the metric $g_\mathcal{C}$. Since, by assumption, the subspace $\mathcal{D}_A(q)$ is orthogonal to $T_q\mathcal{C}$, the projection $\sigma_q : T_q\mathcal{Q} \to T_q\mathcal{C}$ along $\mathcal{D}_A(q)$ is a map whose kernel, $\mathcal{D}_A(q)$, is orthogonal to its image, $T_q\mathcal{C}$. This fact implies that σ_q is a self-adjoint map. Thus, for all $v_q \in T_q\mathcal{C}$,

$$g(\sigma_q(\operatorname{grad} P(q)), v_q) = g(\operatorname{grad} P(q), \sigma_q(v_q)) = g(\operatorname{grad} P(q), v_q). \tag{45}$$

Using the definition of grad, we have

$$g(\operatorname{grad} P(q), v_q) = dP_q(v_q). \tag{46}$$

Since $P|_\mathcal{C} = P \circ \iota$ and, for all $v_q \in T_q\mathcal{C}$, $d\iota_q(v_q) = v_q$, we may write

$$dP_q(v_q) = dP_q \circ d\iota_q(v_q) = d(P \circ \iota)_q(v_q) = (dP|_\mathcal{C})_q(v_q). \tag{47}$$

Substituting (46) and (47) into (45) and using the definition of $\operatorname{grad}_{\mathcal{C}}$, we get

$$g(\sigma_q(\operatorname{grad} P(q)), v_q) = (dP|_{\mathcal{C}})_q(v_q) = g_{\mathcal{C}}(\operatorname{grad}_{\mathcal{C}} P|_{\mathcal{C}}(q), v_q),$$

for all $v_q \in T_q\mathcal{C}$. Since $\sigma_q(\operatorname{grad} P(q))$ and v_q lie in $T_q\mathcal{C}$, $g(\sigma_q(\operatorname{grad} P(q)), v_q) = g_{\mathcal{C}}(\sigma_q(\operatorname{grad} P(q)), v_q)$, and so

$$g_{\mathcal{C}}(\sigma_q(\operatorname{grad} P(q)), v_q) = g_{\mathcal{C}}(\operatorname{grad}_{\mathcal{C}} P|_{\mathcal{C}}(q), v_q),$$

for all $v_q \in T_q\mathcal{C}$. In conclusion, $\sigma_q(\operatorname{grad} P(q)) = \operatorname{grad}_{\mathcal{C}} P|_{\mathcal{C}}(q)$, proving that $P|_{\mathcal{C}}$ satisfies condition (ii) of Theorem 5.1. $\square$

Remark 5.3. As mentioned earlier, the foregoing proposition states that the constrained dynamics associated with a VHC satisfying condition (41) coincide with the constrained dynamics one would have if the Lagrangian system $(\mathcal{Q}, g, P)$ (without control) were subjected to an ideal holonomic constraint. The result in Proposition 5.2 is not new once placed in the context of geometric mechanics, and we do not claim it to be original. For instance, [14, Theorem 2.7] states an analogous result. The value of Proposition 5.2 lies in the that it connects the concept of virtual holonomic constraint in control theory with the concept of ideal holonomic constraint in mechanics in the special case when the control accelerations are fibrewise orthogonal to $T\mathcal{C}$. To gain further understanding of the relationship between Proposition 5.2 and established concepts in geometric mechanics, it is worth comparing it with Proposition 4.97 in [7]. In [7], a holonomic constraint $\mathcal{C}$ is a maximal integral manifold of a distribution $\mathcal{D}$ representing a linear velocity constraint. This distribution is used to define a *constrained connection* $\overset{\mathcal{D}}{\nabla}$ on $\mathcal{Q}$. Proposition 4.97 in [7] states that the restriction of $\overset{\mathcal{D}}{\nabla}$ to $\mathfrak{X}(\mathcal{C}) \times \mathfrak{X}(\mathcal{C})$ is the Riemannian connection of $\iota^* g$, and Proposition 4.85 in [7] implies that $\overset{\mathcal{D}}{\nabla}_X Y$ coincides with our $\overset{\mathcal{C}}{\nabla}_X Y$ for all $X, Y \in \mathfrak{X}(\mathcal{C})$. Taken together, these results recover the proof of the first part of Proposition 5.2, illustrating the strong analogy between VHCs satisfying condition (41) and holonomic constraints in [7]. As a caveat, we remark that, unlike the framework in [7], we do not require an integrable distribution to define the induced connection $\overset{\mathcal{C}}{\nabla}$, for $\overset{\mathcal{C}}{\nabla}$ is only required to be defined on $\mathfrak{X}(\mathcal{C}) \times \mathfrak{X}(\mathcal{C})$.

6. **Conditions for metrizability of affine connections.** Theorem 5.1 establishes that in the absence of a potential function, assessing whether or not the constrained dynamics induced by a regular VHC are Lagrangian amounts to assessing the metrizability of the induced connection. In this section we review the main results on metrizability of affine connections, presenting concrete results for the cases $\dim \mathcal{C} = 1$ (already covered in Theorem 3.5) and $\dim \mathcal{C} = 2$. To keep the notation simple, throughout the section we will consider a symmetric affine connection $\nabla : \mathfrak{X}(\mathcal{C}) \times \mathfrak{X}(\mathcal{C}) \to \mathfrak{X}(\mathcal{C})$ with the understanding that all result will apply to the induced connection $\overset{\mathcal{C}}{\nabla}$. We also assume throughout that the submanifold $\mathcal{C}$ is connected.

6.1. **The holonomy group of an affine connection.** A vector field $X(t)$ along a smooth curve γ on $\mathcal{C}$ is said to be *parallel* if its covariant derivative along γ vanishes, i.e., $D_t X \equiv 0$. Given a point $q \in \mathcal{C}$ and a tangent vector $v_q \in T_q\mathcal{C}$, the equation $D_t X = 0$ with initial conditions $X(0) = v_q$ uniquely determines a parallel vector field $X(t)$ along γ (see [19, Chapter II, Proposition 3.3]). This vector field is called the *parallel translation of* v_q *along* γ. In local coordinates, the equation $D_t X = 0$ is the linear time-varying ODE

$$\dot{X}^k = -\sum_{i,j} \dot{\gamma}^i(t)\Gamma^k_{ij}(\gamma(t))X^j, \ k = 1, \dots, n-m, \tag{48}$$

where Γ^k_{ij} are the Christoffel symbols of ∇ and $(X^1, \dots, X^{n-m})$ is the coordinate representation of X.

In what follows, if $q, p \in \mathcal{C}$, a piecewise smooth curve in $\mathcal{C}$ starting at q and ending at p will be denoted γ^p_q. More precisely, $\gamma^p_q : [0, T] \to \mathcal{C}$ will be a piecewise smooth map such that $\gamma^p_q(0) = q$ and $\gamma^p_q(T) = p$. On the other hand, a *loop at* q, i.e., a piecewise smooth closed curve through q will be denoted by γ_q. The set of all loops at q will be denoted by L_q.

For a curve γ^p_q, the *parallel transport map along* γ^p_q, denoted $\mathsf{P}_{\gamma^p_q} : T_q\mathcal{C} \to T_p\mathcal{C}$, is defined as $\mathsf{P}_{\gamma^p_q}(v_q) := X(T)$, where X is the parallel translation of v_q along γ^p_q. For $\gamma_q \in L_q$, P_{γ_q} maps $T_q\mathcal{C}$ onto itself. The parallel transport map associated with a loop is illustrated in Figure 7.

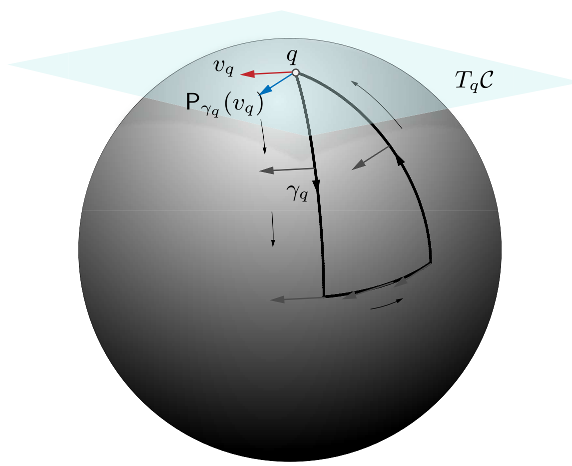

FIGURE 7. The parallel transport map at the north pole of the unit sphere in $\mathbb{R}^3$, with Riemannian connection induced by the Euclidean metric in $\mathbb{R}^3$. The loop γ_q is a triangle on the sphere.

Denoting by $(\gamma^p_q)^{-1}$ the curve obtained by reversing the orientation of γ^p_q, and by $\gamma^p_q \cdot \gamma^r_p$ the concatenation of γ^p_q with γ^r_p, we have the following result.

Proposition 6.1 ([17], Chapter II, Proposition 3.3)**.** *For each* $q \in \mathcal{C}$ *and any piecewise smooth curves* γ^p_q, γ^r_p, *the parallel transport map* $\mathsf{P}_{\gamma^p_q} : T_q\mathcal{C} \to T_p\mathcal{C}$ *is an isomorphism enjoying the following properties:*

(i) *If* $\gamma_q \in L_q$ *is the constant loop* $\gamma_q(t) \equiv q$, *then* P_{γ_q} *is the identity map on* $T_q\mathcal{C}$.
(ii) $\mathsf{P}_{(\gamma^p_q)^{-1}} = (\mathsf{P}_{\gamma^p_q})^{-1}$.
(iii) $\mathsf{P}_{\gamma^p_q \cdot \gamma^r_p} = \mathsf{P}_{\gamma^p_q} \circ \mathsf{P}_{\gamma^r_p}$.

In particular, the set of all isomorphisms $\{\mathsf{P}_{\gamma_q} : \gamma_q \in L_q\}$ *forms a group under composition.*

The group $\mathrm{Hol}_q(\nabla) := \{\mathsf{P}_{\gamma_q} : \gamma_q \in L_q\}$ of all parallel transport maps along loops at q is called the *holonomy group of* ∇ *with reference point* q, while the subgroup $\mathrm{Hol}^0_q(\nabla) := \{\mathsf{P}_{\gamma_q} : \gamma_q \in L_q \text{ is contractible to } q\}$ is the *restricted holonomy group.* A remarkable property of the holonomy groups is that they possess a Lie group structure.

Theorem 6.2 ([17], Chapter II, Theorem 4.2)**.** *Let $\mathcal{C}$ be a connected manifold, and let $q \in \mathcal{C}$. Then the following are true:*

(i) $\mathrm{Hol}^0_q(\nabla)$ *is a connected Lie subgroup of* $\mathrm{GL}(n)$*.*
(ii) $\mathrm{Hol}_q(\nabla)$ *is a Lie subgroup of* $\mathrm{GL}(n)$ *whose identity component is* $\mathrm{Hol}^0_q(\nabla)$*.*

6.2. **Schmidt's metrizability theorem.** The significance of the holonomy group at it pertains to the metrizability of ∇ rests upon the following consideration. If ∇ is Riemannian with respect to a metric g, then it is a basic fact of Riemannian geometry that for any two vector fields V, W that are parallel along a curve γ, $g(V, W)$ is constant along γ. In particular, for any $\gamma_q \in \mathrm{Hol}_q(\nabla)$, $g_q(\mathsf{P}_{\gamma_q}(v_q), \mathsf{P}_{\gamma_q}(w_q)) = g_q(v_q, w_q)$. Also, by definition, g_q is a positive definite, symmetric bilinear form on $T_q\mathcal{C}$. In conclusion, a necessary condition for ∇ to be metrizable is that there exists a symmetric, positive definite bilinear form $T_q\mathcal{C} \times T_q\mathcal{C} \to \mathbb{R}$ that is invariant under the holonomy group $\mathrm{Hol}_q(\nabla)$. This condition is also sufficient.

Theorem 6.3 ([30])**.** *Let ∇ be a symmetric affine connection on a connected manifold $\mathcal{C}$ and let $q \in \mathcal{C}$ be arbitrary. Then ∇ is metrizable if and only there exists a symmetric positive definite bilinear form $g_q : T_q\mathcal{C} \times T_q\mathcal{C} \to \mathbb{R}$ that is invariant under $\mathrm{Hol}_q(\nabla)$, i.e., for all $\gamma_q \in \mathrm{Hol}_q(\nabla)$ and all $v_q, w_q \in T_q\mathcal{C}$,*

$$g_q(\mathsf{P}_{\gamma_q}(v_q), \mathsf{P}_{\gamma_q}(w_q)) = g_q(v_q, w_q). \tag{49}$$

Schmidt's theorem only requires one to determine whether or not a bilinear form on $T_q\mathcal{C} \times T_q\mathcal{C}$ exists which is invariant under $\mathrm{Hol}_q(\gamma_q)$. It then guarantees that the form in question can be extended to a Riemannian metric defined on the whole of $T\mathcal{C} \times T\mathcal{C}$. We have already outlined the necessity part of the proof. The idea of the sufficiency proof is to extend the bilinear form g_q to the entire tangent bundle $T\mathcal{C}$ by parallel translation along curves connecting q to arbitrary points in $\mathcal{C}$. Specifically, for arbitrary $p \in \mathcal{C}$, pick an arbitrary piecewise smooth γ^p_q connecting p and q, and define

$$g_p(v_p, w_p) := g_q\big(\mathsf{P}_{\gamma^p_q}(v_p), \mathsf{P}_{\gamma^p_q}(w_p)\big). \tag{50}$$

The invariance of g_q under $\mathrm{Hol}_q(\nabla)$ guarantees that g_p is path-independent, giving rise to a Riemannian metric on $\mathcal{C}$. One can easily show that the extension so obtained is a Riemannian metric associated with ∇.

The result in Theorem 6.3 is of difficult application because the group $\mathrm{Hol}_q(\nabla)$ is generally hard to find. The Ambrose-Singer theorem [1] characterizes $\mathrm{Hol}^0_q(\nabla)$ in terms of the curvature form of the connection, but it requires the knowledge of the so-called holonomy bundle, an object which is not readily available. In special cases, however, the computations are more manageable, as we discuss next.

6.3. **Flat connections.** The *curvature endomorphism induced by an affine connection* ∇ is the map $\mathfrak{X}(\mathcal{C}) \times \mathfrak{X}(\mathcal{C}) \times \mathfrak{X}(\mathcal{C}) \to \mathfrak{X}(\mathcal{C})$ defined as

$$R(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

If ∇ is a flat connection, i.e., the curvature R induced by ∇ is zero, then the Ambrose-Singer theorem implies that $\mathrm{Hol}_q^0(\nabla)$ is trivial, and by [4, Theorem 2] there exists a surjective homomorphism $\pi_1(\mathcal{C}, q) \to \mathrm{Hol}_q(\nabla)$, where $\pi_1(\mathcal{C}, q)$ is the first homotopy group of $\mathcal{C}$ with reference point q. The homomorphism in question sends an equivalence class of loops $[\gamma_q] \in \pi_1(\mathcal{C}, q)$ to a parallel transport map $\mathsf{P}_{\gamma_q} \in \mathrm{Hol}_q(\nabla)$. In this case, to apply Theorem 6.3 it suffices to compute the transport maps associated with the generators of $\pi_1(\mathcal{C}, q)$, as stated next.

Proposition 6.4. *Let ∇ be a symmetric affine connection on a connected manifold $\mathcal{C}$, and suppose that ∇ is flat. Let $q \in \mathcal{C}$ be arbitrary, and let S_q be a set of generators of $\pi_1(\mathcal{C}, q)$. Then ∇ is metrizable if and only if there exists a symmetric positive definite bilinear form $g_q : T_q\mathcal{C} \times T_q\mathcal{C} \to \mathbb{R}$ such that, for each equivalence class $E_q \in S_q$, there exist a piecewise smooth curve $\gamma_q \in E_q$ for which*

$$g_q(\mathsf{P}_{\gamma_q}(v_q), \mathsf{P}_{\gamma_q}(w_q)) = g_q(v_q, w_q),$$

for any $v_q, w_q \in T_q\mathcal{C}$.

6.4. **Simply connected manifolds.** When $\mathcal{C}$ is simply connected, $\mathrm{Hol}_q(\nabla) = \mathrm{Hol}_q^0(\nabla)$ because, by definition, all loops at q in $\mathcal{C}$ are contractible to q. By Theorem 6.2, $\mathrm{Hol}_q^0(\nabla)$ is a connected Lie group, implying that it is entirely characterized by its Lie algebra, the so-called *holonomy algebra of* ∇. We will denote by $\mathfrak{h}$ the holonomy algebra. For simply connected manifolds, one may express the invariance condition in (49) in infinitesimal form, giving rise to a Lie algebraic metrizability criterion.

Lemma 6.5 ([35])**.** *Let ∇ be a symmetric affine connection on a simply connected manifold $\mathcal{C}$ and let $q \in \mathcal{C}$ be arbitrary. A symmetric positive definite bilinear form $g_q : T_q\mathcal{C} \times T_q\mathcal{C} \to \mathbb{R}$ is invariant under* $\mathrm{Hol}_q(\nabla)$ *if and only if for all $A \in \mathfrak{h}$ and all $v_q, w_q \in T_q\mathcal{C}$,*

$$g_q(Av_q, w_q) + g_q(v_q, Aw_q) = 0. \tag{51}$$

Remark 6.6. If ∇ is real analytic[6], the holonomy algebra $\mathfrak{h}$ is entirely characterized by the curvature and its covariant derivatives (see [17, Chapter II, Proposition 10.4 and Theorem 10.8]). Therefore, $\mathfrak{h}$ can be computed in local coordinates. Then, a consequence of Lemma 6.5 is that if $\mathcal{C}$ is simply connected and ∇ is a real analytic affine connection on $\mathcal{C}$, ∇ is metrizable if and only if it is locally metrizable (i.e., metrizable in local coordinates).

Exploiting Lemma 6.5 and the de Rham decomposition of a Riemannian manifold, Kovalski in [18] gave an effective decision algorithm for metrizability of real analytic affine connections on simply connected manifolds. We will not review the algorithm here, but we refer the reader to the review in [35].

6.5. **One-dimensional manifolds.** If $\mathcal{C}$ is one-dimensional, then it is diffeomorphic to either $\mathbb{R}$ or $\mathbb{S}^1$. This situation occurs in Lagrangian control systems with degree of underactuation one, when a regular VHC of codimension one is enforced. In Theorem 3.5 of Section 3, we reviewed necessary and sufficient conditions for the constrained dynamics to be Lagrangian. Now we show that the conditions of Theorem 3.5 have an elegant interpretation in the context of induced connections. We will recover Theorem 3.5 as a corollary of Theorem 5.1 and Proposition 6.4.

[6] And in other cases when ∇ is smooth, see [17].

We begin with the observation that any affine connection on a one-dimensional manifold is flat, so if $\dim \mathcal{C} = 1$, we may apply Proposition 6.4 to assess the metrizability of the induced connection. By comparing (20) and (35), we deduce that

$$\overset{\mathcal{C}}{\Gamma}{}^1_{11}(s) = -\Psi_2(s), \ \sigma_s(\operatorname{grad} P(s)) = -\Psi_1(s), \quad s \in \Theta \tag{52}$$

where we recall that Θ, defined in Proposition 3.4, is $\mathbb{R}$ or $\mathbb{S}^1$ depending on whether $\mathcal{C} \simeq \mathbb{R}$ or $\mathcal{C} \simeq \mathbb{S}^1$.

If $\mathcal{C} \simeq \mathbb{R}$, then $\pi_1(\mathcal{C}, q)$ is trivial, and Proposition 6.4 is trivially satisfied. Thus a connection on $\mathbb{R}$ is always metrizable. Using $x \in \mathbb{R}$ as coordinate for $\mathcal{C}$, the Riemannian metric on $\mathbb{R}$ will have the form $g_x(v, w) = (1/2)k(x)vw$, with $k(x) > 0$. By Theorem 5.1, the constrained dynamics are Lagrangian if and only if there exists a function $P_{\mathcal{C}} : \mathbb{R} \to \mathbb{R}$ such that $\sigma_x(\operatorname{grad} P(x)) = k^{-1}(x)P'_{\mathcal{C}}(x)$. This identity is satisfied by letting $P_{\mathcal{C}}$ be an antiderivative of the function $k(x)\sigma_x(\operatorname{grad} P(x))$. Having established the metrizability of the induced connection and the existence of $P_{\mathcal{C}}$, by Theorem 5.1 the constrained dynamics are always Lagrangian. This recovers the result of Theorem 3.5 when $\mathcal{C} \simeq \mathbb{R}$.

Now consider the case $\mathcal{C} \simeq \mathbb{S}^1$, so that $\Theta = \mathbb{S}^1$. Since $\pi_1(\mathbb{S}^1, 0) = (\mathbb{Z}, +)$, $\pi_1(\mathbb{S}^1, 0)$ is generated by the loop $\gamma_0 : [0 \ 2\pi] \to \mathbb{S}^1$, $t \mapsto t \bmod 2\pi$. By Proposition 6.4, the induced connection is metrizable if and only if there exists a positive definite quadratic form that is invariant under the transport map P_{γ_0}. Recall the coordinate representation of the parallel transport map in (48). Using $t \in \mathbb{R}$ as local coordinates for $\mathbb{S}^1$, we have that $\mathsf{P}_{\gamma_0}(v) = X(2\pi)$, where X is the solution of the linear time-varying ODE

$$\begin{aligned} \dot{X} &= \big(\Psi_2 \circ \pi(t)\big) X \quad t \in [0, 2\pi) \\ X(0) &= v. \end{aligned}$$

To obtain the above ODE, we substituted the first identity of (52) into (48), and used the fact that the coordinate representation of $\Psi_2(s)$ is $\Psi_2 \circ \pi(t)$, where $\pi(t) = t \bmod 2\pi$. The solution of the above scalar linear system is

$$\mathsf{P}_{\gamma_0}(v) = \left(\exp \int_0^{2\pi} \Psi_2 \circ \pi(z) dz \right) v.$$

We pick $s = 0$ as reference point on $\mathbb{S}^1$. Then, modulo a multiplicative positive scalar, the only positive definite bilinear form on $T_0\mathbb{S}^1 \times T_0\mathbb{S}^1$ is $g_0(v_0, w_0) = v_0 w_0$, and the invariance condition in Proposition 6.4 reads as

$$\exp \left(2 \int_0^{2\pi} \Psi_2 \circ \pi(z) dz \right) v_0 w_0 = v_0 w_0.$$

The above identity holds for arbitrary v_0, w_0 if and only if

$$\int_0^{2\pi} \Psi_2 \circ \pi(z) dz = 0, \tag{53}$$

or, equivalently, if the function $\hat{M}_{\mathcal{C}}(x)$ in Theorem 3.5 is 2π-periodic. Thus, the periodicity requirement on $\hat{M}_{\mathcal{C}}$ in part (b) of Theorem 3.5 is equivalent to the requirement, in part (i) of Theorem 5.1, that the induced connection be metrizable.

To find the Riemannian metric on $\mathbb{S}^1$ (denoted g in what follows), we extend the inner product $g_0 : T_0\mathbb{S}^1 \times T_0\mathbb{S}^1 \to \mathbb{R}$ to the whole $T\mathbb{S}^1 \times T\mathbb{S}^1$ through parallel transport, as in (50). For any $s \in \mathbb{S}^1$, set $g_s(v_s, w_s) := g_0(\mathsf{P}_{\gamma_s^0}(v_s), \mathsf{P}_{\gamma_s^0}(w_s))$, where γ_s^0 is an arbitrary curve from s to 0 in $\mathbb{S}^1$. For instance, pick any $x \in \pi^{-1}(s)$, and define $\gamma_s^0 : [0,x] \to \mathbb{S}^1$ as $\gamma_s^0(t) = s - (t \bmod 2\pi)$. Then, $\mathsf{P}_{\gamma_s^0}(v_s) = \exp\big(-\int_0^x \Psi_2 \circ \pi(z)dz\big)v_s$. Using $\mathsf{P}_{\gamma_s^0}$, the Riemannian metric on $\mathbb{S}^1$ is

$$g_s(v_s, w_s) = \exp\left(-2\int_0^{\pi^{-1}(s)} \Psi_2 \circ \pi(z)dz\right).$$

The above metric on $\mathbb{S}^1$ gives the kinetic energy of the constrained dynamics in Theorem 3.5, since it can be expressed as $g_s(v_s, w_s) = M_{\mathcal{C}}(s)v_s w_s$.

Next, we turn our attention to condition (ii) of Theorem 5.1, namely the existence of $P_{\mathcal{C}} : \mathbb{S}^1 \to \mathbb{R}$ such that $\sigma(\operatorname{grad} P) = \operatorname{grad}_{\mathcal{C}} P_{\mathcal{C}}$, or

$$-\Psi_1(s) = \frac{1}{M_{\mathcal{C}}(s)} P'_{\mathcal{C}}(s)$$

Equivalently, we need to check when is it that the one-form on $\mathbb{S}^1$ $-\Psi_1(s)M_{\mathcal{C}}(s)ds$ is exact. This is the case if and only if the integral of the form along $\mathbb{S}^1$ is zero,

$$0 = \int_{\mathbb{S}^1} \Psi_1(s) M_{\mathcal{C}}(s) ds = \int_0^{2\pi} \big(\Psi_1 \circ \pi(\tau)\big)\hat{M}_{\mathcal{C}}(\tau)d\tau.$$

This is precisely the condition that the function $\hat{P}_{\mathcal{C}}(x)$ in Theorem 3.5 be 2π-periodic. Thus, the periodicity requirement on $\hat{P}_{\mathcal{C}}$ in part (b) of Theorem 3.5 is equivalent to the requirement, in part (ii) of Theorem 5.1, that $\sigma(\operatorname{grad} P) = \operatorname{grad}_{\mathcal{C}} P_{\mathcal{C}}$. We have thus shown that Theorem 3.5 is a corollary of Theorem 5.1 and Proposition 6.4.

6.6. **Two-dimensional manifolds.** When $\dim \mathcal{C} = 2$, the metrizability of an affine connection has a powerful characterization in terms of the Ricci tensor [34, 36]. We remark that the situation $\dim \mathcal{C} = 2$ arises in Lagrangian control systems with degree of underactuation two, when a regular VHC of codimension two is enforced. The Ricci curvature tensor (see, e.g., [19]) is the $(0,2)$ tensor defined as

$$\operatorname{Ric}(Y_p, Z_p) = \operatorname{trace}(X_p \mapsto R(X_p, Y_p)Z_p),\ p \in \mathcal{C}, X_p, Y_p \in T_p\mathcal{C},$$

where $\operatorname{trace}(\cdot)$ denotes the trace of a linear map. If the affine connection ∇ is Riemannian with respect to a metric g, then the Ricci tensor is proportional to the metric,

$$\operatorname{Ric}(X, Y) = g(X, Y)K, \tag{54}$$

where $K \in C^\infty(\mathcal{C})$ denotes the Gaussian curvature of $\mathcal{C}$ (see [19, Lemma 8.7]). Recall from (5) that the compatibility of ∇ with g means that $\nabla g = 0$. If ∇ has nonvanishing curvature, then K is nonvanishing, and setting $\alpha = 1/K$, we have

$$\nabla(\alpha \operatorname{Ric}) = \nabla g = 0. \tag{55}$$

For each $X, Y, Z \in \mathfrak{X}(M)$, we have

$$\nabla_X(\alpha \operatorname{Ric}(Y, Z)) = X(\alpha)\operatorname{Ric}(Y, Z) + \alpha \nabla_X \operatorname{Ric}(Y, Z),$$

and using (55) we deduce that

$$\nabla_X \operatorname{Ric}(Y, Z) = -\frac{d\alpha(X)}{\alpha}\operatorname{Ric}(Y, Z) = d(-\ln|\alpha|)(X)\operatorname{Ric}(Y, Z).$$

The above may be rewritten concisely using the total covariant derivative and the tensor product as

$$\nabla \operatorname{Ric} = d(-\ln|\alpha|) \otimes \operatorname{Ric}.$$

A tensor field F whose total covariant derivative satisfies $\nabla F = \omega \otimes F$, where ω is a one-form, is said to be *recurrent*. Thus, a necessary condition for metrizability of

∇ is that the Ricci tensor induced by ∇ be recurrent, with a one-form ω given by the exact differential of a function in $C^\infty(\mathcal{C})$. Further, in light of the fact that when ∇ is metrizable identity (54) holds, another necessary condition for metrizability is that the Ricci tensor be definite (positive definite if $K > 0$, negative definite if $K < 0$). Together, these conditions are also sufficient.

Theorem 6.7 ([34],[36][7]). *Let $\mathcal{C}$ be a two-dimensional connected manifold and ∇ a symmetric affine connection on $\mathcal{C}$ such that the curvature induced by $\mathcal{C}$ is nowhere zero. Then ∇ is metrizable if and only if the Ricci tensor induced by ∇ is definite and recurrent, with the corresponding one-form being exact. If this is the case, and $\nabla \operatorname{Ric} = df \otimes \operatorname{Ric}$ holds for some $f \in C^\infty(\mathcal{C})$, then all Riemannian metrics compatible with ∇ are given by*

$$g = \pm \exp(-f + b) \operatorname{Ric}, \;\; b \in \mathbb{R} \textit{ arbitrary},$$

with plus sign if Ric *is positive definite, and minus sign otherwise.*

The idea behind the proof of sufficiency rests upon the fact that if $\nabla \operatorname{Ric} = df \otimes \operatorname{Ric}$, then $\nabla\big(\exp(-f+b) \operatorname{Ric}\big) = 0$, and therefore both type $(0, 2)$ tensor fields given by $\pm \exp(-f+b) \operatorname{Ric}$ are compatible with ∇. Since ∇ is symmetric, so is Ric. Since Ric is definite, g in the theorem statement is positive definite.

7. **Examples.** We now illustrate the results of Sections 5 and 6 with three examples. First, we revisit Examples 1 and 2. Then we investigate the Lagrangian structure of a double pendulum on a cart subject to a regular VHC of order 1.

Example 1 (Continued)**.** Consider again the dynamics of the planar point-mass of Example 1, a Lagrangian control system with underactuation degree one. The constrained dynamics on the unit circle $\mathcal{C}$ depicted in Figure 4 are

$$\ddot{s} = -(\tan\alpha)\dot{s}^2, \;\; s \in \mathbb{S}^1.$$

We want to determine the existence of a Lagrangian structure for these constrained dynamics. For this, we may use Theorem 3.5, with $\Psi_1(s) = 0$ and $\Psi_2(s) = -\tan\alpha$. We have

$$\begin{aligned}\hat{M}_{\mathcal{C}}(x) &= \exp\left(-2\int_0^x \tan(\alpha) dz\right) = \exp(-2\tan(\alpha)x),\\ \hat{P}_{\mathcal{C}}(x) &= 0.\end{aligned}$$

The constrained dynamics are Lagrangian if and only if $\hat{M}_{\mathcal{C}}$ is 2π-periodic, or $\alpha = \pm\pi/2 \bmod 2\pi$. Thus the constrained point-mass is Lagrangian if and only if the control force is orthogonal to the circle $\mathcal{C}$, in which case the Lagrangian is $L(s, \dot{s}) = (1/2)\dot{s}^2$. As predicted by Proposition 5.2, $L(s, \dot{s})$ is the restriction of the original Lagrangian to $\mathcal{C}$, i.e., $L(s, \dot{s}) = (1/2)g(\dot{q}, \dot{q})|_{\dot{q}=\varphi'(s)\dot{s}}$. △

Example 2 (Continued)**.** We return now to the unit mass of Example 2, a Lagrangian control system with degree of underactuation two. We seek to determine whether or not the constrained dynamics with coordinate representation given in (38) are Lagrangian. Since this Lagrangian control system has no potential function, by Theorem 5.1 we only need to check whether or not the induced connection on $\mathcal{C}$ is metrizable. To this end, we will use Theorem 6.7. Since $\mathcal{C}$ is simply connected and the induced connection is real analytic, it suffices to check the recurrence condition of Theorem 6.7 in local coordinates (see Remark 6.6).

[7] In these papers, the authors investigate the existence of non-degenerate metrics, whereas we look for positive definite metrics. The statement of the theorem has been adapted accordingly.

In local coordinates $(s^1, s^2) \in W \subset \mathbb{R}^2$, we have the frame $\{\partial_1, \partial_2\}$ given by the natural basis of $\mathbb{R}^2$. We first compute the coefficients R^l_{ijk} of the curvature endomorphism associated with $\overset{c}{\nabla}$ in (s^1, s^2)-coordinates via the formula[8] $R^l_{ijk} = ds^l \left(\overset{c}{\nabla}_{\partial_i} \overset{c}{\nabla}_{\partial_j} \partial_k - \overset{c}{\nabla}_{\partial_j} \overset{c}{\nabla}_{\partial_i} \partial_k \right)$, $i, j, k, l \in \{1, 2\}$. Using (6) for the evaluation of $\overset{c}{\nabla}$ we obtain

$$R^l_{ijk} = \partial_{s^i} \overset{c}{\Gamma}{}^l_{jk} - \partial_{s^j} \overset{c}{\Gamma}{}^l_{ik} + \sum_m \left(\overset{c}{\Gamma}{}^m_{jk} \overset{c}{\Gamma}{}^l_{im} - \overset{c}{\Gamma}{}^m_{ik} \overset{c}{\Gamma}{}^l_{jm} \right), \tag{56}$$

where $\overset{c}{\Gamma}{}^k_{ij}$ are the Christoffel symbols in (33).

The coefficients Ric_{ij} of the Ricci tensor are then given by $\mathrm{Ric}_{ij} = \sum_k R^k_{kij}$. Performing these computations, we get

$$\mathrm{Ric}_{11} = \frac{1}{\cos^2(s^1) + 1}, \ \mathrm{Ric}_{12} = \mathrm{Ric}_{21} = 0, \ \mathrm{Ric}_{22} = \frac{2\sin^2(s^1)}{(\sin^2(s^1) - 2)^2}.$$

Next, using the total covariant differentiation of tensors, (2), (3), and the Christoffel symbols (33), we compute the coefficients $(\overset{c}{\nabla} \mathrm{Ric})_{ijk}$ of $\overset{c}{\nabla} \mathrm{Ric}$ by means of

$$(\overset{c}{\nabla} \mathrm{Ric})_{ijk} = \overset{c}{\nabla}_{\partial_i} \mathrm{Ric}(\partial_j, \partial_k).$$

By so doing, we find that the only nonzero coefficients $(\overset{c}{\nabla} \mathrm{Ric})_{ijk}$ are

$$(\overset{c}{\nabla} \mathrm{Ric})_{111} = \frac{2\sin(2s^1)}{(\cos^2(s^1) + 1)^2}, \ (\overset{c}{\nabla} \mathrm{Ric})_{122} = -\frac{4\sin(2s^1)\sin^2(s^1)}{(\sin^2(s^1) - 2)^3}. \tag{57}$$

Next, we check whether or not $\overset{c}{\nabla} \mathrm{Ric} = df \otimes \mathrm{Ric}$ for a suitable smooth function f. Consider a generic one-form ω on $W \subset \mathbb{R}^2$, $\omega = \omega_1 ds^1 + \omega_2 ds^2$. The nonzero coefficients $(\omega \otimes \mathrm{Ric})_{ijk}$ of the tensor product $\omega \otimes \mathrm{Ric}$ are

$$\begin{aligned} (\omega \otimes \mathrm{Ric})_{111} &= \frac{\omega_1}{\cos^2(s^1) + 1}, \quad (\omega \otimes \mathrm{Ric})_{122} = \frac{2\omega_1 \sin^2(s^1)}{(\sin^2(s^1) - 2)^2}, \\ (\omega \otimes \mathrm{Ric})_{211} &= \frac{\omega_2}{\cos^2(s^1) + 1}, \quad (\omega \otimes \mathrm{Ric})_{222} = \frac{2\omega_2 \sin^2(s^1)}{(\sin^2(s^1) - 2)^2}. \end{aligned} \tag{58}$$

By comparing (57) and (58), we see that $\overset{c}{\nabla} \mathrm{Ric} = \omega \otimes \mathrm{Ric}$ if and only if $\omega_1 = 2\sin(2s^1)/(\cos^2(s^1) + 1)$ and $\omega_2 = 0$, i.e.,

$$\omega = \frac{2\sin(2s^1)}{\cos^2(s^1) + 1} ds^1.$$

The one-form ω is exact, $\omega = df$, with $f(s) = -4\,\mathrm{atanh}(\sin^2(s^1)/(\sin^2(s^1) - 4))$. By Theorem 6.7, the constrained dynamics on $T\mathcal{C}$ are Lagrangian, and a Lagrangian function in local coordinates is given by

$$L(s, \dot{s}) = (1/2)\dot{s}^\top D_{\mathcal{C}}(s)\dot{s},$$

where $D_{\mathcal{C}}(s)$ is the matrix $\exp(-f(s))[\mathrm{Ric}]$, with $[\cdot]$ denoting the matrix representation of the Ricci tensor. Specifically, we have

$$D_{\mathcal{C}}(s) = \begin{bmatrix} 1/2 - \sin^2(s^1)/4 & 0 \\ 0 & \sin^2(s^1)/2 \end{bmatrix}.$$

[8]Here we use the fact that $[\partial_i, \partial_j] = 0$.

In applying Theorem 6.7, we used the plus sign in the metric because the matrix [Ric] is positive definite. One may check that the Euler-Lagrange equation with L as above gives the constrained dynamics (38). We stress once again that although our computations are done in local coordinates, by the argument in Remark 6.6 the foregoing considerations imply that the constrained dynamics are *globally Lagrangian.* $\triangle$

Example 3. Consider the double pendulum on a cart depicted in Figure 8, a Lagrangian control system with three degrees-of-freedom and one input. We investigate two cases.

- **Case (a):** the control input is the force imparted to the cart.
- **Case (b):** the control input is the torque imparted on the second revolute joint.

We assume that the pendulum rods are massless, that the masses of the two pendulums are unitary, and that the rod lengths are unitary as well. Using $q =$

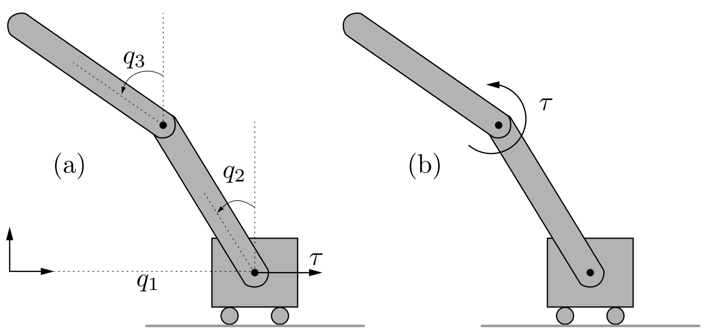

FIGURE 8. The double pendulum on a cart of Example 3. Case (a): control force on the cart. Case (b): control torque on the last joint. The orthogonal frame in the figure is the inertial reference frame.

$(q_1, q_2, q_3) \in \mathbb{R} \times \mathbb{S}^1 \times \mathbb{S}^1$ as generalized coordinates, the Lagrangian of the system is $L(q, \dot{q}) = \frac{1}{2} g_q(\dot{q}, \dot{q}) - P(q)$, with

$$g_q(\dot{q}, \dot{q}) = \dot{q}^\top \begin{bmatrix} 3 & -2\cos(q_2) & -\cos(q_3) \\ -2\cos(q_2) & 2 & \cos(q_2 - q_3) \\ -\cos(q_3) & \cos(q_2 - q_3) & 1 \end{bmatrix} \dot{q},$$

and

$$P(q) = (2\cos(q_2) + \cos(q_3))G,$$

where G is the gravitational constant. In generalized coordinates, the control force is the vector $B(q)\tau = [1\ 0\ 0]^\top \tau$ (case (a)) or $[0\ 0\ 1]^\top \tau$ (case (b)), where $\tau \in \mathbb{R}$ is the control input. Letting $\mathcal{Q} = \mathbb{R} \times \mathbb{S}^1 \times \mathbb{S}^1$, we have a Lagrangian control system $(\mathcal{Q}, g, P, F)$, where $F = dq_1$ or $F = dq_3$, respectively.

Consider the embedded submanifold of $\mathcal{Q}$,

$$\mathcal{C} = \{q \in \mathcal{Q} : q_3 = \rho(q_2)\},$$

where $\rho : \mathbb{S}^1 \to \mathbb{S}^1$ is the smooth function

$$\rho(q_2) = q_2 + 2\arctan\big((1 + \sqrt{2})\tan(-q_2/2)\big).$$

The configuration of the double pendulum on $\mathcal{C}$ is illustrated in Figure 9. The function ρ above was already used in [10] for path following control of a PVTOL aircraft, and in [21] for pendubot swing-up. In the context of the double pendulum on a cart of Figure 8, the function ρ induces the interesting property that, on $\mathcal{C}$, the last link does not perform full revolutions and remains confined to the upper half-plane.

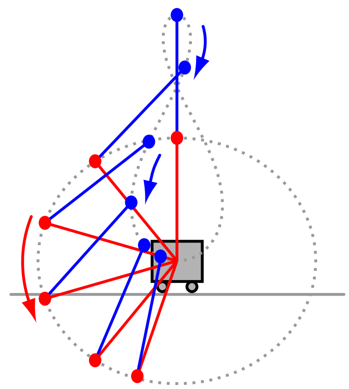

FIGURE 9. Configurations of the double pendulum on the VHC $\mathcal{C}$ of Example 3. The missing configurations on the right-hand side are deduced by symmetry with respect to the vertical axis.

Letting $h(q) = q_3 - \rho(q_3)$, one may check that, in both cases (a) and (b), $dh_q D^{-1}(q)B(q) \neq 0$ for all $q \in \mathcal{C}$, so by the equivalence of Definitions 3.2 and 4.3, $\mathcal{C}$ is a regular VHC. The set $\mathcal{C}$ is diffeomorphic to a cylinder via the diffeomorphism $\mathbb{R} \times \mathbb{S}^1 \to \mathcal{C}$, $(s^1, s^2) \mapsto (s^1, s^2, \rho(s^2))$. Using this global parametrization and the formulas in (33), one may show that, in both cases (a) and (b), the only nonzero Christoffel symbols of $\overset{\mathcal{C}}{\nabla}$ are $\overset{\mathcal{C}}{\Gamma}{}^1_{22}(s^2)$ and $\overset{\mathcal{C}}{\Gamma}{}^2_{22}(s^2)$, and they are functions of s^2 only. Similarly, the representation of $\sigma(\operatorname{grad} P)$ in (s^1, s^2)-coordinates is a function of s^2 only. Therefore, in both cases (a) and (b) the constrained dynamics (34) have the form

$$\begin{aligned} \ddot{s}^1 &= -\overset{\mathcal{C}}{\Gamma}{}^1_{22}(s^2)(\dot{s}^2)^2 - \lambda_1(s^2) \\ \ddot{s}^2 &= -\overset{\mathcal{C}}{\Gamma}{}^2_{22}(s^2)(\dot{s}^2)^2 - \lambda_2(s^2), \end{aligned} \tag{59}$$

where $(\lambda_1(s^2), \lambda_2(s^2))$ is the coordinates representation of $\sigma(\operatorname{grad} P)$. The precise expressions are easy to determine with the formulas in Section 4.3, but they are too long to report here. The function ρ is odd, i.e., $\rho(-s^2) = -\rho(s^2)$, and as a result, the functions $\overset{\mathcal{C}}{\Gamma}{}^1_{22}(s^2), \overset{\mathcal{C}}{\Gamma}{}^2_{22}(s^2), \lambda_i(s^2)$ are odd as well.

Now we investigate the Lagrangian nature of the constrained dynamics. We will show that in case (a) (force on the cart), the constrained dynamics are *not* Lagrangian, while in case (b) (torque on the second revolute joint), they are. We begin by checking condition (i) of Theorem 5.1, i.e., the metrizability of $\overset{\mathcal{C}}{\nabla}$. The coefficients of the curvature endomorphism in (s^1, s^2) coordinates may be computed using the formula (56). Owing to the fact that only the symbols $\overset{\mathcal{C}}{\Gamma}{}^k_{22}$ are nonzero, and that they are functions of s^2 only, we see from (56) that the curvature endomorphism is identically zero, i.e., the induced connection is flat. We can then use Proposition 6.4 to determine whether or not the induced connection is metrizable. Recall that $(s^1, s^2) \in \mathbb{R} \times \mathbb{S}^1$. In what follows, the point $(0,0) \in \mathbb{R} \times \mathbb{S}^1$ will be denoted 0 in subscripts.

The generator of the first homotopy group $\pi_1(\mathbb{R} \times \mathbb{S}^1, (0,0))$ is $[\gamma_0]$, where $\gamma_0 : [0, 2\pi] \to \mathbb{R} \times \mathbb{S}^1$ is the curve $t \mapsto (0, t \bmod 2\pi)$. In light of Proposition 6.4, we seek a positive definite quadratic form that is invariant under the transport map P_{γ_0}. With reference to (48), to find P_{γ_0} we solve the linear time-varying system

$$\dot{X} = \begin{bmatrix} 0 & -\overset{\mathcal{C}}{\Gamma}{}^1_{22}(t) \\ 0 & -\overset{\mathcal{C}}{\Gamma}{}^2_{22}(t) \end{bmatrix} X, \quad X(0) = v,$$

and put $\mathsf{P}_{\gamma_0}(v) = X(2\pi)$. The solution is $X(t) = (X^1(t), X^2(t))$ with

$$\begin{aligned} X^1(t) &= v_1 - v_2 \int_0^t \overset{\mathcal{C}}{\Gamma}{}^1_{22}(z) \exp\left(\int_0^z -\overset{\mathcal{C}}{\Gamma}{}^2_{22}(u) du \right) dz \\ X^2(t) &= v_2 \exp\left(\int_0^t -\overset{\mathcal{C}}{\Gamma}{}^2_{22}(z) dz \right). \end{aligned} \tag{60}$$

Denote $I_1(t) := -\int_0^t \overset{\mathcal{C}}{\Gamma}{}^2_{22}(z) dz$ and $I_2(t) := \int_0^t \overset{\mathcal{C}}{\Gamma}{}^1_{22}(z) \exp\big(I_1(z)\big) dz$. Then,

$$\mathsf{P}_{\gamma_0}(v) = \begin{bmatrix} 1 & -I_2(2\pi) \\ 0 & \exp\big(I_1(2\pi)\big) \end{bmatrix} v.$$

Recall that the functions Γ^k_{22} are odd and 2π-periodic. The integral over one period of an odd periodic function is zero. Using this fact, we have $I_1(2\pi) = 0$. Since the integral of an odd function is an even function, $\exp(I_1(t))$ is even, and its product with $\overset{\mathcal{C}}{\Gamma}{}^1_{22}(t)$ is odd. Thus $I_2(2\pi) = 0$. In conclusion, P_{γ_0} is the identity map, implying that any positive definite quadratic form on $(T_{(0,0)}\mathbb{R} \times \mathbb{S}^1) \times (T_{(0,0)}\mathbb{R} \times \mathbb{S}^1)$ is invariant under P_{γ_0}. By Proposition 6.4, $\overset{\mathcal{C}}{\nabla}$ is metrizable. This result holds for both cases (a) and (b).

Next, we find all Riemannian metrics on $\mathbb{R} \times \mathbb{S}^1$ compatible with $\overset{\mathcal{C}}{\nabla}$. Just like in Section 6.5, we will use the notation g (in place of $g_{\mathcal{C}}$) for such a metric. Modulo scalar multiples, the generic positive definite quadratic form on $(T_{(0,0)}\mathbb{R} \times \mathbb{S}^1) \times (T_{(0,0)}\mathbb{R} \times \mathbb{S}^1)$ is

$$g_0(v_0, w_0) = v_0^\top \begin{bmatrix} 1 & a \\ a & b \end{bmatrix} w_0,$$

with $a, b \in \mathbb{R}$ such that $b > a^2$. As in Section 6.5, the Riemannian metric on $\mathbb{R} \times \mathbb{S}^1$ is found by parallel transporting g_0 by means of (50). The construction is illustrated in Figure 10.

Let $s = (s^1, s^2) \in \mathbb{R} \times \mathbb{S}^1$ be arbitrary, denote $\bar{s} := (0, s^2)$. Pick any $x^2 \in \pi^{-1}(s^2)$ and define a path $\gamma^s_0(t)$ as the concatenation of paths $\gamma^{\bar{s}}_0$ and $\gamma^s_{\bar{s}}$, where $\gamma^{\bar{s}}_0 : [0, x^2] \to \mathbb{R} \times \mathbb{S}^1$ is defined as $\gamma^{\bar{s}}_0(t) := (0, t \bmod 2\pi)$, and $\gamma^s_{\bar{s}} : [0, s^1] \to \mathbb{R} \times \mathbb{S}^1$ is defined as $\gamma^s_{\bar{s}}(t) := (t, s^2)$. Then, $\gamma^s_0 : [0, s^1 + x^2] \to \mathbb{R} \times \mathbb{S}^1$ is a piecewise smooth path connecting $(0,0)$ to (s^1, s^2). See Figure 10. By Proposition 6.1, we have $\mathsf{P}_{\gamma^s_0} = \mathsf{P}_{\gamma^{\bar{s}}_0} \circ \mathsf{P}_{\gamma^s_{\bar{s}}}$. Since $\gamma^s_{\bar{s}}$ is a translation along the real line, $\mathsf{P}_{\gamma^s_{\bar{s}}}$ is the identity

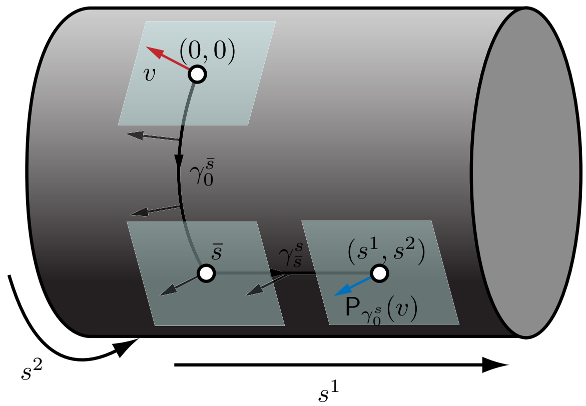

FIGURE 10. Parallel transport on $\mathbb{R} \times \mathbb{S}^1$ from $(0, 0)$ to (s^1, s^2).

map. On the other hand, $\mathsf{P}_{\gamma_0^{\bar{s}}}$ is given by (60) at time s^2. In conclusion, the parallel transport map from $(0, 0)$ to (s^1, s^2) is

$$\mathsf{P}_{\gamma_0^s} = \begin{bmatrix} 1 & -I_2(s^2) \\ 0 & \exp\big(I_1(s^2)\big) \end{bmatrix}.$$

By Proposition 6.1, the parallel transport map from (s^1, s^2) to $(0, 0)$ is $\mathsf{P}_{\gamma_0^s}^{-1}$. Now we define a Riemannian metric g on $\mathbb{R} \times \mathbb{S}^1$ by transporting tangent vectors in $T_{(s^1,s^2)}(\mathbb{R} \times \mathbb{S}^1)$ to $T_{(0,0)}(\mathbb{R} \times \mathbb{S}^1)$ (cf. (50)):

$$\begin{aligned} g(v_s, w_s) &:= g_0\big(\mathsf{P}_{\gamma_0^s}^{-1}(v_s), \mathsf{P}_{\gamma_0^s}^{-1}(w_s)\big) \\ &= v_s^\top \begin{bmatrix} 1 & -I_2(s^2) \\ 0 & \exp\big(I_1(s^2)\big) \end{bmatrix}^{-\top} \begin{bmatrix} 1 & a \\ a & b \end{bmatrix} \begin{bmatrix} 1 & -I_2(s^2) \\ 0 & \exp\big(I_1(s^2)\big) \end{bmatrix}^{-1} w_s \\ &= v_s^\top D_{\mathcal{C}}(s^2) w_s, \end{aligned} \tag{61}$$

where

$$D_{\mathcal{C}}(s^2) := \begin{bmatrix} 1 & \exp(-I_1(s^2))(I_2(s^2)+a) \\ \exp(-I_1(s^2))(I_2(s^2)+a) & \exp(-2I_1(s^2))(I_2(s^2)+2aI_2(s^2)+b) \end{bmatrix}.$$

By this construction, for any (a, b) with $b > a^2$, g is a Riemannian metric on $\mathbb{R} \times \mathbb{S}^1$, and $\overset{\mathcal{C}}{\nabla}$ is the Levi-Civita connection of g. This result holds for both cases (a) and (b).

Next we turn to condition (ii) of Theorem 5.1, namely the existence of a smooth function $P_{\mathcal{C}} : \mathcal{C} \to \mathbb{R}$ such that $\sigma(\operatorname{grad} P) = \operatorname{grad}_{\mathcal{C}} P_{\mathcal{C}}$ or, in (s^1, s^2)-coordinates, a function $P_{\mathcal{C}} : \mathbb{R} \times \mathbb{S}^1 \to \mathbb{R}$ such that

$$\begin{bmatrix} \lambda_1(s^2) \\ \lambda_2(s^2) \end{bmatrix} = D_{\mathcal{C}}^{-1}(s^2) \nabla_s P_{\mathcal{C}},$$

where λ_1, λ_2 are the functions in the constrained dynamics (59). Equivalently, we seek a function $P_{\mathcal{C}}$ such that

$$\begin{aligned} (dP_{\mathcal{C}})_s = {} & \big[\lambda_1(s^2) + \exp(-I_1(s^2))(I_2(s^2)+a)\lambda_2(s^2)\big] ds^1 \\ & + \big[\exp(-I_1(s^2))(I_2(s^2)+a)\lambda_1(s^2) \\ & + \exp(-2I_1(s^2))(I_2(s^2)+2aI_2(s^2)+b)\lambda_2(s)\big] ds^2. \end{aligned} \tag{62}$$

Such a function $P_{\mathcal{C}}$ exists if and only if the one-form on the right-hand side of the above identity is exact. For this, we need to check whether this one-form is closed, and whether its integral over the loop γ_0 defined earlier is zero. As far as closedness of the form is concerned, we need to check whether or not there exists $a \in \mathbb{R}$ such that

$$\partial_{s^2}\left[\lambda_1(s^2) + \exp(-I_1(s^2))(I_2(s^2)+a)\lambda_2(s^2)\right] = 0. \tag{63}$$

In case (a) when the control force is on the cart, one can show that there is no value of a for which (63) holds, whereas in case (b), when there is a control torque on the last joint, (63) is satisfied with $a = -1/2$. In the latter case, the one-form on the right-hand side of (62) is also exact because its components are odd functions of s^2. We choose any $b > a^2$, for instance $b = 1$, and obtain that the constrained dynamics in (59) are a Lagrangian system $(\mathbb{R} \times \mathbb{S}^1, g, P_{\mathcal{C}})$, with g given in (61) and

$$P_{\mathcal{C}}(s) = \int_{\gamma_0^s} dP_{\mathcal{C}},$$

with $dP_{\mathcal{C}}$ given in (62), and γ_0^s defined earlier.

We summarize our results for this example in the following table.

	$q_3 = \rho(q_2)$ regular?	$\overset{\mathcal{C}}{\nabla}$ metr'le?	$P_{\mathcal{C}}$ exists?	Lagr. exists?
Force on cart	yes	yes	no	no
Force on last joint	yes	yes	yes	yes

Figure 11 depicts the $(s^2, \dot{s}^2)$ orbits (equivalently, the $(q_2, \dot{q}_2)$ orbit) of a few solutions of the constrained system in cases (a) and (b). In both cases, we observe two types of behaviours: there are trajectories along which q_2 exhibits a rocking motion around π, and others along which q_2 performs full revolutions. The behaviour of q_1, not shown in the figure, is a drifting motion with bounded, sign-definite speed. $\triangle$

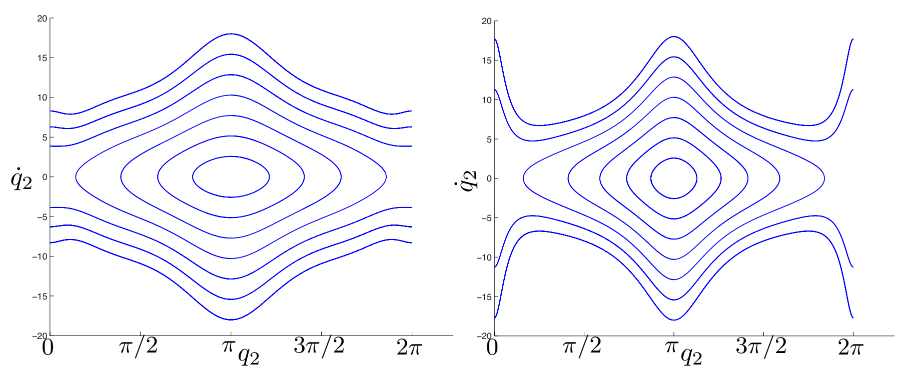

FIGURE 11. $(q_2, \dot{q}_2)$ orbits of a few solutions of the double pendulum on a cart subject to the VHC $q_3 = \rho(q_2)$. On the left, case (a) (force on cart). On the right, case (b) (torque on last joint).

8. **Conclusions.** We introduced a coordinate-free framework of virtual holonomic constraints for underactuated Lagrangian control systems, exposing the role of induced connections in the characterization of constrained dynamics. In this framework, the classical mechanics notion of ideal holonomic constraint becomes the special case in which the acceleration distribution is orthogonal to the VHC. We showed that, generally, the constrained dynamics are not Lagrangian, and the metrizability of the induced connection is key for the existence of a Lagrangian structure. When the constrained dynamics are forced (i.e., when the order of the regular VHC is less than the number of control inputs), the problem remains open of determining when the constrained dynamics are feedback equivalent to a Lagrangian control system. One possible avenue of investigation for the solution of this latter problem is to globalize the local theory of [29] in the context of affine connection control systems.

Appendix A. **Proof of Lemma 4.5.** We begin by observing that if $X \in \mathfrak{X}(\mathcal{Q})$, then for each $Y_q \in T\mathcal{Q}$, $\mathrm{vlft}(X)(Y_q) \in \operatorname{Ker} d\pi_{Y_q}$. Indeed,

$$\begin{aligned} d\pi_{Y_q}\big(\,\mathrm{vlft}(X)(Y_q)\big) &= d\pi_{Y_q}\big((d/dt)|_{t=0}(Y_q + tX(q))\big) \\ &= \frac{d}{dt}\Big|_{t=0}\pi(Y_q + tX(q)) = \frac{d}{dt}\Big|_{t=0} q = 0. \end{aligned}$$

Using the above and the property of geodesic sprays that $d\pi_{X_q}(S(X_q)) = X_q$, for each $X_q \in T\mathcal{C}$ we have

$$\begin{aligned} d\pi_{X_q}\big(S(X_q) - \mathrm{vlft}(\operatorname{grad} P)(X_q)\big) &= d\pi_{X_q}\big(S(X_q)\big) - d\pi_{X_q}\big(\,\mathrm{vlft}(\operatorname{grad} P)(X_q)\big) \\ &= X_q \in T_q\mathcal{C}. \end{aligned} \tag{64}$$

From (64) we deduce that

$$(\forall X_q \in T\mathcal{C})\ S(X_q) - \mathrm{vlft}(\operatorname{grad} P)(X_q) \in (d\pi_{X_q})^{-1}(T_q\mathcal{C}), \tag{65}$$

thus the proof of the lemma will be complete if we show that

$$(\forall X_q \in T\mathcal{C})\ (d\pi_{X_q})^{-1}(T_q\mathcal{C}) = T_{X_q}T\mathcal{C} \oplus \mathrm{vlft}(\mathcal{D}_A)(X_q). \tag{66}$$

Let $X_q \in T\mathcal{C}$ be arbitrary. Since $\dim((d\pi_{X_q})^{-1}(T_q\mathcal{C})) = 2n - m = (2n - 2m) + m = \dim(T_{X_q}T\mathcal{C}) + \dim(\mathrm{vlft}(\mathcal{D}_A)(X_q))$, to prove (66) we need to show that the subspaces $T_{X_q}T\mathcal{C}$ and $\mathrm{vlft}(\mathcal{D}_A)(X_q)$ are independent and contained in the subspace $(d\pi_{X_q})^{-1}(T_q\mathcal{C})$.

First we show that $T_{X_q}T\mathcal{C} \subset (d\pi_{X_q})^{-1}(T_q\mathcal{C})$. Let $Y_{X_q} \in T_{X_q}T\mathcal{C}$, then there exists a smooth curve in $T\mathcal{C}$, $\gamma : \mathbb{R} \to T\mathcal{C}$, such that $\gamma(0) = X_q$ and $\dot\gamma(0) = Y_{X_q}$. We have $d\pi_{X_q}(Y_{X_q}) = (d/dt)|_{t=0}\pi(\gamma(t))$. Since γ is a curve on $T\mathcal{C}$, $\pi(\gamma(t))$ is a curve on $\mathcal{C}$, and thus $(d/dt)|_{t=0}\pi(\gamma(t)) \in T_q\mathcal{C}$, which proves that $Y_{X_q} \in (d\pi_{X_q})^{-1}(T_q\mathcal{C})$.

Next we show that $\mathrm{vlft}(\mathcal{D}_A)(X_q) \subset (d\pi_{X_q})^{-1}(T_q\mathcal{C})$. This follows directly from the fact that $\mathrm{vlft}(\mathcal{D}_A)(X_q) \in \operatorname{Ker} d\pi_{X_q} \subset (d\pi_{X_q})^{-1}(T_q\mathcal{C})$.

Finally, let $Y_{X_q} \in T_{X_q}T\mathcal{C} \cap \mathrm{vlft}(\mathcal{D}_A)(X_q)$. Then there exists $F_q \in \mathcal{D}_A(q)$ such that $Y_{X_q} = (d/dt)|_{t=0}(X_q + tF_q)$. Moreover, $(d/dt)|_{t=0}(X_q + tF_q) \in T_{X_q}T\mathcal{C}$, which implies that $F_q \in T_q\mathcal{C}$. Since $F_q \in \mathcal{D}_A(q)$, the regularity condition (25) implies that $F_q = 0$, and therefore $Y_{X_q} = 0$. Thus the subspaces $T_{X_q}T\mathcal{C}$ and $\mathrm{vlft}(\mathcal{D}_A)(X_q)$ are independent, which shows that (66) holds. Together, (65) and (66) prove the lemma. □

REFERENCES

[1] W. Ambrose and I. Singer, A theorem on holonomy, *Transactions of the American Mathematical Society*, **75** (1953), 428–443.

[2] P. Appell, Exemple de mouvement d'un point assujetti à une liaison exprimèe par une relation non-linéaire entre les composantes de la vitesse, *Rend. Circ. Mat. Palermo*, **32** (1911), 48–50.

[3] V. I. Arnol'd, *Mathematical Methods of Classical Mechanics (Graduate Texts in Mathematics, Vol. 60)*, 2nd edition, Springer, 1989.

[4] L. Auslander and L. Markus, Holonomy of flat affinely connected manifolds, *Annals of Mathematics*, **62** (1955), 139–151.

[5] H. Beghin, *Étude Théorique des Compas Gyrostatiques Anschütz et Sperry*, PhD thesis, Faculté des sciences de Paris, 1922.

[6] W. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, vol. 120, 2nd edition, Academic press, 1986.

[7] F. Bullo and A. Lewis, *Geometric Control of Mechanical Systems*, Texts in Applied Mathematics, Springer, 2005.

[8] C. Chevallereau, J. Grizzle and C. Shih, Asymptotically stable walking of a five-link underactuated 3D bipedal robot, *IEEE Transactions on Robotics*, **25** (2008), 37–50.

[9] F. Clarke, Y. S. Ledyaev, R. Stern and P. Wolenski, *Nonsmooth Analyisis and Control Theory*, Graduate Texts in Mathematics, 178. Springer-Verlag, New York, 1998.

[10] L. Consolini, M. Maggiore, C. Nielsen and M. Tosques, Path following for the PVTOL aircraft, *Automatica*, **46** (2010), 1284–1296.

[11] L. Consolini and A. Costalunga, Induced connections on virtual holonomic constraints, in *IEEE Conference on Decision and Control (CDC)*, IEEE, 2015, 139–144.

[12] M. do Carmo, *Riemannian Geometry*, Birkhäuser Boston, 1992.

[13] L. Freidovich, A. Robertsson, A. Shiriaev and R. Johansson, Periodic motions of the pendubot via virtual holonomic constraints: Theory and experiments, *Automatica*, **44** (2008), 785–791.

[14] L. Godinho and J. Natário, *An Introduction to Riemannian Geometry: With Applications to Mechanics and Relativity*, Springer, 2014.

[15] A. Isidori, *Nonlinear Control Systems*, 3rd edition, Springer, New York, 1995.

[16] D. Jankuloski, M. Maggiore and L. Consolini, Further results on virtual holonomic constraints, in *Proceedings of the 4^{th} IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control, Bertinoro, Italy*, 2012.

[17] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, no. v. 1 in A Wiley Publication in Applied Statistics, Wiley, 1996.

[18] O. Kowalski, Metrizability of affine connections on analytic manifolds, *Note di Matematica*, **8** (1988), 1–11.

[19] J. Lee, *Riemannian Manifolds. An Introduction to Curvature*, Graduate texts in mathematics 0176, Springer, 1997.

[20] J. Lee, *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics, Springer, 2013.

[21] M. Maggiore and L. Consolini, Virtual holonomic constraints for Euler-Lagrange systems, *Automatic Control, IEEE Transactions on*, **58** (2013), 1001–1008.

[22] A. Mohammadi, M. Maggiore and L. Consolini, When is a Lagrangian control system with virtual holonomic constraints Lagrangian?, in *Proceedings of NOLCOS 2013*, **9** (2013), 512–517.

[23] A. Mohammadi, M. Maggiore and L. Consolini, On the lagrangian structure of reduced dynamics under virtual holonomic constraints, *ESAIM: Control, Optimisation and Calculus of Variations*, **23** (2017), 913–935.

[24] J. Nakanishi, T. Fukuda and D. Koditschek, A brachiating robot controller, *IEEE Transactions on Robotics and Automation*, **16** (2000), 109–123.

[25] H. Nijmeijer and A. van der Schaft, *Nonlinear Dynamical Control Systems.*, Springer - Verlag, New York, 1990.

[26] K. Nomizu and T. Sasaki, *Affine Differential Geometry: Geometry of Affine Immersions*, Cambridge University Press, 1994.

[27] F. Plestan, J. Grizzle, E. Westervelt and G. Abba, Stable walking of a 7-DOF biped robot, *IEEE Transactions on Robotics and Automation*, **19** (2003), 653–668.

[28] W. Poor, *Differential Geometric Structures*, McGraw-Hill Book Co., New York, 1981.

[29] S. Ricardo and W. Respondek, When is a control system mechanical?, *Journal of Geometric Mechanics*, **2** (2010), 265–302.

[30] B. Schmidt, Conditions on a connection to be a metric connection, *Communications in Mathematical Physics*, **29** (1973), 55–59.

[31] A. Shiriaev, J. Perram and C. Canudas-de-Wit, Constructive tool for orbital stabilization of underactuated nonlinear systems: Virtual constraints approach, *IEEE Transactions on Automatic Control.*, **50** (2005), 1164–1176.

[32] A. Shiriaev, A. Robertsson, J. Perram and A. Sandberg, Periodic motion planning for virtually constrained Euler-Lagrange systems, *Systems & Control Letters*, **55** (2006), 900–907.

[33] A. Shiriaev, L. Freidovich and S. Gusev, Transverse linearization for controlled mechanical systems with several passive degrees of freedom, *IEEE Transactions on Automatic Control*, **55** (2010), 893–906.

[34] G. Thompson, Local and global existence of metrics in two-dimensional affine manifolds, *Chinese Journal of Physics*, **29** (1991), 529–532.

[35] A. Vanžurová, Metrization problem for linear connections and holonomy algebras, *Archivum Mathematicum (BRNO)*, **44** (2008), 511–521.

[36] A. Vanžurová and P. Žáčková, Metrizability of connections on two-manifolds, *Acta Universitatis Palackianae Olomucensis. Facultas Rerum Naturalium. Mathematica*, **48** (2009), 157–170.

[37] E. Westervelt, J. Grizzle, C. Chevallereau, J. Choi and B. Morris, *Feedback Control of Dynamic Bipedal Robot Locomotion*, Taylor & Francis, CRC Press, 2007.

[38] E. Westervelt, J. Grizzle and D. Koditschek, Hybrid zero dynamics of planar biped robots, *IEEE Transactions on Automatic Control*, **48** (2003), 42–56.

Received September 2017; revised September 2018.

E-mail address: luca.consolini@unipr.it
E-mail address: alessandro.costalunga@studenti.unipr.it
E-mail address: maggiore@ece.utoronto.ca