CHARACTERIZATION OF GRADIENT CONTROL SYSTEMS[*]

JORGE CORTÉS[†], ARJAN VAN DER SCHAFT[‡], AND PETER E. CROUCH[§]

Abstract. Given a general nonlinear affine control system with outputs and a torsion-free affine connection defined on its state space, we investigate the gradient realization problem: we give necessary and sufficient conditions under which the control system can be written as a gradient control system corresponding to some pseudo-Riemannian metric whose Levi-Civita connection is equal to the given affine connection. The results rely on a suitable notion of compatibility of the system with respect to the given affine connection, and on the output behavior of the prolonged system and the gradient extension. The symmetric product associated with an affine connection plays a key role throughout the discussion.

Key words. gradient control systems, symmetric product, prolongation and gradient extension of a nonlinear system, externally equivalent systems.

AMS subject classifications. 93C10, 93B29, 53B05, 93B15

1. Introduction. A physically motivated class of nonlinear systems are *gradient control systems*, see [4, 9, 21, 22, 23, 24, 25] and the references quoted therein. These systems are described in the following way: they are nonlinear affine control systems, which are endowed with a pseudo-Riemannian metric on the state-space manifold. The drift vector field of the system is the gradient vector field associated with an internal potential function with respect to the pseudo-Riemannian metric, and the input vector fields are the gradient vector fields associated with the output functions of the system. Examples of gradient control systems include nonlinear electrical RLC networks, and dissipative systems where the inertial effects are neglected. In the case of RL or RC networks, the pseudo-Riemannian metric is positive-definite, and thus is a usual Riemannian metric, while for general RLC networks the metric is indefinite. We refer to [4, 9, 23, 24] for more background on the modeling of nonlinear networks as gradient systems.

Another relevant class of nonlinear systems is the family formed by the *Hamiltonian control systems*. In this case, the state-space manifold is equipped with a symplectic form. The drift vector field and the input vector fields are the Hamiltonian vector fields associated, respectively, to an internal energy function and the output functions of the system with respect to the symplectic form. Hamiltonian equations are of central importance in the modeling of physical systems as they are the starting point to describe the dynamics of a very large class of phenomena, including mechanical, electrical and electromagnetic systems.

Apart from their physical and engineering importance, gradient and Hamiltonian systems also possess very peculiar mathematical properties. For instance, an

[†]Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1308 W. Main St., Urbana, IL 61801, United States, Ph. +1 217 244-8734, Fax. +1 217 244-1653, `jcortes@uiuc.edu`, `http://motion.csl.uiuc.edu/~jorge`

[‡]Department of Applied Mathematics, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands, Ph. +31 53 489-3449, Fax +31 53 434-0733, `a.j.vanderschaft@math.utwente.nl`, `http://www.math.utwente.nl/~twarjan`

[§]Department of Electrical and Computer Engineering, Arizona State University, Tempe, AZ 85287, United States, Ph. +1 480 965-1722, Fax. +1 480 965-2267, `peter.crouch@asu.edu`, `http://www.eas.asu.edu/~sserc/people/crouch/crouch.html`

observable and controllable linear input-state-output system is a Hamiltonian control system [5] (respectively a gradient control system [25]) if and only if its impulse response matrix $W(t)$ satisfies $W(t) = -W^T(-t)$ (respectively $W(t) = W^T(t)$). Although they are typically not amenable to linearization techniques, their rich geometric structure makes possible to combine powerful tools from Nonlinear Control Theory, Differential Geometry and Classical Mechanics in the study of a variety of problems including stability and stabilization, input-output decoupling, structural synthesis and interconnection.

Their theoretical and practical relevance, together with their meaningful geometric properties and the wide range of results available for them, make the classes of Hamiltonian and gradient systems *distinct* within the family of nonlinear affine control systems. This explains the interest in identifying those systems that can be written as either Hamiltonian or gradient. This *characterization problem* is motivated by the Realization Problem in Systems Theory and the Inverse Problem in Mechanics. The Realization Problem addresses the question of when the input-output map of a system can be realized as the external behavior of a Hamiltonian (respectively gradient) input-output system. The Inverse Problem, which has a longstanding history in mathematical physics, poses the question of when a second-order differential equation can be realized as the Euler-Lagrange equations corresponding to certain Lagrangian function. For further reference on these problems, the reader is referred to [10, 14, 19, 20].

In [11, 12], necessary and sufficient conditions were given under which a minimal nonlinear affine control system with an equal number of inputs and outputs is a Hamiltonian control system with respect to some symplectic structure, which turned out to be unique. As we discuss below, there are a number of key differences in the treatment of the characterization problem for the Hamiltonian and the gradient case, which make the latter more involved. A fundamental observation is that, while every input-state-output system admits a natural extension to a Hamiltonian system living on the cotangent bundle of its state space, the construction of a gradient extension on the cotangent bundle relies on the selection of a torsion-free affine connection on the state space. This is why our starting point in the gradient setting is the selection of an *appropriate* torsion-free affine connection. This appropriateness is defined in terms of a novel *compatibility condition* of the given nonlinear system with the selected affine connection, guaranteeing an appropriate choice of the latter one. The compatibility condition is expressed as a relation of the symmetric products of the drift vector field and the input vector fields with the output functions of the system. As a further remark, the role played in the Hamiltonian setting by the Lie bracket and the Hamiltonian vector fields is taken here by the symmetric product associated with the given affine connection and the gradient vector fields.

The question solved by the main result of this paper (cf. Theorem 5.4 below) is the following: given a torsion-free affine connection which is compatible with the nonlinear control system, find necessary and sufficient conditions under which the system is gradient with respect to a pseudo-Riemannian metric whose Levi-Civita connection is the given affine connection. The question that still remains to be addressed in order to solve the full characterization problem for gradient control systems is the following: given a nonlinear control system, when does it exist an affine connection such that these necessary and sufficient conditions are satisfied?

The paper is organized as follows. In Section 2 we present the class of nonlinear systems considered along the paper. We also introduce the notions of prolongation

and gradient extension of a nonlinear system. The observability properties of these systems, studied in Section 3, together with the concept of (weakly) externally equivalent systems, introduced in Section 4, turn out to be key in establishing Theorem 5.4. In Section 5, we introduce the important notion of compatibility between a nonlinear system and a given affine connection. At this point, we are ready to state and prove the main result of the paper, namely the characterization of when a general nonlinear control system is gradient. This characterization can be roughly described as follows: under certain technical conditions, a nonlinear affine control system is gradient if and only if its prolongation and its gradient extension *behave* similarly (i.e., have the same input-output behavior). In Section 6 we investigate the uniqueness (up to isometry) of gradient realizations with the same input-output behavior and we give an alternative proof of a result in [1, 2]. We present our conclusions in Section 7. Finally, an appendix in Section 8 contains a simplifying result concerning the checkability of the compatibility condition for a nonlinear affine control system.

2. Setting. Let M be an n-dimensional (real-)analytic manifold. We will denote by TM, T^*M the tangent and cotangent bundles of M, by $\mathfrak{X}(M)$ the set of analytic vector fields on M, by $\Omega^1(M)$ the set of analytic one-forms on M, and by $C^\omega(M)$ the set of analytic functions on M. Consider a nonlinear control system Σ with state space M, affine in the inputs, and with an equal number of inputs and outputs,

$$\Sigma : \begin{cases} \dot{x} = g_0(x) + \sum_{j=1}^{m} u_j g_j(x) \,, \\ y_j = V_j(x) \,, \quad j = 1, \ldots, m \,, \end{cases} \tag{2.1}$$

where $x \in M$, $x(0) = x_0$ and $u = (u_1, \ldots, u_m) \in U \subset \mathbb{R}^m$. The vector fields $g_0, g_1, \ldots, g_m$ on M are assumed to be complete and $V_1, \ldots, V_m$ are real-valued functions on M. The set U is the control space, which for simplicity is assumed to be an open subset of $\mathbb{R}^m$, containing 0. The function $t \mapsto u(t) = (u_1(t), \ldots, u_m(t))$, that we will commonly denote as $u(\cdot)$, belongs to a certain class of functions of time, denoted by $\mathcal{U}$, called the *admissible controls*. For our purposes, we may restrict the admissible controls to be the piecewise constant right continuous functions.

An important subclass of the family of nonlinear systems (2.1) is formed by the Hamiltonian control systems, see [12]. Here, we will instead focus our attention on the family of gradient control systems. Let $\mathcal{G}$ be a pseudo-Riemannian metric on M, i.e. a non-degenerate symmetric (0,2)-tensor on M (not necessarily positive definite). Consider the 'musical' isomorphisms associated with $\mathcal{G}$, $\flat_{\mathcal{G}} : \mathfrak{X}(M) \to \Omega^1(M)$, $\sharp_{\mathcal{G}} : \Omega^1(M) \to \mathfrak{X}(M)$ defined by

$$\flat_{\mathcal{G}}(X)(Y) = \mathcal{G}(X, Y) \,, \quad \sharp_{\mathcal{G}}(\mu) = \flat_{\mathcal{G}}^{-1}(\mu) \,,$$

where $X, Y \in \mathfrak{X}(M)$ and $\mu \in \Omega^1(M)$. The *gradient vector field* associated with a function $V \in C^\omega(M)$ is given by $\mathrm{grad}_{\mathcal{G}} V = \sharp_{\mathcal{G}}(dV)$. Reciprocally, a vector field $X \in \mathfrak{X}(M)$ is said to be *locally gradient* if the one-form $\flat_{\mathcal{G}}(X)$ is closed. By Poincaré's lemma, this is equivalent to saying that there exists a locally defined function $V \in C^\omega(M)$ such that $\flat_{\mathcal{G}}(X) = dV$. If this equality holds globally, X is called *gradient* and will be denoted by $X = \mathrm{grad}_{\mathcal{G}} V$. Along the paper, we will drop the subindex when it is clear from the context the pseudo-Riemannian metric with respect to which the gradient vector field is computed. If we fix coordinates $(x^1, \ldots, x^n)$ on M, then the pseudo-Riemannian metric can be locally expressed as $\mathcal{G} = \mathcal{G}_{ab} dx^a \otimes dx^b$, where $(\mathcal{G}_{ab} = \mathcal{G}(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}))$ is a symmetric matrix. The musical isomorphisms are then given

by $\flat_{\mathcal{G}} = \mathcal{G}_{ab} dx^a \otimes dx^b$, $\sharp_{\mathcal{G}} = \mathcal{G}^{ab} \frac{\partial}{\partial x^a} \otimes \frac{\partial}{\partial x^b}$, where $(\mathcal{G}^{ab})$ is the inverse matrix of $(\mathcal{G}_{ab})$. Finally, the gradient vector field associated with V reads

$$\mathrm{grad}_{\mathcal{G}} V = \mathcal{G}^{ab} \frac{\partial V}{\partial x^b} \frac{\partial}{\partial x^a} \,.$$

Now, assume that the state space M in (2.1) is a pseudo-Riemannian manifold, $(M, \mathcal{G})$. Furthermore, assume that the drift vector field g_0 is locally gradient and the input vector fields g_j, $j = 1, \ldots, m$ are gradient with respect to the functions $V_1, \ldots, V_m$, i.e. $g_j = \mathrm{grad}_{\mathcal{G}} V_j$, $j = 1, \ldots, m$. Then, the resulting system

$$\Sigma : \begin{cases} \dot{x} = g_0(x) + \sum_{j=1}^{m} u_j(t) \, \mathrm{grad}_{\mathcal{G}} V_j(x) \,, \\ y_j = V_j(x) \,, \quad j = 1, \ldots, m \,, \end{cases} \tag{2.2}$$

is called a *locally gradient control system on M*. If the drift g_0 is a gradient vector field, then the system is called a *gradient control system on M*.

Given an affine connection on M, our objective is to characterize when a nonlinear system of the form (2.1) is a locally gradient control system (2.2), i.e. find necessary and sufficient conditions for the existence of a pseudo-Riemannian metric $\mathcal{G}$ on the state space M whose Levi-Civita connection is the given affine connection such that the system (2.1) equals system (2.2). These conditions will be given in terms of the output behavior of the so-called prolonged system and the gradient extension of Σ, which we describe next.

2.1. The prolongation of a nonlinear system. Given an initial state $x(0) = x_0$, take a coordinate neighborhood of M containing x_0. Let $t \in [0, T] \mapsto x(t)$ be the solution of (2.1) corresponding to the input function $t \in [0, T] \mapsto u(t) = (u_1(t), \ldots, u_m(t))$ and the initial state $x(0) = x_0$, such that $x(t)$ remains within the selected coordinate neighborhood. Denote the resulting output by $t \in [0, T] \mapsto y(t) = (y_1(t), \ldots, y_m(t))$, with $y_j(t) = V_j(x(t))$. Then the *variational system* along the input-state-output trajectory $t \in [0, T] \mapsto (x(t), u(t), y(t))$ is given by the following time-varying system,

$$\dot{v}(t) = \frac{\partial g_0}{\partial x}(x(t))v(t) + \sum_{j=1}^{m} u_j(t) \frac{\partial g_j}{\partial x}(x(t))v(t) + \sum_{j=1}^{m} u_j^p g_j(x(t)) \,,$$

$$y_j^p(t) = \frac{\partial V_j}{\partial x}(x(t))v(t) \,, \quad j = 1, \ldots, m \,, \tag{2.3}$$

where $v(0) = v_0 \in \mathbb{R}^n$, and $u^p = (u_1^p, \ldots, u_m^p)$, $y^p = (y_1^p, \ldots, y_m^p)$ denote the inputs and the outputs of the variational system. The reason behind the terminology 'variational' comes from the following fact: let $(x(t, \epsilon), u(t, \epsilon), y(t, \epsilon))$, $t \in [0, T]$ be a family of input-state-output trajectories of (2.1) parameterized by $\epsilon \in (-\delta, \delta)$, with $x(t, 0) = x(t)$, $u(t, 0) = u(t)$ and $y(t, 0) = y(t)$, $t \in [0, T]$. Then, the infinitesimal variations

$$v(t) = \frac{\partial x}{\partial \epsilon}(t, 0) \,, \quad u^p(t) = \frac{\partial u}{\partial \epsilon}(t, 0) \,, \quad y^p(t) = \frac{\partial y}{\partial \epsilon}(t, 0) \,,$$

satisfy equation (2.3). Additionally, if the initial state is the same for the whole family of trajectories, $x(0, \epsilon) = x_0$, then the variational state $v(0)$ at time 0 is necessarily 0.

The *prolongation or prolonged system* of (2.1) corresponds to considering together the original system (2.1) and the variational system,

$$\dot{x} = g_0(x) + \sum_{j=1}^{m} u_j g_j(x),$$

$$\dot{v}(t) = \frac{\partial g_0}{\partial x}(x(t))v(t) + \sum_{j=1}^{m} u_j(t)\frac{\partial g_j}{\partial x}(x(t))v(t) + \sum_{j=1}^{m} u_j^p g_j(x(t)),$$

$$y_j = V_j(x), \quad y_j^p(t) = \frac{\partial V_j}{\partial x}(x(t))\, v(t), \quad j = 1,\dots,m, \tag{2.4}$$

with inputs u_j, u_j^p, outputs y_j, y_j^p and state (x, v). To state a coordinate-free definition of the prolonged system (2.4) on the whole tangent space TM, we need to introduce the notions of vertical and complete lifts of functions and vector fields. We do this following [26]. Given a function V on M, the *complete lift* of V to TM, $V^{\mathrm{C}} : TM \to \mathbb{R}$ is defined by $V^{\mathrm{C}}(v) = \langle dV, v \rangle$. In the induced local coordinates on TM, $(x^1, \dots, x^n, v^1, \dots, v^n)$, this reads

$$V^{\mathrm{C}}(x, v) = \sum_{a=1}^{n} \frac{\partial V}{\partial x^a}(x)\, v_a.$$

The *vertical lift* of V to TM, $V^{\mathrm{V}} : TM \to \mathbb{R}$, is defined by $V^{\mathrm{V}} = V \circ \tau_M$, where τ_M denotes the tangent bundle projection. Given a vector field X on M, the *complete lift* of X to TM, $X^{\mathrm{C}} \in \mathfrak{X}(TM)$ is defined as the unique vector field verifying $X^{\mathrm{C}}(f^{\mathrm{C}}) = (Xf)^{\mathrm{C}}$, for any $f \in C^\omega(M)$. Alternatively, if $\Phi_t : M \to M$, $t \in [0, \epsilon)$ denotes the flow of X, then we can define X^{C} as the vector field whose flow is given by $(\Phi_t)_* : TM \to TM$. In local coordinates,

$$X^{\mathrm{C}}(x, v) = \sum_{a=1}^{n} X_a(x)\frac{\partial}{\partial x^a} + \sum_{a,b=1}^{n} \frac{\partial X_a}{\partial x^b}(x)v^b \frac{\partial}{\partial v^a}. \tag{2.5}$$

The *vertical lift* of X to TM, $X^{\mathrm{V}} \in \mathfrak{X}(TM)$ is the unique vector field such that $X^{\mathrm{V}}(f^{\mathrm{C}}) = (Xf)^{\mathrm{V}}$, for any $f \in C^\omega(M)$. In local coordinates,

$$X^{\mathrm{V}}(x, v) = \sum_{a=1}^{n} X_a(x)\frac{\partial}{\partial v^a}. \tag{2.6}$$

The following definition provides an intrinsic way of pasting together the system (2.1) with the variational systems associated with its input-state-output trajectories.

DEFINITION 2.1. *The prolonged system Σ^p of a nonlinear system Σ of the form (2.1) is defined by*

$$\Sigma^p : \begin{cases} \dot{x}_p = g_0^{\mathrm{C}}(x_p) + \sum_{j=1}^{m} u_j(t)g_j^{\mathrm{C}}(x_p) + \sum_{j=1}^{m} u_j^p(t)g_j^{\mathrm{V}}(x_p) \\ y_j = V_j^{\mathrm{V}}(x_p), \quad y_j^p = V_j^{\mathrm{C}}(x_p), \quad j = 1,\dots,m, \end{cases} \tag{2.7}$$

where $x_p = (x, v) \in TM$, and $x_p(0) = (x_0, v_0)$.

One can easily check that in the induced tangent bundle coordinates, the local expression of the system (2.7) is precisely (2.4).

REMARK 2.2. In the same way as we have presented above, one can also introduce the notions of adjoint variational system and Hamiltonian extension of the nonlinear system (2.1). These notions play a key role in the characterization of when a general system admits a Hamiltonian description, see [12].   •

2.2. The gradient extension of a nonlinear system. When dealing with the Hamiltonian extension of a nonlinear system, one relies on the fact that the cotangent bundle is endowed with a canonical symplectic structure. However, this is not the case when treating gradient systems, since a canonical pseudo-Riemannian structure on the cotangent bundle does not exist. In order to define the gradient extension of a nonlinear system of the form (2.1), we will first select a torsion-free affine connection ∇ on M, and then consider its Riemannian extension to T^*M (cf. [18]).

Let us briefly present some basic notions on affine connections and Riemannian geometry. An *affine connection* [15] on a manifold M is defined as an assignment

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M)$$
$$(X, Y) \longmapsto \nabla_X Y$$

which is $\mathbb{R}$-bilinear and satisfies $\nabla_{fX} Y = f \nabla_X Y$ and $\nabla_X(fY) = f \nabla_X Y + X(f)Y$, for any $X, Y \in \mathfrak{X}(M)$, $f \in C^\omega(M)$. This implies that $\nabla_X Y(x)$ only depends on $X(x)$ and the value of Y along a curve which is tangent to X at x. Let $c : t \in [t_0, t_1] \mapsto c(t) = (x^1(t), \ldots, x^n(t)) \in M$ be a curve on M and W a vector field along c, i.e. a map $W : [t_0, t_1] \rightarrow TM$ such that $\tau_M(W(t)) = c(t)$ for all $t \in [a, b]$. Let V be a vector field that satisfies $V(c(t)) = W(t)$. The *covariant derivative of W along c* is defined by

$$\frac{DW(t)}{dt} = \nabla_{\dot{c}(t)} W(t) = \nabla_{\dot{c}(t)} V(x)\big|_{x=c(t)}.$$

This definition makes sense because of the defining properties of the affine connection. Now, we may take $W(t) = \dot{c}(t)$ and set up $\nabla_{\dot{c}(t)} \dot{c}(t) = 0$. This equation is called the *geodesic equation*, and its solutions are termed the *geodesics* of ∇. In local coordinates, this condition can be expressed as $\ddot{x}^a + \Gamma^a_{bc}(x)\dot{x}^b \dot{x}^c = 0$, $1 \le a \le n$, where the $\Gamma^a_{bc}(x)$ are the *Christoffel symbols* of the affine connection, defined by

$$\nabla_{\frac{\partial}{\partial x^b}} \frac{\partial}{\partial x^c} = \Gamma^a_{bc}(x) \frac{\partial}{\partial x^a}.$$

The vector field S on TM describing the geodesic equation is called the *geodesic spray* associated with the affine connection ∇. In local coordinates,

$$S = v^a \frac{\partial}{\partial x^a} - \Gamma^a_{bc}(x) v^b v^c \frac{\partial}{\partial v^a}.$$

Therefore, the integral curves of the geodesic spray S are the solutions of the geodesic equation. The torsion tensor of an affine connection is defined by

$$T : \mathfrak{X}(M) \times \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M)$$
$$(X, Y) \longmapsto \nabla_X Y - \nabla_Y X - [X, Y],$$

Locally, we have

$$T\Big(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\Big) = (\Gamma^c_{ab} - \Gamma^c_{ba}) \frac{\partial}{\partial x^c},$$

An affine connection is *torsion-free* if T is identically zero. Given an affine connection, the *symmetric product* [17] of two vector fields $X, Y \in \mathfrak{X}(M)$ is defined by the operation

$$\langle X : Y \rangle = \nabla_X Y + \nabla_Y X \,.$$

The geometric meaning of the symmetric product is the following [16]: a distribution $\mathcal{D}$ on M is geodesically invariant (meaning that each geodesic of ∇ whose initial velocity is in $\mathcal{D}$ has all its velocities in $\mathcal{D}$) if and only if $\langle X : Y \rangle \in \mathcal{D}$, for all $X, Y \in \mathcal{D}$. The symmetric product plays a crucial role within the so-called affine connection formalism for mechanical control systems in the study of a variety of aspects such as controllability, series expansions, motion planning and optimal control [6]. Note that if the affine connection ∇ is torsion-free, then $\nabla_X Y = \frac{1}{2}(\langle X : Y \rangle + [X, Y])$ for all $X, Y \in \mathfrak{X}(M)$, i.e., there is a one-to-one correspondence between the covariant derivative and the symmetric product.

Associated with th'e metric $\mathcal{G}$ there is a natural affine connection, called the *Levi-Civita* connection. The Levi-Civita connection $\nabla^\mathcal{G}$ is determined by the formula

$$2\mathcal{G}(\nabla^\mathcal{G}_X Y, Z) = X(\mathcal{G}(Y, Z)) + Y(\mathcal{G}(Z, X)) - Z(\mathcal{G}(X, Y))$$
$$+ \mathcal{G}(Y, [Z, X]) - \mathcal{G}(X, [Y, Z]) + \mathcal{G}(Z, [X, Y]) \,, \quad X, Y, Z \in \mathfrak{X}(M) \,.$$

One can compute the Christoffel symbols of $\nabla^\mathcal{G}$ to be

$$\Gamma^a_{bc} = \frac{1}{2}\mathcal{G}^{ad}\left(\frac{\partial \mathcal{G}_{db}}{\partial x^c} + \frac{\partial \mathcal{G}_{dc}}{\partial x^b} - \frac{\partial \mathcal{G}_{bc}}{\partial x^d}\right) \,. \tag{2.8}$$

The Levi-Civita connection is torsion-free, that is $T(X, Y) = 0$, for any $X, Y \in \mathfrak{X}(M)$.

Therefore, a pseudo-Riemannian metric on M defines a unique affine torsion-free connection on M. The converse is, however, not true. Also, note that given an affine torsion-free connection which is the Levi-Civita connection corresponding to some pseudo-Riemannian metric, then there exist many more metrics that give rise to the same affine connection. For instance, any constant metric on the Euclidean space gives rise to the affine connection with Christoffel symbols $\Gamma^a_{bc} = 0$, $1 \le a, b, c \le n$.

Given a pseudo-Riemannian metric $\mathcal{G}$ on M, we can define the so-called Beltrami bracket [9, 21] of functions on M,

$$\{f : g\}_\mathcal{G} = \mathcal{G}(\mathrm{grad}_\mathcal{G} f, \mathrm{grad}_\mathcal{G} g) \,, \quad f, g \in C^\omega(M) \,.$$

In local coordinates, one has the expression,

$$\{f : g\}_\mathcal{G} = \frac{\partial f}{\partial x^a}\mathcal{G}^{ab}\frac{\partial g}{\partial x^b} \,.$$

It is interesting to note that the mapping

$$\mathrm{grad}_\mathcal{G} : (C^\omega(M), \{\cdot : \cdot\}_\mathcal{G}) \to (\mathfrak{X}(M), \langle \cdot : \cdot \rangle_{\nabla^\mathcal{G}})$$

is a homomorphism of symmetric algebras, i.e., $\mathrm{grad}_\mathcal{G}\{f : g\}_\mathcal{G} = \langle \mathrm{grad}_\mathcal{G} f : \mathrm{grad}_\mathcal{G} g \rangle_{\nabla^\mathcal{G}}$, for all $f, g \in C^\omega(M)$.

REMARK 2.3. The latter observation is the gradient analog of the following fact in the Hamiltonian setting: consider the mapping $(C^\omega(M), \{\cdot, \cdot\}) \to (\mathfrak{X}(M), [\cdot, \cdot])$

(where $\{\cdot, \cdot\}$ denotes the Poisson bracket and $[\cdot, \cdot]$ denotes the Lie bracket), associating to each function f its Hamiltonian vector field X_f. Then this mapping is a homomorphism of Lie algebras, i.e. $X_{\{f,g\}} = [X_f, X_g]$. •

Let us now turn our discussion to the cotangent bundle of M. First, we introduce the construction that associates to each vector field X on M, a function V^X on T^*M, defined by $V^X(p) = \langle p, X \rangle$. In the induced local coordinates $(x^1, \ldots, x^n, p_1, \ldots, p_n)$ on T^*M, this reads $V^X(x, p) = \sum_{a=1}^n p_a X_a(x)$. The *complete lift* of X to T^*M, $X^{\mathrm{c}} \in \mathfrak{X}(T^*M)$ is defined as the Hamiltonian vector field (with respect to the canonical symplectic form on T^*M) associated with the function V^X. In local coordinates,

$$X^{\mathrm{c}}(x, p) = \sum_{a=1}^n X_a \frac{\partial}{\partial x^a} - \sum_{a,b=1}^n \frac{\partial X_b}{\partial x^a}(x) p_b \frac{\partial}{\partial p_a}$$

The notion of *vertical lift* of a function V on M to a function V^{v} on T^*M is given by $V^{\mathrm{v}} = V \circ \pi_M$, where π_M is the cotangent bundle projection. An object which will play a key role in the subsequent discussion is the *Riemannian extension* [18, 26] of a torsion-free affine connection. Let ∇ be a torsion-free affine connection on M. Then, ∇ defines a pseudo-Riemannian metric on T^*M, denoted $\mathcal{G}^\nabla$, as the unique (0,2)-tensor on T^*M which satisfies

$$\mathcal{G}^\nabla(X^{\mathrm{c}}, Y^{\mathrm{c}}) = -V^{\langle X:Y \rangle}.$$

The fact that this single equality completely determines the Riemannian extension $\mathcal{G}^\nabla$ is a consequence of the result in Proposition 4.2 in Chapter VII of [26], which asserts that any $(0, s)$-tensor field on T^*M is univocally defined by its action on the complete lifts of vector fields of M. The matrix representations of the musical isomorphisms defined by $\mathcal{G}^\nabla$ in local coordinates are given by

$$\flat_{\mathcal{G}^\nabla} \equiv \begin{pmatrix} -2p_c \Gamma^c_{ab} & I_n \\ I_n & 0 \end{pmatrix}, \quad \sharp_{\mathcal{G}^\nabla} \equiv \begin{pmatrix} 0 & I_n \\ I_n & 2p_c \Gamma^c_{ab} \end{pmatrix}. \tag{2.9}$$

As for the gradient vector fields associated with the functions V^X, $V^{\mathrm{v}} \in C^\omega(T^*M)$, $X \in \mathfrak{X}(M)$, $V \in C^\omega(M)$, one has the local expressions

$$\mathrm{grad}_{\mathcal{G}^\nabla} V^X = X^a \frac{\partial}{\partial x^a} + p_a \left(\frac{\partial X^a}{\partial x^b} + 2\Gamma^a_{bc} X^c \right) \frac{\partial}{\partial p_b}, \quad \mathrm{grad}_{\mathcal{G}^\nabla} V^{\mathrm{v}} = \frac{\partial V}{\partial x^a} \frac{\partial}{\partial p_a}. \tag{2.10}$$

Given a metric $\mathcal{G}$ on M, one can also verify that the pseudo-Riemannian metric on T^*M defined by $\mathcal{G}^{\nabla^\mathcal{G}}$ corresponds to the pullback by $\sharp_\mathcal{G}$ of the complete lift $\mathcal{G}^{\mathrm{c}}$ to TM of $\mathcal{G}$, i.e., $\mathcal{G}^{\nabla^\mathcal{G}} = \sharp_\mathcal{G}^*(\mathcal{G}^{\mathrm{c}})$ (see [26]).

DEFINITION 2.4. *The gradient extension* Σ^e *of a nonlinear system* Σ *of the form* (2.1) *with respect to a torsion-free affine connection* ∇ *on M is given by*

$$\Sigma^e : \begin{cases} \dot{x}_e = \mathrm{grad}_{\mathcal{G}^\nabla} V^{g_0}(x_e) + \sum_{j=1}^m u_j(t) \, \mathrm{grad}_{\mathcal{G}^\nabla} V^{g_j}(x_e) + \sum_{j=1}^m u_j^e(t) \, \mathrm{grad}_{\mathcal{G}^\nabla} V_j^{\mathrm{v}}(x_e), \\ y_j = V_j^{\mathrm{v}}(x_e), \quad y_j^e = V^{g_j}(x_e), \quad j = 1, \ldots, m, \end{cases} \tag{2.11}$$

where $x_e = (x, p) \in T^*M$, $x_e(0) = (x_0, p_0)$, $u = (u_1, \ldots, u_m) \in U \subset \mathbb{R}^m$, *and* $u^e = (u_1^e, \ldots, u_m^e) \in \mathbb{R}^m$.

REMARK 2.5. Note that the gradient extension Σ^e is itself a gradient control system. •

3. Observability of the prolongation and the gradient extension. In this section, we investigate the observability properties of the prolonged system and the gradient extension of a nonlinear system. Roughly speaking, the observability properties of a given system determine to what extent one can *observe* the actual state of the system from its input-output behavior, i.e., to what extent the knowledge of the input-output response allows us to infer things about the evolution of the state. This study will later be key in establishing the characterization of when a nonlinear control system can be written as a gradient control system.

We start by briefly reviewing some notions such as distinguishable points and local observability. Let $\mathcal{Y}$ denote the space of absolutely continuous functions defined on $\mathbb{R}_+ = [0, +\infty)$ with values in $\mathbb{R}^m$. For a nonlinear system of the form (2.1), the *input-output map* $\mathcal{R}_\Sigma : M \times \mathcal{U} \to \mathcal{Y}$, $\mathcal{R}_\Sigma(x_0, u(\cdot)) = y(\cdot)$ is defined by assigning to each initial condition $x_0 \in M$ and any admissible control $u \in \mathcal{U}$ the output of the system,

$$y(\cdot) = (V_1(x(\cdot, x_0, u(\cdot))), \ldots, V_m(x(t, x_0, u(\cdot)))),$$

where $x(\cdot, x_0, u(\cdot))$ denotes the solution of $\dot{x}(t) = g_0(x(t)) + \sum_{j=1}^m u_j(t) g_j(x(t))$ starting at x_0. Now, two points $x_1, x_2 \in M$ are said to be *indistinguishable*, $x_1 \sim x_2$, if $\mathcal{R}_\Sigma(x_1, u(\cdot)) = \mathcal{R}_\Sigma(x_2, u(\cdot))$ for any $u(\cdot) \in \mathcal{U}$.

DEFINITION 3.1. *A system Σ is observable if for any $x_1, x_2 \in M$, one has that $x_1 \sim x_2 \Rightarrow x_1 = x_2$. Alternatively, for any $x_1 \neq x_2$, there exists an admissible control such that the output functions resulting from the initial conditions $x(0) = x_1$, resp. $x(0) = x_2$, are different. The system is* locally observable *at x_0 if there exists a neighborhood $\mathcal{N}$ of x_0 such that this holds for points in $\mathcal{N}$.*

Denote by $\mathcal{H}$ the $\mathbb{R}$-linear space in $C^\omega(M)$ spanned by the functions of the form $\mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_s} V_j$, with $\{X_r\}_{r=1}^s \subset \{g_i \mid i = 0, 1, \ldots, m\}$, and $j \in \{1, \ldots, m\}$. Alternatively, we may take X_r to be arbitrary elements of the accessibility algebra corresponding to the vector fields $g_0, g_1, \ldots, g_m$. $\mathcal{H}$ is called the *observation space* of Σ. It follows from the analyticity assumption that the system is observable if and only if $\mathcal{H}$ distinguishes points in M, i.e. for every $x_1, x_2 \in M$ with $x_1 \neq x_2$, there exists $V \in \mathcal{H}$ such that $V(x_1) \neq V(x_2)$, cf. [13].

PROPOSITION 3.2 ([12]). *Consider a nonlinear system Σ of the form (2.1), with observation space $\mathcal{H}$. Then, the observation space $\mathcal{H}^p$ of the prolongation Σ^p is given by $\mathcal{H}^p = \mathcal{H}^C + \mathcal{H}^V$, where $\mathcal{H}^C = \{V^C \mid V \in \mathcal{H}\}$ and $\mathcal{H}^V = \{V^V \mid V \in \mathcal{H}\}$.*

The following corollary is a modified statement of Corollary 3.3 in [12].

COROLLARY 3.3. *Assume the codistribution $d\mathcal{H}$ is of constant rank. Then the system Σ is (locally) observable if and only if its prolongation is (locally) observable.*

Proof. Following [13], Σ is locally observable if and only $\mathrm{rk}(d\mathcal{H}) = \dim M$. In addition, the codistribution $d\mathcal{H}$ on M has constant rank if and only if the codistribution $d\mathcal{H}^p$ on TM has constant rank. Therefore, $\mathrm{rk}(d\mathcal{H}) = \dim M$ if and only if $\mathrm{rk}(d\mathcal{H}^p) = \dim TM$ if and only if Σ^p is locally observable. The statement regarding observability is proved as in Corollary 3.3 in [12]. $\square$

Let us turn our attention to the observability properties of the gradient extension of a nonlinear system of the form (2.1). The following lemma will be most helpful.

LEMMA 3.4. *Let ∇ be a torsion-free affine connection on a manifold M, and let $\mathcal{G}^\nabla$ denote its Riemannian extension to T^*M. Then, for any vector fields $X, Y \in \mathfrak{X}(M)$, and any functions $f, g \in C^\omega(M)$, the following identities hold*

(i) $(\mathrm{grad}_{\mathcal{G}^\nabla} V^X)(V^Y) = \{V^X : V^Y\}_{\mathcal{G}^\nabla} = V^{\langle X:Y \rangle} = -\mathcal{G}^\nabla(X^C, Y^C).$

(ii) $(\mathrm{grad}_{\mathcal{G}^\nabla} V^X)(f^\mathrm{v}) = (\mathrm{grad}_{\mathcal{G}^\nabla} f^\mathrm{v})(V^X) = \{V^X : f^\mathrm{v}\}_{\mathcal{G}^\nabla} = X(f)^\mathrm{v}$.

(iii) $(\mathrm{grad}_{\mathcal{G}^\nabla} f^\mathrm{v})(g^\mathrm{v}) = \{f^\mathrm{v} : g^\mathrm{v}\}_{\mathcal{G}^\nabla} = 0$.

Proof. The first equality in (i) is the definition of the Beltrami bracket associated with $\mathcal{G}^\nabla$. For the second one, we resort to the local expressions in (2.9) to compute

$$\{V^X : V^Y\}_{\mathcal{G}^\nabla} = \left(p_a \frac{\partial X^a}{\partial x^b}, X^b\right) \begin{pmatrix} 0 & I \\ I & 2p_e \Gamma^e_{cd} \end{pmatrix} \left(p_a \frac{\partial Y^a}{\partial x^b}, Y^b\right)^T$$

$$= p_a \left(\frac{\partial X^a}{\partial x^b} Y^b + \frac{\partial Y^a}{\partial x^b} X^b + 2\Gamma^a_{bc} X^b Y^c\right) = V^{\langle X : Y\rangle}.$$

The third equality corresponds to the definition of $\mathcal{G}^\nabla$. The first and second equalities in (ii) follow again by definition. As for the third one, note that

$$\mathrm{grad}_{\mathcal{G}^\nabla} f^\mathrm{v}(V^X) = \frac{\partial f}{\partial x^a} \frac{\partial}{\partial p_a}(p_b X^b) = \frac{\partial f}{\partial x^a} X^a = X(f)^\mathrm{v}$$

Finally, the equalities in (iii) are straightforward. $\quad\square$

Denote by S_0 the $\mathbb{R}$-linear space in $\mathfrak{X}(M)$ spanned by the vector fields of the form $\langle X_1 : \langle X_2 : \langle\ldots : \langle X_s : g_j\rangle\rangle\ldots\rangle\rangle$, with $\{X_r\}^s_{r=1} \subset \{g_i \mid i = 0, 1, \ldots, m\}$, and $j \in \{1, \ldots, m\}$. Alternatively, one can define S_0 as the smallest subspace of $\mathfrak{X}(M)$ such that (i) $g_1, \ldots, g_m \in S_0$; and (ii) if $X \in S_0$, then $\langle g_i : X\rangle \in S_0$ for all $i = 0, 1, \ldots, m$. We denote by $\mathcal{S}_0$ the distribution on M generated by the space S_0,

$$\mathcal{S}_0(x) = \mathrm{span}\{X(x) \mid X \in S_0\}, \quad x \in M. \tag{3.1}$$

PROPOSITION 3.5. *Consider a nonlinear system Σ of the form* (2.1), *with observation space $\mathcal{H}$. Let ∇ be a torsion-free affine connection on M. Then, the observation space $\mathcal{H}^e$ of the gradient extension Σ^e is given by $\mathcal{H}^e = V^{S_0} + (\mathcal{H} + \mathfrak{h})^\mathrm{v}$, where $V^{S_0} = \{V^X \mid X \in S_0\}$ and $\mathfrak{h}$ is spanned by $\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}\mathcal{L}_X V_j$, with X_r, $r = 1, \ldots, s$, equal to g_i, $i = 0, 1, \ldots, m$, $X \in S_0$ and $j = 1, \ldots, m$.*

Proof. The observation space of the gradient extension of Σ is spanned by

$$\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s} V^\mathrm{v}_j, \quad \mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s} V^{g_j},$$

where X_r, $r = 1, \ldots, s$ is equal to $\mathrm{grad}_{\mathcal{G}^\nabla} V^{g_i}$, $\mathrm{grad}_{\mathcal{G}^\nabla} V^\mathrm{v}_j$, $i = 0, 1, \ldots, m$, $j = 1, \ldots, m$. Now, using Lemma 3.4, we have that

$$\mathcal{L}_{\mathrm{grad}_{\mathcal{G}^\nabla} V^{g_i}} V^\mathrm{v}_j = (\mathcal{L}_{g_i} V_j)^\mathrm{v}, \qquad \mathcal{L}_{\mathrm{grad}_{\mathcal{G}^\nabla} V^{g_i}} V^{g_j} = V^{\langle g_i : g_j\rangle},$$

$$\mathcal{L}_{\mathrm{grad}_{\mathcal{G}^\nabla} V^\mathrm{v}_j} V^\mathrm{v}_k = 0, \qquad \mathcal{L}_{\mathrm{grad}_{\mathcal{G}^\nabla} V^\mathrm{v}_j} V^{g_k} = (\mathcal{L}_{g_k} V_j)^\mathrm{v},$$

with $i = 0, 1, \ldots, m$ and $j, k = 1, \ldots, m$. Considering the next step of Lie derivatives yields

$$\mathcal{L}_{\mathrm{grad}_{\mathcal{G}^\nabla} V^{g_h}} V^{\langle g_i : g_j\rangle} = V^{\langle g_h : \langle g_i : g_j\rangle\rangle}, \qquad \mathcal{L}_{\mathrm{grad}_{\mathcal{G}^\nabla} V^{g_h}} (\mathcal{L}_{g_i} V_j)^\mathrm{v} = (\mathcal{L}_{g_h}\mathcal{L}_{g_i} V_j)^\mathrm{v},$$

$$\mathcal{L}_{\mathrm{grad}_{\mathcal{G}^\nabla} V^\mathrm{v}_k} V^{\langle g_i : g_j\rangle} = (\mathcal{L}_{\langle g_i : g_j\rangle} V_k)^\mathrm{v}, \qquad \mathcal{L}_{\mathrm{grad}_{\mathcal{G}^\nabla} V^\mathrm{v}_k} (\mathcal{L}_{g_i} V_j)^\mathrm{v} = 0,$$

with $h = 0, 1, \ldots, m$. Further iterating this process, we get to the desired result. $\quad\square$

COROLLARY 3.6. *Consider a nonlinear system Σ of the form* (2.1), *with observation space $\mathcal{H}$. Assume the codistribution $d\mathcal{H}$ is of constant rank. Let ∇ be a torsion-free affine connection on M and further assume that the distribution $\mathcal{S}_0$ is full-rank. Then, Σ is (locally) observable implies that Σ^e is (locally) observable.*

Proof. Since the codistribution $d\mathcal{H}$ has constant rank, Σ is locally observable if and only if $\dim d\mathcal{H}(x) = \dim M$. Since $\mathcal{S}_0$ is full-rank, it is clear that Σ locally observable implies that $\mathcal{H}^e$ has constant maximal rank, and therefore Σ^e is locally observable. With respect to observability, let (x_1, p_1), $(x_2, p_2) \in T^*M$ and assume that $V^e(x_1, p_1) = V^e(x_2, p_2)$ for all $V^e \in \mathcal{H}^e$. Since $\mathcal{H}^\vee \subset \mathcal{H}^e$, this yields $V(x_1) = V(x_2)$ for any $V \in \mathcal{H}$. So, under observability of Σ, we conclude that $x_1 = x_2 = x$. Then, we have that $V^X(x, p_1) = V^X(x, p_2)$, for all $X \in S_0$, which finally implies that $p_1 = p_2$. □

4. Externally equivalent systems. In this section we introduce the notion of (weakly) externally equivalent systems, which will be instrumental in the statement of the main result in Section 5. Consider two nonlinear systems $\alpha = 1, 2$, of the form,

$$
\Sigma^\alpha : \begin{cases} \dot{x}^\alpha = g_0^\alpha(x^\alpha) + \sum_{j=1}^m u_j g_j^\alpha(x^\alpha), & x^\alpha \in M^\alpha, \\ y_j = V_j^\alpha(x^\alpha), & j = 1, \ldots, m, \ u = (u_1, \ldots, u_m) \in U \subset \mathbb{R}^m. \end{cases}
$$

Denote by $\mathcal{H}^\alpha$, $\alpha = 1, 2$, the associated observation spaces. Take a function $H^1 \in \mathcal{H}^1$, $H^1 = \mathcal{L}_{X_1} \ldots \mathcal{L}_{X_s} V_j^1$, with $X_r = g_{i_r}^1$, $i_r \in \{0, 1, \ldots, m\}$, $r = 1, \ldots, s$ and $j \in \{1, \ldots, m\}$. Consider the function in $\mathcal{H}^2$ defined by $H^2 = \mathcal{L}_{Y_1} \ldots \mathcal{L}_{Y_s} V_j^2$, with $Y_r = g_{i_r}^2$, $r = 1, \ldots, s$. Then we say that H^1 and H^2 formally *correspond* to each other. This notion is useful to define the concept of *weakly externally equivalent* systems.

DEFINITION 4.1. *The systems Σ^1 and Σ^2 are weakly externally equivalent if and only if for all $x^1 \in M^1$, there exists $x^2 \in M^2$ such that $H^1(x^1) = H^2(x^2)$ for all corresponding $H^1 \in \mathcal{H}^1$, $H^2 \in \mathcal{H}^2$, and reciprocally, for all $x^2 \in M^2$, there exists $x^1 \in M^1$ such that $H^1(x^1) = H^2(x^2)$ for all corresponding $H^1 \in \mathcal{H}^1$, $H^2 \in \mathcal{H}^2$.*

DEFINITION 4.2. *The systems Σ^1 and Σ^2 are externally equivalent if and only if for all $x^1 \in M^1$, there exists $x^2 \in M^2$ such that the input-output maps corresponding to x^1 and x^2 coincide, i.e. $\mathcal{R}_{\Sigma^1}(x^1, u(\cdot)) = \mathcal{R}_{\Sigma^2}(x^2, u(\cdot))$, for all $u(\cdot) \in \mathcal{U}$, and reciprocally, for all $x^2 \in M^2$, there exists $x^1 \in M^1$ such that $\mathcal{R}_{\Sigma^1}(x^1, u(\cdot)) = \mathcal{R}_{\Sigma^2}(x^2, u(\cdot))$, for all $u(\cdot) \in \mathcal{U}$.*

Equivalently, Σ^1 and Σ^2 are externally equivalent if and only if their behaviors are equal. Clearly, if two systems are externally equivalent, then they are weakly externally equivalent.

PROPOSITION 4.3. *Assume that Σ^1 and Σ^2 are weakly externally equivalent, observable and that the codistributions $d\mathcal{H}^\alpha$, $\alpha = 1, 2$, have constant rank. Then there exists a unique diffeomorphism $\varphi : M^1 \to M^2$ with $\varphi^*(\mathcal{H}^2) = \mathcal{H}^1$.*

Proof. Let $x^1 \in M^1$. By definition, there exists $x^2 \in M^2$ such that $H^1(x^1) = H^2(x^2)$ for all corresponding $H^1 \in \mathcal{H}^1$, $H^2 \in \mathcal{H}^2$. Since $\mathcal{H}^2$ distinguishes points in M^2, it follows that x^2 is unique. Define $\varphi : M^1 \to M^2$, $\varphi(x^1) = x^2$. Using $\dim d\mathcal{H}^2 = \dim M^2$ and the inverse function theorem, it follows that φ is smooth. Indeed, for each $x^2 \in M^2$, there exists a neighborhood V of M^2 at x_2 and $\dim M^2$ independent functions $H_1^2, \ldots, H_{\dim M^2}^2$ on V, such that φ is given by

$$
x^2 = (H_1^2, \ldots, H_{\dim M^2}^2)^{-1}(H_1^1, \ldots, H_{\dim M^2}^1)(x^1).
$$

Analogously, we can construct the inverse mapping $\varphi^{-1} : M^2 \to M^1$, making use of the fact that Σ_1 is observable, which concludes the proof. □

COROLLARY 4.4. *Let the systems Σ^1 and Σ^2 be observable and the codistributions $d\mathcal{H}^\alpha$, $\alpha = 1, 2$, have constant rank. Then Σ^1 and Σ^2 are weakly externally equivalent if and only if they are externally equivalent.*

Proof. We already know that if the systems are externally equivalent, then they are weakly externally equivalent. Conversely, assume that Σ^1 and Σ^2 are weakly externally equivalent. From Proposition 4.3, we have that there exists a diffeomorphism $\varphi : M^1 \to M^2$ with $\varphi^*(\mathcal{H}^2) = \mathcal{H}^1$. Using this latter fact, and since the vector fields g_0^i, g_j^i are determined by their action as derivations on $\mathcal{H}^\alpha$, $\alpha = 1, 2$, we conclude that $\varphi_* g_0^1 = g_0^2$, $\varphi_* g_j^1 = g_j^2$, $j = 1, \ldots, m$. $\square$

REMARK 4.5. The map φ in the previous proof is called a *state-space diffeomorphism*.   ●

5. Gradient realization of a nonlinear control system. This section contains the main result of the paper. Under certain technical conditions, Theorem 5.4 below characterizes when a nonlinear control systems admits a gradient realization. Before stating this result, we need to introduce the novel notion of *compatibility* between a nonlinear system and an affine connection.

DEFINITION 5.1 (Compatibility). *Let ∇ be an affine connection on M. A nonlinear control system Σ of the form (2.1) is compatible with ∇ if and only if the following two conditions hold:*

    (a) *For all vector fields $X_1, \ldots, X_{s_1}, Y_1, \ldots, Y_{s_2} \in \{g_0, g_1, \ldots, g_m\}$, and all indexes $j, k = 1, \ldots, m$,*

$$\mathcal{L}_{\langle X_1 : \langle X_2 : \langle \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{Y_1} \mathcal{L}_{Y_2} \ldots \mathcal{L}_{Y_{s_2}} V_k \right]$$
$$= \mathcal{L}_{\langle Y_1 : \langle Y_2 : \langle \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_{s_1}} V_j \right] .$$

    (b) *For all vector fields $X_1, \ldots, X_{s_1}, Y_1, \ldots, Y_{s_2}, Z_1, \ldots, Z_{s_3} \in \{g_0, g_1, \ldots, g_m\}$, and all indexes $j, k, l = 1, \ldots, m$,*

$$\mathcal{L}_{\langle \langle X_1 : \langle X_2 : \langle \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle : \langle Y_1 : \langle Y_2 : \langle \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle \rangle} \left[\mathcal{L}_{Z_1} \mathcal{L}_{Z_2} \ldots \mathcal{L}_{Z_{s_3}} V_l \right]$$
$$= \mathcal{L}_{\langle Z_1 : \langle Z_2 : \langle \ldots : \langle Z_{s_3} : g_l \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{\langle X_1 : \langle X_2 : \langle \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{Y_1} \mathcal{L}_{Y_2} \ldots \mathcal{L}_{Y_{s_2}} V_k \right] \right] .$$

REMARK 5.2. In case the distribution $\mathcal{S}_0$ (cf. equation (3.1)) is full-rank, note that property (b) in the above definition implies property (a) up to a constant on each connected component of M. To see this, one can use the symmetry of the symmetric product to deduce from (b) that

$$\mathcal{L}_{\langle Z_1 : \langle Z_2 : \langle \ldots : \langle Z_{s_3} : g_l \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{\langle X_1 : \langle X_2 : \langle \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{Y_1} \mathcal{L}_{Y_2} \ldots \mathcal{L}_{Y_{s_2}} V_k \right] \right]$$
$$= \mathcal{L}_{\langle Z_1 : \langle Z_2 : \langle \ldots : \langle Z_{s_3} : g_l \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{\langle Y_1 : \langle Y_2 : \langle \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_{s_1}} V_j \right] \right] .$$

Now, one concludes the result from the full-rankness of $\mathcal{S}_0$. Another interesting observation in this case is that the checkability of the compatibility condition can be performed taking a basis of vector fields in $\mathcal{S}_0$ as we discuss later in Lemma 8.1.   ●

REMARK 5.3. Note that a locally gradient control system of the form (2.2) is compatible with the Levi-Civita connection associated with the pseudo-Riemannian metric $\mathcal{G}$. Indeed, let $\langle \cdot : \cdot \rangle$, $\{ \cdot : \cdot \}$ denote, respectively, the symmetric product and the Beltrami bracket corresponding to $\nabla^{\mathcal{G}}$ and $\mathcal{G}$. Take $X_{r_1} = \text{grad}\, V_{\alpha_{r_1}}$, $Y_{r_2} = \text{grad}\, V_{\beta_{r_2}}$,

$Z_{r_3} = \operatorname{grad} V_{\gamma_{r_3}}$, $r_i \in \{1, \dots, s_i\}$ (which can always be written at least locally), then

$$
\begin{aligned}
\mathcal{L}_{\langle X_1 : \langle X_2 : \langle \dots : \langle X_{s_1} : g_j \rangle \rangle \dots \rangle \rangle} & \left[\mathcal{L}_{Y_1} \mathcal{L}_{Y_2} \dots \mathcal{L}_{Y_{s_2}} V_k \right] \\
&= \mathcal{L}_{\operatorname{grad}\{V_{\alpha_1} : \{V_{\alpha_2} : \{\dots : \{V_{\alpha_{s_1}} : V_j\}\} \dots \}\}} \left[\{V_{\beta_1} : \{V_{\beta_2} : \{\dots : \{V_{\beta_{s_2}} : V_k\}\} \dots \}\} \right] \\
&= \mathcal{L}_{\operatorname{grad}\{V_{\beta_1} : \{V_{\beta_2} : \{\dots : \{V_{\beta_{s_2}} : V_k\}\} \dots \}\}} \left[\{V_{\alpha_1} : \{V_{\alpha_2} : \{\dots : \{V_{\alpha_{s_1}} : V_j\}\} \dots \}\} \right] \\
&\qquad\qquad = \mathcal{L}_{\langle Y_1 : \langle Y_2 : \langle \dots : \langle Y_{s_2} : g_k \rangle \rangle \dots \rangle \rangle} \left[\mathcal{L}_{X_1} \mathcal{L}_{X_2} \dots \mathcal{L}_{X_{s_1}} V_j \right],
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{L}_{\langle\langle X_1 : \langle X_2 : \langle \dots : \langle X_{s_1} : g_j \rangle \rangle \dots \rangle \rangle : \langle Y_1 : \langle Y_2 : \langle \dots : \langle Y_{s_1} : g_k \rangle \rangle \dots \rangle \rangle \rangle} & \left[\mathcal{L}_{Z_1} \mathcal{L}_{Z_2} \dots \mathcal{L}_{Z_{s_3}} V_l \right] \\
&= \mathcal{L}_{\operatorname{grad}\{\{V_{\alpha_1} : \{V_{\alpha_2} : \{\dots : \{V_{\alpha_{s_1}} : V_j\}\} \dots \}\} : \{V_{\beta_1} : \{V_{\beta_2} : \{\dots : \{V_{\beta_{s_2}} : V_k\}\} \dots \}\}\}} \\
&\qquad \left[\{V_{\gamma_1} : \{V_{\gamma_2} : \{\dots : \{V_{\gamma_{s_3}} : V_l\}\} \dots \}\} \right] \\
&= \mathcal{L}_{\operatorname{grad}\{V_{\gamma_1} : \{V_{\gamma_2} : \{\dots : \{V_{\gamma_{s_3}} : V_l\}\} \dots \}\}} \\
\left[\{\{V_{\alpha_1} : \{V_{\alpha_2} : \{\dots : \{V_{\alpha_{s_1}} : V_j\}\} \dots \}\} : \{V_{\beta_1} : \{V_{\beta_2} : \{\dots : \{V_{\beta_{s_2}} : V_k\}\} \dots \}\}\} \right]& \\
&= \mathcal{L}_{\langle Z_1 : \langle Z_2 : \langle \dots : \langle Z_{s_3} : g_l \rangle \rangle \dots \rangle \rangle} \left[\mathcal{L}_{\langle X_1 : \langle X_2 : \langle \dots : \langle X_{s_1} : g_j \rangle \rangle \dots \rangle \rangle} \left[\mathcal{L}_{Y_1} \mathcal{L}_{Y_2} \dots \mathcal{L}_{Y_{s_2}} V_k \right] \right],
\end{aligned}
$$

as claimed. •

Now, we come to the main result of the paper.

THEOREM 5.4. *Let Σ be a nonlinear control system of the form* (2.1). *Let ∇ be a torsion-free affine connection defined on the state manifold M. Assume Σ is observable with $\dim d\mathcal{H}$ constant, compatible with ∇ and that the distribution $\mathcal{S}_0$ is full-rank. Then, Σ is a locally gradient control system with respect to a pseudo-metric whose Levi-Civita connection is ∇ if and only if its prolonged system Σ^p and its gradient extension Σ^e are weakly externally equivalent.*

Proof. $\Rightarrow$) Consider a locally gradient control system Σ on $(M, \mathcal{G})$ (cf. (2.2)), together with its prolongation Σ^p on TM and its gradient extension Σ^e on T^*M. Recall that in the induced bundle coordinates (x^a, v^a) on TM, (x^a, p_a) on T^*M, the musical isomorphisms associated with $\mathcal{G}$ read $\flat_{\mathcal{G}}(x^a, v^a) = (x^a, \mathcal{G}_{ab} v^b)$ and $\sharp_{\mathcal{G}}(x^a, p_a) = (x^a, \mathcal{G}^{ab} p_b)$. We are going to show that $\flat_{\mathcal{G}}$ is actually an isomorphism between the prolongation and the gradient extension, i.e. we will prove that $\flat_{\mathcal{G}}(x_p(\cdot)) = x_e(\cdot)$ along the solutions of (2.7) and (2.11) respectively. This will be a consequence of the following equalities

$$
\begin{aligned}
(\flat_{\mathcal{G}})_* g_i^{\mathrm{C}} &= \operatorname{grad}_{\mathcal{G}\nabla} V^{g_i} \circ \flat_{\mathcal{G}}, & V^{g_j} \circ \flat_{\mathcal{G}} &= V_j^{\mathrm{C}}, \\
(\flat_{\mathcal{G}})_* g_j^{\mathrm{V}} &= \operatorname{grad}_{\mathcal{G}\nabla} V_j^{\mathrm{V}} \circ \flat_{\mathcal{G}}, & V_j^{\mathrm{V}} \circ \flat_{\mathcal{G}} &= V_j^{\mathrm{V}},
\end{aligned}
\tag{5.1}
$$

for all $i = 0, 1, \dots, m$, $j = 1, \dots, m$. In order to show (5.1), we will make use of the following identities,

$$
(\flat_{\mathcal{G}})_* \left(\frac{\partial}{\partial x^a} \right) = \frac{\partial}{\partial x^a} + \frac{\partial \mathcal{G}_{cb}}{\partial x^a} v^b \frac{\partial}{\partial p_c}, \quad (\flat_{\mathcal{G}})_* \left(\frac{\partial}{\partial v^a} \right) = \mathcal{G}_{ab} \frac{\partial}{\partial p_b}.
$$

Let $g \in \mathfrak{X}(M)$. In local coordinates, $g = g^a \partial/\partial x^a$. Using (2.5), we get

$$
(\flat_{\mathcal{G}})_* \left(g^{\mathrm{C}} \right) = g^a \frac{\partial}{\partial x^a} + \left\{ g^c \frac{\partial \mathcal{G}_{ab}}{\partial x^c} + \mathcal{G}_{ac} \frac{\partial g^c}{\partial x^b} \right\} v^b \frac{\partial}{\partial p_a}.
$$

On the other hand, we have that

$$\mathrm{grad}_{\mathcal{G}^\nabla}\, V^g \circ \flat_{\mathcal{G}} = g^a \frac{\partial}{\partial x^a} + \left\{ \mathcal{G}_{bc}\frac{\partial g^c}{\partial x^a} + 2\mathcal{G}_{bc}\Gamma^c_{ad}g^d \right\} v^b \frac{\partial}{\partial p_a}\,.$$

Now, suppose that g is a locally gradient vector field. In local coordinates, this means that $\mathcal{G}_{ac}g^c = \partial V/\partial x^a$, for a certain function V, which in turn implies that $\partial\{\mathcal{G}_{ac}g^c\}/\partial x^b = \partial\{\mathcal{G}_{bc}g^c\}/\partial x^a$, that is

$$\mathcal{G}_{ac}\frac{\partial g^c}{\partial x^b} = \frac{\partial \mathcal{G}_{bc}}{\partial x^a}g^c + \mathcal{G}_{bc}\frac{\partial g^c}{\partial x^a} - \frac{\partial \mathcal{G}_{ac}}{\partial x^b}g^c\,.$$

Substituting into the above expression for $(\flat_{\mathcal{G}})_*\left(g^\mathrm{C}\right)$,

$$\begin{aligned}
(\flat_{\mathcal{G}})_*\left(g^\mathrm{C}\right) &= g^a\frac{\partial}{\partial x^a} + \left\{ g^c\frac{\partial \mathcal{G}_{ab}}{\partial x^c} + \frac{\partial \mathcal{G}_{bc}}{\partial x^a}g^c + \mathcal{G}_{bc}\frac{\partial g^c}{\partial x^a} - \frac{\partial \mathcal{G}_{ac}}{\partial x^b}g^c \right\} v^b\frac{\partial}{\partial p_a}\\
&= g^a\frac{\partial}{\partial x^a} + \left\{ g^c\left(\frac{\partial \mathcal{G}_{ab}}{\partial x^c} + \frac{\partial \mathcal{G}_{bc}}{\partial x^a} - \frac{\partial \mathcal{G}_{ac}}{\partial x^b} \right) + \mathcal{G}_{bc}\frac{\partial g^c}{\partial x^a} \right\} v^b\frac{\partial}{\partial p_a}\\
&= g^a\frac{\partial}{\partial x^a} + \left\{ 2g^c\mathcal{G}_{bd}\Gamma^d_{ac} + \mathcal{G}_{bc}\frac{\partial g^c}{\partial x^a} \right\} v^b\frac{\partial}{\partial p_a} = \mathrm{grad}_{\mathcal{G}^\nabla}\, V^g \circ \flat_{\mathcal{G}}\,.
\end{aligned}$$

Therefore, the first equality in (5.1) holds for every $i = 0,1,\dots,m$. The equality $(\flat_{\mathcal{G}})_* g_j^\mathrm{V} = \mathrm{grad}_{\mathcal{G}^\nabla}\, V_j^\mathrm{V} \circ \flat_{\mathcal{G}}$, $j = 1,\dots,m$, follows by considering (2.6) and the fact that the vector fields g_j are gradient by hypothesis,

$$(\flat_{\mathcal{G}})_*\left(g^\mathrm{V}\right) = \mathcal{G}_{ab}g^b\frac{\partial}{\partial p_a} = \frac{\partial V}{\partial x^a}\frac{\partial}{\partial p_a} = \mathrm{grad}_{\mathcal{G}^\nabla}\, V^\mathrm{V} \circ \flat_{\mathcal{G}}\,.$$

As for $V^{g_j} \circ \flat_{\mathcal{G}} = V_j^\mathrm{C}$, for each $v \in T_x M$, we compute $V^{g_j} \circ \flat_{\mathcal{G}}(v) = \mathcal{G}_{ab}v^b g_j^a = \partial V_j/\partial x^b \cdot v^b = <dV_j, v> = V_j^\mathrm{C}(v)$. The last equality follows trivially. Consequently, the prolongation and the gradient extension of a nonlinear system Σ which is itself gradient are externally equivalent, in particular weakly externally equivalent systems.

To prove the converse implication, we need some intermediate steps that we describe in what follows. $\square$

LEMMA 5.5. *Let Σ be a nonlinear system of the form (2.1). Under the hypothesis of Theorem 5.4, assume that the prolongation Σ^p and the gradient extension Σ^e are weakly externally equivalent. Then there exists a unique diffeomorphism $\varphi : TM \to T^*M$ such that*

$$\begin{aligned}
(\varphi)_* g_i^\mathrm{C} &= \mathrm{grad}_{\mathcal{G}^\nabla}\, V^{g_i} \circ \varphi\,, & V^{g_j} \circ \varphi &= V_j^\mathrm{C}\,,\\
(\varphi)_* g_j^\mathrm{V} &= \mathrm{grad}_{\mathcal{G}^\nabla}\, V_j^\mathrm{V} \circ \varphi\,, & V_j^\mathrm{V} \circ \varphi &= V_j^\mathrm{V}\,,
\end{aligned} \tag{5.2}$$

for all $i = 0,1,\dots,m$, $j = 1,\dots,m$. Moreover, φ is a bundle morphism over the identity $\mathrm{Id}_M : M \to M$, i.e. in natural coordinates $\varphi(x,v) = (x,\phi(x,v))$, for certain map $\phi : T_x M \to T_x^* M$, $x \in M$.

Proof. By Proposition 3.2 and Corollary 3.6, we have that both the prolongation and the gradient extension are observable systems. Since they are also weakly externally equivalent by assumption, Corollary 4.4 ensures that there exists a unique diffeomorphism $\varphi : TM \to T^*M$ verifying (5.2). Applying now Corollary 4.4 to $\Sigma^1 = \Sigma = \Sigma^2$, we deduce that there exists a unique diffeomorphism from M to M

mapping the original nonlinear system to itself, namely the identity mapping. Using uniqueness and the fact that φ satisfies (5.2), it then follows that φ is of the form $\varphi(x, v) = (x, \phi(x, v))$, for certain map $\phi : T_x M \to T_x^* M$, $x \in M$. □

LEMMA 5.6. *Under the same assumptions as in Lemma 5.5, there exists a unique pseudo-Riemannian metric $\mathcal{G}$ on M such that $\flat_{\mathcal{G}} = \varphi$, i.e. $\flat_{\mathcal{G}}(v) = \phi(x, v)$ for all $v \in T_x M$.*

Proof. It follows from $V^{g_j} \circ \varphi = V_j^{\mathsf{C}}$ (cf. equation (5.2)) and the structure of the diffeomorphism φ that

$$< \phi(x, v), g_j(x) >=< dV_j(x), v >, \quad \forall v \in T_x M, \ j = 1, \ldots, m.$$

Furthermore, from $(\varphi)_* g_i^{\mathsf{C}} = \operatorname{grad} V^{g_i} \circ \varphi$ (see eq. (5.2)), it follows that

$$\mathcal{L}_{\operatorname{grad} V^{g_i}} V^{g_j} \circ \varphi = \mathcal{L}_{g_i^{\mathsf{C}}} V_j^{\mathsf{C}}, \quad i = 0, 1, \ldots, m, \ j = 1, \ldots, m.$$

Using now Lemma 3.4 (i), we get $< \phi(x, v), \langle g_i : g_j \rangle(x) >=< d\left(\mathcal{L}_{g_i} V_j\right)(x), v >$. In general for all $v \in T_x M$,

$$< \phi(x, v), \langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_s : g_j \rangle \rangle \ldots \rangle \rangle(x) >=$$
$$=< d\left(\mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_s} V_j\right)(x), v >, \quad (5.3)$$

with the X_r, $r = 1, \ldots, s$ equal to some g_i, $i = 0, 1, \ldots, m$. Since the right-hand side of this equation is linear in v and the distribution generated by the space S_0 is full-rank by hypothesis, it follows that for each $x \in M$ there exists a unique matrix $\mathcal{G}(x)$ such that $\phi(x, v) = \mathcal{G}(x)v$. Since φ is a diffeomorphism, $\mathcal{G}(x)$ is non-singular for every x and depends smoothly on the base point. Consider the adjoint mapping of φ, $\varphi^T : TM \to T^*M$, defined by $< \varphi(v), w >=< v, \varphi^T(w) >$, $v, w \in T_x M$, $x \in M$. Then, $\varphi^T(x, v) = (x, \mathcal{G}^T(x)v)$. It follows from (5.3) that $\mathcal{G}(x)$ satisfies

$$\varphi^T(\langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_s : g_j \rangle \rangle \ldots \rangle \rangle(x)) = d\left(\mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_s} V_j\right)(x), \quad (5.4)$$

with the X_r as above. Let us see now that $\mathcal{G}(x) = \mathcal{G}^T(x)$. Note that in local coordinates $(\varphi)_* g_j^{\mathsf{V}} = \operatorname{grad}_{\mathcal{G}^\nabla} V_j^{\mathsf{V}} \circ \varphi$ yields,

$$\begin{pmatrix} I & 0 \\ \frac{\partial}{\partial x}(\mathcal{G}(x)v) & \mathcal{G}(x) \end{pmatrix} \begin{pmatrix} 0 \\ g_j(x) \end{pmatrix} = \begin{pmatrix} 0 \\ \left(\frac{\partial V_j}{\partial x}\right)^T (x) \end{pmatrix},$$

or, equivalently, $\mathcal{G}(x)g_j(x) = (\partial V_j / \partial x)^T (x)$, $j = 1, \ldots, m$, which in intrinsic terms, can be written as $\varphi(g_j) = dV_j$. Now,

$$< \varphi\left(\langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle\right), \langle Y_1 : \langle Y_2 : \langle Y_3, \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle >$$
$$=< \langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle, \varphi^T\left(\langle Y_1 : \langle Y_2 : \langle Y_3, \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle\right) >.$$

Using (5.4), the latter is equal to

$$< \langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle, d\mathcal{L}_{Y_1} \mathcal{L}_{Y_2} \ldots \mathcal{L}_{Y_{s_2}} V_k >$$
$$=< d\mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_{s_1}} V_j, \langle Y_1 : \langle Y_2 : \langle \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle >,$$

where in the last equality we have used the property (a) of the compatibility definition between the nonlinear system Σ and the affine connection ∇. Finally,

$$< \varphi\left(\langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle\right), \langle Y_1 : \langle Y_2 : \langle Y_3, \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle >$$
$$=< \varphi^T\left(\langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle\right), \langle Y_1 : \langle Y_2 : \langle \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle >,$$

By the assumption on the full-rankness of the distribution $\mathcal{S}_0$, we conclude that

$$\varphi\left(\langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_{s_1} : g_j\rangle\rangle \ldots\rangle\rangle\right) =$$
$$= \varphi^T\left(\langle X_1 : \langle X_2 : \langle X_3, \ldots : \langle X_{s_1} : g_j\rangle\rangle \ldots\rangle\rangle\right),$$

which in turn implies that $\varphi(x) = \varphi^T(x)$, i.e., the matrix $\mathcal{G}(x)$ is symmetric. $\square$

LEMMA 5.7. *Under the same assumptions as in Lemma 5.5, the torsion-free affine connection ∇ is the Levi-Civita connection corresponding to the pseudo-Riemannian metric $\mathcal{G}$.*

Proof. First of all, note that

$$< V^{\langle\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\rangle\ldots\rangle\rangle:\langle Y_1:\langle Y_2:\langle\ldots:\langle Y_{s_2}:g_k\rangle\rangle\rangle\ldots\rangle\rangle\rangle}, d(\mathcal{L}_{Z_1}\mathcal{L}_{Z_2}\ldots\mathcal{L}_{Z_{s_3}}V_l) >$$
$$= \mathcal{L}_{\langle\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\rangle\ldots\rangle\rangle:\langle Y_1:\langle Y_2:\langle\ldots:\langle Y_{s_2}:g_k\rangle\rangle\rangle\ldots\rangle\rangle\rangle}\left[\mathcal{L}_{Z_1}\mathcal{L}_{Z_2}\ldots\mathcal{L}_{Z_{s_3}}V_l\right]$$
$$= \mathcal{L}_{\langle Z_1:\langle Z_2:\langle\ldots:\langle Z_{s_3}:g_l\rangle\rangle\rangle\ldots\rangle\rangle}\left[\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\rangle\ldots\rangle\rangle}\left[\mathcal{L}_{Y_1}\mathcal{L}_{Y_2}\ldots\mathcal{L}_{Y_{s_2}}V_k\right]\right] \qquad (5.5)$$
$$= < \left(\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\rangle\ldots\rangle\rangle}\left[\mathcal{L}_{Y_1}\mathcal{L}_{Y_2}\ldots\mathcal{L}_{Y_{s_2}}V_k\right]\right)^{\text{C}}\circ\varphi^{-1}, d(\mathcal{L}_{Z_1}\mathcal{L}_{Z_2}\ldots\mathcal{L}_{Z_{s_3}}V_l) >,$$

where in the second equality we have used the property (b) of the compatibility definition between the nonlinear system Σ and the affine connection ∇. Since the observation space of the nonlinear system Σ is generated by the functions of the form $\mathcal{L}_{Z_1}\mathcal{L}_{Z_2}\ldots\mathcal{L}_{Z_{s_3}}V_l$, and Σ is observable by hypothesis, we conclude that

$$V^{\langle\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\rangle\ldots\rangle\rangle:\langle Y_1:\langle Y_2:\langle\ldots:\langle Y_{s_2}:g_k\rangle\rangle\rangle\ldots\rangle\rangle\rangle}\circ\varphi =$$
$$= \left(\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\rangle\ldots\rangle\rangle}\left[\mathcal{L}_{Y_1}\mathcal{L}_{Y_2}\ldots\mathcal{L}_{Y_{s_2}}V_k\right]\right)^{\text{C}}.$$

Given the structure of the mapping φ (cf. Lemmas 5.5 and 5.6), and equation (5.4), this equality can be rewritten as,

$$\mathcal{G}(\langle\langle X_1 : \langle X_2 : \langle\ldots : \langle X_s : g_j\rangle\rangle \ldots\rangle\rangle : \langle Y_1 : \langle Y_2 : \langle\ldots : \langle Y_{s_2} : g_k\rangle\rangle \ldots\rangle\rangle\rangle, \cdot)$$
$$= d < \varphi(\langle Y_1 : \langle Y_2 : \langle\ldots : \langle Y_{s_2} : g_k\rangle\rangle \ldots\rangle\rangle), \langle X_1 : \langle X_2 : \langle\ldots : \langle X_s : g_j\rangle\rangle \ldots\rangle\rangle >$$
$$= d\left(\mathcal{G}(\text{grad}_{\mathcal{G}}\left(\mathcal{L}_{Y_1}\mathcal{L}_{Y_2}\ldots\mathcal{L}_{Y_{s_2}}V_k\right), \text{grad}_{\mathcal{G}}\left(\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}V_j\right)\right))$$
$$= d\{\mathcal{L}_{Y_1}\mathcal{L}_{Y_2}\ldots\mathcal{L}_{Y_{s_2}}V_k : \mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}V_j\}_{\mathcal{G}}.$$

Since $\text{grad}_{\mathcal{G}}\{f : g\}_{\mathcal{G}} = \langle\text{grad}_{\mathcal{G}} f : \text{grad}_{\mathcal{G}} g\rangle_{\nabla^{\mathcal{G}}}$, we conclude

$$\langle\langle X_1 : \langle X_2 : \langle\ldots : \langle X_s : g_j\rangle\rangle \ldots\rangle\rangle : \langle Y_1 : \langle Y_2 : \langle\ldots : \langle Y_{s_2} : g_k\rangle\rangle \ldots\rangle\rangle\rangle =$$
$$= \langle\langle X_1 : \langle X_2 : \langle\ldots : \langle X_s : g_j\rangle\rangle \ldots\rangle\rangle : \langle Y_1 : \langle Y_2 : \langle\ldots : \langle Y_{s_2} : g_k\rangle\rangle \ldots\rangle\rangle\rangle_{\mathcal{G}}.$$

Using the fact that $\mathcal{S}_0$ is full-rank, we deduce that $\langle X : Y\rangle = \langle X : Y\rangle_{\mathcal{G}}$ for all X, $Y \in \mathfrak{X}(M)$. Finally, using the fact that ∇ is torsion-free, we compute

$$\nabla_X Y = \frac{1}{2}\left(\langle X : Y\rangle + [X, Y]\right) = \frac{1}{2}\left(\langle X : Y\rangle_{\mathcal{G}} + [X, Y]\right) = \nabla_X^{\mathcal{G}} Y, \quad \forall X, Y \in \mathfrak{X}(M),$$

which concludes the result. $\square$

We are now ready to conclude the proof of Theorem 5.4.

Proof of Theorem 5.4. $\Leftarrow$) Assume the prolongation Σ^p and the gradient extension Σ^e are weakly externally equivalent. From Lemmas 5.5, 5.6 and 5.7, we deduce the existence of a pseudo-Riemannian metric $\mathcal{G}$ on M such that $\nabla = \nabla^{\mathcal{G}}$ and the unique diffeomorphism between TM and T^*M relating Σ^p and Σ^e and verifying (5.2) is $\flat_{\mathcal{G}}$.

From $(\flat_\mathcal{G})_* g_j^{\text{V}} = \text{grad}_{\mathcal{G}^\nabla} V_j^{\text{V}} \circ \flat_\mathcal{G}$, we deduce $\flat_\mathcal{G}(g_j) = dV_j$, and hence $\text{grad}_\mathcal{G} V_j = g_j$, $j = 1, \ldots, m$. Finally, we show that g_0 is a locally gradient vector field. From $(\flat_\mathcal{G})_* g_0^{\text{C}} = \text{grad}_{\mathcal{G}^\nabla} V^{g_0} \circ \flat_\mathcal{G}$ and the local expression (2.8) of the Christoffel symbols of the Levi-Civita connection $\nabla^\mathcal{G}$, we deduce that

$$\frac{\partial}{\partial x^b}\left(\mathcal{G}_{ac}g_0^c\right) = \frac{\partial}{\partial x^a}\left(\mathcal{G}_{bc}g_0^c\right), \quad \forall a, b = 1, \ldots, n,$$

which implies that the one-form $\flat_\mathcal{G}(g_0)$ is closed. ☐

EXAMPLE 5.8. Consider a linear input-state-output system Σ on $M = \mathbb{R}^n$, i.e., $\dot{x} = Ax + Bu$, $y = Cx$, $x \in \mathbb{R}^n$, with A a $(n \times n)$-matrix, B a $(n \times m)$-matrix and C a $(m \times n)$-matrix. Assume Σ is observable and controllable. Consider the trivial connection ∇ on $\mathbb{R}^n$ whose Christoffel symbols are given by $\Gamma_{bc}^a = 0$, $1 \leq a, b, c \leq n$. One can easily verify that Σ is compatible with ∇, and, using the hypothesis of controllability, that the distribution $\mathcal{S}_0$ has full rank. The prolonged system consists of the system itself together with the variational equations $\dot{v} = Av + Bu^p$, $y^p = Cv$, and the gradient extension consists of the system itself together with the equations $\dot{p} = A^T p + C^T u^e$, $y^e = B^T p$. Hence the prolonged system and the gradient extension are weakly externally equivalent if and only if the impulse responses of $\dot{v} = Av + Bu^p$, $y^p = Cv$ and $\dot{p} = A^T p + C^T u^e$, $y^e = B^T p$ are equal, that is, $W(t) = W^T(t)$, with $W(t) := Ce^{At}B$. Thus from Theorem 5.4 we recover the classical result (see e.g. [25]) that an observable and controllable linear system is a gradient system (with respect to the trivial connection) if and only if $W(t) = W^T(t)$. •

REMARK 5.9. Note that, given the torsion-free affine connection ∇, the pseudo-Riemannian metric $\mathcal{G}$ obtained in the proof of Theorem 5.4 is unique such that Σ is locally gradient with respect to it. In Section 6 below, we investigate the uniqueness (up to isometry) of gradient realizations with the same input-output behavior. •

REMARK 5.10. In general, we cannot ensure that the drift vector field g_0 is globally gradient, unless we impose some additional conditions on the topology of the state space M (for instance, that the first Betti number of M is zero). This is analogous to the situation in the Hamiltonian setting [12]. •

REMARK 5.11. As noted in Section 2.2, one can verify that the pseudo-Riemannian metric on T^*M defined by $\mathcal{G}^\nabla$ corresponds to the pullback by $\sharp_\mathcal{G}$ of the complete lift $\mathcal{G}^\text{C}$ to TM of the original metric $\mathcal{G}$ on M. •

REMARK 5.12. A different way to prove the same result which indeed keeps a closer parallelism with the proof for the Hamiltonian case [12] would be the following. Once one has proved Lemmas 5.5 and 5.6, instead of proving Lemma 5.7, one can show that

$$(\varphi)_* \langle X_1 : \langle X_2 : \langle \ldots : \langle X_s : g_j \rangle \rangle \ldots \rangle \rangle^\text{C} = \text{grad } V^{\langle X_1 : \langle X_2 : \langle \ldots : \langle X_s : g_j \rangle \rangle \ldots \rangle \rangle} \circ \varphi, \quad (5.6)$$

for any $j \in \{1, \ldots, m\}$ and $X_r \in \{g_0, g_1, \ldots, g_m\}$, $r = 1, \ldots, s$. This can be done by considering the following vector fields on T^*M,

$$\mathcal{Z}_1 = (\varphi)_* \langle X_1 : \langle X_2 : \langle \ldots : \langle X_s : g_j \rangle \rangle \ldots \rangle \rangle^\text{C} \circ \varphi^{-1},$$
$$\mathcal{Z}_2 = \text{grad } V^{\langle X_1 : \langle X_2 : \langle \ldots : \langle X_s : g_j \rangle \rangle \ldots \rangle \rangle},$$

and showing that their action on the observation space $\mathcal{H}^e$ of Σ^e is the same. To see this, recall from Proposition 3.5 that $\mathcal{H}^e = V^{\mathcal{S}_0} + (\mathcal{H} + \mathfrak{h})^\text{V}$. Consider a function of the form $\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}V_j$, with X_r, $r = 1, \ldots, s$, equal to g_i, $i = 0, 1, \ldots, m$, and

$j = 1, \ldots, m$. Then,

$$
\mathcal{L}_{\mathcal{Z}_1}\left[(\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}V_j)^{\mathsf{V}}\right] =
$$
$$
= \left(\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\ldots\rangle\rangle^{\mathsf{C}}}\left[(\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}V_j)^{\mathsf{V}}\circ\varphi\right]\right)\circ\varphi^{-1} =
$$
$$
= \left(\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\ldots\rangle\rangle}\left[\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}V_j\right]\right)^{\mathsf{V}} ,
$$

where we have used twice the fact that φ is the identity mapping on the base mani-fold M. On the other hand,

$$
\mathcal{L}_{\mathcal{Z}_2}\left[(\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}V_j)^{\mathsf{V}}\right] = \left(\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\ldots\rangle\rangle}\left[\mathcal{L}_{X_1}\mathcal{L}_{X_2}\ldots\mathcal{L}_{X_s}V_j\right]\right)^{\mathsf{V}} ,
$$

using property (ii) in Lemma 3.4. The same argument also guarantees that the action of $\mathcal{Z}_1$ and $\mathcal{Z}_2$ is the same over the vertical lifts of the functions spanning $\mathfrak{h}$. Finally, let $\langle Y_1 : \langle Y_2 : \langle\ldots : \langle Y_{s_2} : g_k\rangle\rangle\ldots\rangle\rangle \in S_0$ and consider the corresponding function on T^*M, $V^{\langle Y_1:\langle Y_2:\langle\ldots:\langle Y_{s_2}:g_k\rangle\rangle\ldots\rangle\rangle}$. Then,

$$
\mathcal{L}_{\mathcal{Z}_1}\left[V^{\langle Y_1:\langle Y_2:\langle\ldots:\langle Y_{s_2}:g_k\rangle\rangle\ldots\rangle\rangle}\right] =
$$
$$
= \left(\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\ldots\rangle\rangle^{\mathsf{C}}}\left[V^{\langle Y_1:\langle Y_2:\langle\ldots:\langle Y_{s_2}:g_k\rangle\rangle\ldots\rangle\rangle}\circ\varphi\right]\right)\circ\varphi^{-1} =
$$
$$
= \left(\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\ldots\rangle\rangle^{\mathsf{C}}}\left[(\mathcal{L}_{Y_1}\mathcal{L}_{Y_2}\ldots\mathcal{L}_{Y_{s_2}}V_k)^{\mathsf{C}}\right]\right)\circ\varphi^{-1}
$$
$$
= \left(\mathcal{L}_{\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\ldots\rangle\rangle}\left[\mathcal{L}_{Y_1}\mathcal{L}_{Y_2}\ldots\mathcal{L}_{Y_{s_2}}V_k\right]\right)^{\mathsf{C}}\circ\varphi^{-1} , \quad (5.7)
$$

where we have used equation (5.4). In addition,

$$
\mathcal{L}_{\mathcal{Z}_2}\left[V^{\langle Y_1:\langle Y_2:\langle\ldots:\langle Y_{s_2}:g_k\rangle\rangle\ldots\rangle\rangle}\right] = V^{\langle\langle X_1:\langle X_2:\langle\ldots:\langle X_s:g_j\rangle\rangle\ldots\rangle\rangle:\langle Y_1:\langle Y_2:\langle\ldots:\langle Y_{s_2}:g_k\rangle\rangle\ldots\rangle\rangle\rangle} , \quad (5.8)
$$

where we have used property (i) in Lemma 3.4. Now, equation (5.5) implies that (5.7) and (5.8) coincide. Therefore, $\mathcal{Z}_1$ and $\mathcal{Z}_2$ coincide over $\mathcal{H}^e$, and this concludes the proof of (5.6).

Now, one can proceed by taking local coordinates $(x^1, \ldots, x^n)$ in M such that every coordinate function x^i is of the form $\mathcal{L}_{X_1}\ldots\mathcal{L}_{X_s}V_j$ for a certain $j \in \{1, \ldots, m\}$ and certain vector fields $X_r \in \{g_0, g_1, \ldots, g_m\}$, $r = 1, \ldots, s$. It follows from (5.4) that there exists n independent vector fields $k^1, \ldots, k^n$ of the form $\langle X_1 : \langle X_2 : \langle\ldots : \langle X_s : g_j\rangle\rangle\ldots\rangle\rangle$ such that $\flat_{\mathcal{G}}(k^i) = dx^i$. Finally, spelling out eq. (5.6) for the vector fields k^i and making use of the symmetry of $\mathcal{G}$, one obtains that the Christoffel symbols of the affine connection ∇ are precisely given by (2.8), which concludes the result.   ●

6. Uniqueness of the gradient realization. In this section, we investigate the gradient analog of the following well-known result for Hamiltonian systems: if two minimal Hamiltonian systems have the same input-output map, then they are sym-plectomorphic [3, 21]. We will see how the setting of Theorem 5.4 also provides sufficient conditions under which a similar result holds for gradient realizations.

In [23], P. Varaiya conjectured that if there exists a state-space diffeomorphism between two locally controllable gradient systems, then the diffeomorphism is actu-ally an isometry between the underlying pseudo-Riemannian manifolds (see also [24]). Subsequently, in [1, 2], J. Basto Gonçalves produced an example of two locally con-trollable and observable gradient systems living on the same state space with state-space diffeomorphism given by the identity mapping, where however the Rieman-nian metrics are different; thus providing a counterexample to the conjecture by Varaiya. For the sake of completeness, we review it in the following.

EXAMPLE 6.1 ([1, 2]). Consider two gradient systems Σ^1 and Σ^2 on $M^1 = M^2 = \mathbb{R}^4$ with Riemannian metrics $\mathcal{G}^1$ and $\mathcal{G}^2$ given respectively by

$$\mathcal{G}^1(x_1, x_2, x_3, x_4) = dx_1 \otimes dx_1 + e^{-x_4} dx_2 \otimes dx_2 + e^{-x_1} dx_3 \otimes dx_3 + e^{-x_3} dx_4 \otimes dx_4\,,$$

$$\mathcal{G}^2(x_1, x_2, x_3, x_4) = dx_1 \otimes dx_1 + e^{-x_4} dx_2 \otimes dx_2 + (e^{-x_1} + e^{x_3}) dx_3 \otimes dx_3$$
$$+ e^{-x_3}(1 + e^{2x_1}) dx_4 \otimes dx_4 - e^{x_1}(dx_3 \otimes dx_4 + dx_4 \otimes dx_3)\,.$$

Furthermore, let Σ^1 and Σ^2 have both zero drift vector fields and the same output functions given by

$$y_1 = V_1(x) := x_1\,, \quad y_2 = V_2(x) := x_2 + x_3 + x_4\,.$$

From the definition of $\mathcal{G}^1$ and $\mathcal{G}^2$, it easily follows that the input vector fields of both systems are the same, i.e.,

$$\text{grad}_{\mathcal{G}^1} V_1 = \text{grad}_{\mathcal{G}^2} V_1 = \frac{\partial}{\partial x_1}\,,$$

$$\text{grad}_{\mathcal{G}^1} V_2 = \text{grad}_{\mathcal{G}^2} V_2 = e^{x_4} \frac{\partial}{\partial x_2} + e^{x_1} \frac{\partial}{\partial x_3} + e^{x_3} \frac{\partial}{\partial x_4}\,.$$

Therefore, Σ^1 and Σ^2 are externally equivalent with state-space diffeomorphism given by the identity mapping $\text{Id} : \mathbb{R}^4 \to \mathbb{R}^4$. However, the metrics $\mathcal{G}^1$ and $\mathcal{G}^2$ are different, and hence the identity mapping is *not* an isometry. It should also be noted that Σ^1 and Σ^2 are both controllable and observable.   •

The following result shows that, under the hypotheses of Theorem 5.4, a state-space diffeomorphism linking two gradient systems is an isometry, *provided* the state-space diffeomorphism is already known to respect the affine connections determined by their respective pseudo-Riemannian metrics. A similar statement is already contained in [1, 2]. Here we make use of an argument given in [12], p. 58 for the case of Hamiltonian systems.

PROPOSITION 6.2. *Let Σ^1 and Σ^2 be two gradient systems with state spaces $(M^1, \mathcal{G}^1)$ and $(M^2, \mathcal{G}^2)$, respectively. For $i = 1, 2$, assume that Σ^i is observable with $\dim d\mathcal{H}^i$ constant, and that the distribution $\mathcal{S}_0^i$ is full-rank. Furthermore, let Σ^1 and Σ^2 be externally equivalent with the corresponding state-space diffeomorphism $\psi : M^1 \to M^2$ satisfying*

$$\psi_*(\nabla_X^{\mathcal{G}^1} Y) \circ \psi^{-1} = \nabla_{\psi_* X \circ \psi^{-1}}^{\mathcal{G}^2}(\psi_* Y \circ \psi^{-1})\,, \quad \text{for all } X, Y \in \mathfrak{X}(M^1)\,. \tag{6.1}$$

Then $\psi^* \mathcal{G}^2 = \mathcal{G}^1$, that is, ψ is an isometry.

Proof. By Lemmas 5.5 and 5.6, the map $\varphi^i = \flat_{\mathcal{G}^i}$ is the unique diffeomorphism satisfying (5.2) for system Σ^i, $i = 1, 2$. It is easily checked that since Σ^1 and Σ^2 are externally equivalent with state-space diffeomorphism ψ, then their prolongations Σ^{1p} and Σ^{2p} are externally equivalent with uniquely determined state-space diffeomorphism given by $\psi_* : TM^1 \to TM^2$. Furthermore, it can be readily checked that the gradient extensions Σ^{1e} and Σ^{2e} are externally equivalent with state-space diffeomorphism $\psi^* : T^*M^2 \to T^*M^1$, *provided* ψ satisfies (6.1). This is because equation (6.1) implies that ψ^* respects the Riemannian extensions $\mathcal{G}^{\nabla^{\mathcal{G}^1}}$ and $\mathcal{G}^{\nabla^{\mathcal{G}^2}}$ determined, respectively, by the affine connections $\nabla^{\mathcal{G}^1}$ and $\nabla^{\mathcal{G}^2}$. Therefore, by the uniqueness of all these state-space diffeomorphisms, we obtain the following commutative diagram

$$
\begin{array}{ccc}
TM^1 & \xrightarrow{\ \psi_* \ } & TM^2 \\
\varphi^1 \downarrow & & \downarrow \varphi^2 \\
T^*M^1 & \xleftarrow{\ \psi^* \ } & T^*M^2
\end{array}
$$

that is,

$$\psi^* \circ \varphi^2 \circ \psi_* = \varphi^1 \,. \tag{6.2}$$

Recalling that $\varphi^i = \flat_{\mathcal{G}^i}$, $i = 1, 2$, it is readily seen that (6.2) is equivalent to

$$\psi^* \mathcal{G}^2 = \mathcal{G}^1 \,. \tag{6.3}$$

that is, $\psi : \left(M^1, \mathcal{G}^1 \right) \to \left(M^2, \mathcal{G}^2 \right)$ is an isometry. □

REMARK 6.3. Note that in Example 6.1 the torsion-free connections determined by $\mathcal{G}^1$ and $\mathcal{G}^2$ are *different*, and hence the identity map does not respect them.   •

REMARK 6.4. Since eq. (6.2) is equivalent to eq. (6.3), one may also conclude that under the conditions of Theorem 5.4, the state-space diffeomorphism $\psi : M^1 \to M^2$ is an isometry if and only if $\psi^* : T^*M^2 \to T^*M^1$ is a state-space diffeomorphism between Σ^{1e} and Σ^{2e}.   •

7. Conclusions. We have discussed necessary and sufficient conditions for a nonlinear control system to be realizable as a gradient control system with respect to a pseudo-Riemannian metric whose Levi-Civita connection coincides with a given affine connection. The results rely on a suitable notion of compatibility of the system with respect to the given affine connection, and on the input-output behavior of the prolonged system and the gradient extension. The symmetric product associated with an affine connection plays a key role in the discussion. We believe that the developments in this paper do not only give insight in the system-theoretic properties of the physically motivated class of gradient control systems, but also shed light on the differential-geometric properties of gradient and Lagrangian control systems. Future work will include the investigation of necessary and sufficient conditions that guarantee the existence of an affine connection such that the hypothesis of Theorem 5.4 are satisfied, the development of equivalent characterizations in terms of the input-output behavior of the original nonlinear system and the study of the application of the results to specific classes of nonlinear systems, such as bilinear, homogeneous, and polynomial systems.

REFERENCES

[1] J. Basto Gonçalves: Equivalence of gradient systems. *Control Theory Centre Report*, no. 84, University of Warwick, UK, 1979.

[2] J. Basto Gonçalves: Equivalencia de sistemas de gradiente. *Portugal. Math.* **40** (3) (1981), 263-277.

[3] J. Basto Gonçalves: Realization theory for Hamiltonian systems. *SIAM J. Control Optim.* **25** (1) (1987), 63-73.

[4] R.K. Brayton, J.K. Moser: A theory of nonlinear networks, Part I and Part II. *Quart. Appl. Math.* **22** (1964), 1-33 and 81-104.

[5] R.W. Brockett, A. Rahimi: Lie algebras and linear differential equations. In *Ordinary Differential Equations*, L. Weiss ed., Academic Press, New York, 1972.

[6] F. Bullo, A.D. Lewis: *Geometric Control of Simple Mechanical Systems*. Text in Applied Mathematics Series, Springer Verlag, 2004, incomplete work under contract.

[7] M.P. do Carmo: *Riemannian geometry*. Birkhäuser, Boston-Basel-Berlin, 1992.

[8] J. Cortés, A.J. van der Schaft, P.E. Crouch: Gradient realization of nonlinear control systems. *Proc. IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control,* Seville, Spain, 2003, pp. 73-78.

[9] P.E. Crouch: Geometric structures in systems theory. *Proc. IEE. D. Control Theory and Applications* **128** (5) (1981), 242-252.

[10] P.E. Crouch, M. Irving: On finite Volterra series which admit Hamiltonian realizations. *Math. Systems Theory* **17** (1984), 293-318.

[11] P.E. Crouch, F. Lamnabhi-Lagarrigue, A.J. van der Schaft: Adjoint and Hamiltonian input-output differential equations. *IEEE Trans. Automat. Contr.* **AC-40** (1995), 603-615.

[12] P.E. Crouch, A.J. van der Schaft: *Variational and Hamiltonian control systems*. Lectures Notes in Control and Information Sciences **101**, Springer-Verlag, New York, 1987.

[13] R. Hermann, A.J. Krener: Nonlinear controllability and observability. *IEEE Trans. Automat. Control* **AC-22** (1977), 728-740.

[14] B. Jakubczyk: Hamiltonian realizations of nonlinear systems. In *Theory and Applications of Nonlinear Control Systems*, C.I. Byrnes, A. Lindquist eds., North-Holland, 1986, pp. 261-271.

[15] S. Kobayashi, K. Nomizu: *Foundations of Differential Geometry*. Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers, Wiley, New-York, 1963.

[16] A.D. Lewis: Affine connections and distributions with applications to nonholonomic mechanics. *Rep. Math. Phys.* **42** (1/2) (1998), 135-164.

[17] A.D. Lewis, R.M. Murray: Configuration controllability of simple mechanical control systems. *SIAM J. Control Optim.* **35** (3) (1997), 766-790.

[18] E.M. Patterson, A.G. Walker: Riemann extensions. *Quart. J. Math.* **3** (1952), 19-28.

[19] R.M. Santilli: *Foundations of Theoretical Mechanics I*. Springer-Verlag, New York, 1978.

[20] W. Sarlet, G. Thompson, G.E. Prince: The inverse problem of the calculus of variations: the use of geometrical calculus in Douglas's analysis. *Trans. Amer. Math. Soc.* **354** (7) (2002), 2897-2919.

[21] A.J. van der Schaft: *System theoretic descriptions of physical systems*. CWI Tract, no. 3, CWI, Amsterdam, 1984.

[22] A.J. van der Schaft: Linearization of Hamiltonian and gradient systems. *IMA J. Math. Contr. & Inf.* **1** (1984), 185-198.

[23] P. Varaiya: Equivalent non-linear networks. In *SIAM-AMS Proceedings III*, AMS, Providence, 1971, pp. 141-147.

[24] T.P. Verma: Equivalence of nonlinear networks. Ph.D. dissertation, University of California, Berkeley, 1969.

[25] J.C. Willems: Dissipative Dynamical Systems, Part II. *Arch. Rat. Mech. Anal.* **45** (1972), 352-292.

[26] K. Yano, S. Ishihara: *Tangent and cotangent bundles*. Marcel Dekker, New York, 1973.

8. Appendix. In this appendix we present a simplifying result concerning the compatibility hypothesis in the statement of Theorem 5.4. In general, checking conditions (a) and (b) in the definition of compatibility between the affine connection ∇ and the nonlinear system Σ cannot be performed for every possible choice of vector fields in $\{g_0, g_1, \ldots, g_m\}$ and $\{V_1, \ldots, V_m\}$. The following result shows that it is enough to check the compatibility condition on a basis of vector fields and the corresponding associated functions once we know that the prolongation and the gradient extension of Σ are weakly externally equivalent.

LEMMA 8.1. *Let ∇ be a torsion-free affine connection. Assume Σ is observable with* $\dim d\mathcal{H}$ *constant, and that the distribution $\mathcal{S}_0$ is full-rank. Assume the prolongation Σ^p and the gradient extension Σ^e of Σ are weakly externally equivalent. Then Σ is compatible with* ∇ *if and only if properties (a) and (b) are verified by a basis of vector fields in S_0.*

Proof. Let $R_1, \ldots, R_n$ be linearly independent vector fields of the form $R_i = \langle X_1^i : \langle X_2^i : \langle \ldots : \langle X_{s_i}^i : g_{j_i} \rangle \ldots \rangle \rangle$, $i = 1, \ldots, n$. Let V_{R_i} denote the function on M given by $\mathcal{L}_{X_1^i} \ldots \mathcal{L}_{X_{s_i}^i} V_{j_i}$. From equation (5.4), we know that $\varphi^T(R_i) = dV_{R_i}$. Assume properties (a) and (b) in the definition of the compatibility condition (cf. Definition 5.1) are verified by any combination of the vector fields $R_1, \ldots, R_n$ and

the functions $V_{R_1}, \ldots, V_{R_n}$. Let $X = \langle X_1 : \langle X_2 : \langle \ldots : \langle X_s : g_k \rangle \rangle \ldots \rangle \rangle$ be any element of S_0, and $V_X = \mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_s} V_k$ the associated function on M. Since S_0 is full-rank, we have that $X = \sum_{i=1}^n f_X^i R_i$. Then,

$$dV_X = d\mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_s} V_k = \varphi^T(X) = \sum_{i=1}^n f_X^i \varphi^T(R_i) = \sum_{i=1}^n f_X^i dV_{R_i} \, .$$

Now, let us see that properties (a) and (b) are naturally verified by all possible choices of vector fields in S_0 and generating functions in $\mathcal{H}$. First,

$$\mathcal{L}_{\langle X_1 : \langle X_2 : \langle \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{Y_1} \mathcal{L}_{Y_2} \ldots \mathcal{L}_{Y_{s_2}} V_k \right]$$

$$= \sum_{i=1}^n f_Y^i dV_{R_i} \left(\sum_{j=1}^n f_X^j R_j \right) = \sum_{i,j=1}^n f_Y^i f_X^j dV_{R_j}(R_i) = \sum_{j=1}^n f_X^j dV_{R_j} \left(\sum_{i=1}^n f_Y^i R_i \right)$$

$$= \mathcal{L}_{\langle Y_1 : \langle Y_2 : \langle \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle} \left[\mathcal{L}_{X_1} \mathcal{L}_{X_2} \ldots \mathcal{L}_{X_{s_1}} V_j \right] \, ,$$

where we have used that condition (a) is verified by the vector fields $R_1, \ldots, R_n$ and the functions $V_{R_1}, \ldots, V_{R_n}$. Secondly,

$$\mathcal{L}_{\langle \langle X_1 : \langle X_2 : \langle \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle : \langle Y_1 : \langle Y_2 : \langle \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle \rangle} \left[\mathcal{L}_{Z_1} \mathcal{L}_{Z_2} \ldots \mathcal{L}_{Z_{s_3}} V_l \right]$$

$$= \sum_{i=1}^n f_Z^i dV_{R_i} \left(\sum_{j=1}^n f_{\langle X : Y \rangle}^j R_j \right) = \sum_{i,j=1}^n f_Z^i f_{\langle X : Y \rangle}^j dV_{R_j}(R_i)$$

$$= \sum_{j=1}^n f_{\langle X : Y \rangle}^j dV_{R_j} \left(\sum_{i=1}^n f_Z^i R_i \right) = < \sum_{j=1}^n f_{\langle X : Y \rangle}^j dV_{R_j}, Z > \, . \quad (8.1)$$

Let us compute the coefficients $f_{\langle X : Y \rangle}^j$. We have

$$\langle \langle X_1 : \langle X_2 : \langle \ldots : \langle X_{s_1} : g_j \rangle \rangle \ldots \rangle \rangle : \langle Y_1 : \langle Y_2 : \langle \ldots : \langle Y_{s_2} : g_k \rangle \rangle \ldots \rangle \rangle \rangle =$$

$$\sum_{i,j=1}^n \langle f_X^i R_i : f_Y^j R_j \rangle = \sum_{i,j=1}^n \left(f_X^i f_Y^j \langle R_i : R_j \rangle + f_X^i R_i [f_Y^j] R_j + f_Y^j R_j [f_X^i] R_i \right)$$

$$= \sum_{k=1}^n \left(\sum_{i,j=1}^n f_X^i f_Y^j f_{\langle R_i : R_j \rangle}^k + \sum_{i=1}^n f_X^i R_i [f_Y^k] + \sum_{j=1}^n f_Y^j R_j [f_X^k] \right) R_k \, .$$

Now, note that $\sum_{k=1}^n < f_{\langle R_i : R_j \rangle}^k dV_{R_k}, R_l > = \sum_{k=1}^n < f_{\langle R_i : R_j \rangle}^k dV_{R_l}, R_k >$ using condition (a) for the vector fields $R_1, \ldots, R_n$ and the functions $V_{R_1}, \ldots, V_{R_n}$. Moreover, using condition (b), $\sum_{k=1}^n < f_{\langle R_i : R_j \rangle}^k dV_{R_l}, R_k > = < dV_{R_l}, \langle R_i : R_j \rangle > = < d(dV_{R_i}[R_j]), R_l >$. Hence, $\sum_{k=1}^n f_{\langle R_i : R_j \rangle}^k dV_{R_k} = d(dV_{R_i}[R_j])$. On the other hand,

$$f_X^i R_i [f_Y^k] dV_{R_k} = f_X^i < df_Y^k, R_i > dV_{R_k}$$

$$= f_X^i < dV_{R_k}, R_i > df_Y^k + f_X^i \left(df_Y^k \wedge dV_{R_k} \right) (R_i, \cdot)$$

Since $f_X^i \left(df_Y^k \wedge dV_{R_k} \right) (R_i, \cdot) = f_X^i \left(d \left(f_Y^k dV_{R_k} \right) \right) (R_i, \cdot) = f_X^i \left(d \left(dV_Y \right) \right) (R_i, \cdot) = 0$, we have

$$f_X^i R_i [f_Y^k] dV_{R_k} = f_X^i < dV_{R_k}, R_i > df_Y^k \, .$$

Analogously, one can see that $f_Y^j R_j[f_X^k]dV_{R_k} = f_Y^j < dV_{R_k}, R_j > df_X^k$. Finally,

$$\sum_{k=1}^{n} f_{\langle X:Y \rangle}^k dV_{R_k} = \sum_{k=1}^{n} \left(\sum_{i,j=1}^{n} f_X^i f_Y^j f_{\langle R_i:R_j \rangle}^k + \sum_{i=1}^{n} f_X^i R_i[f_Y^k] + \sum_{j=1}^{n} f_Y^j R_j[f_X^k] \right) dV_{R_k}$$

$$= \sum_{i,j=1}^{n} f_X^i f_Y^j d(dV_{R_i}[R_j]) + \sum_{i,j=1}^{n} f_X^i < dV_{R_j}, R_i > df_Y^j + \sum_{i,j=1}^{n} f_Y^j < dV_{R_i}, R_j > df_X^i$$

$$= d\left(\sum_{i,j=1}^{n} f_X^i f_Y^j dV_{R_i}[R_j] \right) = d(\mathcal{L}_Y [V_X]).$$

Plugging this equality in (8.1), we get the desired result. $\square$