An intrinsic observer for a class of Lagrangian systems

Article *in* IEEE Transactions on Automatic Control · July 2003

DOI: 10.1109/TAC.2003.812778 · Source: IEEE Xplore

Pierre Rouchon
MINES ParisTech

An Intrinsic Observer for a Class of Lagrangian Systems

Nasradine Aghannan and Pierre Rouchon

Abstract—**We propose a new design method of asymptotic observers for a class of nonlinear mechanical systems: Lagrangian systems with configuration (position) measurements. Our main contribution is to introduce a state (position and velocity) observer that is invariant under any changes of the configuration coordinates. The observer dynamics equations, as the Euler–Lagrange equations, are intrinsic. The design method uses the Riemannian structure defined by the kinetic energy on the configuration manifold. The local convergence is proved by showing that the Jacobian of the observer dynamics is negative definite (contraction) for a particular metric defined on the state–space, a metric derived from the kinetic energy and the observer gains. From a practical point of view, such intrinsic observers can be approximated, when the estimated configuration is close to the true one, by an explicit set of differential equations involving the Riemannian curvature tensor. These equations can be automatically generated via symbolic differentiations of the metric and potential up to order two. Numerical simulations for the ball and beam system, an example where the scalar curvature is always negative, show the effectiveness of such approximation when the measured positions are noisy or include high frequency neglected dynamics.**

Index Terms—**Asymptotic observers, contraction, intrinsic equations, Lagrangian systems, mechanical systems, Riemannian metric.**

I. Introduction

OBSERVERS for nonlinear systems were much studied in the last decade, and real advances were made during this period (see, e.g., [27] and [16]). For the control of mechanical systems, symmetries play an important role (see, e.g., [25], [6], and [21]). In this paper, we show how to exploit "symmetries" (as in [1], where chemical systems are considered) in the design of asymptotic observers for a class of nonlinear systems: Lagrangian mechanical systems with position measurements. The Euler–Lagrange equations are indeed intrinsic: their expression does not depend on the choice of a particular set of configuration coordinates. That represents, roughly speaking, the "symmetry" we are dealing with. Such an invariance has been fully used in optimal control (see, e.g., [33]) and in the design of intrinsic controllers for fully actuated mechanical systems (see, e.g., [9]). Preserving such invariance is the guideline of the observer design presented in this paper. As in [9], our method uses the Riemannian structure and tools (geodesic distance, covariant derivation, curvature; see [15] and [32]) defined by the kinetic energy. Some important work concerning control theory for mechanical systems already use Riemannian geometry. This includes work on controllability [23], [10], motion planning [10]–[12], optimal control [28], [22], and underactuated system stabilization [34], [7], [13], [4].

The local convergence is based on two key points.

- Intrinsic computations using covariant derivatives for the first variation of the observer dynamics. These computations are closely related to the Jacobi equation where curvature terms appear naturally.
- Contraction behavior [24], [18] for a well chosen metric on the phase space. This metric is an extension, on the state–space, of the Riemannian structure defined on the configuration space only. Such extension depends on the observer gains. This metric is closely related to the Sasaki metric [30], [31].

To explain the main idea, let us give a short summary when the dynamics corresponds to geodesics (i.e., no potential and no exterior forces).

For an Euclidian configuration space (no curvature) the geodesic equation reads $\ddot{q} = 0$ where q are Euclidian coordinates. In general, the equation reads $\nabla_{\dot{q}}\dot{q} = 0$ where ∇ is the Riemannian connection. We assume that the configuration q is measured. We want to construct a noiseless estimation $\hat{q}$ and $\hat{v}$ of the position q and velocity $\dot{q} = v$. When $\ddot{q} = 0$, this is very simple. It is sufficient to take the following Luenberger observer:

$$\dot{\hat{q}} = \hat{v} - \alpha(\hat{q} - q) \quad \dot{\hat{v}} = -\beta(\hat{q} - q)$$

with α and β constant and positive to ensure exponential convergence. For $\nabla_{\dot{q}}\dot{q} = 0$, we replace the error injection term $\hat{q}-q$ by an intrinsic error term: the gradient $\mathrm{grad}_{\hat{q}} F$, where $F(\hat{q}, q)$ is the half of the square of the geodesic distance between $\hat{q}$ and q. So, we are led to guess that a good candidate for $\nabla_{\dot{q}}\dot{q} = 0$ could be the following observer:

$$\dot{\hat{q}} = \hat{v} - \alpha\, \mathrm{grad}_{\hat{q}} F \quad \nabla_{\dot{\hat{q}}}\hat{v} = -\beta\, \mathrm{grad}_{\hat{q}} F.$$

It is invariant with respect to a change of coordinates on q. The observer is well-defined for $\hat{q}$ close enough to q. Since $\mathrm{grad}_{\hat{q}} F$ belongs to the tangent space at $\hat{q}$ and since the vector $\hat{v}$ is defined along the curve $t \mapsto \hat{q}(t)$, its covariant derivative $\nabla_{\dot{\hat{q}}}\hat{v}$ is geometrically well-defined along this curve. Nevertheless, it does not ensure convergence for any α and β positive. It is known that negative curvature implies exponential instability of the geodesic flow (see, e.g., Anosov ergodic results on compact manifold with strictly negative curvature [3]). Thus, one has at least to compensate via clever injection of error terms for such

Manuscript received October 11, 2002; revised February 5, 2003. Recommended by Associate Editor M. Reyhanoglu.

The authors are with the École des Mines de Paris, Centre Automatique et Systèmes, 75272 Paris Cedex 06, France (e-mail: aghannan@cas.ensmp.fr; pierre.rouchon@ensmp.fr).

Digital Object Identifier 10.1109/TAC.2003.812778

intrinsic instablity. In fact, our convergent observer slightly differs from the previous one via a curvature term, namely

$$\dot{\hat{q}} = \hat{v} - \alpha \operatorname{grad}_{\hat{q}} F \quad \nabla_{\dot{\hat{q}}} \hat{v} = -\beta \operatorname{grad}_{\hat{q}} F + R(\hat{v}, \operatorname{grad}_{\hat{q}} F)\hat{v}$$

where R is the Riemann curvature tensor. Since $R(\hat{v}, \operatorname{grad}_{\hat{q}} F)\hat{v}$ is linear versus $\operatorname{grad}_{\hat{q}} F$ and quadratic versus $\hat{v}$, we have in fact an automatic gain scheduling with respect to the velocity. Such gains will compensate the divergence or oscillations due to curvature effect (gyroscopic terms).

We prove here by using contraction techniques [24], [18] that, for any positive gains α and β, such an observer is exponentially convergent locally around any geodesic $t \mapsto q(t)$. Indeed, when $\hat{q}$ is close to q, the first variation of the observer dynamics reads

$$\nabla_{\dot{\hat{q}}} \xi = \zeta - \alpha \xi \quad \nabla_{\dot{\hat{q}}} \zeta = -\beta \xi$$

where $\xi = \hat{q} - q$. It corresponds exactly to the classical error dynamics in the Euclidian case ($\ddot{q} = 0$):

$$\dot{\tilde{q}} = \tilde{v} - \alpha \tilde{q} \quad \dot{\tilde{v}} = -\beta \tilde{q}$$

where $\tilde{q} = \hat{q} - q$ and $\tilde{v} = \hat{v} - v$. When $R = 0$, the covariant derivation $\nabla_{\dot{\hat{q}}}$ coincides with the standard operator d/dt in the Euclidian coordinates. The addition of potential and known exterior forces changes the design slightly and requires the use of parallel transport [see (2)].

The paper is organized as follows. Section II is devoted to notations and definitions. In Section III, we describe the design of the intrinsic observer in the general case, and we illustrate the invariance on a tutorial example with exterior forces. In Section IV, we prove the local exponential stability (contraction) around any trajectory (local convergence). We illustrate, on the ball and beam system, the effectiveness of the method with a numerical simulation in Appendix I. Appendix II is devoted to contraction properties.

II. Lagrangian System and Riemannian Metric

We consider a Lagrangian mechanical system with an n-dimensional configuration manifold M equipped with a Riemannian metric; see, e.g., [15]. The local coordinates of $q \in M$ will be denoted by $(q^i)_{i=1,\ldots,n}$. The Lagrangian is given by

$$\mathcal{L}(q, \dot{q}) = \tfrac{1}{2} g_{ij}(q)\dot{q}^i \dot{q}^j - U(q)$$

where the positive–definite symmetric matrix $g(q) = (g_{ij}(q))_{i=1,\ldots,n,\, j=1,\ldots,n}$ defines the metric (inertia matrix) and the scalar function $U(q)$ the potential energy. The Euler–Lagrange equations are, in the local coordinates

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}^i}\mathcal{L}\right) = \frac{\partial}{\partial q^i}\mathcal{L} + u^i(q, t), \qquad i = 1, \ldots, n$$

where $u(q, t) = (u^i(q, t))_{i=1,\ldots,n}$ is a known function of t and q, that corresponds, in general, to some known inputs. The Riemannian formulation of the Euler–Lagrange equations is

$$\nabla_{\dot{q}} \dot{q} = -\operatorname{grad}_q U(q) + g(q)^{-1} u(q, t) \equiv S(q, t)$$

where ∇, grad_q and $g(q)^{-1}$ are, respectively, the Levi–Civita connection, the gradient operator associated to the Riemannian structure and the inverse of the metric matrix $g(q)$. As the position q is measured, the source term $S(q, t)$ is a known time-varying vector-field on M. In local coordinates, this formulation reads:

$$\ddot{q}^i = -\Gamma^i_{jk}(q)\dot{q}^j \dot{q}^k + S^i(q, t) \tag{1}$$

where the connection terms Γ^i_{jk} (Christoffel symbols) are given by

$$\Gamma^i_{jk} = \frac{1}{2} g^{il}\left(\frac{\partial g_{lk}}{\partial q^j} + \frac{\partial g_{jl}}{\partial q^k} - \frac{\partial g_{jk}}{\partial q^l}\right)$$

with g^{il} the entries of g^{-1}. We use here the summation convention: when an index appears both as a subscript and a superscript, the summation according to this index is to be taken.

Recall that $\nabla_{\dot{q}} v$ is the covariant derivative of the vector field v along $\dot{q}$. In local coordinates, it reads

$$\{\nabla_{\dot{q}} v\}^i = \dot{v}^i + \Gamma^i_{jk}(q) v^j \dot{q}^k$$

where $\{\ \}^i$ means coordinate i.

III. Intrinsic Observer

In this section, we define an intrinsic observer for the Lagrangian systems described in the previous section. As the Euler–Lagrange equations are coordinate-free (or intrinsic), our guideline for the state observer design consists in preserving this property. After defining the observer dynamics, we will check that its expression is intrinsic and illustrate it with a simple example. The observer convergence will be considered in the next section.

A. Design

Assume that we measure the position (i.e., the q^is) and that we do not measure the velocity (i.e., the $\dot{q}^i = v^i$). Denote by $\hat{q}$ and $\hat{v}$ the estimations of the position q and the velocity v. They are defined by the following coordinate-free dynamics:

$$\begin{aligned}
&\nabla_{\dot{\hat{q}}}\left(\dot{\hat{q}} + \alpha \operatorname{grad}_{\hat{q}} F(\hat{q}, q)\right) \\
&\quad = \mathcal{T}_{//q\to\hat{q}} S(q, t) - \beta \operatorname{grad}_{\hat{q}} F(\hat{q}, q) \\
&\qquad + R\left(\left(\dot{\hat{q}} + \alpha \operatorname{grad}_{\hat{q}} F(\hat{q}, q)\right), \operatorname{grad}_{\hat{q}} F(\hat{q}, q)\right) \\
&\qquad \cdot \left(\dot{\hat{q}} + \alpha \operatorname{grad}_{\hat{q}} F(\hat{q}, q)\right)
\end{aligned}$$

that also reads

$$\begin{aligned}
\dot{\hat{q}} &= \hat{v} - \alpha \operatorname{grad}_{\hat{q}} F(\hat{q}, q) \\
\nabla_{\dot{\hat{q}}} \hat{v} &= \mathcal{T}_{//q\to\hat{q}} S(q, t) - \beta \operatorname{grad}_{\hat{q}} F(\hat{q}, q) \\
&\quad + R(\hat{v}, \operatorname{grad}_{\hat{q}} F(\hat{q}, q))\hat{v}
\end{aligned} \tag{2}$$

where the following holds true.

- α and β are positive design parameters.
- $F(\hat{q}, q)$ is half of the square of the geodesic distance between q and $\hat{q}$. This function is well defined and regular when $\hat{q}$ and q are close enough.

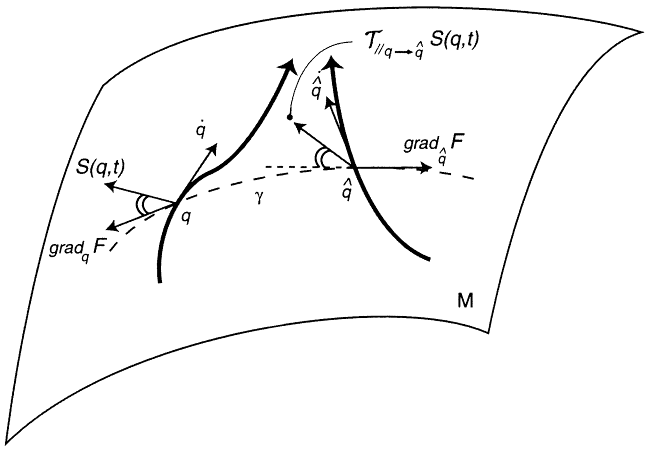

Fig. 1. Parallel transport of the source term $S(q,\, t)$ from the tangent space at q to the tangent space at $\hat{q}$, along the geodesic γ.

- $\mathcal{T}_{//q\to\hat{q}}$ is the parallel transport from q to $\hat{q}$ along the geodesic between q and $\hat{q}$. It is a linear isometry from the tangent space at q to the tangent space at $\hat{q}$. As for F, this operator is well defined for q and $\hat{q}$ close enough.
- R is the curvature tensor.

In local coordinates (q^i), the observer dynamics reads

$$\begin{aligned}\dot{\hat{q}}^i &= \hat{v}^i - \alpha\left\{\mathrm{grad}_{\hat{q}} F(\hat{q},\, q)\right\}^i\\ \dot{\hat{v}}^i &= -\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\hat{q}}^k + \left\{\mathcal{T}_{//q\to\hat{q}} S(q,\, t)\right.\\ &\quad \left. -\, \beta\,\mathrm{grad}_{\hat{q}} F(\hat{q},\, q) + R(\hat{v},\, \mathrm{grad}_{\hat{q}} F(\hat{q},\, q))\hat{v}\right\}^i.\end{aligned}$$

This observer does not depend on the choice of a particular set of coordinates for q: the connection ∇, the function F, the operator $\mathcal{T}_{//q\to\hat{q}}$ and the tensor R are intrinsic objects attached to the Riemannian structure on M.

Fig. 1 illustrates the observer dynamics (2). As the configuration space M has a Riemannian structure, we cannot compare vectors living in tangent spaces at different points on M, as it is usually done in Euclidian spaces. Indeed, we have to take into account the curvature introduced by the metric: the output injection term $\mathcal{T}_{//q\to\hat{q}} S(q,\, t)$ belongs to the tangent space $T_{\hat{q}}M$, whereas $S(q,\, t)$ belongs to T_qM and cannot be combined to $\nabla_{\dot{\hat{q}}}\hat{v} \in T_{\hat{q}}M$. We also replace the often used error term $(\hat{q}-q)$ by $\mathrm{grad}_{\hat{q}} F(\hat{q},\, q)$ to deal with the curvature since it gives the direction by which q can be "joined" from $\hat{q}$ by taking the shortest path. The term $-\beta\,\mathrm{grad}_{\hat{q}} F(\hat{q},\, q)$ can be interpreted as a spring term if we consider its counterpart in the Euclidian case as described in the introduction. The term $R(\hat{v},\, \mathrm{grad}_{\hat{q}} F(\hat{q},\, q))\hat{v}$ is also a spring term, with a stiffness quadratic in the velocity, that represents the minimum compensation term needed to eliminate the possible curvature instability effect (see [3]). In Fig. 1, we represent the operation of parallel transport along the geodesic γ that joins the system position q and the estimate position $\hat{q}$ on the manifold M. We can see for instance that the angle between $S(q,\, t)$ and $\mathrm{grad}_q F(\hat{q},\, q)$ is the same as that between their parallel transported counterparts $\mathcal{T}_{//q\to\hat{q}} S(q,\, t)$ and $\mathcal{T}_{//q\to\hat{q}}\,\mathrm{grad}_q F(\hat{q},\, q) = -\mathrm{grad}_{\hat{q}} F(\hat{q},\, q)$.

B. *Invariance on a Tutorial Example*

This is just to show that once the gains α and β are chosen, (2) defines a unique observer independent of the choice of a particular set of coordinates on the configuration manifold M.

1) Dynamics in q-Coordinate: We consider the one degree of freedom mechanical system whose Lagrangian is given by

$$\mathcal{L}(q,\, \dot{q}) = \tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}q^2$$

which represents the dynamics of the standard oscillator with $q \in \mathbb{R}$: $\ddot{q} = -q$. For this system, as the configuration space is Euclidian, the intrinsic observer equation (2) reads

$$\begin{aligned}\dot{\hat{q}} &= \hat{v} - \alpha(\hat{q} - q)\\ \dot{\hat{v}} &= -q - \beta(\hat{q} - q).\end{aligned} \tag{3}$$

If the gains α and β are chosen positive, we have convergence of $\hat{q}$ and $\hat{v}$ to q and $\dot{q}$.

2) Dynamics in r-Coordinate: Consider now a change of coordinate $r = \exp(q)$. The Lagrangian becomes

$$\mathcal{L}(r,\, \dot{r}) = \frac{1}{2}\frac{\dot{r}^2}{r^2} - \frac{1}{2}(\ln r)^2$$

and the system dynamics then writes

$$\begin{aligned}\dot{r} &= w\\ \dot{w} &= \frac{w^2}{r} - r\ln r \qquad \text{for } r \in]0,\, +\infty[.\end{aligned}$$

We are now going to compute the observer (2)

$$\begin{aligned}\dot{\hat{r}} &= \hat{w} - \alpha\,\mathrm{grad}_{\hat{r}} F(r,\, \hat{r})\\ \dot{\hat{w}} &= \frac{\hat{w}\dot{\hat{r}}}{\hat{r}} + \mathcal{T}_{//r\to\hat{r}}(-r\ln r) - \beta\,\mathrm{grad}_{\hat{r}} F(r,\, \hat{r})\\ &\quad + R(\hat{w},\, \mathrm{grad}_{\hat{r}} F(\hat{r},\, r))\hat{w}.\end{aligned} \tag{4}$$

The metric is given by $g = g_{11}$ with $g_{11}(r) = 1/r^2$. The Christoffel symbol is $\Gamma^1_{11}(r) = -(1/r)$. The equation for the geodesic joining r_1 and r_2 is

$$\gamma(s) = \exp\left(\ln\left(\frac{r_2}{r_1}\right)s + \ln r_1\right).$$

We have then $\gamma(0) = r_1$ and $\gamma(1) = r_2$. So, the geodesic distance between r_1 and r_2 is

$$d_G(r_1,\, r_2) = \int_0^1 \sqrt{g(\gamma(s))(\gamma'(s))^2}\, ds = |\ln r_2 - \ln r_1|$$

where $'$ stands for d/ds. The term $F(\hat{r},\, r)$ is then given by

$$F(r,\, \hat{r}) = \tfrac{1}{2}(\ln r - \ln\hat{r})^2$$

and its gradient by

$$\mathrm{grad}_{\hat{r}} F(r,\, \hat{r}) = \hat{r}^2\frac{(\ln\hat{r} - \ln r)}{\hat{r}} = \hat{r}(\ln\hat{r} - \ln r).$$

The parallel transport equation along the geodesic

$$s \mapsto \gamma(s) = \exp\left(\ln\left(\frac{\hat{r}}{r}\right)s + \ln r\right)$$

joining r at $s=0$ and $\hat{r}$ at $s=1$ reads

$$u' - \frac{1}{\gamma(s)}\gamma'(s)u = 0$$
$$u(0) = -r\ln r$$

for which the solution is given by

$$u(s) = -r\ln r \exp((\ln\hat{r} - \ln r)s).$$

Then, we have

$$\mathcal{T}_{//r\to\hat{r}}(-r\ln r) = u(1) = -\hat{r}\ln r.$$

So, (2) in the r coordinates gives

$$\dot{\hat{r}} = \hat{w} - \alpha\hat{r}(\ln\hat{r} - \ln r)$$
$$\dot{\hat{w}} = \frac{\hat{w}\dot{\hat{r}}}{\hat{r}} - \hat{r}\ln r - \beta\hat{r}(\ln\hat{r} - \ln r).$$

Notice that curvature is zero here. This vanishing is independent of the choice of configuration coordinates, whereas it is false for the Christoffel symbols.

In this set of coordinates, we see that this observer expression is not so intuitive: the error term $\hat{r}(\ln\hat{r} - \ln r)$ is nonlinear and is different from the often used error term $\hat{r} - r$. The convergence is clear since it can be checked that it is just the expression of (3) in r coordinates. When the metric g component is $g_{11}(q) = 1$, we have indeed

$$\mathrm{grad}_{\hat{q}} F(q, \hat{q}) = \hat{q} - q$$
$$\mathcal{T}_{//q\to\hat{q}} S(q, t) = \mathcal{T}_{//q\to\hat{q}}(-q) = -q$$
$$R(\hat{v}, \mathrm{grad}_{\hat{q}} F(\hat{q}, q))\hat{v} = 0.$$

So, the observer dynamics (3) and (4) are two expressions of the same observer, written in different configuration coordinate sets.

C. First-Order Approximation

In general, we have no explicit formula for F and $\mathcal{T}_{//q\to\hat{q}}$ once the metric is given. Nevertheless, the curvature terms are explicit

$$\left\{R(\hat{v}, \mathrm{grad}_{\hat{q}} F(\hat{q}, q))\hat{v}\right\}^i = R^i_{jkl}\hat{v}^k\{\mathrm{grad}_{\hat{q}} F(\hat{q}, q)\}^j\hat{v}^l$$

where R^i_{jkl} are the components of the curvature tensor

$$R^i_{jkl} = \frac{\partial\Gamma^i_{kl}}{\partial q^j} - \frac{\partial\Gamma^i_{jl}}{\partial q^k} + \Gamma^i_{pj}\Gamma^p_{kl} - \Gamma^i_{pk}\Gamma^p_{jl}.$$

However, for $\hat{q}$ close to q, F and $\mathcal{T}_{//q\to\hat{q}}$ admits the following approximations:

$$2F = g_{ij}(q)(\hat{q}^i - q^i)(\hat{q}^j - q^j) + O(\|\hat{q} - q\|^3)$$
$$\{\mathrm{grad}_{\hat{q}} F\}^i = \hat{q}^i - q^i + O(\|\hat{q} - q\|^2)$$
$$\{\mathcal{T}_{//q\to\hat{q}} w\}^i = w^i - \Gamma^i_{jl}(q)w^j(\hat{q}^l - q^l) + O(\|\hat{q} - q\|^2)$$

for any w belonging to the tangent space at q to M. The first equality comes from the definition of the geodesic distance. The second one is derived from the definition of the gradient for a scalar function. And the last one is derived from the expression of the parallel transport (see [32], [3] for more precisions). Remark that the "O-terms" will retain their forms when coordinates are changed in a differentiable manner.

Thus, we can construct an explicit approximation of (2) up to order 2. In local coordinates, this gives the following second-order approximate observer that can be integrated numerically:

$$\begin{aligned}\dot{\hat{q}}^i &= \hat{v}^i - \alpha(\hat{q}^i - q^i)\\ \dot{\hat{v}}^i &= -\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\hat{q}}^k + S^i(q, t) - \Gamma^i_{jl}(q)S^j(q, t)(\hat{q}^l - q^l)\\ &\quad - \beta(\hat{q}^i - q^i) + R^i_{jkl}(q)\hat{v}^k(\hat{q}^j - q^j)\hat{v}^l. \end{aligned} \tag{5}$$

In the term $\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\hat{q}}^k$, it is important to consider $\Gamma^i_{jk}(\hat{q})$ instead of $\Gamma^i_{jk}(q)$ since it is one of the terms of the covariant derivative of $\hat{v}$ with respect to $\dot{\hat{q}}$. Nevertheless in the terms $\Gamma^i_{jl}(q)S^j(q, t)(\hat{q}^l - q^l)$ and $R^i_{jkl}(q)\hat{v}^k(\hat{q}^j - q^j)\hat{v}^l$, we could have used $\Gamma^i_{jk}(\hat{q})$ and $R^i_{jkl}(\hat{q})$, since this represents a second order perturbation. The value of (5) relies on two facts

- the gains are explicit and can be computed via the inertia matrix (g_{ij}) and its q derivatives up to order 2;
- we will prove in the sequel the local convergence of (5) as soon as α and β are strictly positive.

IV. Observer Convergence

The observer dynamics (2) is locally ($\hat{q} \approx q$) contracting in the sense of [24], [18]: some insight on this property is given in Appendix II. As the system dynamics (1) is a solution of (2), this will give the local convergence.

More precisely, we are going to demonstrate the following result.

Theorem 1: Take (1) defining a dynamical system on the tangent bundle TM. Consider a compact subset K of TM and two positive parameters α and β (the observer gains).

Then, there exist $\varepsilon > 0$ (depending only on K, α and β), $\mu > 0$ and a Riemannian metric G on TM (depending only on α and β) such that, for any solution of (1) remaining in K

$$[0, T[\ni t \mapsto X(t) = (q(t), \dot{q}(t)) \in K$$

with $T \leq +\infty$, the solution of (2), $t \mapsto \hat{X}(t) = (\hat{q}(t), \hat{v}(t))$ with $\hat{X}(0)$ satisfying $d_G(\hat{X}(0), X(0)) \leq \varepsilon$, is defined for all $t \in [0, T[$ and, moreover, $\forall\, t \in [0, T[$

$$d_G\left(\hat{X}(t), X(t)\right) \leq d_G\left(\hat{X}(0), X(0)\right)\exp(-\mu t).$$

Here d_G is the geodesic distance associated to the metric G on TM.

The metric G is, in fact, a modified version of the Sasaki metric [30], i.e., the lift of the kinetic energy metric on TM. The observer gains α and β are involved in the definition of G in order to get the convergence estimation and the fact that, locally, the geodesic distance $d_G(\hat{X}(t), X(t))$ is a decreasing function of t.

Proof: The demonstration follows in two steps.

- For each q, we compute intrinsically (as for the second variation of geodesic) the first variation with respect to $\hat{q}$ and $\hat{v}$ of (2). If we denote by $(\hat{q} + \delta\hat{q}, \hat{v} + \delta\hat{v})$ a point defined in a neighborhood of $(\hat{q}, \hat{v})$, we are looking for the intrinsic formulation of the $(\delta\hat{q}, \delta\hat{v})$-dynamics.

- We will deduce from this intrinsic formulation, a metric G on the tangent bundle TM for which the observer dynamics (2) is a contraction for $\hat{q}$ close to q.

A. First Step: First Variation of the Observer Dynamics

We will just mimic here the method that has been used to derive the Jacobi equation [15], that is to say the first variation of the geodesic equation $\nabla_{\dot{q}}\dot{q}=0$. All the calculations presented here are done in a particular set of coordinates (q^i), but the final results are given in an intrinsic formulation. As the first variation $\delta\hat{v}$ is not an intrinsic term, we are going first to define its intrinsic equivalent ζ that belongs to the tangent space at $\hat{q}$ to M. Then, we are going to determine the dynamics of the intrinsic vectors $(\xi=\delta\hat{q},\,\zeta)$ associated to $(\delta\hat{q},\,\delta\hat{v})$.

1) Intrinsic Vectors ξ and ζ: We introduce another set of coordinates (an "intrinsic" one)

$$\begin{aligned}\xi^i &= \delta\hat{q}^i\\ \zeta^i &= \delta\hat{v}^i+\Gamma^i_{jk}(\hat{q})\delta\hat{q}^k\hat{v}^j, \qquad i=1,\,\ldots,\,n.\end{aligned}\tag{6}$$

One can check that the ξ^i and also the ζ^i correspond to the coordinates of two vectors ξ and ζ belonging to the same linear space, the tangent space at $\hat{q}$ to M. Indeed, we have up to second-order terms

- $\xi=\epsilon\dot{\gamma}(0)$, for some small real $\epsilon>0$ with γ a geodesic that joins the points $\gamma(0)=\hat{q}$ and $\gamma(\epsilon)=\hat{q}+\delta\hat{q}$;
- $\zeta=\mathcal{T}_{//\hat{q}+\delta\hat{q}\to\hat{q}}(\hat{v}+\delta\hat{v})-\hat{v}$.

The tangent vector ξ and ζ are defined along the curve $t\mapsto\hat{q}(t)$. Thus, we can consider their covariant derivatives $\nabla_{\dot{\hat{q}}}\xi$ and $\nabla_{\dot{\hat{q}}}\zeta$ still belonging to the tangent space at $\hat{q}$ to M.

Notice that the first variation of (2) gives

$$\begin{aligned}\frac{d}{dt}\left(\delta\hat{q}^i\right) &= \delta\hat{v}^i-\alpha\left\{\delta\left(\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\right)\right\}^i\\ \frac{d}{dt}\left(\delta\hat{v}^i\right) &= -\delta\left(\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\hat{q}}^k\right)\\ &\quad+\left\{\delta\left(\mathcal{T}_{//q\to\hat{q}}S(q,t)\right)-\beta\delta\big(\mathrm{grad}_{\hat{q}}F(\hat{q},q)\big)\right.\\ &\quad\left.+\,\delta\big(R(\hat{v},\mathrm{grad}_{\hat{q}}F(\hat{q},q))\hat{v}\big)\right\}^i.\end{aligned}\tag{7}$$

2) Computation of $\nabla_{\dot{\hat{q}}}\xi$: In local coordinates we have, for $i=1,\,\ldots,\,n$

$$\begin{aligned}\left\{\nabla_{\dot{\hat{q}}}\xi\right\}^i &= \dot{\xi}^i+\Gamma^i_{jk}(\hat{q})\dot{\hat{q}}^k\xi^j\\ &= \delta\hat{v}^i-\alpha\left\{\delta\left(\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\right)\right\}^i\\ &\quad+\Gamma^i_{jk}(\hat{q})\left(\hat{v}^k-\alpha\{\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\}^k\right)\xi^j\\ &= \delta\hat{v}^i+\Gamma^i_{jk}(\hat{q})\hat{v}^k\xi^j-\alpha\left(\left\{\delta\left(\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\right)\right\}^i\right.\\ &\quad\left.+\,\Gamma^i_{jk}(\hat{q})\{\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\}^k\xi^j\right).\end{aligned}$$

Then, we get

$$\boxed{\nabla_{\dot{\hat{q}}}\xi=\zeta-\alpha\nabla_{\xi}\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)}$$

with

$$\begin{aligned}&\left\{\nabla_{\xi}\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\right\}^i\\ &\quad=\left\{\delta\left(\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\right)\right\}^i+\Gamma^i_{jk}(\hat{q})\{\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\}^k\xi^j.\end{aligned}$$

3) Computation of $\nabla_{\dot{\hat{q}}}\zeta$: The covariant derivative of ζ is given by

$$\left\{\nabla_{\dot{\hat{q}}}\zeta\right\}^i=\dot{\zeta}^i+\Gamma^i_{jk}(\hat{q})\dot{\hat{q}}^k\zeta^j.\tag{8}$$

According to (7) and (6), we have

$$\begin{aligned}\dot{\zeta}^i &= \frac{d}{dt}\left(\delta\hat{v}^i\right)+\frac{d}{dt}\left(\Gamma^i_{jk}(\hat{q})\hat{v}^j\xi^k\right)\\ &= -\delta\Big(\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\hat{q}}^k\Big)+\left\{\delta\left(\mathcal{T}_{//q\to\hat{q}}S(q,t)-\beta\,\mathrm{grad}_{\hat{q}}F(\hat{q},q)\right.\right.\\ &\quad\left.\left.+\,R(\hat{v},\mathrm{grad}_{\hat{q}}F(\hat{q},q))\hat{v}\right)\right\}^i+\frac{d}{dt}\big(\Gamma^i_{jk}(\hat{q})\hat{v}^j\xi^k\big).\end{aligned}$$

So, when we consider only the terms of first order in $\delta\hat{q}$ and $\delta\hat{v}$

$$\begin{aligned}&\delta\left(\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\hat{q}}^k\right)\\ &\quad=\frac{\partial\Gamma^i_{jk}(\hat{q})}{\partial\hat{q}^c}\,\xi^c\hat{v}^j\dot{\hat{q}}^k+\Gamma^i_{jk}(\hat{q})\delta\hat{v}^j\dot{\hat{q}}^k+\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\xi}^k\\ &\quad=\frac{\partial\Gamma^i_{jk}(\hat{q})}{\partial\hat{q}^c}\,\xi^c\hat{v}^j\dot{\hat{q}}^k+\Gamma^i_{jk}(\hat{q})\zeta^j\dot{\hat{q}}^k-\Gamma^i_{jk}(\hat{q})\Gamma^j_{pc}(\hat{q})\hat{v}^p\xi^c\dot{\hat{q}}^k\\ &\qquad+\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\xi}^k\\ &\frac{d}{dt}\left(\Gamma^i_{jk}(\hat{q})\hat{v}^j\xi^k\right)\\ &\quad=\frac{\partial\Gamma^i_{jk}(\hat{q})}{\partial\hat{q}^c}\,\dot{\hat{q}}^c\xi^k\hat{v}^j+\Gamma^i_{jk}(\hat{q})\xi^k\dot{\hat{v}}^j+\Gamma^i_{jk}(\hat{q})\dot{\xi}^k\hat{v}^j\\ &\quad=\frac{\partial\Gamma^i_{jk}(\hat{q})}{\partial\hat{q}^c}\,\dot{\hat{q}}^c\xi^k\hat{v}^j-\Gamma^i_{jk}(\hat{q})\Gamma^j_{pc}(\hat{q})\xi^k\dot{\hat{q}}^c\hat{v}^p\\ &\qquad+\Gamma^i_{jk}(\hat{q})\xi^k\left\{\mathcal{T}_{//q\to\hat{q}}S(q,\,t)-\beta\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\right.\\ &\qquad\left.+\,R(\hat{v},\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q))\hat{v}\right\}^j+\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\xi}^k.\end{aligned}$$

Then, we put together the two previous equations

$$\begin{aligned}&\frac{d}{dt}\left(\Gamma^i_{jk}(\hat{q})\hat{v}^j\xi^k\right)-\delta\left(\Gamma^i_{jk}(\hat{q})\hat{v}^j\dot{\hat{q}}^k\right)\\ &\quad=-\Gamma^i_{jk}(\hat{q})\zeta^j\dot{\hat{q}}^k+\Gamma^i_{jk}(\hat{q})\xi^k\left\{\mathcal{T}_{//q\to\hat{q}}S(q,\,t)\right.\\ &\qquad\left.-\,\beta\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)+R(\hat{v},\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q))\hat{v}\right\}^j\\ &\qquad+\xi^k\dot{\hat{q}}^c\hat{v}^p\left(\frac{\partial\Gamma^i_{pk}(\hat{q})}{\partial\hat{q}^c}-\frac{\partial\Gamma^i_{pc}(\hat{q})}{\partial\hat{q}^k}+\Gamma^i_{jc}(\hat{q})\Gamma^j_{pk}(\hat{q})\right.\\ &\qquad\left.-\,\Gamma^i_{jk}(\hat{q})\Gamma^j_{pc}(\hat{q})\right).\end{aligned}$$

So, we get for $\dot{\zeta}^i$

$$\begin{aligned}\dot{\zeta}^i &= -\Gamma^i_{jk}(\hat{q})\zeta^j\dot{\hat{q}}^k+\left\{\delta\left(\mathcal{T}_{//q\to\hat{q}}S(q,\,t)\right.\right.\\ &\quad\left.\left.-\,\beta\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)+R(\hat{v},\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q))\hat{v}\right)\right\}^i\\ &\quad+\Gamma^i_{jk}(\hat{q})\xi^k\left\{\mathcal{T}_{//q\to\hat{q}}S(q,\,t)-\beta\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)\right.\\ &\quad\left.+\,R(\hat{v},\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q))\hat{v}\right\}^j-\xi^k\dot{\hat{q}}^c\hat{v}^pR^i_{kcp}(\hat{q})\end{aligned}$$

which gives the following "semi-intrinsic" expression

$$\begin{aligned}\nabla_{\dot{\hat{q}}}\zeta &= -R\left(\dot{\hat{q}},\,\xi\right)\hat{v}+\nabla_{\xi}\left(\mathcal{T}_{//q\to\hat{q}}S(q,\,t)\right)\\ &\quad-\beta\nabla_{\xi}\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q)+\text{“}\nabla_{\xi}\text{”}\left(R(\hat{v},\,\mathrm{grad}_{\hat{q}}F(\hat{q},\,q))\hat{v}\right)\end{aligned}$$

where

$$\begin{aligned}&\left\{\nabla_\xi\left(\mathcal{T}_{//q\to\hat q}S(q,t)\right)\right\}^i\\&\quad=\left\{\delta\left(\mathcal{T}_{//q\to\hat q}S(q,t)\right)\right\}^i+\Gamma^i_{jk}(\hat q)\xi^k\left\{\mathcal{T}_{//q\to\hat q}S(q,t)\right\}^j\end{aligned}$$

and

$$\begin{aligned}&\left\{\text{“}\nabla_\xi\text{”}\left(R(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v\right)\right\}^i\\&\quad=\left\{\delta\left(R(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v\right)\right\}^i\\&\qquad+\Gamma^i_{jk}(\hat q)\xi^k\left\{R(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v\right\}^j.\end{aligned}$$

For the curvature term, we use “∇_ξ” instead of ∇_ξ since this is not a true covariant derivation with respect to ξ: the vector $R(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v$ is only defined along the curve $t\mapsto\hat q(t)$. Thus, we cannot define properly its covariant derivative in a direction ξ that is not colinear to $\dot{\hat q}$.

Let us express now the “∇_ξ” term in an intrinsic way. Since

$$\begin{aligned}\delta\hat v^i&=\zeta^i-\Gamma^i_{jk}(\hat q)\delta\hat q^k\hat v^j\\\left\{\delta\left(\mathrm{grad}_{\hat q}F(\hat q,\,q)\right)\right\}^i&=\left\{\nabla_\xi\mathrm{grad}_{\hat q}F(\hat q,\,q)\right\}^i\\&\quad-\Gamma^i_{jk}(\hat q)\delta\hat q^k\{\mathrm{grad}_{\hat q}F(\hat q,\,q)\}^j\end{aligned}$$

standard computations provide the following intrinsic expression:

$$\begin{aligned}&\text{“}\nabla_\xi\text{”}\left(R(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v\right)\\&\quad=(\nabla_\xi R)(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v+R(\zeta,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v\\&\qquad+R(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\zeta+R(\hat v,\,\nabla_\xi\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v\end{aligned}$$

where $\nabla_\xi R$ is the covariant derivative of the curvature tensor along ξ.

Finally, we have the following intrinsic formula for the first variation of (2) with respect to $\hat q$ and $\hat v$:

$$\boxed{\begin{aligned}\nabla_{\dot{\hat q}}\xi&=\zeta-\alpha\nabla_\xi\,\mathrm{grad}_{\hat q}F(\hat q,\,q)\\\nabla_{\dot{\hat q}}\zeta&=-R\left(\dot{\hat q},\,\xi\right)\hat v+\nabla_\xi\left(\mathcal{T}_{//q\to\hat q}S(q,t)\right)\\&\quad-\beta\nabla_\xi\,\mathrm{grad}_{\hat q}F(\hat q,\,q)+(\nabla_\xi R)(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v\\&\quad+R(\zeta,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v+R(\hat v,\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\zeta\\&\quad+R(\hat v,\,\nabla_\xi\,\mathrm{grad}_{\hat q}F(\hat q,\,q))\hat v\end{aligned}}\tag{9}$$

where ξ corresponds to the variation of $\hat q$. When $S=0$, $F=0$ and when we set to zero the curvature terms in (2), we recover the classical Jacobi equation:

$$\frac{D^2\xi}{Dt^2}=-R\left(\dot{\hat q},\,\xi\right)\dot{\hat q}$$

where $D/Dt=\nabla_{\dot{\hat q}}$.

4) First Variation When $\hat q$ Is Close to q: Assume that $\hat q$ is close to q (we do not assume here that $\hat v$ is close to v). Then, the aforementioned first variation becomes much simpler since up to order 1 in $\hat q-q$

$$\mathrm{grad}_{\hat q}F(\hat q,\,q)=0,\quad\nabla_\xi\,\mathrm{grad}_{\hat q}F(\hat q,\,q)=\xi\quad\dot{\hat q}=\hat v.$$

Furthermore

$$\begin{aligned}&\left\{\nabla_\xi\left(\mathcal{T}_{//q\to\hat q}S(q,t)\right)\right\}^i\\&\quad=\left\{\delta\left(\mathcal{T}_{//q\to\hat q}S(q,t)\right)\right\}^i+\Gamma^i_{jk}(\hat q)\left\{\mathcal{T}_{//q\to\hat q}S(q,t)\right\}^k\xi^j.\end{aligned}$$

As

$$\left\{\mathcal{T}_{//q\to\hat q}S(q,t)\right\}^i=S(q,t)^i-\Gamma^i_{jk}(q)(\hat q^j-q^j)S(q,t)^k$$

for q close to $\hat q$, we then have

$$\begin{aligned}\left\{\delta\left(\mathcal{T}_{//q\to\hat q}S(q,t)\right)\right\}^i&=-\Gamma^i_{jk}(q)\xi^jS(q,t)^k\\\Gamma^i_{jk}(\hat q)\left\{\mathcal{T}_{//q\to\hat q}S(q,t)\right\}^k\xi^j&=\Gamma^i_{jk}(q)\xi^jS(q,t)^k\end{aligned}$$

up to terms of order 1 in $(\hat q-q)$. So

$$\nabla_\xi\left(\mathcal{T}_{//q\to\hat q}S(q,t)\right)=0.$$

Thus, for $q=\hat q$, (9) becomes

$$\boxed{\begin{aligned}\nabla_{\dot{\hat q}}\xi&=\zeta-\alpha\xi\\\nabla_{\dot{\hat q}}\zeta&=-\beta\xi.\end{aligned}}\tag{10}$$

In a certain sense, we recover the Euclidian case with the classical Luenberger observer described in the introduction. This is due to the cancellation of the curvature terms.

B. Second Step: Contraction Analysis

Let us prove first that (10) implies that the dynamics is strictly contracting when $\hat q=q$. Elementary continuity arguments show that contraction remains for $\hat q$ close to q. This explains the constraint on the initial condition for the observer dynamics. To speak of contraction, we first need to define a metric on the observer state space, i.e., on the tangent bundle TM.

1) Riemannian Structure on the Tangent Bundle TM: Since the observer gains α and β are positive, the matrix

$$A=\begin{pmatrix}-\alpha&1\\-\beta&0\end{pmatrix}$$

is Hurwitz and there exists a symmetric positive–definite matrix Q such that

$$A^tQ+QA=-I.$$

Set

$$Q=\begin{pmatrix}a&c\\c&b\end{pmatrix}$$

and consider the quantity

$$V(\xi,\,\zeta)=\frac{a}{2}\langle\xi,\,\xi\rangle+c\langle\xi,\,\zeta\rangle+\frac{b}{2}\langle\zeta,\,\zeta\rangle\tag{11}$$

where $\langle,\rangle$ is the scalar product associated to the metric g deduced from the kinetic energy. This quantity endows TM with a metric, since Q is positive definite: in local coordinates $(q^i,\,v^i)$, the length of the small vector $(\delta q^i,\,\delta v^i)$ tangent to $(q,\,v)$ at TM is

$$\begin{aligned}&V\left(\delta q,\,(\delta v^i+\Gamma^i_{kl}(q)v^k\delta q^l)_{i=1\cdots n}\right)\\&\quad=\frac{a}{2}g_{ij}\delta q^i\delta q^j+cg_{ij}(\delta v^i+\Gamma^i_{kl}(q)v^k\delta q^l)\delta q^j\\&\qquad+\frac{b}{2}g_{ij}(\delta v^i+\Gamma^i_{kl}(q)v^k\delta q^l)(\delta v^j+\Gamma^j_{kl}(q)v^k\delta q^l).\end{aligned}$$

In the local coordinates $(q^i,\,v^i)$, the metric is a $2n\times 2n$ matrix with entries function of q and v. Using notation $X=(q,\,v)$, we

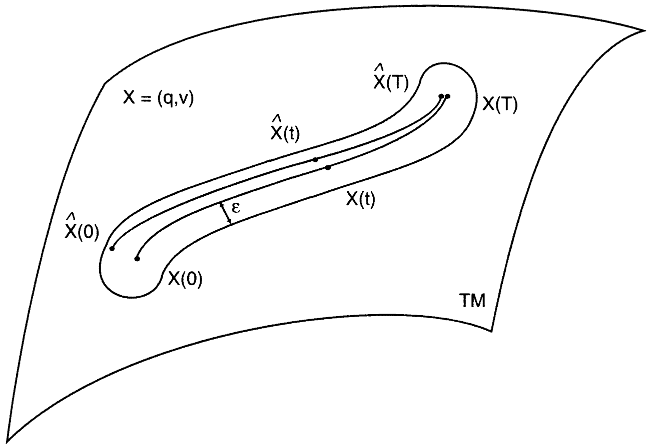

Fig. 2. Contraction tube.

denote by $G(X)$ the matrix defining this metric on TM. This is just a slightly modified version of the Sasaki metric on TM (see [30] and [31]); we get the Sasaki metric when $(a, b, c) = (2, 2, 0)$.

2) Convergence Analysis: When ξ and ζ satisfy (10), simple computations give

$$\begin{aligned}\frac{dV}{dt} &= a\langle\nabla_{\dot{\hat{q}}}\xi, \xi\rangle + c\langle\nabla_{\dot{\hat{q}}}\xi, \zeta\rangle + c\langle\nabla_{\dot{\hat{q}}}\zeta, \xi\rangle + b\langle\nabla_{\dot{\hat{q}}}\zeta, \zeta\rangle \\ &= -\langle\xi, \xi\rangle - \langle\zeta, \zeta\rangle.\end{aligned}$$

Thus, there exists $\lambda > 0$ such that

$$\frac{dV}{dt} \leq -\lambda V.$$

This means that the observer dynamics (2) is a strict contraction with respect to the metric $G(X)$ when $\hat{q} = q$ whatever $\hat{v}$ is.

Otherwise stated, denote by $\dot{\hat{X}} = \Upsilon(X, \hat{X})$ the observer (2). By construction $\dot{X} = \Upsilon(X, X)$ corresponds to the true dynamics (1). The inequality $dV/dt \leq -\lambda V$ just means that we have the following matrix inequality:

$$\begin{aligned}\left.\frac{\partial G}{\partial X}\right|_{\hat{X}} \Upsilon\left(X, \hat{X}\right) + \left(\left.\frac{\partial \Upsilon}{\partial \hat{X}}\right|_{(X,\hat{X})}\right)^t G\left(\hat{X}\right) \\ + G\left(\hat{X}\right)\left(\left.\frac{\partial \Upsilon}{\partial \hat{X}}\right|_{(X,\hat{X})}\right) \leq -\lambda G\left(\hat{X}\right) \quad (12)\end{aligned}$$

for $X = (q, v)$ and $\hat{X} = (q, \hat{v})$, q, v and $\hat{v}$ arbitrary. G is positive definite and the dependence of (12) versus X and $\hat{X}$ is smooth. Thus, for any $0 < \rho < \lambda$, there exists $\varepsilon > 0$ such that, for any X in the compact K and any $\hat{X}$ satisfying $d_G(X, \hat{X}) \leq \varepsilon$, we have

$$\begin{aligned}\left.\frac{\partial G}{\partial X}\right|_{\hat{X}} \Upsilon\left(X, \hat{X}\right) + \left(\left.\frac{\partial \Upsilon}{\partial \hat{X}}\right|_{(X,\hat{X})}\right)^t G\left(\hat{X}\right) \\ + G\left(\hat{X}\right)\left(\left.\frac{\partial \Upsilon}{\partial \hat{X}}\right|_{(X,\hat{X})}\right) \leq -\rho G\left(\hat{X}\right).\end{aligned}$$

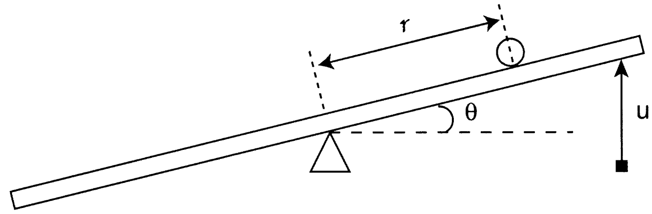

Fig. 3. Ball and beam system.

Assume that $\hat{X}(0)$ is close enough to $X(0)$, i.e., $d_G(X(0), \hat{X}(0)) \leq \varepsilon$. According to Appendix II, we have for $t > 0$ small

$$d_G\left(\hat{X}(t), X(t)\right) \leq d_G\left(\hat{X}(0), X(0)\right)\exp\left(-\frac{\rho}{2}t\right).$$

Thus, as displayed in Fig. 2, for any time $t \in [0, T[$, $d_G(\hat{X}(t), X(t)) \leq \varepsilon$, and $\hat{X}(t)$ remains in a region of contraction. Moreover, we have an exponential convergence with $\mu = \rho/2$. The proof of Theorem 1 is completed. ■

V. Conclusion

Simulations tests (see, e.g., the ball and beam example treated in Appendix I) tend to indicate that the region of convergence of our intrinsic observer (2) is quite large. This could be related to the fact that we have contraction when the estimated position $\hat{q}$ is close to the actual position q, even if the velocity estimation error is large. In our convergence analysis, we do not have fully exploited such nonlocal property. It appears that, combined with some additional structure, say, e.g., M is a Lie-group equipped with a right-invariant metric, one can prove stronger convergence results. Observer (2) is expressed without coordinates and thus could be extended, at least formally, to infinite dimensional mechanical systems such as a perfect incompressible fluid where the curvature tensor defined in [3] and [2] is explicitly given in [29].

Appendix I
Ball and Beam System

We have chosen the well-known ball and beam system [19] as an illustration since the scalar curvature of the metric given by its inertia matrix is strictly negative. The simulation results show then the interest of the invariant asymptotic observer: we can indeed choose small gains that reject noise while still cancelling the effects of the negative curvature.

A. System Dynamics

We consider a reduced ball and beam system, as shown in Fig. 3, with r the distance of the ball to the center of the beam, and θ the angle of the beam with the horizontal. A torque u is applied to control the system.

The kinetic energy is given by

$$T = \tfrac{1}{2}\left(\dot{r}^2 + (1+r^2)\dot{\theta}^2\right)$$

and the potential of the gravitation force by

$$U = r\sin\theta.$$

We get then the following normalized dynamics:

$$\begin{aligned}\dot{r} &= v_r\\ \dot{\theta} &= v_\theta\\ \dot{v}_r &= r{v_\theta}^2 - \sin\theta\\ \dot{v}_\theta &= \frac{-2r}{1+r^2}\, v_r v_\theta + \frac{u - r\cos\theta}{1+r^2}.\end{aligned} \tag{13}$$

B. Invariant Observer

1) Metric Elements: The matrix of components of the metric g defined by the kinetic energy in these coordinates is

$$\begin{pmatrix}1 & 0\\ 0 & 1+r^2\end{pmatrix}.$$

The nonzero Christoffel symbols are

$$\begin{aligned}\Gamma^1_{22} &= -r\\ \Gamma^2_{12} = \Gamma^2_{12} &= \frac{r}{1+r^2}.\end{aligned}$$

The nonzero components of the Riemannian curvature tensor are

$$\begin{aligned}R^1_{122} &= -\frac{1}{1+r^2}\\ R^1_{212} &= +\frac{1}{1+r^2}\\ R^2_{121} &= +\frac{1}{(1+r^2)^2}\\ R^2_{211} &= -\frac{1}{(1+r^2)^2}.\end{aligned}$$

The scalar curvature is then

$$Rs = g^{li} R^p_{pli} = \frac{-2}{(1+r^2)^2}.$$

The ball and beam system has a strictly negative scalar curvature.

2) Observer Expression: We consider the approximate intrinsic observer (5)

$$\begin{aligned}\dot{\hat{r}} &= \hat{v}_r - \alpha(\hat{r} - r)\\ \dot{\hat{\theta}} &= \hat{v}_\theta - \alpha(\hat{\theta} - \theta)\\ \dot{\hat{v}}_r &= \hat{r}\dot{\hat{\theta}}\hat{v}_\theta + \left(-\sin\theta + \hat{r}\left(\hat{\theta} - \theta\right)\frac{u - r\cos\theta}{1+r^2}\right) - \beta(\hat{r} - r)\\ &\quad + \left(\frac{-1}{1+\hat{r}^2}\,\hat{v}_r\hat{v}_\theta\left(\hat{\theta} - \theta\right) + \frac{1}{1+\hat{r}^2}\,\hat{v}_\theta^2(\hat{r} - r)\right)\\ \dot{\hat{v}}_\theta &= \frac{-\hat{r}}{1+\hat{r}^2}\left(\dot{\hat{r}}\hat{v}_\theta + \hat{v}_r\dot{\hat{\theta}}\right) + \left(\frac{u - r\cos\theta}{1+r^2} - \frac{\hat{r}}{1+\hat{r}^2}\right.\\ &\quad \left.\cdot\left((\hat{r} - r)\frac{u - r\cos\theta}{1+r^2} + \left(\hat{\theta} - \theta\right)(-\sin\theta)\right)\right)\\ &\quad - \beta\left(\hat{\theta} - \theta\right) + \left(\frac{1}{(1+\hat{r}^2)^2}\,\hat{v}_r^2\left(\hat{\theta} - \theta\right)\right.\\ &\quad \left. + \frac{-1}{(1+\hat{r}^2)^2}\,\hat{v}_r\hat{v}_\theta(\hat{r} - r)\right).\end{aligned} \tag{14}$$

C. Numerical Simulation

We have chosen for the simulation presented in the Fig. 4, a control that maintains the ball in oscillation near the unstable equilibrium point ($r = 0$, $\theta = 0$, $v_r = 0$, $v_\theta = 0$): $u = -20\theta + 101r$. As r remains small, the scalar curvature keeps a value close to -2. Furthermore, we have added high-frequency signals b_r and b_θ, respectively, to the measurements r and θ to simulate sensors imperfections and neglected high-frequency dynamics.

To show the importance of the parallel transport and the curvature compensation, we have compared the invariant observer (14), to the following one:

$$\begin{aligned}\dot{\hat{r}} &= \hat{v}_r - \alpha(\hat{r} - r)\\ \dot{\hat{\theta}} &= \hat{v}_\theta - \alpha\left(\hat{\theta} - \theta\right)\\ \dot{\hat{v}}_r &= \hat{r}\hat{v}_\theta^2 - \sin\theta - \beta(\hat{r} - r)\\ \dot{\hat{v}}_\theta &= \frac{-2\hat{r}}{1+\hat{r}^2}\,\hat{v}_r\hat{v}_\theta - \frac{r\cos\theta - u}{1+r^2} - \beta\left(\hat{\theta} - \theta\right).\end{aligned} \tag{15}$$

This observer is a standard one with nonlinear input injection for r and θ. It is proved to be convergent for large enough gain assuming bounded velocities. This observer is very efficient for low velocities where gyroscopic terms are not too big.

The initial conditions for the simulation are

	Real System	Observers (14) and (15)
r	0.02	0.021
v_r	0.05	0.525
θ	0	0.2
v_θ	0	2.

If the gains α and β are chosen large enough, the observers (14) and (15) are both convergent. Nevertheless, the high frequencies b_r and b_θ are not filtered.

For the simulation presented in Fig. 4, we have taken the following values for the gains:

$$\begin{aligned}\omega_o &= \tfrac{1}{2}\sqrt{2}\\ \alpha &= 2\omega_o\\ \beta &= \omega_o^2\end{aligned}$$

since in absolute value, the scalar curvature maximum is 2.

In Fig. 4, the pictures c) and d) are copies of the pictures a) and b), where the real system position r and θ are presented without the high frequency signals b_r and b_θ introduced by the sensors. We can see that the observer (15) does not converge: the parameter β is not large enough to compensate the effects of the negative curvature. However, the invariant observer (14) is convergent. It shows the importance of the curvature term, quadratic in velocities, in the observer expression.

APPENDIX II
CONTRACTION INTERPRETATION

The contraction [24], [18] for a system, with the dynamics $\dot{x} = f(x, t)$, can be understood as the exponential decay, with time, of the length of any segment of initial conditions transported by the flow.

Definition 1 (Strict Contraction): Let $\dot{x} = f(x, t)$ be a regular ($\mathcal{C}^1$ for instance) dynamical system defined on some smooth

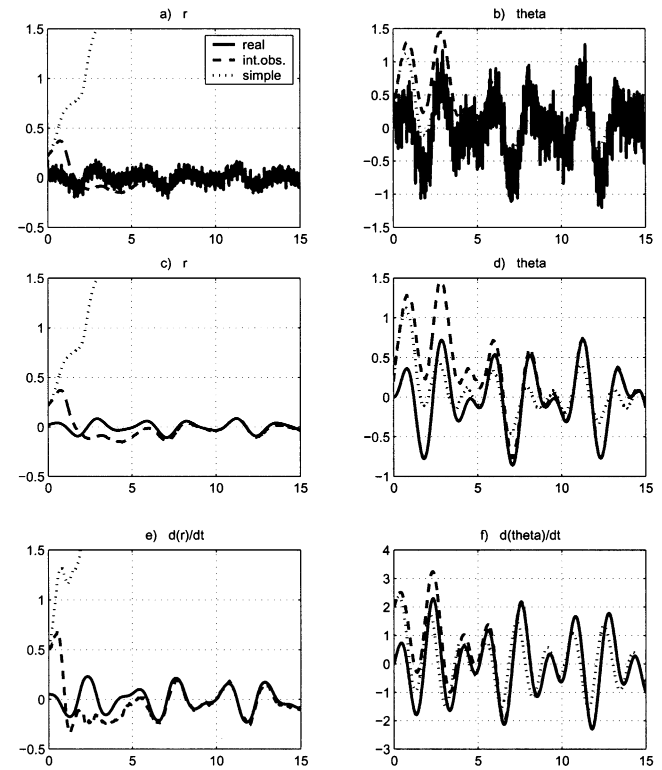

Fig. 4. Ball and beam observer simulation: Real = real system (13), int obs = intrinsic observer (14), simple = observer (15). a) r. b) θ. c) r. d) θ. e) $d(r)/dt$. f) $d(\theta)/dt$.

manifold M. Let g be a metric on M. Let $U \subset M$ be a set in M. The dynamics f is said to be a strict contraction in U with respect to the metric g, if the symmetric part of its Jacobian is negative definite, that is to say, if there exists some $\lambda > 0$ such that, in local coordinates x on U, we have for any t

$$\frac{\partial f}{\partial x}^T g(x) + g(x)\,\frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}\, f(x,\,t) \le -\lambda g(x).$$

We have the following result that justifies such definition and terminology.

Theorem 2: Let $\dot{x} = f(x,\,t)$ be a smooth dynamical system defined on a smooth manifold M. Let $g(x)$ be a metric on M. Let $X(x,\,t)$ be the flow associated to f

$$\begin{aligned} \frac{d}{dt} X(x,\,t) &= f(X(x,\,t),\,t) \qquad \forall\, t \in [0,\,T[\;\text{ with }\; T \le +\infty \\ X(x,\,0) &= x. \end{aligned}$$

Consider two points x_0 and x_1 in M and a geodesic $\gamma(s)$ joining $x_0 = \gamma(0)$ and $x_1 = \gamma(1)$. If

- f is a strict contraction on some subset $U \subset M$, with λ the constant defined in the definition 1;
- $X(\gamma(s),\,t)$ belongs to U for all $s \in [0,\,1]$; and for all $t \in [0,\,T[$

then

$$d_g(X(x_0,\,t),\,X(x_1,\,t)) \le e^{-(\lambda/2)t} d_g(x_0,\,x_1) \qquad \forall\, t \in [0,\,T[$$

where d_g is the geodesic distance associated to the metric g.

Proof: The proof is inspired by computations of L. Praly. Let $l(t)$ be the length of the curve $(X(\gamma(s),\,t),\, s \in [0,\,1])$ with respect to the metric g

$$l(t) = \int_0^1 \sqrt{\frac{dX(\gamma(s),\,t)}{ds}^T g(X(\gamma(s),\,t))\,\frac{dX(\gamma(s),\,t)}{ds}}\, ds.$$

We have

$$\frac{d}{dt}\, l(t) = \int_0^1 \frac{\frac{d}{dt}\left\{\frac{dX(\gamma(s),\,t)}{ds}^T g(X(\gamma(s),\,t))\,\frac{dX(\gamma(s),\,t)}{ds}\right\}}{2\sqrt{\frac{dX(\gamma(s),\,t)}{ds}^T g(X(\gamma(s),\,t))\,\frac{dX(\gamma(s),\,t)}{ds}}}\, ds.$$

As

$$\begin{aligned} \frac{d}{dt}\,\frac{dX(\gamma(s),\,t)}{ds} &= \frac{d}{ds}\, f(X(\gamma(s),\,t),\,t) \\ &= \frac{\partial}{\partial X}\, f(X(\gamma(s),\,t),\,t)\,\frac{d}{ds}\, X(\gamma(s),\,t) \end{aligned}$$

we get

$$\begin{aligned} &\frac{d}{dt}\left\{\frac{dX(\gamma(s),\,t)}{ds}^T g(X(\gamma(s),\,t))\,\frac{dX(\gamma(s),\,t)}{ds}\right\} \\ &\qquad = \frac{dX(\gamma(s),\,t)}{ds}^T P(s,\,t)\,\frac{dX(\gamma(s),\,t)}{ds} \end{aligned}$$

with

$$P(s,\,t) = \frac{\partial f(X)}{\partial X}^T g(X) + g(X)\,\frac{\partial f(X)}{\partial X} + \frac{\partial g(X)}{\partial X}\, f(X,\,t)$$

and

$$X = X(\gamma(s),\,t).$$

Since f is a contraction on U, there exists $\lambda > 0$ such that

$$P(s,\,t) \le -\lambda g(X(\gamma(s),\,t)).$$

We can then write the following inequality for the derivative $(d/dt) l(t)$:

$$\frac{d}{dt}\, l(t) \le -\frac{\lambda}{2}\, l(t)$$

which leads to

$$l(t) \le l(0) e^{-(\lambda/2)t} \qquad \forall\, t \in [0,\,T[.$$

Since $d_g(X(x_0,\,t),\,X(x_1,\,t)) \le l(t)$ and $l(0) = d_g(x_0,\,x_1)$ (indeed γ is a geodesic that joins the two points x_0 and x_1), the result is proved. ■

ACKNOWLEDGMENT

This work has benefited from the useful comments of L. Praly on stability issues and observer concepts. The authors would like to thank J. J. Slotine for interesting discussions on contraction theory. They are indebted to the fruitful suggestions and comments of N. Petit.

REFERENCES

[1] N. Aghannan and P. Rouchon, “On invariant asymptotic observers,” in *Conf. Decision Control*, vol. 2, Las Vegas, NV, 2002, pp. 1479–1484.

[2] V. I. Arnold, “Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l’hydrodynamique des fluides parfaits,” *Ann. Inst. Fourier*, vol. 16, pp. 319–361, 1966.

[3] ——, *Mathematical Methods of Classical Mechanics*, 2nd ed. New York: Springer-Verlag, 1989.
[4] G. Blankenstein, R. Ortega, and A. J. Van Der Schaft, "The matching conditions of controlled Lagrangians and interconnection and damping assignment passivity based control," *Int. J. Control*, vol. 75, no. 9, pp. 645–665, 2002.
[5] A. M. Bloch and P. E. Crouch, "Nonholonomic control systems on Riemannian manifolds," *SIAM J. Control Optim.*, vol. 37, no. 1, pp. 126–148, 1995.
[6] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, and R. Murray, "Nonholonomic mechanical systems with symmetry," *Arch. Ration. Mech. Anal.*, vol. 136, pp. 21–99, 1996.
[7] A. M. Bloch, N. E. Leonard, and J. E. Marsden, "Controlled Lagrangians and the stabilization of mechanical systems I: The first matching theorem," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 2253–2270, 2000.
[8] W. M. Boothby, *An Introduction to Differential Manifolds and Riemannian Geometry*, 2nd ed. New York: Academic, 1986.
[9] F. Bullo and R. M. Murray, "Tracking for fully actuated mechanical systems: A geometric framework," *Automatica*, vol. 35, no. 1, pp. 17–34, 1999.
[10] F. Bullo, N. E. Leonard, and A. D. Lewis, "Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 1437–1454, Aug. 2000.
[11] F. Bullo, "Series expansions for the evolution of mechanical control systems," *SIAM J. Control Optim.*, vol. 40, no. 1, pp. 166–190, 2001.
[12] F. Bullo and K. M. Lynch, "Kinematic controllability and decoupled trajectory planning for underactuated mechanical systems," *IEEE Trans. Robot. Automat.*, vol. 17, pp. 402–412, Apr. 2001.
[13] F. Bullo, "Averaging and vibrational control of mechanical systems," *SIAM J. Control Optim.*, vol. 35, no. 3, pp. 766–790, 1997.
[14] M. Perdigão do Carmo, *Riemannian Geometry*, 2nd ed. Boston, MA: Birkhäuser, 1992.
[15] B. Doubrovine, S. Novikov, and A. Fomenko, *Géométrie Contemporaine Méthodes et Applications*. Moscow, Russia: Mir, 1982.
[16] J. P. Gauthier and I. Kupka, *Deterministic Observation Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2001.
[17] H. Hammouri and J. P. Gauthier, "Global time-varying linearization up to output injection," *SIAM J. Control Optim.*, vol. 30, pp. 1295–1310, 1992.
[18] P. Hartman, *Ordinary Differential Equations*, 2nd ed. Boston, MA: Birkhäuser, 1982.
[19] N. H. Jo and J. H. Seo, "A state observer for nonlinear systems and its application to ball and beam system," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 968–373, May 2000.
[20] A. J. Krener and W. Respondek, "Nonlinear observers with linearizable error dynamics," *SIAM J. Control Optim.*, vol. 23, pp. 197–216, 1985.
[21] W. S. Koon and J. E. Marsden, "Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction," *SIAM J. Control Optim.*, vol. 35, pp. 901–929, 1997.
[22] F. Silva Leite, M. Camarinha, and P. E. Crouch, "Elastic curves as solutions of Riemannian and sub-Riemannian control problems," *Math. Control, Signals, Syst.*, vol. 13, no. 2, pp. 140–155, 2000.
[23] A. D. Lewis and R. M. Murray, "Configuration controllability of simple mechanical control systems," *SIAM J. Control Optim.*, vol. 35, no. 3, pp. 766–790, 1997.
[24] W. Lohmiller and J. J. E. Slotine, "On contraction analysis for nonlinear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.
[25] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*. New York: Springer-Verlag, 1994.
[26] F. Morgan, *Riemannian Geometry, a Beginner's Guide*: Jones and Bartlett Publishers, 1993.
[27] H. Nijmeijer and T. I. Fossen, *New Directions in Nonlinear Observer Design*. New York: Springer-Verlag, 1999.
[28] L. Noakes, G. Heinzinger, and B. Paden, "Cubic splines on curved spaces," *Inst. Math. Applicat.: J. Math. Control Inform.*, vol. 6, no. 4, pp. 465–473, 1989.
[29] P. Rouchon, "Jacobi equation, Riemannian curvature and the motion of a perfect incompressible fluid," *Eur. J. Mech. B Fluids*, vol. 11, pp. 317–336, 1992.
[30] S. Sasaki, "On the differential geometry of tangent bundles of Riemannian manifolds," *Tôhoku Math. J.*, ser. 2nd, vol. 10, pp. 338–354, 1958.
[31] ——, "On the differential geometry of tangent bundles of Riemannian manifolds II," *Tôhoku Math. J.*, ser. 2nd, vol. 14, pp. 146–155, 1962.
[32] J. J. Stoker, *Differential Geometry*. New York: Wiley, 1989.
[33] H. J. Sussmann, "Geometry and optimal control," in *Mathematical Control Theory*, J. Baillieul and J. C. Willems, Eds. New York: Springer-Verlag, October 1998, pp. 140–198.
[34] S. P. Weibel and J. Baillieul, "Averaging and energy methods for robust open-loop control of mechanical systems," in *Essays on Mathematical Robotics*. New York: Springer-Verlag, 1998, pp. 203–269.
[35] K. Yano and S. Ishihara, "Tangent and cotangent bundles," in *Pure and Applied Mathematics*. New York: Marcel Dekker, 1973.

Nasradine Aghannan was born in Nador, Morocco, in 1977. He graduated from École Polytechnique, Paris, France, in 2000. He is currently working toward the Ph.D. degree in mathematics and control at École des Mines de Paris, Paris, France.

His research interests are in the area of observer design for systems with symmetries and industrial process control. He has developed the controllers of several industrial chemical reactors, including polyethylene and polypropylene reactors.

Pierre Rouchon was born in Saint-Étienne, France, in 1960. He graduated from École Polytechnique, Paris, France, in 1983, and received the Ph.D. degree in chemical engineering from École des Mines de Paris, Paris, France, in 1990. In 2000, he obtained his "habilitation à diriger des recherches" in mathematics from the Université Paris-Sud Orsay, Paris, France.

Since 1993, he has been an Associate Professor of applied mathematics at École Polytechnique. From 1998 to 2002, he was the Head of the Centre Automatique et Systèmes, École des Mines de Paris. His fields of interest include the theory and applications of nonlinear control and, in particular, differential flatness and its extension to infinite-dimensional systems. He has worked on many industrial applications, such as distillation columns, electrical drives, car equipments, and chemical reactors.

Dr. Rouchon is an Associate Editor for *ESAIM-COCV*.