Partial Linearization of Mechanical Systems with Application to Observer Design

Sarras, Ioannis; Venkatraman, Aneesh; Ortega, Romeo; Schaft, Arjan van der

Partial Linearization of Mechanical Systems with Application to Observer Design

Ioannis Sarras, Aneesh Venkatraman, Romeo Ortega, and Arjan van der Schaft

Abstract—We consider general mechanical systems and establish a necessary and sufficient condition for the existence of a suitable change in the generalized momentum coordinates such that the new dynamics become linear in the transformed momenta. The class of systems which can be (partially) linearized by the proposed approach is characterized by (the solvability of) a set of partial differential equations and is shown to be larger than the class reported in all the previous works on linearization. We employ this linearization procedure to design an observer for mechanical systems where, we first (partially) linearize the system to make it affine in the new momenta and then construct a globally exponentially stable reduced order observer (which estimates the new momenta) by using the Immersion and Invariance approach.

I. Introduction

We consider n degree of freedom mechanical systems modeled in Hamiltonian form as

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \begin{pmatrix} \frac{\partial H}{\partial q} \\ \frac{\partial H}{\partial p} \end{pmatrix}, \qquad (1)$$

where $q \in \mathbb{R}^n$, $p \in \mathbb{R}^n$ are the generalized positions and momenta respectively. Further, the Hamiltonian function $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is the total energy of the system and is given as

$$H(q,p) = \frac{1}{2} p^\top M^{-1}(q) p + V(q). \qquad (2)$$

where $M = M^\top > 0$ is the mass matrix and V the potential energy function.

The problem of transforming the dynamical equations of a given mechanical system into a form that admits a relatively simple controller or an observer design, is of great practical interest and has henceforth been extensively studied in the literature. References [5], [12], [14], [18] for instance, consider mechanical systems and propose a coordinate transformation that renders the system linear. In the context of observer design, more particularly, the problem of velocity reconstruction has been treated exhaustively and many interesting solutions have been reported —we refer the reader to the recent books [2], [6] for a list of references.

The contributions of this paper are:

- Characterization, by a set of partial differential equations (PDEs), of the class of mechanical systems that can be rendered linear in the generalized momenta under a

I. Sarras and R. Ortega are with the Laboratoire des Signaux et Systèmes, Université Paris-Sud XI/Supelec - 3, rue Joliot-Curie, 91192 Gif-sur-Yvette cedex (France). Email: sarras@lss.supelec.fr, Romeo.Ortega@lss.supelec.fr

A. Venkatraman and A.van der Schaft are with the Institute for Mathematics and Computing Science, University of Groningen, P.O.Box 407, 9700 AK Groningen, The Netherlands. E-mail: aneesh@math.rug.nl ,A.J.van.der.Schaft@math.rug.nl

(partial) change of coordinates of the form $(q, P) = (q, \mathcal{T}^\top(q)p)$, with $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ full rank. We also prove that the results reported in the control literature on linearization are particular cases of our result and that the new characterization covers a large class of practical examples.

- Identification, in terms of two sets of PDEs (including the one mentioned previously), depending on M, of the class of systems for which we can construct a globally (exponentially) convergent reduced order observer for p.

The remaining of the paper is organized as follows. In Section II the main result on linearization via a partial change of coordinates is stated and proved. Section III is devoted to the system theoretic interpretation of some special choices of the matrix $\mathcal{T}$. In Section IV we present the observer design methodology and identify—in terms of two key assumptions that yield the two sets of PDEs—the class of systems for which we can generate a stable observer error dynamics. We wrap up the paper with some concluding remarks and future work in Section V.

II. A necessary and sufficient condition for partial linearization

Given a inertia matrix M, we introduce the following assumption.

Assumption 1: There exists a full rank matrix $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ such that the parameterized vector

$$\mathbf{D}_\mathcal{T}(q, P) = 0, \qquad (3)$$

where

$$\mathbf{D}_\mathcal{T} := \{P, P\}_\mathbb{J} \frac{\partial}{\partial P}(\frac{1}{2} P^\top \bar{M}^{-1} P) + \{P, q\}_\mathbb{J} \frac{\partial}{\partial q}(\frac{1}{2} P^\top \bar{M}^{-1} P) \qquad (4)$$

with $P = \mathcal{T}^\top(p)q$, $\bar{M}(q) = \mathcal{T}^\top(q)M(q)\mathcal{T}(q)$ and the quantity $\{.,.\}_\mathbb{J}$ is the matrix Poisson bracket of vector fields [5], [9] with respect to the standard skew symmetric matrix $\mathbb{J}$, being defined for any given $Q(q,p) : \mathbb{R}^{2n} \to \mathbb{R}^n$, $P(q,p) : \mathbb{R}^{2n} \to \mathbb{R}^n$ as,

$$\{Q, P\}_\mathbb{J} = \begin{bmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \end{bmatrix} \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^\top P}{\partial q} \\ \frac{\partial^\top P}{\partial p} \end{bmatrix}. \qquad (5)$$

We will show that Assumption 1 is precisely identifying the class of mechanical systems for which a change of coordinates of the form $(q, P) = (q, \mathcal{T}^\top(q)p)$ exists, that renders the system linear in the transformed momenta P.

Proposition 1: The dynamics of the system (1) expressed in the coordinates (q, P), where $P = \mathcal{T}^\top(q)p$, is linear in P if

and only if Assumption 1 holds, in which case, the dynamics becomes

$$
\begin{aligned}
\dot{q} &= M^{-1}\mathcal{T}^{-\top}P, \\
\dot{P} &= -\mathcal{T}^{\top}\frac{\partial V}{\partial q}.
\end{aligned}
\tag{6}
$$

Proof: The transformed system can be expressed in the form (refer to [5], [9]):

$$
\begin{pmatrix} \dot{q} \\ \dot{P} \end{pmatrix} = \begin{bmatrix} \{q,q\}_{\mathbb{J}} & \{q,P\}_{\mathbb{J}} \\ \{P,q\}_{\mathbb{J}} & \{P,P\}_{\mathbb{J}} \end{bmatrix} \begin{pmatrix} \frac{\partial \bar{H}}{\partial q} \\ \frac{\partial \bar{H}}{\partial P} \end{pmatrix}, \tag{7}
$$

with the new energy function being given as

$$
\bar{H}(q,P) := \frac{1}{2}P^{\top}\bar{M}^{-1}P + V(q),
$$

and the new inertia matrix $\bar{M}(q) := \mathcal{T}^{\top}(q)M(q)\mathcal{T}(q)$. Using the definition of the matrix Poisson bracket (5), we can compute

$$
\begin{aligned}
\{q,q\}_{\mathbb{J}} &= 0_{n\times n} \\[4pt]
\{q,P\}_{\mathbb{J}} &= \frac{\partial P}{\partial p}^{\top} \\[4pt]
\{P,q\}_{\mathbb{J}} &= -\frac{\partial P}{\partial p} \\[4pt]
\{P,P\}_{\mathbb{J}} &= -\frac{\partial P}{\partial p}\frac{\partial P}{\partial q}^{\top} + \frac{\partial P}{\partial q}\frac{\partial P}{\partial p}^{\top}
\end{aligned}
\tag{8}
$$

where, to obtain the third identity we have invoked the skew-symmetry property of the matrix Poisson bracket i.e. $\{P,q\}_{\mathbb{J}} = -\{q,P\}_{\mathbb{J}}^{\top}$. The dynamics of q follows trivially from the definition of P. By performing some simplifications, we get

$$
\begin{aligned}
\dot{P} &= \{P,q\}_{\mathbb{J}}\frac{\partial \bar{H}}{\partial q} + \{P,P\}_{\mathbb{J}}\frac{\partial \bar{H}}{\partial P} \\[4pt]
&= \mathbf{D}_{\mathcal{T}}(q,P) - \mathcal{T}^{\top}\frac{\partial V}{\partial q},
\end{aligned}
\tag{9}
$$

where we have used (4) and (8). We see from (9) that, if and only if Assumption 1 holds, that is if and only if $\mathbf{D}_{\mathcal{T}} = 0$, the dynamics is linear in P, and assumes the form (6). ∎

When Assumption 1 does not hold, the transformed dynamics in the coordinates (q, P) is given by

$$
\begin{pmatrix} \dot{q} \\ \dot{P} \end{pmatrix} = \begin{bmatrix} 0 & \{q,P\}_{\mathbb{J}} \\ \{P,q\}_{\mathbb{J}} & \{P,P\}_{\mathbb{J}} \end{bmatrix} \begin{pmatrix} \frac{\partial \bar{H}}{\partial q} \\ \frac{\partial \bar{H}}{\partial P} \end{pmatrix}. \tag{10}
$$

To streamline the presentation in the sequel we find it convenient at this point to recall the Lagrangian model of the mechanical system (1)

$$
M(q)\ddot{q} + C(q,\dot{q})\dot{q} + \frac{\partial V}{\partial q}(q) = 0, \tag{11}
$$

where $C(q,\dot{q})\dot{q}$ is the vector of Coriolis and centrifugal forces, with the ik–th element of the matrix $C : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ defined by

$$
C_{ik}(q,\dot{q}) = \sum_{j=1}^{n} C_{ij}^{k}(q)\dot{q}_j.
$$

The elements $C_{ij}^{k} : \mathbb{R}^n \to \mathbb{R}$ are the Christoffel symbols of the second kind, of the inertia matrix M, given by

$$
C_{ij}^{k}(q) := \frac{1}{2}\Big\{\frac{\partial M_{\mathrm{ik}}}{\partial q_{\mathrm{j}}} + \frac{\partial M_{\mathrm{jk}}}{\partial q_{\mathrm{i}}} - \frac{\partial M_{\mathrm{ij}}}{\partial q_{\mathrm{k}}}\Big\}, \quad \forall\, i,j,k \in \bar{n} := \{1,\dots,n\}, \tag{12}
$$

where M_{ij} is the ij–th element of M. It is well known that, for all vectors $x, y \in \mathbb{R}^n$, we have

$$
C(q,x)y = \begin{bmatrix} x^{\top}C^{1}(q)y \\ x^{\top}C^{2}(q)y \\ \vdots \\ x^{\top}C^{n}(q)y \end{bmatrix}
$$

where the ij element of the symmetric matrices $C^{k} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is precisely C_{ij}^{k}. We also recall the well–known fact that

$$
\frac{\partial}{\partial q}\Big\{\frac{1}{2}\dot{q}^{\top}M\dot{q}\Big\} = (C - \dot{M})\dot{q} \tag{13}
$$

See [13] for other properties of mechanical systems that are relevant in control applications.

III. SOME PARTICULAR CASES OF $\mathcal{T}$

In this section we discuss some particular selections of $\mathcal{T}$ (and hence P) that have some nice physical or geometrical interpretation and have been considered in the literature.

A. $\mathcal{T} = M^{-1}$: A Strong Condition for (Partial) Linearizability

With the choice $\mathcal{T} = M^{-1}$, we have

$$
\mathbf{D}_{M^{-1}} = \{P,P\}_{\mathbb{J}}MP + \{P,q\}_{\mathbb{J}}\,\frac{\partial}{\partial q}(\frac{1}{2}P^{\top}MP) \tag{14}
$$

and $P = \dot{q}$. Furthermore, (6) becomes

$$
\ddot{q} = -M^{-1}\frac{\partial V}{\partial q}. \tag{15}
$$

Unfortunately, Assumption 1 in this case is extremely restrictive as shown in the following proposition.

Proposition 2: Consider the parameterized vector $\mathbf{D}_{\mathcal{T}}$ introduced in Assumption (1). The following statements are equivalent:

(i) Assumption 1 holds with $\mathcal{T} = M^{-1}$, that is, $\mathbf{D}_{M^{-1}} = 0$.
(ii) The Christoffel symbols of the second kind of the inertia matrix M, (12), are all equal to zero.
(iii) The Coriolis and centrifugal forces $C(q,\dot{q})\dot{q}$ are equal to zero.

Proof: The proof follows directly comparing (11) with (15). ∎

B. $\mathcal{T}\mathcal{T}^{\top} = M^{-1}$: A Weaker Condition for (Partial) Linearizability

In this subsection we propose—as suggested in [7], [8]—to take $\mathcal{T}$ equal to a factor of M^{-1}. More precisely, we set

$$
M^{-1}(q) = T(q)T^{\top}(q), \tag{16}
$$

with $T : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ being a full rank matrix, and select $\mathcal{T} = T$.[1] Our motivation for this choice is threefold. Firstly,

---

[1]Since M is positive definite this factorization always exists, see, e.g., Corollary 7.2.9 of [10]. Further, T can be taken to be triangular with the diagonal terms strictly positive—the so–called, Cholesky factorization [11], which is uniquely defined. Also, notice that T is necessarily full rank—a fact that follows from rank $\{AB\} \leq$ min $\{$rank $\{A\}$, rank $\{B\}\}$.

Assumption 1 takes a particularly simple form that can be verified without solving any PDEs. Secondly, it leads to a relatively simple observer design. Thirdly, it allows us to establish some connections of our results with the existing literature on linearization and observer design.

In this case, the new energy function becomes

$$\bar{H}(q, P) = \frac{1}{2}|P|^2 + V(q),$$

and the dynamics (10) takes the form

$$\begin{pmatrix} \dot{q} \\ \dot{P} \end{pmatrix} = \begin{bmatrix} 0 & \{q, P\}_{\mathbb{J}} \\ \{P, q\}_{\mathbb{J}} & \{P, P\}_{\mathbb{J}} \end{bmatrix} \begin{pmatrix} \frac{\partial V}{\partial q} \\ P \end{pmatrix}. \qquad (17)$$

Further, if Assumption 1 holds, the transformed system becomes

$$\begin{aligned} \dot{q} &= TP \\ \dot{P} &= -T^{\top}\frac{\partial V}{\partial q}. \end{aligned}$$

Compare with (15). We now prove that Assumption 1 in this case is strictly weaker than the absence of Coriolis and centrifugal forces and, furthermore, has a nice geometric interpretation.

Proposition 3: Consider the factorization $M^{-1} = TT^{\top}$. Assumption 1 holds with $\mathcal{T} = T$ if and only if $\{P, P\}_{\mathbb{J}}P = 0$.

Proof: Upon substituting in the parameterized vector $\mathbf{D}_{\mathcal{T}}$, $\mathcal{T} = T$, it becomes

$$\mathbf{D}_T = \{P, P\}_{\mathbb{J}}P, \qquad (18)$$

from which the proof follows directly. ∎

Remark 1: A sufficient condition for $\{P, P\}_{\mathbb{J}}P = 0$, is clearly that, $\{P, P\}_{\mathbb{J}} = 0$ where, the (i,j) element of $\{P, P\}_{\mathbb{J}}$ is given by

$$\{P, P\}_{\mathbb{J}}^{i,j} = \{P_i, P_j\},$$

with $\{P_i, P_j\}$ being the standard poisson bracket [16] of smooth functions being given as,

$$\{P_i, P_j\} = \sum_{k=1}^{n}\left[\frac{\partial P_i}{\partial p_k}\frac{\partial P_j}{\partial q_k} - \frac{\partial P_i}{\partial q_k}\frac{\partial P_j}{\partial p_k}\right] \qquad (19)$$

where P_i, P_j are the ith and jth elements of the vector P.

Hence, we can see that, $\{P, P\}_{\mathbb{J}} = 0$ implies that for all $0 \leq i, j \leq n$, we have $\{P_i, P_j\} = 0$ — when it is said that the Poisson bracket commutes.

Poisson commutativity of two functions $P_i(q, p)$, $P_j(q, p)$ has an interesting geometrical interpretation. If $\{P_i, P_j\} = 0$, then the vectors P_i, P_j are said to be in involution, that is, each is constant along the integral curves of the other's Hamiltonian vector field which is in fact equivalent to the right hand side of (19) becoming equal to zero.

Further, if T_i, T_j are the ith and jth columns of the matrix T, then we have the relation (from differential geometry), $\{P_i, P_j\} = \{p^{\top}T_i, p^{\top}T_j\} = -p^{\top}[T_i, T_j]$, where $[T_i, T_j]$ is the standard Lie bracket [16] of vector fields defined as,

$$[T_i, T_j] = \frac{\partial T_j}{\partial q}T_i - \frac{\partial T_i}{\partial q}T_j.$$

We can thus see that $\{P, P\}_{\mathbb{J}} = 0$ if and only if the column vectors of the matrix T satisfy, $[T_i, T_j] = 0$ – when it is said that the columns of T commute among each other.

We show in the next subsection that (for $n \geq 3$) the condition $\{P, P\}_{\mathbb{J}} = 0$ is not necessary for $\{P, P\}_{\mathbb{J}}P = 0$ to hold.

Remark 2: Referring to the sufficient condition for linearizability we notice that selecting $\mathcal{T} = T$ and rewriting the dynamics in the Lagrangian form, we see that the Coriolis and centrifugal forces matrix should satisfy

$$C = T^{-\top}\frac{d}{dt}(T^{-1}).$$

Although the expression above relates the Coriolis matrix with the factor of M, the geometric interpretation is far from clear.

C. Proposition 3 is Strictly Weaker than Commutativity

The case when the columns of the matrix T (which satisfies (16)) commute has been extensively studied in analytical mechanics and has a deep geometric significance—stemming from Theorem 2.36 in [16]. It is widely accepted that this condition is quite restrictive and a natural question is whether Proposition 3 is strictly weaker than commutativity. In this subsection we show that this is indeed the case for $n \geq 3$.

Before presenting the result we find convenient to recall the following well–known fact of Riemannian geometry that has been used in the context of linearization, in the references [4], [18].

Fact 1: Given an inertia matrix M. The following statements are equivalent:

i) There exists a matrix T verifying (16) such that the vector $P = T^{\top}p$ satisfies $\{P, P\}_{\mathbb{J}} = 0$ or equivalently (as shown in the remark 1) that the columns of T satisfy, $[T_i, T_j] = 0$ for all $0 \leq i, j \leq n$.

ii) There exists a vector function $Q : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\frac{\partial Q}{\partial q} = T^{-1}(q). \qquad (20)$$

iii) The Riemann symbols (that can be computed directly from M with the formulas given in (22)) vanish identically.

If the conditions in the above fact are satisfied, then the system is said to be *Euclidean* [4], where the qualifier stems from the fact that the dynamics expressed in the coordinates (Q, P) reduces to a "linear double integrator" of the form

$$\begin{aligned} \dot{Q} &= P \\ \dot{P} &= -\frac{\partial \tilde{V}}{\partial Q}, \end{aligned}$$

where $\tilde{V}(Q) := V(Q^I(Q))$, with $Q^I : \mathbb{R}^n \to \mathbb{R}^n$ a left inverse of $Q(q)$, that is, $Q(Q^I(x)) = x$ for all $x \in \mathbb{R}^n$.

We next state the following interesting result.

Proposition 4: Given an inertia matrix M. The fact that there exists a factorization (16) such that $\mathbf{D}_T = 0$ *does not imply* that the system is Euclidean for $n \geq 3$. On the other hand, for $n \leq 2$ both conditions are equivalent.

Proof: First, we prove that for $n \leq 2$, the conditions $\{P, P\}_{\mathbb{J}}P = 0$ and $\{P, P\}_{\mathbb{J}} = 0$ are equivalent for any non

trivial vector P. For $n = 1$ the equivalence is, of course, obvious. For $n = 2$ this can be easily shown using the fact that $\mathbf{D}_T = \{P, P\}_{\mathbb{J}} P$ takes the form

$$\mathbf{D}_T = \begin{bmatrix} 0 & \{P_1, P_2\} \\ -\{P_1, P_2\} & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}.$$

We now construct an inertia matrix whose Riemann symbols are not all zero, but for which we can find a factorization that satisfies $\mathbf{D}_T = 0$. Towards this end, set $n = 3$ and consider

$$M^{-1} = \begin{bmatrix} 1 + q_2^2 & 0 & q_2\sqrt{1 + q_2^2} \\ 0 & (1 + q_2^2)^2 & 0 \\ q_2\sqrt{1 + q_2^2} & 0 & 1 + q_2^2 \end{bmatrix}. \quad (21)$$

We now compute the Riemann symbols, defined in page (4D-7) of [17] as

$$R_{ijlk} := \frac{1}{2}\left[\frac{\partial^2 M_{\mathrm{ik}}}{\partial q_{\mathrm{j}}\partial q_{\mathrm{l}}} + \frac{\partial^2 M_{\mathrm{jl}}}{\partial q_{\mathrm{i}}\partial q_{\mathrm{k}}} - \frac{\partial^2 M_{\mathrm{il}}}{\partial q_{\mathrm{j}}\partial q_{\mathrm{k}}} - \frac{\partial^2 M_{\mathrm{jk}}}{\partial q_{\mathrm{i}}\partial q_{\mathrm{l}}}\right]$$
$$+ \sum_{a,b=1}^{n} (M^{-1})_{ab}\left[C_{jl}^a C_{ik}^b - C_{il}^a C_{jk}^b\right] \quad (22)$$

where C_{ij}^k are the Christoffel symbols of the second kind as defined in (12) and $(M^{-1})_{\mathrm{ij}}$ is the ij-th element of the inertia matrix inverse. Notice that we only need to calculate $R_{1212}, R_{1213}, R_{1223}, R_{1313}, R_{1323}, R_{2323}$ because of he symmetries of the tensor. After some computations we verify that $R_{1212}, R_{1323}, R_{2323} \neq 0$ for all q and $R_{1223} \neq 0$ for $q_2 \neq 0$, and hence we conclude from Fact 1 that the system is not Euclidean.

On the other hand, it can be easily verified that the matrix M^{-1} admits a factorization (16) with

$$T = \begin{bmatrix} \sin(q_1)q_2 & \cos(q_1)q_2 & 1 \\ (1 + q_2^2)\cos(q_1) & -(1 + q_2^2)\sin(q_1) & 0 \\ \sqrt{1 + q_2^2}\sin(q_1) & \sqrt{1 + q_2^2}\cos(q_1) & 0 \end{bmatrix}. \quad (23)$$

Using (19), we compute

$$\{P_1, P_2\} = -P_3, \quad \{P_2, P_3\} = -P_1, \quad \{P_3, P_1\} = -P_2. \quad (24)$$

We conclude the proof by checking that $\mathbf{D}_T = \{P, P\}_{\mathbb{J}} P = 0$. ∎

An explanation regarding the construction of the example used in the proof above is in order. First, we observe that (24) is a sufficient condition for $\mathbf{D}_T = 0$. Condition (24) is satisfied by the vectors $T_i = A_i x$ where[2] $x \in \mathbb{R}^3$ and $A_i \in \mathbb{R}^{3\times 3}$ are the rotation matrices

$$A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

However, the resulting matrix $T = [T_1|T_2|T_3]$ has zero determinant, hence cannot qualify as a factor of M^{-1}.

[2]For reasons that will become clear below we find convenient to, temporarily, use the notation x instead of q.

To complete the example we must invoke some concepts from Lie group theory see, e.g., [17], [15]. The first observation is that the matrices A_i are tangent vectors at the identity point of the Lie group $SO(3)$ and, furthermore, form a basis for its associated Lie algebra $so(3)$. We then extend these vectors to left–invariant vector fields on the group $SO(3)$ using a push–forward of the left multiplication map $L_g(h) = gh$, where $g, h \in SO(3)$. The push–forward is defined as $(L_g)_*(A_i) = gA_i$, where g is taken to be the matrix

$$R(x) = \begin{bmatrix} \cos x_1 & \sin x_1 & 0 \\ -\sin x_1 & \cos x_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos x_2 & \sin x_2 \\ 0 & -\sin x_2 & \cos x_2 \end{bmatrix}$$
$$\begin{bmatrix} \cos x_3 & \sin x_3 & 0 \\ -\sin x_3 & \cos x_3 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which is a parametrization (using the Euler angles) of $SO(3)$. The question is then to find the vectors $\tilde{T}_i$, whose push–forward by R_*, that is $R_*(\tilde{T}_i)$, will equal $(L_R)_*(A_i)$. This leads to the following set of equations

$$\frac{\partial R}{\partial x_1}\tilde{T}_{i1}(x) + \frac{\partial R}{\partial x_2}\tilde{T}_{i2}(x) + \frac{\partial R}{\partial x_3}\tilde{T}_{i3}(x) = R(x)A_i, \quad i = 1, 2, 3.$$

Solving these equations we obtain the matrix $\tilde{T}$

$$\tilde{T}(x) = \begin{bmatrix} -\sin(x_1)\cot(x_2) & -\cos(x_1)\cot(x_2) & 1 \\ \cos(x_1) & -\sin(x_1) & 0 \\ \sin(x_1)\mathrm{cosec}(x_2) & \cos(x_1)\mathrm{cosec}(x_2) & 0 \end{bmatrix}.$$

Some simple computations show that the matrix $\tilde{T}$ has full rank (almost everywhere) and verifies (24) as desired.

The matrix $\tilde{T}$ above has a singularity at zero that can be easily "removed" introducing an homeomorphism $F : \mathbb{R} \times (0, \pi) \times \mathbb{R} \to \mathbb{R}^3 : x \mapsto q$. For instance, $F(x) = [x_1, \tan(x_2 - \frac{\pi}{2}), x_3]^\top$, which has an inverse map $F^I : \mathbb{R}^3 \to \mathbb{R}^3$, $F^I(q) = [q_1, \frac{\pi}{2} + \tan^{-1}(q_2), q_3]^\top$. We then define the transformed vectors,

$$T_{\mathrm{i}}(q) = [\frac{\partial F}{\partial x}(x)\tilde{T}_{\mathrm{i}}(x)]_{x=F^I(q)}, \ i = 1, 2, 3,$$

that, after some simple calculations, yields (23).

IV. IMMERSION AND INVARIANCE OBSERVERS

In this section, we use our main result on (partial) linearization to characterize a class of mechanical systems that admit a globally exponentially convergent, reduced order observer which estimates the (unmeasurable) momenta.

A. Problem Formulation and Proposed Approach

In this section we adopt the observer design framework proposed in [8], which follows the Immersion and Invariance (I&I) principles first articulated in [3]—see [2] for a tutorial account of this method and its applications. In the context of observer design the objective of I&I is to generate an attractive invariant manifold, defined in the extended state-space of the plant and the observer. This manifold is defined by an invertible function in such a way that the unmeasurable part of the state can be reconstructed by inversion of this

function. We thus introduce the definition of an I&I observer for the system (7), which is a particular case of the one given in [8], see also [19].

Definition 1: The dynamical system

$$\dot{\eta} = \alpha(q, \eta), \tag{25}$$

with $\eta \in \mathbb{R}^n$, is called an I&I observer for the system (7) if there exist a full rank matrix $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ and a vector function $\beta : \mathbb{R}^n \to \mathbb{R}^n$, such that the manifold

$$\mathcal{M} := \{(\eta, q, P) : \beta(q) = \eta + P\} \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \tag{26}$$

is positively invariant and attractive.[3] In this way, an asymptotic estimate of P, which we will denote by $\hat{P}$, is given by

$$\hat{P} = \beta - \eta.$$

Assumption 2: There exists a matrix $\mathcal{P} : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ satisfying the following two conditions:

(i)   The matrix inequality

$$\mathcal{A}(q) + \mathcal{A}^\top(q) \geq \epsilon I_n, \tag{27}$$

holds, uniformly in q, for some $\epsilon > 0$, where

$$\mathcal{A}(q) := \mathcal{P}(q)\mathcal{T}(q)\bar{M}^{-1}(q). \tag{28}$$

(ii)   The rows of $\mathcal{P}$, denoted $\mathcal{P}^j$, satisfy the integrability condition

$$\frac{\partial \mathcal{P}^j}{\partial q} = \left(\frac{\partial \mathcal{P}^j}{\partial q}\right)^\top, \quad j \in \bar{n}. \tag{29}$$

Assumption 1 defines a set of PDEs that has to be solved for the unknown $\mathcal{T}$. Similarly, for a given $\mathcal{T}$, the matrix $\mathcal{P}$ of Assumption 2 can be computed from the solution of the PDEs (29), subject to the inequality constraint (27).

B. I&I Observer

Proposition 5: If the matrices $\mathcal{T}$ and $\mathcal{P}$ satisfy Assumptions 1 and 2, the dynamical system

$$\dot{\eta} = \mathcal{P}\mathcal{T}\bar{M}^{-1}(\beta - \eta) + \mathcal{T}^\top \frac{\partial V}{\partial q}$$
$$\hat{P} = \beta - \eta \tag{30}$$

with

$$\frac{\partial \beta}{\partial q} = \mathcal{P}, \tag{31}$$

is a globally exponentially convergent reduced order observer for (7)—with the estimation error verifying

$$|\hat{P}(t) - P(t)|^2 \leq \exp^{-\epsilon t}|\hat{P}(0) - P(0)|^2,$$

where $|\cdot|$ is the Euclidean norm.

---

[3]We recall that the set $\mathcal{M}$ is positively invariant if $(\eta(0), q(0), P(0)) \in \mathcal{M} \Rightarrow (\eta(t), q(t), P(t)) \in \mathcal{M}$ for all $t \geq 0$. It is said to be globally attractive if, for all $(\eta(0), q(0), P(0))$, the distance of the state vector to the manifold asymptotically goes to zero, i.e., $\lim_{t \to \infty} \text{dist}\{(\eta(t), q(t), P(t)), \mathcal{M}\} = 0$.

Proof: Following the I&I procedure [2], to prove that the manifold $\mathcal{M}$, defined in (26), is attractive and invariant we show that the off–the–manifold coordinate

$$z = \beta - \eta - P, \tag{32}$$

that determines the distance from the state vector to the manifold $\mathcal{M}$, verifies: (i) $(z(0) = 0 \Rightarrow z(t) = 0)$ for all $t \geq 0$, and (ii) $z(t)$ asymptotically (actually, exponentially) converges to zero.

We first differentiate (32) to obtain the dynamics of z as

$$
\begin{aligned}
\dot{z} &= \dot{\beta} - \dot{\eta} - \dot{P} \\
&= \frac{\partial \beta}{\partial q}\{q, P\}_{\mathbb{J}}\frac{\partial \bar{H}}{\partial P} - [\frac{\partial \beta}{\partial q}\mathcal{T}\bar{M}^{-1}(\beta - \eta) + \mathcal{T}^\top \frac{\partial V}{\partial q}] \\
&\quad -\{P, q\}_{\mathbb{J}}\frac{\partial \bar{H}}{\partial q} - \{P, P\}_{\mathbb{J}}\frac{\partial \bar{H}}{\partial P} \\
&= -\frac{\partial \beta}{\partial q}\mathcal{T}\bar{M}^{-1}z - \mathbf{D}_{\mathcal{T}}(q, P)
\end{aligned}
$$

where we have used (7), (9), (30), and the fact $\{P, q\}_{\mathbb{J}} = \mathcal{T}^\top(q)$.

Upon invoking Assumption 1, we get $\mathbf{D}_{\mathcal{T}} = 0$. Now, with the integrability condition (29) of Assumption 2 and (31), the error dynamics reduces to,

$$\dot{z} = -\mathcal{P}\mathcal{T}\bar{M}^{-1}z. \tag{33}$$

The manifold $\mathcal{M}$ is clearly positively invariant. To establish global exponential attractivity of $\mathcal{M}$ consider the Lyapunov function $V(z) = \frac{1}{2}|z|^2$. Condition (27) ensures that $\dot{V} \leq -\epsilon V$, which proves the global exponential convergence to zero of z, hence of $\hat{P} - P$—with exponential rate ϵ. ∎

Remark 3: The manifold $\mathcal{M}$ proposed above is a particular case of the one considered in [8], where it is defined as $\{(\eta, q, P) : \tilde{\beta}(q, \eta) = P\}$, with $\tilde{\beta} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$—notice that we have taken $\tilde{\beta}$ to be a linear function of η in (26). It is clear that considering the more general manifold expression it is possible—in principle—to handle a larger class of systems. However, for the purposes of our work, which is to explicitly define (in terms of PDEs) a class of inertia matrices for which the construction works, this is done without loss of generality—see Remark 4.

Remark 4: Some connections between our observer and the one proposed in [8] may be established at this point. Towards this end we refer to the function $\tilde{\mathbf{D}}_{\mathcal{T}}(q, \dot{q}) := \mathbf{D}_{\mathcal{T}}(q, p)$ which can be written (by performing some simplifications) as

$$
\begin{aligned}
\tilde{\mathbf{D}}_{\mathcal{T}} &= \mathcal{T}^\top \frac{\partial}{\partial q}\{\frac{1}{2}\dot{q}^\mathrm{T} M \dot{q}\} - \dot{\mathcal{T}}^\top M \dot{q}, \\
&= [\mathcal{T}^\top C - \frac{d}{dt}(\mathcal{T}^\top M)]\dot{q}, \tag{34}
\end{aligned}
$$

where, to obtain the second identity, we have used (13). Evaluating it for $\mathcal{T} = T$ we obtain

$$\tilde{\mathbf{D}}_T = [T^\top C - \frac{d}{dt}T^{-1}]\dot{q} =: \bar{C}(q, \dot{q})\dot{q}.$$

It can easily be shown that the matrix $\bar{C}(q, \dot{q})T(q)$ is linear in $\dot{q}$ and furthermore, exploiting the fact that $\dot{M} = C + C^\top$, we can also prove that it is skew–symmetric. These properties are used in [8] to, adding to the observer a "certainty–equivalent"

term $\bar{C}(q, T\tilde{\beta})T\tilde{\beta}$, generate an error dynamics of the form $\dot{z} = [\Gamma(q, \eta) - \bar{C}(q, \dot{q})T(q)]z$, where $\Gamma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n \times n}$ is a matrix that can be shaped selecting the function $\tilde{\beta}$. A constructive solution is given for some particular cases of systems with $n = 2$, namely: diagonal inertia matrix and inertia matrix with bounded elements. However, the extension of the proposed techniques beyond these cases does not seem to be immediate. Of course, it may be argued that the route taken in the present paper (that aims at eliminating the term $\mathbf{D}_{\mathcal{T}}$), although leading to the explicit identification of some PDEs to be solved, is also not constructive—given our inability to guarantee their solution in general.

Remark 5: The proposed linearization technique and the observer design has been successfully applied to some well known physical examples like the inverted pendulum on a cart, the 3-link planar manipulator, the planar redundant manipulator with one elastic degree of freedom. The Assumption 1, for these examples, was satisfied by choosing $\mathcal{T}$ equal to a lower triangular cholesky factorization T of M^{-1}, and further, the observer design was performed by following a constructive procedure [1].

V. CONCLUDING REMARKS AND FUTURE WORK

We have identified a class of mechanical systems, characterized by (the solvability of) a set of PDEs, that contains all systems that can be rendered linear in (the unmeasurable) momenta via a (partial) change of coordinates $P = \mathcal{T}(q)p$. We have shown that this class is larger than the one reported in the literature of linearization and observer design. We have also shown that the class admits a globally exponentially stable reduced order observer.

Several open questions are currently under investigation:

- Similar to the well–known characterization of Euclidean systems in terms of the Riemann symbols (as told in fact 1), it would be interesting to derive necessary and sufficient conditions on M to verify the condition $\{P, P\}_{\mathbb{J}} P = 0$.
- In Remark 4 we have explained the difference between our approach to observer design and the one used used in [8]. Namely, the incorporation of the term $\bar{C}(q, \dot{q})$ in the observer, which is absent in our design. Some preliminary calculations show that, as expected, adding this term modifies the perturbing term $\mathbf{D}_T$ leading to alternative conditions for it to be zero.
- It is possible to show that the condition of Proposition 3, using the Cholesky factorization, is not verified for manipulators with more than one rotational joint. However, it is not clear whether other factorizations may exist of and whether they can be handled imposing the weaker Assumption 1.
- The solvability of the PDEs arising in Assumption 1 is a widely open question. These PDEs are, in general, nonlinear and quite involved. For instance, for the classical two-dimensional Ball–and–Beam example, they take the

form

$$
\begin{aligned}
\frac{\partial T_{11}}{\partial q_1} &= 0 \\
\frac{m_1 q_1 T_{11}}{1 + m_1 q_1^2} + \frac{\partial T_{21}}{\partial q_2} &= 0 \\
\frac{m_2}{1 + m_1 q_1^2} \frac{\partial T_{11}}{\partial q_2} + \frac{\partial T_{21}}{\partial q_1} &= 0 \\
\frac{\partial T_{12}}{\partial q_1} &= 0 \\
\frac{m_1 q_1 T_{12}}{1 + m_1 q_1^2} + \frac{\partial T_{22}}{\partial q_2} &= 0 \\
\frac{m_2}{1 + m_1 q_1^2} \frac{\partial T_{12}}{\partial q_2} + \frac{\partial T_{22}}{\partial q_1} &= 0.
\end{aligned}
$$

It is possible to show that these equations do not admit an explicit solution in separable variables form $T_{ij}(q) = a_{ij}(q_1) b_{ij}(q_2)$.

REFERENCES

[1] A. Venkatraman, R. Ortega, I. Sarras, A. J. van der Schaft, "Control of Underactuated Mechanical Systems: Observer Design and Position Feedback Stabilization", *Internal Report of Laboratoire des Signaux et Systèmes*, 2008.

[2] A. Astolfi, D. Karagiannis, R. Ortega, *Nonlinear and Adaptive Control with Applications*, Springer-Verlag, Berlin, Communications and Control Engineering, 2007.

[3] A. Astolfi, R. Ortega, "Immersion and Invariance: A new tool for stabilization and adaptive control of nonlinear systems", *IEEE Trans. Automat. Contr.*, Vol. 48, No. 4, pp. 590-606, April 2003.

[4] N. S. Bedrossian, "Linearizing coordinate transformation and Riemannian curvature", *Proc. IEEE Conf. Dec. & Contr.*, Tuscon, Arizona, pp. 80-85, December 1992.

[5] N. S. Bedrossian, *Nonlinear Control Using Linearizing Transformations*, Ph.D. Thesis, Massachusetts Institute of Technology, 1991.

[6] G. Besançon (Ed.), *Nonlinear Observers and Applications*, Lecture Notes in Control and Information Science 363, Springer Verlag, 2007

[7] G. Besançon, "Global output feedback tracking control for a class of Lagrangian systems", *Automatica*, Vol. 36, pp. 1915-1921, April 2000.

[8] D. Carnevale, D. Karagiannis, A. Astolfi, "Reduced-order observer design for nonlinear systems", *Proc. European Contr. Conf.*, Kos, Greece, pp. 559-564, July 2007.

[9] H. Goldstein, C. P. Poole, J. L. Safko, *Classical Mechanics*, 3rd edition, Addison Wesley, 2002.

[10] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.

[11] A. Jain and G. Rodriguez, "Diagonalized Lagrangian Robot Dynamics," *Proc. IEEE Int. Conf. Robot. & Autom.*, Vol. 11, no. 4, pp. 571-583, 1995.

[12] B. Jakubczyk and W. Respondek, "On linearization of control systems", *Bull. Acad. Pol. Sci. Ser. Sci. Math.*, 28:517522, 1980.

[13] R. Kelly, V. Santibanez, A. Loria, *Control of Robot Manipulators in Joint Space*, Springer, London, 2005.

[14] D.E. Koditchek, "Robot Kinematics and Coordinate Transformations", *Proc. IEEE Conf. Dec. & Contr.*, Ft. Lauderdale FL, pp. 1-4, 1985.

[15] J. M. Lee, *Introduction to Smooth Manifolds*, GTM 218, Springer, 2003.

[16] H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer, London, 1990.

[17] M. Spivak, *A comprehensive Introduction to Differential Geometry*, Vol. 2, Publish or Perish, Inc, 3rd ed edition, 1999.

[18] M. W. Spong, "Remarks on robot dynamics: canonical transformations and Riemannian geometry", *Proc. IEEE Int. Conf. Robot. & Autom.*, Nice, France, pp. 554-559, May 1992.

[19] J. Tsinias, "Further results on the observer design problem," *Systems & Control Letters*, Vol. 14, no. 5, pp. 411-418, 1990.