Tracking for Fully Actuated Mechanical Systems: A Geometric Framework*

Francesco Bullo† and Richard M. Murray†
Control and Dynamical Systems
California Institute of Technology
Pasadena, CA 91125
bullo@indra.caltech.edu murray@indra.caltech.edu

CIT–CDS Technical Report 97–003,
February 10, 1997
Revised October 23, 1997

Abstract

We present a general framework for the control of Lagrangian systems with as many inputs as degrees of freedom. Relying on the geometry of mechanical systems on manifolds, we propose a design algorithm for the tracking problem. The notions of error function and transport map lead to a proper definition of configuration and velocity error. These are the crucial ingredients in designing a proportional derivative feedback and feedforward controller. The proposed approach includes as special cases a variety of results on control of manipulators, pointing devices and autonomous vehicles. Our design provides particular insight into both aerospace and underwater applications where the configuration manifold is a Lie group.

Keywords: mechanical systems, nonlinear control, Lie groups, Riemannian geometry

Contents

*Modified versions of this paper have been submitted to *Automatica* and to the European Control Conference, Brussels 1997.

†Research supported in part by the National Science Foundation under grant CMS-9502224.

1 Introduction

Mechanical control systems provide an important and challenging research area that falls between the study of classical mechanics and modern nonlinear control theory. From a theoretical standpoint, the geometric structure of mechanical systems gives way to stronger control algorithms than those obtained for generic nonlinear systems. Recent promising results in this area are surveyed in (Murray 1995). The driving applications are motion control problems arising in the study of underwater and aerospace autonomous vehicles. In these environments, relevant Lagrangian models are available and a sharp nonlinear analysis can successfully exploit this structure.

This paper deals with the trajectory tracking problem for a class of Lagrangian systems. The control objective is to track a trajectory with exponential convergence rates in order to guarantee performance and robustness. The mechanical systems we consider have Lagrangian equal to the kinetic energy and are fully actuated, that is, they have as many independent input forces as degrees of freedom. A wide variety of aerospace and underwater vehicles, as well as robot manipulators, fulfill these assumptions. The main emphasis is on the fact that the configuration space of these systems is a generic manifold. In particular, the group of rotations $SO(3)$ and the group of rigid rotations and translations $SE(3)$ are commonly encountered examples.

The tracking problem for robot manipulators has received much attention in the literature. Examples are the contributions in (Takegaki & Arimoto 1981), (Wen & Bayard 1988) and (Slotine & Li 1989), where asymptotic, exponential and adaptive tracking are achieved via a nonlinear analysis. These results are now standard in textbooks on control (Nijmeijer & van der Schaft 1990) and robotics (Murray, Li & Sastry 1994). Since then, similar techniques have been applied to the attitude control problem for satellites (Wen &

Kreutz-Delgado 1991), and likewise to the attitude and position control for underwater vehicles (Fossen 1994, Section 4.5.4). A further example is the spin axis stabilization problem for satellites (Tsiotras & Longuski 1994). A common feature in all these works is the preliminary choice of a parametrization, i.e. a choice of coordinates for the configuration manifold. The synthesis of both control law and corresponding Lyapunov function is performed in this specific parametrization. This set of coordinate plays then an important role, when the control system is characterized in terms of, for example, singularities and exponential convergence, and when adaptive capabilities are included.

In this paper we propose a unifying framework that applies to a large class of mechanical systems. In the spirit of (Koditschek 1989), this is achieved by avoiding the parametrization step. Our design algorithm focuses on basic, intrinsic issues such as how to define a state error and how to exploit the Lagrangian dynamics. The notions of "error function" and "transport map" yield to a coordinate-free definition of errors between configurations and between velocities. Together with a dissipation function these ingredients determine the feedback law. The feedforward control is devised using the theory of Riemannian connections. Provided a compatibility condition between error function and transport map holds, our control strategy achieves globally stable tracking. As discussed in (Koditschek 1989), (possible) topological properties of the configuration manifold preclude global asymptotic stabilization. However, we prove local exponential stability under some boundedness conditions and we provide an estimate of the region of attraction. Useful extensions to adaptive control and to more general mechanical systems can be included via standard techniques. We remark that that the design process, the statement and the proof of the main theorem are all performed without choosing coordinates the configuration manifold.

The resulting design algorithm is then set to work in a variety of applications, recovering previous controllers and suggesting new ones. Examples are the standard "augmented PD control" for robot manipulators, see (Murray et al. 1994) , and the novel tracking controller for systems on the two sphere. Most instructive is the treatment of the tracking problem on the group of rigid rotations $SO(3)$ and on the group of rigid motions $SE(3)$. In the latter case for example, we design a large set of error functions with matrix gains and we characterize transport maps as changes of reference frame. These ideas lead to a comparison of various previous approaches and to new results. Finally, some computationally simple feedforward controls are derived via an extension of the main theorem.

The paper is organized as follows. Section 2 reviews some required tools from Riemannian geometry and some concepts from mechanical control systems. Section 3 introduces the notions of error function and transport map. The two sphere example illustrates these concepts and the section ends with an additional study of the transport map. All these ideas lead to the main theorem, with proof and comments, in Section 4. Many examples and applications of the main result are finally discussed in Section 5. A preliminary version of this paper appeared at the European Control Conference, Brussels 1997 (Bullo & Murray 1997).

2 Mathematical Preliminaries

In this section we introduce the mathematical machinery needed for the remainder of the paper. For an introduction to Riemannian geometry we refer to (Boothby 1986), (Do Carmo 1992) and (Kobayashi & Nomizu 1963). For an introduction to mechanics we refer to (Arnold 1989) and (Marsden & Ratiu 1994).

2.1 Elements of Riemannian geometry

A *Riemannian metric* on a manifold Q is a smooth map that associates to each tangent space T_qQ an inner product $\langle\langle \cdot\,,\,\cdot\rangle\rangle_q$. Given a pair of smooth vector fields X, Y, we let $[X, Y]$ denote their Lie bracket and equivalently $\mathcal{L}_X Y$ denote the Lie derivative of Y with respect to X. An *affine connection* on Q is a smooth map that assigns to each pair of smooth vector fields X, Y a smooth vector field $\nabla_X Y$ such that for all functions f on Q

1. $\nabla_{fX} Y = f\nabla_X Y,$ and
2. $\nabla_X fY = f\nabla_X Y + (\mathcal{L}_X f)\, Y$

where $\mathcal{L}_X f$ denotes the Lie derivative of f with respect to X. In system of local coordinates $(q^1, \ldots, q^n)$ we define the *Christoffel symbols* Γ^k_{ij} by

$$\nabla_{\frac{\partial}{\partial q^i}}\left(\frac{\partial}{\partial q^j}\right) = \Gamma^k_{ij}\frac{\partial}{\partial q^k}.$$

where the summation convention is enforced here and in what follows. Given any three vector fields X, Y, Z on Q, we say that the affine connection ∇ on Q is *torsion-free* if

$$[X, Y] = \nabla_X Y - \nabla_Y X \tag{1}$$

and is *compatible* with the metric $\langle\langle \cdot\,,\,\cdot\rangle\rangle$ if

$$\mathcal{L}_X \langle\langle Y\,,\,Z\rangle\rangle = \langle\langle \nabla_X Y\,,\,Z\rangle\rangle + \langle\langle Y\,,\,\nabla_X Z\rangle\rangle. \tag{2}$$

The Levi-Civita theorem states that on the Riemannian manifold Q there exists a unique affine connection ∇, which is torsion-free and compatible with the metric. Indeed, combining equations (1), (2) and their permutations, one obtains the equality

$$\begin{aligned} 2\langle\langle X\,,\,\nabla_Z Y\rangle\rangle = \mathcal{L}_Z \langle\langle X\,,\,Y\rangle\rangle + \langle\langle Z\,,\,[X, Y]\rangle\rangle + \mathcal{L}_Y \langle\langle X\,,\,Z\rangle\rangle& \\ + \langle\langle Y\,,\,[X, Z]\rangle\rangle - \mathcal{L}_X \langle\langle Y\,,\,Z\rangle\rangle - \langle\langle X\,,\,[Y, Z]\rangle\rangle,& \end{aligned} \tag{3}$$

which uniquely determines the connection ∇ as a function of the metric $\langle\langle \cdot\,,\,\cdot\rangle\rangle$. We call this ∇ the Riemannian (or Levi-Civita) connection on Q.

We conclude with two useful definitions. Given a real valued function f on Q, the *gradient* of f is the vector field ∇f such that

$$\langle\langle \nabla f\,,\,X\rangle\rangle \triangleq \mathcal{L}_X f.$$

Given a one form ω and a vector field X, the covariant derivative of ω with respect to X is the one form $\nabla_X \omega$ such that

$$(\nabla_X \omega) \cdot Y = \mathcal{L}_X(\omega \cdot Y) - \omega \cdot \nabla_X Y,$$

for all vector fields Y.

2.2 Computing covariant derivatives

Loosely speaking, covariant derivatives are directional derivatives of quantities defined on manifolds. Equation (1) relates them to the notion of Lie differentiation, whereas equation (2) plays the role of the Leibniz rule. In the following we present some useful approaches on how to compute covariant derivatives.

A first instructive case is when the manifold Q is a submanifold of $\mathbb{R}^n$ and the Riemannian metric on Q is the one induced by the Euclidean metric on $\mathbb{R}^n$. Then we can denote with π_q the orthogonal projection from $\mathbb{R}^n$ onto the tangent space T_qQ. Given any two vector fields X, Y on Q, it holds that

$$(\nabla_X Y)(q_0) = \pi_{q_0}\left(\frac{d}{dt}\bigg|_{t=0} Y(q(t))\right), \tag{4}$$

where $\{q(t), t \in \mathbb{R}\}$ is any curve on Q with $q(0) = q_0$ and $\dot{q}(0) = X(q_0)$. We refer to Boothby (1986, Chapter VII) for more details on this description of covariant differentiation.

In the general case, e.g. whenever the previous assumptions are not satisfied, we can express covariant derivatives in a system of local coordinates. The Christoffel symbols Γ_{ij}^k of a Riemannian connection are computed from equation (3) as follows. Denoting with $M_{ij} = \langle\langle \frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j} \rangle\rangle$, we have

$$\Gamma_{ij}^k = \frac{1}{2} M^{mk}\left(\frac{\partial M_{mj}}{\partial q^i} + \frac{\partial M_{mi}}{\partial q^j} - \frac{\partial M_{ij}}{\partial q^m}\right), \tag{5}$$

where M^{ij} is the inverse of the tensor M_{ij}. The covariant derivative of a vector field is then written as

$$\nabla_X Y = \left(\frac{\partial Y^i}{\partial q^j} X^j + \Gamma_{jk}^i X^j Y^k\right) \frac{\partial}{\partial q^i}, \tag{6}$$

and of a one form as

$$\nabla_X \omega = \left(\frac{\partial \omega_i}{\partial q^j} X^j - \Gamma_{ij}^k \omega_k X^j\right) dq^i. \tag{7}$$

Finally, we describe Riemannian connections within the context of Lie groups. For an introduction, see (Marsden & Ratiu 1994, Chapter 9). We denote with G a Lie group and with $\mathfrak{g}$ its Lie algebra. An example is the group of special orthogonal matrices $SO(3)$ and the set of skew symmetric matrices $\mathfrak{so}(3)$. The letters g and h denote elements in G, e is the identity. The Greek letters ξ and η denote elements in $\mathfrak{g}$ and $\mathrm{ad}_\xi \eta = [\xi, \eta]$ denotes the Lie bracket operation on $\mathfrak{g}$. The map $L_g : G \to G; h \mapsto gh$ is called left translation. A left invariant vector field satisfies the equality

$$X(gh) = T_h L_g X(h),$$

where $T_h L_g$ is the tangent map to L_g at h. If, in the last equality, we set $h = e$, it follows that $X(g) = T_e L_g \xi$, where $\xi = X(e)$. Thanks to this, $\mathfrak{g}$ can be identified with T_eG. Additionally, to simplify notation, we define $g \cdot \xi = T_e L_g \xi$. Left invariance is preserved by the Lie bracket operation, since

$$[g \cdot \xi \ , \ g \cdot \eta] = g \cdot [\xi, \eta].$$

Let $\mathfrak{g}^*$ denote the dual space of $\mathfrak{g}$, that is the set of covectors α such that $\langle \alpha, \xi \rangle$ is a linear function of $\xi \in \mathfrak{g}$. An inner product on the Lie algebra $\mathfrak{g}$, that is a tensor $\mathbb{I} : \mathfrak{g} \to \mathfrak{g}^*$, induces a left invariant metric on G by left translation. The Riemannian connection ∇ associated to this metric is of interest to us. A straightforward application of equation (3) shows that this connection satisfies

$$\nabla_{(g\cdot\xi)}\,(g\cdot\eta) = g\cdot\left({}_{\mathfrak{g}}\nabla_{\xi}\eta\right),$$

where the map ${}_{\mathfrak{g}}\nabla : \mathfrak{g}\times\mathfrak{g} \to \mathfrak{g}$ is defined by

$${}_{\mathfrak{g}}\nabla_{\xi}\eta = \frac{1}{2}[\xi,\eta] - \frac{1}{2}\mathbb{I}^{-1}\big(\operatorname{ad}^*_{\xi}\mathbb{I}\eta + \operatorname{ad}^*_{\eta}\mathbb{I}\xi\big), \tag{8}$$

and where $\operatorname{ad}^*_{\xi} : \mathfrak{g}^* \to \mathfrak{g}^*$ is the dual operator of ad_{ξ} defined by $\langle \operatorname{ad}^*_{\xi}\alpha, \eta\rangle = \langle \alpha, [\xi,\eta]\rangle$ for all $\alpha \in \mathfrak{g}^*$. Invariant connections on Lie groups are useful in various fields like hydrodynamic of ideal fluids (Arnold 1989, Appendix 1 and 2) and nonholonomic control systems (Bloch & Crouch 1995).

2.3 Mechanical systems in a Riemannian context

Here we describe a mechanical system and its equations of motion in a coordinate free fashion. Key ideas are regarding the system's kinetic energy as a Riemannian metric and writing the Euler-Lagrange's equations in terms of the associated Riemannian connection. For a more complete treatment, see (Lewis 1995).

We start with some definitions. A *simple mechanical control system* is defined by a Riemannian metric on a configuration manifold Q (defining the kinetic energy), a function V on Q (defining the potential energy), and m one-forms, $F^1, \ldots, F^m$, on Q (defining the inputs).

A simple mechanical system is said to be *fully actuated* if for all $q \in Q$, the family of vectors $\{F^1(q), \ldots, F^m(q)\}$ spans the whole vector space T^*_qQ, that is if there exists an independent input one form corresponding to each degree of freedom.

Let $M_q :\ T_qQ \to T^*_qQ$ denote the metric tensor associated to the kinetic energy and ∇ the corresponding Riemannian connection. Let $q(t) \in Q$ be the configuration of the system and $\dot{q}(t) \in T_qQ$ its velocity. Using the formalism introduced in the previous section, the forced Euler-Lagrange equations can be written as

$$\nabla_{\dot{q}}\dot{q} \;=\; M_q^{-1}\big(\mathrm{d}\,V(q) + F(t,q,\dot{q})\big), \tag{9}$$

where $\mathrm{d}\,V(q)$ is the differential of the potential function V and where the resultant force $F(q,t) = \sum F^a(q)u_a(t)$ is the input. In a system of local coordinates $(q^1, \ldots, q^n)$ the previous equations read

$$\ddot{q}^i + \Gamma^i_{jk}\dot{q}^j\dot{q}^k = M^{ij}\left(\frac{\partial V}{\partial q^j} + F_j\right).$$

Note that the Euler-Lagrange's equations are coordinate independent (intrinsic), in the sense that they are satisfied in every system of local coordinates.

Finally, we describe mechanical systems within the context of Lie groups. A *simple mechanical control system on a Lie group* is defined by a Lie group G with its algebra $\mathfrak{g}$, an inertia tensor $\mathbb{I} : \mathfrak{g} \to \mathfrak{g}^*$ (defining the kinetic energy) and m covectors $f^1, \dots, f^m$, on $\mathfrak{g}^*$ (defining the body-fixed forces).

If $g \in G$ denotes the configuration of the system and $\xi \in \mathfrak{g}$ the body-fixed velocity, then the equations of motion (9) reduce to two sets of equations, (kinematic and dynamic)

$$\begin{aligned} \dot{g} &= g \cdot \xi \\ \mathbb{I}\dot{\xi} &= \mathrm{ad}^*_\xi \, \mathbb{I}\xi + f, \end{aligned} \tag{10}$$

where $f = \sum f^a u_a(t)$ is the resultant force acting on the system and where ad^*_ξ is defined above. The previous equations are called the Euler-Poincaré equations (Marsden & Ratiu 1994).

3 Geometric Description of Configuration and Velocity Error

In this section we study the geometric objects involved in the controller design. To measure the distance between reference and actual configuration, we introduce the notion of error function. To measure the distance between reference and actual velocity, we introduce the notion of transport map. A design on two sphere manifold provides an example of our definitions. Finally we study the time derivative of the transport map. Together with a dissipation function, these ingredients are crucial in designing a tracking controller.

3.1 Error function and configuration error

Let φ be a smooth real valued function on $Q \times Q$. We shall call φ an *error function* if it is *positive definite*, that is $\varphi(q, r) \geq 0$ for all q and r, and $\varphi(q, r) = 0$ if and only if $q = r$. We shall say that the error function φ is *symmetric*, if $\varphi(q, r) = \varphi(r, q)$ for all q and r.

Let $\mathrm{d}_1 \varphi$ and $\mathrm{d}_2 \varphi$ denote the gradient of $\varphi(q, r)$ with respect to its first and second argument. We shall say that the error function φ is *(uniformly) quadratic with constant L* if for all $\epsilon > 0$ there exist two constants $b_1 \geq b_2 > 0$ such that $\varphi(q, r) < L - \epsilon$ implies

$$b_1 \| \mathrm{d}_1 \varphi(q, r) \|^2_{M_q} \;\geq\; \varphi(q, r) \;\geq\; b_2 \| \mathrm{d}_1 \varphi(q, r) \|^2_{M_q}. \tag{A1}$$

Here and in the following, the tag (An) denotes some design assumptions which will play a crucial role in later sections.

Theorem 3.1. *The quadratic assumption on the error function is necessary in order to prove exponential convergence rates. This is a weak requirement, since positive definite functions are always of at least quadratic order in a neighborhood of their critical point.*

When q and r are actual and reference configuration, we will sometimes call the quantity $\varphi(q, r)$ configuration error. As mentioned above, the error function φ will be instrumental in designing the proportional action.

3.2 Transport map and velocity error

Given two points $q, r \in Q$, we shall call a linear map $\mathcal{T}_{(q,r)} : T_r Q \to T_q Q$ a *transport map* if it is *compatible with the error function*, that is if

$$\mathrm{d}_2\, \varphi(q,r) = -\mathcal{T}^*_{(q,r)}\, \mathrm{d}_1\, \varphi(q,r), \tag{A2}$$

where $\mathcal{T}^*_{(q,r)} : T^*_q Q \to T^*_r Q$ is the dual map of $\mathcal{T}_{(q,r)}$. The transport map $\mathcal{T}$ is also required to be *smooth*, i.e., for all points r in Q and tangent vectors Y_r in $T_r Q$, the vector field $\mathcal{T}_{(q,r)} Y_r$ is smooth.

Using a transport map, velocities belonging to different tangent bundles can be compared. In the following, we shall call *velocity error* the quantity

$$\dot{e} \triangleq \dot{q} - \mathcal{T}_{(q,r)} \dot{r} \quad \in T_q Q. \tag{11}$$

Note the slight abuse of terminology, given that the velocity error is not the time derivative of a position error. Also note that since the definition of $\mathcal{T}$ and $\dot{e}$ are equivalent, we will sometimes talk about compatibility between configuration and velocity errors. The next lemma provides some insight into the meaning of the velocity error and of condition (A2).

Lemma 3.2 (Time derivative of an error function). *Let $\{q(t), t \in \mathbb{R}_+\}$ and $\{r(t), t \in \mathbb{R}_+\}$ be two smooth curves in Q. Let φ be an error function and $\mathcal{T}$ a compatible transport map. Then*

$$\frac{d}{dt}\varphi\big(q(t), r(t)\big) = \mathrm{d}_1\, \varphi\big(q(t), r(t)\big) \cdot \dot{e}(t), \qquad \forall t \in \mathbb{R}_+.$$

Proof. Applying the compatibility condition (A2), we have:

$$\begin{aligned}\frac{d}{dt}\varphi(q(t), r(t)) &= \mathrm{d}_1\, \varphi(q,r) \cdot \dot{q} \;+\; \mathrm{d}_2\, \varphi(q,r) \cdot \dot{r} \\ &= \mathrm{d}_1\, \varphi(q,r) \cdot \dot{q} \;+\; \big(- \mathcal{T}^*_{(q,r)}\, \mathrm{d}_1\, \varphi(q,r)\big) \cdot \dot{r} \\ &= \mathrm{d}_1\, \varphi(q,r) \cdot \big(\dot{q} - \mathcal{T}_{(q,r)} \dot{r}\big).\end{aligned}$$

□

The result can be restated as follows. As both q and r are functions of time, the time derivative of $\varphi : Q \times Q \to \mathbb{R}$ reduces to a derivative only with respect to the first argument

$$\mathcal{L}_{(\dot{q},\, \dot{r})}\varphi \;=\; \mathcal{L}_{(\dot{e},\, 0)}\varphi, \tag{12}$$

where $(X,\ Y)$ denotes a vector field on the product manifold $Q \times Q$.

Last, we introduce the notion of dissipation function, which will be useful in defining a derivative action. We define a *(linear Rayleigh) dissipation function* as a smooth, self-adjoint, positive definite tensor field $(K_d)(q) : T_q Q \to T^*_q Q$. We shall say that K_d is bounded if there exist $d_2 \geq d_1 > 0$ such that

$$d_2 \geq \sup_{q \in Q} \|K_d(q)\|_{M_q} \geq \inf_{q \in Q} \|K_d(q)\|_{M_q} \geq d_1, \tag{B1}$$

where $\|\cdot\|_{M}$ is the operator norm for $(1,1)$ type tensors on $T_q Q$ induced by the metric M_q on $T_q Q$. Here and in the following, the tag (Bn) denotes some boundedness assumptions which will play a crucial role in later sections.

3.3 Example design for the two sphere $\mathbb{S}^2$

To illustrate the previous ideas we apply them to the two sphere $\mathbb{S}^2 \triangleq \{p \in \mathbb{R}^3 | \; p^T p = 1\}$. Note that the Euclidean norm $\| \cdot \|$ on $\mathbb{R}^3$ induces a metric on the submanifold $\mathbb{S}^2$. Given an error function $\varphi : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}_+$, the norm of its differential $\| \, \mathrm{d}_1 \varphi \|$ is therefore well defined. In what follows, we let $a \times b$ denote the outer product between the two vectors $a, b \in \mathbb{R}^3$, and we let $\widehat{a} = a^\frown$ denote the 3×3 skew symmetric matrix such that $\widehat{a}\, b = a \times b$.

Lemma 3.3 (Design on the sphere). *Let q and r belong to $\mathbb{S}^2$. It holds that*

1. $\varphi(q, r) \triangleq 1 - q^T r$ *is a symmetric error function with differential* $\mathrm{d}_1 \varphi(q, r) = \widehat{q}^2 r = -r + (q^T r) q$*,*
2. $\varphi(q, r)$ *is a quadratic error function with constant* $L = 2$*, and*
3. $\mathcal{T}_{(q,r)} \triangleq (q^T r) I_3 + (r \times q)^\frown$ *is a compatible transport map.*

Proof. Since the orthogonal projection of $r \in \mathbb{S}^2$ onto $\operatorname{span}\{q\}^\perp$ is $r - (q^T r) q = -\widehat{q}^2 r$, we have

$$\mathcal{L}_{(\dot{q},0)} \varphi = -\dot{q}^T r = -\dot{q}^T (r - (q^T r) q) = (\widehat{q}^2 r)^T \dot{q}.$$

This proves (1). We prove (2) as follows. By assumption we are given an $\epsilon > 0$ such that $0 \le \varphi(q, r) \le 2 - \epsilon$, or equivalently $1 \ge q^T r \ge -1 + \epsilon$. The differential of the error function satisfies

$$\| \, \mathrm{d}_1 \varphi \|^2 = \| r - (q^T r) q \|^2 = 1 - (q^T r)^2 = (1 + q^T r) \varphi(q, r).$$

Since at $\varphi(q, r) = 0$ the bounds in assumption (A1) are verified, we only need to check that there exist $b_1 \ge b_2 > 0$ such that

$$b_1 (1 + q^T r) \ge 1 \ge b_2 (1 + q^T r).$$

This holds true for $b_1 = 1/2$ and $b_2 = 1/\epsilon$, proving (2). Next we show that φ and $\mathcal{T}$ are compatible (A2). This is verified with some algebraic simplifications based on the equality $v \times (w \times z) = (v^T z) w - (v^T w) z$. We have

$$\begin{aligned}
\mathcal{T}^* \, \mathrm{d}_1 \varphi(q, r) &= -(q^T r) r + (q^T r)^2 q - \big(r - (q^T r) q\big) \times (r \times q) \\
&= -(q^T r) r + (q^T r)^2 q - \Big(\big(r - (q^T r) q\big)^T q\Big) r - \Big(\big(r - (q^T r) q\big)^T r\Big) q \\
&= -(q^T r) r + (q^T r)^2 q + \big(1 - (q^T r)^2\big) q \\
&= q - (q^T r) r \;\equiv\; -\mathrm{d}_2 \varphi(q, r).
\end{aligned}$$

□

Next, we present some figures to compare our design with a traditional one. To warn of the effects of a design performed in local coordinates, Fig. 1 shows various paths connecting the same two points on a sphere. In each figure we employ a different projection, that is a different set of coordinates $x(q)$, and we draw the flow of the gradient of the (error) function $\|x(q) - x(r)\|^2$, that is a straight line in the particular set of coordinates. Note how the resulting paths depend on the choice of projection.

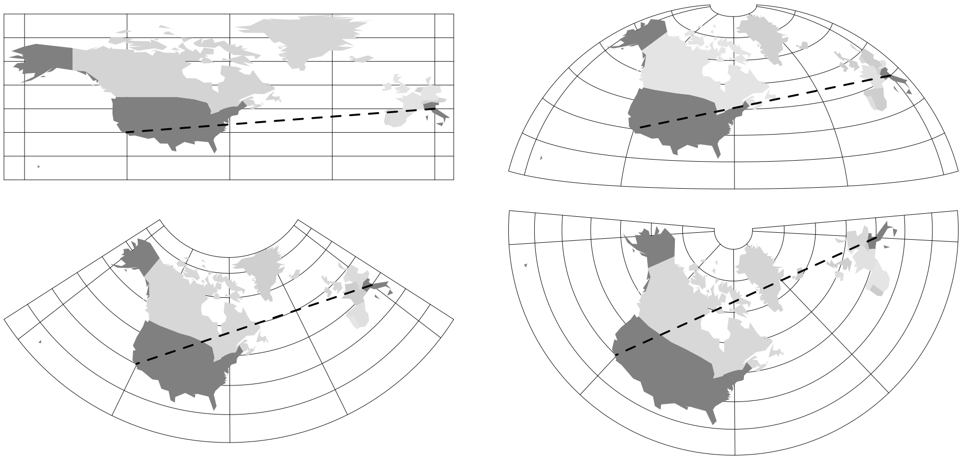

Figure 1: Straight lines on different projections of $\mathbb{S}^2$ are different curves.

In Fig. 2, we focus on two different choices of transport map and velocity error. Given a fixed reference velocity $\dot{r}$ (which is represented in both pictures by a thick arrow on top of the sphere), we draw for various points q the vector field $\mathcal{T}_{(q,r)}\dot{r}$. The left picture portrays the global, smooth design described above. On the right picture, we show the velocity error computed in a latitude, longitude parametrization. This is the procedure: if (θ_1, θ_2) are the local coordinates, then we can write

$$\dot{r} = \dot{r}_1 \frac{\partial}{\partial \theta_1}(r) + \dot{r}_2 \frac{\partial}{\partial \theta_2}(r).$$

We computed the "velocity error" vector field as

$$\mathcal{T}_{\text{lat/long}}\dot{r} = \dot{r}_1 \frac{\partial}{\partial \theta_1}(q) + \dot{r}_2 \frac{\partial}{\partial \theta_2}(q).$$

At the north pole of the latitude, longitude chart the singularity is evident.

3.4 Derivatives of the transport map and boundedness assumptions

So far we have introduced configuration and velocity errors that will be key ingredients in designing a proportional and derivative feedback in the next section. We now study how the transported reference velocity $(\mathcal{T}_{(q,r)}\dot{r})$ varies as a function of both $q(t)$ and $(r, \dot{r})(t)$. This will be useful in designing the feedforward action. We denote the total derivative of $(\mathcal{T}_{(q,r)}\dot{r})$ with

$$\frac{D\left(\mathcal{T}\dot{r}\right)}{dt} = \nabla_{\dot{q}}\left(\mathcal{T}\dot{r}\right) + \left.\frac{d}{dt}\right|_{q \text{ fixed}}\left(\mathcal{T}\dot{r}\right),$$

where the two terms are described as follows:

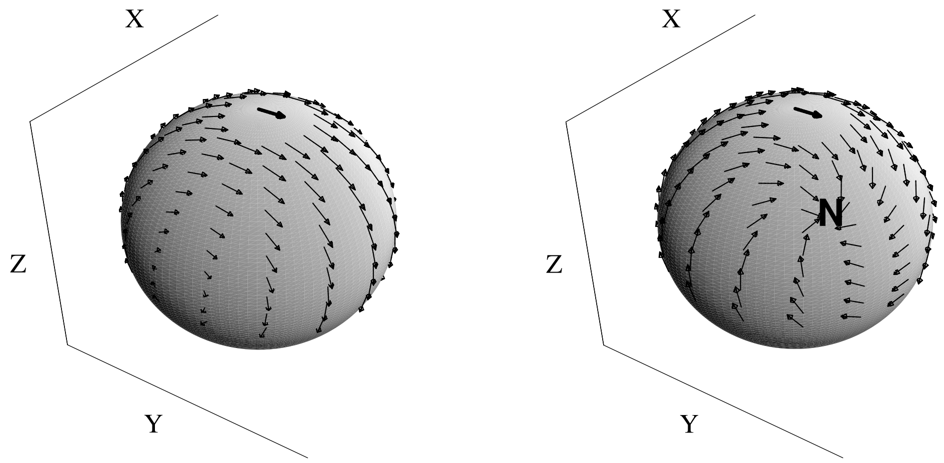

Figure 2: Transport maps on $\mathbb{S}^2$. We depict the vector $\mathcal{T}_{(q,r)}\dot{r}$ for two different transport maps: on the left our smooth global design, on the right a design based on a latitude, longitude chart, with the north pole denoted by the letter N.

1. At $(r, \dot{r})$ fixed, $\mathcal{T}_{(q,r)}\dot{r}$ is a vector field on Q and therefore its covariant derivative $\nabla_{\dot{q}}(\mathcal{T}\dot{r})$ is well-defined on Q. We call *covariant derivative of the transport map* the map $\nabla\mathcal{T} : T_qQ \times T_rQ \to T_qQ$ defined as
$$(\nabla_X\mathcal{T})Y_r \triangleq \nabla_X(\mathcal{T}Y_r),$$
for all tangent vectors $X \in T_qQ$ and $Y_r \in T_rQ$.

2. At q fixed, $\mathcal{T}_{(q,r)}\dot{r}$ is a vector on the vector space T_qQ and therefore its time derivative is well-defined. We denote it with the symbol:
$$\frac{d}{dt}\Big|_{q \text{ fixed}}(\mathcal{T}\dot{r}) \quad \in T_qQ.$$

Next, we compute coordinate expressions for the previous quantities. We denote with $\left\{\frac{\partial}{\partial q^1}, \dots \frac{\partial}{\partial q^n}\right\}$ a basis for T_qQ and with $\left\{\frac{\partial}{\partial r^1}, \dots \frac{\partial}{\partial r^n}\right\}$ a basis for T_rQ. Then we have the decompositions
$$\dot{r} = \dot{r}^\alpha \tfrac{\partial}{\partial r^\alpha}, \quad \text{and} \quad \mathcal{T}\dot{r} = \mathcal{T}^k_\alpha \dot{r}^\alpha \tfrac{\partial}{\partial q^k}.$$
If Γ^k_{ij} are the Christoffel symbols of ∇ and if X is a tangent vector in T_qQ, then we have
$$(\nabla_X\mathcal{T})^k_\alpha = \frac{\partial \mathcal{T}^k_\alpha}{\partial q^j} X^j + \Gamma^k_{ij}\mathcal{T}^i_\alpha X^j. \tag{13}$$
Regarding the time derivative at q fixed, we have
$$\left(\frac{d}{dt}\Big|_{q \text{ fixed}}(\mathcal{T}\dot{r})\right)^k = \frac{\partial \mathcal{T}^k_\alpha}{\partial r^\beta}\dot{r}^\alpha\dot{r}^\beta + \mathcal{T}^k_\alpha \ddot{r}^\alpha. \tag{14}$$

Remark 3.4. *Assume for an instant that the reference trajectory $r(t)$ obeys the same equations of motion as the actual mechanical system, that is*

$$\nabla_{\dot r}\dot r \;=\; M_r^{-1}F_r(t),$$

for some appropriate reference force $F_r(t) \in T_rQ$. Since in coordinates we have $(\nabla_{\dot r}\dot r)^\alpha = \ddot r^\alpha + \Gamma^\alpha_{\beta\gamma}(r)\dot r^\beta \dot r^\gamma$, then we can rewrite the equation (14) as

$$\begin{aligned}\frac{d}{dt}\Big|_{q\text{ fixed}}(\mathcal{T}\dot r) =\;& \mathcal{T}\nabla_{\dot r}\dot r \;+\; \left(\frac{\partial \mathcal{T}^k_\alpha}{\partial r^\beta} - \mathcal{T}^k_\gamma\Gamma^\gamma_{\alpha\beta}(r)\right)\dot r^\alpha \dot r^\beta \tfrac{\partial}{\partial q^k}.\\ \triangleq\;& \mathcal{T}\left(M_r^{-1}F_r(t)\right) \;+\; \left(\nabla_{(0,\dot r)}\mathcal{T}\right)\dot r,\end{aligned}$$

where the last equality defines implicitly the map $\nabla_{(0,\dot r)}\mathcal{T} : T_rQ \to T_qQ$. Note that the definition is coordinate independent, hence well-posed. Roughly speaking, this map is the covariant derivative of $\mathcal{T}$ with respect to $\dot r$. This statement can be made precise by defining $\mathcal{T}$ as a tensor on the product Riemannian manifold $Q \times Q$. We do not pursue this direction here for reasons of economy.

We conclude the section with some boundedness assumptions. We shall say that the transport map $\mathcal{T}$ has bounded covariant derivative and that the error function φ has bounded second covariant derivative if

$$\sup_{(q,r)\in Q\times Q} \|\nabla \mathcal{T}_{(q,r)}\|_M < \infty, \tag{B2}$$

and

$$\sup_{(q,r)\in Q\times Q} \|\nabla\, \mathrm{d}_1\, \varphi\,(q,r)\|_M < \infty, \tag{B3}$$

where $\|\cdot\|_M$ is the induced operator norm on the inner product space $(T_qQ,\, M_q)$. We shall say that the twice differentiable curve $\{r(t), t\in\mathbb{R}_+\} \subset Q$ is a reference trajectory with bounded time derivative if

$$\sup_{t\in\mathbb{R}} \|\dot r\|_{M_r} < \infty. \tag{B4}$$

Given the equalities (7) and (13), a sufficient condition for the bounds (B2) and (B3) to hold, is that the quantities M_{ij}, M^{ij}, Γ^k_{ij}, $\partial\mathcal{T}^i_\alpha/\partial q^k$ and $\partial^2\varphi/(\partial q^i \partial q^j)$ are bounded over $(q,r) \in Q\times Q$ for all i,j,k,α. On a compact manifold these conditions are implied by the smoothness of M, K_d, $\mathcal{T}$ and φ.

4 Tracking on Manifolds

4.1 Problem statement and main result

In this section we state and solve the exponential tracking problem. In what follows, we let $\{r(t), t\in\mathbb{R}_+\}$ denote a reference trajectory, $(\varphi,\mathcal{T})$ denote a pair of error function and transport map and we focus on a simple mechanical control system with no potential energy:

$$\nabla_{\dot q}\dot q \;=\; M_q^{-1}F, \qquad\qquad q\in Q. \tag{15}$$

We loosely state the control objective as follows:

Problem 4.1. *Design a control law $F = F(q, \dot{q};\ r, \dot{r})$ such that the configuration $q(t)$ tracks $r(t)$ with an exponentially decreasing error.*

Special care is needed to make this statement precise, as no trivial definition of exponential stability exists for systems on manifolds. We start by introducing a total energy function, defined as the sum of a generalized potential (the configuration error) and a kinetic energy function (the norm of the velocity error):

$$W_{\text{total}}(q, \dot{q};\ r, \dot{r}) \triangleq \varphi(q, r) + \frac{1}{2}\|\dot{q} - \mathcal{T}_{(q,r)}\dot{r}\|^2_{M_q}. \tag{16}$$

Alternatively we will write $W_{\text{total}}(t)$ for $W_{\text{total}}(q(t), \dot{q}(t);\ r(t), \dot{r}(t))$. Next we introduce the following definitions:

1. the curve $q(t) = r(t)$ is *stable with Lyapunov function* W_{total} if it holds $W(t) \leq W_{\text{total}}(0)$ from all initial conditions $(q(0), \dot{q}(0))$.
2. the curve $q(t) = r(t)$ is *exponentially stable with Lyapunov function* W_{total} if there exist two positive constants λ, k such that $W_{\text{total}}(t) \ \leq\ k\, W_{\text{total}}(0)\mathrm{e}^{-\lambda t}$, from all initial conditions $(q(0), \dot{q}(0))$.

We are now ready to state the main result.

Theorem 4.2 (Exponential tracking). *Consider the mechanical control system* (15)*, and let $\{r(t), t \in \mathbb{R}_+\}$ be a twice differentiable reference trajectory. Let φ be an error function, $\mathcal{T}$ be a transport map satisfying the compatibility condition* (A2) *and K_d be a dissipation function.*

If the control input is defined as $F = F_{\text{PD}} + F_{\text{FF}}$ with

$$\begin{aligned} F_{\text{PD}}(q, \dot{q};\ r, \dot{r}) \ &=\ -\,\mathrm{d}_1\,\varphi(q, r) - K_d\,\dot{e} \\ F_{\text{FF}}(q, \dot{q};\ r, \dot{r}) \ &=\ M_q\left((\nabla_{\dot{q}}\mathcal{T}_{(q,r)})\dot{r} \ +\ \frac{d}{dt}\Big|_{q\ \text{fixed}}\big(\mathcal{T}_{(q,r)}\dot{r}\big) \right), \end{aligned}$$

then the curve $q(t) = r(t)$ is stable with Lyapunov function W_{total}.

In addition, if the error function φ satisfies the quadratic assumption (A1) *with a constant L, and if the boundedness assumptions* (B1–B4) *hold, then the curve $q(t) = r(t)$ is exponentially stable with Lyapunov function W_{total} from all initial conditions* $(q(0), \dot{q}(0))$ *such that*

$$\varphi\big(q(0), r(0)\big) + \frac{1}{2}\|\dot{e}(0)\|^2_{M_q} < L.$$

4.2 Proof of the main theorem

Proof. The proof is divided into three parts: first we prove Lyapunov stability using the total energy as a Lyapunov function. Second, we add an additional “cross” term to the Lyapunov function. Finally, we conclude local exponential stability with a bounding argument.

The proof is based on the properties of covariant derivatives described in Section 2.1 and on the definitions in Section 3.4. This approach makes the proof straightforward and independent from the choice of local coordinates: the Lyapunov function, its time derivative and the final bounding argument are coordinate-free.

Part I: Lyapunov stability from total energy

We employ the total energy function $W_{\text{total}} = \varphi + \frac{1}{2}\|\dot{e}\|^2_{M_q}$ as candidate Lyapunov function. By Lemma 3.2 the time derivative of the first term is $\dot{\varphi} = \mathrm{d}_1\,\varphi \cdot \dot{e}$. We compute the time derivative of the second term in two steps. At r fixed, the equality (2) allows us to write

$$\begin{aligned}\frac{d}{dt}\Big|_{r\text{ fixed}}\frac{1}{2}\|\dot{e}\|^2_{M_q} &= \frac{1}{2}\mathcal{L}_{\dot{q}}\langle\!\langle \dot{e}\,,\,\dot{e}\rangle\!\rangle = \langle\!\langle \dot{e}\,,\,\nabla_{\dot{q}}\dot{e}\rangle\!\rangle \\ &= \langle\!\langle \dot{e}\,,\,\nabla_{\dot{q}}\,(\dot{q} - \mathcal{T}\dot{r})\rangle\!\rangle \\ &= \langle\!\langle \dot{e}\,,\,M_q^{-1}(F_{\mathrm{PD}} + F_{\mathrm{FF}}) - (\nabla_{\dot{q}}\mathcal{T})\dot{r}\rangle\!\rangle.\end{aligned}$$

At q fixed, we have instead

$$\frac{d}{dt}\Big|_{q\text{ fixed}}\frac{1}{2}\|\dot{e}\|^2_{M_q} = \langle\!\langle \dot{e}\,,\,\frac{d}{dt}\Big|_{q\text{ fixed}}(\dot{q} - \mathcal{T}\dot{r})\rangle\!\rangle = -\langle\!\langle \dot{e}\,,\,\frac{d}{dt}\Big|_{q\text{ fixed}}(\mathcal{T}\dot{r})\rangle\!\rangle.$$

Plugging in we have

$$\begin{aligned}\frac{d}{dt}W_{\text{total}} &= \mathrm{d}_1\,\varphi \cdot \dot{e} + \langle\!\langle \dot{e}\,,\,M_q^{-1}(F_{\mathrm{PD}} + F_{\mathrm{FF}}) - (\nabla_{\dot{q}}\mathcal{T})\dot{r} - \frac{d}{dt}\Big|_{q\text{ fixed}}(\mathcal{T}\dot{r})\rangle\!\rangle \\ &= \mathrm{d}_1\,\varphi \cdot \dot{e} + \langle\!\langle M_q^{-1}F_{\mathrm{PD}}\,,\,\dot{e}\rangle\!\rangle \\ &= \mathrm{d}_1\,\varphi \cdot \dot{e} + (-\,\mathrm{d}_1\,\varphi - K_d\,\dot{e}) \cdot \dot{e} \;=\; -K_d\dot{e} \cdot \dot{e}\end{aligned}$$

so that $\frac{d}{dt}W_{\text{total}}$ is negative semidefinite and Lyapunov stability as defined in Theorem 4.2 is proven.

Part II: Introduction of cross term

To construct a strict Lyapunov function (i.e. a function with a time derivative strictly definite), we add a "small" cross term to W_{total}. Let ϵ be a positive constant, let

$$W_{\text{cross}}(t) \;=\; \dot{\varphi} \;=\; \mathrm{d}_1\,\varphi \cdot \dot{e}$$

and consider the candidate Lyapunov function

$$W \;\triangleq\; W_{\text{total}} + \;\epsilon\,W_{\text{cross}}.$$

We need to show that there exists a sufficiently small ϵ, such that W is positive definite in φ and $\|\dot{e}\|_{M_q}$. We start by noting that from Part I and the assumptions on the initial conditions we have

$$W_{\text{total}}(t) \leq W_{\text{total}}(0) < L \quad\Longrightarrow\quad \varphi(t) < L,$$

which implies that the bounds (A1) on the differential of the error function hold for all time. Then we have

$$\begin{aligned}W &\geq \varphi + \frac{1}{2}\|\dot{e}\|^2_{M_q} - \epsilon\|\,\mathrm{d}_1\,\varphi\|_{M_q} \cdot \|\dot{e}\|_{M_q} \\ &\geq \varphi + \frac{1}{2}\|\dot{e}\|^2_{M_q} - \epsilon(1/\sqrt{b_2})\sqrt{\varphi}\,\|\dot{e}\|_{M_q},\end{aligned}$$

and therefore

$$W \geq \frac{1}{2}\begin{bmatrix}\sqrt{\varphi}\\ \|\dot{e}\|_{M_q}\end{bmatrix}^T \begin{bmatrix}2 & -\epsilon/\sqrt{b_2}\\ -\epsilon/\sqrt{b_2} & 2\end{bmatrix}\begin{bmatrix}\sqrt{\varphi}\\ \|\dot{e}\|_{M_q}\end{bmatrix} \triangleq \begin{bmatrix}\sqrt{\varphi}\\ \|\dot{e}\|_{M_q}\end{bmatrix}^T \mathcal{P}\begin{bmatrix}\sqrt{\varphi}\\ \|\dot{e}\|_{M_q}\end{bmatrix}.$$

By choosing $\epsilon < 2\sqrt{b_2}$, the matrix $\mathcal{P}$ and the function W are positive definite with respect to $\sqrt{\varphi}$ and $\|\dot{e}\|_{M_q}$.

Next, we compute the time derivative of W_{cross}. At r fixed, we have

$$\begin{aligned}\frac{d}{dt}\Big|_{r \text{ fixed}}\dot{\varphi} &= \nabla_{\dot{q}}\left(\mathrm{d}_1\,\varphi\cdot\dot{e}\right) = \left(\nabla_{\dot{q}}\,\mathrm{d}_1\,\varphi\right)\cdot\dot{e} + \mathrm{d}_1\,\varphi\cdot\left(\nabla_{\dot{q}}\dot{e}\right)\\ &= \left(\nabla_{\dot{q}}\,\mathrm{d}_1\,\varphi\right)\cdot\dot{e} + \mathrm{d}_1\,\varphi\cdot\left(M_q^{-1}F - (\nabla_{\dot{q}}\mathcal{T})\dot{r}\right). \end{aligned} \tag{17}$$

At q fixed, we have

$$\frac{d}{dt}\Big|_{q \text{ fixed}}(\mathrm{d}_1\,\varphi) = \mathrm{d}_1\left(\frac{d}{dt}\Big|_{q \text{ fixed}}\varphi\right) = \mathrm{d}_1\left(\mathrm{d}_1\,\varphi\cdot(-\mathcal{T}\dot{r})\right),$$

and therefore

$$\begin{aligned}\frac{d}{dt}\Big|_{q \text{ fixed}}\dot{\varphi} &= \left(\frac{d}{dt}\Big|_{q \text{ fixed}}\mathrm{d}_1\,\varphi\right)\cdot\dot{e} + \mathrm{d}_1\,\varphi\cdot\left(\frac{d}{dt}\Big|_{q \text{ fixed}}\dot{e}\right)\\ &= \mathrm{d}_1\left(\mathrm{d}_1\,\varphi\cdot(-\mathcal{T}\dot{r})\right)\cdot\dot{e} + \mathrm{d}_1\,\varphi\cdot\left(-\frac{d}{dt}\Big|_{q \text{ fixed}}(\mathcal{T}\dot{r})\right)\\ &= -\mathfrak{L}_{\dot{e}}\left(\mathrm{d}_1\,\varphi\cdot(\mathcal{T}\dot{r})\right) - \mathrm{d}_1\,\varphi\cdot\left(\frac{d}{dt}\Big|_{q \text{ fixed}}(\mathcal{T}\dot{r})\right)\\ &= -\left(\nabla_{\dot{e}}\,\mathrm{d}_1\,\varphi\right)\cdot(\mathcal{T}\dot{r}) - \mathrm{d}_1\,\varphi\cdot\left(\nabla_{\dot{e}}(\mathcal{T}\dot{r})\right) - \mathrm{d}_1\,\varphi\cdot\left(\frac{d}{dt}\Big|_{q \text{ fixed}}(\mathcal{T}\dot{r})\right).\end{aligned}$$

Summing the previous equation with equation (17) we obtain

$$\begin{aligned}\frac{d}{dt}\dot{\varphi} = &\left(\nabla_{\dot{e}}\,\mathrm{d}_1\,\varphi\right)\cdot\dot{e} + \mathrm{d}_1\,\varphi\cdot\left(M_q^{-1}F - (\nabla_{\dot{q}}\mathcal{T})\dot{r} - \frac{d}{dt}\Big|_{q \text{ fixed}}(\mathcal{T}\dot{r})\right)\\ &- \mathrm{d}_1\,\varphi\cdot\left(\nabla_{\dot{e}}(\mathcal{T}\dot{r})\right)\end{aligned}$$

and substituting the control force F

$$\begin{aligned}&= \left(\nabla_{\dot{e}}\,\mathrm{d}_1\,\varphi\right)\cdot\dot{e} + \mathrm{d}_1\,\varphi\cdot\left(M_q^{-1}F_{\text{PD}}\right) - \mathrm{d}_1\,\varphi\cdot\left(\nabla_{\dot{e}}(\mathcal{T}\dot{r})\right)\\ &= -\|\,\mathrm{d}_1\,\varphi\|^2_{M_q} + \left(\nabla_{\dot{e}}\,\mathrm{d}_1\,\varphi\right)\cdot\dot{e} + \mathrm{d}_1\,\varphi\cdot\left(M_q^{-1}K_d\dot{e}\right) - \mathrm{d}_1\,\varphi\cdot\left((\nabla_{\dot{e}}\mathcal{T})\dot{r}\right).\end{aligned}$$

Next, by means of the quadratic assumption (A1) on φ, we can express $\dot{W}_{\text{cross}} = \ddot{\varphi}$ as a function of φ and $\|\dot{e}\|_{M_q}$. It holds that

$$\dot{W}_{\text{cross}} \;=\; \ddot{\varphi} \;\leq\; -\begin{bmatrix}\sqrt{\varphi}\\ \|\dot{e}\|_{M_q}\end{bmatrix}^T \mathcal{Q}_{\text{cross}}\begin{bmatrix}\sqrt{\varphi}\\ \|\dot{e}\|_{M_q}\end{bmatrix},$$

where the symmetric matrix $\mathcal{Q}_{\text{cross}}$ has the following entries:

$$
\begin{aligned}
(\mathcal{Q}_{\text{cross}})_{1,1} &= 1/b_1, \\
(\mathcal{Q}_{\text{cross}})_{2,1} &= -\left(\sup_{q\in Q} \|K_d\|_{M_q} \;+\; \sup_t \|\dot r\|_{M_r} \cdot \sup_{(q,r)\in Q\times Q} \|\nabla \mathcal{T}\|_M \right) / \sqrt{b_1} \\
(\mathcal{Q}_{\text{cross}})_{2,2} &= - \sup_{(q,r)\in Q\times Q} \|\nabla \, \mathrm{d}_1 \, \varphi\|_M .
\end{aligned}
$$

Note that the operators in $\mathcal{Q}_{\text{cross}}$ are bounded: $(\mathcal{Q}_{\text{cross}})_{1,2}$ is upper bounded due to assumptions (B1), (B4) and (B2), $(\mathcal{Q}_{\text{cross}})_{2,2}$ is upper bounded due to the assumption (B3).

Part III: Bounding arguments

As last step, we bound the time derivative of the Lyapunov function $W = W_{\text{total}} + \epsilon W_{\text{cross}}$. We have

$$
\frac{d}{dt} W \;\le\; - \begin{bmatrix} \sqrt{\varphi} \\ \|\dot e\|_{M_q} \end{bmatrix}^T \mathcal{Q} \begin{bmatrix} \sqrt{\varphi} \\ \|\dot e\|_{M_q} \end{bmatrix},
$$

where the symmetric matrix $\mathcal{Q}$ is positive definite for small enough ϵ, since

$$
\begin{aligned}
\mathcal{Q}_{1,1} &= \epsilon(\mathcal{Q}_{\text{cross}})_{1,1} \\
\mathcal{Q}_{1,2} &= \epsilon(\mathcal{Q}_{\text{cross}})_{1,2} \\
\mathcal{Q}_{2,2} &= \inf_{q\in Q} \|K_d\|_{M_q} + \epsilon(\mathcal{Q}_{\text{cross}})_{2,2},
\end{aligned}
$$

and $\mathcal{Q}_{2,2}$ is bounded away from zero thanks to (B1). Hence, there exist a $\lambda > 0$ such that $\dot W < -\lambda W$. Finally it holds that

$$
W_{\text{total}} = \begin{bmatrix} \sqrt{\varphi} \\ \|\dot e\|_{M_q} \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \sqrt{\varphi} \\ \|\dot e\|_{M_q} \end{bmatrix}
$$

and for an appropriate positive k_1

$$
\begin{aligned}
&\le\; k_1 \begin{bmatrix} \sqrt{\varphi} \\ \|\dot e\|_{M_q} \end{bmatrix}^T \mathcal{P} \begin{bmatrix} \sqrt{\varphi} \\ \|\dot e\|_{M_q} \end{bmatrix} \\
&\le\; k_1 W(t) \;\le\; k_1 W(0) \mathrm{e}^{-\lambda t} \;\le\; 2 k_1 W_{\text{total}}(0) \mathrm{e}^{-\lambda t},
\end{aligned}
$$

where we used the fact that $W_{\text{total}}(0) + \epsilon W_{\text{cross}}(0) \le 2 W_{\text{total}}(0)$. □

4.3 Remarks and extensions

The design process and the theorem's results are global in the reference position $r(t)$ but only local in the configuration q (the error function $\varphi(q, r)$ must remain smaller than the parameter L). This cannot be avoided because of possible topological properties of the manifold Q. For additional details we refer to (Koditschek 1989), where the author discusses the global aspects of the point stabilization problem.

Theorem 4.2 achieves Lyapunov and exponential stability with respect to the particular total energy W_{total} we devised. Therefore the design of error function and transport map plays a central role in imposing performance requirements. For example the choice of error function $\varphi(q,r)$ affects the type of convergence we obtain: the configuration q converges to the reference r in the topology induced by φ. Additionally, the choice of $(\varphi, \mathcal{T})$ determines the (computational) complexity of the control action. For example, one particular transport map might be desirable since it generates a "simple" velocity error and a "simple" feedforward control. However, the compatibility condition (A2) constitutes a constraint on the set of admissible pairs $(\varphi, \mathcal{T})$. The next section, and in particular the $SO(3)$ and $SE(3)$ cases, illustrates some of the tradeoffs involved in the control design.

As expected, the final control law is sum of a feedback and a feedforward term. This is in agreement with the ideas exposed in (Murray 1995) on "two degree of freedom system design" for mechanical systems. While the feedforward term depends on the geometry of both the manifold and the mechanical system, the feedback term is designed knowing only the configuration manifold Q. We expect the ideas of configuration and velocity error to be relevant for more general second order nonlinear systems on manifolds.

Given a mechanical system on a Riemannian manifold, Theorem 4.2 focuses on the geometric aspects of the tracking problem. The result can be extended in various directions, for example allowing for potential, gyroscopic and viscous forces. Also, should some inertia's parameters be unknown, adaptive capabilities can be added to the control design via some standard techniques, see for example (Slotine & Li 1989). We do not present these extensions here for brevity's sake and because they are outside the focus of the present treatment.

5 Applications and Extensions

In what follows we describe examples of the design techniques and of the stability results presented so far.

5.1 A pointing device on $\mathbb{S}^2$

In this section we apply the main theorem to the sphere example described in Section 3.3. Motivating applications are the so called "spin axis stabilization" problem for a satellite and workspace control of a robot manipulator such as a pan tilt unit.

Recall from Section 2.2 that, since $\mathbb{S}^2$ is a submanifold of $\mathbb{R}^3$, the Euclidean metric on $\mathbb{R}^3$ induces a Riemannian connection ∇ on $\mathbb{S}^2$. In particular this connection ∇ can be described in terms of the orthogonal projection π_q from $\mathbb{R}^3$ to $T_q\mathbb{S}^2$ as follows. If $\{q(t)\}$ is a curve and $X(q)$ is a vector field on $\mathbb{S}^2 \subset \mathbb{R}^3$, then

$$\big(\nabla_{\dot q} X\big)(q) = \pi_q\left(\ddot X(q(t))\right) = \ddot X(q(t)) - \left(q(t)^T \ddot X(q(t))\right) q(t),$$

where both $q(t)$ and $X(q(t))$ are thought of as vectors on $\mathbb{R}^3$. In the following we consider a mechanical system defined by

$$\nabla_{\dot q}\dot q = F, \tag{18}$$

where the input force F lives on the cotangent bundle $T_q^*\mathbb{S}^2$, which we identify with $T_q\mathbb{S}^2 \subset \mathbb{R}^3$. Last, recall that in Section 3.3 we designed a quadratic error function and a compatible transport map as

$$\varphi(q,r) \triangleq 1 - q^T r \qquad \text{and} \qquad \mathcal{T}_{(q,r)} \triangleq (q^T r) I_3 + (r \times q)\widehat{\ },$$

where r is the reference configuration on $\mathbb{S}^2$.

Lemma 5.1 (Tracking on the sphere). *Consider the system in equation* (18) *and let* $\{r(t), t \in \mathbb{R}_+\}$ *be a reference trajectory with* $\sup_t \|\dot r\|$ *bounded . Let* k_p *and* k_d *be two positive constants. Then the control law* $F = F_{\text{PD}} + F_{\text{FF}}$*, with*

$$\begin{aligned} F_{\text{PD}} &= -k_p \widehat{q}^2 r \; - \; k_d\big(\dot q - q \times (r \times \dot r)\big) \\ F_{\text{FF}} &= \big(\dot r^T r \times q\big)(q \times \dot q) \; + \; q \times (r \times \nabla_{\dot r} \dot r), \end{aligned}$$

exponentially stabilizes $k_p\varphi(q,r) + \frac{1}{2}\|\dot q - \mathcal{T}_{(q,r)}\dot r\|^2$ *to zero from any initial condition* $q(0) \neq -r(0)$ *and for all* $\dot q(0), \dot r(0), k_p$ *such that*

$$k_p > \frac{\|\dot q(0) - \mathcal{T}_{(q,r)}\dot r(0)\|^2}{2\big(1 + q(0)^T r(0)\big)}.$$

Proof. In Section 3.3 we proved that φ is quadratic (A1) and $\mathcal{T}$ is compatible (A2). Additionally, since $\mathbb{S}^2$ is compact, the conditions (B1–3) are satisfied, because of the smoothness of the metric, of k_d, of $\mathcal{T}$ and of φ. The assumption (B4) is explicitly made in the text.

Hence we only need to prove that the F_{PD} and the F_{FF} above are designed as prescribed by Theorem 4.2. Applying twice the equality $v \times (w \times z) = (v^T z)w - (v^T w)z$, we have

$$\begin{aligned} \mathcal{T}\dot r &= (q^T r)\dot r + (r \times q) \times \dot r = (q^T r)\dot r - (\dot r^T q) r \\ &= q \times (r \times \dot r), \end{aligned}$$

and

$$\frac{d}{dt}\Big|_{q \text{ fixed}} \big(\mathcal{T}_{(q,r)}\dot r\big) = q \times (r \times \ddot r) = q \times (r \times \nabla_{\dot r}\dot r).$$

Finally, following the description in Section 2.2, we compute the covariant derivative of the vector field $\big(\mathcal{T}\dot r\big)(q)$ by differentiating it with respect to time and then projecting the result onto the tangent plane at q. In formulas this reads as:

$$\big(\nabla_{\dot q}\mathcal{T}\big)\dot r = \pi_q\left(\frac{d}{dt}\Big|_{r \text{ fixed}} \mathcal{T}\dot r\right) = -\widehat{q}^2\left(\frac{d}{dt}\Big|_{r \text{ fixed}} \dot{\mathcal{T}}\dot r\right).$$

Summarizing some algebraic equalities, we have

$$\begin{aligned} \big(\nabla_{\dot q}\mathcal{T}\big)\dot r &= -\widehat{q}^2\big(\frac{d}{dt}\Big|_{r \text{ fixed}} q \times (r \times \dot r)\big) = -\widehat{q}\big(q \times (\dot q \times (r \times \dot r))\big) \\ &= -\widehat{q}\big(q^T(r \times \dot r)\dot q - q^T \dot q (r \times \dot r)\big) = (q^T r \times \dot r)(\dot q \times q). \end{aligned}$$

This completes the proof. ∎

5.2 A robot manipulator on $\mathbb{R}^n$

In this section, we shall recover the standard results on tracking control of manipulators contained in (Murray et al. 1994). Let $q \in \mathbb{R}^n$ be the joint variables and $M(q)$ be the inertia matrix of the manipulator. The design described in Section 3 is performed as follows.

Let K be a symmetric positive definite matrix and let $\varphi(q,r) = \frac{1}{2}(q-r)^T K_p (q-r)$ be a quadratic error function. Thanks to the identification $T_q\mathbb{R}^n = T_r\mathbb{R}^n$, we let the transport map be equal to the identity matrix: $\mathcal{T}_{(q,r)} = I_n$. Assumptions (A1) and (A2) are easily verified. To design the feedforward action, we compute the covariant derivative of $\mathcal{T}$. Let $\left\{\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}\right\}$ be the standard basis in $\mathbb{R}^n$, let $\{i,j,k,\dots\}$ be indices over q and $\{\alpha,\beta,\dots\}$ be indices over r. Then, from equation (13)

$$(\nabla I_n)^i_{\alpha j} = \frac{\partial (I_n)^i_\alpha}{\partial q^j} + \Gamma^i_{jk}(I_n)^k_\alpha = \Gamma^i_{j\alpha},$$

Therefore, in contrast to a naive guess, the covariant derivative of the identity map is different from zero. Given a symmetric positive definite K_d, the control law is

$$\begin{aligned} F_{\mathrm{PD}} &= -K_p(q-r) - K_d(\dot q - \dot r) \\ F_{\mathrm{FF}} &= M(q)\Big((\nabla_{\dot q} I_n)\dot r + \frac{d}{dt}\Big|_{q\ \mathrm{fixed}}\dot r\Big) \\ &= M(q)\big(\Gamma^i_{j\alpha}\dot q^j \dot r^\alpha \tfrac{\partial}{\partial q^i} + \ddot r\big) \ \equiv\ M(q)\ddot r + C(q,\dot q)\dot r, \end{aligned} \tag{19}$$

where $C(\cdot,\cdot)$ is the *Coriolis matrix* typically encountered in robotics. The control law $F = F_{\mathrm{PD}} + F_{\mathrm{FF}}$ agrees with the one presented in (Murray et al. 1994, Chapter 4, Section 5.3) under the name of "augmented PD control". The assumptions (B1–B4) can be written in terms of Γ^k_{ij} and $\dot r$ being bounded over $t \in \mathbb{R}$ and $q \in \mathbb{R}^n$.

Linearization by state transformations and by feedback

Sometimes a simple state transformation suffices for the linearization of the Euler-Lagrange's equations. This happens when there exists a choice of local coordinates such that the Christoffel symbols vanish. If the designed described above is performed in this specific set of coordinates, the expression (19) for the feedforward control simplifies considerably since the cross term $(\dot q, \dot r)$ vanishes. More details on this case are discussed in (Bedrossian & Spong 1995).

More generally the Euler-Lagrange's equations can be linearized by means of a feedback transformation. By setting $F = M^{-1}(q)\,(U - C(q,\dot q))\,\dot q$, we have that the equations of motion

$$M(q)\ddot q + C(q,\dot q)\dot q = F \qquad \text{become} \qquad \ddot q = U. \tag{20}$$

A tracking controller is then designed using linear techniques. The design procedure is the so-called computed torque method, (Murray et al. 1994, Chapter 4, Section 5.2). Note that a controller designed this way depends on the initial choice of the coordinates system $(q^1, \dots, q^n)$.

We reconcile this method with our framework as follows. We denote with $\overline{\nabla}$ the connection characterized by vanishing Christoffel symbols in the chart $\{q^1, \dots, q^n\}$. Then the equality $\ddot{q} = U$ can be written as $\overline{\nabla}_{\dot{q}}\dot{q} = U$, hence as a mechanical system. In other words, we regard the feedback transformation (20) as a "change of connection" from ∇ to $\overline{\nabla}$. This idea is described in some theoretical details in (Kobayashi & Nomizu 1963, Proposition 7.10). Summarizing, the computed torque method falls within the scope of Theorem 4.2 if feedback pre-transformations are allowed.

5.3 A satellite on the rotation group $SO(3)$

In the next two sections we design tracking controllers for mechanical systems defined on the group of rotations $SO(3)$ and on the group of rigid motions $SE(3)$. We focus on rigid bodies with body-fixed forces and invariant kinetic energy, as satellites and underwater vehicles. Nevertheless our treatment is relevant also for workspace control of robot manipulators. This section presents the attitude control problem for a satellite.

The configuration of the satellite (rigid body) is the rotation matrix R representing the position of a frame fixed with the rigid body with respect to an inertially fixed frame. A rotation matrix on $\mathbb{R}^3$ is an element on the special orthogonal group $SO(3) = \{R \in \mathbb{R}^{3\times3} | \, RR^T = I_3, \ \det(R) = +1\}$. The kinematic equation describing the evolution of $R(t)$ is

$$\dot{R} = R\widehat{\Omega} \tag{21}$$

where $\Omega \in \mathbb{R}^3$ is the body angular velocity expressed in the body frame. Recall that the matrix $\widehat{\Omega}$ is defined such that $\widehat{\Omega}x = \Omega \times x$ for all $x \in \mathbb{R}^3$ and it belongs to the space of skew symmetric matrices $\mathfrak{so}(3) = \{S \in \mathbb{R}^{3\times3} | \, S^T = -S\}$. We refer to (Murray et al. 1994) for additional details.

The kinetic energy of the rigid body is $\frac{1}{2}\Omega^T \mathbb{J}\Omega$, where the inertia matrix $\mathbb{J}$ is symmetric and positive definite. The Euler equations describing the time evolution of Ω are

$$\mathbb{J}\dot{\Omega} = \mathbb{J}\Omega \times \Omega + f, \tag{22}$$

where $f \in \left(\mathbb{R}^3\right)^*$ is the resultant torque acting on the body.

Error functions

Let $\{R_d(t), t \in \mathbb{R}_+\}$ denote the reference trajectory corresponding to a desired or reference frame and let $\widehat{\Omega}_d = R_d^T \dot{R}_d$ denote the reference velocity in the reference frame. Using the group operation, we define right and left attitude errors as

$$R_{e,r} \triangleq R_d^T R \qquad \text{and} \qquad R_{e,\ell} \triangleq R R_d^T. \tag{23}$$

The matrix $R_{e,r}$ is the relative rotation from the body frame to the reference frame. Two error functions are then defined as $\varphi_r(R, R_d) \triangleq \phi(R_{e,r})$ and $\varphi_\ell(R, R_d) \triangleq \phi(R_{e,\ell})$, where $\phi : \, SO(3) \to \mathbb{R}_+$ is defined as (Koditschek 1989)

$$\phi(R_e) \triangleq \frac{1}{2} \operatorname{tr}\left(K_p(I_3 - R_e)\right).$$

If the eigenvalues $\{k_1, k_2, k_3\}$ of the symmetric matrix K_p satisfy $k_i + k_j > 0$ for $i \neq j$, then both error functions φ_ℓ and φ_R are symmetric, positive definite and quadratic with constant $L = \min_{i \neq j}(k_i + k_j)$. Locally near the identity the function ϕ assigns a weight $k_2 + k_3$ to a rotation error about the first axis (and similarly for the other axes). Appendix A contains the proof of these facts and the expression of ϕ in the unit quaternion representation.

Velocity errors

To define compatible velocity errors, we compute the time derivative of the two error functions. Let the matrix skew(A) denote $\frac{1}{2}(A - A^T)$ and let $\cdot^\vee$ denote the inverse operator to $\widehat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$. We have

$$\frac{d}{dt}\varphi_r = \left(\text{skew}(K_p R_{e,r})^\vee\right)^T \Omega_{e,r} \tag{24}$$

$$\frac{d}{dt}\varphi_\ell = \left(\text{skew}(K_p R_{e,\ell})^\vee\right)^T R_d \, \Omega_{e,\ell}, \tag{25}$$

where we define right and left velocity errors in the body frame as

$$\Omega_{e,r} \triangleq \Omega - R_{e,r}^T \Omega_d \qquad \text{and} \qquad \Omega_{e,\ell} \triangleq \Omega - \Omega_d.$$

Note the slightly improper wording, since a velocity error $\dot{e} = \dot{R} - \mathcal{T}\dot{R}_d$ lives on the tangent bundle $T_R SO(3)$. A precise statement is

$$\begin{aligned} \dot{e}_\ell &= R\widehat{\Omega}_{e,\ell} \equiv \dot{R} - \left(R R_d^T\right) \dot{R}_d \\ \dot{e}_r &= R\widehat{\Omega}_{e,r} \equiv \dot{R} - \dot{R}_d \left(R_d^T R\right). \end{aligned}$$

These equalities also motivate the names "left" and "right". A left (right) velocity error is obtained by left (right) translation of the velocity $\dot{R}_d$.

Next we describe compatible couples of configuration and velocity errors. Equation (24) suggests that a right attitude error $R_d^T R$ and a right velocity error $\Omega - R^T R_d \Omega_d$ are compatible. This couple is the most common choice in the literature, see for example (Meyer 1971), (Koditschek 1989), (Wen & Kreutz-Delgado 1991), and (Egeland & Godhavn 1994).

Left attitude and velocity error appear less frequently (Luh, Walker & Paul 1980). With this choice both the velocity error and, as we show below, the feedforward control have a simple expression. Remarkably, when the gain K_p is a scalar multiple of the identity $k_p I_3$, the left and right error functions are equal and the couple $(\varphi_{e,r}, \Omega_{e,\ell})$ is compatible. Finally, coordinate based approaches are also possible. The velocity error in (Slotine & Di Benedetto 1990) is taken to be the difference between the rate of change of the Gibbs vectors for actual and reference attitude. Similarly, in the flight control literature, Euler angles and their rates are often used (Etkin 1982).

Control laws and simulations

Finally we summarize the design process.

Lemma 5.2. *Consider the system in equation* (22)*. Let* $\{R_d(t), t \in \mathbb{R}_+\}$ *denote the reference trajectory and let* $\widehat{\Omega}_d = R_d^T \dot{R}_d$ *denote its bounded body-fixed velocity. Corresponding to the two choices of attitude error, we define*

$$\begin{aligned} f_r &= -\operatorname{skew}(K_p R_{e,r})^\vee \quad - K_d \Omega_{e,r} + \Omega \times \mathbb{J}(R_{e,r}^T \Omega_d) + \mathbb{J}(R_{e,r}^T \dot{\Omega}_d) \\ f_\ell &= -R_d^T \operatorname{skew}(K_p R_{e,\ell})^\vee - K_d \Omega_{e,\ell} + \Omega_d \times \mathbb{J}\Omega + \mathbb{J}\dot{\Omega}_d \end{aligned}$$

where K_d *is a positive definite matrix and* K_p *is a symmetric matrix with eigenvalues* $\{k_1, k_2, k_3\}$ *such that* $k_i + k_j > 0$ *for* $i \neq j$*.*

Then, for both choices of attitude error, the total energy $\phi(R_e) + \frac{1}{2}\|\Omega_e\|_{\mathbb{J}}^2$ *converges exponentially to zero from all initial conditions* $(R(0), \Omega(0))$ *such that*

$$\phi(R_e(0)) + \frac{1}{2}\|\Omega_e(0)\|_{\mathbb{J}}^2 \; < \; \min_{i \neq j}(k_i + k_j).$$

This lemma is a direct consequence of Theorem 4.2, except for the design of the feedforward control which is discussed in the next section. To the authors' knowledge, both control laws are novel: f_ℓ in the choice of velocity error, f_r in the expression of the feedforward control.

To illustrate the difference between the two velocity errors, we run simulations without the PD action. The reference trajectory is a 2π radians rotation about the vertical Z axis performed in 10 seconds with velocity profile of $2\pi(3t^2 - t)/100$ radians per second. The initial attitude error is a rotation of $\pi/4$ radians about the X axis. Both the angular velocity and the reference angular velocity are zero at time $t = 0$ and therefore the velocity error is zero for all times. Indeed, the latter property characterizes the two simulations completely: on the left side of Figure 4 we have $\dot{R}(t) = R(t)\widehat{\Omega}_d(t)$, on the right side $\dot{R}(t) = (R_d(t)\Omega_d(t))^\wedge R(t)$.

5.4 An underwater vehicle on the group of rigid motions $SE(3)$

In this section we extend the treatment of the attitude tracking problem to the group of rigid rotations and translations $SE(3) = SO(3) \times \mathbb{R}^3$. Motivated by recent interest in the area (Fossen 1994, Chapter 2) and (Leonard 1997), we focus on the idealized model of an underwater vehicle.

The configuration of the underwater vehicle (rigid body) is the rigid motion $g = (R, p)$ representing the position and attitude of a body frame with respect to an inertial frame. The kinematic equations are

$$\begin{aligned} \dot{R} &= R\widehat{\Omega} \\ \dot{p} &= RV, \end{aligned} \tag{26}$$

where $\xi = (\widehat{\Omega}, V) \in \mathfrak{se}(3) = \mathfrak{so}(3) \times \mathbb{R}^3$ is the body velocity expressed in the body frame. Introducing the homogeneous coordinates,

$$g = \begin{bmatrix} R & p \\ 0_{1\times 3} & 1 \end{bmatrix} \quad \text{and} \quad \xi = \begin{bmatrix} \widehat{\Omega} & V \\ 0_{1\times 3} & 0 \end{bmatrix},$$

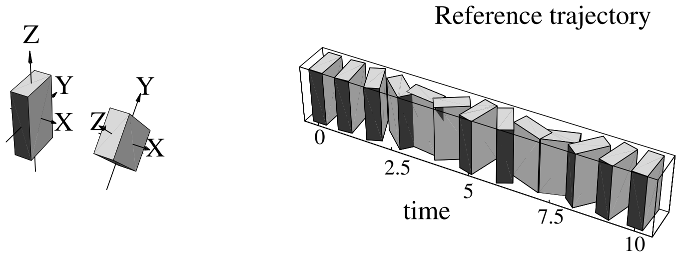

Figure 3: Bricks represent rotation matrices. On the left we depict initial reference attitude and initial error (i.e. a rotation of $\pi/4$ about the X axis). On the right we depict the reference attitude trajectory.

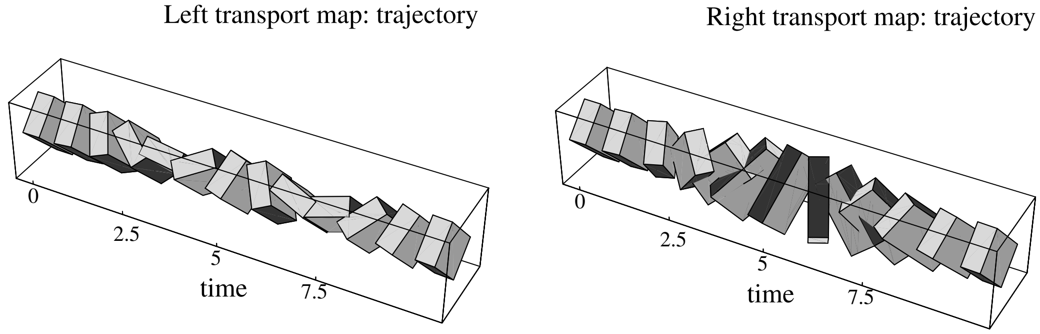

Figure 4: Trajectories in the closed loop help us compare feedforward policies. Left and right velocity errors are employed correspondingly on the left and right picture.

the kinematic equations read $\dot{g} = g\cdot\xi$. As described in Section 2.2, the matrix multiplication $(g\cdot)$ can be interpreted as the tangent map to the left translation on the Lie group $SE(3)$.

The motion of a rigid body in incompressible, irrotational and inviscid fluid satisfies the Euler-Lagrange equations with an inertia tensor which includes added masses and inertias, see (Leonard 1997). If the underwater vehicle is an ellipsoidal body with uniformly distributed mass, the kinetic energy of the body-fluid system is $\frac{1}{2}\Omega^T\mathbb{J}\Omega + \frac{1}{2}V^T\mathbb{M}V \equiv \frac{1}{2}\xi^T\mathbb{I}\xi$, where $\mathbb{M}$ and $\mathbb{J}$ are the (positive definite) mass and inertia matrices. The Kirchhoff equations describing the time evolution of the body velocity ξ are

$$\begin{aligned} \mathbb{J}\dot{\Omega} &= \mathbb{J}\Omega \times \Omega + \mathbb{M}V \times V \quad + \quad f_\Omega \\ \mathbb{M}\dot{V} &= \mathbb{M}V \times \Omega \qquad\qquad\quad + \quad f_V, \end{aligned} \tag{27}$$

where $f = [f_\Omega \ f_V] \in \mathfrak{se}(3)^*$ is the resultant generalized force acting on the body. As described in Section 2.3, the equations (27) are the Euler-Poincaré equations (10) for a simple mechanical system on a Lie group

$$\mathbb{I}\dot{\xi} = \mathrm{ad}^*_\xi\, \mathbb{I}\xi + f$$

where the adjoint operator on $\mathfrak{se}(3) = \mathbb{R}^6$ is

$$\mathrm{ad}_{(\Omega,V)} = \begin{bmatrix} \widehat{\Omega} & 0 \\ \widehat{V} & \widehat{\Omega} \end{bmatrix}.$$

Error functions

Let $\{g_d = (R_d, p_d), t \in \mathbb{R}_+\}$ denote the reference trajectory corresponding to a desired frame and let $\xi_d = (\Omega_d, V_d)$ denote the reference velocity expressed in the desired frame, that is $\dot{g}_d = g_d \cdot \xi_d$. As in the $SO(3)$ case, we design an error function φ by composing a group error $g_e(g, g_d)$ and a positive definite function $\phi : SE(3) \to \mathbb{R}$.

The group operation on $SE(3)$ provides us with right and left group errors

$$\begin{aligned} g_{e,r} &\triangleq g_d^{-1} g = \left(R_d^T R,\ R_d^T(p - p_d)\right), \\ g_{e,\ell} &\triangleq g g_d^{-1} = \left(R R_d^T,\ p - R R_d^T p_d\right). \end{aligned}$$

The group element $g_{e,r}$ is the relative motion from the body frame to the desired frame. Disregarding the group structure, two other group errors are

$$g_{e,1} \triangleq \left(R_d^T R,\ p - p_d\right) \quad \text{and} \quad g_{e,2} \triangleq \left(R R_d^T,\ R^T p - R_d^T p_d\right).$$

Next we design some positive definite functions on $SE(3)$. We set

$$\begin{aligned} \phi_1(R,p) &= \frac{1}{2}\operatorname{tr}(K_1(I_3 - R)) + \frac{1}{2} p^T K_2 p \quad \triangleq \quad \phi_1(R) + \frac{1}{2}\|p\|_{K_2}^2 \quad, \\ \phi_2(R,p) &= \phi_1(R) + \frac{1}{2}\|R^T p\|_{K_2}^2 \quad, \\ \phi_3(R,p) &= \phi_1(R) + \frac{1}{2}\|(I_3 + R^T)p\|_{K_2}^2 \quad, \end{aligned}$$

where the eigenvalues $\{k_1, k_2, k_3\}$ of the symmetric matrix K_1 satisfy $k_i + k_j > 0$ for $i \neq j$, and where K_2 is positive definite. The presence of matrix gains in both the attitude and position variables is useful in applications.

In Fig. 5 we attempt to portray these functions restricted to $SE(2)$, the group of rigid motions on the plane. We equip this space with an invariant metric (kinetic energy) of the form $J\omega^2 + m_x v_x^2 + m_y v_y^2$, where (ω, v_x, v_y) is the velocity in the body frame. Then we compute the gradient vector field for each of the three error functions and we draw their flow fields. The gains on rotational and translational components are chosen equal to (J, m_x, m_y).

Finally we design error functions by combining a group error g_e with a function ϕ. For all choices of g_e and ϕ, the resulting error function φ is quadratic with constant $\min_{i\neq j}(k_i + k_j)$, where $\{k_1, k_2, k_3\}$ are the eigenvalues of the matrix K_1. Since many combinations are possible, we report only the most instructive ones in the first column of Table 1. In the third column we characterize the error functions in terms of various properties. For example, we call φ *invariant* if it is invariant under changes in the inertial coordinate frame. Also recall that φ is symmetric if $\varphi(g, g_d) = \varphi(g_d, g)$. Additionally, we specify the frame in which the proportional gains K_1 and K_2 are expressed.

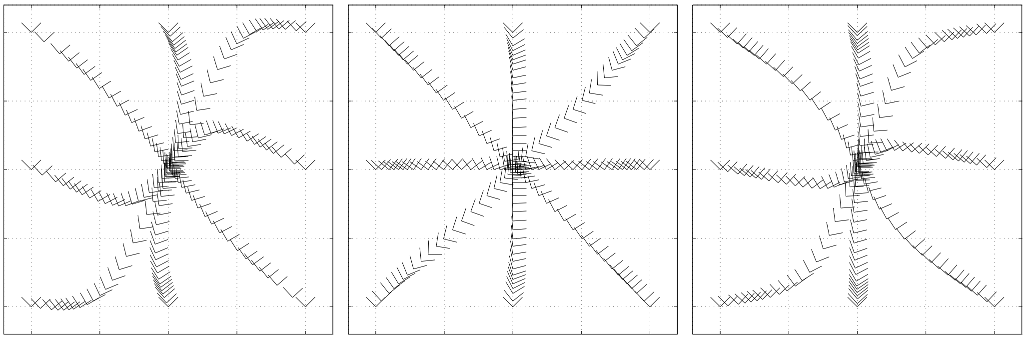

Figure 5: From left to right the functions ϕ_1, ϕ_2 and ϕ_3 are compared in terms of the flow of their gradient. Each frame on the plane represents a configuration on $SE(2)$.

Velocity errors

We start by recalling some kinematics (Murray et al. 1994, Chapter 2). We are interested in the adjoint map $\mathrm{Ad}_g : \mathfrak{se}(3) \to \mathfrak{se}(3)$ that transforms velocity vectors (elements in $\mathfrak{se}(3)$) from the body coordinate frame to the inertial coordinate frame. Identifying $\mathfrak{se}(3)$ with $\mathbb{R}^6$, this map is

$$\mathrm{Ad}_g = \mathrm{Ad}_{(R,p)} = \begin{bmatrix} R & 0 \\ \widehat{p}R & R \end{bmatrix}.$$

More generally, since $g_{e,r} = g_d^{-1} g$ is the relative motion from the body frame to the desired frame, the reference velocity in the desired frame ξ_d is expressed in the body frame via the map $\mathrm{Ad}_{g^{-1}g_d} = \mathrm{Ad}_{g_{e,r}^{-1}}$. These ideas lead to a natural definition of velocity error as

$$\xi_{e,r} = \xi - \mathrm{Ad}_{g_{e,r}^{-1}} \xi_d,$$

where the body and the reference velocities are expressed in the same frame. We call $\xi_{e,r}$ the right velocity error. This is a useful definition since, with the aid of the homogeneous representation and some matrix algebra, we have

$$\begin{aligned} \dot{g}_{e,r} &= g_d^{-1} \left(\frac{d}{dt} g \right) + \left(\frac{d}{dt} g_d^{-1} \right) g = g_d^{-1} g \cdot \xi - \xi \cdot g_d^{-1} g \\ &\equiv g_{e,r} \left(\xi - \mathrm{Ad}_{g_{e,r}^{-1}} \xi_d \right). \end{aligned}$$

Therefore, every error function that relies on the right group error $g_{e,r}$ is compatible with the right velocity error.

More generally the adjoint map is useful in describing transport maps. In what follows, we parametrize the set of transport maps with the set of change of frames, that is with $SE(3)$. For each transport map $\mathcal{T}$, we call *transport element* the unique motion $\tau \in SE(3)$ such that

$$\dot{g} - \mathcal{T} \dot{g}_d = g \cdot (\xi - \mathrm{Ad}_\tau \xi_d).$$

Table 1: Error functions and transport elements on $SE(3)$.

Error function	Transport element	Comments
$\phi_1(R_d^T R) + \frac{1}{2}\|p - p_d\|^2_{K_2}$	$(R^T R_d, 0)$	$\phi_1(g_{e,1})$, not invariant, symmetric, gains expressed in inertial frame.
$\phi_1(R_d^T R) + \frac{1}{2}\|R_d^T(p - p_d)\|^2_{K_2}$	$g^{-1}g_d$	$\phi_1(g_{e,r})$, invariant, not symmetric, gains expressed in reference frame.
$\phi_1(R_d^T R)+\frac{1}{2}\|R^T(p - p_d)\|^2_{K_2}$	$g^{-1}g_d$	$\phi_2(g_{e,r})$, invariant, not symmetric, gains expressed in body frame.
$\phi_1(R_d^T R)+\frac{1}{2}\|(R^T+R_d^T)(p - p_d)\|^2_{K_2}$	$g^{-1}g_d$	$\phi_3(g_{e,r})$, invariant, symmetric.
$\phi_1(RR_d^T) + \frac{1}{2}\|(R+R_d)(p - p_d)\|^2_{K_2}$	$(I_3, 0)$	$\phi_3(g_{e,\ell})$, not invariant, symmetric.
$\phi_1(RR_d^T) + \frac{1}{2}\|R^T p - R_d^T p_d\|^2_{K_2}$	$(I_3, R_d^T p_d - R^T p)$	$\phi_1(g_{e,2})$, not invariant, symmetric.

In the Table 1, we report compatible transport elements for each error function. For each couple (φ, τ), the compatibility is verified with some straightforward algebra. Note that the choice of τ depends only on the group error g_e employed to define φ.

Control laws

We here summarize the ideas exposed so far and design a proportional derivative feedback. Additionally we devise a set of feedforward control laws by means of a minor extension of Theorem 4.2. Let $\langle\cdot,\cdot\rangle$ denote the natural pairing between $\mathfrak{se}(3)$ and its dual $\mathfrak{se}(3)^*$, and let (φ, τ) be a compatible pair of error function and transport element. We define $f_\mathrm{P}, f_\mathrm{D} \in \mathfrak{se}(3)^*$ by means of

$$\begin{aligned} \langle f_\mathrm{P}\ ,\ \eta\rangle &= -\mathcal{L}_{(g\cdot\eta,\ 0)}\varphi(g, g_d), && \forall\eta \in \mathfrak{se}(3),\\ f_\mathrm{D} &= -K_d\left(\xi - \mathrm{Ad}_\tau\, \xi_d\right), \end{aligned} \tag{28}$$

where $K_d : \mathfrak{se}(3) \mapsto \mathfrak{se}(3)^*$ is a self-adjoint (symmetric) and positive definite. For example, from the first row of Table 1 we compute

$$f_\mathrm{P} + f_\mathrm{D} = -\begin{bmatrix}\mathrm{skew}(K_1 R_e)^\vee \\ R^T K_2 p_e\end{bmatrix} - K_d\begin{bmatrix}\Omega - R_e^T\Omega_d \\ V - R_e^T V_d\end{bmatrix},$$

where $(R_e, p_e) = (R_d^T R,\ p - p_d)$, and likewise from the third row

$$f_\mathrm{P} + f_\mathrm{D} = -\begin{bmatrix}\mathrm{skew}(K_1 R_e)^\vee + (K_2 p_e)\times p_e \\ R^T K_2 p_e\end{bmatrix} - K_d\begin{bmatrix}\Omega - R_e^T\Omega_d \\ V - R_e^T\left(V_d + \Omega_d \times p_e\right)\end{bmatrix},$$

where $(R_e, p_e) = (R_d^T R,\ R_d^T(p - p_d))$. Next, we define a family of feedforward control laws as

$$f_{\mathrm{FF}} = -\operatorname{ad}^*_{(\operatorname{Ad}_\tau \xi_d)} \xi + \mathbb{I}\frac{d}{dt}(\operatorname{Ad}_\tau \xi_d) + S_\tau(\xi_e, \xi_d), \tag{29}$$

where the bilinear operator $S_\tau : \mathfrak{se}(3) \times \mathfrak{se}(3) \mapsto \mathfrak{se}(3)^*$ is skew symmetric with respect to its first argument, i.e. it holds

$$\langle S_\tau(\xi_e, \eta), \xi_e \rangle = 0 \qquad \forall \eta \in \mathfrak{se}(3). \tag{30}$$

For example, corresponding to $\tau = g^{-1} g_d$, (second, third and fourth row in Table 1, right group error $g_{e,r}$) and $\tau = (I_3, 0)$, (fifth row in Table 1, left group error $g_{e,\ell}$), an appropriate choice of S_τ leads to the simple feedforward controls:

$$\begin{aligned} f_{\mathrm{FF},r} &= -\operatorname{ad}^*_\xi \mathbb{I}\operatorname{Ad}_{g^{-1}g_d} \xi_d + \mathbb{I}\operatorname{Ad}_{g^{-1}g_d} \dot{\xi}_d, \\ f_{\mathrm{FF},\ell} &= -\operatorname{ad}^*_{\xi_d} \mathbb{I}\xi + \mathbb{I}\dot{\xi}_d. \end{aligned}$$

Note that, with the corresponding definition of Ad and ad operators, these choices are the same employed for the attitude tracking problem in Lemma 5.2.

Lemma 5.3. *Consider the system in equation* (26) *and* (27). *Let* $\{g_d(t), t \in \mathbb{R}_+\}$ *denote the reference trajectory and let* $\xi_d = g_d^{-1}\dot{g}_d \in \mathfrak{se}(3)$ *denote its bounded body-fixed velocity. From Table 1, let* φ *be a quadratic error function with constant* $\min_{i\neq j}(k_i + k_j)$, *and let* τ *be a compatible transport element. Also, let* S_τ *be a bilinear operator satisfying* (30), *and according to equations* (28) *and* (29), *let*

$$f = f_{\mathrm{P}} + f_{\mathrm{D}} + f_{\mathrm{FF}} \quad \in \mathfrak{se}(3)^*.$$

Then the total energy $\varphi(g, g_d) + \frac{1}{2}\|\xi - \operatorname{Ad}_\tau \xi_d\|^2_{\mathbb{I}}$ *converges exponentially to zero from all initial conditions* $(g(0), \xi(0))$ *such that*

$$\varphi\big(g(0), g_d(0)\big) + \frac{1}{2}\|\xi(0) - \operatorname{Ad}_{\tau(0)} \xi_d(0)\|^2_{\mathbb{I}} < \min_{i\neq j}(k_i + k_j).$$

In what follows we present a sketch of the proof. First, the proportional and derivative feedback are devised according to the design procedure in Section 3, so that the only difference with the design in Theorem 4.2 regards the feedforward control. In fact, the latter theorem can be extended as follows.

Lemma 5.4. *Let the map* $S_{(q,r)} : T_qQ \times T_rQ \to T_q^*Q$ *satisfy*

$$S_{(q,r)}(X_q, Y_r) \cdot X_q = 0$$

for all $X_q \in T_qQ$ *and* $Y_r \in T_rQ$. *Also consider the boundedness condition:*

$$\sup_{(q,r)\in Q\times Q} \|\nabla \mathcal{T}_{(q,r)} + M_q^{-1} S\|_M < \infty. \tag{B2'}$$

The statement of Theorem 4.2 holds true if we set $F = F_{\mathrm{PD}} + F_{\mathrm{FF}} + S(\dot{e}, \dot{r})$ *instead of* $F = F_{\mathrm{PD}} + F_{\mathrm{FF}}$, *and if we assume* (B2') *instead of condition* (B2).

The proof of this statement is a straightforward modification of the proof of Theorem 4.2. Thus we only need to show that feedforward action f_{FF} in equation (29) differs from the one defined in the main theorem, call it f'_{FF}, by a skew symmetric operator. Indeed, using some of the tools introduced in Section (2.2), we compute

$$f'_{\mathrm{FF}} = \mathbb{I}\,{}_{\mathfrak{g}}\nabla_{\xi}\left(\mathrm{Ad}_t\,\xi_d\right) + \frac{d}{dt}\left(\mathrm{Ad}_t\,\xi_d\right),$$

where the map ${}_{\mathfrak{g}}\nabla : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is defined in equation (8), and $f_{\mathrm{FF}} = f'_{\mathrm{FF}}$ when the operator S_τ is defined as

$$S_\tau(\xi_e, \xi_d) = \frac{1}{2}\left(\mathbb{I}[\xi_e, \mathrm{Ad}_t\,\xi_d] + \mathrm{ad}^*_{(\mathrm{Ad}_t\,\xi_d)}\,\mathbb{I}\xi_e - \mathrm{ad}^*_{\xi_e}(\mathrm{Ad}_t\,\xi_d)\right).$$

6 Summary and Conclusions

This work unveils the geometry and the mechanics of the tracking problem for fully actuated Lagrangian systems. The design process in Section 3 allows us to characterize in an intrinsic way a tracking controller. The basic answered questions concern how to define configuration and velocity errors and how to compute the feedforward control. Almost global stability and local exponential convergence are proven in full generality. Our framework successfully unifies a variety of examples: a robot manipulator on the Euclidean space $\mathbb{R}^n$, a pointing device on the two sphere $\mathbb{S}^2$, a satellite on the group of rotations $SO(3)$ and an underwater vehicle on the group of rigid motions $SE(3)$. Case by case, we provide new insight into previous results and introduce novel viewpoints and control laws.

Relying on concepts from Riemannian geometry this work provides coordinate free design techniques for nonlinear mechanical systems. Other recent papers on modeling (Bloch & Crouch 1995), controllability (Lewis & Murray 1997), interpolation (Noakes, Heinzinger & Paden 1989) and dynamic feedback linearization (Rathinam & Murray 1998) share the same theoretical tools. A parallel avenue of research relies on the Hamiltonian formulation of mechanical systems, see for example (Nijmeijer & van der Schaft 1990, Chapter 12) and (Simo, Lewis & Marsden 1991). All these geometric techniques are a promising starting point in the design of control policies for underactuated systems.

Acknowledgments

The first author would like to thank Dr. Andrew D. Lewis and Professor Jerrold E. Marsden for insightful discussions on geometric mechanics. The authors would like to thank the reviewers for the useful comments and corrections.

References

Arnold, V. I. (1989). *Mathematical Methods of Classical Mechanics*, GTM 60, second edn, Springer Verlag, New York, NY.

Bedrossian, N. S. & Spong, M. W. (1995). Feedback linearization of robot manipulators and Riemannian curvature, *Journal of Robotic Systems* **12**(8): 541–552.

Bloch, A. M. & Crouch, P. E. (1995). Nonholonomic control systems on Riemannian manifolds, *SIAM Journal of Control and Optimization* **33**(1): 126–148.

Boothby, W. M. (1986). *An Introduction to Differentiable Manifolds and Riemannian Geometry*, second edn, Academic Press, New York.

Bullo, F. & Murray, R. M. (1997). Trajectory tracking for fully actuated mechanical systems, *European Control Conference*, Brussels, Belgium. Available electronically via `http://www.cds.caltech.edu/~bullo`.

Do Carmo, M. P. (1992). *Riemannian Geometry*, Birkhäuser, Boston.

Egeland, O. & Godhavn, J.-M. (1994). Passivity-based adaptive attitude control of a rigid spacecraft, *IEEE Transactions on Automatic Control* **39**(4): 842–846.

Etkin, B. (1982). *Dynamics of Flight: Stability and Control*, John Wiley and Sons, New York, NY.

Fossen, T. I. (1994). *Guidance and Control of Ocean Vehicles*, John Wiley and Sons, New York, NY.

Kobayashi, S. & Nomizu, K. (1963). *Foundations of Differential Geometry*, Vol. I, Interscience Publishers, New York, NY.

Koditschek, D. E. (1989). The application of total energy as a Lyapunov function for mechanical control systems, *in* J. E. Marsden, P. S. Krishnaprasad & J. C. Simo (eds), *Dynamics and Control of Multibody Systems*, Vol. 97, AMS, pp. 131–157.

Leonard, N. E. (1997). Stability of a bottom-heavy underwater vehicle, *Automatica* **33**(3): 331–346.

Lewis, A. D. (1995). *Aspects of Geometric Mechanics and Control of Mechanical Systems*, PhD thesis, California Institute of Technology, Pasadena, California, USA. Technical report CIT-CDS 95-017, available electronically via `http://avalon.caltech.edu/cds/`.

Lewis, A. D. & Murray, R. M. (1997). Controllability of simple mechanical control systems, *SIAM Journal of Control and Optimization* **35**(3): 766–790.

Luh, J., Walker & Paul, R. (1980). Resolved acceleration control, *IEEE Transactions on Automatic Control* **25**(3): 468–474.

Marsden, J. E. & Ratiu, T. S. (1994). *Introduction to Mechanics and Symmetry*, Springer Verlag, New York, NY.

Meyer, G. (1971). Design and global analysis of spacecraft attitude control systems, *Technical report*, NASA. Tech. Report R-361.

Murray, R. M. (1995). Nonlinear control of mechanical systems: a Lagrangian perspective, *IFAC Symposium on Nonlinear Control Systems (NOLCOS)*, Lake Tahoe, CA, pp. 378–389.

Murray, R. M., Li, Z. X. & Sastry, S. S. (1994). *A Mathematical Introduction to Robotic Manipulation*, CRC Press, Boca Raton, Florida.

Nijmeijer, H. & van der Schaft, A. (1990). *Nonlinear Dynamical Control Systems*, Springer Verlag, New York, NY.

Noakes, L., Heinzinger, G. & Paden, B. (1989). Cubic splines on curved spaces, *IMA Journal of Mathematical Control & Information* **6**: 465–473.

Rathinam, M. & Murray, R. M. (1998). Configuration flatness of Lagrangian systems underactuated by one control, *SIAM Journal of Control and Optimization* . To appear. A short version appeared at CDC, 1996 in Kobe, Japan.

Simo, J. C., Lewis, D. R. & Marsden, J. E. (1991). Stability of relative equilibria I: The reduced energy momentum method, *Archive for Rational Mechanics and Analysis* **115**: 15–59.

Slotine, J.-J. E. & Di Benedetto, M. D. (1990). Hamiltonian adaptive control of spacecraft, *IEEE Transactions on Automatic Control* **35**: 848–852.

Slotine, J.-J. E. & Li, W. (1989). Composite adaptive control of robot manipulators, *Automatica* **25**: 509–519.

Takegaki, M. & Arimoto, S. (1981). A new feedback method for dynamic control of manipulators, *Journal of Dynamic Systems, Measurement, and Control* **102**: 119–125.

Tsiotras, P. & Longuski, Y. M. (1994). Spin-axis stabilization of symmetrical spacecraft with 2 control torque, *Systems & Control Letters* **23**(6): 395–402.

Wen, J. T.-Y. & Bayard, D. S. (1988). A new class of control laws for robotic manipulators. Part I: Non–adaptive case, *International Journal of Control* **47**(5): 1361–1385.

Wen, J. T.-Y. & Kreutz-Delgado, K. (1991). The attitude control problem, *IEEE Transactions on Automatic Control* **36**: 1148–1162.

A The Error Function on $SO(3)$

We here study the modified trace function on $SO(3)$ introduced in Section 5.3. We refer to (Koditschek 1989) for additional details. Given a 3×3 symmetric matrix K, recall that we defined $\phi : SO(3) \to \mathbb{R}_+$ as

$$\phi(R) \triangleq \frac{1}{2} \operatorname{tr} \big(K(I_3 - R)\big),$$

and, given any 3×3 matrix A, we defined $\operatorname{skew}(A) = \frac{1}{2}(A - A^T)$.

Lemma A.1. *Let the eigenvalues $\{k_1, k_2, k_3\}$ of the matrix K satisfy $k_i + k_j > 0$ for all $i \neq j$, and define $\mathrm{d}\,\phi \in \mathfrak{so}(3)^*$ such that $\dot{\phi} = \mathrm{d}\,\phi \cdot (R^T \dot{R})$. It holds*

1. *$\phi(R) = \phi(R^T) \geq 0$ and $\phi(R) = 0$ if and only if $R = I_3$,*
2. *$\mathrm{d}\,\phi = \operatorname{skew}(KR)$, and*
3. *for all $\epsilon > 0$ there exist $b_1 \geq b_2 > 0$ such that $\phi(R) < \min_{i \neq j}(k_i + k_j) - \epsilon$ implies $b_1 \| \mathrm{d}\,\phi \|^2 \ \geq \ \phi \ \geq \ b_2 \| \mathrm{d}\,\phi \|^2$,*

In addition we have the following coordinate expressions. Let R be a rotation of angle θ about the unit vector $\mathbf{k}$ and define the unit quaternion representation of R by $q^T = [q_0 \ q_1 \ q_2 \ q_3] \equiv [q_0, \ \mathbf{q}_v^T]$, where

$$q_0 = \cos(\theta/2) \quad \textit{and} \quad \mathbf{q}_v = \sin(\theta/2)\mathbf{k}.$$

Finally define $K^{[2]}$ as the matrix with the same eigenvectors as K and with eigenvalues $\{(k_2 + k_3), (k_1 + k_3), (k_1 + k_2)\}$. Then it holds

4. *$\phi(R) = \|\mathbf{q}_v\|^2_{K^{[2]}}$, and*
5. *$\mathrm{d}\,\phi = \frac{1}{2}\big(q_0 K^{[2]}\mathbf{q}_v - \widehat{\mathbf{q}}_v K^{[2]}\mathbf{q}_v\big)\widehat{\ }$.*

Proof. We start by proving (2). It holds:

$$\dot{\phi} = \frac{1}{2} \operatorname{tr} \big(K(-\dot{R})\big) = -\frac{1}{2} \operatorname{tr} \big(KR\, R^T \dot{R}\big).$$

Recall that the linear space of 3×3 matrix decomposes into the direct sum of symmetric and skew symmetric matrices with respect to the trace inner product. If we let $\operatorname{skew}(A) = \frac{1}{2}(A - A^T)$ and $\operatorname{sym}(A) = \frac{1}{2}(A + A^T)$, it holds

$$\begin{aligned}\dot{\phi} &= -\frac{1}{2} \operatorname{tr} \Big(\big(\operatorname{skew}(KR) + \operatorname{sym}(KR)\big)\ (R^T \dot{R})\Big) \\ &= -\frac{1}{2} \operatorname{tr} \big(\operatorname{skew}(KR)\ (R^T \dot{R})\big).\end{aligned}$$

Finally we recall the matrix pairing between $\mathfrak{so}(3)$ and its dual: $\alpha \cdot \xi = \frac{1}{2} \operatorname{tr}(\alpha^T \xi)$, where α is in $\mathfrak{so}(3)^*$ and ξ in $\mathfrak{so}(3)$. This pairing corresponds to the standard pairing $\widehat{x} \cdot \widehat{y} = x^T y$ for all x, y in $\mathbb{R}^3$. This leads us to

$$\dot{\phi} = \operatorname{skew}(KR) \cdot (R^T \dot{R}),$$

which proves (2). Next we introduce the unit quaternion representation. By Rodrigues' formula, it holds that $R = I_3 + 2q_0\widehat{\mathbf{q}}_v + 2\widehat{\mathbf{q}}_v^2$. Hence we have

$$\begin{aligned}\phi_{SO(3)} &= -\operatorname{tr}(Kq_0\mathbf{q}_v) - \operatorname{tr}(K\,\widehat{\mathbf{q}}^2)\\ &= -\operatorname{tr}(K\,\widehat{\mathbf{q}}^2) \;\equiv\; \mathbf{q}^T K^{[2]}\mathbf{q},\end{aligned}$$

where the second equality can be proved in coordinates. This proves (4) and (5) can be verified by recalling the kinematic equation for R in terms of the unit quaternion representation. Regarding (1), it is straightforward that $\phi(R) = \phi(R^T) \geq 0$. Also if $\phi(R) = 0$, then $\mathbf{q}_v$ is the zero vector and R is the identity matrix.

Last, we prove the claim in (3), that is that ϕ is quadratic with constant equal to the minimum eigenvalue of $K^{[2]}$. Since the two terms in the equation in (5) are orthogonal, we have

$$2\|\operatorname{d}\phi\|^2 = \|q_0K^{[2]}\mathbf{q}_v\|^2 + \|\widehat{\mathbf{q}}_vK^{[2]}\mathbf{q}_v\|^2.$$

Since $\|\widehat{\mathbf{q}}_vK^{[2]}\mathbf{q}_v\|^2 \leq \|\mathbf{q}_v\|^2\|K^{[2]}\mathbf{q}_v\|^2 \leq \lambda_{\max}\big(K^{[2]}\big)\|\mathbf{q}_v\|^2_{K^{[2]}}$, we have

$$\begin{aligned}\phi &= \|\mathbf{q}_v\|^2_{K^{[2]}}\\ &\geq \frac{1}{2}q_0^2\|\mathbf{q}_v\|^2_{K^{[2]}} + \frac{1}{2}\|\mathbf{q}_v\|^2_{K^{[2]}}\\ &\geq \frac{1}{2}q_0^2\|\mathbf{q}_v\|^2_{K^{[2]}} + \frac{1}{2\lambda_{\max}\big(K^{[2]}\big)}\|\widehat{\mathbf{q}}_vK^{[2]}\mathbf{q}_v\|^2\\ &\geq \min\big(1, 1/\lambda_{\max}\big(K^{[2]}\big)\big)\|\operatorname{d}\phi\|^2.\end{aligned}$$

This proves one direction of the bound. Next, recall that we are assuming that, given an $\epsilon > 0$, we have the inequality $\phi \leq \lambda_{\min}\big(K^{[2]}\big) - \epsilon$. Hence it holds that

$$\exists\epsilon_1 > 0 \qquad \text{s.t.} \qquad \|\mathbf{q}_v\|^2 \leq 1 - \epsilon_1,$$

and this implies that

$$\exists\epsilon_1 > 0 \qquad \text{s.t.} \qquad \|q_0\|^2 \geq \epsilon_1.$$

However it holds that $\|\operatorname{d}\phi\|^2 \geq \frac{1}{2}\|q_0K^{[2]}\mathbf{q}_v\|^2 = \frac{1}{2}q_0^2\phi$, and therefore

$$\|\operatorname{d}\phi\|^2 \geq \frac{1}{2}\epsilon_1\phi.$$

This completes the proof of (3) and of the whole Lemma. $\square$