Motion Control along Relative Equilibria

Thesis by

Nikolaj Nordkvist

Supervisors:
Poul G. Hjorth
Francesco Bullo

Ph.D. degree awarded January 2008

Department of Mathematics
Technical University of Denmark

Acknowledgements

First of all I would like thank my supervisors Poul G. Hjorth and Francesco Bullo.

Poul was also supervisor on my bachelor project and later on my master project. After finishing my master degree Poul asked me to do a Ph.D. with him as supervisor. Poul has never shown any doubt, not even when I felt insecure of my academic potential, that he believed I was capable of attaining the Ph.D. degree. His broad range of knowledge has been a continuous source of inspiration and his support has been endless.

After attending a summer school where Francesco was giving lectures he invited me to come to Santa Barbara to do research in collaboration with him; for this opportunity I am grateful. Later he offered to become co-supervisor for my project. He has been persistent in the guidance of me, even after I left Santa Barbara, and with his overwhelming knowledge in control theory he has always been able to get me unstuck when needed.

I would also like to thank all the great people from Francesco's lab who, with their friendliness, made me feel welcome at UCSB. In particular thanks go to Sara Susca, Anurag Ganguli, Ketan Savla, Chunkai Gao, Stephen L. Smith, Shaunak D. Bopardikar, Ruggero Carli, and Sonia Martínez. Thanks is also in order for Jeff Oakes who helped me out numerous times with computer issues while at UCSB. Furthermore I want to express my gratitude to the wonderful people I shared the house with at 6048 Passeo Palmilla and to the many people hanging out at our house; rest in piece Eric. Though far away from home I always felt at home.

I also want to thank Marie-Louise Højlund Rasmussen and Mikael Sonne Hansen, with whom I have shared office with at DTU, for the great company and many interesting discussions. Thanks is also in order for the department Ph.D. confidant Tom Høholdt for always looking out for, and caring about the well-being of, the Ph.D. students.

My parents have always encouraged me to do with my life what I wanted and have always believed in me—I have never felt any pressure to pursue a certain career path—and for that I am thankful. Also, I would like to thank my brother Mikkel for always helping me out with all kinds of bureaucracy issues, e.g., filling out tax forms. Without his help, in particular in connection with my stay in Santa Barbara, I would have experienced a lot more moments of utter despair during the course of my Ph.D.

I am indebted to my great crew of friends, "banden," back in my hometown Thisted. Without them I would probably have been beaten up a lot more by rednecks not liking my, back then, long hair. Most of these people have now moved to Copenhagen and I am happy we still stick together after all these years. I am sure my years of study in Copenhagen would have been a lot harder without the great company of these guys.

Resume

Emnet for denne afhandling er kontrol af mekaniske systemer under en bevægelse, der kaldes en relativ ligevægtskurve. En sådan bevægelse er karakteriseret ved at hastigheden, i et koordinatsystem i legemet, er konstant. Et stift legeme, der roterer med konstant hastighed omkring en af hovedakserne, er et eksempel på en relativ ligevægtskurve.

I afhandlingen fokuseres på såkaldte simple mekaniske kontrolsystemer på Lie grupper. Denne klasse af systemer er defineret ved følgende: konfigurationsmangfoldigheden er en Lie gruppe, den totale energi er givet ved den kinetiske energi (d.v.s. ingen potentiel energi) og den kinetiske energi samt kontrolkrafterne er invariante i en bestemt betydning.

Afhandlingen indeholder to hovedresultater. Først udledes tilstrækkelige betingelser, af algebraisk karakter, under hvilke et simpelt mekanisk kontrolsystem på en Lie gruppe er lokalt kontrollerbart langs en relativ ligevægtskurve. Disse betingelser omfatter de velkendte betingelser for lokal kontrollerbarhed af et ligevægtspunkt for et simpelt mekanisk system på en Lie gruppe. Dernæst præsenteres en ny kontrolalgoritme for systemer med færre kontrolkrafter end frihedsgrader. Forudsat nogle antagelser er opfyldt, beregner denne algoritme kontrolkræfter, der får systemet til at accelerere, decelerere eller stabiliseres langs en relativ ligevægtskurve; valget af bevægelse bestemmes af fortegnet af en parameter i algoritmen. Algoritmen anvendes konkret på et stift legeme i planen samt en satellit med to kraftmomenter.

Abstract

The subject of this thesis is control of mechanical systems as they evolve along the steady motions called relative equilibria. These trajectories are of interest in theory and applications and have the characterizing property that the system's body-fixed velocity is constant. For example, constant-speed rotation about a principal axis is a relative equilibrium of a rigid body in three dimensions.

We focus our study on simple mechanical control systems on Lie groups, i.e., mechanical systems with the following properties: the configuration manifold is a matrix Lie group, the total energy is equal to the kinetic energy (i.e., no potential energy is present), and the kinetic energy and control forces both satisfy an invariance condition.

The novel contributions of this thesis are twofold. First, we develop sufficient conditions, algebraic in nature, that ensure that a simple mechanical control system on a Lie group is locally controllable along a relative equilibrium. These conditions subsume the well-known local controllability conditions for equilibrium points. Second, for systems that have fewer controls than degrees of freedom, we present a novel algorithm to control simple mechanical control systems on Lie groups along relative equilibria. Under some assumptions, we design iterative small-amplitude control forces to accelerate along, decelerate along, and stabilize relative equilibria. The technical approach is based upon perturbation analysis and the design of inversion primitives and composition methods. We finally apply the algorithms to a planar rigid body and a satellite with two thrusters.

Contents

Chapter 1

Introduction

Differential geometry applied to the analysis of mechanical systems and, in particular, to nonlinear control of mechanical systems, provides a fruitful way to gain insight into the intrinsic properties of the system, such as a number of control and controllability properties. Control problems for mechanical systems are known to be challenging when the number of independent control actuators is strictly less than the degrees of freedom of the system.

The main focus of this work is motion control along relative equilibria for simple mechanical control systems on Lie groups. A simple mechanical control system on a Lie group is a mechanical system for which the configuration manifold is a matrix Lie group, the kinetic energy and the control forces are invariant under the application of the group action, and the total energy is equal to the kinetic energy. A relative equilibrium is a steady motion for which the body-fixed velocity is constant when applying no control forces. For example, the configuration of a satellite is the matrix Lie group $SO(3)$ and a steady motion about any of its principal axes is a relative equilibrium.

Using geometry, primarily Lie group theory, we develop novel results regarding local controllability along relative equilibria for simple mechanical control systems on Lie groups. Specifically, we obtain two sets of results. First, we establish a theorem giving sufficient conditions, of algebraic nature, for a mechanical system on a Lie group to be locally controllable along a relative equilibrium. Second, for simple mechanical control systems on Lie groups with fewer control actuators than degrees of freedom, we design an algorithm producing acceleration along, deceleration along, and stabilization of a relative equilibrium using small-amplitude control forces.

In this chapter we start out with a short description of the history of the theory of mechanical control systems. We then describe the main contributions of this thesis. We end with a short outline of the thesis.

1.1 Geometry, Nonlinear Control, and Mechanics

Since Sir Isaac Newton published the ground-breaking "Philosophiae Naturalis Principia Mathematica" in 1687, the mathematical theory of mechanics has continuously attracted tremendous scientific interest. Remarkable breakthroughs were the introduction of Lagrangian mechanics in 1788 and Hamiltonian mechanics in 1833. The theory of differential geometry was established in the beginning of the 20th century, but it was not until 1967 that the first book in English treating mechanics in a geometric manner was published by Abraham and Marsden; for a later revised edition see [1]. Another classic text in the field is the work by Arnol'd [4] first published in English in 1978. Geometric mechanics has since been an active field of research in both the Lagrangian setting, using calculus of variations and Riemannian geometry, as well as the Hamiltonian setting, using symplectic geometry and Poisson geometry. A description of some of the modern theory can be found in [31].

The use of geometry in nonlinear control began in the late 1970s and important early contributions include [10], [18], [23], and [41]. The modern nonlinear control theory now relies on concepts from differential geometry: the aim is to provide intrinsic descriptions of various control theory concepts and to avoid arbitrary choices of coordinates. The modern geometric approach to control theory is described in the books [3] [21], [22], [25], [34], and [39].

The paper [9] from 1977 by Brockett is one of the earliest accounts where the differential geometric link between mechanics and control theory is stressed. During the 1980s there was only limited research activity on mechanical control theory; a prominent work being a series of papers on control theory for Hamiltonian systems by van der Schaft, see chapter 12 in the book [34] for an account of this theory. From around 1990 until today, the field gained interest and since then much new insight has been gained and sophisticated theoretical results have emerged. The books [6] and [13] describe some of these approaches.

1.2 Contributions of this Thesis

The contributions of this thesis are twofold. The first contribution is a general result providing sufficient conditions for a simple mechanical control system on a Lie group to be locally controllable along a relative equilibrium. The second result is the design of a control algorithm to compute control inputs to speed up a system along a relative equilibrium. In other words, the first result is an existence result, whereas the second is a constructive result. Though closely related in nature, the analysis leading to the two results differs considerably.

Local Controllability along a Relative Equilibrium

In [19] a result giving sufficient conditions for small time local controllability of general nonlinear control systems was conjectured. A stronger version of this result was later proved in [41]. Finally in [5] this approach was extended to address local controllability problems along an uncontrolled reference trajectory. This latter work contains the strongest known theorems providing sufficient conditions for local controllability along trajectories.

In [29] the main theorem of [41] was used to give local controllability results for mechanical control systems whose Lagrangian is kinetic energy, given by a Riemannian metric, minus potential energy. In [11], these results were used to give sufficient conditions for local controllability results for simple mechanical control systems on Lie groups. In particular, sufficient conditions for a system to be locally accessible at zero velocity, locally configuration accessible, small-time local controllable at zero velocity, and small-time local configuration controllable are given. These results are all for zero initial velocity. The sufficient conditions involves only algebraic analysis of the fixed input vectors defining the control directions and are much simpler to verify than the general conditions in [41].

In Chapter 4 we apply the results in [5] to prove a new proposition regarding local controllability along a relative equilibrium for a simple mechanical control system on a Lie group. As for the zero velocity results the sufficient conditions to ensure local controllability along a relative equilibrium are algebraic. To be more precise the sufficient conditions requires examining the space of symmetric products and the space of Lie brackets of the fixed input vector fields defining the control directions. In the special case when the relative equilibrium in fact is an equilibrium the result reduce to the proposition in [11] regarding small-time local controllability at zero velocity.

Motion Control Algorithm along a Relative Equilibrium

In the design of controls for a mechanical system the number of control forces is an important factor. If the system has as many control actuators as degrees of freedom it is called fully actuated; otherwise it is called underactuated. The motivation for studying underactuated systems is twofold; it gives rise to other design possibilities than a fully actuated system and it is appropriate in the situation of an actuator failure, meaning that such an analysis improves robustness to actuator failures which, e.g., is crucial in case the system is an a hazardous invironment such as outer space for a satellite.

Extensive research has focused on underactuated mechanical systems, especially in the context of controlled Lagrangians and Hamiltonians, e.g., see [7], [36], and subsequent works. In [11] motion control algorithms to reconfigure and exponentially stabilize simple mechanical systems on Lie groups using small amplitude periodic forcing are proposed. These algorithms are also applicable, under some conditions, in the case there are fewer actuators than degrees of freedom. The constructive approach is the same as in [27] and [28] where it is applied to a class of kinematic systems on Lie groups. The method is similar to that applied in [38] and [37] to different classes of mechanical systems. The results of [11] was later in [32] extended to include the more general class of mechanical systems where the configuration manifold has a principal bundle structure and the kinetic energy is given by a Riemannian metric. Another approach to motion planning of mechanical systems is to use oscillatory controls combined with an analysis using averaging theory, see for example [17], [42], and [44].

Less research has been done on controlling systems along relative equilibria; a related spin-up problem of a rolling ball is considered in [20]. In case the relative equilibrium is aligned with one of the vectors defining the directions of control, the theory of kinematic reductions can be applied to generate motion along a relative equilibrium. For an account of the theory of kinematic reduction see [13] or the series of papers [12], [14], and [33].

Using an approach resembling that of [11], Chapter 5 presents a new motion control algorithm for an invariant class of underactuated simple mechanical control systems on Lie groups. This motion algorithm produces control forces which make the system accelerate along, decelerate along, and stabilize a relative equilibrium; which type of motion is determined by the sign of a parameter in the algorithm. The main limitation of the algorithm is that it only applies to n dimensional systems with $n-1$ control forces. The results of Chapter 5 will be published in [35].

1.3 Outline of the Thesis

A short outline of the thesis is as follows:

Chapter 2: In this chapter we review the necessary elements from the theory of differential geometry and Lie groups.

Chapter 3: Here we derive the equations of motion for the so called simple mechanical systems on Lie groups. This is done using calculus of variations and Lie group theory.

Chapter 4: In this chapter we review elements of control theory and give a control analysis of simple mechanical systems on Lie groups. Most importantly we provide a new result giving sufficient conditions for a simple mechanical system on a Lie group to be locally controllable along a relative equilibrium.

Chapter 5: This chapter is devoted to an exposition of control algorithms for simple mechanical systems on Lie groups with fewer actuators than degrees of freedom. In particular we design an algorithm which is able speed up a system along a relative equilibrium; this is illustrated by applying the theory to two example systems.

Chapter 2

Lie Groups

The purpose of this chapter is to introduce some elements from differential geometry necessary to understand the material presented later on. Most of this material can be found in e.g. the books [8], [45] and [2]. The notation in this thesis follows most closely the one in [2], which is the notation most often used in geometric mechanics literature.

We start out by reviewing elements from differential geometry, without any proofs, in order to introduce notation and clarify what is assumed knowledge of the reader. The more thorough presentation of the theory starts by introducing Lie brackets. This is followed by a section on general Lie groups including the important concept of the Lie algebra corresponding to a Lie group. We end the chapter with a section on matrix Lie groups which are the special case of Lie groups we will focus on after this chapter.

2.1 Preliminaries in Differential Geometry

A map $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **analytic**, or C^ω, on an open set $D \subset \mathbb{R}^n$ if f in an open neighbourhood of every point in D is expressible as a convergent power series, i.e. it's Taylor series expansion around an arbitrary point in D converges in an open neighbourhood of it.

A set M is said to be a locally Euclidean space of dimension n if M is a Hausdorff topological space for which every point $m \in M$ has an open neighborhood U homeomorphic via x to an open subset of Euclidean space $\mathbb{R}^n$, i.e. $x : U \to \mathbb{R}^n$. The pair (U, x) is called a **coordinate system** or a **chart**. A differentiable structure of class C^k, $k \in \mathbb{N} \cup \{\infty\} \cup \{\omega\}$, on a locally Euclidean space M is a collection of coordinate systems $\{(U_\alpha, x_\alpha) \mid \alpha \in A\}$ satisfying

$$\bigcup_{\alpha \in A} U_\alpha = M,$$

and

$$x_\alpha \circ x_\beta^{-1} \quad \text{is } C^k \text{ for all } \alpha, \beta \text{ with } U_\alpha \cap U_\beta \neq \emptyset,$$

and the collection contains all coordinate systems enjoying this property. If $k = \infty$ the space M, with this differentiable structure, is called a **differentiable manifold** or just a **manifold** and if $k = \omega$ then M, with this differentiable structure, is called an **analytic manifold**.

The **tangent space** to M at $m \in M$, denoted by T_mM, is the vector space which in a coordinate system $(U,x) = (U, x^1, \dots, x^n)$ it is given by

$$T_mM = \operatorname{span}\left\{ \left.\frac{\partial}{\partial x^1}\right|_m, \dots, \left.\frac{\partial}{\partial x^n}\right|_m \right\}.$$

Let N denote a (analytic) manifold and let $f : M \to N$, then f is said to be smooth (analytic) if for local coordinate systems (U,x) around $m \in M$ and (V,y) around $f(m) \in N$ the map $y \circ f \circ x^{-1}$ is smooth (analytic). If f is a smooth bijection with a smooth inverse it is called a **diffeomorphism**. For the special case $N = \mathbb{R}$ we denote the class of smooth functions on M by $C^\infty(M)$. The **tangent map** of f at $m \in M$ is the linear map

$$T_mf : T_mM \to T_{f(m)}N,$$

defined by

$$T_mf(v)(g) = v(g \circ f),$$

where $v \in T_mM$ and g is a smooth function in a neighbourhood of $f(m)$. The dual of T_mf is the linear map

$$T_m^*f : T_{f(m)}^*N \to T_m^*M,$$

defined by

$$T_m^*f(\alpha)(v) = \alpha(T_mf(v)),$$

where $\alpha \in T_{f(m)}^*N$ and $v \in T_mM$.

Locally T_mf is seen to be given by the Jacobian matrix as

$$T_mf\left(\left.\frac{\partial}{\partial x^j}\right|_m \right) = \sum_{i=1}^{\dim(N)} \left.\frac{\partial y^i \circ f}{\partial x^j}\right|_m \left.\frac{\partial}{\partial y^i}\right|_{f(m)}, \quad i \in \{1, \dots, n\}.$$

For the special case $N = \mathbb{R}$ and f a smooth function the tangent map is denoted the differential which for $m \in M$ and $v \in T_mM$ is defined

$$\mathrm{d}f_m(v) = v(f).$$

Thus we have $\mathrm{d}f_m \in T_m^*M$, where T_m^*M is the dual of T_mM, and in the coordinate system (U,x) we get

$$T_m^*M = \operatorname{span}\left\{ \mathrm{d}x_m^1, \dots, \mathrm{d}x_m^n \right\}.$$

Let $f : M \to N$ be smooth. Then if f is injective and T_mf is nonsingular for all $m \in M$ the pair (M, f) is called a **submanifold** of N. If for $p_0 = f(m)$, $m \in M$, T_pf has full rank for all $p \in P = f^{-1}(p_0)$ then (P, i), where $i : P \hookrightarrow M$ is the inclusion map, is a submanifold and we have $T_pP = \{v \in T_pM |\ T_pf(v) = 0\}$. Furthermore, if M and N are analytic and f is an analytic map then P can be given the structure of an analytic manifold.

The sets

$$TM = \bigcup_{m \in M} T_m M, \qquad T^*M = \bigcup_{m \in M} T_m^* M,$$

can be given a natural differentiable structure induced by the differentiable structure $\{(U_\alpha, x_\alpha) \mid \alpha \in A\}$ for M. For TM it is the differential structure containing $\{(TU_\alpha, \phi_\alpha) \mid \alpha \in A\}$ where $\phi_\alpha(v) = (x_\alpha(m), \mathrm{d}(x_\alpha^1)_m(v), \ldots, \mathrm{d}(x_\alpha^n)_m(v))$, for $v \in T_m U_\alpha$, and for T^*M it is the differentiable structure containing $\{(T^*U_\alpha, \psi_\alpha) \mid \alpha \in A\}$ where $\psi_\alpha(w) = (x_\alpha(m), w(\frac{\partial}{\partial x_\alpha^1}|_m), \ldots, w(\frac{\partial}{\partial x_\alpha^n}|_m))$ for $w \in T_m^* U_\alpha$. TM and T^*M are thus seen to be manifolds of dimension $2n$. Equipped with these differential structures the manifold TM is called the **tangent bundle** of M and the manifold T^*M is called the **cotangent bundle** of M. We denote by $\tau : TQ \to Q$ and $\pi : T^*Q \to Q$ the natural projections given by $\tau(v) = m$, for $v \in T_m M$, and $\pi(w) = m$, for $w \in T_m^* M$.

The **tangent map**

$$Tf : TM \to TN,$$

is the map defined by $Tf|_{T_m M} = T_m f$.

In geometric mechanics the manifold describing the possible configurations of a mechanical system is called the **configuration manifold** of the system. The dimension of the configuration manifold is called the **degrees of freedom** for the mechanical system.

A **vector field** $X : M \to TM$ on M is a lifting of M into TM, that is

$$\tau(X) = \mathrm{id}_M,$$

where id_M is the identity on M. In a local coordinate system $(U, x) = (U, x^1, \ldots, x^n)$ a vector field X is given by

$$X(x) = \sum_{i=1}^{n} X^i(x) \frac{\partial}{\partial x^i}.$$

If $X^i(x) \in C^\infty(U)$, $i \in \{1, \ldots, n\}$, for every coordinate system X is called smooth or C^∞. We denote by $\mathfrak{X}(M)$ the set of smooth vector fields on M. If M is an analytic manifold X is called analytic or C^ω if it in every coordinate system is analytic.

Let $X \in \mathfrak{X}(M)$ and let $\Phi_t^X(m)$ be the solution to the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \Phi_t^X(m) = X(\Phi_t^X(m)),$$

on M with $\Phi_0^X(m) = m$, then Φ_t^X is called the **flow** of X. The flow exist and is unique by a translation of the fundamental existence and uniqueness theorem for first order differential equations in $\mathbb{R}^n$ to the language of manifolds.

2.1.1 The Lie Bracket

We start by defining the Lie bracket of vector fields.

Definition 1. *The **Lie bracket** of two smooth vector fields X and Y on a manifold M is for $f \in C^\infty(M)$ the vector field operating as*

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

Locally $X(x) = \sum_{i=1}^{n} X^i(x)\frac{\partial}{\partial x^i}$ and $Y(x) = \sum_{i=1}^{n} Y^j(x)\frac{\partial}{\partial x^j}$ which gives the local expression for the Lie bracket as

$$[X,Y](x) = \sum_{i,j=1}^{n} X^i(x)\frac{\partial Y^j}{\partial x^i}(x)\frac{\partial}{\partial x^j} - \sum_{i,j=1}^{n} Y^j(x)\frac{\partial X^i}{\partial x^j}(x)\frac{\partial}{\partial x^i} \tag{2.1}$$

$$= \sum_{i,j=1}^{n} \left(X^i(x)\frac{\partial Y^j}{\partial x^i}(x) - Y^i(x)\frac{\partial X^j}{\partial x^i}(x) \right) \frac{\partial}{\partial x^j}. \tag{2.2}$$

If we write $X = (X^1, \dots, X^n)^T$ and $Y = (Y^1, \dots, Y^n)^T$ then (2.1) becomes

$$[X,Y](x) = \frac{\partial Y}{\partial x}(x)X(x) - \frac{\partial X}{\partial x}(x)Y(x),$$

where $\frac{\partial X}{\partial x}(x)$ is the Jacobian of X. If X and Y are smooth $[X,Y]$ is clearly seen to be smooth.

The following is a straightforward result (though the Jacobi identity is tedious to prove) of the definition.

Proposition 2. *Let M be a manifold. The Lie bracket $[\cdot,\cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ is bilinear, skew-symmetric, i.e. $[X,Y] = -[Y,X]$, and satisfies the Jacobi identity*

$$[[X,Y],Z] + [[Z,X],Y] + [[Y,Z],X] = 0,$$

for $X, Y, Z \in \mathfrak{X}(M)$.

A connection between composition of integral curves and the Lie bracket of the vector fields is given by the following proposition.

Proposition 3. *Let X and Y be smooth vector fields on a manifold $M \ni m$. Then we have that the Lie bracket can be computed as*

$$[X,Y](m) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(T_{\Phi_t^X(m)}\Phi_{-t}^X\big(Y(\Phi_t^X(m))\big)\right).$$

Proof. Let X and Y be given in the coordinate chart (U,x) by $X(x) = (X^1(x), \dots, X^n(x))^T$ and $Y(x) = (Y^1(x), \dots, Y^n(x))^T$. Then we get, when using Taylor expansions and leaving out all terms of order $\mathcal{O}(t^2)$ and $\mathcal{O}(s^2)$, the following

$$\begin{aligned}
&\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(T_{\Phi_t^X(x)}\Phi_{-t}^X Y(\Phi_t^X(x))\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \Phi_{-t}^X \circ \Phi_s^Y(\Phi_t^X(x)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \Phi_{-t}^X \circ \Phi_s^Y(x + tX(x)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \Phi_{-t}^X(x + tX(x) + sY(x) + st\frac{\partial Y}{\partial x}(x)X(x)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} (x + tX(x) + sY(x) + st\frac{\partial Y}{\partial x}(x)X(x) - tX(x) - st\frac{\partial X}{\partial x}(x)Y(x)) \\
&= \frac{\partial Y}{\partial x}(x)X(x) - \frac{\partial X}{\partial x}(x)Y(x).
\end{aligned}$$

□

This proposition also provides a convenient method for calculating the Lie bracket of vector fields in a coordinate free way.

We conclude this section by a proposition which will be needed in the following section in connection with so called left-invariant vector fields.

Proposition 4. *Let M and N be manifolds and let $\phi : M \to N$ be smooth. Assume that $X, X_1 \in \mathfrak{X}(M)$ and $Y, Y_1 \in \mathfrak{X}(N)$ satisfies $T\phi \circ X = Y \circ \phi$ and $T\phi \circ X_1 = Y_1 \circ \phi$. Then we have $T\phi \circ [X, X_1] = [Y, Y_1] \circ \phi$.*

Proof. Let $m \in M$ and $f \in C^\infty(N)$. By use of the definition of the tangent map and the Lie bracket we get

$$\begin{aligned} T_m\phi([X, X_1](m))(f) &= [X, X_1](m)(f \circ \phi) \\ &= X(m)(X_1(f \circ \phi)) - X_1(m)(X(f \circ \phi)) \\ &= X(m)((T\phi \circ X_1)(f)) - X_1(m)((T\phi \circ X)(f)) \\ &= X(m)(Y_1(f) \circ \phi) - X_1(m)(Y(f) \circ \phi) \\ &= T_m\phi(X(m))(Y_1(f)) - T_m\phi(X_1(m))(Y(f)) \\ &= Y \circ \phi(m)(Y_1(f)) - Y_1 \circ \phi(m)(Y(f)) \\ &= [Y, Y_1] \circ \phi(m)(f), \end{aligned}$$

and the result has been obtained. □

2.2 Lie Groups

We start this section with the definition of a Lie group.

Definition 5. *A **Lie group** is a differentiable manifold G which is endowed with a group structure such that the the product map $G \times G \to G$, $(x, y) \mapsto xy$, and the inverse map $G \to G$, $x \mapsto x^{-1}$, are C^∞. If in addition G is an analytic manifold and the product map and the inverse map are analytic G is called an analytic Lie group.*

In this chapter we will use $e \in G$ to denote the identity element of G.

For $g, h \in G$ left translation by g and right translation by g, denoted L_g and R_g respectively, are defined by

$$L_g(h) = gh, \qquad R_g(h) = hg.$$

A **left-invariant vector field** X on a Lie group G is a vector field satisfying

$$X(gh) = T_hL_g(X(h)),$$

for all $g, h \in G$. We denote by $\mathcal{L}(G)$ the space of all left-invariant vector fields on G which is seen to be a vector space.

Proposition 6. *Let G be a Lie group and $X \in \mathcal{L}(G)$. Then X is C^∞.*

Proof. Let G be a Lie group. Let $X \in \mathcal{L}(G)$ and $g \in G$, then $X(g) = T_eL_g(X(e))$. Let (W, q_1) be a coordinate system around g. Let (V, q_0) be a coordinate system around

e and (U, q_1) be a coordinate system around g which satisfy $\bigcup_{g\in U} L_g(V) \subset W$. Such charts exist: let (V_0, q_0) be a chart around e and $\sigma : G \times G \to G$ the product map, then (U, q_1) and (V, q_0) with $U \times V = (W \times V_0) \cap \sigma^{-1}(W)$ satisfies the condition since U and V are open and $g \in U$ and $e \in V$. Let $\iota : q_1(U) \times q_0(V) \to q_1(W)$ be the coordinate expression for the product map in these charts. Let $\phi(t)$ be a smooth curve on V with $\phi(0) = e$ and $\dot\phi(0) = X(e)$ then we have that $X(g)$ in coordinates for $T_g q_1(W)$ is given by

$$\begin{aligned} T_g q_1(X(g)) &= T_g q_1(T_e L_g(X(e))) = T_g q_1 \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} L_g\phi(t) \\ &= T_g q_1 \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} q_1^{-1} \circ \iota(q_1(g), q_0(\phi(t))) \right) = T_{(q_1(g),q_0(e))}\iota(0, T_e q_0(X(e))). \end{aligned}$$

Since ι is smooth its Jacobian $T_{(q_1,q_0)}\iota$ is a smooth function of q_1 and q_0 and thus X depends smoothly on g in coordinates and it is therefore smooth. □

Remark 7. If G is an analytic Lie group we can replace the word "smooth" with "analytic" in the proof of Proposition 6. Therefore we get that a left-invariant vector field on an analytic Lie group is analytic. •

Proposition 6 gives the following result.

Corollary 8. *Let G be a Lie group. If $X, Y \in \mathcal{L}(G)$ then $[X, Y] \in \mathcal{L}(G)$.*

Proof. This is a direct consequence of the definition of a left-invariant vector field and Propositions 4 and 6. □

Let $g \in G$, $\xi \in T_eG$, and $(\cdot)_L : T_eG \to \mathfrak{X}(G)$ be the map defined by

$$\xi_L(g) = T_e L_g(\xi),$$

which since

$$\xi_L(gh) = T_e L_{gh}(\xi) = T_h L_g(T_e L_h(\xi)) = T_h L_g(\xi_L(h)),$$

is a map $(\cdot)_L : T_eG \to \mathcal{L}(G)$. Then we have.

Proposition 9. *Let G be a Lie group. Then $(\cdot)_L : T_eG \to \mathcal{L}(G)$ is an isomorphism with inverse $X \mapsto X(e)$.*

Proof. Let $\beta : \mathcal{L}(G) \to T_eG$ denote the map $X \mapsto X(e)$. Let $X, Y \in \mathcal{L}(G)$ and assume $\beta(X) = \beta(Y)$ then

$$X(g) = T_e L_g(X(e)) = T_e L_g(Y(e)) = Y(g),$$

for all $g \in G$ so β is injective. Let $\xi \in T_eG$ then

$$\beta(\xi_L) = \xi_L(e) = \xi,$$

so β is surjective and thus bijective and we see that β is the inverse of $(\cdot)_L$. □

Next we define a Lie algebra.

Definition 10. *Let V be a vector space (over $\mathbb{R}$) and let the map $[\cdot,\cdot]_V : V \times V \to V$, for all $\xi, \eta, \zeta \in V$, satisfy:*

1. *bilinearity,*
2. *skew symmetry, i.e., $[\xi,\eta]_V = -[\eta,\xi]_V$,*
3. *the Jacobi identity, i.e.,*

$$[[\xi,\eta]_V,\zeta]_V + [[\zeta,\xi]_V,\eta]_V + [[\eta,\zeta]_V,\xi]_V = 0.$$

Then $(V, [\cdot,\cdot]_V)$ is called a ***Lie algebra****.*

Let $(V, [\cdot,\cdot]_V)$ be a Lie algebra and W a nonempty subset of V. If $(W, [\cdot,\cdot]_W)$, where $[\cdot,\cdot]_W$ is the restriction of $[\cdot,\cdot]_V$ to W, is a Lie algebra it is called a Lie subalgebra of $(V, [\cdot,\cdot]_V)$. Given a subset $S \subset V$ the **Lie algebra generated by** S is the smallest Lie subalgebra of $(V, [\cdot,\cdot]_V)$ containing S.

From Corollary 8 we have that we can make the following definition.

Definition 11. *Let G be a Lie group. For $\xi, \eta \in T_eG$ define the* ***bracket*** *$[\cdot,\cdot] : T_eG \times T_eG \to T_eG$ by*

$$[\xi,\eta] = [\xi_L, \eta_L](e).$$

By construction this bracket inherits the properties of the Lie bracket for vector fields; see Proposition 2. Therefore we have the following.

Corollary 12. *Let G be a Lie group. Then $(T_eG, [\cdot,\cdot])$ is a Lie algebra.*

Definition 13. *Let G be a Lie group. Then we denote the corresponding Lie algebra $(T_eG, [\cdot,\cdot])$ by $\mathfrak{g}$.*

We will denote $[\cdot,\cdot]$ by $[\cdot,\cdot]_{\mathfrak{g}}$ when risk of confusing the bracket of a Lie algebra corresponding to a Lie group with the Lie bracket of vector fields on $\mathfrak{g}$.

Proposition 14. *For $g, h \in G$ and $X \in \mathcal{L}(G)$ we have*

$$L_g \Phi_t^X(h) = \Phi_t^X(L_g h),$$

for $|t| < \delta$ for some $\delta > 0$.

Proof. From the existence and uniqueness theorem for differential equations (see e.g. [15]) we know there exists a $\delta > 0$ such that $\Phi_t^X(h)$ and $\Phi_t^X(L_g h)$ are defined.

We have $L_g \Phi_0^X(h) = \Phi_0^X(L_g h) = L_g h$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} L_g \Phi_t^X(h) = T_{\Phi_t^X(h)} L_g \big(X(\Phi_t^X(h))\big) = X(L_g \Phi_t^X(h)),$$
$$\frac{\mathrm{d}}{\mathrm{d}t} \Phi_t^X(L_g h) = X(\Phi_t^X(L_g h)).$$

Since $L_g \Phi_t^X(h)$ and $\Phi_t^X(L_g h)$ satisfy the same differential equation and are equal for $t = 0$ we know from the existence and uniqueness theorem for differential equations that they are equal for all $|t| < \delta$. □

This gives the next useful corollary.

Corollary 15. *For a Lie group $G \ni g$ and $X \in \mathcal{L}(G)$ we have*

$$\Phi^X_{t_1+t_2}(g) = \Phi^X_{t_1}(g)\Phi^X_{t_2}(e),$$

for $|t_1| < \delta$ and $|t_2| < \delta$ for some $\delta > 0$.

Proof. This is an immediate consequence of the group property of flows ($\Phi^X_t \circ \Phi^X_s = \Phi^X_{t+s}$) and Proposition 14 with $h = e$ and $g = \Phi^X_{t_1}(g)$. □

Proposition 14 leads to the following result on left-invariant vector fields.

Proposition 16. *A left-invariant vector field X on a Lie group is complete, i.e., the flow Φ^X_t is defined for all $t \in \mathbb{R}$.*

Proof. From the existence and uniqueness theorem for differential equations we know that there exist a neighborhood V of e and a $\delta > 0$ such that $\Phi^X_t(g)$ is defined for $g \in V$ and $|t| < \delta$. From Proposition 14 we have that $\Phi^X_t(g)$ is defined for $g \in L_h(V)$ and $|t| < \delta$ for all $h \in G$. Thus $\Phi^X_t(g)$ is defined for all $g \in G$ for $|t| < \delta$ but because of the group property of the flow this means that $\delta = \infty$. □

Since left-invariant vector fields are complete we can make the following definition.

Definition 17. *For a Lie group G we define the map* $\exp : \mathfrak{g} \to G$ *by*

$$\exp(\xi) := \Phi^{\xi_L}_1(e),$$

for $\xi \in \mathfrak{g}$.

This map is called the **exponential map** for reasons which will become clear later. From the definition we see that we have $\exp(t\xi) = \Phi^{t\xi_L}_1(e) = \Phi^{\xi_L}_t(e)$ for $t \in \mathbb{R}$ and thus $\exp(t\xi)$ is the integral curve of ξ_L which at $t = 0$ is e and has tangent ξ. From Proposition 14 we have that $L_g \exp(t\xi) = \Phi^{\xi_L}_t(g)$ so this is the integral curve of ξ_L which at $t = 0$ is equal to g and has tangent $\xi_L(g)$.

Proposition 18. *For a Lie group G and $\xi \in \mathfrak{g}$ the map $t \mapsto \exp(t\xi)$, $t \in \mathbb{R}$, is a 1-parameter subgroup of G, i.e., $\exp(0\xi) = e$ and $\exp((t_1 + t_2)\xi) = \exp(t_1\xi)\exp(t_2\xi)$.*

Proof. We immediately get $\exp(0\xi) = \Phi^{\xi_L}_0(e) = e$ and $\exp((t_1+t_2)\xi) = \exp(t_1\xi)\exp(t_2\xi)$ is a consequence of Corollary 15. □

The next result gives the smoothness properties of the exponential map.

Proposition 19. *Let G be a Lie group. The exponential map* $\exp : \mathfrak{g} \to G$ *is C^∞ and we have $T_0 \exp = \mathrm{id}_{\mathfrak{g}}$.*

Proof. Consider the smooth complete vector field on $G \times \mathfrak{g}$ given by $X(g, \xi) = (\xi_L(g), 0)$. Since X is smooth so is Φ^X_1. Let $\pi : G \times \mathfrak{g} \to G$ be the projection onto G. Then $\exp(\xi) = \pi \circ \Phi^X_1(e, \xi)$ is a composition of smooth maps and it is therefore smooth.

Since $\xi = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \Phi^{\xi_L}_t(e) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \exp(t\xi) = T_0 \exp(\xi)$ we have $T_0 \exp = \mathrm{id}_{\mathfrak{g}}$.

□

Remark 20. Let G be an analytic Lie group. Due to the Remark 7 we can replace the word "smooth" with "analytic" in the proof of Proposition 19. Therefore we get that the exponential map is analytic for an analytic Lie group. •

Proposition 19, together with the inverse function theorem, shows that in a neighborhood $U \subset G$ of e and a neighborhood $V \subset \mathfrak{g}$ of 0 there exist a C^∞ (C^ω) inverse, denoted log, to exp, i.e., $\log(\exp(\xi)) = \xi$ and $\exp(\log(g)) = g$ for $\xi \in V$ and $g \in U$. The coordinates in the chart $(U, \log)$ are the so called exponential coordinates of the first kind.

Proposition 21. *Let G and H be Lie groups. Let $\phi : G \to H$ be a Lie group homomorphism, i.e. ϕ is smooth and $\phi(ab) = \phi(a)\phi(b)$ for $a, b \in G$. Then, for $\xi \in \mathfrak{g}$ we have*

$$\phi(\exp(t\xi)) = \exp\Big(t\big(T_e\phi(\xi)\big)\Big).$$

Proof. Since ϕ is a homomorphism we have $L_{\phi(g)} \circ \phi = \phi \circ L_g$ for $g \in G$. This fact is used in the following

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\phi(\exp(t\xi)) &= T_{\exp(t\xi)}\phi\big(\xi_L(\exp(t\xi))\big) = T_{\exp(t\xi)}\phi\big(T_eL_{\exp(t\xi)}(\xi)\big) \\ &= T_e\phi \circ L_{\exp(t\xi)}(\xi) = T_eL_{\exp(t\xi)} \circ \phi(\xi) \\ &= T_eL_{\phi(\exp(t\xi))}(T_e\phi(\xi)) = (T_e\phi(\xi))_L\big(\phi(\exp(t\xi))\big).\end{aligned}$$

Since $\exp(tT_e\phi(\xi))$ satisfies the same differential equation and the initial conditions are the same for $t = 0$ the result follows. □

For a Lie group G and $g \in G$ the inner automorphism $I_g : G \to G$ is given by $I_g = L_g \circ R_{g^{-1}} = R_{g^{-1}} \circ L_g$. It is easily checked that this in fact is a homomorphism. Since R_g and L_g are diffeomorphisms I_g is a diffeomorphism with inverse $I_g^{-1} = L_g^{-1} \circ R_{g^{-1}}^{-1} = L_{g^{-1}} \circ R_g = I_{g^{-1}}$. Since I_g is a homomorphism we have $T_eI_g : \mathfrak{g} \to \mathfrak{g}$. We denote by Ad the adjoint map given by

$$\mathrm{Ad}_g := T_eI_g.$$

From Proposition 21 we get the following result.

Corollary 22. *For a Lie group G, $g \in G$, and $\xi \in \mathfrak{g}$, we have*

$$I_g \exp(t\xi) = \exp(t\mathrm{Ad}_g(\xi)).$$

Another result relating the bracket $[\cdot,\cdot]$ to the adjoint map and the exponential map is the following.

Proposition 23. *Let G be a Lie group. Then for $\xi, \eta \in \mathfrak{g}$ we have*

$$[\xi, \eta] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}_{\exp(t\xi)}(\eta).$$

Proof. Let $g \in G$, $v \in T_gG$ and $\gamma = T_eL_{g^{-1}}(v) \in \mathfrak{g}$. This means that $v = \gamma_L(g)$. From Proposition 14 with $h = e$ and g replaced with $\Phi_s^{\gamma_L}(g)$ we get $\Phi_t^{\xi_L} \circ \Phi_s^{\gamma_L}(g) = \Phi_s^{\gamma_L}(g)\Phi_t^{\xi_L}(e)$. Using this we get

$$\begin{aligned}
T_g\Phi_t^{\xi_L}(v) &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \Phi_t^{\xi_L} \circ \Phi_s^{\gamma_L}(g) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \Phi_s^{\gamma_L}(g)\Phi_t^{\xi_L}(e) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} L_g\Phi_s^{\gamma_L}(e)\Phi_t^{\xi_L}(e) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} L_g \circ R_{\Phi_t^{\xi_L}(e)}(\Phi_s^{\gamma_L}(e)) \\
&= T_eL_g \circ R_{\Phi_t^{\xi_L}(e)}(\gamma) = T_eL_g \circ R_{\Phi_t^{\xi_L}(e)}(T_eL_{g^{-1}}(v)) \\
&= T_gR_{\Phi_t^{\xi_L}(e)}(v).
\end{aligned}$$

From this and Proposition 3 we get

$$\begin{aligned}
[\xi, \eta] = [\xi_L, \eta_L](e) &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Big(T_{\Phi_t^{\xi_L}(e)}\Phi_{-t}^{\xi_L}(\eta_L(\Phi_t^{\xi_L}(e)))\Big) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Big(T_{\Phi_t^{\xi_L}(e)}R_{\Phi_{-t}^{\xi_L}(e)}(T_eL_{\Phi_t^{\xi_L}(e)}(\eta))\Big) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Big(T_eR_{\Phi_{-t}^{\xi_L}(e)} \circ L_{\Phi_t^{\xi_L}(e)}(\eta)\Big) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathrm{Ad}_{\Phi_t^{\xi_L}(e)}(\eta) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathrm{Ad}_{\exp(t\xi)}(\eta).
\end{aligned}$$

□

For a Lie group G and $\xi, \eta \in \mathfrak{g}$ we define the adjoint operator $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ as

$$\mathrm{ad}_\xi(\eta) := [\xi, \eta].$$

The dual of ad_ξ is the map $\mathrm{ad}_\xi^* : \mathfrak{g} \to \mathfrak{g}$ defined for $\xi, \eta \in \mathfrak{g}$ and $\alpha \in \mathfrak{g}^*$ by $\mathrm{ad}_\xi(\alpha)(\eta) = \alpha(\mathrm{ad}_\xi(\eta))$. In a given basis for $\mathfrak{g}$ the matrix representation of ad_ξ^* is the transpose of the matrix representation for ad_ξ.

The definition of the adjoint operator leads to the following result regarding the relation between Ad_g, ad_ξ, and exp.

Proposition 24. *Let G be a Lie group and $\xi \in \mathfrak{g}$. Then* Ad *is a group homomorphism and it satisfies*

$$\mathrm{Ad}_{\exp(\xi)} = \exp(\mathrm{ad}_\xi).$$

Proof. Let $g, h \in G$. Then from the definition of Ad we get

$$\mathrm{Ad}_{gh} = T_eI_{gh} = T_eI_g \circ I_h = T_eI_g \circ T_eI_h = \mathrm{Ad}_g \circ \mathrm{Ad}_h,$$

so Ad is a group homomorphism.

Using Proposition 23 and the fact that exp and Ad are homomorphisms we get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Ad}_{\exp(t\xi)}(\eta) &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathrm{Ad}_{\exp((s+t)\xi)}(\eta) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathrm{Ad}_{\exp(s\xi)}(\mathrm{Ad}_{\exp(t\xi)}(\eta)) \\
&= \mathrm{ad}_\xi(\mathrm{Ad}_{\exp(t\xi)}(\eta)).
\end{aligned}$$

Since $\mathfrak{g}$ is a vector space and ad_ξ is linear this shows that

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathrm{Ad}_{\exp(t\xi)}(\eta) = \mathrm{ad}_\xi^k(\mathrm{Ad}_{\exp(t\xi)}(\eta)) \Rightarrow \left.\frac{\mathrm{d}^k}{\mathrm{d}t^k}\right|_{t=0}\mathrm{Ad}_{\exp(t\xi)}(\eta) = \mathrm{ad}_\xi^k(\eta).$$

Thus formally a Taylor expansion gives

$$\mathrm{Ad}_{\exp(t\xi)} = \exp(\mathrm{ad}_{t\xi}).$$

Since $\|\mathrm{ad}_\xi\| \leq 2\|\xi\| < \infty$ we have that $\exp(\mathrm{ad}_\xi)$ converges for all ξ. □

2.3 Matrix Lie Groups

The class of Lie groups we will consider in the remainder of this thesis is that of matrix Lie groups. Therefore we start by defining some of the classical matrix Lie groups.

2.3.1 Some Classical Matrix Lie Groups

For a vector space $V \ni v$ we use in the following the natural identification of T_vV with V itself. Since $R^{n\times n}$ is a Banach space and $Gl(n)$ is an open subset of $\mathbb{R}^{n\times n}$ we will in this section use the differential D instead of the tangent map when convenient.

The General Linear Group $Gl(n)$

For $A, B \in \mathbb{R}^{n\times n}$ we have $\det(AB) = \det(A)\det(B)$, and $\det(e) = 1$, so the space

$$Gl(n) = \{X \in \mathbb{R}^{n\times n} |\ \det(X) \neq 0\},$$

equipped with the matrix product is a group. This group is called the general linear group. Since $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ is a continuous function $Gl(n)$ becomes a manifold as the open subset of $\mathbb{R}^{n^2}$, identified with $\mathbb{R}^{n\times n}$, where the determinant is non-zero. Since $(AB)_{ij} = \sum_{k=1}^n A_{ik}B_{kj}$ and the entries of A^{-1} is a rational function of the entries of A both the matrix product and the inverse map are analytic and thus $Gl(n)$ with the matrix product can be given the structure of an analytic Lie group. A subgroup of $Gl(n)$ which is also a submanifold of $Gl(n)$, hence a Lie group, is called a **matrix Lie group**. Since det is continuous there is a neighborhood $U \subset \mathbb{R}^{n\times n}$ of e such that $\det(U) > 0$ and thus $U \subset Gl(n)$ which shows that

$$T_eGl(n) = T_eU = \mathbb{R}^{n\times n} =: \mathfrak{gl}(n).$$

The Special Linear Group $Sl(n)$

Since $\det : Gl(n) \to \mathbb{R} \setminus \{0\}$ is a homomorphism the space

$$Sl(n) = \{X \in Gl(n) |\ \det(X) = 1\},$$

is a subgroup of $Gl(n)$. This group is called the special linear group. Define the map $F : Gl(n) \to \mathbb{R} \setminus \{0\}$ by $F(X) = \det(X)$. Let $A \in Gl(n)$, $a = \det(A)$, then $F(x) = L_a \circ F \circ L_{A^{-1}}(X)$ giving

$$DF(X) = D(L_a \circ F \circ L_{A^{-1}})(X) = aDF(A^{-1}X)DL_{A^{-1}}(X).$$

Since L_g is a diffeomorphism we thus get choosing $A = X$

$$\text{Rank}(DF(X)) = \text{Rank}(DF(e)),$$

and thus the rank i constant. We therefore get that $Sl(n) = F^{-1}(1)$ is a submanifold of $Gl(n)$. Since $Sl(n)$ is a subgroup of $Gl(n)$, and F is analytic, it is therefore an analytic matrix Lie group.

The Orthogonal Group $O(n)$

Define the analytic map $H : Gl(n) \to Gl(n)$ according to $H(X) = X^T X$. Since H is a homomorphism the space

$$O(n) = \{X \in Gl(n) |\ X^T X = e\},$$

is a group. This group is denoted the orthogonal group. Let $A \in Gl(n)$, then $H(X) = R_{A^{-1}} \circ L_{A^{-T}} \circ H \circ R_A(X)$ giving

$$DH(X) = DR_{A^{-1}}(L_{A^{-T}} \circ H \circ R_A(X))DL_{A^{-T}}(H \circ R_A(X))DH(XA)DR_A(X).$$

Since L_g and R_g are diffeomorphisms we get when choosing $A = X^{-1}$

$$\text{Rank}(DF(X)) = \text{Rank}(DF(e)),$$

so the rank is constant. We therefore have that $O(n) = H^{-1}(e)$ is an analytic submanifold of $Gl(n)$ and since it is a subgroup of $Gl(n)$ it is an analytic matrix Lie group. Let $\xi \in T_e Gl(n)$, then

$$\begin{aligned}
T_e H(\xi) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} H(\Phi_t^{\xi_L}(e)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\Phi_t^{\xi_L}(e)(\Phi_t^{\xi_L}(e))^T\right) \\
&= \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^{\xi_L}(e)\right)(\Phi_0^{\xi_L}(e))^T + \Phi_0^{\xi_L}(e)\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^{\xi_L}(e)\right)^T \\
&= \xi + \xi^T.
\end{aligned}$$

Thus we have

$$T_e O(n) = \{A \in \mathbb{R}^{n\times n} |\ A + A^T = 0\} =: \mathfrak{o}(n).$$

This also shows that $\dim(O(n)) = \frac{1}{2}n(n-1)$.

The Special Orthogonal Group $SO(n)$

Let

$$SO(n) = Sl(n) \cap O(n).$$

Since $SO(n) = (F \times H)^{-1}(1, e)$ we know from the previous discussions that $SO(n)$ is an analytic matrix Lie group. This group is called the special orthogonal group. Since for

$A \in O(n)$ we have $\det(A) = \pm 1$ and since the determinant is continuous we can choose a neighborhood $V \subset O(n)$ of e where we must have $\det(V) = 1$ and thus $V \subset SO(n)$. Therefore $T_e SO(n) = T_e V = T_e O(n) =: \mathfrak{so}(n)$.

For $n = 3$ Rodrigues' formula gives, with $\hat{x} \in \mathfrak{so}(3)$, that

$$\exp(\hat{x}) = \mathrm{id} + \sin(\|x\|)\frac{\hat{x}}{\|x\|} + (1 - \cos(\|x\|))\frac{\hat{x}^2}{\|x\|^2},$$

with the isomorphism $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ given by $\hat{x}y = x \times y$, $x, y \in \mathbb{R}^3$, where $\times$ is the cross product and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^3$. If $R \in SO(3)$ then for $\mathrm{trace}(R) \neq -1$ we have

$$\log(R) = \frac{\phi}{2\sin(\phi)}(R - R^T),$$

where $\phi \in (-\pi, \pi)$ satisfies $2\cos(\phi) = \mathrm{trace}(R) - 1$. For a derivation of exp and log on $SO(3)$ see, e.g., [13].

The Special Euclidean Group $SE(n)$

The subset of $Sl(n+1)$

$$SE(n) = \left\{ \begin{bmatrix} A & v \\ 0_{1\times n} & 1 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} \;\middle|\; A \in SO(n),\ v \in \mathbb{R}^n \right\},$$

is a subgroup of $Sl(n+1)$ since

$$\begin{bmatrix} A_1 & v_1 \\ 0_{1\times n} & 1 \end{bmatrix} \begin{bmatrix} A_2 & v_2 \\ 0_{1\times n} & 1 \end{bmatrix} = \begin{bmatrix} A_1 A_2 & A_1 v_2 + v_1 \\ 0_{1\times n} & 1 \end{bmatrix}.$$

It is a analytic manifold since it can be identified with the analytic product manifold $SO(n) \times \mathbb{R}^n$ and it becomes a submanifold of $Gl(n+1)$ by inclusion. Thus it is an analytic matrix Lie group; it is called the special Euclidean group. It is seen that we have

$$T_e SE(n) = \left\{ \begin{bmatrix} A & v \\ 0_{1\times n} & 0 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} \;\middle|\; A \in \mathfrak{so}(n),\ v \in \mathbb{R}^n \right\} =: \mathfrak{se}(n).$$

By construction $SE(n)$ is isomorphic (meaning there is a group homomorphism between the sets as groups which is a diffeomorphism between the sets as manifolds) to $SO(n) \times \mathbb{R}^n$ with product $(R_1, v_1)(R_2, v_2) = (R_1 R_2, R_1 v_v + v_1)$ for $(R_1, v_1), (R_2, v_2) \in SO(n) \times \mathbb{R}^n$.

Explicit formulas for the exponential and the logarithm, in the case of $SE(3)$, can be found in, e.g., [13].

Some Useful Formulas for Matrix Lie Groups

Since the product on a matrix Lie group is just the ordinary matrix product we easily obtain the following result.

Lemma 25. *Let $G \ni g$ be a matrix Lie group and $\xi \in \mathfrak{g}$. Then*

$$T_e L_g(\xi) = g\xi, \qquad T_e R_g(\xi) = \xi g, \qquad \mathrm{Ad}_g(\xi) = g\xi g^{-1}.$$

Proof. By straightforward calculations we get

$$T_eL_g(\xi) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} L_g\Phi_t^{\xi_L}(e) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g\Phi_t^{\xi_L}(e) = g\xi.$$

$$T_eR_g(\xi) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} R_g\Phi_t^{\xi_L}(e) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^{\xi_L}(e)g = \xi g.$$

$$\mathrm{Ad}_g(\xi) = T_eI_g(\xi) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} I_g(\Phi^{\xi_L}(e)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g\Phi_t^{\xi_L}(e)g^{-1} = g\xi g^{-1}.$$

□

With this result we are able to give an exact formula for the exponential map for a matrix Lie group.

Proposition 26. *Let G be a matrix Lie group and $\xi \in \mathfrak{g}$. Then*

$$\exp(\xi) = \sum_{k=0}^{\infty} \frac{\xi^k}{k!},$$

i.e., exp *is the matrix exponential.*

Proof. $\exp(\xi) = \Phi_t^{\xi_L}(e)$ is the unique solution to the differential equation $\dot{g} = \xi_L(g) = T_eL_g(\xi) = g\xi$, $g(0) = e$, but the solution to this problem is $\sum_{k=0}^{\infty} \frac{(t\xi)^k}{k!}$ which is the matrix exponential. □

This result enables us to show a result which greatly simplifies the calculation of the bracket for the Lie algebra of a matrix Lie group.

Proposition 27. *Let G be a matrix Lie group. Then for $\xi, \eta \in \mathfrak{g}$ we have*

$$[\xi, \eta] = \xi\eta - \eta\xi,$$

i.e., $[\cdot,\cdot]$ is the matrix commutator.

Proof. From Proposition 23, Lemma 25, and Proposition 26 we get

$$[\xi, \eta] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}_{\exp(t\xi)}(\eta) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(t\xi)\eta\exp(-t\xi) = \xi\eta - \eta\xi.$$

□

2.3.2 The Magnus Expansion and the Baker-Campbell-Hausdorff Formula

We start out by proving some results needed to give the first terms in the Magnus series and the Baker-Campbell-Hausdorff formula.

Lemma 28. *For $\Omega, H \in \mathbb{R}^{n\times n}$ and $k \in \mathbb{N}$ the differential of the map $\Omega \mapsto \Omega^k$ operates as*

$$D\Omega^k(H) = \sum_{n=0}^{k-1} \binom{k}{n+1} (\mathrm{ad}_\Omega^n H)\Omega^{k-n-1},$$

where $\binom{i}{j} = \frac{i!}{j!(i-j)!}$ is the binomial coefficient.

Proof. The expression $\Omega \mathrm{ad}_\Omega^n H = (\mathrm{ad}_\Omega^n H)\Omega + \mathrm{ad}_\Omega^{n+1} H$ is obviously true for $n = 0$. By induction we get

$$\begin{aligned} \Omega \mathrm{ad}_\Omega^{n+1} H &= \Omega \mathrm{ad}_\Omega^n (\mathrm{ad}_\Omega H) \\ &= (\mathrm{ad}_\Omega^n (\mathrm{ad}_\Omega H))\Omega + \mathrm{ad}_\Omega^{n+1}(\mathrm{ad}_\Omega H) \\ &= (\mathrm{ad}_\Omega^{n+1} H)\Omega + \mathrm{ad}_\Omega^{n+2} H, \end{aligned}$$

and therefore it must be true for all $n \in \mathbb{N}$.
The expression $D\Omega^k(H) = \sum_{n=0}^{k-1} \binom{k}{n+1} (\mathrm{ad}_\Omega^n H)\Omega^{k-n-1}$ is seen to be true for $k = 1$ and the proof proceeds by induction assuming the validity of it for k

$$\begin{aligned} D\Omega^{k+1}(H) &= D(\Omega \cdot \Omega^k)(H) \\ &= D\Omega(H) \cdot \Omega^k + \Omega \cdot D\Omega^k(H) \\ &= H\Omega^k + \Omega \cdot \sum_{n=0}^{k-1} \binom{k}{n+1} (\mathrm{ad}_\Omega^n H)\Omega^{k-n-1} \\ &= H\Omega^k + \sum_{n=0}^{k-1} \binom{k}{n+1} \left((\mathrm{ad}_\Omega^n H)\Omega + \mathrm{ad}_\Omega^{n+1} H\right) \Omega^{k-n-1} \\ &= H\Omega^k + \sum_{n=0}^{k-1} \binom{k}{n+1} (\mathrm{ad}_\Omega^n H)\Omega^{k-n} + \sum_{n=1}^{k} \binom{k}{n+1} (\mathrm{ad}_\Omega^n H)\Omega^{k-n} \\ &= H\Omega^k + \sum_{n=0}^{k} \binom{k+1}{n+1} \frac{k-n}{k+1} (\mathrm{ad}_\Omega^n H)\Omega^{k-n} + \sum_{n=1}^{k} \binom{k+1}{n+1} \frac{n+1}{k+1} (\mathrm{ad}_\Omega^n H)\Omega^{k-n} \\ &= \sum_{n=0}^{k} \binom{k+1}{n+1} (\mathrm{ad}_\Omega^n H)\Omega^{k-n}. \end{aligned}$$

So if the expression holds for k it will also hold for $k + 1$. □

With this lemma we are able to prove the following proposition.

Proposition 29. *Let G be a matrix Lie group. For $\Omega \in \mathfrak{g}$ we have*

$$T_\Omega(R_{\exp(-\Omega)} \circ \exp) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \mathrm{ad}_\Omega^k \ ,$$

which converges for all $\Omega \in \mathfrak{g}$.

Proof. Let $H \in T_\Omega \mathfrak{g} = \mathfrak{g}$. Using Lemma 25 we get

$$\begin{aligned} T_\Omega \exp(H) &= T_\Omega(R_{\exp(\Omega)} \circ R_{\exp(-\Omega)} \circ \exp)(H) \\ &= T_e R_{\exp(\Omega)}(T_\Omega(R_{\exp(-\Omega)} \circ \exp))(H) \\ &= T_\Omega(R_{\exp(-\Omega)} \circ \exp)(H) \exp(\Omega). \end{aligned}$$

We therefore calculate $T_\Omega \exp(H)$. By the definition of the tangent map of the exponential map we get using Lemma 28

$$\begin{aligned} T_\Omega \exp(H) &= \left(D \sum_{k=0}^{\infty} \frac{1}{k!} \Omega^k \right) (H) \\ &= \sum_{k=1}^{\infty} \frac{1}{k!} (D\Omega^k)(H) \\ &= \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{n=0}^{k-1} \binom{k}{n+1} (\mathrm{ad}_\Omega^n H) \Omega^{k-n-1} \\ &= \sum_{k=1}^{\infty} \sum_{n=0}^{k-1} \frac{1}{(n+1)!(k-n-1)!} (\mathrm{ad}_\Omega^n H) \Omega^{k-n-1}. \end{aligned}$$

Putting $l = k - n - 1$ gives

$$\begin{aligned} T_\Omega \exp(H) &= \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{(n+1)!} \frac{1}{l!} (\mathrm{ad}_\Omega^n H) \Omega^l \\ &= \sum_{n=0}^{\infty} \frac{1}{(n+1)!} (\mathrm{ad}_\Omega^n H) \sum_{l=0}^{\infty} \frac{1}{l!} \Omega^l \\ &= \left(\sum_{n=0}^{\infty} \frac{1}{(n+1)!} \mathrm{ad}_\Omega^n H \right) \exp(\Omega), \end{aligned}$$

which gives the desired expression. Since the linear operator ad_Ω is bounded, $\|\mathrm{ad}_\Omega\| \leq 2\|\Omega\|$, and since the series $\sum_{k=0}^{\infty} \frac{1}{(k+1)!} x^k$, $x \in \mathbb{R}$, has infinite radius of convergence the series $\sum_{n=0}^{\infty} \frac{1}{(n+1)!} \mathrm{ad}_\Omega^n$ converges for all $\Omega \in \mathfrak{g}$. □

The next proposition concerns the inverse of the linear operator $T_\Omega \exp$.

Proposition 30. *Let G be a matrix Lie group. For $\Omega \in \mathfrak{g}$ the linear map $T_\Omega(R_{\exp(-\Omega)} \circ \exp) : \mathfrak{g} \to \mathfrak{g}$ is bijective if and only if no eigenvalue of the operator ad_Ω is of the form $2l\pi\sqrt{-1}$ for a nonzero integer l. If this is the case we have for $\|\Omega\| < \pi$*

$$\left(T_\Omega(R_{\exp(-\Omega)} \circ \exp) \right)^{-1} = \sum_{k=0}^{\infty} \frac{b_k}{k!} \mathrm{ad}_\Omega^k \ ,$$

where $b_k = \left. \frac{\mathrm{d}^k}{\mathrm{d}x^k} \right|_{x=0} \frac{x}{e^x - 1}$ are the Bernoulli numbers .

Proof. Since $\sum_{k=0}^{\infty} \frac{1}{(k+1)!} x^k = \frac{e^x - 1}{x}$, $x \in \mathbb{R}$, the eigenvalues for $T_\Omega(R_{\exp(-\Omega)} \circ \exp) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \mathrm{ad}_\Omega^k$ will be $\mu = \frac{e^\lambda - 1}{\lambda}$, where λ is an eigenvalue for ad_Ω. This means that if and only if no eigenvalue of ad_Ω is of the form $2l\pi\sqrt{-1}$ then all eigenvalues for $T_\Omega(R_{\exp(-\Omega)} \circ \exp)$ will be nonzero which is equivalent to it being invertible. By the definition of the Bernoulli numbers $\sum_{k=0}^{\infty} \frac{b_k}{k!} \mathrm{ad}_\Omega^k$ will be the inverse of $\sum_{k=0}^{\infty} \frac{1}{(k+1)!} \mathrm{ad}_\Omega^k$, and since the radius of convergence for $\sum_{k=0}^{\infty} \frac{b_k}{k!} x^k$ is 2π the series $\sum_{k=0}^{\infty} \frac{b_k}{k!} \mathrm{ad}_\Omega^k$ will converge for $\|\mathrm{ad}_\Omega\| < 2\pi$ but since $\|\mathrm{ad}_\Omega\| \leq 2\|\Omega\|$ this means that the series converges for $\|\Omega\| < \pi$. □

The first Bernoulli numbers can be calculated to be $b_0 = 1$, $b_1 = -\frac{1}{2}$, and $b_{2k+1} = 0$ for $k \in \mathbb{N}$.

Proposition 31. *Let G be an analytic matrix Lie group and let $\xi : [0, t^*] \to \mathfrak{g}$, $t^* > 0$, be piecewise smooth. Then there exists $\delta > 0$ such that for $|\epsilon| < \delta$ the differential equation on G*

$$\dot{g} = \epsilon g \xi(t), \quad g(0) = e, \tag{2.3}$$

has, for $t \in [0, t^*]$, the solution

$$g(t) = \exp\left(\epsilon \int_0^t \xi(s)\mathrm{d}s - \epsilon^2 \tfrac{1}{2} \int_0^t [\xi(s), \int_0^s \xi(\tau)\mathrm{d}\tau]\mathrm{d}s + \mathcal{O}(\epsilon^3)\right).$$

Proof. Let $(U, \log)$, $e \in U$, be a chart with exponential coordinates. Since equation (2.3) depends continuously on ϵ there exists $\delta > 0$ such that for $|\epsilon| < \delta$ the solution to (2.3) stays in U for $t \in [0, t^*]$. Since the differential equation (2.3) depends analytically on ϵ so will the solution $g(t)$ meaning in particular that in the chart $(U, \log)$ the solution is analytic, i.e. $g(t) = \exp(x(t))$ where $x(t) = \sum_{j=1}^{\infty} \epsilon^j x_j(t)$. Inserting $g(t) = \exp(x(t))$ into the differential equation (2.3) gives

$$\begin{aligned}\exp(x(t))\epsilon\xi(t) &= \dot{g}(t) \\ &= T_{x(t)} \exp(\dot{x}(t)) \\ &= T_{x(t)}(R_{\exp(-x(t))} \circ \exp)(\dot{x}(t)) \exp(x(t)).\end{aligned}$$

By Proposition 30 we can ensure by possibly choosing δ smaller that $T_{x(t)}(R_{\exp(-x(t))} \circ \exp)$ is invertible for $t \in [0, t^*]$. Thus we get

$$\begin{aligned}\dot{x}(t) &= \big(T_{x(t)}(R_{\exp(-x(t))} \circ \exp)\big)^{-1}\big(\epsilon \mathrm{Ad}_{\exp(x(t))}(\xi(t))\big) \\ &= \big(T_{x(t)}(R_{\exp(-x(t))} \circ \exp)\big)^{-1}\big(\epsilon \exp(\mathrm{ad}_{x(t)})(\xi(t))\big).\end{aligned}$$

Inserting $x(t) = \sum_{j=1}^{\infty} \epsilon^j x_j(t)$ on both sides of this equation and using Proposition 30 shows that

$$\begin{aligned}\dot{x}_1 &= \xi(t), \\ \dot{x}_2 &= -\tfrac{1}{2}[\xi(t), x_1(t)],\end{aligned}$$

which gives the result. □

The differential equation

$$\dot{x}(t) = \big(T_{x(t)}(R_{\exp(-x(t))} \circ \exp)\big)^{-1}(\xi(t)),$$

is called the **Magnus equation** after W. Magnus who first treated it in 1954, see [30]. The Magnus equation gives the solution $g(t) = \exp(x(t))$ to $\dot{g} = \xi(t)g$. The expansion of $x(t)$ in terms of integrals of repeated Lie brackets, as in Proposition 31, is called the **Magnus expansion**.

Corollary 32. *Let G be an analytic matrix Lie group and let $\eta, \zeta \in \mathfrak{g}$. Then there exists $\delta > 0$ such that for $|\epsilon| < \delta$ we have*

$$\exp(\epsilon\eta)\exp(\epsilon\zeta) = \exp\left(\epsilon(\eta+\zeta) + \epsilon^2 \tfrac{1}{2}[\eta,\zeta] + \mathcal{O}(\epsilon^3)\right).$$

Proof. We use Proposition 31 with $t^* = 2$ and

$$\xi(t) = \begin{cases} \eta, & t \in [0,1) \\ \zeta, & t \in [1,2] \end{cases}.$$

This gives

$$g(2) = \exp(\epsilon\eta)\exp(\epsilon\zeta).$$

But we have

$$\int_0^s \xi(\tau)\mathrm{d}\tau = \begin{cases} \eta s, & s \in [0,1) \\ \eta + \zeta(s-1), & s \in [1,2] \end{cases},$$

and thus

$$\int_0^2 \xi(\tau)\mathrm{d}\tau = \eta + \zeta,$$

and

$$\begin{aligned} \int_0^2 [\xi(s), \int_0^s \xi(\tau)\mathrm{d}\tau]\mathrm{d}s &= \int_0^1 [\eta, \int_0^s \xi(\tau)\mathrm{d}\tau]\mathrm{d}s + \int_1^2 [\zeta, \int_0^s \xi(\tau)\mathrm{d}\tau]\mathrm{d}s \\ &= \int_0^1 [\eta, \eta s]\mathrm{d}s + \int_1^2 [\zeta, \eta + \zeta(s-1)]\mathrm{d}s \\ &= [\zeta, \eta], \end{aligned}$$

which gives the result when inserted in Proposition 31. □

The full Taylor expansion of $\log(\exp(\epsilon\eta)\exp(\epsilon\zeta))$ is given recursively by the so called **Baker-Campbell-Hausdorff formula**, see e.g. [43].

Chapter 3

Simple Mechanical Control Systems on Lie Groups

The subject of this chapter is simple mechanical control systems on Lie groups which is the particular class of mechanical systems we will focus on in the remainder of this thesis.

We start by introducing some concepts from the theory of calculus of variations, since the approach leading to the equations of motions for mechanical systems, in particular the equations of motion for simple mechanical systems on Lie groups, is of a variational nature. We then derive the forced Euler-Lagrange equations, which are the equations of motion for a forced mechanical system, and the Euler-Poincaré equations which, along with the kinematic equations, are the equations of motion for a mechanical system when the configuration manifold is a Lie group. We define simple mechanical control systems on Lie groups which are a special class of mechanical systems, with a Lie group as configuration manifold, and we give the Euler-Poincaré equations for this special case. We end the chapter with some examples of simple mechanical systems on Lie groups.

A more exhaustive treatment of the material covered in this chapter can be found, e.g, in [6],[13], and [31].

3.1 Elements of Calculus of Variations

We start with a standard definition from calculus of variations.

Definition 33. *Let Q be a manifold and $q : [a, b] \to Q$, $a, b \in \mathbb{R}$, $b > a$, a smooth curve on Q. A **variation** of the curve $q : [a, b] \to Q$ is a smooth map $(t, \epsilon) \mapsto q_\epsilon(t) \in Q$, $\epsilon \in [c, d]$, $d > 0$, $c < 0$, satisfying*

1. $q_0(t) = q(t)$.

2. $q_\epsilon(a) = q(a)$ *and* $q_\epsilon(b) = q(b)$ *for all* $\epsilon \in [c, d]$.

*The corresponding **infinitesimal variation** is given by*

$$\delta q(t) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} q_\epsilon(t) \in T_{q(t)}Q.$$

For a smooth function $L : TQ \to \mathbb{R}$ the ***variation*** *of the functional $I(q) = \int_a^b L(\dot{q}(t))\mathrm{d}t$ is defined as*

$$\delta \int_a^b L(\dot{q}(t))\mathrm{d}t = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \int_a^b L(\dot{q}_\epsilon(t))\mathrm{d}t$$

and the ***functional derivative*** *of L, $\frac{\delta L}{\delta q} : TQ \to T^*Q$, is the bundle map over id_Q, i.e., $\pi \circ \frac{\delta L}{\delta q} = \mathrm{id}_Q \circ \tau$, given by*

$$\delta \int_a^b L(\dot{q}(t))\mathrm{d}t = \int_a^b \frac{\delta L}{\delta q}(\dot{q}(t)) \cdot \delta q(t)\mathrm{d}t,$$

if it exists.

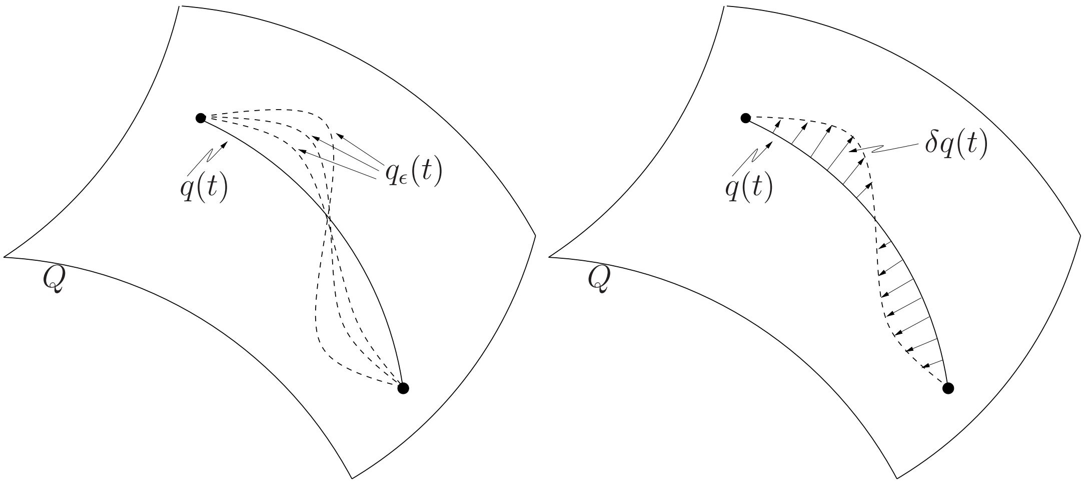

Figure 3.1: The curve $q(t)$ and a variation $q_\epsilon(t)$ of it (left). The infinitesimal variation $\delta q(t)$, of $q(t)$, given by $q_\epsilon(t)$ (right).

From Definition 33 we see that $\delta \int_a^b L(\dot{q}(t))\mathrm{d}t$ also depends on the choice of variation $q_\epsilon(t)$. This is not the case for $\frac{\delta L}{\delta q}$ which is intrinsically defined.

We have the following result regarding the coordinate expression for the functional derivative of L.

Proposition 34. *Let Q be an n-dimensional manifold and (U, q) a coordinate chart on Q. Let $L : TQ \to \mathbb{R}$ be a smooth function. Then $\frac{\delta L}{\delta q}$ exists and we have in TU*

$$\frac{\delta L}{\delta q}(q, \dot{q}) = \sum_{i=1}^n \left(\frac{\partial L}{\partial q^i}(q, \dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i}(q, \dot{q}) \right) \mathrm{d}q^i,$$

where $(q, \dot{q})$ are the natural coordinates on TU corresponding to the coordinates q for U.

Proof. Since we want to calculate a local expression for $\frac{\delta L}{\delta q}$ we may assume for simplicity that $q(t) \in U$ for $t \in [a, b]$. Since U is open we can choose $|\epsilon|$ small enough to ensure

$q_\epsilon([a,b]) \subset U$. Thus we get

$$\begin{aligned}
\delta \int_a^b L(q(t),\dot q(t))\mathrm{d}t &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \int_a^b L(q_\epsilon(t),\dot q_\epsilon(t))\mathrm{d}t \\
&= \int_a^b \left(\frac{\partial L}{\partial q}(q(t),\dot q(t))\cdot \delta q(t) + \frac{\partial L}{\partial \dot q}(q(t),\dot q(t))\cdot \dot{\delta q}(t)\right)\mathrm{d}t \\
&= \int_a^b \left(\frac{\partial L}{\partial q}(q(t),\dot q(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot q}(q(t),\dot q(t))\right)\cdot \delta q(t)\mathrm{d}t \\
&\quad + \left[\frac{\partial L}{\partial \dot q}(q(t),\dot q(t))\cdot \delta q(t)\right]_a^b \\
&= \int_a^b \left(\frac{\partial L}{\partial q}(q(t),\dot q(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot q}(q(t),\dot q(t))\right)\cdot \delta q(t)\mathrm{d}t \\
&=: \int_a^b \frac{\delta L}{\delta q}(q(t),\dot q(t))\cdot \delta q(t)\mathrm{d}t.
\end{aligned}$$

To ensure that we have that $\frac{\delta L}{\delta q}$ is defined globally we must check that this local expression for $\frac{\delta L}{\delta q}$ behaves correctly under a change of coordinates.

Let $(V,\tilde q)$ be a coordinate chart for Q with $U \cap V \neq \emptyset$. From the chain rule we have

$$\dot{\tilde q}^j = \sum_{k=1}^n \frac{\partial \tilde q^j}{\partial q^k}\dot q^k \qquad \Rightarrow \qquad \frac{\partial \dot{\tilde q}^j}{\partial \dot q^i} = \frac{\partial \tilde q^j}{\partial q^i}.$$

Thus we get using the chain rule on $L(q,\dot q) = \widetilde L(\tilde q(q), \dot{\tilde q}(q,\dot q))$

$$\frac{\partial L}{\partial \dot q^i} = \sum_{j=1}^n \frac{\partial \widetilde L}{\partial \dot{\tilde q}^j}\frac{\partial \dot{\tilde q}^j}{\partial \dot q^i} = \sum_{j=1}^n \frac{\partial \widetilde L}{\partial \dot{\tilde q}^j}\frac{\partial \tilde q^j}{\partial q^i},$$

giving

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot q^i}\right) = \sum_{j=1}^n \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \widetilde L}{\partial \dot{\tilde q}^j}\right)\frac{\partial \tilde q^j}{\partial q^i} + \sum_{j,k=1}^n \frac{\partial \widetilde L}{\partial \dot{\tilde q}^j}\frac{\partial^2 \tilde q^j}{\partial q^i \partial q^k}\dot q^k.$$

The second part of $\frac{\delta L}{\delta q}$ is similarly calculated using the chain rule

$$\begin{aligned}
\frac{\partial L}{\partial q^i} &= \sum_{j=1}^n \frac{\partial \widetilde L}{\partial \tilde q^j}\frac{\partial \tilde q^j}{\partial q^i} + \sum_{j=1}^n \frac{\partial \widetilde L}{\partial \dot{\tilde q}^j}\frac{\partial \dot{\tilde q}^j}{\partial q^i} \\
&= \sum_{j=1}^n \frac{\partial \widetilde L}{\partial \tilde q^j}\frac{\partial \tilde q^j}{\partial q^i} + \sum_{j,k=1}^n \frac{\partial \widetilde L}{\partial \dot{\tilde q}^j}\frac{\partial^2 \tilde q^j}{\partial q^i \partial q^k}\dot q^k.
\end{aligned}$$

Combining these we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot q^i} - \frac{\partial L}{\partial q^i} = \sum_{j=1}^n \frac{\partial \tilde q^j}{\partial q^i}\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \widetilde L}{\partial \dot{\tilde q}^j} - \frac{\partial \widetilde L}{\partial \tilde q^j}\right).$$

Since $\pi \circ \frac{\delta L}{\delta q} = \mathrm{id}_Q \circ \tau$ and since $\frac{\delta L}{\delta q}$ under a change of coordinates behaves as a one form we have that $\frac{\delta L}{\delta q} : TQ \to T^*Q$ indeed is a bundle map over id_Q. □

3.2 The Euler-Poincaré Equations

We start with the definition of the Lagrange-d'Alembert principle.

Definition 35. *Let Q be a manifold, $q : [a, b] \to Q$ a smooth curve on Q, and $F : [a, b] \times TQ \to T^*Q$ a bundle map over* id_Q. *Then q is said to satisfy the* ***Lagrange-d'Alembert*** *principle if for every variation $q_\epsilon(t)$ of $q(t)$ with corresponding infinitesimal variation $\delta q(t)$ we have*

$$\delta \int_a^b L(\dot{q}(t))\mathrm{d}t + \int_a^b F(t, \dot{q}(t)) \cdot \delta q(t)\mathrm{d}t = 0.$$

It can be shown, see e.g. [16], that Newtons equations of motion are equivalent to the Lagrange-d'Alembert principle where F is the resultant force and $L = T - V$ is the so-called **Lagrangian** consisting of the kinetic energy T minus the potential energy V.

Proposition 36 (Forced Euler-Lagrange equations)**.** *Let Q be an n-dimensional manifold, $q : [a, b] \to Q$ a smooth curve on Q, and $F : [a, b] \times TQ \to T^*Q$ a bundle map over* id_Q. *The Lagrange-d'Alembert principle is satisfied if and only if*

$$\frac{\delta L}{\delta q}(\dot{q}(t)) + F(t, \dot{q}(t)) = 0, \tag{3.1}$$

which in a coordinate system (U, q) is equivalent to $q(t)$ satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i}(q, \dot{q}) - \frac{\partial L}{\partial q^i}(q, \dot{q}) = F_i(t, q, \dot{q}), \qquad i \in \{1, \ldots, n\}. \tag{3.2}$$

Proof. From Definition 33 we have

$$\delta \int_a^b L(\dot{q}(t))\mathrm{d}t + \int_a^b F(t, \dot{q}(t)) \cdot \delta q(t)\mathrm{d}t = \int_a^b \left(\frac{\delta L}{\delta q}(\dot{q}(t)) + F(t, \dot{q}(t))\right) \cdot \delta q(t)\mathrm{d}t.$$

From Definition 35 and the fundamental lemma of the calculus of variations, see, e.g., [24], we thus obtain equation (3.1). Equation (3.2) is a direct consequence of Proposition 34 and equation (3.1). □

Equation 3.2 is referred to as the forced Euler-Lagrange equations. If $F = 0$ equations (3.2) are the well known Euler-Lagrange equations, see ,e.g., [4].

Using the chain rule the forced Euler-Lagrange equations can be written

$$\frac{\partial^2 L}{\partial \dot{q}^2}\ddot{q} + \frac{\partial^2 L}{\partial \dot{q}\partial q}\dot{q} - \frac{\partial L}{\partial q} = F,$$

which when written out completely is

$$\sum_{k=1}^{n} \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^k}\ddot{q}^k + \sum_{k=1}^{n} \frac{\partial^2 L}{\partial \dot{q}^j \partial q^k}\dot{q}^k - \frac{\partial L}{\partial q_j} = F_j, \quad j \in \{1, \ldots, n\},$$

where $q = (q^1, \ldots, q^n)$, $F = (F_1, \ldots, F_n)$.

Let G denote a Lie group. Then a Lagrangian $L : TG \to \mathbb{R}$ is called left-invariant if $L(g, \dot{g}) = L(L_h(g), T_g L_h(\dot{g}))$, $\dot{g} \in T_g G$, for all $g, h \in G$. For a matrix Lie group this means that

$$L(g, \dot{g}) = L(L_{g^{-1}}(g), T_g L_{g^{-1}}(\dot{g})) = L(e, g^{-1}\dot{g}) = L(e, \xi) =: l(\xi),$$

where $\xi := T_g L_{g^{-1}}(\dot{g}) = g^{-1}\dot{g} \in T_e G = \mathfrak{g}$. l is called the restriction of L to $\mathfrak{g}$.

Proposition 37 (The Euler-Poincaré equations)**.** *Let G be a matrix Lie group and $L : TG \to \mathbb{R}$ a left invariant Lagrangian and l its restriction to $\mathfrak{g}$. For a curve $g(t) \in G$ define the curve $\xi(t) \in \mathfrak{g}$ by*

$$\xi(t) = g(t)^{-1}\dot{g}(t). \tag{3.3}$$

Let the force $F : \mathbb{R} \times G \to T^*G$ be given by $F(t, g) = T_g^* L_{g^{-1}} f(t) = f(t)g^{-1}$ where $f(t) \in \mathfrak{g}^*$ is the body-fixed force. Then $g(t)$ satisfies the Lagrange-d'Alembert principle if and only if $\xi(t)$ satisfies the Euler-Poincaré equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta \xi} = \mathrm{ad}_\xi^* \frac{\delta l}{\delta \xi} + f(t). \tag{3.4}$$

Proof. Let $g(t) \in G$ be a curve in G and $g_\epsilon(t)$ a variation of $g(t)$. This gives a variation $\xi_\epsilon(t) = \mathfrak{g}_\epsilon(t)^{-1}\dot{g}_\epsilon(t) \in \mathfrak{g}$ of $\xi(t)$. The infinitesimal variation of $g(t)$ is given by $\delta g(t) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\big|_{\epsilon=0} g_\epsilon(t) \in T_{g(t)}G$ and the infinitesimal variation of $\xi(t)$ is given by

$$\delta\xi(t) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \xi_\epsilon(t) \in \mathfrak{g}.$$

If we define $\eta(t) = g^{-1}(t)\delta g(t) \in \mathfrak{g}$ we get

$$\begin{aligned}
\delta\xi(t) - \dot{\eta}(t) &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \left(g_\epsilon(t)^{-1}\dot{g}_\epsilon(t)\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(g_0(t)^{-1} \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} g_\epsilon(t)\right) \\
&= -g_0(t)^{-1}\delta g(t) g_0(t)^{-1}\dot{g}_0(t) + g_0(t)^{-1}\dot{\delta g}(t) + g_0(t)^{-1}\dot{g}_0(t)g_0(t)^{-1}\delta g(t) - g_0(t)^{-1}\dot{\delta g}(t) \\
&= -g(t)^{-1}\delta g(t) g(t)^{-1}\dot{g}(t) + g(t)^{-1}\dot{g}(t)g(t)^{-1}\delta g(t) \\
&= \xi(t)\eta(t) - \eta(t)\xi(t) \\
&= \mathrm{ad}_{\xi(t)}\eta(t).
\end{aligned}$$

This means that we have

$$\delta\xi(t) = \mathrm{ad}_{\xi(t)}\eta(t) + \dot{\eta}(t)$$

where $\eta(t) \in \mathfrak{g}$ vanishes at the endpoints. Since $F \cdot \delta g = fg^{-1}\delta g = f \cdot \eta$ we therefore have that the Lagrange-d'Alembert equations

$$\delta \int_a^b L(g(t), \dot{g}(t))\mathrm{d}t + \int_a^b F(t, g(t)) \cdot \delta g(t)\mathrm{d}t = 0$$

are equivalent to

$$\delta \int_a^b l(\xi(t))\mathrm{d}t + \int_a^b f(t) \cdot \eta(t)\mathrm{d}t = 0 \tag{3.5}$$

using variations $\delta\xi(t) = \mathrm{ad}_{\xi(t)}\eta(t) + \dot{\eta}(t)$ and where η vanishes at the endpoints. Calculating

$$\begin{aligned}
\delta \int_a^b l(\xi(t))\mathrm{d}t &= \int_a^b \frac{\delta l}{\delta \xi}\delta\xi(t)\mathrm{d}t \\
&= \int_a^b \frac{\delta l}{\delta \xi}(\dot{\eta}(t) + \mathrm{ad}_{\xi(t)}\eta(t))\mathrm{d}t \\
&= \int_a^b \left(-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta \xi}\eta(t) + \frac{\delta l}{\delta \xi}\mathrm{ad}_{\xi(t)}\eta(t)\right)\mathrm{d}t + \left[\frac{\delta l}{\delta \xi}\eta(t)\right]_a^b \\
&= \int_a^b \left(-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta \xi} + \mathrm{ad}^*_{\xi(t)}\frac{\delta l}{\delta \xi}\right)\eta(t)\mathrm{d}t,
\end{aligned}$$

and inserting this expression into equation (3.5) the result is a consequence of the fundamental lemma from the calculus of variations, see, e.g., [24]. □

Equation (3.3) is referred to as the kinematic equations for obvious reasons. Together the kinematic equations (3.3) and the Euler-Poincaré equations (3.4) give the equation of motion for a forced mechanical system on a Lie group.

Definition 38. *A* ***simple mechanical control system on a Lie group*** *is a mechanical system described by the following:*

1. *the configuration manifold G is a matrix Lie group, with Lie algebra $\mathfrak{g}$,*
2. *the total energy is equal to the kinetic energy which is given by an inertia tensor $\mathbb{I} : \mathfrak{g} \to \mathfrak{g}^*$,*
3. *a set of body-fixed vectors $\{f_1, \ldots, f_m\} \subset \mathfrak{g}^*$ and $u : \mathbb{R} \to \mathbb{R}^m$, bounded and measurable, defining the resultant body-fixed force according to $f(t) = \sum_{i=1}^m f_i u_i(t)$.*

$\Sigma = (G, \mathbb{I}, \{f_1, \ldots, f_m\})$ denotes this mechanical control system.

For simple mechanical control systems on Lie groups we have.

Proposition 39. *Let $\Sigma = (G, \mathbb{I}, \{f_1, \ldots, f_m\})$ be a simple mechanical control system, then the equations of motion for this system are*

$$\dot{g} = g \cdot \xi, \tag{3.6}$$

$$\mathbb{I}\dot{\xi} = \mathrm{ad}^*_\xi \mathbb{I}\xi + \sum_{i=1}^m f_i u_i(t). \tag{3.7}$$

Proof. This follows directly from Proposition 37 noticing that $l(\xi) = \frac{1}{2}\mathbb{I}(\xi) \cdot \xi$ and using Proposition 34 giving the functional derivative as $\frac{\delta l}{\delta \xi} = \frac{\partial l}{\partial \xi} = \mathbb{I}\xi$. □

We define the **symmetric product** $\langle \cdot : \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ by

$$\langle \xi : \eta \rangle := -\mathbb{I}^{-1}(\mathrm{ad}^*_\xi \mathbb{I}\eta + \mathrm{ad}^*_\eta \mathbb{I}\xi),$$

which is seen to be bilinear and symmetric. Using the symmetric product the Euler-Poincaré equations (3.7) can be written

$$\dot{\xi} = -\tfrac{1}{2}\langle \xi : \xi \rangle + \sum_{i=1}^{m} b_i u_i(t),$$

where $b_i = \mathbb{I}^{-1} f_i$, for $i \in \{1, \ldots, m\}$.

A **relative equilibrium** for Σ is a curve $t \mapsto g_0 \exp(t\xi_{\mathrm{re}}) \in G$, for $g_0 \in G$ and $\xi_{\mathrm{re}} \in \mathfrak{g}$, that is a solution to the dynamics (3.6), (3.7) for zero input u. It is easy to see that $t \mapsto g_0 \exp(t\xi_{\mathrm{re}})$ is a relative equilibrium if and only if $\langle \xi_{\mathrm{re}} : \xi_{\mathrm{re}} \rangle = 0$. It is convenient to call relative equilibrium both the curve $t \mapsto g_0 \exp(t\xi_{\mathrm{re}})$ and the vector ξ_{re}.

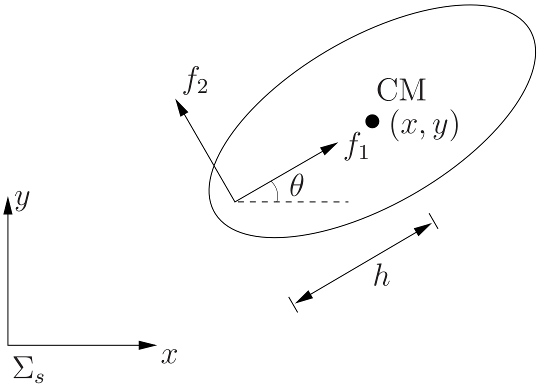

Figure 3.2: The planar rigid body with two forces applied at a point a distance h from the center of mass CM. Σ_s denotes an inertial reference frame. $(\theta, x, y) \in S \times \mathbb{R}^2$ denotes the configuration of the body. The body reference frame (not depicted) is aligned with the direction of application of f_1 and f_2.

3.3 Examples

Example 1 (Underactuated planar rigid body)**.** Consider a rigid body moving in the plane without friction, see figure 3.2. This is, e.g., a model of a hovercraft when disregarding friction.

The configuration space for this system is the matrix Lie group $SE(2)$, with Lie algebra $\mathfrak{se}(2)$. Since $SE(2)$ is isomorphic to $SO(2) \times \mathbb{R}^2$ and $SO(2)$ is isomorphic to $S \subset \mathbb{C} \setminus \{0\}$, equipped with the product induced from the Lie group $\mathbb{C} \setminus \{0\}$, $SE(2)$ is isomorphic to $S \times \mathbb{R}^2$, with product

$$(e^{i\theta_1}, (x_1, y_1))(e^{i\theta_2}, (x_2, y_2)) = (e^{i(\theta_1+\theta_2)}, (\mathrm{RE}(e^{i\theta_1}(x_2 + iy_2)) + x_1), \mathrm{IM}(e^{i\theta_1}(x_2 + iy_2)) + y_1)),$$

where $i = \sqrt{-1}$. The diffeomorphism giving this isomorphism is for $(\theta, (x, y)) \in S \times \mathbb{R}^2$

$$P = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & x \\ \sin(\theta) & \cos(\theta) & y \\ 0 & 0 & 1 \end{bmatrix} \in SE(2).$$

If we take $(\theta, x, y, \omega, v_1, v_2) \in T(S \times \mathbb{R}^2) = (S \times \mathbb{R}^2) \times \mathbb{R}^3$ and define the vector space isomorphism $\widehat{\cdot} : \mathbb{R}^3 \to \mathfrak{se}(2)$ given by

$$\widehat{\begin{bmatrix} \omega \\ v_1 \\ v_2 \end{bmatrix}} = \begin{bmatrix} 0 & -\omega & v_1 \\ \omega & 0 & v_2 \\ 0 & 0 & 0 \end{bmatrix},$$

we thus have a diffeomorphism between $SE(2) \times \mathfrak{se}(2)$ and $(S \times \mathbb{R}^2) \times \mathbb{R}^3$. The Lie bracket on $\mathbb{R}^3$ induced by $\widehat{\cdot}$ gives

$$\mathrm{ad}_{(\omega, v_1, v_2)^T} = \begin{bmatrix} 0 & 0 & 0 \\ v_2 & 0 & -\omega \\ -v_1 & \omega & 0 \end{bmatrix}$$

With controls as in the figure we have

$$f_1 = e_2, \qquad f_2 = -he_1 + e_3,$$

where h is the distance from the center of mass to the control forces. Denote by m the mass of the body and by J the moment of inertia about its center of mass. Since the kinetic energy is $\frac{1}{2}m(v_1^2 + v_2^2) + \frac{1}{2}J\omega^2$ the inertia tensor is $\mathbb{I} = \mathrm{diag}(J, m, m)$ and the equations of motion become, using proposition 39

$$\dot{P} = P \begin{bmatrix} 0 & -\omega & v_1 \\ \omega & 0 & v_2 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} J\dot{\omega} &= -hu_2 \\ m\dot{v}_1 &= m\omega v_2 + u_1 \\ m\dot{v}_2 &= -m\omega v_1 + u_2 \end{aligned}$$

The kinematic equation can be rewritten to

$$\begin{aligned} \dot{\theta} &= \omega \\ \dot{x} &= \cos(\theta)v_1 - \sin(\theta)v_2 \\ \dot{y} &= \sin(\theta)v_1 + \cos(\theta)v_2 \end{aligned}$$

on $S \times \mathbb{R}^2$.

We see that the vectors

$$\xi = e_1, \qquad \xi = \alpha e_2 + \beta e_3,$$

where $\alpha, \beta \in \mathbb{R}$, are relative equilibria for the system.

Example 2 (Satellite with two thrusters)**.** Consider a satellite, i.e., a rigid body floating in space, subject to a torque around the first and second principal axes, see figure 3.3.

The configuration manifold for a rigid body is the matrix Lie group $G = SO(3)$ with Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$. The isomorphism $\widehat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ given by $\widehat{x}y := x \times y$, $x, y \in \mathbb{R}^3$, that is,

$$\widehat{x} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix},$$

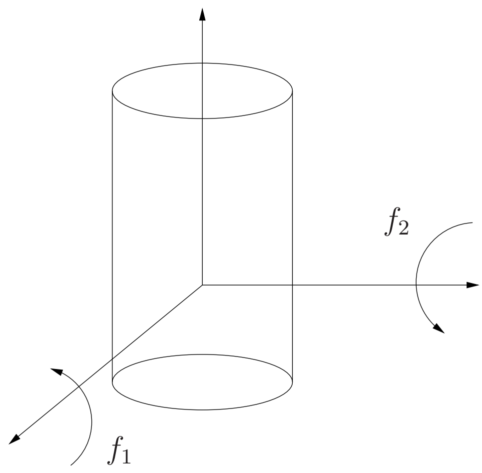

Figure 3.3: The satellite with two thrusters.

is a Lie algebra isomorphism between $\mathbb{R}^3$ with the cross product and $\mathfrak{so}(3)$ with the matrix commutator; thus $\mathrm{ad}_x = \widehat{x}$. With forces as in figure 3.3 the control forces are

$$f_1 = e_1, \qquad\qquad f_2 = e_2.$$

Using proposition 39 then gives that the dynamics of this system is given by

$$\dot{R} = R\widehat{\Omega},$$
$$\mathbb{J}\dot{\Omega} = (\mathbb{J}\Omega) \times \Omega + e_1 u_1(t) + e_2 u_2(t),$$

where $R \in SO(3)$, $\Omega \in \mathbb{R}^3$, and $\mathbb{J} = \mathrm{diag}(J_1, J_2, J_3)$, J_i being the moment of inertia around the ith principal axis.

A vector along any of the principal axes is seen to be a relative equilibrium.

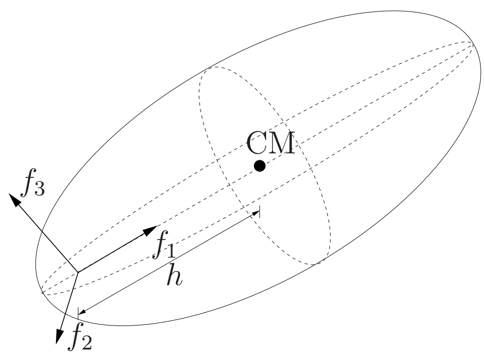

Figure 3.4: A schematic of the underwater vehicle.

Example 3 (Underwater vehicle in Ideal Fluid)**.** Consider a rigid body submerged in an incompressible, irrotational, and inviscid fluid. The configuration manifold for this system is $SE(3)$. The motion of this system is Hamiltonian, see e.g. [26], meaning that the theory in this chapter can be applied. Let $g \in SE(3)$ and $\xi \in \mathfrak{se}(3)$ be given by

$$g = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix}, \qquad \xi = \begin{bmatrix} \widehat{\Omega} & v \\ 0 & 0 \end{bmatrix},$$

where $\Omega, v \in \mathbb{R}^3$ and $\widehat{\cdot}: \mathfrak{so}(3) \to \mathbb{R}^3$ is the isomorphism given in the previous example. The kinematic equation (3.6) then reduces to

$$\begin{aligned} \dot{R} &= R\widehat{\Omega}, \\ \dot{p} &= Rv. \end{aligned}$$

The kinetic energy for this system is given by $\frac{1}{2}\Omega^T\mathbb{J}\Omega + \frac{1}{2}v^T\mathbb{M}v$, where $\mathbb{J} = \mathrm{diag}(J_1, J_2, J_3)$ is the inertia matrix and $\mathbb{M} = \mathrm{diag}(m_1, m_2, m_3)$ comprises the added masses, which describes the inertia added to the system due to the fact that moving the body also means moving some of the surrounding fluid. This means that $\mathbb{I} = \mathrm{diag}(\mathbb{J}, \mathbb{M})$ and since

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

the Euler-Poincaré equations (3.7) for this system are

$$\begin{aligned} \mathbb{J}\dot{\Omega} &= (\mathbb{J}\Omega) \times \Omega + (\mathbb{M}v) \times v + f_\Omega, \\ \mathbb{M}\dot{v} &= (\mathbb{M}v) \times \Omega + f_v, \end{aligned}$$

where $f = (f_\Omega, f_v) \in \mathfrak{se}(3)^*$ is the resultant body-fixed force. With forces as in figure 3.4 we have

$$f_1 = e_4, \qquad f_2 = -he_3 + e_5, \qquad f_3 = he_2 + e_6.$$

Any vector $\xi \in \mathfrak{se}(3)$ of the form

$$\xi = \alpha e_i + \beta e_{i+3}, \quad i \in \{1, 2, 3\},$$

where $\alpha, \beta \in \mathbb{R}$, is seen to be a relative equilibrium.

Chapter 4

Elements of Controllability Theory

In this chapter we present some elements from controllability theory in order to do a controllability analysis of simple mechanical control systems on Lie groups.

We begin by introducing some concepts from the theory of controllability of affine control systems and present some of the strongest theorems available regarding local controllability properties of these systems. We then review control results for simple mechanical control systems on Lie groups and prove an additional result regarding local controllability along a relative equilibrium for a simple mechanical system on a Lie group; this result is new and one of the main contributions of this thesis. We end by applying the theory to three example systems and thus provide a controllability analysis of these systems.

Standard references in nonlinear control theory include [21], [34], and [40].

4.1 Controllability of Affine Control Systems

Let Q denote a smooth manifold and consider the affine control system on Q given by

$$\dot{q} = X(q) + \sum_{j=1}^{m} Y_j(q) u_j, \tag{4.1}$$

where $X, Y_1, \ldots, Y_m$ are C^∞ vector fields on Q and the controls $u_1, \ldots, u_m$ are bounded and measurable functions defined for some time interval $[0, T]$, $T > 0$.

Let $\overline{\mathrm{Lie}}(X, Y_1, \ldots, Y_m)$ denote the Lie algebra generated by the elements of $\{X, Y_1, \ldots, Y_m\}$. The system (4.1) is said to satisfy the **Lie algebra rank condition** (LARC) at $q \in Q$ if $\overline{\mathrm{Lie}}(X, Y_1, \ldots, Y_m)(q) = T_q Q$.

Let $q_0 \in Q$ and let $W \subset Q$ be a neighborhood of q_0. For $T > 0$ we define

$$\begin{aligned} \mathcal{R}_Q^W(q_0, T) = \{q_1 \in Q |\ & \text{there exists a solution } (q, u)(t) \text{ of the system (4.1)} \\ & \text{such that } q(0) = q_0,\ q(t) \in W \text{ for } t \in [0, T],\ \text{and } q(T) = q_1\}, \end{aligned}$$

and

$$\mathcal{R}_Q^W(q_0, \leq T) = \bigcup_{t \in [0,T]} \mathcal{R}_Q^W(q_0, t).$$

Then we have the following definition.

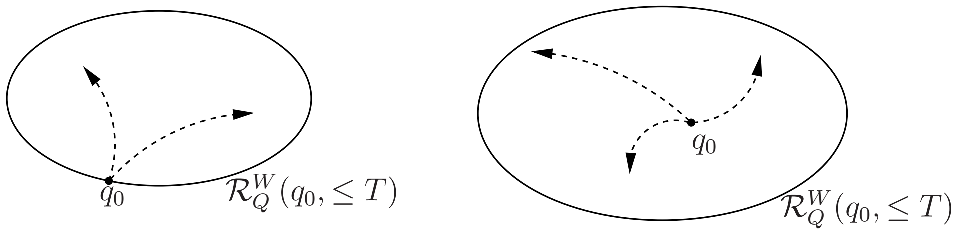

Figure 4.1: (left) locally accessible at q_0. (right) small-time locally controllable (STLC) at q_0.

Definition 40. *The system* (4.1) *is called* ***locally accessible*** *at $q_0 \in Q$ if there exists $T > 0$ such that $\mathcal{R}_Q^W(q_0, \leq t)$ contains a nonempty open set of Q for all neighborhoods W of q_0 for all $t \in (0, T]$. If the system is locally accessible for all $q_0 \in Q$ it is said to be* ***locally accessible****.*

Let $q_0 \in Q$ satisfy $X(q_0) = 0$. Then the system (4.1) *is said to be* ***small-time locally controllable*** *(STLC) at $q_0 \in Q$ if it is locally accessible at q_0 and q_0 belongs to the interior of $\mathcal{R}_Q^W(q_0, \leq t)$ for all $t \in (0, T]$.*

We have the following theorem regarding local accessibility.

Theorem 41. *Consider the system* (4.1)*. It is locally accessible at $q_0 \in Q$ if the LARC is satisfied at q_0. Conversely if the system is locally accessible then the LARC is satisfied in an open and dense subset of Q.*

The proof of this can be found in, e.g., [34].

Let $\mathrm{Br}(X, Y_1, \ldots, Y_m)$ denote the smallest subset of $\overline{\mathrm{Lie}}(X, Y_1, \ldots, Y_m)$ which contains $\{X, Y_1, \ldots, Y_m\}$ and is closed under the operation of taking brackets of its elements, i.e. if $B_1, B_2 \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$ then $[B_1, B_2] \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$. For $B \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$ we define $\delta^0(B)$ to be the number of times X occurs in B and $\delta^j(B)$, $j \in \{1, \ldots, m\}$, the number of times Y_j occurs in B. A Lie bracket $B \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$ is said to be **bad** if $\delta^0(B)$ is odd and $\delta^1(B), \ldots, \delta^m(B)$ are even; otherwise it is said to be **good**. For $\theta \in [0, 1]$ define the **order** of a bracket $B \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$ as the number

$$\delta_\theta(B) = \theta\delta^0(B) + \sum_{j=1}^{m} \delta^j(B).$$

Remark 42. The above definitions can be made more precise using the notion of a free Lie algebra. We have chosen to avoid the notion of a free algebra in this chapter for simplicity of the presentation. •

This enables us to state one of the strongest results regarding STLC of a system. A weaker version of this result was first conjectured in [19].

Theorem 43 (Sussmann [41])**.** *Consider the system* (4.1) *and a point $q_0 \in Q$ satisfying $X(q_0) = 0$. Assume that the LARC is satisfied at q_0. Assume there exists a $\theta \in [0, 1]$,* *giving the order δ_θ, such that every bad Lie bracket $B \in \mathrm{Br}(X, Y_1, \ldots, Y_m)(q_0)$ is a linear combination of lower order good Lie brackets from $\mathrm{Br}(X, Y_1, \ldots, Y_m)(q_0)$.*

Then the system is STLC at q_0.

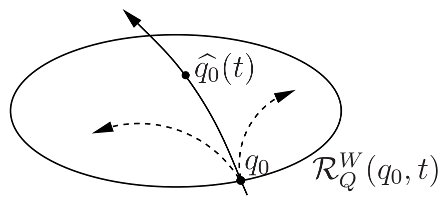

Figure 4.2: Locally controllability along the trajectory $\widehat{q_0}(t)$.

Let $\widehat{q_0}(t)$ be the solution of (4.1) for $u = 0$ satisfying $\widehat{q_0}(0) = q_0$ for $q_0 \in Q$. Then the system is said to be **locally controllable along the trajectory** $\widehat{q_0}(t)$ if there exists $T > 0$ such that $\widehat{q_0}(t)$ lies in the interior of $\mathcal{R}^W_Q(q_0, t)$ for all $t \in (0, T]$. This reduces to STLC when q_0 is an equilibrium point.

We have the following result regarding local controllability along a trajectory.

Theorem 44 (Bianchini and Stefani [5])**.** *Consider the system (4.1) and a point $q_0 \in Q$. Assume that the LARC is satisfied at q_0. Take as weight $\theta = 0$ giving the order δ_θ of a bracket. Assume that every Lie bracket $B \in \mathrm{Br}(X, Y_1, \ldots, Y_m)(q_0)$ with $\delta^i(B) = \text{even}$, $i \in \{1, \ldots, m\}$, is a linear combination of lower order Lie brackets from $\mathrm{Br}(X, Y_1, \ldots, Y_m)(q_0)$.*

Then the system is locally controllable along $\widehat{q_0}(t)$.

This theorem with q_0 being an equilibrium point for X is seen to be contained in Theorem 43.

Another theorem is the following.

Theorem 45 (Bianchini and Stefani [5])**.** *Consider the system (4.1) and a point $q_0 \in Q$. Assume that the LARC is satisfied at q_0. Take as weight $\theta = 1$ defining the order δ_θ of a bracket. Assume that every subspace of $\overline{\mathrm{Lie}}(X, Y_1, \ldots, Y_m)$ has constant rank along $\widehat{q_0}(t)$. Assume furthermore that every bad Lie bracket $B \in \mathrm{Br}(X, Y_1, \ldots, Y_m)(\widehat{q_0}(t))$ is a linear combination of lower order good Lie brackets from $\mathrm{Br}(X, Y_1, \ldots, Y_m)(\widehat{q_0}(t))$ at each point $\widehat{q_0}(t)$ of the reference trajectory.*

Then the system is locally controllable along $\widehat{q_0}(t)$.

Also this theorem is seen to be contained in Theorem 43 when q_0 is an equilibrium point for X. Actually the main result in [5] contains the main result in [41] which has Theorem 43 as a corollary.

Remark 46. Theorem 45 is not in this exact form in [5] but is a consequence of Theorem 1.1 and Lemma 1.2 in [5] with $l = (1, \ldots, 1)$. •

4.2 Controllability of Simple Mechanical Control Systems on Lie Groups

Let $\Sigma = (G, \mathbb{I}, \{f_1, \ldots, f_m\})$ be a simple mechanical control system on a Lie group. Then we consider the Euler-Poincare equations along with the kinematic equations

$$\dot{g} = g \cdot \xi, \tag{4.2}$$

$$\dot{\xi} = -\tfrac{1}{2}\langle \xi : \xi \rangle + \sum_{i=1}^{m} b_i u_i(t), \tag{4.3}$$

as given in chapter 3. Let $g_0 \in G$ and $\xi_0 \in \mathfrak{g}$ and let $U \subset G$ be a neighborhood of g_0. For $T > 0$ we define

$$\begin{aligned}\mathcal{R}_G^U(g_0, T) = \{g_1 \in G|\ &\text{there exists a solution } (g,u)(t) \text{ of the system (4.2)-(4.3)}\\ &\text{such that } \dot{g}(0) = 0,\ g(t) \in U \text{ for } t \in [0,T], \text{ and } g(T) = g_1\},\end{aligned}$$

and

$$\mathcal{R}_G^U(g_0, \leq T) = \bigcup_{t \in [0,T]} \mathcal{R}_G^U(g_0, t).$$

Then we have the following definition.

Definition 47. *The system* (4.2)-(4.3) *is* ***locally configuration accessible*** *at g_0 if there exists $T > 0$ such that $\mathcal{R}_G^U(g_0, \leq t)$ contains a nonempty open set of G for all neighborhoods U of g_0 for all $t \in (0,T]$. If g_0 belongs to the interior of the open set the system is called* ***small-time locally configuration controllable*** *(STLCC) at g_0. If the system is locally configuration accessible (STLCC) for all $g_0 \in G$ it is said to be* ***locally configuration accessible*** *(small-time locally configuration controllable (STLCC)).*

Let $V \subset G \times \mathfrak{g}$ be a neighborhood of (g_0, ξ_0). Converting the definition of $\mathcal{R}_Q^W$ from the previous section to the system (4.2)-(4.3) gives

$$\begin{aligned}\mathcal{R}_{G\times\mathfrak{g}}^V\big((g_0,\xi_0), T\big) = \{(g_1,\xi_1) \in G \times \mathfrak{g}|\ &\text{there exists a solution } (g,u)(t) \text{ of the system (4.2)-(4.3)}\\ &\text{such that } (g,\xi)(0) = (g_0,\xi_0),\ (g,\xi)(t) \in V \text{ for } t \in [0,T],\\ &\text{and } (g,\xi)(T) = (g_1,\xi_1)\},\end{aligned}$$

and

$$\mathcal{R}_{G\times\mathfrak{g}}^V\big((g_0,\xi_0), \leq T\big) = \bigcup_{t \in [0,T]} \mathcal{R}_{G\times\mathfrak{g}}^V\big((g_0,\xi_0), t\big).$$

Then we have the following definition.

Definition 48. *If the system system* (4.2)-(4.3) *is locally accessible at $(g_0, 0)$ and $(g_0, 0)$ belongs to the interior of $\mathcal{R}_{G\times\mathfrak{g}}^V\big((g_0,0), \leq t\big)$, for all $t \in (0,T]$, the system is called* ***small-time locally controllable*** *(STLC)* ***at g_0 and at zero velocity***.

If the system (4.2)-(4.3) *is STLC at g_0 and at zero velocity for all $g_0 \in G$ it is said to be* ***small-time locally controllable at zero velocity*** *(STLC at zero velocity).*

A symmetric algebra is an algebra where the multiplication, denoted by $(x, y) \mapsto \langle x : y \rangle$, satisfies $\langle x : y \rangle = \langle y : x \rangle$. We denote by $\overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ the symmetric algebra generated by the vectors $b_1, \ldots, b_m \in \mathfrak{g}$ and the symmetric product $\langle \cdot : \cdot \rangle$ on $\mathfrak{g}$.

Proposition 49. *Consider the system* (4.2)-(4.3). *The system satisfies the LARC if the subspace defined by* $\overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ *has full rank.*

Proof. We calculate brackets. Let $\eta, \zeta \in \mathfrak{g}$ be fixed. Exploiting the bilinearity of $\langle \cdot : \cdot \rangle$ gives

$$\begin{aligned}\left[\begin{bmatrix} g \cdot \xi \\ -\frac{1}{2}\langle \xi : \xi \rangle \end{bmatrix}, \begin{bmatrix} 0 \\ \eta \end{bmatrix}\right] &= 0 - \begin{bmatrix} * & g \\ 0 & -\langle \xi : \cdot \rangle \end{bmatrix} \begin{bmatrix} 0 \\ \eta \end{bmatrix} \\ &= \begin{bmatrix} -g \cdot \eta \\ \langle \xi : \eta \rangle \end{bmatrix}, \end{aligned} \tag{4.4}$$

and

$$\begin{aligned}\left[\begin{bmatrix} 0 \\ \zeta \end{bmatrix}, \left[\begin{bmatrix} g \cdot \xi \\ -\frac{1}{2}\langle \xi : \xi \rangle \end{bmatrix}, \begin{bmatrix} 0 \\ \eta \end{bmatrix}\right]\right] &= \begin{bmatrix} * & 0 \\ 0 & \langle \eta : \cdot \rangle \end{bmatrix} \begin{bmatrix} 0 \\ \zeta \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ \langle \eta : \zeta \rangle \end{bmatrix}. \end{aligned}$$

Thus we have with $\eta, \zeta \in \mathrm{span}\{b_1, \ldots, b_m\}$ and $\xi \in \mathfrak{g}$ that for an arbitrary $\kappa \in \overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ there exists a $X \in \overline{\mathrm{Lie}}((g \cdot \xi, -\frac{1}{2}\langle \xi : \xi \rangle)^T, (0, b_1)^T, \ldots, (0, b_m)^T)$ of the form

$$X = \begin{bmatrix} 0 \\ \kappa \end{bmatrix}.$$

This and equation (4.4) combined with the assumption that $\overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ has full rank show that there must also exist a $Y \in \overline{\mathrm{Lie}}((g \cdot \xi, -\frac{1}{2}\langle \xi : \xi \rangle)^T, (0, b_1)^T, \ldots, (0, b_m)^T)$ of the form

$$Y = \begin{bmatrix} g \cdot \kappa \\ 0 \end{bmatrix}.$$

Since g is nonsingular the Lie algebra rank condition is therefore satisfied. □

Let $\xi \in \mathfrak{g}$ and denote by $\mathrm{Pr}(\xi, b_1, \ldots, b_m)$ the smallest subset of $\overline{\mathrm{Sym}}(\xi, b_1, \ldots, b_m)$ which contains $\{\xi, b_1, \ldots, b_m\}$ and is closed under the operation of taking symmetric products of its elements, i.e., if $S_1, S_2 \in \mathrm{Pr}(\xi, b_1, \ldots, b_m)$ then $\langle S_1 : S_2 \rangle \in \mathrm{Pr}(\xi, b_1, \ldots, b_m)$. For $S \in \mathrm{Pr}(\xi, b_1, \ldots, b_m)$ we define $\Delta^i(S)$, $i \in \{1, \ldots, m\}$, to be the number of times b_i occurs in S. Similarly we define $\Delta^0(S)$ to be the number of times ξ occurs in S. A symmetric product $S \in \mathrm{Pr}(\xi, b_1, \ldots, b_m)$ is said to be **bad** if $\Delta^i(S)$ is even for all $i \in \{1, \ldots, m\}$; otherwise it is said to be **good**. We define the **order** of a symmetric product $S \in \mathrm{Pr}(\xi, b_1, \ldots, b_m)$ to be the number

$$\Delta_0(S) = \sum_{i=1}^{m} \Delta^i(S).$$

The following result is, as pointed out in [11], a direct consequence of the results in [41] and [29] applied to the system (4.2)-(4.3).

Theorem 50. *Consider the system (4.2)-(4.3). Assume that every bad symmetric product* $S \in \mathrm{Pr}(b_1, \ldots, b_m)$ *is a linear combination of lower order good symmetric products from* $\mathrm{Pr}(b_1, \ldots, b_m)$. *Then*

1. *The system is STLC at zero velocity if the subspace defined by* $\overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ *has full rank.*

2. *The system is STLCC if the subspace defined by* $\overline{\mathrm{Lie}}(\overline{\mathrm{Sym}}(b_1, \ldots, b_m))$ *has full rank.*

To prove a similar theorem regarding local controllability along a relative equilibrium of (4.2)-(4.3) we first need two lemmas.

Lemma 51. *Let* G *be a matrix Lie group with corresponding Lie algebra* $\mathfrak{g}$ *and let* $\xi \in \mathfrak{g}$. *Consider the vector fields*

$$Z_1 = \begin{bmatrix} g \cdot f_1(\xi) \\ S_1(\xi) \end{bmatrix}, \qquad\qquad Z_2 = \begin{bmatrix} g \cdot f_2(\xi) \\ S_2(\xi) \end{bmatrix},$$

on $G \times \mathfrak{g}$, *where* $f_1, f_2, S_1, S_2 : \mathfrak{g} \to \mathfrak{g}$ *are differentiable. Then we have*

$$[Z_1, Z_2] = \begin{bmatrix} g \cdot \left(\mathrm{ad}_{f_1(\xi)}(f_2(\xi)) + Df_2(\xi)(S_1(\xi)) - Df_1(\xi)(S_2(\xi)) \right) \\ DS_2(\xi)(S_1(\xi)) - DS_1(\xi)(S_2(\xi)) \end{bmatrix},$$

where D *is the differential.*

Proof. Using Proposition 31 gives

$$\Phi_t^{X_i}\left(\begin{bmatrix} g_0 \\ \xi_0 \end{bmatrix} \right) = \begin{bmatrix} g_0 \exp(f_i(\xi_0)t + \mathcal{O}(t^2)) \\ \xi_0 + S_i(\xi_0)t + \mathcal{O}(t^2) \end{bmatrix}.$$

Using this, and doing Taylor expansions leaving out terms of order $\mathcal{O}(t^2)$ and $\mathcal{O}(s^2)$, the result then follows as an application of Proposition 3. □

Define

$$\begin{aligned} A_0(\xi, b_1, \ldots, b_m) &= \{\xi, b_1, \ldots, b_m\}, \\ A_{i+1}(\xi, b_1, \ldots, b_m) &= A_i(\xi, b_1, \ldots, b_m) \cup \{\langle v_1 : v_2 \rangle, \mathrm{ad}_{v_1}(v_2) |\; v_1, v_2 \in A_i(\xi, b_1, \ldots, b_m)\} \\ A(\xi, b_1, \ldots, b_m) &= A_\infty(\xi, b_1, \ldots, b_m). \end{aligned}$$

Then we have the following useful lemma

Lemma 52. *Consider the system (4.2)-(4.3). Every bracket*

$$B \in \mathrm{Br}\left(\begin{bmatrix} g \cdot \xi \\ -\frac{1}{2}\langle \xi : \xi \rangle \end{bmatrix}, \begin{bmatrix} 0 \\ b_1 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ b_m \end{bmatrix} \right),$$

when evaluated is of the form

$$B = \begin{bmatrix} g \cdot \left(\sum_j \gamma_j f_j(\xi) \right) \\ \sum_i \alpha_i S_i \end{bmatrix},$$

where $\gamma_j, \alpha_i \in \mathbb{R}$, $S_i \in \mathrm{Pr}(\xi, b_1, \ldots, b_m) \setminus \{\xi\}$, *and* $f_j(\xi) \in A(\xi, b_1, \ldots, b_m)$ *if* $\Delta^0(S_i) > 0$ *and* $f_j = 0$ *if* $\Delta^0(S_i) = 0$. *Furthermore if* $\Delta^0(S_i) > 0$ *then* $f_j(\xi)$ *is homogeneous of order* $\Delta^0(S_i) - 1$, *i.e.,* $f_j(a\xi) = a^{\Delta^0(S_i)-1} f_j(\xi)$ *for all* $a \in \mathbb{R}$ *and all* $\xi \in \mathfrak{g}$.

Proof. For $f_i : \mathfrak{g} \to \mathfrak{g}$ smooth and $S_i(\xi) \in \mathrm{Pr}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)$, $i \in \{1, 2\}$, we have from Lemma 51

$$\left[\begin{bmatrix} g \cdot f_1(\xi) \\ S_1(\xi) \end{bmatrix}, \begin{bmatrix} g \cdot f_2(\xi) \\ S_2(\xi) \end{bmatrix}\right] = \begin{bmatrix} g \cdot \left(\mathrm{ad}_{f_1(\xi)}(f_2(\xi)) + Df_2(\xi)(S_1) - Df_1(\xi)(S_2)\right) \\ [S_1(\xi), S_2(\xi)] \end{bmatrix}, \quad (4.5)$$

where

$$\begin{aligned}[S_1(\xi), S_2(\xi)] =& DS_2(\xi)(S_1(\xi)) - DS_1(\xi)(S_2(\xi)) \\ =& \sum_j S_2(\xi\text{−entry } \# \, j \text{ replaced with } S_1) \\ &- \sum_k S_1(\xi\text{−entry } \# \, k \text{ replaced with } S_2).\end{aligned} \quad (4.6)$$

This shows that $[S_1(\xi), S_2(\xi)] = \sum_j \beta_j S_{12j}(\xi)$ where $\beta_j \in \mathbb{R}$ and $S_{12j}(\xi) \in \mathrm{Pr}(\xi, b_1, \ldots, b_m) \setminus \{\xi\}$.

Let

$$X = \begin{bmatrix} g \cdot \xi \\ -\frac{1}{2}\langle \xi : \xi \rangle \end{bmatrix}, \qquad Y_i = \begin{bmatrix} 0 \\ b_i \end{bmatrix}, \quad i \in \{1, \ldots, m\}.$$

The statement is seen to be true for $B \in \{X, Y_1, \ldots, Y_m\}$. We proceed by induction. Assume it is true for $Z_1, Z_2 \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$. Thus

$$Z_1 = \begin{bmatrix} g \cdot (\sum_i \gamma_{1i} f_{1i}(\xi)) \\ \sum_i \alpha_{1i} S_{1i}(\xi) \end{bmatrix}, \qquad Z_2 = \begin{bmatrix} g \cdot \left(\sum_j \gamma_{2j} f_{2j}(\xi)\right) \\ \sum_j \alpha_{2j} S_{2j}(\xi) \end{bmatrix}.$$

Let $k_1 = \Delta^0(S_{1i}(\xi))$ and $k_2 = \Delta^0(S_{2j}(\xi))$.

Since the bracket is bilinear we get using (4.5)

$$[Z_1, Z_2] = \begin{bmatrix} g \cdot \left(\sum_{ij} (\gamma_{1i}\gamma_{2j} \mathrm{ad}_{f_{1i}(\xi)}(f_{2j}(\xi)) + \gamma_{2j}\alpha_{1i} Df_{2j}(\xi)(S_{1i}(\xi)) - \gamma_{1i}\alpha_{2j} Df_{1i}(\xi)(S_{2j}(\xi)))\right) \\ \sum_{ij} \alpha_{1i}\alpha_{2j} [S_{1i}(\xi), S_{2j}(\xi)] \end{bmatrix},$$

where $[S_{1i}(\xi), S_{2j}(\xi)] = \sum_k \beta_{12ijk} S_{12ijk}$ and $\beta_{12ijk} \in \mathbb{R}$ and $S_{12ijk}\mathrm{Pr}(\xi, b_1, \ldots, b_m) \setminus \{\xi\}$. Because of equation (4.6) we have

$$\Delta^0(S_{12ijk}) = k_1 + k_2 - 1 =: k.$$

By definition we get $S_{12ijk} \in \mathrm{Pr}(\xi, b_1, \ldots, b_m) \setminus \{\xi\} \subset A(\xi, b_1, \ldots, b_m)$ and $\mathrm{ad}_{f_{1i}(\xi)}(f_{2j}(\xi)) \in A(\xi, b_1, \ldots, b_m)$. Similarly we have by definition that $Df_{2j}(\xi)(S_{1i}(\xi)) = \sum_k \beta_{21jik} f_{21jik}(\xi)$ and $Df_{1i}(\xi)(S_{2j}(\xi)) = \sum_k \beta_{12ijk} f_{12ijk}(\xi)$ where $\beta_{21jik}, \beta_{12ijk} \in \mathbb{R}$ and $f_{21jik}(\xi), f_{12ijk}(\xi) \in A(\xi, b_1, \ldots, b_m)$.

For $k_1 = 0$ or $k_2 = 0$ we have $\mathrm{ad}_{f_{1i}(\xi)}(f_{2j}(\xi)) = 0$. If $k_1, k_2 > 0$ $\mathrm{ad}_{f_{1i}(\xi)}(f_{2j}(\xi))$ is homogeneous of order $(k_1 - 1) + (k_2 - 1) = k - 1$.

For $k_1 \leq 1$ we have $Df_{1i}(\xi)(S_{2j}(\xi)) = 0$. For $k_1 > 1$ $Df_{1i}(\xi)(S_{2j}(\xi))$ is homogeneous of order $(k_1 - 1) - 1 + k_2 = k - 1$.

When $k_2 \leq 1$ we have $Df_{2j}(\xi)(S_{1i}(\xi)) = 0$. For $k_2 > 1$ $Df_{2j}(\xi)(S_{1i}(\xi))$ is homogeneous of order $(k_2 - 1) - 1 + k_1 = k - 1$.

If the statement is true for $Z_1, Z_2 \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$ it is therefore true for $[Z_1, Z_2]$, and since it is true for $Z_1, Z_2 \in \{X, Y_1, \ldots, Y_m\}$ it is therefore true for all $Z \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$.

□

With these two lemmas we are able to prove the following result regarding local controllability along a relative equilibrium.

Proposition 53. *Consider the system* (4.2)-(4.3). *Let* ξ_{re} *satisfy* $\langle \xi_{\mathrm{re}} : \xi_{\mathrm{re}} \rangle = 0$. *Assume that* $\overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ *has full rank. Assume:*

1. *Every bad symmetric product* $S \in \mathrm{Pr}(b_1, \ldots, b_m)$ *is a linear combination of lower order good symmetric products from* $\mathrm{Pr}(b_1, \ldots, b_m)$.
2. *Every symmetric product* $S \in \mathrm{Pr}(\xi_{\mathrm{re}}, b_1, \ldots, b_m) \setminus \{\xi_{\mathrm{re}}\}$ *is a linear combination of equal and lower order good symmetric products from* $\mathrm{Pr}(b_1, \ldots, b_m)$.
3. *Every bracket* $B \in \mathrm{Br}(\xi_{\mathrm{re}}, b_1, \ldots, b_m) \setminus \{\xi_{\mathrm{re}}\}$, *given by* $[\cdot, \cdot]_{\mathfrak{g}}$, *with order given by* δ_0 *is a linear combination of equal and lower order products from* $\mathrm{Pr}(b_1, \ldots, b_m)$.

Then the system is locally controllable along $(g, \xi)(t) = (g_0 \exp(t\xi_{\mathrm{re}}), \xi_{\mathrm{re}})$ *for all* $g_0 \in G$.

Proof. Since we assume that $\overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ has full rank we know from Proposition 49 that the LARC is satisfied.

From Lemma 51 we have that for $B_1, B_2 \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)$ that B_1, B_2, and $[B_1, B_2]$ when evaluated are

$$B_1 = \sum_i \alpha_{1i} S_{1i}, \qquad B_2 = \sum_i \alpha_{2i} S_{2i}, \qquad [B_1, B_2] = \sum_i \alpha_{12i} S_{12i},$$

where $\alpha_{1i}, \alpha_{2i}, \alpha_{12i} \in \mathbb{R}$ and $S_{1i}, S_{2i}, S_{12i} \in \mathrm{Pr}(\xi, b_1, \ldots, b_m)$. From Lemma 51 we also get

$$\Delta^0(S_{12i}) = \Delta^0(S_{1j}) + \Delta^0(S_{2k}) - 1. \tag{4.7}$$

We choose $\theta = 1$ meaning that for $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)$ the order $\delta_1(B)$ of B is the total number of factors in B. Taking $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)(\xi_{\mathrm{re}})$, which when evaluated is

$$B = \sum_i \alpha_i S_j,$$

where $\alpha_i \in \mathbb{R}$ and $S_j \in \mathrm{Pr}(\xi, b_1, \ldots, b_m)$, we get using (4.7) recursively

$$\Delta^0(S_j) = \begin{cases} 2\delta^0(B) - \delta_1(B) + 1 & , \quad \text{if } 2\delta^0(B) - \delta_1(B) + 1 > 0 \\ 0 & , \quad \text{otherwise} \end{cases}, \tag{4.8}$$

where we recall that $\delta^0(B)$ is the number of times $-\frac{1}{2}\langle \xi : \xi \rangle$ occurs in B. We recall that we have defined the order of a product $S \in \mathrm{Pr}(\xi_{\mathrm{re}}, b_1, \ldots, b_m)$ as $\Delta_0(S) := \sum_{j=1}^m \Delta^j(S)$. Equation 4.8 can be rewritten to

$$\delta_1(B) = \Delta^0(S_j) + 2\Delta_0(S_j) - 1. \tag{4.9}$$

Define

$$X = \begin{bmatrix} g \cdot \xi \\ -\frac{1}{2}\langle \xi : \xi \rangle \end{bmatrix}, \qquad Y = \begin{bmatrix} 0 \\ b_i \end{bmatrix}, \quad i \in \{1, \ldots, m\}.$$

Using Lemma 51 we get

$$[X, Y_i] = \begin{bmatrix} -g \cdot b_i \\ \langle \xi : b_i \rangle \end{bmatrix}, \qquad [Y_i, [X, Y_j]] = \begin{bmatrix} 0 \\ \langle b_i : b_j \rangle \end{bmatrix}.$$

This means that

$$\forall\, S \in \mathrm{Pr}(b_1, \ldots, b_m)\ \exists\, B \in \mathrm{Br}(X, Y_1, \ldots, Y_m) \text{ s.t. } B = \begin{bmatrix} 0 \\ S \end{bmatrix}, \tag{4.10}$$

we denote this bracket by B_S.

From Lemma 52 we have that every bracket $B \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$ when evaluated is of the form

$$B = \begin{bmatrix} g \cdot \left(\sum_j \gamma_j f_j(\xi) \right) \\ \sum_i \alpha_{2i} S_{2i} \end{bmatrix}, \tag{4.11}$$

where $\gamma_j, \alpha_{2i} \in \mathbb{R}$, $S_{2i} \in \mathrm{Pr}(\xi, b_1, \ldots, b_m)$, and $f_j \in A(\xi, b_1, \ldots, b_m)$ if $\Delta^0(S_i) > 0$ and $f_j = 0$ if $\Delta^0(S_i) = 0$. Using Lemma 51 recursively we furthermore have that b_k, $k \in \{1, \ldots, m\}$, occurs the same number of times in S_i as in f_j, if $\Delta^0(S_i) > 0$, and $\delta^k(B) = \Delta^k(S_i)$.

Let B be a bad bracket and take $\xi = \xi_{\mathrm{re}}$. Then there are two situations, here denoted (a) and (b).

(a). $\Delta^0(S_{2i}) > 0$.
According to assumption 2 we have

$$\alpha_{2i} S_{2i} = \sum_j \alpha_{2ij} S_{2ij},$$

where $\alpha_{2ij} \in \mathbb{R}$ and $S_{ij} \in \mathrm{Pr}(b_1, \ldots, b_m)$ is of equal and lower order as S_{2i} and according to assumption 1 we can assume that S_{2ij} is good.

According to assumptions 2 and 3 we have

$$\sum_j \gamma_j f_j(\xi) = \sum_k \beta_k S_{1k},$$

where $\beta_k \in \mathbb{R}$ and $S_{1k} \in \mathrm{Pr}(b_1, \ldots, b_m)$ is of equal and lower order as S_{2i} and according to assumption 1 we can assume that S_{1k} is good. Using Lemma 51 and (4.10) we get

$$[X, B_{S_{1k}}] = \begin{bmatrix} -g \cdot S_{1k} \\ \langle \xi : S_{1k} \rangle \end{bmatrix}.$$

Because of assumption 2 we have that

$$\langle \xi_{\mathrm{re}} : S_{1k} \rangle = \sum_i \alpha_{1ki} S_{1ki}$$

where $\alpha_{1ki} \in \mathbb{R}$ and $S_{1ki} \in \mathrm{Pr}(b_1, \ldots, b_m)$ is of equal and lower order as S_{1k} and because of assumption 1 we can assume that S_{1ki} is good. Therefore

$$\begin{bmatrix} g \cdot S_{1k} \\ 0 \end{bmatrix} = -[X, B_{S_{1k}}] + \sum_i \alpha_{1ki} B_{S_{1ki}}.$$

From the above we get

$$\begin{aligned} B &= \begin{bmatrix} g\cdot(\sum_k \beta_k S_{1k}) \\ \sum_{ij}\alpha_{2ij}S_{2ij}\end{bmatrix} \\ &= \sum_k \beta_k \left(-[X, B_{S_{1k}}] + \sum_i \alpha_{1ki} B_{S_{1ki}}\right) + \sum_{ij}\alpha_{2ij}B_{S_{2ij}}, \end{aligned}$$

where $B_{S_{1k}}$, $B_{S_{1ki}}$, and $B_{S_{2ij}}$ are good since S_{1k}, S_{1ki}, and S_{2ij} are good.

We have from (4.9) that

$$\begin{aligned} \delta_1(B) &= \Delta^0(S_{2i}) + 2\Delta_0(S_{2i}) - 1, \\ \delta_1(B_{S_{1k}}) &= \Delta^0(S_{1k}) + 2\Delta_0(S_{1k}) - 1 = 2\Delta_0(S_{1k}) - 1, \\ \delta_1(B_{S_{1ki}}) &= \Delta^0(S_{1ki}) + 2\Delta_0(S_{1ki}) - 1 = 2\Delta_0(S_{1ki}) - 1, \\ \delta_1(B_{S_{2ij}}) &= \Delta^0(S_{2ij}) + 2\Delta_0(S_{2ij}) - 1 = 2\Delta_0(S_{2ij}) - 1. \end{aligned}$$

Since $\Delta_0(S_{1k}) \leq \Delta_0(S_{2i})$, $\Delta_0(S_{1ki}) \leq \Delta_0(S_{2i})$, $\Delta_0(S_{2ij}) \leq \Delta_0(S_{2i})$, and $\Delta^0(S_{2i}) > 0$ we therefore get

$$\begin{aligned} \delta_1(B_{S_{1k}}) &< \delta_1(B), \\ \delta_1(B_{S_{1ki}}) &< \delta_1(B), \\ \delta_1(B_{S_{2ij}}) &< \delta_1(B). \end{aligned}$$

Finally

$$\delta_1([X, B_{S_{1k}}]) = 1 + \delta_1(B_{S_{1k}}) = 2\Delta_0(S_{1k}) \leq \delta_1(B),$$

But since B is bad $\delta_1(B)$ is odd so we must have $\delta_1([X, B_{S_{1k}}]) < \delta_1(B)$.

(b). $\Delta^0(S_{2i}) = 0$.

By assumption we have $S_{2i} \in \Pr(b_1, \ldots, b_m)$. From Lemma 52 we get

$$B = \begin{bmatrix} 0 \\ \sum_i \alpha_{2i} S_{2i} \end{bmatrix}.$$

Since $\delta^q(B) = \Delta^q(S_{2i})$, $q \in \{1, \ldots, m\}$, B bad means that S_{2i} is bad but then according to assumption 1 we have that

$$B = \begin{bmatrix} 0 \\ \sum_{ij} \alpha_{2ij} S_{2ij} \end{bmatrix},$$

where $\alpha_{2ij} \in \mathbb{R}$ and $S_{2ij} \in \Pr(b_1, \ldots, b_m)$ is good of order $\Delta_0(S_{2ij}) < \Delta_0(S_{2i})$. Thus we get

$$B = \sum_{ij} \alpha_{2ij} B_{S_{2ij}},$$

where $B_{S_{2ij}}$ is good, since $\delta^q(B_{S_{2ij}}) = \Delta^q(S_{2ij})$, $q \in \{1, \ldots, m\}$, and the order of $B_{S_{2ij}}$ is $\delta_1(B_{S_{ij}}) = 2\Delta_0(S_{2ij}) - 1 < \delta_1(B) = 2\Delta_0(S_{2i}) - 1$.

Since every bracket $B \in \mathrm{Br}(X, Y_1, \ldots, Y_m)$ is of the form 4.11 and since $\xi(t) = \xi_{\mathrm{re}}$ we get that every subspace of $\overline{\mathrm{Lie}}(X, Y_1, \ldots, Y_m)$ has constant rank along $(g_0 \exp(\xi_{\mathrm{re}} t), \xi_{\mathrm{re}})$ for all $g_0 \in G$. The result then follows as an application of Theorem 45. □

Remark 54. For $\xi_{\mathrm{re}} = 0$ assumption 2 is automatically satisfied. From Lemma 52 we have that $f_j(\xi)$, in the proof of Proposition 53, is homogeneous of order $k = \Delta^0(S_{2i}) - 1$ for $\Delta^0(S_{2i}) > 0$ and $f_j = 0$ for $\Delta^0(S_{2i}) = 0$. Since

$$\begin{aligned}\Delta^0(S_{2i}) &= 2\delta^0(B) - \delta_1(B) + 1\\ &= \delta^0(B) - \delta_0(B) + 1,\end{aligned}$$

we have that if B is bad then $\Delta^0(S_{2i}) = \mathrm{odd} - \mathrm{even} + 1 = \mathrm{even}$. For $\Delta^0(S_{2i}) > 0$ we thus have that k is odd and $k > 0$. This gives that for a bad bracket we have $f_j(0) = 0$ and therefore assumption 3 is superfluous in the case that $\xi_{\mathrm{re}} = 0$. Therefore when $\xi_{\mathrm{re}} = 0$ Proposition 53 can be simplified to statement 1 of Proposition 50. •

Proposition 53 has the following corollary which is useful in the analysis of a real mechanical system since its assumptions are easy to verify .

Corollary 55. *Consider the system* (4.2)-(4.3)*. Let* ξ_{re} *satisfy* $\langle \xi_{\mathrm{re}} : \xi_{\mathrm{re}} \rangle = 0$*. Assume that* $\mathrm{span}\{b_i, \langle b_i : b_j \rangle \mid i, j \in \{1, \ldots, m\}\}$ *is full rank and* $\langle \xi_{\mathrm{re}} : b_i \rangle, \langle b_i : b_i \rangle \in \mathrm{span}\{b_1, \ldots, b_m\}$*, for* $i \in \{1, \ldots, m\}$*.*

Then the system is locally controllable along $(g, \xi)(t) = (g_0 \exp(t\xi_{\mathrm{re}}), \xi_{\mathrm{re}})$ *for all* $g_0 \in G$*.*

Consider now instead the Euler-Poincare equations only, without the kinematic equation. Then we prove two propositions regarding STLC of this system. The first result shows what can be omitted in Proposition 53 when disregarding the kinematic equation.

Proposition 56. *Consider the system* (4.3)*. Let* ξ_{re} *satisfy* $\langle \xi_{\mathrm{re}} : \xi_{\mathrm{re}} \rangle = 0$*. Assume that* $\overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ *has full rank. Assume:*

1. *Every bad symmetric product* $S \in \mathrm{Pr}(b_1, \ldots, b_m)$ *is a linear combination of lower order good symmetric products from* $\mathrm{Pr}(b_1, \ldots, b_m)$*.*
2. *Every symmetric product* $S \in \mathrm{Pr}(\xi_{\mathrm{re}}, b_1, \ldots, b_m) \setminus \{\xi_{\mathrm{re}}\}$ *is a linear combination of equal and lower order good symmetric products from* $\mathrm{Pr}(b_1, \ldots, b_m)$*.*

Then the system is STLC at ξ_{re}*.*

Proof. Since we assume that $\overline{\mathrm{Sym}}(b_1, \ldots, b_m)$ has full rank we know from Proposition 49 that the LARC is satisfied.

For $S_1, S_2 \in \mathrm{Pr}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)$ we have, due to the bilinearity of the symmetric product, that

$$\begin{aligned}[S_1, S_2] = &\sum_j S_2(\xi\mathrm{-entry}\ \#\ j \text{ replaced with } S_1)\\ &- \sum_j S_1(\xi\mathrm{-entry}\ \#\ j \text{ replaced with } S_2).\end{aligned} \tag{4.12}$$

Using this recursively we can define the map $\widetilde{\cdot} : \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m) \to \overline{\mathrm{Sym}}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)$, which we extend linearly so $\widetilde{\cdot} : \overline{\mathrm{Lie}}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m) \to \overline{\mathrm{Sym}}(\xi, b_1, \ldots, b_m)$.

Since $[b_i, [-\frac{1}{2}\langle \xi : \xi \rangle, b_j]] = \langle b_i : b_j \rangle$ we have

$$\forall\ S \in \mathrm{Pr}(b_1, \ldots, b_m)\ \exists\ B \in \mathrm{Br}(-\tfrac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m) \text{ s.t. } S = \widetilde{B}.$$

This $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)$ we denote B_S, i.e. $\widetilde{B_S} = S$ for $S \in \mathrm{Pr}(b_1, \ldots, b_m)$.

Let $S \in \mathrm{Pr}(\xi_{\mathrm{re}}, b_1, \ldots, b_m)$, and now $\Delta^0(S)$ denotes the number of times ξ_{re} occurs in S and $\Delta^i(S)$, $i \in \{1, \ldots, m\}$, the number of times b_i occurs in S as before. From (4.12) we have for $B_1, B_2 \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)(\xi_{\mathrm{re}})$ that

$$\Delta^0(\widetilde{[B_1, B_2]}) = \Delta^0(\widetilde{B_1}) + \Delta^0(\widetilde{B_2}) - 1, \tag{4.13}$$

where we define Δ^i, $i \in \{0, 1, \ldots, m\}$, on a sum of products with the same factors to be Δ^i applied to one of these products.

We choose $\theta = 1$ meaning that for $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)$ the order $\delta_1(B)$ of B is the total number of factors in B. Using (4.13) recursively gives for $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)(\xi_{\mathrm{re}})$

$$\Delta^0(\widetilde{B}) = \begin{cases} 2\delta^0(B) - \delta_1(B) + 1 & , \quad \text{if } 2\delta^0(B) - \delta_1(B) + 1 > 0 \\ 0 & , \quad \text{otherwise} \end{cases}, \tag{4.14}$$

where we recall that $\delta^0(B)$ is the number of times $X = -\frac{1}{2}\langle \xi : \xi \rangle$ occurs in B. We recall that we have defined the order of a product $S \in \mathrm{Pr}(\xi_{\mathrm{re}}, b_1, \ldots, b_m)$ as $\Delta_0(S) := \sum_{j=1}^m \Delta^j(S)$.

Let $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \ldots, b_m)(\xi_{\mathrm{re}})$ be a bad bracket of order $\delta_1(B)$. Then we have from (4.12) that

$$\widetilde{B} = \sum_i \alpha_i S_i, \tag{4.15}$$

where $\alpha_i \in \mathbb{R}$ and $S_i \in \mathrm{Pr}(\xi_{\mathrm{re}}, b_1, \ldots, b_m) \setminus \{\xi_{\mathrm{re}}\}$ is a bracket of order l. From (4.14) we have

$$\delta_1(B) = \Delta^0(S_i) + 2l - 1.$$

Then there are two situations, here denoted (a) and (b).

(a). $\Delta^0(S_i) > 0$.

According to assumption 2 equation (4.15) becomes

$$\widetilde{B} = \sum_{ij} \alpha_{ij} S_{ij},$$

where $\alpha_{ij} \in \mathbb{R}$ and $S_{ij} \in \mathrm{Pr}(b_1, \ldots, b_m)$ has order $\Delta_0(S_{ij}) = l' \le l$. Then according to assumption 1 we have

$$\widetilde{B} = \sum_{ijk} \alpha_{ijk} S_{ijk},$$

where $\alpha_{ijk} \in \mathbb{R}$ and $S_{ijk} \in \mathrm{Pr}(b_1, \ldots, b_m)$ is good of order $\Delta_0(S_{ijk}) = l'' \le l' \le l$. Thus we get

$$B = \sum_{ijk} \alpha_{ijk} B_{S_{ijk}},$$

where $B_{S_{ijk}}$ is good, since $\delta^q(B_{S_{ijk}}) = \Delta^q(S_{ijk})$, $q \in \{1,\ldots,m\}$, and the order of $B_{S_{ijk}}$ is according to (4.14) $\delta_1(B_{S_{ijk}}) = 2l'' - 1 < \delta_1(B)$.

(b). $\Delta^0(S_i) = 0$.

By assumption we have $S_i \in \mathrm{Pr}(b_1,\ldots,b_m)$. Since $\delta^q(B) = \Delta^q(S_i)$, $q \in \{1,\ldots,m\}$, B bad means that S_i is bad but then according to assumption 1 we have that equation 4.15 becomes

$$\widetilde{B} = \sum_{ij} \alpha_{ij} S_{ij},$$

where $\alpha_{ij} \in \mathbb{R}$ and $S_{ij} \in \mathrm{Pr}(b_1,\ldots,b_m)$ is good of order $\Delta_0(S_{ij}) = l' < l$. Thus we get

$$B = \sum_{ij} \alpha_{ij} B_{S_{ij}},$$

where $B_{S_{ij}}$ is good, since $\delta^q(B_{S_{ij}}) = \Delta^q(S_{ij})$, $q \in \{1,\ldots,m\}$, and the order of $B_{S_{ij}}$ is according to (4.14) $\delta_1(B_{S_{ij}}) = 2l' - 1 < \delta_1(B)$.

The result then follows as an application of Theorem 43. □

Denote by $\mathrm{inv}_{\xi_{\mathrm{re}}}\overline{\mathrm{Sym}}(b_1,\ldots,b_m)$ the vector space spanned by the elements from $K(\xi_{\mathrm{re}}, b_1,\ldots,b_m) \subset \mathrm{Pr}(\xi_{\mathrm{re}}, b_1,\ldots,b_m)$ where

$$\begin{aligned} K_0(\xi_{\mathrm{re}}, b_1,\ldots,b_m) &= \mathrm{Pr}(b_1,\ldots,b_m), \\ K_{i+1}(\xi_{\mathrm{re}}, b_1,\ldots,b_m) &= K_i(\xi_{\mathrm{re}}, b_1,\ldots,b_m) \cup \{\langle \xi_{\mathrm{re}} : v\rangle \mid v \in K_i(\xi_{\mathrm{re}}, b_1,\ldots,b_m)\}, \\ K(\xi_{\mathrm{re}}, b_1,\ldots,b_m) &= K_\infty(\xi_{\mathrm{re}}, b_1,\ldots,b_m) \end{aligned}$$

Then we have the following result.

Proposition 57. *Consider the system* (4.3). *Let* ξ_{re} *satisfy* $\langle \xi_{\mathrm{re}} : \xi_{\mathrm{re}}\rangle = 0$. *Assume that the subspace* $\mathrm{inv}_{\xi_{\mathrm{re}}}\overline{\mathrm{Sym}}(b_1,\ldots,b_m)$ *has full rank and that every bad product in* $\mathrm{Pr}(\xi_{\mathrm{re}}, b_1,\ldots,b_m) \setminus \{\xi_{\mathrm{re}}\}$ *is a linear combination of lower order good products from* $K(\xi_{\mathrm{re}}, b_1,\ldots,b_m)$. *Then the system is STLC at* ξ_{re}.

Proof. Equation (4.12) in the proof of Proposition 56 shows that

$$\widetilde{[b_i, [-\tfrac{1}{2}\langle \xi : \xi\rangle, b_j]]} = \langle b_i : b_j\rangle,$$

and

$$\widetilde{[-\tfrac{1}{2}\langle \xi : \xi\rangle, [-\tfrac{1}{2}\langle \xi : \xi\rangle, [\ldots, [-\tfrac{1}{2}\langle \xi : \xi\rangle, b_i]\ldots]]]}(\xi_{\mathrm{re}}) = \langle \xi_{\mathrm{re}} : \langle \xi_{\mathrm{re}} : \langle \ldots : \langle \xi_{\mathrm{re}} : b_i\rangle \ldots\rangle\rangle\rangle$$

where $-\frac{1}{2}\langle \xi : \xi\rangle$ appears on the left hand side the same number of times as ξ_{re} appears on the right hand side. This shows that the LARC is satisfied if $\mathrm{inv}_{\xi_{\mathrm{re}}}\overline{\mathrm{Sym}}(b_1,\ldots,b_m)$ has full rank. We also see from the above that for every $S \in K(\xi_{\mathrm{re}}, b_1,\ldots,b_m)$ there exists a bracket $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi\rangle, b_1,\ldots,b_m)$ such that $\widetilde{B} = S$, we denote this bracket B_S. Since the number of times b_k, $k \in \{1,\ldots,m\}$, occurs in B_S is the same as in S, B_S will be good if S is good.

Take $\theta = 0$ to define the order of a bracket $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi\rangle, b_1,\ldots,b_m)$, this means that the order of a bracket $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi\rangle, b_1,\ldots,b_m)$ is the same as the order of the elements in the sum of products that $\widetilde{B}$ is composed of. Let $B \in \mathrm{Br}(-\frac{1}{2}\langle \xi : \xi \rangle, b_1, \dots, b_m)$ be a bad bracket. Then we have from the calculations in Proposition 56 that

$$\widetilde{B}(\xi_{\mathrm{re}}) = \sum_i \alpha_i S_i$$

where $\alpha_i \in \mathbb{R}$ and $S_i \in \mathrm{Pr}(\xi_{\mathrm{re}}, b_1, \dots, b_m) \setminus \{\xi_{\mathrm{re}}\}$ is bad and of order same order as B. By assumption we have

$$\widetilde{B}(\xi_{\mathrm{re}}) = \sum_{ij} \alpha_{ij} S_{ij},$$

where $\alpha_{ij} \in \mathbb{R}$ and $S_{ij} \in K(\xi_{\mathrm{re}}, b_1, \dots, b_m)$ is good and of lower order than S_i. Thus we have

$$B = \sum_{ij} \alpha_{ij} B_{S_{ij}},$$

where $B_{S_{ij}}$ is good and of lower order than B.

The result thus again follows as an application of Theorem 43, but this time with $\theta = 0$. □

This proposition has the following corollary.

Corollary 58 (Linear controllability)**.** *Consider the system* (4.3)*. Let* ξ_{re} *satisfy* $\langle \xi_{\mathrm{re}} : \xi_{\mathrm{re}} \rangle = 0$*. Assume that the space*

$$\mathrm{span}\{b_1, \dots, b_m, \langle \xi_{\mathrm{re}} : b_1 \rangle, \dots, \langle \xi_{\mathrm{re}} : b_m \rangle, \langle \xi_{\mathrm{re}} : \langle \xi_{\mathrm{re}} : b_1 \rangle\rangle, \dots, \langle \xi_{\mathrm{re}} : \langle \xi_{\mathrm{re}} : b_m \rangle\rangle, \dots\}$$

has full rank. Then the system is STLC at ξ_{re}*.*

From the theory of linear systems, see, e.g., [40], we know that in Corollary 58 we only need to include products where ξ_{re} appears less than or equal to $n-1$ times, where n is the number of degrees of freedom for the system.

4.3 Examples

Example 4 (Planar rigid body)**.** Reconsider the planar rigid body as described in the previous chapter. The configuration manifold is the matrix Lie group $G = SE(2)$ which is isomorphic to $S \times \mathbb{R}^2 \ni (\theta, x, y)$. m denotes the mass of the body, J its moment of inertia, and h the distance from the center of mass to the control forces. The inertia tensor has the representation $\mathbb{I} = \mathrm{diag}(J, m, m)$. With controls as in Figure 3.2 we therefore have

$$b_1 = \frac{1}{m} e_2, \qquad b_2 = -\frac{h}{J} e_1 + \frac{1}{m} e_3,$$

For $(\omega, v_1, v_2)^T \in \mathbb{R}^3 \simeq \mathfrak{se}(2)$ the adjoint operator is given by

$$\mathrm{ad}_{(\omega, v_1, v_2)^T} = \begin{bmatrix} 0 & 0 & 0 \\ v_2 & 0 & -\omega \\ -v_1 & \omega & 0 \end{bmatrix},$$

and the symmetric product is, for $\omega, \lambda \in \mathbb{R}$ and $v, w \in \mathbb{R}^2$, given by

$$\langle (\omega, v) : (\lambda, w) \rangle = \begin{bmatrix} 0 \\ \hat{\omega} w + \hat{\lambda} v \end{bmatrix},$$

where $\hat{\omega} = \left[\begin{smallmatrix} 0 & -\omega \\ \omega & 0 \end{smallmatrix}\right]$. This gives

$$\langle b_1 : b_1 \rangle = 0, \qquad \langle b_2 : b_2 \rangle = \frac{2h}{Jm} e_2, \qquad \langle b_1 : b_2 \rangle = -\frac{h}{Jm} e_3.$$

Since $\langle b_1 : b_1 \rangle, \langle b_2 : b_2 \rangle \in \operatorname{span}\{b_1, b_2\}$ and $\operatorname{span}\{b_1, b_2, \langle b_1 : b_2 \rangle\}$ has full rank we have from Theorem 50 that the system is STLC at zero velocity.

The relative equilibrium e_1 satisfies

$$\langle e_1 : b_1 \rangle = \frac{1}{m} e_3, \qquad \langle e_1 : b_2 \rangle = -\frac{1}{m} e_2, \tag{4.16}$$

meaning that the system does not satisfy the sufficient condition 2 of Proposition 53 which can therefore not be used to determine whether (4.2)-(4.3) for this system is locally controllable for the relative equilibrium e_1. Instead equations (4.16) show the assumption of Corollary 58 is satisfied for $\xi_{\mathrm{re}} = e_1$ and (4.3) is therefore STLC for this equilibrium.

The relative equilibrium $\alpha e_2 + \beta e_3$ satisfies

$$\langle \alpha e_2 + \beta e_3 : b_1 \rangle = 0, \qquad \langle \alpha e_2 + \beta e_3 : b_2 \rangle = \frac{\beta h}{J} e_2 - \frac{\alpha h}{J} e_3,$$

so the condition 2 of Proposition 53 is satisfied if and only if $\alpha = 0$. Since $\langle b_1 : b_1 \rangle, \langle b_2 : b_2 \rangle \in \operatorname{span}\{b_1, b_2\}$ and $\operatorname{span}\{b_1, b_2, \langle b_1 : b_2 \rangle\}$ has full rank we therefore know from Corollary 55 that (4.2)-(4.3) for this system is locally controllable for the relative equilibrium $\xi_{\mathrm{re}} = e_3$.

Example 5 (Satellite with two thrusters)**.** Reconsider the satellite with two thrusters aligned with the first and second principal axes. The configuration manifold is the matrix Lie group $G = SO(3)$ and the equations of motion are given by (4.2)-(4.3). From the previous chapter we know that for $x \in \mathbb{R}^3 \simeq \mathfrak{so}(3)$ we have $\mathrm{ad}_x = \widehat{x}$ and the inertia tensor is given by $\mathbb{I} = \operatorname{diag}(J_1, J_2, J_3)$, where J_i is the moment of inertia around the ith principal axis. Therefore the symmetric product $\langle \xi : \eta \rangle = -\mathbb{I}^{-1}(\mathrm{ad}^*_\xi \mathbb{I}\eta + \mathrm{ad}^*_\eta \mathbb{I}\xi)$, $\xi, \eta \in \mathbb{R}^3 \simeq \mathfrak{so}(3)$, is given by

$$\langle \xi : \eta \rangle = \mathbb{I}^{-1}\big(\xi \times (\mathbb{I}\eta) + \eta \times (\mathbb{I}\xi)\big),$$

where $\times$ is the cross product.

With controls as in figure 3.3 we have

$$b_1 = \frac{1}{J_1} e_1, \qquad b_2 = \frac{1}{J_2} e_2,$$

giving

$$\langle b_1 : b_1 \rangle = 0, \qquad \langle b_2 : b_2 \rangle = 0, \qquad \langle b_1 : b_2 \rangle = \frac{J_2 - J_1}{J_1 J_2 J_3} e_3.$$

Thus $\langle b_1 : b_1\rangle, \langle b_2 : b_2\rangle \in \operatorname{span}\{b_1, b_2\}$ and since $\operatorname{span}\{b_1, b_2, \langle b_1 : b_2\rangle\}$ has full rank we know from Theorem 50 that the system is STLC at zero velocity.

Since

$$\langle e_1 : b_1\rangle = 0, \qquad \langle e_1 : b_2\rangle = \frac{J_2 - J_1}{J_2 J_3} e_3,$$
$$\langle e_2 : b_1\rangle = \frac{J_2 - J_1}{J_1 J_3} e_3, \qquad \langle e_2 : b_2\rangle = 0,$$

Proposition 53 can not be used to determine whether 4.2-4.3 for this system is locally controllable along the relative equilibria $(g_0 \exp(te_1), e_1)$ and $(g_0 \exp(te_2), e_2)$. We see instead that the assumption of Corollary 58 is satisfied so the Euler-Poincare equation (4.3) is STLC for the equilibria e_1 and e_2.

We have

$$\langle e_3 : b_1\rangle = \frac{J_1 - J_3}{J_1 J_2} e_2, \qquad \langle e_3 : b_2\rangle = \frac{J_3 - J_2}{J_1 J_2} e_1,$$

so Corollary 55 gives that equations (4.2)-(4.3) for the system is locally controllable along the relative equilibrium $(g_0 \exp(te_3), e_3)$, $g_0 \in SO(3)$, since $\langle b_1 : b_1\rangle, \langle b_2 : b_2\rangle \in \operatorname{span}\{b_1, b_2\}$, $\langle e_3 : b_1\rangle, \langle e_3 : b_2\rangle \in \operatorname{span}\{b_1, b_2\}$, and $\operatorname{span}\{b_1, b_2, \langle b_1 : b_2\rangle\}$ has full rank.

Example 6 (Underwater vehicle in ideal fluid)**.** We re-examine the underwater vehicle in an ideal fluid as described in the previous chapter. The inertia tensor is given by $\mathbb{I} = \operatorname{diag}(\mathbb{J}, \mathbb{M})$, where $\mathbb{J} = \operatorname{diag}(J_1, J_2, J_2)$ is the inertia matrix for the body and $\mathbb{M} = \operatorname{diag}(m_1, m_2, m_3)$ includes added masses. Since, for $\Omega, v \in \mathbb{R}^3$, we have

$$\operatorname{ad}_{(\Omega,v)} = \begin{bmatrix} \widehat{\Omega} & 0 \\ \widehat{v} & \widehat{\Omega} \end{bmatrix},$$

the symmetric product is given by

$$\langle(\Omega, v) : (\Gamma, w)\rangle = \mathbb{I}^{-1} \begin{bmatrix} \Omega \times (\mathbb{J}\Gamma) + \Gamma \times (\mathbb{J}\Omega) + v \times (\mathbb{M}w) + w \times (\mathbb{M}v) \\ \Omega \times (\mathbb{M}w) + \Gamma \times (\mathbb{M}v) \end{bmatrix}.$$

With forces as in figure 3.4 we have

$$b_1 = \frac{1}{m_1} e_4, \qquad b_2 = -\frac{h}{J_3} e_3 + \frac{1}{m_3} e_5, \qquad b_3 = \frac{h}{J_2} e_2 + \frac{1}{m_3} e_6.$$

Calculating symmetric products gives

$$\langle b_1 : b_1\rangle = 0, \qquad \langle b_1 : b_2\rangle = \frac{m_2 - m_1}{J_3 m_1 m_2} e_3 - \frac{h}{J_3 m_2} e_5,$$
$$\langle b_2 : b_2\rangle = \frac{2h}{J_3 m_1} e_4, \qquad \langle b_1 : b_3\rangle = \frac{m_1 - m_3}{J_2 m_1 m_3} e_2 - \frac{h}{J_2 m_3} e_6,$$
$$\langle b_3 : b_3\rangle = \frac{2h}{J_2 m_1} e_4, \qquad \langle b_2 : b_3\rangle = \frac{1}{J_1} \left(\frac{h^2}{J_3} - \frac{h^2}{J_2} - \frac{1}{m_3} + \frac{1}{m_2} \right) e_1.$$

The space $\mathrm{span}\{b_1, b_2, b_3, \langle b_1 : b_2\rangle, \langle b_1 : b_3\rangle, \langle b_2 : b_3\rangle\}$ has full rank if

$$h^2 m_1 m_2 + J_3(m_1 - m_2) \neq 0,$$
$$h^2 m_1 m_3 + J_2(m_1 - m_3) \neq 0,$$
$$h^2\left(\frac{1}{J_3} - \frac{1}{J_2}\right) - \frac{1}{m_3} + \frac{1}{m_2} \neq 0,$$

and since $\langle b_1 : b_1\rangle, \langle b_2 : b_2\rangle, \langle b_3 : b_3\rangle \in \mathrm{span}\{b_1, b_2, b_3\}$ the system is therefore STLC at zero velocity, according to Theorem 50, if this is satisfied.

Since

$$\langle \alpha_1 e_1 + \beta_1 e_4 : b_1\rangle = 0,$$
$$\langle \alpha_1 e_1 + \beta_1 e_4 : b_2\rangle = \alpha_1\left(\frac{h}{J_2} - \frac{J_1 h}{J_2 J_3}\right) e_2 + \beta_1\left(\frac{1}{J_3} - \frac{m_1}{J_3 m_2}\right) e_3 - \beta_1 \frac{m_1 h}{m_2 J_3} e_5 + \alpha_1 \frac{1}{m_3} e_6,$$
$$\langle \alpha_1 e_1 + \beta_1 e_4 : b_3\rangle = \beta_1\left(\frac{m_1}{J_2 m_3} - \frac{1}{J_2}\right) e_2 + \alpha_1\left(\frac{h}{J_3} - \frac{J_1 h}{J_2 J_3}\right) e_3 - \alpha_1 \frac{1}{m_2} e_5 - \beta_1 \frac{m_1 h}{m_3 J_2} e_6,$$

are not all in $\mathrm{span}\{b_1, b_2, b_3\}$ Proposition 53 can not be used to determine whether the system is locally controllable along $(g, \xi)(t) = (g_0 \exp(t(\alpha_1 e_1 + \beta_1 e_4)), \alpha_1 e_1 + \beta_1 e_4)$. Since the $\langle \alpha_1 e_1 + \beta_1 e_4 : (x, y)\rangle$ has no component in the e_1 direction the assumption of Corollary 58 cannot be satisfied.

Similarly we have that

$$\langle \alpha_2 e_2 + \beta_2 e_5 : b_1\rangle = \beta_2\left(\frac{m_2}{J_3 m_1} - \frac{1}{J_3}\right) e_3 - \alpha_2 \frac{1}{m_3} e_6,$$
$$\langle \alpha_2 e_2 + \beta_2 e_5 : b_2\rangle = \alpha_2\left(\frac{J_2 h}{J_1 J_3} - \frac{h}{J_1}\right) e_1 + \beta_2 \frac{m_2 h}{m_1 J_3} e_4,$$
$$\langle \alpha_2 e_2 + \beta_2 e_5 : b_3\rangle = \beta_2\left(\frac{1}{J_1} - \frac{m_2}{J_1 m_3}\right) e_1 + \alpha_2 \frac{1}{m_1} e_4,$$

are not in $\mathrm{span}\{b_1, b_2, b_3\}$ and therefore Proposition 53 can not be used to determine whether the system is locally controllable along the relative equilibrium $\alpha_2 e_2 + \beta_2 e_5$. Calculating

$$\langle \alpha_2 e_2 + \beta_2 e_5 : \langle \alpha_2 e_2 + \beta_2 e_5 : b_2\rangle\rangle =$$
$$\frac{h(\alpha_2^2(J_1 m_1(J_3 - J_2) + J_2 m_1(J_2 - J_3)) + \beta_2^2 J_1 m_2(m_2 - m_1))}{J_3^2 J_1 m_1} e_3 + \alpha_2\beta_2 \frac{m_2 h(J_2 - J_1 - J_3)}{m_3 J_1 J_3} e_6,$$

gives that

$$\mathrm{span}\{b_1, b_2, b_3, \langle \alpha_2 e_2 + \beta_2 e_5 : b_1\rangle, \langle \alpha_2 e_2 + \beta_2 e_5 : b_2\rangle, \langle \alpha_2 e_2 + \beta_2 e_5 : \langle \alpha_2 e_2 + \beta_2 e_5 : b_2\rangle\rangle\}$$

generically has full rank if $\alpha_2 \neq 0$. Thus the condition for Corollary 58 to be applied is satisfied.

Likewise, for the relative equilibrium $\alpha_3 e_3 + \beta_3 e_6$, we get

$$\langle \alpha_3 e_3 + \beta_3 e_6 : b_1 \rangle = \beta_3 \left(\frac{1}{J_2} - \frac{m_3}{J_2 m_1} \right) e_2 + \alpha_3 \frac{1}{m_2} e_5,$$
$$\langle \alpha_3 e_3 + \beta_3 e_6 : b_2 \rangle = \beta_3 \left(\frac{m_3}{J_1 m_2} - \frac{1}{J_1} \right) e_1 - \alpha_3 \frac{1}{m_1} e_4,$$
$$\langle \alpha_3 e_3 + \beta_3 e_6 : b_3 \rangle = \alpha_3 \left(\frac{J_3 h}{J_1 J_2} - \frac{h}{J_1} \right) e_1 + \beta_3 \frac{m_3 h}{m_1 J_2} e_4,$$

which are not in $\operatorname{span}\{b_1, b_2, b_3\}$ so also for this relative equilibrium the assumptions of Proposition 53 are not satisfied. Since

$$\operatorname{span}\{b_1, b_2, b_3, \langle \alpha_3 e_3 + \beta_3 e_6 : b_1 \rangle, \langle \alpha_3 e_3 + \beta_3 e_6 : b_2 \rangle, \langle \alpha_3 e_3 + \beta_3 e_6 : \langle \alpha_3 e_3 + \beta_3 e_6 : b_2 \rangle \rangle \}$$

generically has full rank for $\beta_3 \neq 0$ this equilibrium point for the Euler-Poincare equations (4.3) is STLC according to Corollary 58.

Chapter 5

Control Algorithms along Relative Equilibria

In this chapter we study control of underactuated mechanical systems on Lie groups. We focus on the construction of an algorithm which, depending on the sign of a parameter, generates small-amplitude control forces to accelerate along, decelerate along, or stabilize a relative equilibrium of a system. Perturbation analysis and Lie group theory play a crucial role in the analysis. The main limitation of the presented theory is that part of the results are applicable only to n-dimensional systems with $(n-1)$ controls. Example systems to which the theory applies are an underactuated planar rigid body and a satellite with two thrusters.

This chapter is organized as follows. First, we review the mathematical model of simple mechanical control systems on Lie groups and perform perturbation analysis for small amplitude forcing and initial velocity close to a relative equilibrium and give a similar result obtained in [11] for small initial velocity. Following we review the theory from [11] regarding motion algorithms for small velocity. Based on the perturbation analysis for the case when the initial velocity is close to a relative equilibrium we construct two "inversion maps" and combine them into a "motion primitive." After an application of the motion primitive, the velocity has changed in the direction of a relative equilibrium, while the configuration has changed as if the velocity was a relative equilibrium throughout the execution of the primitive. Using this motion primitive iteratively we design an algorithm which gives small-amplitude control forces which make the system accelerate along or decelerate along a relative equilibrium or stabilizes the motion along a relative equilibrium. We illustrate the approach by applying the algorithm to an underactuated planar rigid body and the satellite with two thrusters.

5.1 Mathematical Model and Perturbation Analysis

We consider a simple mechanical control system on a matrix Lie group given by $\Sigma = (G, \mathbb{I}, \{f_1, \ldots, f_m\})$. We let id denote the identity element and let $A \mapsto e^A$ denote the matrix exponential of a square matrix A. With $g \in G$ and $\xi \in \mathfrak{g}$, $\mathfrak{g}$ being the Lie algebra corresponding to G, we have from chapter 3 that the equations of motion for this system are

$$\dot{g} = g \cdot \xi, \tag{5.1}$$

$$\dot{\xi} = -\tfrac{1}{2}\langle \xi : \xi \rangle + \sum_{i=1}^{m} b_i u_i(t), \tag{5.2}$$

where $b_i = \mathbb{I}^{-1} f_i$, and the symmetric product $\langle \cdot : \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is given by

$$\langle \eta : \zeta \rangle = -\mathbb{I}^{-1}(\mathrm{ad}^*_\eta \mathbb{I}\zeta + \mathrm{ad}^*_\zeta \mathbb{I}\eta),$$

for $\eta, \zeta \in \mathfrak{g}$.

Recall that a relative equilibrium for Σ is a curve $t \mapsto g_0 \exp(t\xi_{\mathrm{re}}) \in G$, for $g_0 \in G$ and $\xi_{\mathrm{re}} \in \mathfrak{g}$, that is a solution to the dynamics (5.1), (5.2) for zero input u, i.e. ξ_{re} satisfies $\langle \xi_{\mathrm{re}} : \xi_{\mathrm{re}} \rangle = 0$. We call relative equilibrium both the curve $t \mapsto g_0 \exp(t\xi_{\mathrm{re}})$ and the vector ξ_{re}.

Given a relative equilibrium ξ_{re}, we define the linear map $A_{\mathrm{re}} : \mathfrak{g} \to \mathfrak{g}$ by $A_{\mathrm{re}}\eta := -\langle \xi_{\mathrm{re}} : \eta \rangle$, for all $\eta \in \mathfrak{g}$.

Remark 59 (Simplifying convention)**.** It is well known that $\mathfrak{g}$ is an n-dimensional vector space. In what follows, we make no distinction between $\mathfrak{g}$ and $\mathbb{R}^n$. This is done in order to be able to express a vector in $\mathfrak{g}$ as a column vector and thus being able to collect vectors in a matrix and in order to represent a linear map on $\mathfrak{g}$ as a matrix. This choice of notation is not to be confused with the claim that the Lie algebra structure on $\mathfrak{g}$ is insignificant since this is far from being the case. •

We are interested in bounded control signals $u \in C^0([0, 2\pi], \mathbb{R}^m)$ of the form

$$u(t) = \epsilon u^1(t) + \epsilon^2 u^2(t), \qquad 0 < \epsilon \ll 1,$$

where $u^i \in C^0([0, 2\pi], \mathbb{R}^m)$, $i \in \{1, \ldots, 2\}$. Accordingly, we define

$$b^j(t) := \sum_{i=1}^{m} b_i u_i^j(t), \quad j \in \{1, 2\},$$

and equation (5.2) thus becomes

$$\dot{\xi} = -\tfrac{1}{2}\langle \xi : \xi \rangle + \epsilon b^1(t) + \epsilon^2 b^2(t).$$

For $f \in C^0([0, 2\pi], \mathbb{R}^n)$ and $\sigma \in \mathbb{R}$ it will be convenient to make the definition

$$\overline{f}^\sigma(t) := \int_0^t e^{\sigma A_{\mathrm{re}}(t-s)} f(s)\mathrm{d}s, \qquad \overline{f}(t) := \overline{f}^0(t).$$

In what follows, s and τ will be used as integration variables only.

In [11] the following perturbation result is obtained.

Theorem 60 (Perturbation analysis for small velocity)**.** *For $0 < \epsilon \ll 1$ and for inputs of the form $\sum_{i=1}^m b_i u_i(t) = \epsilon b^1(t) + \epsilon^2 b^2(t)$, let $(g(t), \xi(t))$ be the solutions of* (5.1) *and* (5.2)*. Let $x(t)$ be the exponential coordinates of $g(t)$ with initial condition $g(0) = \mathrm{id}$.*

Also, assume that the initial velocity is $\xi(0) = \epsilon\xi_0^1 + \epsilon^2\xi_0^2$*, where* ξ_0^1 *and* ξ_0^2 *are of order* $\mathcal{O}(1)$.

Then for $t \in [0, 2\pi]$ *it holds that* $\xi(t,\epsilon) = \epsilon\xi^1(t) + \epsilon^2\xi^2(t) + \mathcal{O}(\epsilon^3)$*, with*

$$\begin{aligned}\xi^1(t) &= \xi_0^1 + \overline{b^1}(t),\\ \xi^2(t) &= \xi_0^2 - \tfrac{1}{2}\langle\xi_0^1 : \xi_0^1\rangle t - \langle\xi_0^1 : \overline{\overline{b^1}}(t)\rangle + \left(\overline{b^2 - \tfrac{1}{2}\langle\overline{b^1} : \overline{b^1}\rangle}\right)(t),\end{aligned}$$

and $x(t,\epsilon) = \epsilon x^1(t) + \epsilon^2 x^2(t) + \mathcal{O}(\epsilon^3)$*, with*

$$\begin{aligned}x^1(t) &= \xi_0^1 t + \overline{\overline{b^1}}(t),\\ x^2(t) &= \xi_0^2 t - \tfrac{1}{4}\langle\xi_0^1 : \xi_0^1\rangle t^2 + \left(\overline{\overline{b^2 - \tfrac{1}{2}\langle\overline{b^1} : \overline{b^1}\rangle}}\right)(t) - \langle\xi_0^1 : \overline{\overline{\overline{b^1}}}(t)\rangle - \tfrac{1}{2}\overline{[\xi_0^1 + \overline{b^1}, \xi_0^1 t + \overline{\overline{b^1}}]}(t).\end{aligned}$$

Instead of the velocity being small, i.e. of order $\mathcal{O}(\epsilon)$, we let the velocity be of arbitrary size but aligned with a relative equilibrium with a deviance of order $\mathcal{O}(\epsilon^2)$.

Proposition 61 (Perturbation analysis for a relative equilibrium)**.** *Let* Σ *be a simple mechanical control system on a Lie group with a relative equilibrium* ξ_{re} *and corresponding matrix* A_{re}*. For* $0 < \epsilon \ll 1$ *and* $\sigma > 0$*, let* $[0, 2\pi] \ni t \mapsto (g(t), \xi(t))$ *be the solution to* (5.1) *and* (5.2) *with* $t \mapsto \sum_i^m b_i u_i(t) = \epsilon b^1(t) + \epsilon^2 b^2(t)$ *and from initial velocity* $\xi(0) = \sigma\xi_{\mathrm{re}} + \epsilon^2\xi_0^2$*, for* $\xi_0^2 = \mathcal{O}(1)$*, and initial configuration* $g(0) = \mathrm{id}$*. Let* $h(t) := g(t)\cdot\exp(-t\sigma\xi_{\mathrm{re}})$ *and let* $x(t) := \log(h(t))$ *be the exponential coordinates of* h*. Then, for* $t \in [0, 2\pi]$*, it holds that* $\xi(t,\epsilon) = \xi^0(t) + \epsilon\xi^1(t) + \epsilon^2\xi^2(t) + \mathcal{O}(\epsilon^3)$ *with*

$$\begin{aligned}\xi^0(t) &= \sigma\xi_{\mathrm{re}},\\ \xi^1(t) &= \overline{b^1}^{\sigma}(t),\\ \xi^2(t) &= e^{\sigma A_{\mathrm{re}} t}\xi_0^2 - \tfrac{1}{2}\overline{\langle\overline{b^1}^{\sigma} : \overline{b^1}^{\sigma}\rangle}^{\sigma}(t) + \overline{b^2}^{\sigma}(t),\end{aligned}$$

and $x(t,\epsilon) = \epsilon x^1(t) + \epsilon^2 x^2(t) + \mathcal{O}(\epsilon^3)$ *with*

$$\begin{aligned}x^1(t) &= \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(s))}(t),\\ x^2(t) &= \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(e^{\sigma A_{\mathrm{re}} s}\xi_0^2)}(t) - \tfrac{1}{2}\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{\langle\overline{b^1}^{\sigma} : \overline{b^1}^{\sigma}\rangle}^{\sigma}(s))}(t)\\ &\quad + \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^{\sigma}(s))}(t) - \tfrac{1}{2}\overline{[\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(s)), \overline{\mathrm{Ad}_{\exp(\tau\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(\tau))}(s)]}(t).\end{aligned}$$

Proof. Since the input is analytic in ϵ so is the solution $\xi(t) = \sum_{j=0}^{+\infty}\epsilon^j\xi^j(t)$. Inserting the expansions for ξ into equation (5.2) and collecting terms of same order we compute

$$\begin{aligned}\dot{\xi}^0 &= -\tfrac{1}{2}\langle\xi^0 : \xi^0\rangle,\\ \dot{\xi}^1 &= -\langle\xi^0 : \xi^1\rangle + b^1(t),\\ \dot{\xi}^2 &= -\langle\xi^0 : \xi^2\rangle - \tfrac{1}{2}\langle\xi^1 : \xi^1\rangle + b^2(t).\end{aligned}$$

Inserting the initial condition then gives

$$\begin{aligned}\xi^0(t) &= \sigma\xi_{\mathrm{re}},\\ \xi^1(t) &= \overline{b^1}^{\sigma}(t),\\ \xi^2(t) &= e^{\sigma A_{\mathrm{re}} t}\xi_0^2 - \tfrac{1}{2}\overline{\langle\xi^1 : \xi^1\rangle}^{\sigma}(t) + \overline{b^2}^{\sigma}(t)\\ &= e^{\sigma A_{\mathrm{re}} t}\xi_0^2 - \tfrac{1}{2}\overline{\langle\overline{b^1}^{\sigma} : \overline{b^1}^{\sigma}\rangle}^{\sigma}(t) + \overline{b^2}^{\sigma}(t).\end{aligned}$$

Since g is a solution to the kinematic equation (5.1), it follows that

$$\begin{aligned}\dot{h} &= \dot{g}\cdot\exp(-t\sigma\xi_{\mathrm{re}}) - g\cdot\exp(-t\sigma\xi_{\mathrm{re}})\cdot\sigma\xi_{\mathrm{re}}\\ &= g\cdot\xi\cdot\exp(-t\sigma\xi_{\mathrm{re}}) - h\cdot\sigma\xi_{\mathrm{re}}\\ &= h\cdot(\exp(t\sigma\xi_{\mathrm{re}})\cdot\xi\cdot\exp(-t\sigma\xi_{\mathrm{re}}) - \sigma\xi_{\mathrm{re}})\\ &= h\cdot(\mathrm{Ad}_{\exp(t\sigma\xi_{\mathrm{re}})}(\xi) - \sigma\xi_{\mathrm{re}})\\ &= h\cdot\left(\mathrm{Ad}_{\exp(t\sigma\xi_{\mathrm{re}})}(\sigma\xi_{\mathrm{re}} + \epsilon\xi^1 + \epsilon^2\xi^2 + \mathcal{O}(\epsilon^3)) - \sigma\xi_{\mathrm{re}}\right)\\ &= h\cdot\mathrm{Ad}_{\exp(t\sigma\xi_{\mathrm{re}})}(\epsilon\xi^1 + \epsilon^2\xi^2 + \mathcal{O}(\epsilon^3)).\end{aligned}$$

If we define $\zeta(t) := \mathrm{Ad}_{\exp(t\sigma\xi_{\mathrm{re}})}(\epsilon\xi^1 + \epsilon^2\xi^2 + \mathcal{O}(\epsilon^3))$, then we have, according to Proposition 31, that

$$x(t) = \overline{\zeta}(t) - \tfrac{1}{2}\overline{[\zeta,\overline{\zeta}]}(t) + \mathcal{O}(\epsilon^3). \tag{5.3}$$

Using $x = \epsilon x^1 + \epsilon^2 x^2 + \mathcal{O}(\epsilon^3)$ we achieve the result on x^1 and x^2 by inserting the expression for ζ into equation (5.3). □

When comparing Proposition 61 with Proposition 60 we see that when $\sigma = 0$ Proposition 61 simplifies to Proposition 60, with $\xi_0^1 = 0$, as expected. In both propositions we see that $\xi^2(t)$ not is restricted to move in $\mathrm{span}\{b_1,\ldots,b_m\}$ but new directions of motion are possible in particular due to the symmetric product term $\overline{\langle\overline{b^1}^\sigma : \overline{b^1}^\sigma\rangle}^\sigma(t)$. It is precisely this term we will utilize for generation of motion in the directions not lying in $\mathrm{span}\{b_1,\ldots,b_m\}$.

5.2 Small Velocity Motion Algorithms

In this section we recapitulate the motion planning algorithms obtained in [11]. These results are included in order to give a full perspective of what motion algorithms, giving small-amplitude control forces, are available for simple mechanical control systems on Lie groups. In these algorithms the velocity is small, that is, of order $\mathcal{O}(\epsilon)$ where $0 < \epsilon \ll 1$ and the results are therefore built on Proposition 60.

As suggested by Theorem 50 the following assumption is needed.

Assumption 1. The subspace $\mathrm{span}\{b_i, \langle b_i : b_j\rangle \mid i,j \in \{1,\ldots,m\}\}$ is full rank and $\langle b_i : b_i\rangle \in \mathrm{span}\{b_1,\ldots,b_m\}$, for $i \in \{1,\ldots,m\}$.

Using this assumption the following theorem, having character of a lemma for the proposed motion primitives, is proved.

Theorem 62. *Let Assumption 1 hold and let $\eta \in \mathfrak{g}$ be arbitrary. Define the inputs $(b^1(t), b^2(t))$ as follows:*

1. *Set $N = \frac{1}{2}m(m-1)$ and let P denote the ordered set of pairs $\{(j,k)|\ 1 \le j < k \le m\}$. Identify the elements in P with the set $\{1,\ldots,N\}$, and let $a(j,k)$ be the integer associated with the pair (j,k). For $\alpha \in \{1,\ldots,N\}$ define the scalar functions*

$$\psi_\alpha(t) = \frac{1}{\sqrt{2\pi}}(\alpha\sin(\alpha t) - (\alpha+N)\sin((\alpha+N)t)).$$

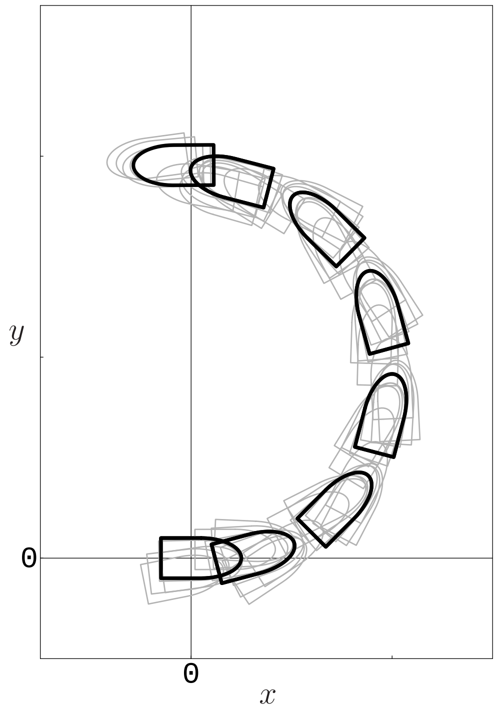

Figure 5.1: The constant velocity algorithm applied to the planar rigid body. The bullet shaped objects illustrates the planar rigid body, the darker ones correspond to the body at the beginning and the end of a primitive. Figure taken from [11] with permission.

2. *By means of the pseudoinverse compute* $(m+N)$ *real numbers* z_i *and* z_{jk} *such that*

$$\eta = \sum_{i=1}^{m} z_i b_i + \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} z_{jk} \langle b_j : b_k \rangle.$$

3. *Finally, set*

$$b^1(t) = \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \sqrt{|z_{jk}|}(b_j - \text{sign}(z_{jk}) b_k) \psi_{a(j,k)}(t),$$

$$b^2(t) = \frac{1}{2\pi} \sum_{i=1}^{m} z_i b_i + \frac{1}{4\pi} \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} |z_{jk}|(\langle b_j : b_j \rangle + \langle b_k : b_k \rangle).$$

Then $b^1(t)$ *and* $b^2(t)$ *satisfy*

$$\left(\overline{b^2 - \tfrac{1}{2} \langle \overline{b^1} : \overline{b^1} \rangle} \right) (2\pi) = \eta.$$

We will call this map $(b^1(t), b^2(t)) = \texttt{Inverse}(\eta)$.

Using this theorem, the orthogonality properties of $\psi_\alpha(t)$, and Theorem 60 the motion primitives `Maintain-Velocity` and `Change-Velocity` can be constructed and analysed according to the following.

Proposition 63 (`Maintain-Velocity` motion primitive)**.** *Consider the system 5.1 and 5.2 with inputs of the form* $\sum_{i=1}^{m} b_i u_i(t) = \epsilon b^1(t) + \epsilon^2 b^2(t)$*, where* $0 < \epsilon \ll 1$*. Let Assumption 1 be satisfied. Let* $\epsilon = \sigma$*,* $0 < \sigma \ll 1$*, and assume that*

$$\begin{aligned} g(0) &= g_0, \\ \xi(0) &= \sigma\xi_{\text{ref}} + \sigma^2\xi_{\text{error}}, \end{aligned}$$

for some $g_0 \in G$ *and* $\xi_{\text{ref}}, \xi_{\text{error}} \in \mathfrak{g}$*. If we for* $t \in [0, 2\pi]$ *take*

$$(b^1(t), b^2(t)) = \texttt{Inverse}(\pi\langle\xi_{\text{ref}} : \xi_{\text{ref}}\rangle - \xi_{\text{error}}),$$

then we obtain

$$\begin{aligned} \log(g_0^{-1}g(2\pi)) &= 2\pi\sigma\xi_{\text{ref}} + \pi\sigma^2\xi_{\text{error}} + \mathcal{O}(\sigma^3), \\ \xi(2\pi) &= \sigma\xi_{\text{ref}} + \mathcal{O}(\sigma^3). \end{aligned}$$

We denote this motion primitive `Maintain-Velocity`$(\sigma, \xi_{\text{ref}})$*.*

Proposition 64 (`Change-Velocity` motion primitive)**.** *Consider the system 5.1 and 5.2 with inputs of the form* $\sum_{i=1}^{m} b_i u_i(t) = \epsilon b^1(t) + \epsilon^2 b^2(t)$*, where* $0 < \epsilon \ll 1$*. Let Assumption 1 be satisfied. Let* $\epsilon = \sqrt{\sigma}$*,* $0 < \sigma \ll 1$*, and assume that*

$$\begin{aligned} g(0) &= g_0, \\ \xi(0) &= \sigma\xi_0, \end{aligned}$$

for some $g_0 \in G$ *and* $\xi_0 \in \mathfrak{g}$*. If we for* $t \in [0, 2\pi]$ *take*

$$(b^1(t), b^2(t)) = \texttt{Inverse}(\xi_{\text{final}} - \xi_0),$$

for some $\xi_{\text{final}} \in \mathfrak{g}$*, then we obtain*

$$\begin{aligned} \log(g_0^{-1}g(2\pi)) &= \pi\sigma(\xi_0 + \xi_{\text{final}}) + \mathcal{O}(\sigma^{3/2}), \\ \xi(2\pi) &= \sigma\xi_{\text{final}} + \mathcal{O}(\sigma^2). \end{aligned}$$

We denote this motion primitive `Change-Velocity`$(\sigma, \xi_{\text{final}})$*.*

These two motion primitives are the basic ingredients in the following algorithm which, by keeping a constant velocity, steers the system from one configuration with low velocity (of order $\mathcal{O}(\sigma^2)$) to another configuration with low velocity (of order $\mathcal{O}(\sigma^2)$).

In the algorithm the function $\lfloor\cdot\rfloor : \mathbb{R} \to \mathbb{Z}$, defined by $\lfloor x \rfloor = \max\{n \in \mathbb{Z} | \ n \leq x\}$, is used.

Proposition 65 (Constant velocity algorithm)**.** *Consider the system 5.1 and 5.2 with inputs of the form* $\sum_{i=1}^{m} b_i u_i(t) = \epsilon b^1(t) + \epsilon^2 b^2(t)$*, where* $0 < \epsilon \ll 1$*. Let Assumption 1 be satisfied. Assume* $g_0, g_1 \in G$ *satisfies that* $\log(g_0^{-1}g_1)$ *is well defined. For* $0 < \sigma \ll 1$ *and* $(g, \xi)(0) = (g_0, \mathcal{O}(\sigma^2))$ *define the algorithm*

1: $N = \lfloor \| \log(g_0^{-1}g_1)\| / (2\pi\sigma) \rfloor$
2: $\xi_{\text{nom}} = \log(g_0^{-1}g_1)/(2\pi\sigma N)$
3: `Change-Velocity`$(\sigma, \xi_{\text{nom}})$
4: ***for*** $k \in \{1, \ldots, N-1\}$ ***do***

```
    Maintain-Velocity(σ, ξ_nom)
  end for
  Change-Velocity(σ, 0).
```

Then the final configuration g_{final} *and the final velocity* ξ_{final} *after an execution of the algorithm satisfies*

$$\log(g_{\text{final}}^{-1} g_1) = \mathcal{O}(\sigma^{3/2}),$$
$$\xi_{\text{final}} = \mathcal{O}(\sigma^2).$$

We notice that the velocity throughout the duration of this algorithm is at most of order $\mathcal{O}(\sigma)$. Therefore the time it takes for this algorithm to reconfigure a body is inherently of order $\mathcal{O}(\frac{1}{\sigma})$.

Since, after applying the constant velocity algorithm, the final velocity may be nonzero the system will drift if not stabilized. The following algorithm is able to stabilize the system exponentially if it is close enough to equilibrium.

Proposition 66 (Local exponential stabilization algorithm)**.** *Consider the system 5.1 and 5.2 with inputs of the form* $\sum_{i=1}^{m} b_i u_i(t) = \epsilon b^1(t) + \epsilon^2 b^2(t)$*, where* $0 < \epsilon \ll 1$*. Let Assumption 1 be satisfied. For* $0 < \sigma \ll 1$ *assume that* $g(0) = g_0 \in G$ *and* $\xi(0) = \xi_0 \in \mathfrak{g}$ *satisfies* $\|(\log(g_0), \xi_0)\| \leq \sigma$*. Let* $N \in \mathbb{N}$*. Define the algorithm*

```
  for k ∈ {0, 1, ..., N} do
    t_k = 4kπ
    σ_k = ‖(log(g(t_k)), ξ(t_k))‖
    Change-Velocity(σ_k, −(log(g(t_k)) + πξ(t_k))/(2πσ_k))
    Change-Velocity(σ_k, 0)
  end for
```

Then there exists a $\lambda > 0$*, independent of* N*, such that the final configuration* g_{final} *and the final velocity* ξ_{final} *after an execution of the algorithm satisfies*

$$\|(\log(g_{\text{final}}), \xi_{\text{final}})\| \leq \|(\log(g_0), \xi_0)\| e^{-\lambda N}.$$

In [11] an additional motion algorithm is constructed, called the "static interpolation algorithm", which steers the systems configuration through a sequence of points. This algorithm is, with minor modifications, a repeated application of the constant velocity algorithm between the points in the given sequence.

5.3 A Motion Algorithm along a Relative Equilibrium

For a simple mechanical control system $\Sigma = (G, \mathbb{I}, \{f_1, \ldots, f_m\})$ with relative equilibrium ξ_{re} and corresponding matrix A_{re}, we present the following assumptions. First, we make the standing assumption that $\xi_{\text{re}} \notin \text{span}\{b_1, \ldots, b_m\}$, otherwise the theory of kinematic reductions [13] is readily applicable and the control problems we consider below are trivial.

Assumption 2 (Lack of linear controllability)**.** The subspace $\text{span}\{b_1, \ldots, b_m\}$ is invariant under the linear map A_{re}, that is, $\langle \xi_{\text{re}} : b_i \rangle \in \text{span}\{b_1, \ldots, b_m\}$, for $i \in \{1, \ldots, m\}$.

Assumption 3. $\langle \xi_{\text{re}} : \langle b_j : b_k \rangle \rangle \in \text{span}\{b_1, \ldots, b_m\}$, for $j, k \in \{1, \ldots, m\}$ and $j \neq k$.

Assumption 4. The subspace $\operatorname{span}\{b_1, \dots, b_m, \xi_{\rm re}\}$ is invariant under the linear map $\operatorname{ad}_{\xi_{\rm re}}$.

Assumption 5. The subspace $\operatorname{span}\{b_1, \dots, b_m\}$ is invariant under the linear map $\operatorname{ad}_{\xi_{\rm re}}$.

If we define the matrix $B := \begin{bmatrix} b_1 & \cdots & b_m \end{bmatrix} \in \mathbb{R}^{n\times m}$, then Assumption 2 is equivalent to the existence of a matrix $Q \in \mathbb{R}^{m\times m}$ such that $A_{\rm re}B = BQ$, and in turn $e^{A_{\rm re}}B = Be^{Q}$. Similarly Assumption 4 is equivalent to the existence of a matrix $W \in \mathbb{R}^{(m+1)\times(m+1)}$ such that $\operatorname{ad}_{\xi_{\rm re}} \begin{bmatrix} B & \xi_{\rm re} \end{bmatrix} = \begin{bmatrix} B & \xi_{\rm re} \end{bmatrix} W$. For Assumption 5 this reduces to the existence of a matrix $M \in \mathbb{R}^{m\times m}$ such that $\operatorname{ad}_{\xi_{\rm re}} B = BM$.

Given $Q \in \mathbb{R}^{m\times m}$, define $F_Q : C^0([0, 2\pi], \mathbb{R}^m) \to \{f \in C^1([0, 2\pi], \mathbb{R}^m) \mid f(0) = 0\}$ by

$$F_Q[u](t) := \int_0^t e^{Q(t-s)}u(s)\mathrm{d}s.$$

Lemma 67 (Transformation of controls)**.** *The map F_Q is invertible and its inverse is given as follows: if $w = F_Q[u]$, then $u(t) = -Qw(t) + \dot{w}(t)$. Additionally, as in Assumption 2, let $A_{\rm re}$, B and Q satisfy $A_{\rm re}B = BQ$. If $u \in C^0([0, 2\pi], \mathbb{R}^m)$ and $w = F_{\sigma Q}[u]$, $\sigma \in \mathbb{R}$, then*

$$\overline{Bu}^{\sigma}(t) = Bw(t).$$

Proof. One-to-one correspondence between u and w is readily checked. We compute $\overline{Bu}^{\sigma}(t) = \int_0^t e^{\sigma A_{\rm re}(t-s)}Bu(s)\mathrm{d}s = B\int_0^t e^{\sigma Q(t-s)}u(s)\mathrm{d}s = Bw(t)$. □

From this lemma we see that Assumption 2 ensures that there is a one-to-one correspondence between $\overline{Bu}^{\sigma}(t)$ and $Bw(t)$. Without this assumption $\overline{Bu}^{\sigma}(t)$ would also contain components not lying in $\operatorname{span}\{b_1, \dots, b_m\}$ which we would not be able to steer for all $t \in [0, 2\pi]$.

The function $\lceil\cdot\rceil : \mathbb{R} \to \mathbb{Z}$ defined by $\lceil x\rceil = \min\{n \in \mathbb{Z} | \ n \geq x\}$ is needed in the following.

Definition 68 (Convenient forcing frequencies)**.** *Take $r = \lceil \frac{n}{m} \rceil$. For $(i, h) \in \{1, \dots, m\}\times\{1, \dots, r\}$, select numbers α_{ih} in the set $\{0, \dots, rm+\frac{1}{2}m(m-1)\}$ as follows:*

1: $\mathcal{V} := \emptyset$; $\mathcal{I} := \{1, \dots, rm + \frac{1}{2}m(m-1)\}$
2: ***for*** $h \in \{1, \dots, r\}$ *and* ***for*** $i \in \{1, \dots, m\}$ ***do***
3: $\quad \omega := \min(\mathcal{I})$; $v := \int_0^{2\pi} \operatorname{Ad}_{\exp(s\sigma\xi_{\rm re})}(b_i \sin(\omega s))\mathrm{d}s$
4: $\quad$***if*** $v \in \operatorname{span}(\mathcal{V})$ ***then*** $\alpha_{ih} := 0$ ***else*** $\alpha_{ih} := \omega$; $\mathcal{I} := \mathcal{I} \setminus \{\omega\}$; $\mathcal{V} := \mathcal{V}\cup\{v\}$ ***end if***
5: ***end for***

Define the $n \times rm$ matrix

$$\mathcal{A}_{\sigma,\alpha} := \int_0^{2\pi} \operatorname{Ad}_{\exp(s\sigma\xi_{\rm re})}(B\begin{bmatrix}\operatorname{diag}(\sin(\alpha_{11}s), \dots, \sin(\alpha_{m1}s)) & \cdots & \operatorname{diag}(\sin(\alpha_{1r}s), \dots, \sin(\alpha_{mr}s))\end{bmatrix})\mathrm{d}s.$$

Next, for $(i, j) \in \{1, \dots, m\}^2$, select numbers β_{ij} as follows: for $i < j$ take $\beta_{ij} \in \{1, \dots, rm + \frac{1}{2}m(m-1)\} \setminus \{\alpha_{kh}\}_{(k,h)\in\{1,\dots,m\}\times\{1,\dots,r\}}$ all having distinct values, for $i > j$ take $\beta_{ij} = \beta_{ji}$, and for $i = j$ take $\beta_{ij} = 0$.

Remark 69. In other words, the numbers α_{ij} are selected sequentially in such a way as to maximize the rank of $\mathcal{A}_{\sigma,\alpha}$. Note that, for $i, j, k, l \in \{1, \ldots, m\}$ and $h \in \{1, \ldots, r\}$, we have: (i) all nonzero α_{ih} are distinct, (ii) all nonzero α_{ih} are distinct from all nonzero β_{jk}, and (iii) $\beta_{ij} = \beta_{kl}$ if and only if $(i, j) = (k, l)$ or $(i, j) = (l, k)$. •

Remark 70. If Assumption 5 is satisfied we know that $\text{Image}(\mathcal{A}_{\sigma,\alpha}) \subset \text{span}\{b_1, \ldots, b_m\}$. This means that we only need $r = 1$, and $\mathcal{A}_{\sigma,\alpha}$ reduces to

$$\begin{aligned}\mathcal{A}_{\sigma,\alpha} &= \int_0^{2\pi} \text{Ad}_{\exp(s\sigma\xi_{\text{re}})}(B\text{diag}(\sin(\alpha_1 s), \ldots, \sin(\alpha_m s))\text{d}s \\ &= B\int_0^{2\pi} e^{s\sigma M}\text{diag}(\sin(\alpha_1 s), \ldots, \sin(\alpha_m s))\text{d}s,\end{aligned}$$

where $M \in \mathbb{R}^{m\times m}$ is the matrix satisfying $\text{ad}_{\xi_{\text{re}}} B = BM$. •

Remark 71. The computations required by Definition 68 include checking that a vector belongs to a subspace. In practical numerical implementations it is sufficient to verify this condition up to a specified tolerance. It is convenient to choose this tolerance comparable with the accuracy of the control algorithms. •

For $Z \in \mathbb{R}^{m\times m}$ define $\lambda : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$ by

$$\lambda_{jk}(Z) := \begin{cases} \text{sign}(Z_{jk})\sqrt{|Z_{jk}|} & , \quad j < k, \\ 0 & , \quad j = k, \\ \frac{1}{\pi}\sqrt{|Z_{kj}|} & , \quad j > k. \end{cases}$$

We are now able to obtain the following result.

Proposition 72 (`speed_inversion`). *Let Σ be a simple mechanical control system on a Lie group with a relative equilibrium ξ_{re} and corresponding matrix A_{re} and satisfying Assumptions 1, 2 and 3. Let $Q \in \mathbb{R}^{m\times m}$ satisfy $A_{\text{re}}B = BQ$. Let $\eta \in \mathbb{R}^n$, $\sigma \in \mathbb{R}$, and compute $z \in \mathbb{R}^m$ and $Z \in \mathbb{R}^{m\times m}$ as the pseudoinverse solution to*

$$\eta = \sum_{i=1}^{m} z_i b_i - \sum_{j=1}^{m-1}\sum_{k=j+1}^{m} Z_{jk}\langle b_j : b_k\rangle, \qquad Z_{jk} = 0 \ \text{for } j \geq k.$$

Given r, α, $\mathcal{A}_{\sigma,\alpha}$, and β as in Definition 68, let

$$y_j(t) := \sum_{k=1}^{m} \lambda_{jk}(Z)\sin(\beta_{jk}t), \quad j \in \{1, \ldots, m\},$$

and let $\gamma = (\gamma_{11}, \ldots, \gamma_{m1}, \ldots, \gamma_{1r}, \ldots, \gamma_{mr})^T$ be the unique solution to

$$\begin{aligned}\mathcal{A}_{\sigma,\alpha}\gamma &= -\overline{\text{Ad}_{\exp(s\sigma\xi_{\text{re}})}(By(s))}(2\pi), \\ \gamma_{ih} &= 0 \ \text{ if } \ \alpha_{ih} = 0 \ \text{ for } \ (i, h) \in \{1, \ldots, m\} \times \{1, \ldots, r\}.\end{aligned} \tag{5.4}$$

Additionally, if we take

$$\begin{aligned} w_j^1(t) &= y_j(t) + \sum_{l=1}^{r} \gamma_{jl}\sin(\alpha_{jl}t), \quad j \in \{1, \ldots, m\}, \\ u^1(t) &= F_{\sigma Q}^{-1}[w^1](t), \\ u^2(t) &= \tfrac{1}{2\pi}e^{\sigma Q(t-2\pi)}(\chi + z),\end{aligned}$$

where $\chi \in \mathbb{R}^m$ *is the unique solution to*

$$B\chi = \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \int_0^{2\pi} (e^{\sigma A_{\text{re}}(2\pi - s)} - I) w_j^1(s) w_k^1(s) \mathrm{d}s \; \langle b_j : b_k \rangle$$
$$+ \tfrac{1}{2} \sum_{i=1}^{m} \int_0^{2\pi} e^{\sigma A_{\text{re}}(2\pi - s)} (w_i^1(s))^2 \mathrm{d}s \; \langle b_i : b_i \rangle, \quad (5.5)$$

then $b^1(t) = Bu^1(t)$ *and* $b^2(t) = Bu^2(t)$ *satisfy*

$$-\tfrac{1}{2} \overline{\langle \overline{b^1}^{\sigma} : \overline{b^1}^{\sigma} \rangle}^{\sigma} (2\pi) + \overline{b^2}^{\sigma}(2\pi) = \eta, \quad (5.6)$$
$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\text{re}})}(\overline{b^1}^{\sigma}(s))}(2\pi) = 0. \quad (5.7)$$

We call this map `speed_inversion`$(\sigma, \eta) = (b^1(t), b^2(t))$.

Proof. Existence and uniqueness of the solution to (5.5) is a consequence of Assumptions 3 and 1. Regarding existence and uniqueness of the solution to (5.4), Definition 68 ensures that

$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\text{re}})}(By(s))}(2\pi) \in \mathrm{Image}(\mathcal{A}_{\sigma,\alpha}).$$

Since every nonzero column in $\mathcal{A}_{\sigma,\alpha}$ contributes to the rank of $\mathcal{A}_{\sigma,\alpha}$, the entries of γ corresponding to these will be unique. The remaining γ-values are defined to be 0.

Regarding the proof of equation (5.7), direct calculations show that

$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\text{re}})}(\overline{b^1}^{\sigma}(s))}(2\pi) = \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\text{re}})}(Bw^1(s))}(2\pi) = \mathcal{A}_{\sigma,\alpha}\gamma + \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\text{re}})}(By(s))}(2\pi) = 0.$$

Regarding the proof of equation (5.6), from Lemma 67 we compute

$$\langle \overline{b}^{\sigma} : \overline{b}^{\sigma} \rangle (t) = \langle \sum_{j=1}^{m} w_j^1(t) b_j : \sum_{k=1}^{m} w_k^1(t) b_k \rangle$$
$$= 2 \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} w_j^1(t) w_k^1(t) \langle b_j : b_k \rangle + \sum_{i=1}^{m} (w_i^1(t))^2 \langle b_i : b_i \rangle.$$

Since all nonzero α-values are distinct and are distinct from the β-values we have for $j < k$

$$\int_0^{2\pi} w_j^1(t) w_k^1(t) \mathrm{d}t = \sum_{l,q=1}^{m} \lambda_{jl}(Z) \lambda_{kq}(Z) \int_0^{2\pi} \sin(\beta_{jl} t) \sin(\beta_{kq} t) \mathrm{d}t$$
$$= \sum_{l,q=1}^{m} \lambda_{jl}(Z) \lambda_{kq}(Z) \delta_{\beta_{kq}}^{\beta_{jl}} \pi = \lambda_{jk}(Z) \lambda_{kj}(Z) \pi = Z_{jk}.$$

By straightforward calculations we then obtain

$$
\begin{aligned}
&-\tfrac{1}{2}\overline{\langle \overline{b^1}^\sigma : \overline{b^1}^\sigma \rangle}^\sigma(2\pi) + \overline{b^2}^\sigma(2\pi) && (5.8)\\
&\quad = -\tfrac{1}{2}\int_0^{2\pi} e^{\sigma A_{\mathrm{re}}(2\pi - s)}\langle \overline{b^1}^\sigma : \overline{b^1}^\sigma\rangle(s)\mathrm{d}s + B\int_0^{2\pi} e^{\sigma Q(2\pi-s)}u^2(s)\mathrm{d}s\\
&\quad = -\sum_{j=1}^{m-1}\sum_{k=j+1}^{m}\int_0^{2\pi} w_j^1(s)w_k^1(s)\mathrm{d}s\langle b_j : b_k\rangle && (5.9)\\
&\quad\quad -\sum_{j=1}^{m-1}\sum_{k=j+1}^{m}\int_0^{2\pi}(e^{\sigma A_{\mathrm{re}}(2\pi-s)} - I)w_j^1(s)w_k^1(s)\mathrm{d}s\langle b_j : b_k\rangle && (5.10)\\
&\quad\quad -\tfrac{1}{2}\sum_{j=1}^{m}\int_0^{2\pi} e^{\sigma A_{\mathrm{re}}(2\pi-s)}(w_j^1(s))^2\mathrm{d}s\langle b_j : b_j\rangle && (5.11)\\
&\quad\quad +\sum_{i=1}^{m}(\chi_i + z_i)b_i && (5.12)\\
&\quad = -\sum_{j=1}^{m-1}\sum_{k=j+1}^{m} Z_{jk}\langle b_j : b_k\rangle + \sum_{i=1}^{m} z_i b_i = \eta.
\end{aligned}
$$

□

If we look at the proof of this proposition the roles of Assumptions 1, 2, and 3 become clear.

If we consider equation (5.8) for $\sigma = \mathcal{O}(\epsilon)$, and disregard terms of order $\mathcal{O}(\epsilon)$, it reduces to

$$
\begin{aligned}
-\tfrac{1}{2}\overline{\langle \overline{b^1}^\sigma : \overline{b^1}^\sigma \rangle}^\sigma(2\pi) + \overline{b^2}^\sigma(2\pi) = &-\sum_{j=1}^{m-1}\sum_{k=j+1}^{m}\int_0^{2\pi} w_j^1(s)w_k^1(s)\mathrm{d}s\langle b_j : b_k\rangle\\
&-\tfrac{1}{2}\sum_{j=1}^{m}\int_0^{2\pi}(w_j^1(s))^2\mathrm{d}s\langle b_j : b_j\rangle + \overline{Bu^2}(2\pi).
\end{aligned}
$$

Thus Assumption 1 ensures, for $\sigma = \mathcal{O}(\epsilon)$, first of all that all the necessary directions to span the full space are available and second of all that $\int_0^{2\pi}(w_j^1(s))^2\mathrm{d}s\langle b_j : b_j\rangle$, which can only move in the positive direction of $\langle b_j : b_j\rangle$ and therefore complicates controllability, belongs to the linearly controllable subspace which $b^2(t) = Bu^2(t)$ can compensate for.

Assumption 3 then means that the term (5.10) will lie in $\mathrm{span}\{b_1, \ldots, b_m\}$ and thus controlling the term (5.9) will mean controlling what is not in $\mathrm{span}\{b_1, \ldots, b_m\}$. The control u^2 can then, via term (5.12), compensate for the terms (5.10)-(5.11).

Assumption 2 means according to Lemma 67 that we can design w^1 (with $w^1(0) = 0$) and then calculate the corresponding u^1. This simplifies considerably the control over the term (5.9).

Remark 73. In the proof of Proposition 72 we see that the reason we need Definition 68 is to ensure that

$$\mathcal{A}_{\sigma,\alpha}\gamma = -\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(By(s))}(2\pi),$$

has a solution and that all the frequencies are different. If a solution can be found by other means than using this definition, keeping the condition that the frequencies are different, we can thus disregard it. This will in fact be the case in the examples we consider. •

We now construct a different inversion map.

Proposition 74 (`configuration_inversion`)**.** *Let Σ be a simple mechanical control system on a Lie group with a relative equilibrium ξ_{re} and corresponding matrix A_{re} and satisfying Assumptions 2 and 4. Let $Q \in \mathbb{R}^{m\times m}$ satisfy $A_{\mathrm{re}}B = BQ$ and $W \in \mathbb{R}^{(m+1)\times(m+1)}$ satisfy $\mathrm{ad}_{\xi_{\mathrm{re}}}\begin{bmatrix} B & \xi_{\mathrm{re}}\end{bmatrix} = \begin{bmatrix} B & \xi_{\mathrm{re}}\end{bmatrix} W$. If $\mu \in \mathbb{R}^m$, $\sigma \in \mathbb{R}$ and*

$$u^1(t) = 0,$$
$$u^2(t) = F_{\sigma Q}^{-1}[w^2](t), \qquad w^2(t) = \tfrac{1}{\pi}\begin{bmatrix} I_m & 0_{m\times 1}\end{bmatrix} e^{-\sigma W t}\begin{bmatrix}\mu \\ 0\end{bmatrix}\sin^2(t),$$

then $b^1(t) = Bu^1(t)$ and $b^2(t) = Bu^2(t)$ satisfy

$$-\tfrac{1}{2}\overline{\langle \overline{b^1}^\sigma : \overline{b^1}^\sigma \rangle}^\sigma(2\pi) + \overline{b^2}^\sigma(2\pi) = 0,$$
$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^\sigma(s))}(2\pi) = B\mu + \delta\xi_{\mathrm{re}},$$

for some $\delta \in \mathbb{R}$. We denote this map `configuration_inversion`$(\sigma,\mu) = (b^1(t), b^2(t)) = (0, b^2(t))$.

Proof. For $b^1(t) = 0$ we have, using Lemma 67, that

$$-\tfrac{1}{2}\overline{\langle \overline{b^1}^\sigma : \overline{b^1}^\sigma \rangle}^\sigma(2\pi) + \overline{b^2}^\sigma(2\pi) = \overline{b^2}^\sigma(2\pi) = Bw^2(2\pi) = 0.$$

Since $\mathrm{ad}_{\xi_{\mathrm{re}}}\xi_{\mathrm{re}} = 0$ we have that

$$W = \begin{bmatrix} \widetilde{W} & 0_{m\times 1} \\ *_{1\times m} & 0\end{bmatrix}, \quad \widetilde{W} \in \mathbb{R}^{m\times m},$$

which in turn gives

$$e^{Ws} = \begin{bmatrix} e^{\widetilde{W}s} & 0_{m\times 1} \\ *_{1\times m} & 1\end{bmatrix}.$$

Thus we have

$$e^{Ws}\begin{bmatrix} I_m & 0_{m\times 1} \\ 0_{1\times m} & 0\end{bmatrix} e^{-Ws} = \begin{bmatrix} e^{\widetilde{W}s}e^{-\widetilde{W}s} & 0_{m\times 1} \\ *_{1\times m} & 0\end{bmatrix}$$
$$= \begin{bmatrix} I_m & 0_{m\times 1} \\ *_{1\times m} & 0\end{bmatrix}.$$

This together with Assumption 4 and Lemma 67 enables us to compute

$$\begin{aligned}\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^{\sigma}(s))}(2\pi) &= \overline{\exp(s\sigma \mathrm{ad}_{\xi_{\mathrm{re}}})(Bw^2(s))}(2\pi) \\ &= \overline{\exp(s\sigma \mathrm{ad}_{\xi_{\mathrm{re}}})(\begin{bmatrix} B & \xi_{\mathrm{re}}\end{bmatrix} \begin{bmatrix} I_m \\ 0_{1\times m}\end{bmatrix} w^2(s))}(2\pi) \\ &= \overline{\begin{bmatrix} B & \xi_{\mathrm{re}}\end{bmatrix} e^{\sigma W s} \begin{bmatrix} I_m \\ 0_{1\times m}\end{bmatrix} w^2(s)}(2\pi) \\ &= \tfrac{1}{\pi}\overline{\begin{bmatrix} B & \xi_{\mathrm{re}}\end{bmatrix} e^{s\sigma W} \begin{bmatrix} I_m & 0_{m\times 1} \\ 0_{1\times m} & 0\end{bmatrix} e^{-s\sigma W} \begin{bmatrix} \mu \\ 0\end{bmatrix} \sin^2(s)}(2\pi) \\ &= \tfrac{1}{\pi}\overline{\begin{bmatrix} B & \xi_{\mathrm{re}}\end{bmatrix} \begin{bmatrix} I_m & 0_{m\times 1} \\ *_{1\times m} & 0\end{bmatrix} \begin{bmatrix} \mu \\ 0\end{bmatrix} \sin^2(s)}(2\pi) \\ &= B\mu + \delta\xi_{\mathrm{re}}.\end{aligned}$$

□

If Assumption 5 is satisfied this result, and its proof, can be simplified.

Proposition 75 (`configuration_inversion`). *Let Σ be a simple mechanical control system on a Lie group with a relative equilibrium ξ_{re} and corresponding matrix A_{re} and satisfying Assumptions 2 and 5. Let $Q, M \in \mathbb{R}^{m\times m}$ satisfy $A_{\mathrm{re}}B = BQ$ and $\mathrm{ad}_{\xi_{\mathrm{re}}}B = BM$. If $\mu \in \mathbb{R}^m$, $\sigma \in \mathbb{R}$ and*

$$\begin{aligned} u^1(t) &= 0, \\ u^2(t) &= F^{-1}_{\sigma Q}[w^2](t), & w^2(t) &= \tfrac{1}{\pi} e^{-\sigma M t}\mu \sin^2(t),\end{aligned}$$

then $b^1(t) = Bu^1(t)$ and $b^2(t) = Bu^2(t)$ satisfy

$$\begin{aligned} -\tfrac{1}{2}\overline{\langle \overline{b^1}^{\sigma} : \overline{b^1}^{\sigma}\rangle}^{\sigma}(2\pi) + \overline{b^2}^{\sigma}(2\pi) &= 0, \\ \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^{\sigma}(s))}(2\pi) &= B\mu.\end{aligned}$$

We denote this map `configuration_inversion`$(\sigma,\mu) = (b^1(t), b^2(t)) = (0, b^2(t))$.

Proof. For $b^1(t) = 0$ we have, using Lemma 67 and $w^2(t) = \frac{1}{\pi}e^{-\sigma M t}\mu\sin^2(t)$, that

$$-\tfrac{1}{2}\overline{\langle \overline{b^1}^{\sigma} : \overline{b^1}^{\sigma}\rangle}^{\sigma}(2\pi) + \overline{b^2}^{\sigma}(2\pi) = \overline{b^2}^{\sigma}(2\pi) = Bw^2(2\pi) = 0.$$

Using Assumption 5 and Lemma 67 we compute

$$\begin{aligned}\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^{\sigma}(s))}(2\pi) &= \overline{\exp(s\sigma \mathrm{ad}_{\xi_{\mathrm{re}}})(Bw^2(s))}(2\pi) = \overline{Be^{\sigma M s} w^2(s)}(2\pi) \\ &= \tfrac{1}{\pi}\overline{B\mu\sin^2(s)}(2\pi) = B\mu.\end{aligned}$$

□

The algorithm presented in this section requires the following additional assumption.

Assumption 6. The n dimensional system Σ has $n-1$ control forces, that is, $m = n-1$.

Assumption 6 together with the standing assumption $\xi_{\mathrm{re}} \notin \operatorname{span}\{b_1, \ldots, b_m\}$ implies $\mathbb{R}^n = \operatorname{span}\{b_1, \ldots, b_m, \xi_{\mathrm{re}}\}$—thus Assumption 4 is trivially satisfied. These assumption therefore gives that

$$\langle b_j : b_k \rangle = \sum_{i=1}^{m} \alpha_{jk}^i b_i + \alpha_{jk}^0 \xi_{\mathrm{re}}, \quad j,k \in \{1, \ldots, m\}$$

where $\alpha_{jk}^i \in \mathbb{R}$, $i \in \{0, 1, \ldots, m\}$. Therefore

$$\langle \xi_{\mathrm{re}} : \langle b_j : b_k \rangle \rangle = \sum_{i=1}^{m} \langle \xi_{\mathrm{re}} : \alpha_{jk}^i b_i + \alpha_{jk}^0 \xi_{\mathrm{re}} \rangle = \sum_{i=1}^{m} \alpha_{jk}^i \langle \xi_{\mathrm{re}} : b_i \rangle.$$

Thus Assumption 6 along with Assumption 2 (and the standing assumption $\xi_{\mathrm{re}} \notin \operatorname{span}\{b_1, \ldots, b_m\}$) imply Assumption 3.

Since Assumption 6 gives $\mathbb{R}^n = \operatorname{span}\{b_1, \ldots, b_m, \xi_{\mathrm{re}}\}$ every $\nu \in \mathbb{R}^n$ can be written $\nu = \sum_{i=1}^m \nu_i b_i + \nu_0 \xi_{\mathrm{re}}$ and with this we define the projection operators $\mathcal{P}_B : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathcal{P}_{\xi_{\mathrm{re}}} : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\mathcal{P}_{\xi_{\mathrm{re}}}\Big(\sum_{i=1}^{m} \nu_i b_i + \nu_0 \xi_{\mathrm{re}} \Big) := \nu_0 \xi_{\mathrm{re}}, \qquad \mathcal{P}_B := \mathrm{id} - \mathcal{P}_{\xi_{\mathrm{re}}},$$

where id is the identity. Notice that, under Assumption 5, these projection operators commute with $\mathrm{ad}_{\xi_{\mathrm{re}}}$—this is not the case under Assumption 4. This allows us to construct the following motion primitive.

Proposition 76 (`change_speed` motion primitive)**.** *Let Σ be a simple mechanical control system on a Lie group with a relative equilibrium ξ_{re} and corresponding matrix A_{re} and satisfying Assumptions 1, 2, and 6. For $0 < \epsilon \ll 1$, assume that*

$$\begin{aligned} g(0) &= g_0 \exp(\epsilon^2 \nu_{\mathrm{error}}), \\ \xi(0) &= \sigma \xi_{\mathrm{re}} + \epsilon^2 \xi_{\mathrm{error}}, \end{aligned}$$

for some $g_0 \in G$, $\sigma \in \mathbb{R}$, $\nu_{\mathrm{error}}, \xi_{\mathrm{error}} \in \mathbb{R}^n$ with $\nu_{\mathrm{error}} = \mathcal{O}(1)$ and $\xi_{\mathrm{error}} = \mathcal{O}(1)$. If we take $\rho \in \mathbb{R}$ and

$$(b^1(t), b^2(t)) = \begin{cases} \texttt{speed\_inversion}(\sigma, \rho\xi_{\mathrm{re}} - e^{2\pi\sigma A_{\mathrm{re}}}\xi_{\mathrm{error}}) \;, & t \in [0, 2\pi], \\ \texttt{configuration\_inversion}(\sigma, \mu) \;, & t \in [2\pi, 4\pi], \end{cases}$$

$$B\mu = -\mathcal{P}_B\left(\mathrm{Ad}_{\exp(-2\pi\sigma\xi_{\mathrm{re}})}\Big(\mathcal{P}_B(\nu_{\mathrm{error}}) + \frac{1}{\epsilon^2} \log\big(g(0)^{-1} g(2\pi) \exp(-2\pi\sigma\xi_{\mathrm{re}}) \big) \Big) \right),$$

then we obtain

$$\begin{aligned} g(4\pi) &= g_0^* \exp(\epsilon^2 \nu_{\mathrm{error}}^*), \\ \xi(4\pi) &= (\sigma + \epsilon^2 \rho)\xi_{\mathrm{re}} + \epsilon^2 \xi_{\mathrm{error}}^*, \end{aligned}$$

for some $\nu_{\mathrm{error}}^*, \xi_{\mathrm{error}}^* \in \mathbb{R}^n$ with $\mathcal{P}_{\xi_{\mathrm{re}}}(\nu_{\mathrm{error}}^*) = \mathcal{O}(1)$, $\mathcal{P}_B(\nu_{\mathrm{error}}^*) = \mathcal{O}(\epsilon)$, $\xi_{\mathrm{error}}^* = \mathcal{O}(\epsilon)$ and for

$$g_0^* = g_0 \exp\Big((4\pi\sigma + 2\pi\epsilon^2\rho)\xi_{\mathrm{re}} + \epsilon^2 \mathcal{P}_{\xi_{\mathrm{re}}}(\nu_{\mathrm{error}}) \Big).$$

We denote this control map by $(\sigma+\epsilon^2\rho, g_0^*, \nu_{\mathrm{error}}^*, \xi_{\mathrm{error}}^*) =$ `change_speed`$(\epsilon, \sigma, \rho, g_0, \nu_{\mathrm{error}}, \xi_{\mathrm{error}})$.

Proof. Using Propositions 61 and 72 we compute

$$\xi(2\pi) = \sigma\xi_{\mathrm{re}} + \epsilon^2\left(e^{\sigma A_{\mathrm{re}}2\pi}\xi_{\mathrm{error}} + \rho\xi_{\mathrm{re}} - e^{\sigma A_{\mathrm{re}}2\pi}\xi_{\mathrm{error}}\right) + \mathcal{O}(\epsilon^3) = (\sigma + \rho\epsilon^2)\xi_{\mathrm{re}} + \mathcal{O}(\epsilon^3),$$

and from this, Propositions 61 and 75 we have $\xi(4\pi) = (\sigma + \rho\epsilon^2)\xi_{\mathrm{re}} + \mathcal{O}(\epsilon^3)$. Define $g_{0,1/2} := g_0 \exp\left((2\pi\sigma + \epsilon^2\tilde{\nu})\xi_{\mathrm{re}}\right)$, $\tilde{\nu}\xi_{\mathrm{re}} := \mathcal{P}_{\xi_{\mathrm{re}}}(\nu_{\mathrm{error}})$, and $\nu_B := \mathcal{P}_B(\nu_{\mathrm{error}})$, then we achieve using Proposition 61 and Corollary 32

$$\begin{aligned}
g_{0,1/2}^{-1}g(2\pi) &= \exp\left(-(2\pi\sigma + \epsilon^2\tilde{\nu})\xi_{\mathrm{re}}\right)g_0^{-1}g(0)\exp(\epsilon^2 x^2(2\pi) + \mathcal{O}(\epsilon^3))\exp(2\pi\sigma\xi_{\mathrm{re}})\\
&= \exp\left(-(2\pi\sigma + \epsilon^2\tilde{\nu})\xi_{\mathrm{re}}\right)\exp(\epsilon^2(\tilde{\nu}\xi_{\mathrm{re}} + \nu_B))\exp(\epsilon^2 x^2(2\pi) + \mathcal{O}(\epsilon^3))\exp(2\pi\sigma\xi_{\mathrm{re}})\\
&= \exp(-2\pi\sigma\xi_{\mathrm{re}})\exp(\epsilon^2\nu_B + \mathcal{O}(\epsilon^4))\exp(\epsilon^2 x^2(2\pi) + \mathcal{O}(\epsilon^3))\exp(2\pi\sigma\xi_{\mathrm{re}})\\
&= \exp\left(\epsilon^2\mathrm{Ad}_{\exp(-2\pi\sigma\xi_{\mathrm{re}})}(\nu_B + x^2(2\pi)) + \mathcal{O}(\epsilon^3)\right).
\end{aligned}$$

From Propositions 61 and 72 we know that

$$x^2(2\pi) = \frac{1}{\epsilon^2}\log\left(g(0)^{-1}g(2\pi)\exp(-2\pi\sigma\xi_{\mathrm{re}})\right) + \mathcal{O}(\epsilon).$$

The definition of g_0^* and $g_{0,1/2}$ gives

$$\begin{gathered}
g_0^*\exp(-2\pi(\sigma + \epsilon^2\rho)\xi_{\mathrm{re}})g_{0,1/2}^{-1} =\\
g_0\exp((4\pi\sigma + 2\pi\epsilon^2\rho)\xi_{\mathrm{re}} + \epsilon^2\tilde{\nu}\xi_{\mathrm{re}})\exp(-2\pi(\sigma + \epsilon^2\rho)\xi_{\mathrm{re}})\exp(-(2\pi\sigma + \epsilon^2\tilde{\nu})\xi_{\mathrm{re}})g_0^{-1} =\\
\mathrm{id}.
\end{gathered}$$

Using these results, Propositions 61, 72, 75, and Corollary 32 we obtain

$$\begin{aligned}
g(4\pi) &= g(2\pi)\exp(\epsilon^2(B\mu + \delta\xi_{\mathrm{re}}) + \mathcal{O}(\epsilon^3))\exp\left(2\pi(\sigma + \epsilon^2\rho)\xi_{\mathrm{re}}\right)\\
&= g_0^*\exp\left(-2\pi(\sigma + \epsilon^2\rho)\xi_{\mathrm{re}}\right)g_{0,1/2}^{-1}g(2\pi)\exp(\epsilon^2(B\mu + \delta\xi_{\mathrm{re}}) + \mathcal{O}(\epsilon^3))\exp\left(2\pi(\sigma + \epsilon^2\rho)\xi_{\mathrm{re}}\right)\\
&= g_0^*\exp\left(\epsilon^2\mathrm{Ad}_{\exp(-2\pi(\sigma+\epsilon^2\rho)\xi_{\mathrm{re}})}\left(\mathrm{Ad}_{\exp(-2\pi\sigma\xi_{\mathrm{re}})}(\nu_B + x^2(2\pi)) + B\mu + \delta\xi_{\mathrm{re}}\right) + \mathcal{O}(\epsilon^3)\right)\\
&= g_0^*\exp\left(\epsilon^2\mathrm{Ad}_{\exp(-2\pi(\sigma+\epsilon^2\rho)\xi_{\mathrm{re}})}\left(\mathcal{P}_{\xi_{\mathrm{re}}}\left(\mathrm{Ad}_{\exp(-2\pi\sigma\xi_{\mathrm{re}})}(\nu_B + x^2(2\pi))\right) + \delta\xi_{\mathrm{re}}\right) + \mathcal{O}(\epsilon^3)\right)\\
&= g_0^*\exp\left(\epsilon^2\left(\mathcal{P}_{\xi_{\mathrm{re}}}\left(\mathrm{Ad}_{\exp(-2\pi\sigma\xi_{\mathrm{re}})}(\nu_B + x^2(2\pi))\right) + \delta\xi_{\mathrm{re}}\right) + \mathcal{O}(\epsilon^3)\right).
\end{aligned}$$

□

With this motion primitive we are able to construct the following algorithm that speeds up, slows down, or stabilizes, a system along a relative equilibrium.

Proposition 77 (`speed_control` algorithm)**.** *Let* Σ *be a simple mechanical control system on a Lie group with a relative equilibrium* ξ_{re} *and corresponding matrix* A_{re}*. Assume* Σ *satisfies Assumptions 1, 2, and 6 and take* $0 < \epsilon \ll 1$*. Let* $g(0)$*,* g_0*,* ν_{error}*,* σ*,* ξ_{error}*,* ρ *be as in Proposition 76 and let* $N \in \mathbb{N}$*.*

Define the algorithm $(\sigma+\epsilon^2 N\rho, g_0^*, \nu_{\mathrm{error}}^*, \xi_{\mathrm{error}}^*) =$ `speed_control`$(\epsilon, \sigma, \rho, N, g_0, \nu_{\mathrm{error}}, \xi_{\mathrm{error}})$ *by*

1: $g_{0,1} := g_0$; $\nu_{\mathrm{error},1} := \nu_{\mathrm{error}}$; $\sigma_1 := \sigma$; $\xi_{\mathrm{error},1} := \xi_{\mathrm{error}}$;
2: ***for*** $k \in \{1, \ldots, N\}$ ***do***
3: $\quad(\sigma_{k+1}, g_{0,k+1}, \nu_{\mathrm{error},k+1}, \xi_{\mathrm{error},k+1}) :=$ `change_speed`$(\epsilon, \sigma_k, \rho, g_{0,k}, \nu_{\mathrm{error},k}, \xi_{\mathrm{error},k})$
4: ***end for***
5: $g_0^* = g_{0,N+1}$; $\nu_{\mathrm{error}}^* := \nu_{\mathrm{error},N+1}$; $\xi_{\mathrm{error}}^* := \xi_{\mathrm{error},N+1}$;

The final configuration and velocity after the execution of this algorithm are

$$g(N4\pi) = g_0^* \exp(\epsilon^2 \nu_{\text{error}}^*),$$
$$\xi(N4\pi) = (\sigma + \epsilon^2 N\rho)\xi_{\text{re}} + \epsilon^2 \xi_{\text{error}}^*,$$

where $\nu_{\text{error}}^*, \xi_{\text{error}}^* \in \mathbb{R}^n$, $\mathcal{P}_{\xi_{\text{re}}}(\nu_{\text{error}}^*) = \mathcal{O}(1)$, $\mathcal{P}_B(\nu_{\text{error}}^*) = \mathcal{O}(\epsilon)$, $\xi_{\text{error}}^* = \mathcal{O}(\epsilon)$, *and*

$$g_0^* = g_0 \exp\left(\left(\sigma T_{\text{final}} + \tfrac{1}{2}\rho\epsilon^2 N T_{\text{final}}\right)\xi_{\text{re}} + \epsilon^2 \sum_{k=1}^{N} \mathcal{P}_{\xi_{\text{re}}}(\nu_{\text{error},k})\right).$$

Proof. From Proposition 76 we have $\sigma_k = \sigma + (k-1)\rho\epsilon^2$ so we immediately obtain $\xi(N4\pi) = \sigma_{N+1}\xi_{\text{re}} + \mathcal{O}(\epsilon^3) = (\sigma + \epsilon^2 N\rho)\xi_{\text{re}} + \mathcal{O}(\epsilon^3)$. From Proposition 76 we have $g(N4\pi) = g_0^* \exp(\epsilon^2 \nu_{\text{error}}^*)$ where

$$\begin{aligned}
g_0^* &= g_0 \left(\prod_{k=1}^{N} \exp\left(2\pi(2\sigma_k + \rho\epsilon^2)\xi_{\text{re}} + \epsilon^2 \mathcal{P}_{\xi_{\text{re}}}(\nu_{\text{error},k})\right)\right) \\
&= g_0 \exp\left(\sum_{k=1}^{N} \left(2\pi(2\sigma_k + \rho\epsilon^2)\xi_{\text{re}} + \epsilon^2 \mathcal{P}_{\xi_{\text{re}}}(\nu_{\text{error},k})\right)\right) \\
&= g_0 \exp\left(2\pi N\left(2\sigma + N\rho\epsilon^2\right)\xi_{\text{re}} + \epsilon^2 \sum_{k=1}^{N} \mathcal{P}_{\xi_{\text{re}}}(\nu_{\text{error},k})\right) \\
&= g_0 \exp\left(\left(\sigma T_{\text{final}} + \tfrac{1}{2}\rho\epsilon^2 N T_{\text{final}}\right)\xi_{\text{re}} + \epsilon^2 \sum_{k=1}^{N} \mathcal{P}_{\xi_{\text{re}}}(\nu_{\text{error},k})\right).
\end{aligned}$$

From Proposition 76, its proof, and Proposition 61, we have that `change_speed` gives the map $(\xi_{\text{error},k}, \mathcal{P}_B(\nu_{\text{error},k}), \sigma) \mapsto (\xi_{\text{error},k+1}, \mathcal{P}_B(\nu_{\text{error},k+1}), \sigma + \epsilon^2\rho)$ independent of g_0 and $\mathcal{P}_{\xi_{\text{re}}}(\nu_{\text{error},k})$. Because $(\xi_{\text{error},k}, \mathcal{P}_B(\nu_{\text{error},k})) = \mathcal{O}(1)$ gives $(\xi_{\text{error},k+1}, \mathcal{P}_B(\nu_{\text{error},k+1})) = \mathcal{O}(\epsilon)$ we obtain that $\mathcal{P}_B(\nu_{\text{error},k}) = \mathcal{O}(\epsilon, k) = \mathcal{O}(\epsilon)$, $\mathcal{P}_{\xi_{\text{re}}}(\nu_{\text{error},k}) = \mathcal{O}(1, k) = \mathcal{O}(1)$, and $\xi_{\text{error},k} = \mathcal{O}(\epsilon, k) = \mathcal{O}(\epsilon)$. □

Note that $\rho > 0$ speeds up the system along the relative equilibrium, $\rho < 0$ slows down the system, and $\rho = 0$ stabilizes the system's motion along the relative equilibrium. We may select $N = \mathcal{O}(\frac{1}{\epsilon^2})$ in Proposition 77 so that the absolute change in velocity along the relative equilibrium is of order $\mathcal{O}(1)$. Thus, it is possible to use the algorithm `speed_control` to change the velocity along the relative equilibrium from a given value to another independent of ϵ.

5.3.1 Interlude Regarding the Assumptions

In this section we examine some possibilities for relaxing some of the assumptions needed for the `change_speed` motion primitive. It turns out that an alternative speed inversion map can be created such that the assumption regarding the linearly controllable subspace, i.e. Assumption 2, can be weakened. Also the $m = n - 1$ assumption, i.e. Assumption 6, can be weakened, at least for $m \leq 3$. The replacing assumption seems to be too strict though, but the analysis gives insight into the difficulties that arises when removing the $m = n - 1$ assumption.

The objective of this section is to point out and clarify the difficulties that arises when weakening the assumptions. The theory will not be applied to any example systems.

Weakening the Assumption Concerning Linear Controllability

We define $\{b_{m+1}, \dots, b_l\}$ by $\mathrm{span}\{b_1, \dots, b_m, b_{m+1}, \dots, b_l\} := \mathrm{span}\{b_1, \dots, b_m, \langle \xi_{\mathrm{re}} : b_1 \rangle, \dots, \langle \xi_{\mathrm{re}} : b_m \rangle\}$, where $b_1, \dots, b_m, b_{m+1}, \dots, b_l$ are linearly independent, and let $\widetilde{B} := \begin{bmatrix} b_1 & \dots & b_l \end{bmatrix}$.

Assumption 7 (Linear controllable subspace)**.** The subspace $\mathrm{span}\{b_1, \dots, b_l\}$ is invariant under the linear map A_{re}, that is, $\langle \xi_{\mathrm{re}} : b_i \rangle \in \mathrm{span}\{b_1, \dots, b_l\}$, for $i \in \{1, \dots, l\}$.

Assumption 8. $l > m$ and $\langle b_q : b_v \rangle \in \mathrm{span}\{b_1, \dots, b_l\}$ for $q \in \{1, \dots, l\}$, $v \in \{m+1, \dots, l\}$, and $\langle \xi_{\mathrm{re}} : \langle b_j : b_k \rangle \rangle \in \mathrm{span}\{b_1, \dots, b_l\}$ for $j, k \in \{1, \dots, l\}$.

Assumption 7 means that there exists a matrix $\widetilde{Q} \in \mathbb{R}^{l \times l}$ such that $A_{\mathrm{re}} \widetilde{B} = \widetilde{B} \widetilde{Q}$. We define $\widetilde{Q} := \left[\begin{smallmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{smallmatrix}\right]$ where $Q_{11} \in \mathbb{R}^{m \times m}$, $Q_{12} \in \mathbb{R}^{m \times (l-m)}$, $Q_{21} \in \mathbb{R}^{(l-m) \times m}$, and $Q_{22} \in \mathbb{R}^{(l-m) \times (l-m)}$.

Given $\widetilde{Q} := \left[\begin{smallmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{smallmatrix}\right]$, $Q_{11} \in \mathbb{R}^{m \times m}$, $Q_{12}, Q_{21}^T \in \mathbb{R}^{m \times (l-m)}$, $Q_{22} \in \mathbb{R}^{(l-m) \times (l-m)}$, define $\mathcal{L}_{\widetilde{Q}} : C^0([0, 2\pi], \mathbb{R}^l) \to \{f \in C^1([0, 2\pi], \mathbb{R}^l) \mid f(0) = 0\}$ by

$$\mathcal{L}_{\widetilde{Q}}[u](t) := \int_0^t e^{\widetilde{Q}(t-s)} u(s) \mathrm{d}s,$$

and use this to define $\mathcal{F}_{\widetilde{Q}} : C^0([0, 2\pi], \mathbb{R}^m) \to \{f \in C^1([0, 2\pi], \mathbb{R}^m) \mid f(0) = 0\}$ as

$$\mathcal{F}_{\widetilde{Q}}[u](t) := \begin{bmatrix} I_m & 0 \end{bmatrix} \mathcal{L}_{\widetilde{Q}} \left[\left[\begin{smallmatrix} I_m \\ 0 \end{smallmatrix}\right] u \right] (t).$$

Lemma 78 (Transformation of controls)**.** *The map $\mathcal{F}_{\widetilde{Q}}$ is invertible and its inverse is given as follows: if $w = \mathcal{F}_{\widetilde{Q}}[u]$, then $u(t) = -Q_{11} w(t) + \dot{w}(t) - Q_{12} \int_0^t e^{(t-s)Q_{22}} Q_{21} w(s) \mathrm{d}s$. Additionally, let Assumption 7 be satisfied, and let A_{re}, $\widetilde{B}$ and $\widetilde{Q}$ satisfy $A_{\mathrm{re}} \widetilde{B} = \widetilde{B} \widetilde{Q}$. If $u \in C^0([0, 2\pi], \mathbb{R}^m)$, $w = \mathcal{F}_{\sigma \widetilde{Q}}[u]$, and $z = \left[\begin{smallmatrix} 0 & I_{l-m} \end{smallmatrix}\right] \mathcal{L}_{\sigma \widetilde{Q}} \left[\left[\begin{smallmatrix} I_m \\ 0 \end{smallmatrix}\right] u \right]$, then*

$$\overline{Bu}^{\sigma}(t) = \widetilde{B} \left[\begin{smallmatrix} w(t) \\ z(t) \end{smallmatrix}\right].$$

Proof. One-to-one correspondence between u and w is readily checked. We compute $\overline{Bu}^{\sigma}(t) = \int_0^t e^{\sigma A_{\mathrm{re}}(t-s)} B u(s) \mathrm{d}s = \widetilde{B} \int_0^t e^{\sigma \widetilde{Q}(t-s)} \left[\begin{smallmatrix} I_m \\ 0 \end{smallmatrix}\right] u(s) \mathrm{d}s = \widetilde{B} \mathcal{L}_{\sigma \widetilde{Q}} \left[\left[\begin{smallmatrix} I_m \\ 0 \end{smallmatrix}\right] u \right] (t) = \widetilde{B} \left[\begin{smallmatrix} w(t) \\ z(t) \end{smallmatrix}\right]$. □

Notice that for $\sigma = \mathcal{O}(\epsilon)$, $0 < \epsilon \ll 1$, we get $z(t) = \mathcal{O}(\epsilon)$.

Proposition 79 (alternative speed inversion)**.** *Let Σ be a simple mechanical control system on a Lie group with a relative equilibrium ξ_{re} and corresponding matrix A_{re} and satisfying Assumptions 1, 7 and 8. Let $\widetilde{Q} \in \mathbb{R}^{l \times l}$ satisfy $A_{\mathrm{re}} \widetilde{B} = \widetilde{B} \widetilde{Q}$. Let $\eta \in \mathbb{R}^n$, $\sigma \in \mathbb{R}$, and compute $\tilde{z} \in \mathbb{R}^m$ and $Z \in \mathbb{R}^{n \times n}$ as the pseudoinverse solution to*

$$\eta = \sum_{i=1}^{m} \tilde{z}_i b_i - \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} Z_{jk} \langle b_j : b_k \rangle, \qquad Z_{jk} = 0 \text{ for } j \geq k.$$

Given r, α, $\mathcal{A}_{\sigma,\alpha}$, and β as in Definition 68, let

$$y_j(t) := \sum_{k=1}^{m} \lambda_{jk}(Z)\sin(\beta_{jk}t), \quad j \in \{1,\dots,m\},$$

and let $\gamma = (\gamma_{11},\dots,\gamma_{m1},\dots,\gamma_{1r},\dots,\gamma_{mr})^T$ be the unique solution to

$$\begin{aligned} \mathcal{A}_{\sigma,\alpha}\gamma &= -\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(By(s))}(2\pi), \\ \gamma_{ih} &= 0 \quad \textit{if} \quad \alpha_{ih} = 0 \ \textit{ for } \ (i,h) \in \{1,\dots,m\}\times\{1,\dots,r\}. \end{aligned} \tag{5.13}$$

Additionally, if we take

$$\begin{aligned} w_j^1(t) &= y_j(t) + \sum_{k=1}^{r} \gamma_{jk}\sin(\alpha_{jk}t), \quad j \in \{1,\dots,m\}, \\ u^1(t) &= \mathcal{F}^{-1}_{\sigma\widetilde{Q}}[w^1](t), \\ z^1(t) &= \begin{bmatrix} 0 & I_{l-m} \end{bmatrix} \mathcal{L}_{\sigma\widetilde{Q}}\left[\begin{bmatrix} I_m \\ 0 \end{bmatrix} u^1\right](t) \end{aligned} \tag{5.14}$$

$$u^2(t) = \left(\int_0^{2\pi} e^{\sigma\widetilde{Q}(2\pi-s)}\begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} e^{\sigma\widetilde{Q}^T(2\pi-s)}\mathrm{d}s e^{-\sigma\widetilde{Q}^T(2\pi-t)}\begin{bmatrix} I_m \\ 0 \end{bmatrix}\right)^{\#}\left(\chi + \begin{bmatrix} \tilde{z} \\ 0 \end{bmatrix}\right), \tag{5.15}$$

where $\chi \in \mathbb{R}^l$ is the unique solution to

$$\widetilde{B}\chi = \sum_{j=1}^{m-1}\sum_{k=j+1}^{m}\int_0^{2\pi}(e^{\sigma A_{\mathrm{re}}(2\pi-s)} - I)w_j^1(s)w_k^1(s)\mathrm{d}s\,\langle b_j : b_k\rangle + \int_0^{2\pi} e^{\sigma A_{\mathrm{re}}(2\pi-s)}f(s)\mathrm{d}s, \tag{5.16}$$

$$\begin{aligned} f(t) = \sum_{i=1}^{m}(w_i^1(t))^2\langle b_i : b_i\rangle &+ 2\sum_{j=1}^{m}\sum_{k=1}^{l-m} w_j^1(t)z_k^1(t)\langle b_j : b_{k+m}\rangle \\ &+ \sum_{j=1}^{l-m}\sum_{k=1}^{l-m} z_j^1(t)z_k^1(t)\langle b_{j+m} : b_{k+m}\rangle, \end{aligned}$$

then $b^1(t) = Bu^1(t)$ and $b^2(t) = Bu^2(t)$ satisfy

$$-\tfrac{1}{2}\overline{\langle \overline{b^1}^{\sigma} : \overline{b^1}^{\sigma}\rangle}^{\sigma}(2\pi) + \overline{b^2}^{\sigma}(2\pi) = \eta, \tag{5.17}$$

$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(s))}(2\pi) = 0. \tag{5.18}$$

Proof. Existence and uniqueness of the solution to (5.16) is a consequence of Assumptions 1 and 8. Regarding existence and uniqueness of the solution to (5.13), Definition 68 ensures that

$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(By(s))}(2\pi) \in \mathrm{Image}(\mathcal{A}_{\sigma,\alpha}).$$

Since every nonzero column in $\mathcal{A}_{\sigma,\alpha}$ contributes to the rank of $\mathcal{A}_{\sigma,\alpha}$, the entries of γ corresponding to these will be unique. The remaining γ-values are defined to be 0.

Regarding the proof of equation (5.18), direct calculations show that

$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(s))}(2\pi) = \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(Bw^1(s))}(2\pi) = \mathcal{A}_{\sigma,\alpha}\gamma + \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(By(s))}(2\pi) = 0.$$

Regarding the proof of equation (5.17), from Lemma 78 we compute

$$\begin{aligned}
\langle \overline{b}^{\sigma} : \overline{b}^{\sigma}\rangle(t) &= \langle \sum_{j=1}^{m} w_j^1(t)b_j + \sum_{j=1}^{l-m} z_j^1(t)b_{j+m} : \sum_{k=1}^{m} w_k^1(t)b_k + \sum_{k=1}^{l-m} z_k^1(t)b_{k+m}\rangle \\
&= 2\sum_{j=1}^{m-1}\sum_{k=j+1}^{m} w_j^1(t)w_k^1(t)\langle b_j : b_k\rangle + \sum_{i=1}^{m}(w_i^1(t))^2\langle b_i : b_i\rangle \\
&+ 2\sum_{j=1}^{m}\sum_{k=1}^{l-m} w_j^1(t)z_k^1(t)\langle b_j : b_{k+m}\rangle + \sum_{j=1}^{l-m}\sum_{k=1}^{l-m} z_j^1(t)z_k^1(t)\langle b_{j+m} : b_{k+m}\rangle \\
&= 2\sum_{j=1}^{m-1}\sum_{k=j+1}^{m} w_j^1(t)w_k^1(t)\langle b_j : b_k\rangle + f(t),
\end{aligned}$$

where $f(t) \in \mathrm{span}\{b_1, \ldots, b_l\}$. Furthermore, if we write $f(t) = \widetilde{B}\left[\begin{smallmatrix} f_1(t) \\ f_2(t)\end{smallmatrix}\right]$, $f_1 : [0, 2\pi] \to \mathbb{R}^m$, $f_2 : [0, 2\pi] \to \mathbb{R}^{l-m}$, then for $\sigma = \mathcal{O}(\epsilon)$ we have $f_2(t) = \mathcal{O}(\epsilon)$. This shows that in this limit we have $\chi = \left[\begin{smallmatrix} \chi_0 \\ \mathcal{O}(\epsilon)\end{smallmatrix}\right]$, for $\chi_0 \in \mathbb{R}^m$.

Since all nonzero α-values are distinct and are distinct from the β-values we have for $j < k$

$$\begin{aligned}
\int_0^{2\pi} w_j^1(t)w_k^1(t)\mathrm{d}t &= \sum_{l,q=1}^{m} \lambda_{jl}(Z)\lambda_{kq}(Z)\int_0^{2\pi} \sin(\beta_{jl}t)\sin(\beta_{kq}t)\mathrm{d}t \\
&= \sum_{l,q=1}^{m} \lambda_{jl}(Z)\lambda_{kq}(Z)\delta_{\beta_{kq}}^{\beta_{jl}}\pi = \lambda_{jk}(Z)\lambda_{kj}(Z)\pi = Z_{jk}.
\end{aligned}$$

From the theory of linear control systems, see e.g. [40], we know that when the system $\dot{x} = \sigma Qx + \left[\begin{smallmatrix} I_m \\ 0\end{smallmatrix}\right]u$, $x \in \mathbb{R}^l$, $u \in \mathbb{R}^m$, is controllable, i.e. for $\sigma \neq 0$, then for $x(0) = 0$ and arbitrary $x_f \in \mathbb{R}^l$ the control

$$u(t) = \left[\begin{smallmatrix} I_m & 0\end{smallmatrix}\right]e^{\sigma Q(2\pi - t)}\left(\int_0^{2\pi} e^{\sigma Q(2\pi - s)}\left[\begin{smallmatrix} I_m \\ 0\end{smallmatrix}\right]\left[\begin{smallmatrix} I_m & 0\end{smallmatrix}\right]e^{\sigma Q^T(2\pi - s)}\mathrm{d}s\right)^{-1} x_f$$

gives $x(2\pi) = x_f$; in fact this control is of minimum norm among the controls giving this end point. Using the pseudoinverse instead of the inverse and rearranging the terms this is equation (5.15) with $x_f = \chi + \left[\begin{smallmatrix} \tilde{z} \\ 0\end{smallmatrix}\right]$. Since for $\sigma = \mathcal{O}(\epsilon)$ we have $\chi = \left[\begin{smallmatrix} \chi_0 \\ \mathcal{O}(\epsilon)\end{smallmatrix}\right]$, the control (5.15) is also able to achieve the desired result in this limit without becoming infinite. Notice that that $\overline{b^2}^{\sigma}(2\pi) = x(2\pi)$.

By straightforward calculations we then obtain

$$
\begin{aligned}
&-\tfrac{1}{2}\overline{\langle \overline{b^1}^\sigma : \overline{b^1}^\sigma \rangle}^\sigma(2\pi) + \overline{b^2}^\sigma(2\pi) \\
&= -\tfrac{1}{2}\int_0^{2\pi} e^{\sigma A_{\mathrm{re}}(2\pi - s)} \langle \overline{b^1}^\sigma : \overline{b^1}^\sigma \rangle(s)\mathrm{d}s + \widetilde{B}\int_0^{2\pi} e^{\sigma \widetilde{Q}(2\pi - s)} u^2(s)\mathrm{d}s \\
&= -\sum_{j=1}^{m-1}\sum_{k=j+1}^{m}\Big(\int_0^{2\pi} w_j^1(s) w_k^1(s)\mathrm{d}s\langle b_j : b_k\rangle + \int_0^{2\pi}(e^{\sigma A_{\mathrm{re}}(2\pi - s)} - I) w_j^1(s) w_k^1(s)\mathrm{d}s \langle b_j : b_k\rangle\Big) \\
&\quad - \tfrac{1}{2}\sum_{j=1}^{m}\int_0^{2\pi} e^{\sigma A_{\mathrm{re}}(2\pi - s)} f(s)\mathrm{d}s + \widetilde{B}\left(\chi + \left[\begin{smallmatrix}\tilde{z}\\ 0\end{smallmatrix}\right]\right) \\
&= -\sum_{j=1}^{m-1}\sum_{k=j+1}^{m} Z_{jk}\langle b_j : b_k\rangle + \sum_{i=1}^{m}\tilde{z}_i b_i = \eta.
\end{aligned}
$$

□

From the proof of this proposition we see that replacing Assumption 2 with something more general inevitably leads to order considerations since the algorithms in [11] are not able to generate velocities above the order $\mathcal{O}(\epsilon)$, meaning that any new algorithm at least needs to be able to start with $\sigma = \mathcal{O}(\epsilon)$. In this limit the linearly controllable subspace is $\mathrm{span}\{B, \sigma A_{\mathrm{re}}B, \ldots, \sigma^{n-1}A_{\mathrm{re}}^{n-1}B\}$ so we need to be able to guarantee that every part of the error, which we need to correct with b^2, is of the same or higher order as the part of the controllable subspace it belongs to. Otherwise the control u^2 will grow unbounded as $\sigma \to 0$. In this analysis we could also have included the subspace generated by the vectors in $A_{\mathrm{re}}^2 B$ but this would have complicated the analysis and is therefore left out since the reason for this section is solely to point out some of the complications that arises when weakening the assumptions. We could not have included vectors from $A_{\mathrm{re}}^k B$, $k > 2$, since these would not suffice in the limit $\sigma = \mathcal{O}(\epsilon)$.

Since `configuration_inversion`, see Proposition 74, needs Assumption 2 this new alternative speed inversion map can not be used in the construction of a replacement for `change_speed`, needing weaker assumptions, unless a version of `configuration_inversion` which uses Assumption 7 instead of 2 is constructed.

Removing the $m = n - 1$ Condition

If Assumptions 2 and 5 are satisfied then `configuration_inversion` gives a way to construct b^2, while $b^1 = 0$, such that for any $\mu \in \mathbb{R}^m$ we get

$$
\begin{aligned}
-\tfrac{1}{2}\overline{\langle \overline{b^1}^\sigma : \overline{b^1}^\sigma \rangle}^\sigma(2\pi) + \overline{b^2}^\sigma(2\pi) &= 0, \\
\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^\sigma(s))}(2\pi) &= B\mu.
\end{aligned}
$$

This means that, when $\xi(0) = \sigma\xi_{\mathrm{re}} + \epsilon^2\xi_0^2$, we get

$$
\begin{aligned}
\xi(2\pi) &= \sigma\xi_{\mathrm{re}} + \epsilon^2 e^{\sigma A_{\mathrm{re}}2\pi}\xi_0^2 + \mathcal{O}(\epsilon^3), \\
\log\big(g(0)^{-1}g(2\pi)\exp(-2\pi\sigma\xi_{\mathrm{re}})\big) &= \epsilon^2 B\mu + \mathcal{O}(\epsilon^3).
\end{aligned}
$$

The limitation of `configuration_inversion` is thus that it is only able to correct errors in the configuration that are a result of motion in span$\{b_1, \ldots, b_m\}$. If we take $\sigma = \mathcal{O}(\epsilon)$ we get

$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^{\sigma}(s))}(2\pi) = \overline{\overline{b^2}}(2\pi) + \mathcal{O}(\epsilon)$$

and therefore, when $b^1 = 0$, this problem cannot be avoided by any assumption. In the previous section we therefore assume $m = n-1$ such that $\mathbb{R}^n = \mathrm{span}\{b_1, \ldots, b_m, \xi_{\mathrm{re}}\}$ because we are not interested in how precisely the configuration behaves along ξ_{re}.

If we do not want to assume $m = n-1$ but instead seek to replace `configuration_inversion` with another scheme requiring weaker assumptions we thus need $b^1 \neq 0$ and the equations (which were satisfied by the controls given by `configuration_inversion`) b^1 and b^2 need to satisfy are

$$\begin{aligned}
0 &= \overline{b^1}^{\sigma}(2\pi), \\
0 &= -\tfrac{1}{2}\overline{\langle \overline{b^1}^{\sigma} : \overline{b^1}^{\sigma} \rangle}^{\sigma}(2\pi) + \overline{b^2}^{\sigma}(2\pi), \\
0 &= \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(s))}(2\pi), \\
\nu + \nu_0 &= -\tfrac{1}{2}\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{\langle \overline{b^1}^{\sigma} : \overline{b^1}^{\sigma} \rangle}^{\sigma}(s))}(2\pi) \\
&\quad + \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^{\sigma}(s))}(2\pi) - \tfrac{1}{2}\overline{[\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(s)), \overline{\mathrm{Ad}_{\exp(\tau\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(\tau))}(s)]}(2\pi),
\end{aligned}$$

where $\nu \in \mathrm{span}\{b_1, \ldots, b_m, v_{m+1}, \ldots, v_{n-1}\}$ and $\mathbb{R}^n = \mathrm{span}\{b_1, \ldots, b_m, v_{m+1}, \ldots, v_{n-1}, \xi_{\mathrm{re}}\}$. The value of $\nu_0 \in \mathrm{span}\{\xi_{\mathrm{re}}\}$ is irrelevant. It is fairly apparent that the complexity of the problem increases considerably when $m \neq n-1$.

In the following Assumptions 2 and 5 are implied and the matrix $M \in \mathbb{R}^{m\times m}$ is the one satisfying $\mathrm{ad}_{\xi_{\mathrm{re}}} B = BM$. To replace the $m = n-1$ assumption we instead assume

$$A_{\mathrm{re}}(\mathbb{R}^n) \subset \mathrm{span}\{b_1, \ldots, b_m, \xi_{\mathrm{re}}\}, \tag{5.19}$$

$$\mathrm{ad}_{\xi_{\mathrm{re}}}(\mathbb{R}^n) \subset \mathrm{span}\{b_1, \ldots, b_m, \xi_{\mathrm{re}}\}, \tag{5.20}$$

$$[b_j, b_k] \in \mathrm{span}\{b_1, \ldots, b_m, \xi_{\mathrm{re}}\}, \quad j,k \in \{1, \ldots, m\}, \tag{5.21}$$

which is trivially satisfied for $m = n-1$. Since it is still a strong assumption it is therefore close to the $m = n-1$ assumption in some sense. Assumptions 5.19-5.21 means that

$$\overline{[\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(s)), \overline{\mathrm{Ad}_{\exp(\tau\sigma\xi_{\mathrm{re}})}(\overline{b^1}^{\sigma}(\tau))}(s)]}(2\pi) \in \mathrm{span}\{b_1, \ldots, b_m, \xi_{\mathrm{re}}\},$$

so the error produced by this term can be corrected by b^2 according to Proposition 75. Define $\mathcal{P}_{\{B,\xi_{\mathrm{re}}\}^{\perp}} : \mathbb{R}^n \to \mathbb{R}^n$ to be the projection onto span$\{v_{m+1}, \ldots, v_{n-1}\}$ given by

$$\mathcal{P}_{\{B,\xi_{\mathrm{re}}\}^{\perp}}\left(\sum_{i=1}^{m} a_i b_i + \sum_{i=m+1}^{n-1} a_i v_i + a_n \xi_{\mathrm{re}}\right) = \sum_{i=m+1}^{n-1} a_i v_i,$$

for $a \in \mathbb{R}^n$. The assumptions (5.19)-(5.21) gives

$$\begin{aligned}
&\mathcal{P}_{\{B,\xi_{\mathrm{re}}\}^{\perp}}\left(\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{\langle \overline{b^1}^{\sigma} : \overline{b^1}^{\sigma} \rangle}^{\sigma}(s))}(2\pi)\right) \\
&= \sum_{1\le j<k\le m} \int_0^{2\pi}\int_0^t w_j^1(s) w_k^1(s)\,\mathrm{d}s\mathrm{d}t\ \mathcal{P}_{\{B,\xi_{\mathrm{re}}\}^{\perp}}(\langle b_j : b_k \rangle),
\end{aligned}$$

which is "as nice as possible". By choosing $w^2(t) = e^{-\sigma Mt}\big(\frac{\mu}{\pi}\sin^2(t) + \frac{3\varsigma}{4\pi^2}(t - \frac{4\pi}{3})t\big)$, $\mu, \varsigma \in \mathbb{R}^m$, we get

$$\overline{b^2}^\sigma(2\pi) = Be^{-\sigma M2\pi}\varsigma,$$
$$\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^\sigma(s))}(2\pi) = B\mu,$$

and we can therefore choose μ and ς such that the problem reduces to finding w^1 which satisfies

$$0 = \overline{b^1}^\sigma(2\pi) = Bw^1(2\pi), \tag{5.22}$$
$$\begin{aligned}0 &= -\tfrac{1}{2}\overline{\langle\overline{b^1}^\sigma : \overline{b^1}^\sigma\rangle}^\sigma(2\pi) + \overline{b^2}^\sigma(2\pi)\\ &= -\sum_{1\le j<k\le m}\int_0^{2\pi} w_j^1(s)w_k^1(s)\mathrm{d}s\langle b_j : b_k\rangle,\end{aligned} \tag{5.23}$$
$$\begin{aligned}0 &= \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^\sigma(s))}(2\pi)\\ &= B\int_0^{2\pi} e^{s\sigma M}w^1(s)\mathrm{d}s,\end{aligned} \tag{5.24}$$
$$\begin{aligned}\nu + \nu_0 &= -\tfrac{1}{2}\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{\langle\overline{b^1}^\sigma : \overline{b^1}^\sigma\rangle}^\sigma(s))}(2\pi)\\ &\quad + \overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^2}^\sigma(s))}(2\pi) - \tfrac{1}{2}\overline{[\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(\overline{b^1}^\sigma(s)), \overline{\mathrm{Ad}_{\exp(\tau\sigma\xi_{\mathrm{re}})}(\overline{b^1}^\sigma(\tau))}(s)]}(2\pi)\\ &= \sum_{1\le j<k\le m}\int_0^{2\pi}\int_0^t w_j^1(s)w_k^1(s)\mathrm{d}s\mathrm{d}t\ \mathcal{P}_{\{B,\xi_{\mathrm{re}}\}^\perp}(\langle b_j : b_k\rangle) + B\omega,\end{aligned} \tag{5.25}$$

where, for any $\omega \in \mathbb{R}^m$, we can choose μ such that this is the result.

Define

$$\psi_j(t) = \sin(jt), \qquad j \in \mathbb{N},$$
$$\phi_k(t) = \cos((m+k)t) - \cos(kt), \qquad k \in \mathbb{N},$$

and

$$\Lambda_{jk} = \int_0^{2\pi}\int_0^t \psi_j(s)\phi_k(s)\mathrm{d}s\mathrm{d}t, \quad j, k \in \mathbb{N},$$

which is nonzero. If we use

$$w_j^1(t) = \psi_j(t) + \gamma_{\alpha_j}\phi_{\alpha_j}(t) + \sum_{k=1}^m \delta_{jk}\phi_k(t), \quad j \in \{1, \dots, m\},$$

and choose $\alpha_1, \dots, \alpha_m \in \mathbb{N} \setminus \{1, \dots, m\}$, all distinct, such that the rank of

$$\int_0^{2\pi} e^{s\sigma M}\mathrm{diag}(\phi_{\alpha_1}(s), \dots, \phi_{\alpha_m}(s))\mathrm{d}s,$$

is maximal then we can choose $\gamma_{\alpha_1}, \dots, \gamma_{\alpha_m} \in \mathbb{R}$, such that (5.24) is satisfied. Then for equation (5.23) and (5.25) (equation (5.22) is trivially satisfied) to be solved means solving

$$\int_0^{2\pi} w_j^1(s)w_k^1(s)\mathrm{d}s = 0, \qquad \int_0^{2\pi}\int_0^t w_j^1(s)w_k^1(s)\mathrm{d}s\mathrm{d}t = Z_{jk},$$

for arbitrary $Z_{jk} \in \mathbb{R}$, which reduces to finding $\delta \in \mathbb{R}^{m\times m}$ satisfying

$$\sum_{l=1}^{m} \delta_{jl}\delta_{kl} = 0, \qquad 1 \le j < k \le m, \tag{5.26}$$

$$\sum_{l=1}^{m}(\delta_{jl}\Lambda_{kl} + \delta_{kl}\Lambda_{jl}) = Z_{jk} - \gamma_{\alpha_j}\Lambda_{k\alpha_j} - \gamma_{\alpha_k}\Lambda_{j\alpha_k} =: X_{jk}, \qquad 1 \le j < k \le m. \tag{5.27}$$

For $m = 2$ such a solution is $\delta = \operatorname{diag}(\delta_1, \delta_2)$ with

$$\delta_1 = \delta_2 = \frac{X_{12}}{\Lambda_{12} + \Lambda_{21}} = -\frac{5X_{12}}{8\pi}.$$

For $m = 3$ a diagonal solution $\delta = \operatorname{diag}(\delta_1, \delta_2, \delta_3)$ is given by

$$\begin{bmatrix}\delta_1\\ \delta_2\\ \delta_3\end{bmatrix} = \begin{bmatrix}\Lambda_{21} & \Lambda_{12} & 0\\ \Lambda_{31} & 0 & \Lambda_{13}\\ 0 & \Lambda_{32} & \Lambda_{23}\end{bmatrix}^{-1}\begin{bmatrix}X_{12}\\ X_{13}\\ X_{23}\end{bmatrix} = \frac{1}{\pi}\begin{bmatrix}\frac{12}{5} & -\frac{28}{9} & \frac{8}{9}\\ \frac{60}{7} & -\frac{80}{9} & \frac{160}{63}\\ 20 & -\frac{560}{27} & \frac{200}{27}\end{bmatrix}\begin{bmatrix}X_{12}\\ X_{13}\\ X_{23}\end{bmatrix}.$$

A diagonal solution trivially satisfies (5.27). We can not expect to be able to find a diagonal solution for $m \ge 4$ since we in this situation have m coefficients and $\frac{1}{2}m(m-1)$ equations to be solved but $\frac{1}{2}m(m-1) > m$ for $m \ge 4$. Whether there exists a solution to equations (5.26)-(5.27) for $m \ge 4$ we leave as an open question.

The underwater vehicle example from the last chapter does not satisfy assumptions 5.19-5.21. We have not been able to come up with any examples, where $m \ne n-1$, that satisfy assumption 5.19-5.21. These assumptions are simply still too strict. Thus the main argument of this section is that weakening the assumptions, in particular the $m = n - 1$ assumption, complicates the analysis considerably.

5.4 Examples

The usefulness of the theory is illustrated in the following examples.

Example 7 (Planar rigid body)**.** Reconsider the rigid body moving in the plane as described in the previous chapters. The configuration manifold is $G = SE(2) \simeq S\times\mathbb{R}^2 \ni (\theta, x, y)$. Let m denote the mass of the body, J its moment of inertia and h the distance from the center of mass to the control forces. The symmetric product is, for $\omega, \lambda \in \mathbb{R}$ and $v, w \in \mathbb{R}^2$, given by

$$\langle(\omega, v) : (\lambda, w)\rangle = \begin{bmatrix}0\\ \hat{\omega}w + \hat{\lambda}v\end{bmatrix},$$

where $\hat{\omega} = \left[\begin{smallmatrix}0 & -\omega\\ \omega & 0\end{smallmatrix}\right]$. With controls as in Figure 3.2 we have

$$b_1 = \frac{1}{m}e_2, \qquad b_2 = -\frac{h}{J}e_1 + \frac{1}{m}e_3,$$

which gives

$$\langle b_1 : b_1\rangle = 0, \qquad \langle b_2 : b_2\rangle = \frac{2h}{Jm}e_2, \qquad \langle b_1 : b_2\rangle = -\frac{h}{Jm}e_3.$$

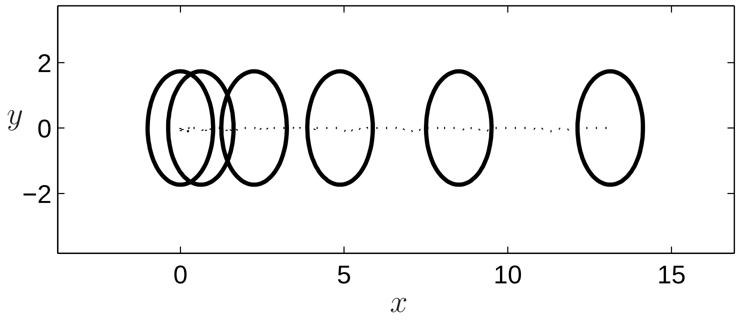

Figure 5.2: `speed_control` applied to the planar rigid body with $\xi_{\mathrm{re}} = e_3$, $\epsilon = 0.1$, and $\rho = 2$ and with initial conditions $(\theta, x, y)(0) = (-\frac{1}{2}\pi, 0, 0)$, $g_0 = g(0)$, and $(\omega, v_1, v_2)(0) = 0$. The dotted curve corresponds to the motion of the center of mass and the ellipses corresponds to the planar body at time equidistant instances.

Assumption 1 is immediately seen to be satisfied. It is straightforward to compute that

$$\langle e_3 : e_3 \rangle = 0,$$

so $\xi_{\mathrm{re}} = e_3$ is a relative equilibrium. Choosing this relative equilibrium we have

$$A_{\mathrm{re}} = \mathrm{ad}_{\xi_{\mathrm{re}}} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and we have $A_{\mathrm{re}} B = BQ$ and $\mathrm{ad}_{\xi_{\mathrm{re}}} B = BM$, $M = Q$, with

$$Q = \begin{bmatrix} 0 & -\frac{hm}{J} \\ 0 & 0 \end{bmatrix},$$

According to Remark 70 we calculate $\mathcal{A}_{\sigma,\alpha}$ as

$$\begin{aligned} \mathcal{A}_{\sigma,\alpha} &= B \int_0^{2\pi} e^{s\sigma M} \mathrm{diag}(\sin(\alpha_1 s), \ldots, \sin(\alpha_m s)) \mathrm{d}s \\ &= \begin{bmatrix} b_1 & b_2 \end{bmatrix} \sigma \begin{bmatrix} 0 & \frac{2\pi}{\alpha_2} \frac{hm}{J} \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Since $\mathcal{A}_{\sigma,\alpha}$ is independent of the frequencies α_1 and α_2 and the rank is constant for $\sigma \neq 0$ we may, according to remark 73, disregard Definition 68. Equation (5.4) is seen to reduce to

$$\mathcal{A}_{\sigma,\alpha}\gamma = -\mathcal{A}_{\sigma,(\beta,\beta)} \begin{bmatrix} \lambda_{12}(Z) \\ \lambda_{21}(Z) \end{bmatrix},$$

which is satisfied if

$$\gamma_2 = -\frac{\alpha_2 \lambda_{21}(Z)}{\beta},$$

and we can for example choose

$$\alpha_2 = 1, \qquad \beta = 3, \qquad \gamma_1 = 0.$$

The components of χ are computed

$$\begin{aligned}\sum_{i=1}^{2}\chi_i b_i &= \int_0^{2\pi}(e^{\sigma A_{\mathrm{re}}(2\pi-s)} - I)w_1^1(s)w_2^1(s)\mathrm{d}s\langle b_1 : b_2\rangle + \tfrac{1}{2}\sum_{i=1}^{2}\int_0^{2\pi} e^{\sigma A_{\mathrm{re}}(2\pi-s)}(w_i^1(s))^2\mathrm{d}s\langle b_i : b_i\rangle \\ &= \int_0^{2\pi}(e^{\sigma A_{\mathrm{re}}(2\pi-s)} - I)\left(-\frac{h}{Jm}\xi_{\mathrm{re}}\right)w_1^1(s)w_2^1(s)\mathrm{d}s + \tfrac{1}{2}\int_0^{2\pi} e^{\sigma A_{\mathrm{re}}(2\pi-s)}\left(\frac{2h}{Jm}e_2\right)(w_2^1(s))^2\mathrm{d}s \\ &= \frac{h}{Jm}\int_0^{2\pi}(w_2^1(s))^2\mathrm{d}s\ e_2 \\ &= \frac{h}{Jm}\int_0^{2\pi}(\lambda_{21}\sin(\beta a) + \gamma_2\sin(\alpha_2 s))^2\mathrm{d}s\ e_2 \\ &= \frac{h}{Jm}\pi(\lambda_{21}(Z)^2 + \gamma_2^2)e_2,\end{aligned}$$

meaning that we have

$$\chi_1 = \frac{\pi h(\lambda_{21}(Z)^2 + \gamma_2^2)}{J}, \qquad \chi_2 = 0.$$

Assumption 6 is immediately seen to be satisfied, so all the assumptions are met, and therefore we can apply the `speed_control` algorithm to speed up the system along e_3. The result of the `speed_control` algorithm applied to the planar rigid body can be seen in Figure 5.2 and 5.3. In the implementation we have chosen $\alpha_2 = 1$, $\beta = 3$, and $\gamma_1 = 0$.

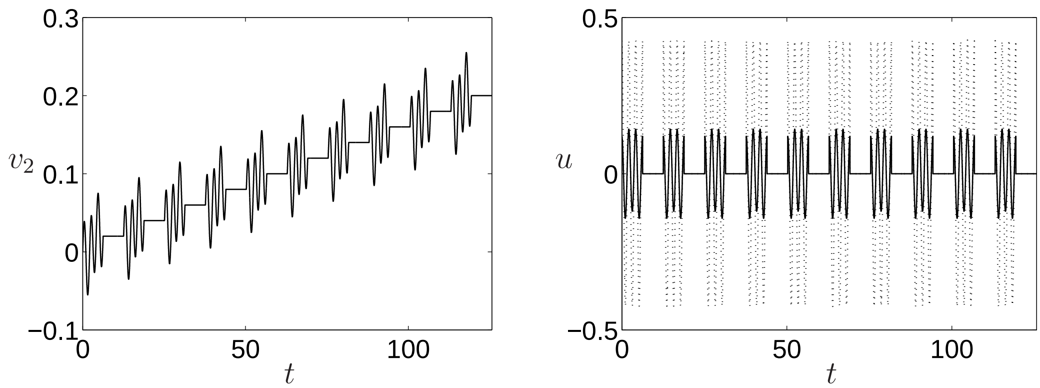

Figure 5.3: `speed_control` applied to the planar rigid body with $\xi_{\mathrm{re}} = e_3$, $\epsilon = 0.1$, and $\rho = 2$ and with initial conditions $(\theta, x, y)(0) = (-\frac{1}{2}\pi, 0, 0)$, $g_0 = g(0)$, and $(\omega, v_1, v_2)(0) = 0$. In the right figure the dashed curve corresponds to $u_1(t)$ and the solid curve corresponds to $u_2(t)$.

Example 8 (Satellite with two thrusters)**.** Consider a satellite with two thrusters aligned with the first and second principal axes. The configuration manifold is $G =$

$SO(3)$ and the equations of motion are of the form (5.1) and (5.2) where the symmetric product is given by

$$\langle \xi : \eta \rangle = \mathbb{I}^{-1}\big(\xi \times (\mathbb{I}\eta) + \eta \times (\mathbb{I}\xi)\big),$$

where $\mathbb{I} = \text{diag}(J_1, J_2, J_3)$, J_i being the moment of inertia along the ith principal axis, and $\times$ is the cross product. We have that

$$\langle e_3 : e_3 \rangle = 0,$$

so e_3 is a relative equilibrium. With controls as in figure 3.3 we have

$$b_1 = \frac{1}{J_1} e_1, \qquad b_2 = \frac{1}{J_2} e_2,$$

so it is not possible to directly control the motion in the e_3 direction. With $\xi_{\text{re}} = e_3$ we compute

$$A_{\text{re}} = \begin{bmatrix} 0 & \frac{J_2 - J_3}{J_1} & 0 \\ \frac{J_3 - J_1}{J_2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

It is straightforward to calculate that $A_{\text{re}} B = BQ$, with

$$Q = \begin{bmatrix} 0 & \frac{J_2 - J_3}{J_2} \\ \frac{J_3 - J_1}{J_1} & 0 \end{bmatrix},$$

so Assumption 2 is satisfied. We have that

$$\langle b_1 : b_1 \rangle = \langle b_2 : b_2 \rangle = 0, \qquad \langle b_1 : b_2 \rangle = \tfrac{J_2 - J_1}{J_1 J_2 J_3} e_3$$

we see that Assumption 1 is fulfilled if $J_1 \neq J_2$. Assumption 3 is satisfied because

$$\langle \xi_{\text{re}} : \langle b_1 : b_2 \rangle \rangle = \frac{J_2 - J_1}{J_1 J_2 J_3} \langle e_3 : e_3 \rangle = 0.$$

Since

$$\text{ad}_\xi \eta = \xi \times \eta$$

we see that

$$\text{ad}_{\xi_{\text{re}}} b_1 = \frac{J_2}{J_1} b_2, \qquad \text{ad}_{\xi_{\text{re}}} b_2 = -\frac{J_1}{J_2} b_1,$$

so Assumption 5 is satisfied and we have $\text{ad}_{\xi_{\text{re}}} B = BM$ with

$$M = \begin{bmatrix} 0 & -\frac{J_1}{J_2} \\ \frac{J_2}{J_1} & 0 \end{bmatrix}$$

Assumption 6 is immediately seen to be met.

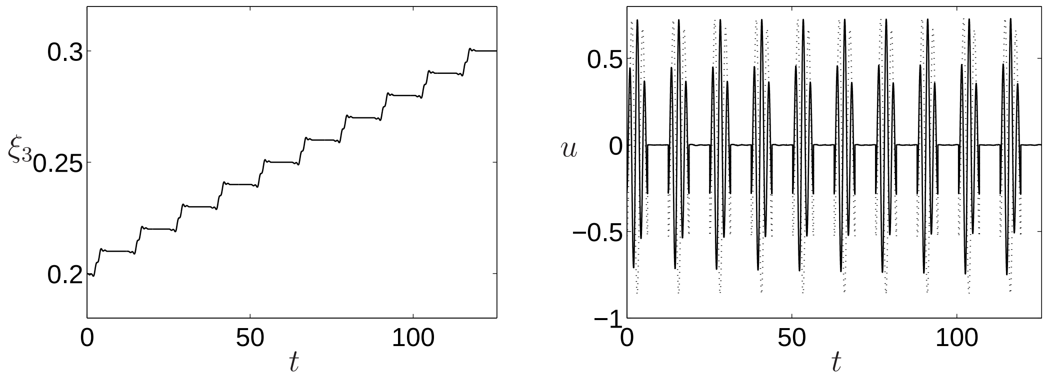

Figure 5.4: `speed_control` applied to the satellite with two thrusters with $\xi_{\mathrm{re}} = e_3$, $\epsilon = 0.1$, and $\rho = 1$ and with initial conditions $\xi(0) = (0, 0, 0.2)$ and $g_0 = g(0)$. In the right figure the dashed curve corresponds to $u_1(t)$ and the solid curve corresponds to $u_2(t)$

Using Remark 70 we calculate $\mathcal{A}_{\sigma,\alpha}$ as

$$\begin{aligned}
\mathcal{A}_{\sigma,\alpha} &= B\int_0^{2\pi} e^{s\sigma M}\mathrm{diag}(\sin(\alpha_1 s), \ldots, \sin(\alpha_m s))\mathrm{d}s \\
&= B\int_0^{2\pi} \begin{bmatrix} \cos(\sigma s) & -\frac{J_1}{J_2}\sin(\sigma s) \\ \frac{J_2}{J_1}\sin(\sigma s) & \cos(\sigma s)\end{bmatrix}\begin{bmatrix}\sin(\alpha_1 s) & 0 \\ 0 & \sin(\alpha_2 s)\end{bmatrix}\mathrm{d}s \\
&= B2\sin(\sigma\pi)\begin{bmatrix} \frac{\alpha_1\sin(\sigma\pi)}{\alpha_1^2-\sigma^2} & \frac{J_1}{J_2}\frac{\alpha_2\cos(\sigma\pi)}{\alpha_2^2-\sigma^2} \\ -\frac{J_2}{J_1}\frac{\alpha_1\cos(\sigma\pi)}{\alpha_1^2-\sigma^2} & \frac{\alpha_2\sin(\sigma\pi)}{\alpha_2^2-\sigma^2}\end{bmatrix} =: B\mathcal{H}_{\sigma,\alpha}.
\end{aligned}$$

Finding a solution to

$$\mathcal{A}_{\sigma,\alpha}\gamma = -\overline{\mathrm{Ad}_{\exp(s\sigma\xi_{\mathrm{re}})}(By(s))}(2\pi),$$

thus amounts to finding a solution to

$$\mathcal{H}_{\sigma,\alpha}\gamma = \mathcal{H}_{\sigma,(\beta,\beta)}\begin{bmatrix}\lambda_{12}(Z) \\ \lambda_{21}(Z)\end{bmatrix}.$$

Since the rank of $\mathcal{H}_{\sigma,\alpha}$ is not full when $\sigma \in \mathbb{N}$ special care is to be taken for these σ-values according to Definition 68. For $\sigma = \beta$ we have that $\mathcal{H}_{\sigma,(\beta,\beta)}$ has full rank but $\mathcal{H}_{\sigma,\alpha} = 0$ (since the α-values are different from β) so using Definition 68 this situation will be avoided. Since

$$\mathcal{H}_{\sigma,\alpha}^{-1}\mathcal{H}_{\sigma,(\beta,\beta)} = \begin{bmatrix}\frac{\beta(\alpha_1^2-\sigma^2)}{\alpha_1(\beta^2-\sigma^2)} & 0 \\ 0 & \frac{\beta(\alpha_2^2-\sigma^2)}{\alpha_2(\beta^2-\sigma^2)}\end{bmatrix},$$

we have, according to Remark 73, that we may in fact disregard Definition 68 and instead take

$$\gamma = -\mathcal{H}_{\sigma,\alpha}^{-1}\mathcal{H}_{\sigma,(\beta,\beta)}\begin{bmatrix}\lambda_{12}(Z) \\ \lambda_{21}(Z)\end{bmatrix} = -\begin{bmatrix}\frac{\beta(\alpha_1^2-\sigma^2)}{\alpha_1(\beta^2-\sigma^2)}\lambda_{12}(Z) \\ \frac{\beta(\alpha_2^2-\sigma^2)}{\alpha_2(\beta^2-\sigma^2)}\lambda_{21}(Z)\end{bmatrix},$$

as long as we ensure $\beta^2 - \sigma^2 \neq 0$ and that α_1, α_2, and β are different. In practical implementations we only need $\frac{\beta}{|\beta^2-\sigma^2|} > c$, for some $c > 0$. During `speed_inversion` we have

$$u^1(t) = -\sigma Q w^1(t) + \dot{w}^1(t), \qquad w^1(t) = \begin{bmatrix} \lambda_{12}(Z)\sin(\beta t) + \gamma_1 \sin(\alpha_1 t) \\ \lambda_{21}(Z)\sin(\beta t) + \gamma_2 \sin(\alpha_2 t) \end{bmatrix}$$

which gives

$$\begin{aligned} \|u^1\|_2^2 = \pi\Bigg(\sigma^2 \left(\left(\tfrac{J_3-J_1}{J_1}\right)^2 (\lambda_{12}(Z)^2 + \gamma_1^2) + \left(\tfrac{J_2-J_3}{J_2}\right)^2 (\lambda_{21}(Z)^2 + \gamma_2^2)\right) \\ + \gamma_1^2\alpha_1^2 + \gamma_2^2\alpha_2^2 + \lambda_{12}(Z)^2\beta^2 + \lambda_{21}(Z)^2\beta^2 \Bigg), \end{aligned}$$

where $\|\cdot\|_2$ is the norm on $L^2([0,2\pi],\mathbb{R}^2)$. We can thus in `speed_inversion` choose α_1, α_2, and β, all different, as to minimize $\|u^1\|_2$.

Calculating the value of χ gives

$$\begin{aligned} \sum_{i=1}^{2} \chi_i b_i &= \int_0^{2\pi} (e^{\sigma A_{\mathrm{re}}(2\pi - s)} - I) w_1^1(s) w_2^1(s) \mathrm{d}s \langle b_1 : b_2 \rangle + \tfrac{1}{2} \sum_{i=1}^{2} \int_0^{2\pi} e^{\sigma A_{\mathrm{re}}(2\pi - s)} (w_i^1(s))^2 \mathrm{d}s \langle b_i : b_i \rangle \\ &= \int_0^{2\pi} (e^{\sigma A_{\mathrm{re}}(2\pi - s)} - I) \left(\tfrac{J_2-J_1}{J_1 J_2 J_3} \xi_{\mathrm{re}}\right) w_1^1(s) w_2^1(s) \mathrm{d}s \\ &= 0, \end{aligned}$$

so we have $\chi = 0$. The result of the `speed_control` algorithm applied to this example can be seen in Figure 5.4. In the implementation giving this figure we have chosen $\alpha_1, \alpha_2, \beta \in \{1, \ldots, 5\}$, all different, as to minimize $\|u^1\|_2$.

Chapter 6

Conclusion

In this thesis we have focused on control of simple mechanical control system on Lie groups. In particular we have developed novel theory regarding local controllability along a relative equilibrium and constructed an algorithm capable of speeding up an invariant simple mechanical control system on a Lie group along a relative equilibrium.

In this chapter we will present a summary of this thesis and give some suggestions to future directions of research.

6.1 Summary of Dissertation

In Chapter 2 we presented, in a rigorous manner, the necessary theory of Lie groups needed to understand and analyse simple mechanical control systems on Lie groups. In particular we presented some of the classical matrix Lie groups describing rigid bodies.

Chapter 3 was concerned with the derivation of the equations of motion for forced mechanical systems, giving the forced Euler-Lagrange equations, and the equations of motion for mechanical systems with a Lie group as configuration manifold. Simple mechanical control systems is a special class of mechanical systems on Lie groups and the equations of motions for this class were deduced. The theory was applied to three example systems.

In Chapter 4 we focused on local controllability issues. We introduced some of the strongest theorems regarding local controllability properties of affine control systems. Previous controllability results for simple mechanical control systems on Lie groups were presented. Using the presented controllability theorems we derived a novel result giving sufficient conditions for a simple mechanical control system on a Lie group to be locally controllable along a relative equilibrium. This result is one of the main contributions of this thesis. The results were applied to give a controllability analysis of three example systems.

Chapter 5 was devoted to the construction of a novel motion algorithm. We presented the previous results regarding small amplitude control of simple mechanical systems in order to give a better understanding of the perspective of the new algorithm. Perturbation analysis was used to construct two inversion maps and a motion primitive was constructed as a composition of these maps. This motion primitive is the basis of the constructed algorithm. Computing small-amplitude control forces, this motion algorithm is capable of speeding up a simple mechanical control system on a Lie group along a relative equilibrium. The content of this chapter is the second main contribution of this thesis and will be published in [35].

6.2 Future Directions

In the following we will give some suggestions to future research related to the main contributions of this dissertation.

Local Controllability along a Relative Equilibrium

The novel result of chapter 4 gives sufficient conditions for a simple mechanical system on a Lie group to be locally controllable along a relative equilibrium. The concept of a relative equilibrium can be extended to simple mechanical control systems; these are mechanical systems for which the kinetic energy is given by a Riemannian metric and the Lagrangian is the kinetic energy minus the potential energy (which is a function of the configuration only), see, e.g., [13]. The results of [29] are controllability results for zero initial velocity (and therefore build upon the work of [41]) for simple mechanical control systems and these results are then in [11] applied to simple mechanical control systems on Lie groups. In this thesis we have worked the other way giving the local controllability result based directly on the theory of [5]. It would be interesting to investigate general results regarding local controllability along a relative equilibrium for a simple mechanical control systems and hopefully obtain the result of chapter 4 as a corollary of a more general result.

Motion Algorithms for Mechanical System

One of the main disadvantages of the new motion algorithm presented in chapter 5 are the strict assumptions. In particular the assumption that the number of independent control forces is $n-1$ where n is the number of degrees of freedom for the system. As illustrated in chapter 5 removing this particular condition complicates the analysis considerably. An important feature of the proposed motion algorithm is that everything is given explicitly meaning that it can be implemented in real time. It would be interesting to examine the possibility of an implicit method not needing the $n-1$ assumption. One might hope that an implicit method utilizing the geometric structure of simple mechanical control systems on Lie groups would produce a fast and efficient means to circumvent the problem.

Another interesting challenge would be to generalize the results of chapter 5 to more general mechanical systems. In the vein the work in [11] is generalized to a bigger class of systems in [32], one could attempt a similar generalization of the results in chapter 5.

Bibliography

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison-Wesley, Reading, MA, 2 edition, 1987.

[2] R. Abraham, J. E. Marsden, and T. S. Ratiu. *Manifolds, Tensor Analysis, and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer Verlag, New York, 2 edition, 1988.

[3] A. A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer Verlag, New York, 2004.

[4] V. I. Arnol'd. *Mathematical Methods of Classical Mechanics*, volume 60 of *GTM*. Springer Verlag, New York, 2 edition, 1989.

[5] R. M. Bianchini and G. Stefani. Graded approximations and controllability along a trajectory. *SIAM Journal on Control and Optimization*, 28(4):903–924, 1990.

[6] A. M. Bloch. *Nonholonomic Mechanics and Control*. Number 24 in Interdisciplinary Texts in Mathematics. Springer Verlag, 2003.

[7] A. M. Bloch, N. E. Leonard, and J. E. Marsden. Controlled Lagrangians and the stabilization of mechanical systems. I. The first matching theorem. *IEEE Transactions on Automatic Control*, 45(12):2253–2270, 2000.

[8] W. M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, New York, 2 revised edition, 2002.

[9] R. W. Brockett. Control theory and analytical mechanics. In C. Martin and R. Hermann, editors, *The 1976 Ames Research Center (NASA) Conference on Geometric Control Theory (Moffett Field, CA)*, pages 1–48. Math Sci Press, Brookline, MA, 1977.

[10] R. W. Brockett. Feedback invariants for nonlinear systems. In *Proceedings of the International Congress of Mathematicians*, Helsinki, Finland, 1978.

[11] F. Bullo, N. E. Leonard, and A. D. Lewis. Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups. *IEEE Transactions on Automatic Control*, 45(8):1437–1454, 2000.

[12] F. Bullo and A. D. Lewis. Kinematic controllability and motion planning for the snakeboard. *IEEE Transactions on Robotics and Automation*, 19(3):494–498, 2003.

[13] F. Bullo and A. D. Lewis. *Geometric Control of Mechanical Systems*, volume 49 of *Texts in Applied Mathematics*. Springer Verlag, New York, 2004.

[14] F. Bullo and K. M. Lynch. Kinematic controllability and decoupled trajectory planning for underactuated mechanical systems. In *IEEE Int. Conf. on Robotics and Automation*, pages 3300–3307, Seoul, Korea, April 2001.

[15] E. Coddington and N. Levinson. *Theory of Ordinary Differential Equations*. McGraw-Hill, 1955.

[16] H. Goldstein. *Classical Mechanics*. Addison-Wesley, Reading, MA, 2 edition, 1980.

[17] L. Gurvits. Averaging approach to nonholonomic motion planning. In *IEEE Int. Conf. on Robotics and Automation*, pages 2541–2546, Nice, France, April 1992.

[18] R. Hermann and A. J. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22:728–740, 1977.

[19] H. Hermes. Control systems which generate decomposable Lie algebras. *Journal of Differential Equations*, 44(2):166–187, 1982.

[20] D. Hristu. The dynamics of a forced sphere-plate system. *IEEE Transactions on Automatic Control*, 46(5):678–86, 2001.

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

[22] A. Isidori. *Nonlinear Control Systems II*. Communications and Control Engineering Series. Springer Verlag, New York, 1999.

[23] A. Isidori and A. J. Krener. On feedback equivalence of nonlinear systems. *Systems & Control Letters*, 2(2):118–121, 1982.

[24] J. Jost and X. Li-Jost. *Calculus of Variations*. Cambridge University Press, 1999.

[25] V. Jurdjevic. *Geometric Control Theory*. Cambridge University Press, Cambridge, UK, 1997.

[26] H. Lamb. *Hydrodynamics*. Dover Publications, New York, 6 edition, 1932.

[27] N. E. Leonard. Control synthesis and adaptation for an underactuated autonomous underwater vehicle. *IEEE Journal of Oceanic Engineering*, 20(9):211–220, 1995.

[28] N. E. Leonard and P. S. Krishnaprasad. Motion control of drift-free, left-invariant systems on Lie groups. *IEEE Transactions on Automatic Control*, 40(9):1539–1554, 1995.

[29] A. D. Lewis and R. M. Murray. Configuration controllability of simple mechanical control systems. *SIAM Journal on Control and Optimization*, 35(3):766–790, 1997.

[30] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics*, VII:649–673, 1954.

[31] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*. Springer Verlag, New York, 2 edition, 1999.

[32] S. Martínez and J. Cortés. Motion control algorithms for simple mechanical systems with symmetry. *Acta Applicandae Mathematicae*, 76(3):221–264, 2003.

[33] S. Martínez, J. Cortés, and F. Bullo. A catalog of inverse-kinematics planners for underactuated systems on matrix Lie groups. In *IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, pages 625–630, Las Vegas, NV, October 2003.

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

[35] N. Nordkvist and F. Bullo. Control algorithms along relative equilibria of underactuated Lagrangian systems on Lie groups. In *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, LA, December 2007. To appear.

[36] R. Ortega, A. J. van der Schaft, B. Maschke, and G. Escobar. Interconnection and damping assignment passivity-based control of port-controlled Hamiltonian systems. *Automatica*, 38(4):585–96, 2002.

[37] J. P. Ostrowski. Steering for a class of dynamic nonholonomic systems. *IEEE Transactions on Automatic Control*, 45(8):1492–1497, 2000.

[38] C. Rui, I. V. Kolmanovsky, P. J. McNally, and N. H. McClamroch. Attitude control of underactuated multibody spacecraft. In *IFAC World Congress*, San Francisco, CA, July 1996.

[39] S. S. Sastry. *Nonlinear Systems: Analysis, Stability and Control*. Interdisciplinary Applied Mathematics. Springer Verlag, New York, 1999.

[40] E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, volume 6 of *TAM*. Springer Verlag, New York, 2 edition, 1998.

[41] H. J. Sussmann. A general theorem on local controllability. *SIAM Journal on Control and Optimization*, 25(1):158–194, 1987.

[42] H. J. Sussmann and W. Liu. Lie bracket extensions and averaging: the single bracket case. In Z. Li and J. F. Canny, editors, *Nonholonomic Motion Planning*, pages 109–148. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[43] V. S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations*, volume 102 of *Graduate Texts in Mathematics*. Springer Verlag, New York, 1984.

[44] P. A. Vela. *Averaging and Control of Nonlinear Systems (with Application to Biomimetic Locomotion)*. PhD thesis, California Institute of Technology, Pasadena, CA, USA, May 2003.

[45] F. W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*, volume 94 of *GTM*. Springer Verlag, New York, 2 edition, 1989.