The Attitude Control Problem

Christopher Geyer

August 8, 2002

Abstract

§0 Introduction

In this paper we discuss an interesting problem encountered in control engineering called the *attitude control problem*. This problem attracted our interest because of recent work in visual servoing; because of its complexity further study was deemed necessary. Atttiude control is the task of manipulating the orientation of a body to a desired attitude. There are many applications in robotics where one would like to control the orientation of an end-effector or free flying object. Perhaps the biggest application of attitude control is in aerospace engineering where one would like to control the attitude of a satellite. Other applications even include submersible vehicles. Note that in this paper we are not interested in the stabilization of angular velocity, which is a slightly different though related problem.

For example in aerospace engineering the Hubble space telescope is an amazing example of the state of the art in attitude control. Designed to fix on distant stars, it is capable of locking on a target to within 2 millionths of a degree over a period of 24 hours. This would be equivalent to fixating on a strand of hair from a mile away. Such a feat, as in any control problem, is accomplished by precision actuation and measurement and a suitable control law. An extensive reference on spacecraft attitude control can be found in [1].

It is important to have some idea of the actuators affecting the attitude of an attitude control system. In satellites usually some combination of the following are used: magneto-torquers, jet thrusters or momentum wheels. Jet thrusters expell small amount of mass at a very high velocity by igniting fuel; positioned symmetrically, jet thrusters can be used control attitude without affecting translation. Momentum wheels are heavy wheels attached to electric motors; turning of the wheel in one direction causes the entire satellite to rotate in the reverse direction so as to conserve angular momentum. Magneto-torquers interact with the earth's graviational field by applying a current to a coil thereby inducing a torque. Planes, helicopters and submarines use flaps, ailerons or rudders to steer. These work by opposing the fluid or gas passing over them, causing a torque.

Equally important are the sensors used to measure the current attitude. In spacecraft control most attitude determination sensors utilize stars, the sun, or the earth to determine the attitude with respect to a coordinate system at the earth or the solar system. Many modern satellites use a wide field of view star tracker to measure the direction of the brightest stars and then apply a robust algorithm to match the measured positions against an on-board star catalog to determine attitude. This is simply an extension of the idea of navigation by stars that humans have used for many millennia. Closer to the ground we are not usually able to rely on stars and we cannot always rely on the sun. In mose cases though we are more interested in determining the orientation with respect to other nearby objects and not the solar system. This can be achieved using computer vision to reconstruct the structure of a scene and from there on estimate what is known as the egomotion and thence the attitude. As stated before, the impetus for this work grew out of designing algorithms for orienting a vehicle with respect to targets obtained using computer vision. If all else fails gyroscopes may be used as a mechanical integrator or dead reckoning system, though their accuracy decays over time and the presence of some initial measurement of attitude. Compasses also provide some information on attitude.

Looking more closely at the sensor examples, the direction of the sun or of the north pole given by a compass are two kinds of measurements which do not provide information about the angle of rotation in the direction of the reference point. As a consequence the state space as measured by such a sensor is not fully observable. This is one variation on the attitude control problem. Another possible variation would be consider the capabilities of a system where one actuator has failed, say one of the momentum wheels. We will consider both of these problems.

The motivation for this work came from a problem in computer vision. We supposed that a flying object used a camera to track a plane on the ground and that the goal was to orient the body frame so as to make the xy-plane parallel to the ground plane. Having devised an algorithm to determine the normal of the plane we required a method of controlling the object's attitude. This lead us to the current investigation into the vast attitude control literature.

The attitude control literture is indeed vast and this is understandable considering that it is not a new problem. The goal of this paper is really to tackle the subject from the ground up and from the point of view of a computer scientist. We therefore start at the very beginning with the physics of rigid body motion and end with examples of control laws. This review centers around three main papers which we think follow naturally. We consider first Jurdjevic and Sussmann's paper [2] which was one of the first to consider control systems on Lie groups after a suggestion by Brockett [3]. In this paper they demonstrate the conditions under which Lie systems are first order controllable. This is our motivation to go control law hunting. However, we are at risk of making impossible claims if we first do not consider that global control laws having a single equilibrium point do not exist on compact manifolds. This is demonstrated in the paper of Bhat and Bernstein [4] and which we prove here. After we are grounded in reality, having determined what is not possible, we consider examples of control laws proposed by Bullo and Murray [5] and their caveats.

What we plan to demonstrate in this paper are

- derivation of Euler's equations of motion;
- a fully contained contained proof that $SO(3)$ is a manifold and further properties of $SO(3)$;
- the differential geometry essential to understanding the phase space of the control problem;
- controllability of first order control systems on $SO(3)$;
- a proof that no first or second order control law with a globally asmyptotically stable equilibrium point exists on $SO(3)$; and
- examples of first and second order locally asmyptotically stable control laws on $SO(3)$.

We make no claim to cover all of the attitude control literature. That is not the purpose of this paper. The paper by Wen and Delgado [6] as well as the previously mentioned book [1] provides numerous references for and information on the attitude control problem.

§1 Rigid Body Motion

At a given instant the pose of a rigid body in space is determined by its position and orientation with respect to an inertial coordinate system. The configuration space is therefore the product of the space of 3-vectors p and the set of rotations R of space respectively specifying the translation and rotation between the inertial and body frames. In all of this paper we do not concern ourselves with the first component of this product; translation is not a relevant variable in attitude control and is ignored. Thus the *configuration space* of our system is the set of rigid rotations called $SO(3)$, to be defined later. We will consistently operate in the body frame so that a point R in this configuration space is the rotation to the inertial frame. In particular if x is vector specifying the direction to the sun then Rx is the direction to the sun in the inertial frame. Alternatively if y points to the sun in the inertial frame then $R^{-1}y$ points to the sun in the body frame.

As time progresses the state of the system will trace out a curve $R(t)$ in the configuration space. For small time intervals the evolution of $R(t)$ is primarily determined from the angular velocity. The angular velocity ω is a vector whose direction specifies the instantaneous axis of rotation in the body frame; its magnitude specifies the rate of rotation in radians per second. We will use the following notation, for any x define

$$\hat{\omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

This matrix has the property that for all a and b that $a \times b = \hat{a}b$.

With this notation let us calculate the rate of change of the matrix R in terms of the body angular velocity. The rotation between the attitudes at times t and $t + \delta t$ is

$$R^{-1}(t)R(t + \delta t) = I + \delta t\, \hat{\omega} + O(\delta t^2).$$

We can manipulate this equation into a difference of rotations on one side, divide both sides by δt and then let $\delta t \to 0$ to find that

$$\dot{R} = R\hat{\omega}. \tag{1}$$

If we are given the curve $\omega(t)$ and an initial condition $R(0)$ we can integrate to find the unique curve $R(t)$ satisfying (1) for all t.

Angular velocity has properties unlike its linear counterpart. Whereas in the absence of forces the velocity of a particle remains constant this is not true of angular velocity. Instead, the appropriate invariant common to both kinds of motion is momentum. Angular momentum is the vector

$$L = \mathcal{I}\,\omega.$$

The 3×3 matrix $\mathcal{I}$ is called the *inertia tensor* defined by,

$$\mathcal{I} \;=\; \int \rho(p)\, \hat{p}^2\, dp$$

where the integration is over the extent of the body and ρ is the body's mass density which might be measured in kilograms per cubic centimeter for example. The inertia tensor is peculiar in that it is dependent on the choice of coordinate system. In particular if two frames differ by a rotation R then in the second frame the inertia tensor is

$$\mathcal{I}' \;=\; R\,\mathcal{I}\,R^T\,,$$

where $\mathcal{I}$ is the inertia tensor in the first frame. Since $\mathcal{I}$ is symmetric, a rotation can be found for which $\mathcal{I}'$ is diagonal. The axes of the frame in which $\mathcal{I}$ is diagonal and $\mathcal{I}_{11} \geq \mathcal{I}_{22} \geq \mathcal{I}_{33}$ are called the *principal axes.*

If the body is not subjected to any torques then the angular momentum is conserved ($\dot{L} = 0$) otherwise the change in angular momentum is proportional to the torque:

$$\frac{dL}{dt} \;=\; \tau\,. \tag{2}$$

From this law we can derive Euler's equations governing the evolution of $\omega(t)$. In calculating the difference between the angular momentums at two points in time we must be careful to use the appropriate coordinate system to evaluate the inertia tensor. To begin with we have

$$L(t) \;=\; \mathcal{I}\,\omega(t)\,.$$

To evaluate the angular momentum at $t + \delta t$ in the body frame at time t we must consider that the body has rotated by $R^{-1}(t)R(t+\delta t) \approx (I + \delta t\,\hat{\omega})$. The angular momentum is therefore

$$\begin{aligned} L(t+\delta t) \;&=\; \Big(I + \delta t\,\widehat{\omega(t)} + O(\delta t^2)\Big)\,\mathcal{I}\,\Big(I - \delta t\,\widehat{\omega(t)} + O(\delta t^2)\Big)\,\omega(t+\delta t) \\ &=\; \mathcal{I}\,\omega(t+\delta t) \;+\; \delta t\,\widehat{\omega(t)}\,\mathcal{I}\,\omega(t+\delta t) \;-\; \delta t\,\mathcal{I}\,\widehat{\omega(t)}\,\omega(t+\delta t) \;+\; O(\delta t^2)\,. \end{aligned}$$

Thus,

$$\frac{dL}{dt} \;=\; \lim_{\delta t \to 0} \frac{L(t+\delta t) - L(t)}{\delta t} \;=\; \mathcal{I}\,\dot{\omega} - \omega \times \mathcal{I}\,\omega$$

Substitute dL/dt into equation (2) thereby obtaining Euler's equations in vector form:

$$\mathcal{I}\,\dot{\omega} - \omega \times \mathcal{I}\,\omega \;=\; \tau\,. \tag{3}$$

When there are no inputs to the system (torques) there exist points of equilibrium of equation (3), i.e. those points where $\omega \times \mathcal{I}\omega = 0$ so that $\dot{\omega} = 0$. If $\mathcal{I}\omega = \lambda\omega$ then the cross product is zero, thus any equilibrium point ω ought to be an eigenvector of $\mathcal{I}$. Since $\mathcal{I}$ is symmetric, the existence of three orthogonal real eigenvectors is guaranteed, then the equilibria lie in the subspaces parallel to the eigenvectors. If some of the eigenvalues are equal then the set of equilibria will consist of an entire plane ($\lambda_1 = \lambda_2$) or all of $\mathbb{R}^3$ ($\lambda_1 = \lambda_2 = \lambda_3$). Curiously not all equilibria are created equal, for example in the case where the eigenvalues of $\mathcal{I}$ are distinct then the eigenspaces of the smaller and larger eigenvalues are all stable equilibria; on the other hand the eigenspace of the middle eigenvalue only contains unstable equilibria.

There exists another constant of motion besides the angular momentum and that is the *angular kinetic energy* defined as

$$K \;=\; \frac{1}{2}\omega^T \mathcal{I}\,\omega\,. \tag{4}$$

Like the kinetic energy in translational motion it express the minimum amount of effort required to bring the object to rest.

§2 State and Measurement Equations

The curves $R(t)$ and $\omega(t)$ live in the *phase space* of the system in which a point is of the form $x = (R, \omega)$. In this phase space the differential equations (1) and (3) are combined to form one system of first order nonlinear differential equations:

$$\dot{x} \;=\; f(x,\tau) \quad \text{or} \quad \begin{pmatrix} \dot{R} \\ \dot{\omega} \end{pmatrix} \;=\; \begin{pmatrix} R\hat{\omega} \\ \mathcal{I}^{-1}(\tau - \omega \times \mathcal{I}\,\omega) \end{pmatrix}\,. \tag{5}$$

In addition to the state equation we need to make some assumption of what information the measurement process gives. If the sensor is able to determine the attitude of the body with respect to a fixed inertial frame then the measurement equation of such a system would be,

$$y(t) \;=\; R(t)\,. \tag{6}$$

This corresponds to the case when the system is fully observable. A star tracker would be able to provide such a measurement. Some sensors, however, are only able to report the direction in which a fixed target lies; the angle of rotation about this direction cannot be measured. A sun sensor would be an example of such a sensor; only the direction of the sun can be measured not any rotation about the axis in the direction of the sun. If the object in question is infinitely far away in the direction of the unit vector $z \in \mathbb{R}^3$ in an inertial frame then the measurement equation is

$$y(t) \quad = \quad R(t)z\,. \tag{7}$$

The observation space is therefore the sphere. This detail is to emphasize that the distance to z (anyway irrelevant to attitude control) is infinite or unknown, and more imporantly that the observation space is of smaller dimension than the configuration space. This follows because S^2 is a 2-manifold, whereas the set of rotations will be shown to be a 3-manifold.

We are now equipped with the knowledge of how the state evolves as well as two models of a measurement process; what is the goal of the control process? We are interested in the following problems.

The fundamental problem is the ability to drive the system from one state to another. We will consider first and second order control systems.

Problem 1. Is there a first order control law ω such that any initial configuration R_0 is driven to a desired configuration R_d? asymptotically? globally?

Problem 2. Is there a second order control law τ such that any initial configuration $(R_0, 0)$ (or maybe (R_0, ω)) is driven to a desired configuration $(R_d, 0)$? asymptotically? exponentially? globally?

Problem 3. What are the reachable states of first and second order control systems? What about in cases of reduced control when only two or fewer dimensions can be actuated?

Problem 4. What are the observable states of a reduced measurement process?

The next section is devoted to a detailed look at the group of rotations, the configuration space of the control system.

§3 The Special Orthogonal Group

In the previous section we stated without comment that the state space of the system is the set of orientation preserving rotations. In this section we explore this consequence further.

Consider a cube rotating in space with the letter **G** painted on one face. A rotation of this cube preserves the distance between any two points on the cube and the orientation of the painted letter so that a **⅁** never appears on the face. All such transformations can be modelled by a set of 3 × 3 matrices $\{R\}$. The first condition, that distances are preserved is satisfied if and only if $RR^T = I$. The preservation of the orientation of the letter **G** is equivalent to the condition that the right-hand rule of the cross product be preserved, i.e. for any u, v we must have $Ru \times Rv = R(u \times v)$. This, in turn, is satisfied if and only if $\det R = 1$. These two conditions formally define the set of rigid rotations; we also formally define the set of all rotations.

Definition 5. The orthogonal group of $\mathbb{R}^3$ is

$$O(3) \quad \equiv \quad \{R \in \mathbb{R}^{3\times 3} \,:\, RR^T = I\}\,.$$

The special orthogonal group of $\mathbb{R}^3$ is defined to be

$$SO(3) \quad \equiv \quad \{R \in \mathbb{R}^{3\times 3} \,:\, RR^T = I \text{ and } \det R = +1\}\,.$$

Not just any set is a group, it must earn that badge by satisfying the following conditions:

1. it must have an associative composition law closed in the group, in this case matrix multiplication is already associative and $(AB)(AB)^T = A(BB^T)A^T$ $= AA^T = I$ so that both $O(3)$ and $SO(3)$ are closed under matrix multiplication;
2. it must contain an identity element whose composition on the left or right with any element yields the original element; here this is satisfied by the identity matrix contained in both $O(3)$ and $SO(3)$; and
3. for every element there exists an inverse whose composition on the left or right with the original element yields the identity and this is true of matrix inversion under which both $O(3)$ and $SO(3)$ are closed.

These concepts are consistent with the fact that rotations can be reversed, composed associatively, and that no rotation is still a rotation.

Historically groups arose in the study of the symmetry of objects. For example, an equilateral triangle has three axes of symmetry. Assuming the vertices are labelled, a reflection about one of these axes swaps the vertices of one edge. Compositions of these reflections generate all six permutations of the vertices; these permutations form a group satisfying the rules laid out above. This is called the *symmetry group* of the triangle. It is finite and clearly any object with finite symmetries has a finite symmetry group.

What of the sphere? Any plane through its center is a plane of symmetry. Hence the compositions of all reflections is the symmetry group of the sphere

— in fact equal to the set of all (not necessarily rigid) rotations! What kind of cardinality does this group have? Since it is at least as numerous as the reflections which are in one-to-two correspondence with the sphere, the set of rotations is uncountably infinite[1]. Furthermore, neighborhoods of points in $O(3)$ or $SO(3)$ look locally like neighborhoods of $\mathbb{R}^3$ even though they are a subset of $\mathbb{R}^9$. We now define what we mean by a manifold; we will only be concerned with manifolds that are subsets of $\mathbb{R}^n$.

Definition 6. A function $f : X \to Y$ is known as a *diffeomorphism* if it is everywhere differentiable (we will almost always assume infinitely differentiable) and its inverse f^{-1} exists and is also differentiable everywhere. If a diffeomorphism exists between spaces then they are *diffeomorphic*. An *n-manifold* is a set M of points such that every point $x \in M$ is contained in open subset U of M which is diffeomorphic to some subset V of $\mathbb{R}^n$.

Given this definition we will now show that $O(3)$ is a manifold.

Proposition 7. The group $O(3)$ is a compact 3-manifold.

Proofs of this fact typically rely on a theorem called in Guillemin's book [7] the "preimage theorem" which states that the inverse image of a regular value is a manifold. We hope to gain greater understanding by proving the specialization of that theorem to this case. We appeal only to the inverse function theorem (a function f with non-singular Jacobian at x is a local diffeomorphism in neighborhoods of x and $f(x)$).

Proof: Note that if $f(R) = RR^T$ then $f^{-1}(I) = O(3)$. We would like to apply the inverse function theorem to show that a local inverse of f provides a local map for $SO(3)$. However, *df* is singular because $f(R)$ is always symmetric and thus a derivative evaluated in the direction of an anti-symmetric matrix must be zero. Solution: add a component to f which leaves the symmetric part unchanged and affects only the anti-symmetric part; then it becomes locally bijective and we can apply the inverse function theorem. Let

$$g(R) \quad = \quad RR^T + R - R^T .$$

We claim that

$$g^{-1}\{I + \mathcal{A}\} \;=\; O(3) \quad \text{where } \mathcal{A} \;=\; \{A : A + A^T = 0\} .$$

---

[1] What are examples of *countable* symmetry groups? The analog to the triangle would be the image of $e^{i\mathbb{Q}}$, i.e. the unit complex numbers with a rational argument which has a countable symmetry group; tilings have countable symmetry groups isomorphic to addition on $\mathbb{Z}$ or $\mathbb{Z} \times \mathbb{Z}$.

For if $R \in SO(3)$ then $g(R) = I + R - R^T \in I + \mathcal{A}$. Contrariwise, if $g(R) = I + A$ for some $A \in \mathcal{A}$ then since $RR^T \in \mathcal{A}^\perp$ we must have $RR^T = I$ implying $R \in O(3)$. Furthermore the Jacobian of g at I is non-singular. Compute the directional derivative in the direction of a matrix A:

$$dg_I(A) \;=\; \tfrac{d}{dt} g(I + tX)\big|_{t=0} \;=\; 2X\,.$$

Indeed, if given a Y we wish to find an X such that $dg_I(X) = Y$ we need only choose $X = Y/2$. Thus by the inverse function theorem there exists in this case neighborhoods $U, V \subset \mathbb{R}^9$ of I such that a smooth local inverse $h : V \to U$ exists satisfying $h(g(X)) = X$ for all $X \in U$. Now take $U' = U \cap O(3)$ and $V' = V \cap \mathcal{A}$ which are respectively open in the relative topologies of $O(3)$ and $\mathcal{A}$ and contain I in the range and domain. Then $h|_{V'}$ is a local diffeomorphism of $O(3)$ and $\mathbb{R}^3$ at $I \in O(3)$.

The preceeding result applies only to $I \in O(3)$, now choose an arbitrary $R \in O(3)$. Note that a left translation $l_R(A) = RA$ is a diffeomorphism of the set of matrices into itself (its inverse exists and both are smooth). Therefore $l_R \cdot h|_{V'}$ is a local diffeomorphism at R in the neighborhood RU' of $O(3)$.

Finally, $O(3)$ it is compact because it is the inverse image of a closed set and its rows (or columns) must be unit vectors, therefore lying on S^2. Hence $O(3)$ is closed and a subset of the compact set $S^2 \times S^2 \times S^2$. ■

Now that we know that $O(3)$ is a manifold what are the properties of the familiar operations of composition and inversion in this group? First the binary composition $(R, S) \mapsto RS$ is smooth in both variables and if one of R or S is constant then the resulting function is a diffeomorphism. An inversion $R \mapsto R^{-1}$ is also a diffeomorphism. Any group which is a manifold and whose composition and inversion operators are smooth is called a *Lie group.* Other examples of Lie groups are $\mathbb{R}$ under addition and most matrix groups.

Looking back at the proof, the diffeomorphism between $O(3)$ and the subspace of symmetric matrices is essentially an oblique projection of R to the vector space of anti-symmetric matrices. As such it is not especially useful since it does not yield information about the group of rotations nor does it give hints to its topological structure. Consider an alternative parameterization, though of $SO(3)$ only, provided by Euler:

$$(a_1, a_2, a_3, a_4) \;\overset{f}{\longmapsto}\; \begin{pmatrix} a_1^2 + a_2^2 - a_3^2 - a_4^2 & 2(a_2a_3 + a_1a_4) & 2(a_2a_4 - a_1a_3) \\ 2(a_2a_3 - a_1a_4) & a_1^2 - a_2^2 + a_3^2 - a_4^2 & 2(a_3a_4 + a_1a_2) \\ 2(a_2a_4 + a_1a_3) & 2(a_3a_4 - a_1a_2) & a_1^2 - a_2^2 - a_3^2 + a_4^2 \end{pmatrix}$$

as long as $a \in S^3$, i.e. $a_1^2+a_2^2+a_3^2+a_4^2 = 1$. This map is smooth and it is a double covering in the sense that there is only one $y \neq x$ for which $f(x) = f(y)$ and that is $y = -x$. This implies that $SO(3)$ is topologically equivalent to $S^3/x \sim -x$, that is S^3 divided by the equivalence relation that antipodal points are equal, this is in turn equal to projective space $\mathbb{P}^3$. Furthermore we have been hiding all this time that f is a closet homomorphism since S^3 is itself equal to the subgroup of unit quaternions where for example $q = a_1 + a_2\,i + a_3\,j + a_4\,k$ and $f(q \cdot r) = f(q)f(r)$. Moreover this equivalence shows that $SO(3)$ is connected since S^3 is connected and the continuous image of a connected set is connected. We have thus far given sufficient justification that:

Proposition 8. $SO(3)$ is a connected subgroup of $O(3)$ diffeomorphic to $\mathbb{P}^3$ and is orientable.

An *orientable manifold* is a connected manifold satisfying the following condition. If there exists a covering of an n-manifold by homeomorphisms with R^n such that in the overlap of any two homeomorphisms φ_1 and φ_2 the Jacobian of $\varphi_1 \circ \varphi_2$ at any point has positive determinant then the manifold is orientable. The projective plane $\mathbb{P}^2$ is an example of on object which is not orientable. It can be shown that all Lie groups are orientable [8].

§4 Differential Geometry

In the last section we proved that $SO(3)$ is a three dimensional manifold. Therefore each point $R \in SO(3)$ has a *tangent space* containing the derivatives[2,3] of all curves through R. In other words

$$T_R SO(3) \quad = \quad \{\, \gamma'(0) \,:\, \gamma(0) = R \text{ and } \gamma : (-\epsilon, \epsilon) \to SO(3) \,\}\,.$$

We claim that this is a vector space. First, the constant curve has zero derivative. Second, any curve can be reparameterized to scale its derivative arbitrarily. Third, if we have curves γ_1 and γ_2 then $\gamma_3(t) = \gamma_1(t)R^T\gamma_2(t) \in SO(3)$ since $SO(3)$ is a group and its derivative is $\gamma_1'(0) + \gamma_2'(0)$. So it is closed under scaling and addition of vectors and it contains a zero vector, therefore it is a vector space. The dimension of the vector space is equal in dimension to that of the manifold, therefore three. In order to be able to tell the difference between an

---

[2] We will have no need to look at non-differentiable functions, we assume all functions are smooth, i.e. infinitely differentiable

[3] This definition is suitable to manifolds which are subsets of R^n; in an abstract setting it is necessary to define the tangent space as the set of directional derivatives of functions on the manifold.

element $A \in T_R SO(3)$ from an element $B \in T_S SO(3)$ we will write $[A]_R$ for the tangent vector A at R.

To calculate $T_R SO(3)$ explicitly note that a curve $\gamma(t) \in SO(3)$ must satisfy $\gamma(t)\gamma(t)^T = I$. This is an invariant that can be differentiated with respect to t and evaluated at $t = 0$:

$$\gamma'(0)\gamma(0)^T + \gamma(0)\gamma'(0)^T \;=\; 0 \;\text{ or }\; AR^T + RA^T \;=\; 0 \;\text{ where }\; \gamma'(0) \;=\; A\,.$$

We can then be more specific about $T_R SO(3)$:

$$T_R SO(3) \;=\; \{\, [A]_R \,:\, AR^T + RA^T = 0 \,\}\,.$$

For a given R, $A \in T_R SO(3)$ if and only if there exists an ω such that $A = R\hat{\omega}$; then $T_I SO(3)$ is the set of all anti-symmetric matrices $\hat{\omega}$, $\omega \in \mathbb{R}^3$. A basis for $T_R SO(3)$ is $R\hat{x}$, $R\hat{y}$ and $R\hat{z}$ where $\{x, y, z\}$ is the canonical basis for $\mathbb{R}^3$.

In this context, but for arbitrary manifolds, a differential df_x of $f : M \to N$ at x is then a linear map from $T_x M$ to $T_{f(x)}N$. In particular we define the notation[4]

$$df([a]_x) \;=\; \left[\frac{d}{dt} f(\gamma(t)) \Big|_{t=0} \right]_{f(x)}, \tag{8}$$

where $\gamma(0) = x$ and $\gamma'(0) = a$. For each x, $df(x, \cdot)$ is a linear map between the tangent spaces. If f is a diffeomorphism then df has an inverse at each f, for then $df^{-1}(f(x), \cdot) : T_{f(x)}M \to T_x M$ and $df^{-1}(x, \cdot) = d(f^{-1})(x, \cdot) = (df(x, \cdot))^{-1}$ for all x. In the latter equation we mean the inverse of the the matrix $df(x, \cdot)$ in some bases for $T_x M$ and $T_{f(x)}M$.

The tangent space exists at every point on the manifold but is everywhere different. The union over all R is called the *tangent bundle* defined as

$$T\,SO(3) \;=\; \{\, [v]_R \,:\, R \in SO(3) \;\text{ and }\; v \in T_R SO(3) \,\}\,.$$

Associated with the tangent bundle is the projection map $\pi : T\,SO(3) \to SO(3)$ which takes $[v]_R \mapsto R$. As another example the tangent bundle of the sphere, $T\,S^2$, would consist of the pairs $[v]_x$ with $x \in S^2$ and v tangent to S^2 at x. As applied to our problem, $T\,SO(3)$ consists of vectors $[\omega]_R$, where R is the orientation and ω is the angular velocity, equivalent to the phase vector $(R, \omega)^T$ used in (5).

Let us return to the differential equation in (1) restated here with initial condition:

$$\tfrac{dR(t)}{dt} \;=\; R(t)\,\hat{\omega} \;\text{ and }\; R(0) \;=\; R_0\,, \tag{9}$$

---

[4] Unfortunately this notation is not standard, but we felt it was necessary so as not to get lost in subscripts, superscripts, etc. such as in df_R.

where we assume that ω is constant. In the discussion of rigid body motion we never verified whether a solution to (9) is actually a curve in $SO(3)$. The reason why this is the case has to do with the tangent bundle. Let

$$X(R) = R\hat{\omega} \tag{10}$$

which is the right hand side of the differential equation. Given any $R \in SO(3)$, $X(R) \in T_R SO(3)$ by the property that $R\hat{\omega} \in T_R SO(3)$ for all ω. Therefore X is a map from $SO(3)$ to its tangent bundle as long as we observe the formality that $X(R) = [R\hat{\omega}]_R$ which normally will be implied from context. Any function $X : SO(3) \to T\,SO(3)$ and such that $\pi(X(R)) = R$ is known as a *vector field*. The second requirement ensures that $X(R) \neq (A, B)$ for some arbitrary A and $B \in T_A SO(3)$. If X is a vector field and a curve $R(t)$ satisfies[5]

$$\frac{dR(t)}{dt} = X(R(t))\,,$$

then $R(t)$ is called an *integral curve* of X. Any integral curve defined on all of $\mathbb{R}$ is called *complete* and it can be shown that all integral curves on compact Lie groups are complete. Since an integral curve of a vector field X is defined for each R, we can define the *flow* of X to be the map $F_X : \mathbb{R} \times SO(3) \to SO(3)$ such that $F_X(\cdot, x)$ is an integral curve for all x with $F_X(0, x) = x$ and $F_{X,t} = F_X(t, \cdot)$ is a diffeomorphism of M into itself whose inverse is $F_{X,-t}$.

In order to represent the entire system (5) in this framework we need to define a vector field *on* the tangent bundle. Such a vector field would have to have a range in the tangent bundle of the tangent bundle. Thus the motions of (5) are the integral curves of the vector field

$$X([\omega]_R) = [\hat{\omega}, \omega \times \mathcal{I}\omega]_{[\omega]_R}\,.$$

One interpretation of a vector in a tangent space is that it is the directional derivative of some function $f : M \to \mathbb{R}$. Then $(Xf)(R) = df(R, X(R))$ according to the notation given in (8). The derivative of a vector field can also be defined but it is necessary to to do so in terms of another vector field. In other words we would like to find the amount of change in a vector field Y in the direction of an integral curve of X, where both X and Y are on M. The term for this kind of derivative is the *Lie derivative* and is defined as

$$L_X Y(R) = \frac{d}{dt} dF_{X,t}^{-1}\big(Y(F_{X,t}(R))\big)\Big|_{t=0}\,.$$

---

[5] dR/dt is now shorthand for the the map from the tangent bundle of an interval to the tangent bundle of $SO(3)$ as given in (8).

Note that it is insufficient to use $d/dtY(F_{X,t}(R))|_{t=0}$ because for each t the vector $Y(F_{X,t}(R))$ lies in a different tangent space. This is the reason for $dF_{X,t}^{-1}$ which is an isomorphism from $T_{F_{X,t}(R)}M$ to T_RM. The Lie derivative is linear in X and Y and is itself a vector field. It is also anti-symmetric about its two arguments $L_XY = -L_XY$ and satisfies the Jacobi identity $L_XL_YZ + L_YL_XZ + L_ZL_XY = 0$. A vector space with a bilinear, anti-symmetric operator satisfying the Jacobi identity is known as a *Lie algebra* whose operator is a *Lie bracket* $[\cdot,\cdot]$ given in this case by

$$[X,Y] \quad = \quad L_XY\,.$$

Finally, a *Riemannian manifold* is a connected n-manifold M equipped with a smooth function $\Phi : M \to P$ where we use P to denote the set of n-dimensional positive definite and symmetric matrices, so that $\langle u,v\rangle_x = u^T\Phi(x)v$ is a dot product on T_xM. This defines a metric space on M in the following way:

$$d(p,q) \quad = \quad \inf_{\substack{\gamma:[0,1]\to M\\ \gamma(0)=p,\gamma(1)=q}} \int_0^1 \sqrt{\langle\gamma'(t),\gamma'(t)\rangle_{\gamma(t)}}\,dt\,, \qquad (11)$$

which is to say that the distance is to defined to be the infimum of the arclengths of all paths from p to q. One example of a dot product on $SO(3)$ is the one given by the angular kinetic energy (4).

This section is not summarized in succinct theorems so let us recap the highlights. Looking back, the key ideas and notation in this section are:

- T_xM, the tangent space of the manifold M at $x \in M$ equal to the set of derivatives of all curves through x;
- a function $f : M \to N$ has differential $df : TM \to TN$, if f is a diffeomorphism then df^{-1} is defined;
- TM, the tangent bundle equal to the set of all (x,v) with $x \in M$ and $v \in T_xM$ and $T\,SO(3)$ is the phase space of (5);
- $\pi : TM \to M$ the projection defined by $\pi(x,V) = x$;
- $X : M \to TM$ such that $\pi(X(x)) = x$ is a vector field; and
- any curve $x(t)$ such that $dx/dt = X(x)$ is an integral curve;
- the Lie derivative L_XY is the change in Y in the direction of X;
- a Riemannian manifold M is a connected manifold with a dot product defined on every tangent space and defines a metric on all of M.

This is hardly a complete account of differential geometry. Further information can be found in [8] which is a pleasure to read. In the next section we show that X defined in (10) is a rather special vector field.

§5 The Lie Algebra of a Lie Group

Let us pause for a moment before defining the Lie algebra naturally associated with a Lie group. We would like to find a specific formula for the flow F_X when X is the vector field for rigid motion defined in (10). The flow must satisfy the differential equation (9). For a linear differential equation one is almost always right to guesse the solution is in the form of an exponential. In this case we have a linear matrix differential equation, is this the right path? In analogy to the exponential map for real numbers define the exponential map for square matrices as

$$e^X \quad = \quad \sum_{k=0}^{\infty} \frac{X^i}{k!}\,,$$

which can be shown to converge for all matrices X. In general $\frac{d}{dt}e^{X(t)} \neq X'(t)e^{X(t)}$, however if $X(t) = tA$ then it is valid and $\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A$ as can be verified directly by differentiating the series. Therefore the solution to (9) is of the form

$$R(t) \quad = \quad R_0\, e^{t\hat{\omega}}\,.$$

One readily verifies that $\exp A^T = (\exp A)^T$ and furthermore if $AB = BA$ then $\exp A + B = \exp A\ \exp B$. These two results imply that $\exp\hat{\omega}\ (\exp\hat{\omega})^T$, which ought to yield I in order that $R(t)$ be a rotation, equals $\exp\hat{\omega} + \hat{\omega}^T = \exp 0 = I$. As a result the flow of the vector field (10) is

$$F(t, R) = Re^{t\hat{\omega}}\,. \tag{12}$$

Then F_t is indeed a diffeomorphism of $SO(3)$.

This follows from the fact that the composition operator of the group is a diffeomorphism if one of the elements is constant. We define the maps $l_S(R) = SR$ and $r_S(R) = RS$ respectively called *left* and *right translations.* By this former property both l_S and r_S are diffeomorphisms. Their derivatives are respectively

$$dl_S(R, A) = (SR, SA) \ \text{ and } \ dr_S(R, A) = (RS, AS)\,,$$

recalling that differentials are maps on the tangent bundle.

What is special about the vector field $X(R) = (R, R\hat{\omega})$ is that

$$dl_S \circ X \quad = \quad X \circ l_S\,, \tag{13}$$

which is to say that the vector field is completely determined by its value at any one point, in the case of $SO(3)$ we have $X(R) = RX(I)$ (in abbreviated notation). Any vector field satisfying (13) for all S is *left invariant.* Similarly if $dr_S \circ X = X \circ r_S$ for all S then X is *right invariant.* Since $X(I)$ must be in $T_I SO(3)$ we find that every left invariant vector field X is of the form $X(R) = R\hat{\omega}$ for some ω.

Any linear combination of left invariant vector field is left invariant and by the above comments we have that the set of left invariant vector fields is isomorphic to $T_I SO(3)$. Next we show that the left invariant vector fields, in addition to being a vector space, is a Lie algebra.

Proposition 9. If X and Y are left invariant vector fields on $SO(3)$ then $L_X Y$ is also a left invariant vector field.

Proof: Assume $X(R) = R\hat{\nu}$ and $Y(R) = \hat{\omega}$. Start by computing $L_X Y(R)$, finding at first that $dF_{X,t}^{-1} = dF_{X,-t}$ and from the computation of the flow that

$$dF_{X,-t}(R, A) = \left(Re^{-t\hat{\nu}}, Ae^{-t\hat{\omega}}\right) .$$

Therefore

$$\begin{aligned} L_X Y(R) &= \tfrac{d}{dt}\, dF_{X,-t}\left(Y(Re^{t\hat{\nu}})\right)\Big|_{t=0} = \tfrac{d}{dt}\, dF_{X,-t}\left(Re^{t\hat{\nu}}, Re^{t\hat{\nu}}\hat{\omega}\right)\Big|_{t=0} \\ &= \tfrac{d}{dt}\left(Re^{t\hat{\nu}}e^{-t\hat{\nu}}, Re^{t\hat{\nu}}\hat{\omega}e^{-t\hat{\nu}}\right)\Big|_{t=0} = (R, R\hat{\nu}\hat{\omega} - R\hat{\omega}\hat{\nu}) \\ &= \left(R, R\,\widehat{\nu \times \omega}\right) , \end{aligned}$$

which is left invariant. ■

This is a consequence of a general theorem that says that on any Lie group the space of left invariant vector fields is closed under Lie differentiation. We therefore make the following definition.

Definition 10. The Lie algebra of a Lie group G is the set $\mathfrak{g}$ of left invariant vector fields. In the case of $SO(3)$ we define

$$\mathfrak{so}(3) = \{\, X : X(R) = R\hat{\omega} \,\}$$

Because of the natural isomorphism $\varphi : \mathfrak{so}(3) \to T_I SO(3)$, where in particular $\varphi(X) = X(I)$, $\mathfrak{so}(3)$ is equal in dimension as a vector space to $T_I SO(3)$ and therefore also $SO(3)$. We can define via φ a Lie bracket on $T_I SO(3)$:

$$[\hat{\nu}, \hat{\omega}]_{T_I SO(3)} = \varphi\left([\varphi^{-1}(\nu), \varphi^{-1}(\omega)]\right) = \hat{\nu}\hat{\omega} - \hat{\omega}\hat{\nu} = \widehat{\nu \times \omega} .$$

making $T_I SO(3)$ a Lie algebra.

We end the section with the following proposition given without proof. The second statement is proven by Bullo and Murray [5].

Proposition 11.

1. The exponential map of $\hat{\omega}$ reduces to

$$e^{\theta\hat{\omega}} = I + \frac{\sin\theta}{\theta}\hat{\omega} + \frac{1-\cos\theta}{\theta^2}\hat{\omega}^2 \quad (\textit{Rodrigues' formula}) ,$$

   where $\theta = \|\omega\|$. It is surjective and its inverse, the logarithmic map, is defined for all R such that $\operatorname{tr} R \neq -1$ and reduces to

$$\log R = \frac{\theta}{\sin\theta}\left(R - R^T\right) \quad \text{where} \quad 2\cos\theta + 1 = \operatorname{tr} R .$$

   with the exception that $\log I = 0$. For all R with $\operatorname{tr} R \neq -1$, $\exp\log R = R$ and for all $\hat{\omega}$ with $\|\omega\| < \pi$, $\log\exp\hat{\omega} = \hat{\omega}$.

2. If $\theta(t) = \log R(t)$, $\operatorname{tr} R(t) \neq -1$ and $\omega(t) = R^{-1}(t)\dot{R}(t)$ then

$$\dot{\theta}(t) = A(\theta(t))\,\hat{\omega}(t) ,$$

   where

$$A(\theta) = I + \frac{1}{2}\hat{\theta} + \left(1 - \frac{\|\theta\|}{2}\cot\frac{\|\theta\|}{2}\right)\frac{\hat{\theta}}{\|\theta\|^2} .$$

Let us step back to take in what we have shown. The Lie algebra $\mathfrak{so}(3)$ is the set of left invariant vector fields on $SO(3)$. Left invariant vector fields can be parameterized by the angular velocity ω. There is a natural association with $\mathfrak{so}(3)$ and $T_I SO(3)$ since a left invariant vector field specified by ω will have an integral curve whose tangent at I is $\hat{\omega} \in T_I SO(3)$. The exponential map takes a left invariant vector field and gives in return the integral curve evaluated in unit time, i.e. $\exp X = \gamma_X(1)$ where γ_X is the integral curve of $X \in \mathfrak{so}(3)$. The logarithm returns the left invariant vector field whose integral curve evaluated at unit time is the given element of $SO(3)$. So the Lie algebra can have two functions: to specify a velocity or to specify an element of the group. For more information on Lie groups and their associated Lie algebra see [9, 10, 11] in increasing level of difficulty.

§6 Controllability of Lie Systems

The exponential coordinates of an element R of $SO(3)$, by which we mean $\log R$, have a distinct *disadvantage* over other representations such as quaternions and that is that there is no compact formula for a product in the representation. In particular we would like to find the Z such that $\exp X \, \exp Y = \exp Z$, which by the way must exist by the surjectivity of exp. Short of letting $Z = \log(\exp X \, \exp Y)$, there is a small time approximation of the product in exponential coordinates which is a truncation of the Campbell-Baker-Haussdorf formula applied to general Lie groups and their Lie algebras. For small t,

$$e^{tX} e^{tY} \;=\; e^{tX+tY+\frac{t^2}{2}[X,Y]+O(t^3)} \,.$$

This has immediate application to control because it tells us that the solution to the following differential equation discontinuous in time

$$R(0) \;=\; I \;\text{ and }\; \dot{R}(s) \;=\; \begin{cases} \hat{\nu} & 0 \le s < \sqrt{t} \\ \hat{\omega} & \sqrt{t} \le s < 2\sqrt{t} \\ -\hat{\nu} & 2\sqrt{t} \le s < 3\sqrt{t} \\ -\hat{\omega} & 3\sqrt{t} \le s \le 4\sqrt{t} \end{cases}$$

is approximated at time $s = 4t$ by

$$\begin{aligned} R(4t) &= e^{-\sqrt{t}\hat{\omega}} e^{-\sqrt{t}\hat{\nu}} e^{\sqrt{t}\hat{\omega}} e^{\sqrt{t}\hat{\nu}} \\ &= e^{-\sqrt{t}\hat{\omega}-\sqrt{t}\hat{\nu}+\frac{t}{2}[\hat{\omega},\hat{\nu}]+O(t^{3/2})} e^{\sqrt{t}\hat{\omega}+\sqrt{t}\hat{\nu}+\frac{t}{2}[\hat{\omega},\hat{\nu}]+O(t^{3/2})} \\ &= e^{t[\hat{\omega},\hat{\nu}]+O(t^{3/2})} \\ &\approx e^{t\,\widehat{\omega\times\nu}} \,. \end{aligned} \tag{14}$$

Let $X(R) = R\hat{\nu}$ and $Y(R) = R\hat{\omega}$. Thus despite the fact that the left invariant vector fields X and Y do not span $\mathfrak{so}(3)$, by appropriately switching between the vector fields the the whole space $\mathfrak{so}(3)$ can effectively be spanned. In particular the vector field $Z(R) = R\,\widehat{\omega\times\nu}$ is not in the span of X and Y and is the vector field of which (14) would be the integral curve.

We can therefore purposefully design systems with a basis of control vectors having dimension less than the dimension the Lie group, called an *underactuated system*, while still being able to steer to any element of the Lie group, thus reducing the number of components and therefore cost and weight of a vehicle. Furthermore, systems which are not underactuated by design, are robust to the failures of some subset of its actuators. For a general Lie group G, as long as the *Lie span*, i.e. the closure of a vector space under all combinations of Lie brackets, of the control basis is equal to the span of $\mathfrak{g}$ then any element of G can is reachable. One can generate vector fields in the Lie span by a concatenation of oscillatory moves in the control basis as demonstrated above. This point of view is taken in work by Leonard and Krishnaprasad [12].

Proposition 12. There exists a $T > 0$ and a measurable function $u : [0, T] \to \mathbb{R}^n$ such that

$$\dot{R}(t) \;=\; R(t) \sum_{i=1}^{n} u_i(t)\,\hat{\omega}_i\,,\;\; R(0) \;=\; R_0 \;\text{ and }\; R(T) \;=\; R_1\,, \tag{15}$$

for any R_0 and R_1 if and only if $\mathcal{L}(\hat{\omega}_i) = \mathfrak{so}(3)$.

Proof: Let

$$\mathcal{A} \;=\; \{\, R(T) \;:\; \text{for all measurable } u \,\}\,,$$

often called the *attainable set* or *reachable set.* We assume the result from [13] that the map from the the set of measurable functions to $\mathcal{A}$ is continuous,[6] and also that because the set of continuous functions is path connected then by this result so is $\mathcal{A}$.

We claim that regardless of $\{\hat{\omega}_i\}$, $H = \mathcal{A}$ is a group. First it is closed under composition, for if u is a control such that $R(T_1) = R_1$ and v is a control such that $R(T_2) = R_2$ then the control u followed in succession by v has endpoint $R(T_1 + T_2) = R_1 R_2$ which by construction is in H. If u is again a control such that $R(T) = R_1$ then the control $v(t) = -u(T - t)$ gives an integral curve having an endpoint $R(T) = R_1^{-1}$. Clearly for any control u, $R(0) = I \in H$. Hence H is a subgroup of $SO(3)$ and since it is path connected subgroup of a Lie group it is itself a Lie subgroup.

Now H is a path connected subgroup of $SO(3)$ and it equals $SO(3)$ if and only its Lie algebra $\mathfrak{h}$ equals $\mathfrak{so}(3)$. ■

Notice that at any given t, the right hand side of system (15) is linear combination of left-invariant vector fields. For this reason it is often called a *left-invariant control system.* In its current form (15) is *driftless*, however if added to the right hand side were a constant ω_0 it would be a system with *drift.*

The proof of second order controllability is beyond the scope of this paper, we state the result from Bullo et al. [14].

---

[6] This requires a topology on the set of measurable functions, which is beyond the scope of this paper; it is not unreasonable to assume the result.

Proposition 13. For any $R_0, R_1 \in SO(3)$ there exists a $T > 0$ and a measurable $u : [0,T] \to \mathbb{R}^2$ such that

$$\begin{aligned} \dot{R}(t) &= \hat{\omega}(t)\,, \ \ R(0) = R_0\,, \ \ R(T) = R_1, \\ \text{and } \ \mathcal{I}\dot{\omega}(t) &= \omega(t) \times \mathcal{I}\,\omega(t) + u_1(t)\,x + u_2(t)\,y \end{aligned}$$

where x and y are the first two basis elements of $\mathbb{R}^3$. The control vectors are therefore aligned with the first two principle axes.

While this section has claimed that it is possible to steer first and second order systems from one point in the configuration space to another, the next shows that there is no vector field which has a globally asmyptotically stable equilibrium point.

§7 Impossibility of Global Asymptotic Stability

The main result of this section is that there exists no equilibrium point which is globally asymptotically stable on $T\,SO(3)$. We show this by proving a more general statement, and that is that no manifold which can be decomposed into a product consisting of at least one compact manifold has a vector field having a single globally asmyptotically stable equilibrium point. Let M be a manifold and F the flow of some smooth vector field X. Using standard terminology [15] we mean a point x to be an equilibrium point if $F(t,x) = x$ for all t. An equilibrum point is stable in the sense of Lyapunov if for every open neighborhood U of x there exists an open neighborhood V such that if $x \in V$ then $F(t,x) \in U$ for all $t \geq 0$. An equilibrium x is globally asmyptotically stable if it is stable in the sense of Lyapunov and if for all x, $\lim_{t\to\infty} F(t,x) = x$.

To show that no globally asmyptotically stable control law exists we require some ideas from differential topology. First, functions $f : M \to N$ and $g : M \to N$, where M and N smooth manifolds, are called *homotopic* if there exists a *homotopy* $h : [0,1] \times M \to N$ which is continuous and is such that $h(0,x) = f(x)$ and $h(1,x) = g(x)$. A manifold M is *contractible* if the identity map $f(x) = x$ is homotopic to a constant map $g(x) = x_0$, the homotopy is then called a *contraction*. Euclidean space R^n is contractible for all n, take for example the homotopy $h(t,x) = (1-t)x$ from the identity to the constant function.

Mathematicians are never short of definitions and we need a few more. Given a function $f : M \to N$ with the dimensions of M and N equal, a *regular point* is a point at which *df* is non-singular. If $f^{-1}(y)$ contains only regular values (or is empty) then y is a *regular value*. If y is not a regular value then it is a *critical value*. One can show that if y is a regular value then $f^{-1}(y)$ is finite.

Proposition 14. No compact manifold M is contractible.

Proof: A standard result in differential topology states that if $f, g : M \to N$ (M compact and N connected) are homotopic and if y is a regular value of both f and g then $|f^{-1}(y)| \cong_2 |g^{-1}(y)|$, that is the cardinalities of the inverse images are congruent modulo 2 (see [16] for example). Choose a point $y, z \in M$ with $y \neq z$. Let $f(x) = x$ and $g(x) = z$. Then the y is a regular value of both f and g. We have

$$f^{-1}(y) = y \ \text{ and } \ g^{-1}(y) = \emptyset \, .$$

If a homotopy between f and g existed it would contradict that $|f^{-1}(y)| \not\cong_2 |g^{-1}(y)|$. Hence M cannot be contractible. ■

This is best illustrated by imagining a sphere covered with a balloon and attempting to pull the balloon into a single point which obviously cannot be accomplished without tearing the balloon. We show next that global asymptotic stability implies contractability.

Proposition 15. If on a manifold M there exists a vector field X having a globally asmymptotically stable equilibrium point then M is contractible.

The following proof is inspired from the proof in [4].

Proof: Assume X is a vector field which has a globally asmyptotically stable equilibrum point at x_0. Define the map

$$h(\tau, x) = F\left(\frac{\tau}{1-\tau}, x\right) \quad \text{and} \quad h(1, x) \ = \ x_0 \, .$$

Clearly h is continuous on $[0, 1) \times M$. However, it is not obvious that it is continuous on $[0, 1] \times M$.

To show that it is continuous we need to prove that for any open neighborhood $U \subset M$ of x there exists an open neighborhood $V \subset [0, 1] \times M$ such that if $(\tau, x) \in V$ then $h(\tau, x) \in U$. First, by the Lyapunov stability of the equilibrium point, there exists a V_1 such that $F(t, x) \in U$ for all $x \in V_1$ and $t > 0$. Second, by the asmyptotic convergence to x_0 for this V_1 there exists a T such that $F(t, x) \in V_1$ for all $t > T$. So let $V = (\frac{T}{1+T}, 1] \times V_1$ which is open in the subspace topology of $[0, 1] \times M$[7]. If $(\tau, x) \in V$ and $\tau < 1$ then $\frac{\tau}{1-\tau} > T$

---

[7] Since $V = \left((\frac{T}{1+T}, 1 + \epsilon) \times M\right) \cap \left((\frac{T}{1+T}, 1] \times M\right)$, i.e. is the intersection of an open set with the space.

implying that $x \in V_1$. Since $x \in V_1$, we have $F(\frac{\tau}{1-\tau}, x) \in U$. Otherwise if $\tau = 1$ then clearly $h(\tau, x) = x_0 \in U$. ■

Because $SO(3)$ is compact a number of corollaries immediately follow from this proposition. The second order system result follows because $T\,SO(3) \cong SO(3) \times \mathfrak{so}(3)$.

Corollary 16. A globally asmyptotically stable first order control law on $SO(3)$ does not exist.

Corollary 17. A globally asmyptotically stable second order control law on $T\,SO(3)$ does not exist.

Many systems have compact configuration spaces so this result has many applications. Any system whose configuration cosists of an angle or orientation is compact. For example it applies to the pendulum whose phase space is $T\,S^1 \cong S^1 \times \mathbb{R}$.

Corollary 18. A globally asmyptotically stable control law for the forced pendulum does not exist.

This does not preclude the existence of control laws with multiple equilibria of multiple types, i.e. stable and unstable, which are discontinuous on a set of measure zero, or which vary with time. The number of such equilibria is determined by the global topology of the manifold and these questions are answered in the study of differential topology.

§8 Proportional Derivative Control Laws

In this section we go about making $SO(3)$ a metric space and then develop control laws which drive the distance between current and desired states to zero. The dot product that we will define will give geodesics which are natural for first order control where the inertia tensor is irrelevant to velocity control, however for second order control some axes may be easier to rotate about and these geodesics will not correspond to paths using the least amount of work.

For $\hat{\nu}, \hat{\omega} \in T_I SO(3)$ let

$$\langle \hat{\nu}, \hat{\omega} \rangle \;=\; \tfrac{1}{2} \operatorname{tr} \hat{\nu}^T \hat{\omega} \;=\; \nu^T \omega\,.$$

Thus if we identity $T_I SO(3)$ with $\mathbb{R}^3$ we see that we have chosen the standard Euclidean product. This dot product extends naturally to all of $T\,SO(3)$, for if we choose for $R\hat{x}$, etc. to be the basis of each $T_R SO(3)$ we have

$$\langle\, [R\,\hat{\nu}]_R,\, [R\,\hat{\omega}]_R \,\rangle \;=\; \nu^T \omega\,. \tag{16}$$

Proposition 19.

1. Using the dot product (16) the distance between R_1 and R_2 in $SO(3)$ is

$$d(R_1, R_2) = \sqrt{\langle \log R_1^{-1} R_2, \log R_1^{-1} R_2 \rangle}$$

   and the geodesic curve between them is

$$\gamma(t) = R_1 e^{t \log R_1^{-1} R_2}$$

   as long as $\operatorname{tr} R_1^{-1} R_2 \neq -1$. If $\operatorname{tr} R_1^{-1} R_2 = -1$ then the unique contiuous extension of d gives $d(R_1, R_2) = 2\pi$.

2. Furthermore if $R(t) \in SO(3)$ then

$$\frac{d}{dt} d(I, R(t))^2 = 2 \langle \log R(t), \hat{\omega}(t) \rangle$$

Proof: We prove only the second statement, for the first we refer to [8]. Letting $\hat{\theta}(t) = \log R(t)$, we have

$$\begin{aligned} \frac{d}{dt} d(I, R(t))^2 &= \frac{d}{dt} \operatorname{tr} \left[\hat{\theta}(t)^T \hat{\theta}(t) \right] \\ &= 2\, \theta(t)^T A(\theta(t))\, \omega(t) \,. \end{aligned}$$

Here $A(x)$ denotes the tangent map such that $A(\theta(t))\omega(t) = \dot{\theta}(t)$. Using the formula stated in Proposition 11, we see that $\theta(t)$ is a left eigenvector of $A(\theta(t))$, so that $\theta(t)^T A(\theta(t)) = \theta(t)^T$. Therefore

$$\begin{aligned} \frac{d}{dt} d(I, R(t))^2 &= 2\, \theta(t)^T \omega(t) \\ &= 2 \langle \log R(t), \hat{\omega}(t) \rangle \,. \end{aligned}$$

■

A consequence of the second part is that we now know that the quantity $\frac{d}{dt} d(I, R(t))^2$ is maximized when $\hat{\omega}(t)$ is parallel to and in the same direction as $\log R(t)$. In other words the direction maximizing the rate at which the distance to the identity decreases is the direction of the tangent to the geodesic from the current point to the identity. This motivates a first order or velocity control law which is proportional to the error,

$$\hat{\omega}(t) = -k \log R(t) \,. \tag{17}$$

for some gain k.

Proposition 20. If $k > 0$ then the first order control law (17) is asymptotically stable for all initial conditions R with $\operatorname{tr} R \neq -1$.

Proof: Let $\theta(t) = \log R(t)$. The system then satisfies the differential equation

$$\begin{aligned}\dot{\theta}(t) &= -k\, A(\theta(t))\, \theta(t) \\ &= -k\, \theta(t)\,.\end{aligned}$$

The latter following because $\theta(t)$ is an eigenvector of $A(\theta(t))$. In fact this law is exponentially stable whenever $k > 0$. ■

For the second order system we choose a proportional plus derivative control law. We will assume that the system is fully actuated, in other words the torque $\tau(t) \in \mathbb{R}^3$ is unrestricted. Let

$$\tau \;=\; -\omega(t) \times \mathcal{I}\, \omega(t) - k_p\, \theta(t) - k_d\, \omega(t)\,, \tag{18}$$

where k_p is a scalar proportional gain, k_d is the derivative gain, and again $\hat{\theta}(t) = \log R(t)$. This control law is inefficient in that it tries to cancel the inertia acting on the body.

Proposition 21. If

$$\frac{1}{2}\|\theta(0)\|^2 + \lambda_{max}(\mathcal{I})\|\omega(0)\|^2 \leq \frac{1}{2}\pi^2 \tag{19}$$

where $\lambda_{max}(\mathcal{I})$ is the largest eigenvalue of $\mathcal{I}$ then the control law (18) is asmyptotically stable. In particular any initial configuration at rest with $\operatorname{tr} R \neq -1$ converges.

To prove this proposition we will apply a technique developed by a Russian engineer by the name of Lyapunov who in 1892 developed a theory to determine the stability of differential equations. The theory that follows is necessarily abbreviated and we refer to [15, 17] for a full description. Assume x_0 is an equilibrium point of a dynamical system $\dot{x} = f(x)$, i.e. $f(x_0) = 0$. Let $V(x)$ be a continuous function defined in a neighborhood U of x_0 for which $V(x_0) = 0$ but $V(x) > 0$ for all $x \neq x_0$, such a function is called *locally positive definite*; a function for which $V(x) < 0$ for all $x \neq x_0$ would be called *locally negative definite*. Let $\dot{V}(x) = (\nabla V)^T f(x)$, so that $\dot{V}$ measures the rate of change of V along an integral curve of f. The function V acts like an energy function, though which is not conserved because of the dissipation of energy by any derivative control law.

The theorem of Lyapunov states that if $\dot{V}$ is locally negative definite in the neighborhood U then x_0 is an asmyptotically stable equilibrium point. For within U there exists an open ball containing x_0 and on the surface of which the minimum value is m. Inside this sphere is a further smaller ball on which V never exceeds m; any integral curve starting here never leaves U since V evaluated at any point on the curve is bounded by m. Thus x_0 is stable in the sense of Lyapunov. The asmyptotic part of the stability goes as follows. Since V is bounded from below and the by the stability, the path of an integral curve $x(t)$ is bounded and thus $\lim_{t\to\infty} V(x(t)) = V_0$ exists. If $V_0 = 0$ then asmyptotic convergence is immediately implied. If, however, $V_0 \neq 0$ then $x(t)$ never enters a ball B about x_0. Because V is locally negative definite, outside of B but within U, $\dot{V}(x)$ must be no greater than $-v_1$, for some $v_1 > 0$. This is a contradiction because $\dot{V}(x(t)) < -v_1$ for all t implies that $V(x(t+h)) \leq V(x(t)) - hv_1$ for all $h > 0$. Now finally to prove the proposition.

Proof (of Proposition 21): If we let $\theta(t) = \log R(t)$ then the differential equations of the system are

$$\dot{\theta}(t) \;=\; A(\theta(t))\,\omega(t) \quad \text{and} \quad \dot{\omega}(t) \;=\; -k_p\,\theta(t) - k_d\,\omega(t)\,.$$

For a sufficiently small $\epsilon > 0$ and any positive gain k_p the function

$$V(\theta,\omega) = \frac{k_p}{2}\,\theta^T\theta + \frac{1}{2}\omega^T\mathcal{I}\,\omega + \epsilon\,\theta^T\mathcal{I}\,\omega\,,$$

is locally positive definite. The first term is analagous to a potential energy and the second is like a kinetic energy. The derivative

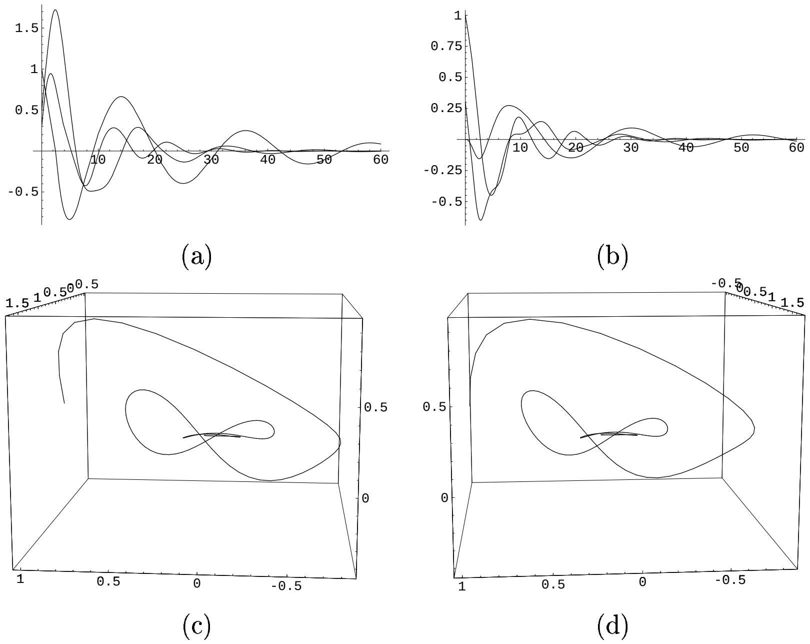
$$\begin{aligned}\dot{V}_\epsilon(\theta,\omega) \;&=\; k_p\,\theta^T\dot{\theta} + \omega^T\mathcal{I}\,\dot{\omega} + \epsilon\,\theta^T\mathcal{I}\dot{\omega} + \epsilon\dot{\theta}^T\mathcal{I}\omega \\ &=\; k_p\,\theta^T A(\theta)\omega + \omega^T(\omega\times\mathcal{I}\omega - k_p\,\theta - k_d\,\omega) + \\ &\quad\;\; \epsilon\,\theta^T(\omega\times\mathcal{I}\omega - k_p\,\theta - k_d\,\omega) + \epsilon\,\omega^T A(\theta)^T\mathcal{I}\,\omega\,.\end{aligned}$$

Recall $\theta A(\theta) = \theta$ and note that $\theta^T\hat{\omega} = -\omega\hat{\theta}$,

$$\begin{aligned}\dot{V}_\epsilon(\theta,\omega) \;&=\; k_p\,\theta^T\omega - k_p\omega^T\theta - k_d\omega^T\omega + \\ &\quad\;\; \epsilon\,\theta^T\hat{\omega}\mathcal{I}\omega - \epsilon k_p\theta^T\,\theta - \epsilon k_d\,\theta^T\omega + \epsilon\,\omega^T A(\theta)^T\mathcal{I}\,\omega \\ &=\; -k_d\omega^T\omega - \epsilon k_p\theta^T\,\theta - \epsilon\,\omega^T\hat{\theta}\mathcal{I}\omega - \epsilon k_d\,\theta^T\omega + \epsilon\,\omega^T A(\theta)^T\mathcal{I}\,\omega\,.\end{aligned}$$

We can still simplify because $-\hat{\theta}+A(\theta) = A(-\theta)$. Furthermore $\omega^T A(-\theta)^T\mathcal{I}\omega \leq \frac{\pi}{2}\omega^T\mathcal{I}\omega$. Then,

$$\begin{aligned}\dot{V}_\epsilon(\theta,\omega) \;&\leq\; -k_d\omega^T\omega - \epsilon k_p\theta^T\,\theta - \epsilon k_d\,\theta^T\omega + \epsilon\frac{\pi}{2}\,\omega^T\mathcal{I}\,\omega \\ &\leq\; -\begin{pmatrix}\theta & \omega\end{pmatrix}\begin{pmatrix}\epsilon\,k_p\,I & \epsilon\,k_d/2\,I \\ \epsilon\,k_d/2\,I & k_d\,I - \epsilon\frac{\pi}{2}\mathcal{I}\end{pmatrix}\begin{pmatrix}\theta \\ \omega\end{pmatrix}\,.\end{aligned}$$



(a) (b)

(c) (d)

Figure 1: (a) Plots of θ_x, θ_y and θ_z of an integral curve. (b) Plots of ω_x, ω_y and ω_z of an integral curve. (c), (d) Stereoscopic view of integral curves of the second order control law projected to exponential coordinates.

For sufficiently small ϵ, the matrix within is positive definite.

With the condition that the integral curve not meet the singularity of log (i.e. $\operatorname{tr} R(t) = -1$) we require that $\|\theta(t)\| < \pi$. Notice that $2V_0(\theta, \omega) \geq \|\theta(t)\|$, and that V_0 on any integral curve is non-decreasing. Therefore

$$\begin{aligned} \frac{1}{2}\|\theta(t)\|^2 &\leq \frac{1}{2}\|\theta(0)\|^2 + \omega(0)^T \mathcal{I} \omega(0) \\ &\leq \frac{1}{2}\|\theta(0)\|^2 + \lambda_{max}(\mathcal{I}) \|\omega(0)\| \,. \end{aligned}$$

And so if the right hand side is bounded by π^2 then under initial conditions satisfying this bound the control law will be asmyptotically stable. ■

Figure 1 shows the integral curve for the second order control law projected from $T\,SO(3)$ down to $SO(3)$ represented in exponential coordinates. The inertia matrix in this simulation is $\operatorname{diag}(6, 3, 1)$. The initial conditions satisfy (19).

§9 Observability in the Case of Reduced Measurement

In this final section we consider the consequences of choosing the reduced measurement equation (7) which we recall is

$$y(t) \quad = \quad R(t)z$$

where z is assumed to be some constant unit vector. It is at once clear that this is an ambiguous measurement equation, for if $y(t_0) = y_0$ when $R(t_0) = R_0$ then we obtain the same measurement for an alternative curve $R'(t) = R_0\, e^{\phi\hat{z}}$ for any ϕ. This follows because by Proposition 11, z is a left eigenvector of $e^{\phi\hat{z}}$.

Now let

$$H_z \quad = \quad \{\, R \in SO(3) \,:\, Rz = z \}$$

Because it is closed under multiplication and inversion and it contains the identity it is a subgroup of $SO(3)$. It is called the *isotropy group* of z. All elements of H_z are of the form $e^{\phi\hat{z}}$ and so $H_z \cong SO(2)$.

For any $R \in SO(3)$ the set

$$R\,H_z = \{\, R\,S \,:\, S \in H_z \,\}\,,$$

is called a *coset* of H_z. For any two R_1 and R_2 the cosets $R_1 H_z$ and $R_2 H_z$ are disjoint and thus the set of cosets partition $SO(3)$. This set of cosets,

$$SO(3)/H_z \quad = \quad \{RH_z \,:\, R \in SO(3)\}\,.$$

is called the *quotient space*. Under normal circumstances, if G is a group and H a subgroup, then G/H would be a group, but H_z is not normal. The isomorphism between H_z and $SO(2)$ implies that $SO(3)/H_z \cong SO(3)/SO(2)$. A topology is naturally induced on the quotient space where if U is an open subset of $SO(3)$ then the union of equivalence classes over elements in U is open in $SO(3)/SO(2)$.

Going back to the measurement equation, if $y(t_0) = y_0$ then $R(t_0) \in R_0 H_z$ where $R_0 z = y_0$ which always exists as long as y_0 and z are both unit vectors. Thus the ambiguity of the measurement made by y is best expressed when we re-express the measurement as an element of $SO(3)/H_z$, i.e. $y(t_0) \equiv R_0 H_z \in SO(3)/H_z$. Such an $R_0(t)$ exists for all t and if prior to our definition $y(t)$ was continuous then so is $y(t) \in SO(3)/SO(2)$.

Proposition 22. If $\operatorname{tr} R(0) \neq -1$ then for any z there exists a control law such that $y(t) \to H_z$ as $t \to \infty$, $R(t) \to R_0$ for some constant R_0, both exponentially.

Proof: Assume without loss of generality that $z = (0, 0, 1)$, if this is not so the inertial system can be rotate so that it is. Since H_z consists of $e^{\phi\hat{z}}$, in exponential coordinates this is a vertical line through the origin. Within the ball of radius π every coset RH_z with $\operatorname{tr} R \neq -1$ intersects the plane $z = 0$ exactly once. Thus the cosets with $\operatorname{tr} R \neq -1$ are in one-to-one correspondence with the points inside the disk of radius π in the plane $z = 0$. By rotating the coordinate system about the axis of z we can assume without loss of generality that $R(0)$ is on the plane $z = 0$. Then the control law of Proposition 20 applies and $R(t) \to I$ exponentially. As $R(t) \to I$, $y(t) \to H_z$ since H_z contains the origin. ∎

The story for second order control is slightly different. Though it is seems that the rotation about the measured axis is unobservable, is this true of the angular velocity? If the body is not symmetric then the angular velocity can be measured even from a reduced measurement because of the presence of the Coriolis force, i.e. the term $\omega \times \omega$ in Euler's equations. Thus if an intial estimate of the angle about the axis is available then the angular velocity can be integrated to find the angle for all future points in time.

§10 Conclusion

In this paper we have tried to give a self-contained presentation of the attitude control problem. Starting with Euler's equations we describe the state equations for rigid body motion. We then cover the differential geometry and theory of Lie groups necessary to understand both global and local properties of the group of rotations. The local properties demonstrate that under certain benign conditions systems defined on Lie groups are controllable. Global properties derived with the help of differential topology show us that global asymptotic stability is not possible for compact Lie groups. We then give examples of control laws and prove their local asmyptotic stability. We hope to have shown some of the many facets of the attitude control problem related to physics, Lie groups, differential geometry, differential topology and stability theory.

References

[1] J. R. Wertz, *Spacecraft Attitude Determination and Control.* D. Reidel, Dordrecht, Netherlands, 1978.

[2] V. Jurdjevic and H. Sussmann, “Control systems on Lie groups,” *J. Diff. Eqns.*, vol. 12, pp. 313–329, 1972.

[3] R. W. Brockett, “System theory on group manifolds and coset spaces,” *SIAM J. Control*, vol. 10, no. 2, pp. 265–284, 1972.

[4] S. Bhat and D. Bernstein, “A topological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon,” *Systems and Control Letters*, vol. 39, pp. 63–70, 2000.

[5] F. Bullo and R. M. Murray, “Proportional derivative (PD) control on the Euclidean group,” *CDS Technical Report 95-010, California Institute of Technology, Pasadena, CA*, 1995.

[6] J. Wen and K. Delgado, “The attitude control problem,” *IEEE Trans. on Automatic Control*, vol. 36, no. 10, pp. 1148–1162, 1991.

[7] V. Guillemin and A. Pollack, *Differential Topology.* Prentice-Hall, Englewood Cliffs, NJ, 1974.

[8] W. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry.* Academic Press, London, 1986.

[9] M. Curtis, *Matrix Groups.* Springer, New York, 1984.

[10] F. Warner, *Foundations of Differentiable Manifolds and Lie Groups.* Scott, Foresman and Company, Glenview, IL, 1970.

[11] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces.* AMS, Providence, RI, 1978.

[12] N. Leonard and P. Krishnaprasad, “Motion control of drift-free, left-invariant systems on Lie groups,” *IEEE Trans. on Automatic Control*, vol. 40, no. 9, pp. 1539–1554, 1995.

[13] V. Jurdjevic, *Geometric Control Theory.* Cambridge University Press, Cambridge, 1997.

[14] N. L. F. Bullo and A. Lewis, “Controllability and motion algorithms for underactuated lagrangian systems on Lie groups,” *IEEE Trans. on Automatic Control*, vol. 45, no. 8, pp. 1437–1454, 2000.

[15] S. Sastry, *Nonlinear Systems.* Springer Verlag, New York, 1999.

[16] J. Milnor, *Topology from the Differentiable Viewpoint.* Princeton University Press, Princeton, NJ, 1997.

[17] J. Slotine and W. Li, *Applied Nonlinear Control.* Prentice-Hall, Englewood Cliffs, NJ, 1991.