UNIVERSITY of CALIFORNIA
SANTA CRUZ

OPTIMAL CONTROL OF A TAEKWONDO KICK

A thesis submitted in partial satisfaction of the
requirements for the degree of

BACHELOR OF SCIENCE

in

PHYSICS

by

Aaron Palmer

31 May 2010

The thesis of Aaron Palmer is approved by:

Professor Debra Lewis Advisor	Professor David P. Belanger Senior Theses Coordinator

Professor David P. Belanger
Chair, Department of Physics

Abstract

Optimal Control of a Taekwondo Kick

by

Aaron Palmer

The tools of physics apply to a larger class of problems then the analysis of physical systems. Mathematical formalisms developed for classical mechanics are used to study a problem of optimal control theory. In the problem of an optimal roundhouse kick in Taekwondo, the Pontryagin Maximum Principle relates the optimal control problem to the framework of Hamiltonian mechanics. The theory of Lie groups provides a structured manifold to analyze the non-linear mechanics of the system of two joints that is represented by connected rotations. The simplified system of a front kick in a single plane is initially analyzed. This planar problem is represented by the manifold of a torus, which is the maximal torus of the manifold for the full problem. The manifold for the Taekwondo roundhouse kick is represented by the non-abelian Lie group $SO(3) \times SO(2)$. A 'diagonalization' of this problem projects it onto the torus and relates back to the planar problem.

A full analytic solution is not possible due to the difficulties of the non-linear system. Instead, methods of approximation, by reduction of the problem to finite dimensional optimization and numerical calculation of optimal trajectories, are used.

Contents

List of Figures

Acknowledgements

I want to thank Debra Lewis for advising me and supporting me on this project. Professor Dejan Milutinovic for broadening my perspective on control theory. The other professors of math and physics who have taught me all of the math I use. Grandmaster Soonho Song who taught me the martial arts that inspired this problem. And of course my family.

1

Introduction

Methods for the solution of problems of optimal control have been well developed to efficiently solve for systems that are approximated well by linear dynamics. In this thesis, I present a problem of optimal control of a mechanical system that is highly non-linear. The problem deals with a physical system that is intuitive from my training in Taekwondo and contains many of the elements of a general mechanical control problem. The theory of non-linear control is understood from a mathematical perspective and is directly related to a generalization of classical mechanics in physics. The ability to effectively solve such a problem of optimal control in order to be used in robotics is of importance to the development of more capable technology. Non-linear optimal control is not widely used in the current state of robotics. Instead more basic controllers are used, and if optimality is wanted either the problem is solved for a linearized system, or a trial and error engineering approach is taken to approximately minimize the cost. There are several difficulties that currently prevent a mainstream use of non-linear optimal control by engineers. Here, I study the approach from the mathematical perspective, and use the analogy in physics to

develop a method to solve the problem of a roundhouse kick in Taekwondo.

The problem of an optimal roundhouse kick, has many aspects of geometric control theory. Naturally, it is a non-linear mechanical system with constraints. The constraints could be added to a larger vector space in the form of Lagrange multipliers. This computation has a number of disadvantages and a geometric approach leads to interesting analysis and computational methods influenced by the geometry and conservation laws. These benefits include a more intuitive expression for the governing equations and conservations. The Cayley transform allows efficiently approximated integration with the ability to stay exactly on the constraint manifold.

Control theory has a long history of use in a variety of fields of engineering. I focus on the Hamiltonian approach to optimal control developed by Lev Pontryagin [7]. Not only is this approach effective in understanding and solving problems in optimal control, but the mechanisms parallel almost exactly similar mechanisms used in physics and suggest a broader mathematical understanding of the concept of momentum. By expressing the kick as a representation otrajectory in a Lie group, we get to make use of the physical geometry and symmetries as well as the well established theory of Lie groups. Of particular importance to the approach is the triviality of the tangent bundle for Lie groups. We shall see that while the Hamiltonian approach to optimal control adds more variables and an even further removed dual to the tangent of the co-tangent space, there is no additional complexity added in the way we must compute the equations.

The full problem has two features that add difficulty. One is the fundamental non-linearity of the important dynamics of the problem. The second is the non-commutativity of the rotation group. Not only does this make the algebra difficult, but it gives the equations

more difficult terms. In the thesis, I attempt to derive these equations in a straight forward manner so that a computer could handle all the algebraic manipulations and derivatives. I also consider if it is beneficial to analyze the non-linearity on the maximal torus of the full manifold, and then use approximation methods to get back to the full manifold. In Chapter 2.1, I develop the first order differential equations of a planar front kick, which is on the maximal torus of the full manifold. In Chapter 2.2, I present an optimal control problem for the full-mechanical system that uses the dynamics from Chapter 2.1. In Chapter 3, I finish with computational results for the planar problem.

1.1 Background

1.1.1 Hamilton's Mechanics and Lie Groups

Hamiltonian mechanics plays a very significant role in physics. Far more than just a convenient way to derive equations of motion, analysis of the Hamiltonian is very useful in understanding many of the properties of the system. Even in the upper division class on mechanics, PHYS 105, the discussion of Hamiltonian mechanics did not move far from discussing the case of an unconstrained space on $\mathbb{R}^n$, or simple constraints such as a few circles or free body rotation. Certainly the ability to express parameterizations of manifolds with scalars is general, but this direct approach to solving problems can lead to tedious calculations for unsolvable equations and dead ends in analysis. The problems of Hamiltonian mechanics that occur on manifolds with symmetry have a direct relation with very deep mathematics through the representation of these manifolds with Lie groups.

A Lie group is a mathematical group, with a multiplicative binary operation, identity, and inverses for all elements, that is also a smooth, differentiable manifold. The theory

of Lie groups combines the powerful mathematical fields of group theory and topology, in a way that leads to many useful results [6, 8]. Not only are Lie groups of mathematical interest, but they are at the forefront of our understanding of physics as they describe continuous symmetries. In this thesis I examine a problem where Lie groups arise from rotations of rigid bodies, a global representation of the system. However, one of the most important uses of Lie groups is in understanding the fundamental forces on the very minute scale. Not only do the Lie groups come up in particle physics, and classical mechanics, but they are prevalently used in quantum mechanics. In classical mechanics, Lie groups are represented as an action on a vector space, whereas in quantum mechanics, the action is on a Hilbert space. The Lie algebra describes the infinitesimal changes in the group representation, and in the classical theory can be represented by a tensor algebra. In quantum mechanics the infinitesimal changes of the function space representation require the Lie algebra to be expressed as differential operators.

The reason Lie groups are important in the problem I am looking at is because they govern the principles of calculus over non-commutative variables. This is equivalent to saying they govern the calculus over non-linear manifolds with symmetry if we emphasize the manifold structure. For any Lie group we can consider the tangent space to the manifold at any point. Because it is a manifold, if it is of finite dimension, n, the tangent space looks locally like $\mathbb{R}^n$ anywhere on the manifold. However there is more to it. In fact the vector space picks up a generally non-commutative product along with the abelian vector addition, which becomes extremely useful in the analysis of changes on the manifold. The mathematical theory of Hamiltonian systems generalizes naturally for a manifold of a Lie group. The only difference is a single extra term to take into account the non-commutativity

of the group. The equations (1.1) give a vector field on the manifold, the equations for the vector field of the Taekwondo kick are developed in Chapter 2. The vectors generate one-parameter subgroups through the exponential mapping from the Lie algebra to the Lie group. Infinitesimally these are the integral of the vector field and give the trajectory on the manifold. We must relate the Lie algebra elements to a vectors space by a trivialization that relates the tangent space at a point on the manifold to the tangent space at the identity element of the group. There exists more than one way to do this trivialization. For a group isomorphic to a finite dimensional matrix group, the choice corresponds to whether the tangent space at a point is related to the tangent space at the identity by left or right multiplication. The physical representation of the system depends on this choice and the trivialization used admits a subtle difference in the mathematics. There is a natural pairing of algebra elements and dual elements, which for our purpose is equivalent to an inner product on the vectors space of tangent and co-tangent vectors.

Throughout this thesis I use lower case English letters for variables in the tangent space, and corresponding Greek letters for the co-tangent space. The terminology of tangent space and Lie algebra is used interchangeably when there is an association of tangent vectors with Lie algebra elements. Given a manifold M, group element g, tangent vector m, and co-tangent vector μ, the Hamiltonian becomes a scalar function of g, μ. Hamilton's equations give the vector fields on the tangent and co-tangent bundles of the Lie group manifold,

$$\begin{aligned} \chi_H^g(g,\mu) &= \frac{\delta H}{\delta \mu}(g,\mu) \\ \chi_H^\mu(g,\mu) &= -\frac{\delta H}{\delta g}(g,\mu) + ad^*_{\frac{\delta H}{\delta g}(g,\mu)}\mu. \end{aligned} \tag{1.1}$$

The evolution of a trajectory from these vector fields depends on the trivialization. For a

right trivialization the evolution takes the form

$$\dot{g}(g,\mu) = g\chi_H^g(g,\mu), \quad \dot{\mu}(g,\mu) = \chi_H^\mu(g,\mu)$$

The partial derivatives in Eq. (1.1) can be defined in a standard manner as the limit of a change in a given direction. For changes on the manifold we take that direction as an vector of the tangent space, relating back to the group through the trivialization and exponential mapping. The partial derivatives are computed via the natural pairing of algebra elements and their dual, or an inner product on the vector space. In the second equation we assume a right trivialization, which will be what is used in this thesis,

$$\begin{aligned} < \frac{\delta H}{\delta \mu}(g,\mu), \gamma > &= \lim_{\epsilon \to 0} \frac{dH(g, \mu + \epsilon\gamma)}{d\epsilon} \\ < m, \frac{\delta H}{\delta g}(g,\mu) > &= \lim_{\epsilon \to 0} \frac{dH(ge^{\epsilon m}, \mu)}{d\epsilon}. \end{aligned} \tag{1.2}$$

The adjoint is a linearization of the conjugate action, and the co-adjoint appearing in the equation is from the pairing with a co-tangent vector,

$$\begin{aligned} ad_n m &= \frac{de^{\epsilon m} n e^{-\epsilon m}}{d\epsilon} \in TM \\ < m, ad_n^* \gamma > &= < ad_n m, \gamma > . \end{aligned} \tag{1.3}$$

Clearly if the group is abelian this term disappears. While notationally complex, after we choose a trivialization of the Lie group the computation is straightforward, and basic identities lead to a simple expression for this term.

1.1.2 Poisson Brackets

Poisson brackets are a very interesting way to analyze the behavior of functionals in a system. They give an alternative method to derive equations of motion, and they are

very important in understanding the transition between classical and quantum mechanics. Abstractly, the Poisson bracket gives an algebraic structure to functionals of T^*M, with the product the Poisson bracket. The Poisson bracket takes a form on the manifold of Lie groups in a manner where the connection with the Hamiltonian approach is obvious. The bracket is defined for two functionals on the system at a point in the co-tangent space

$$\{f, H\} =< \frac{\delta f}{\delta g}, \frac{\delta H}{\delta \mu} > - < \frac{\delta H}{\delta g}, \frac{\delta f}{\delta \mu} > \pm < \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta H}{\delta \mu}\right] > . \tag{1.4}$$

What appears here as the commutator bracket is generally the Lie algebra multiplication operator. The ambiguity in sign of the last term is resolved by choosing a trivialization for the co-tangent bundle. This is equivalent to the co-adjoint operator that came up in the Hamiltonian equations. The physical significance comes when the H is the Hamiltonian for a system, then the Poisson bracket gives the time evolution for a functional f, $\dot{f} = \{f, H\}$. If this quantity is zero for every Hamiltonian, then the function is known as a Casimir. For any function whose bracket with a specific Hamiltonian is zero, represents a conserved quantity for the Hamiltonian system. From the asymmetry of the bracket, $\{H, H\}$ is zero and the Hamiltonian is always conserved.

Notice that all the pairings in the inner products are natural pairings of tangent vectors and co-tangent vectors in the dual space. The dual space is defined mathematically as a space of linear functionals mapping a vector space to $\mathbb{R}$. This has major significance in physics as it represents the difference of velocity and momentum. In fact, if unsure about whether a vector is dual or not, it can be resolved by dimensional analysis of the units. If there is a term of mass, or charge for electromagnetism, the quantity can be considered to lie in some sort of dual space.

1.1.3 Control Theory

Control theory comes in many flavors, but I focus on control of a deterministic, first order system, which is to be done in an optimal way. In many ways a Hamiltonian mechanical system can be thought of as a control problem without any controls and with explicit dependence of the cost function on the dual variables. If the quantity to minimize is the Lagrangian action, the equations of control theory are identical to Hamilton's equations. Control theory generalizes this principle by considering that our system depends in some way on quantities which we have control over. Now we are faced with the problem of how to choose these controls in an optimal manner. This problem has a wide range of uses. There is the obvious problem in industry of what inputs into a system will produce the best output. Control can also be used abstractly in problems like modeling nature. The basic principle that evolution is an optimizing system can be used to gain an understanding of ecosystems or even biological functions.

In optimal control theory the problem is generally stated as a dynamical system, where the first derivatives of the state variables are known up to some dependence on the control variables. The problem can be stated for very general constraint spaces, but then more information is necessary to make use of many of the tools. For the main problems considered here we assume exact conditions on the initial and end points of the trajectory are known, but not the end time. These assumptions are natural for the roundhouse kick in Taekwondo because the initial state of the leg is vertical and the ending state of the leg is the target, but the speed and timing of the kick is unknown. Other approaches are also valid, and in Section 2.2.1 I discuss the tradeoffs in complexity the problem when we adjust the assumptions of the problem. A generic one-dimensional problem with a single control

has the following form:

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = x_0, \quad x(t_1) = x_1,$$
$$u^* = \min_{u \in U} \left\{ \int_0^{t_1(u)} C(x, u) dt \right\}. \tag{1.5}$$

The minimization is understood to be under the dynamic constraints of the system as well as the boundary conditions. This is obviously a difficult problem, as we have the constraint given by a differential equation, or generally a system of ODE's, and we are not even given the time that the system will terminate, but instead the position. For the purposes of this study I only consider when the domain of controls is open and dense. If the domain is not open more care must be taken to consider the possibility of a control on the boundary. If there is no or little cost on the controls, Pontryagin discovered that it is common for a problem with a closed set of controls to exhibit a bang-bang solution. This term refers to how the control jumps from one boundary to another without spending any time in between. While simple solutions, the switching time can still be difficult to calculate for these problems [5]. The Pontryagin Maximum Principal [7] gives us an approach to this general problem by constructing a Hamiltonian and analyzing the problem through the evolution of the adjoint system.

$$H(x, p, u) = p_0 C(x, u) + f(x, u)p,$$
$$\dot{x} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial x}, \quad p_0 < 0,$$
$$u(t)^* = \max_{u(t)} \{H(x(t), p(t), u(t)\}. \tag{1.6}$$

We have introduced additional variables here, so we have some degree of freedom of choice. First of all, p_0 is a constant that we can choose to be less than zero. This allows the expression of the control as a maximization of the Hamiltonian, the variable, p_0, is by

convention set to -1. Some people refuse to use this convention and refer to this as the minimization principle by having $p_0 > 0$. The dual variable p also has a scale factor we can decide on. Under assumptions of continuity of the equations and a certain niceness of the control space (while complicated to make rigorous we certainly stay in this range in this discussion), the Hamiltonian is conserved under the evolution of the equation. Because of this we use the freedom in the scale of p to set $H(t) = 0$ for all t. The advantage to these equations is that now we can determine u at a specific time due to the maximization of the Hamiltonian. In the problems we analyze, the simple dependence on u allows for a simple expression for this maximization. The remaining difficulty is in determining the beginning and end conditions for p. In fact, this is only determined by solving for the dynamical systems dependance on an optimal control with the boundary conditions given for the state system. While these equations can all be expressed, the analytic solution is generally not possible to get even for simple examples. There are numerous approaches taken to approximate this. One is to expand the expression for $u(t)$ to a finite-degree polynomial for which the state system can then be integrated analytically in terms of $p(t)$. Then this analytic expression of $x(t)$ is solved for the boundary conditions giving an approximation for the boundary conditions on the adjoint system. This approach would be difficult in the problems we analyze, and may not capture the non-linearities that are important. From this point a better approximation is still wanted, particularly approximations that converge to the true optimal solution. A basic way to do this is to consider that we have in effect reduced the optimal control problem to the optimization problem on the space on terminal or initial conditions of the dual variable p. This optimization problem is particularly difficult as the constraints are the constraints of the state in Eq. (1.5) and of the evolution from Eq.

(1.6). If we can efficiently compute the optimal trajectories, there are approaches to this optimization problem. This is an advantage, as we have reduced the mathematically infinite dimensional optimization problem, over the function space on $g : [0, t_1] \to U$ to a finite optimization problem on the end space of the dual variable. There is one more difficulty, in that these are necessary but not sufficient conditions for optimality. For example on this end condition space, there may be more than one point that satisfy all the conditions. We have to compute the actual cost of these trajectories in order to find the optimal one. In the problem analyzed in this thesis, intuition suggests where we can look for the actual optimal solution.

These equations clearly resemble Hamilton's equations from physics, but let us look into this in more depth. Consider the problem in classical mechanics of the minimization of the action, $L = KE(x, \dot{x}) - PE(x)$. Although there is no control, we can still put it into Eq. (1.6)

$$H = p_0 L(x, \dot{x}) + \dot{x}p.$$

This looks similar except for the dependance on $\dot{x}$ in the cost. In physics this is resolved by the interpretation of the dual as the momentum, a physically relevant quantity. This dual can be solved by $p = \frac{\partial L}{\partial \dot{x}}$, and then the cost becomes a function of the momentum p. Set $p_0 = -1$ and solve for $\dot{x}$ in terms of p

$$H = -KE(x, \dot{x}(p)) + PE(x) + \dot{x}p.$$

The final step is to realize that we have effectively linearized the kinetic energy dependance on the velocity, so the kinetic energy is actually just $\frac{1}{2} < \dot{x}, p >$ and the Hamiltonian

becomes what we are used to

$$H = KE(\dot{x}, p) + PE(x) = \frac{1}{2}\dot{x}p + PE(x).$$

Hamilton's equations are also identical. There is some hand-waving in this description; in the thesis I am writing for math, I attempt to explore more of the relationship of physics between the state and adjoint systems and the relation with control theory. The co-state variables for a control problem form a sort of momentum with regard to the cost. This momentum hold information so that optimal trajectories won't just go in the direction of instantaneous least cost, but in a direction that also minimizes future cost. For the purpose of this thesis the focus is on how to use theses concepts to approach a problem. In this regard it is important that the generalization of Eq. (1.6) to manifolds in particular Lie groups, is equivalent to the generalization in Hamiltonian mechanics so there is no need for a description as the equations are the same [1].

1.1.4 Geometric Algebra

A primary aim of the approach taken in this thesis is to express the relationship of the geometry of the mechanics and the algebra utilized to solve the problem. One tool to do this is the formulation of a geometric algebra. Geometric algebra, developed by William K. Clifford, is also known as a Clifford algebra. More recent work has been done by David Hestenes and others to develop the generalized calculus on the algebra [4]. The geometric algebra is intrinsically related to the exterior algebra of a vector space. It encompasses the algebra of differential forms and also works with a general metric. The basic idea is to combine all the elements of geometry in an n-dimensional space, including the planes, volumes and scalars, into a single algebra with a new geometrically relevant

product. The basic building blocks are the symmetric and anti-symmetric products of vectors. The elements of the algebra are separated into different grades, which are raised or lowered by the products. The geometric product, represented by simple multiplication, is an associative and non-commutative product among all the elements of the algebra. The geometric product distributes over addition, making it analogous to matrix multiplication. Matrices are typically used when working with tensors in physics. The advantages of working with a geometric algebra is that it only holds the relevant information and makes more obvious the properties of the relevant algebra, and the relations with mathematical ring theory. The greatest benefit of using a geometric algebra in physics is found when doing relativistic physics. The metric signature is incorporated into the algebra, the Lorentz transformations appear just as the standard rotations of the geometric algebra. The Clifford algebra provides a very useful setting for the spin representation of Lie groups, but this is beyond the scope of this discussion.

The geometric relationship exists for non-relativistic mechanics. For the purposes of this thesis we are only concerned with the $Cl_{3,0}$ Clifford algebra of three dimensional space. Construction of this algebra begins with an ortho-normal basis for a Banach vector space, where there is a well defined norm. $a_1, a_2, a_3,\ a_i \cdot a_j = \delta_{ij}$, is the standard inner product in $\mathbb{R}^3$. The outer product, $\wedge$, results in a second grade bi-vector and is anti-commutative. The geometric product among vectors can be expressed as the combination of the inner and outer products, $ab = a \wedge b + a \cdot b$. The product has a grade-0 scalar, and a grade-2 bi-vector. Geometrically the scalar is the magnitude of the projection, $|a||b|cos(\theta_{a,b})$, and the bi-vector contains the information of the orientated plane containing the two vectors, as well as the magnitude of the parallelogram in the plane. The product of higher grade elements has

geometric significance through the projection onto certain grades of the resulting element. The parity of the sum of the grades of two multiplied vectors is preserved by the product. The lowest resultant possible grade is the absolute value of the difference of the grades, and represents an orthogonal projection of the lesser grade onto the higher grade. The highest grade is a sum of the two grades, and gives a volume spanned by the orthogonal components of the two elements multiplied. Some basic facts of the algebra that follow from the above definitions are that the highest grade is the same as the dimension, n of the starting vector space, and that there is only one element known as the pseudo-scalar of grade-n, which may or may not commute with other elements. The number of elements in a grade k is given by the combinatoric

$$\left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{(n-k)!k!}.$$

The total number of elements in the algebra is 2^n. All elements of a single grade square to a scalar, although not necessarily a positive one. We consider the conjugate of an element to be the reversal in order of the vectors composed to get the element. For example, for a bi-vector $M = ab, \overline{M} = ba = -M$. The switch in sign is immediately obvious because two vectors that multiply to form a pure bi-vector must be orthogonal. The conjugate is no more than a switch of sign for single grade elements, and distributes over addition for multi-vectors.

An interesting difference from standard vector formulation is how linear transformations can be described by elements in the algebra. Up to now there was no discussion of the geometric significance of mixed grade elements, multi-vectors, but these are important for the role that they play in describing linear transformations of the algebra. A general linear transformation requires multiplication on both the right and the left by multi-vectors.

Any linear transformation on vectors given in matrix form can be easily converted to the algebra by using the inner product to project onto a basis. More interesting is the case when the linear transformation can be expressed without a basis in a manner that reflects a geometric transformation.

The main focus of this thesis deals with using the geometric algebra to describe rotations. Rotations are special linear transformations that form a Lie group. One construction of the rotations relies on identifying the Lie algebra as a sub-algebra of the geometric algebra. For any geometric algebra, the bi-vectors generate rotations. For $Cl_{3,0}$, the bi-vectors along with the scalar form a sub-algebra isomorphic to the quaternions. Every element squares to -1 and the basis elements permute as quaternions. All of this can be discovered through forming a basis of bi-vectors with an ortho-normal vector basis. The quaternions have a subgroup of unit quaternions that describes the rotational Lie group $SU(2)$ by their adjoint action on pure quaternions. Let $q \in \mathbb{H} \cong (Cl_{3,0})_{2,0}$, meaning the grade 2 and grade 0 elements of the three dimensional Clifford algebra. The the set $\{q : \overline{q}$ is a sub group. If q is expressed as its scalar and bi-vector parts, we can deconstruct the rotation,

$$q = s + B \Rightarrow q = cos(\frac{\theta}{2}) - sin(\frac{\theta}{2})\hat{B} = e^{-\frac{\theta}{2}\hat{B}}.$$

The exponential map can be defined by the standard series expansion, and this result is clear from the association of $\hat{B}$ with the imaginary unit. The factor of $-\frac{1}{2}$ is necessary in the description of this as a rotation through the adjoint action. The bi-vectors are now the generators of the Lie group and describe the change of a rotation through the formula, $\dot{R} = -\frac{1}{2}R\Omega$. Whether the Ω appears on the left or on the right is a choice we make that determines the trivialization of the Lie group. Put simply, appearing on the left the Ω

describes a change in the rotation of the frame, where as on the right it is the change of the rotation of the body. We could use either one. If we were to use the rotation of the frame, then the leg would always stay in the same plane leaving out any term requiring the rotation of the inertial tensor. I choose the right trivialization as a more useful choice. This choice preserves the symmetry that would be broken if the axis of gravitation was rotating. Rotation in the plane of $\hat{B}$ through an angle θ is given by,

$$R_{\theta,\hat{B}}(v) = qv\overline{q} = e^{-\frac{\theta}{2}\hat{B}} v e^{\frac{\theta}{2}\hat{B}}. \tag{1.7}$$

This rotation makes sense for vectors, bi-vectors, and can be extended for linear transformations. Rotation of every term of a linear transformation is equivalent to rotating the frame with which the transformation is defined. This is not the same as the product of the rotational group, which is a composition of rotations and is simply obtained by the product of rotors. The composition of a rotation and a linear transformation is also different from rotation of the transformation as it just rotates the output or the input not the frame.

The adjoint operation for the rotation is a linearization of the rotation,

$$\frac{de^{-\frac{\epsilon}{2}B} A e^{\frac{\epsilon}{2}B}}{d\epsilon}|_{\epsilon=0} = \frac{1}{2}(-BA + AB) = [A, B]. \tag{1.8}$$

This is the product for the Lie algebra and returns another bi-vector.

Note that the wedge product of the algebra is equivalent to the cross product for vectors if we replace the bi-vector by an axial vector. Using the cross product we can also formulate the rotations, but using the geometric algebra does it in a way that generalizes to higher dimension and works for higher grades. The commutator product is also equivalent to the cross product between axial vectors.

The geometric algebra puts many concepts of physics in an interesting context [3].

In classical mechanics the angular velocity, angular momentum are all naturally bi-vectors as they are dealing with infinitesimal rotations, or generators for our rotational Lie group. The torque, as a time derivative of the angular momentum, is also naturally described by the geometry of the bi-vectors. The equations are equivalent to what they are in standard notation, with the cross product replaced by the wedge product. This distinction helps to keep straight that position or velocity vectors are conceptually different from axial rotation vectors with entirely different units. The inertia tensor, is now a linear transformation taking an angular velocity to the angular momentum. The definition in terms of the geometric algebra is fairly simple, take a bi-vector B and a density ρ [3],

$$\mathbb{I}(B) = \int \rho(x) x \wedge (xB)_1.$$

The projection of xB onto the vectors is done using a half of the commutator bracket, $(xB)_1 = [x, B] = \frac{1}{2}(xB - Bx)$. Some authors use this projection as a generalization of the inner product. I don't like this because it lacks symmetry; I only consider the inner product to be defined on objects of the same grade as the scalar part of the geometric product. This orthogonal projection nicely captures the necessary geometry as it returns a vector of the velocity at the point due to a rotation generated by B. For our problem we only consider an inertia tensor with principal axis along the basis axes, which can be defined by the projection of the angular velocity on the bi-vector basis. The inner product of the algebra for bi-vectors returns a scalar that is negative for the inner product of a bi-vector with itself,

$$A \cdot B = \frac{1}{2}(AB + BA). \tag{1.9}$$

In Section 2.3, an inner product denoted by a bracket is used. The bracket inner product refers to the above equation except with the conjugate of one of the vectors, introducing a

negative sign.

We are often concerned with the transpose of a linear operator. In the geometric algebra, linear operators transpose via a reversal of order. The linear transform can be decomposed in terms with only single grade elements on either side. All of these are invertible, but the following holds for when m, n are multi-vectors by considering the decomposition into graded vectors,

$$\begin{aligned} < mXn, Y > &= \frac{1}{2}(mXnY + YmXn) = \frac{1}{2}((mXnY)_0 m^{-1} m + nn^{-1}(YmXn)_0) \\ &= \frac{1}{2}(XnYm + nYmX) =< X, nYm > . \end{aligned} \tag{1.10}$$

Another property that is useful is how we can cycle commutations and inner products among bi-vectors similar to the cycling of the cross product and vectors. Let A, B, C be bi-vectors,

$$< A, [B, C] >= \frac{1}{2} < A, BC - CB >= \frac{1}{2} < AB - BA, C >=< [A, B], C > . \tag{1.11}$$

The geometric algebra can also be used to express a differential operator [4]. The operator can be defined as partial derivatives with respect to vectors, but the great thing is that it operates as an element of the algebra. We use $\square_1$ for the general vector derivative. If we let f be a scalar function over vectors, then $\square_1 f$ takes on the interpretation as the vector valued gradient. In some ways it is more natural to interpret this output as being in the dual space, but this space has the same algebraic properties as our original space. If F is a vector valued function over vectors the derivative is more interesting, $\square_1 F = (\square_1 F)_0 + (\square_1 F)_2$. The scalar part is equivalent to the divergence of F in three dimensions, whereas the bi-vector component represents the curl. It is natural that the curl takes on a bi-vector value as the intuition is that it represents some infinitesimal rotation of the vector field.

The bi-vector derivative can also be defined as partials with respect to the bi-vectors of a function over bi-vectors. This is something that comes up in the problem I am focusing on. When we take the derivatives with respect to the co-tangent vectors, this is sort of a dual bi-vector gradient that takes values in the tangent bi-vector space. If we remember that the derivative is with respect to the co-tangent bi-vectors, we can express the Poisson bracket in terms of this operator. The derivative with respect to a rotation, naturally takes a value in the dual to the tangent bi-vector space. For just the $SO(3)$ part of the problem, let L be the angular momentum, which is a co-tangent bi-vector. Then Eq. (1.4) can be expressed with the dual bi-vector derivative operator, $\overline{\Box}_2$,

$$\{f,g\}(R,L) =< \overline{\Box}_2 g, \frac{\delta f}{\delta R} > - < \overline{\Box}_2 f, \frac{\delta g}{\delta R} > + < \overline{\Box}_2 g, \overline{\Box}_2 f], L > .$$

2

The Problem

This problem considered here is the optimization of the trajectory of a double jointed limb. We have two rigid bodies with known lengths, masses, and moment of inertia tensors. The first joint, the hip, is fixed to a point in space, the second joint, the knee, is connected to the end of the first. The end of the second rigid body is the foot, which we want to strike the target. We first consider the simpler problem of the motion in a plane. The martial arts inspiration is a front snap kick, or 'ap chagi' in Korean, a forward kick where the leg is whipped or snapped to make efficient use of energy. In Taekwondo, this kick is used as an instructional kick for beginners, who may not have the hip flexibility for the full kick, to start to grasp the fundamental motion. For the full problem the first joint, the hip, is free to rotate in three dimensions and the second joint, the knee, is restricted to rotations in a plane. This is inspired by the roundhouse kick, 'dollyo chagi' in Korean (although this spelling does not capture the pronunciation well.) This kick is similar to the front kick, but comes around to strike parallel to the ground.

The controls for the problem are torques around the joints. The goal is to find

a time-optimal trajectory with respect to quadratic costs on the control torques. The beginning condition is with the limb straight down. The end condition has the leg horizontal, with the upper limb stopped and only the lower limb moving; in the first problem this is upwards, while in the final problem the rotation of the second joint is around the vertical axis. This is clearly a simplification of the motion of the kicks in Taekwondo, where such a movement would result in an overextended knee. However, in the real motion there is a transfer of energy and momentum from the leg to the target, which we do not model. The costs for the kick are up for interpretation. A simple cost is simply the square of the magnitude of the torques, along with a constant term that encourages speed. Another approach would be to emphasis power by maximizing the terminal energy through a term maximizing the instantaneous work done on the second limb. A final approach, which is slightly different, is to incorporate the end condition into the cost. This allows the end condition not to be determined exactly allowing the end adjoint condition to be known. It would also allow for the accuracy to be related in a trade off with the other costs of optimal time and minimal exertion. In Taekwondo the major considerations are accuracy, speed, efficiency, and power, often in that order. While other arts put more focus on other things, the understanding in Taekwondo is that if the strike is precise enough, not much power is needed to disable an opponent. Besides most of the power comes from speed, and a slow or inaccurate strike becomes a major vulnerability.

2.1 The Manifold

For the initial planar problem, the configuration manifold is the two dimensional torus, $M = S^1 \times S^1$, representing the rotations of the two joints. Along with the tangent

space $TM \cong M \times \mathbb{R}^2$, representing the angular velocity. The state of controls is the tangent space to T^*M, which represents the change in angular momentum or torque and is isomorphic to $\mathbb{R}^2$. With this respect the configuration is a faithful, surjective representation of the $SO(2) \times SO(2)$ Lie group. In the final problem of the roundhouse kick, the manifold is $S^1 \times (S^3/\pm 1)$. It is also a faithful, surjective representation of $SO(3) \times SO(2)$. It is important to note that the maximal tori is identical to the manifold of the initial problem.

The physical representation is of two connected rigid bodies in $\mathbb{R}^3$. The limbs are constrained to be described by only their rotations, the tangent space to these rotations is isomorphic to the space of pure vector quaternions, vectors with the cross product operation, or the bi-vectors of the three dimensional geometric algebra. For the quaternions or bi-vectors, half of the commutator is the proper operation for the Lie algebra. A choice must be made whether to use a reference frame fixed in space or fixed with the orientation of the leg. This makes a difference in how we trivialize the tangent space. If we chose the frame of the leg we would not have to worry about the rotation of the tangent vectors with respect to the inertia tensor. Then the only term in the change of the momentum is from gravity and the rotation of the knee, but gravity would no longer be in a fixed direction. The real disadvantage to this choice however is the complication of the conserved quantities. Without external torque, the momentum is conserved around the vertical axis, if this axis is rotating then the conserved part of the momentum is more difficult calculate. In terms of how we intend to calculate and analyze the problem, I expect this additional complexity in the momentum conservation to pose a more substantial problem than the extra terms to the change of momentum.

There is a kinetic energy function, $TM \to \mathbb{R}$, that provides a metric on the space,

as well as a defined potential energy functoin to comprise the Lagrangian. From the metric we can calculate the dual vectors, and then from the Lagrangian we can derive equations for the non-linear geodesic curves on the manifold. For the control problem we require first order equations so we refer to Hamilton's equations of motion Eq. (1.1). In the control problem it is also standard to define a metric on the control space as a contributor for the cost function, which is chosen to be quadratic in the controls. The problem can be considered to find the Hamiltonian vector fields on the full space of TT^*M, once we fully develop the relations of the controls with the configuration. The paths generated by these vector fields are the optimal trajectories, it is then be a matter of finding the optimal trajectory with the boundary conditions we require. An additional difficulty is added because we do not know the initial or end conditions for the adjoint system on the dual space T^*T^*M and have to find ones for which the cost is minimal.

Let R_α, R_β denote the rotations of the 'hip' and 'knee' joint respectively. In the planar model, both of these rotations lie in $SO(2)$, in the full model the first is an element of $SO(3)$. Dealing now with only the mechanical system, we get a Hamiltonian structure, with a co-tangent space from the dual to the tangent via the kinetic energy metric. The co-tangent momentum vectors are denoted as η, ν; the corresponding tangent vectors, e, n. Please note that these elements are not raised to any powers, so if there is an e raised to a power this is the exponential map from the Lie algebra to the Lie group. From the Hamiltonian structure, there is a conserved Hamiltonian $H : T^*M \to \mathbb{R}$, and a Poisson bracket given by the manifold for the problem without the external torques. We have a choice to make for the exact representation of the group. I choose the second rotor to be inside the rotated frame from the first rotation. The position of a point on the second

limb is $R_\alpha(\mathbf{r_1} + R_\beta(x\mathbf{r_2}))$, not $R_\alpha(\mathbf{r_1}) + R_\beta(x\mathbf{r_2})$, for $x \in [0,1]$. From here on out, $\mathbf{r_1}, \mathbf{r_2}$ are the vectors with magnitude of the full lengths of the respective limbs. The reason for choosing this manner of rotation is that R_β only has one degree of freedom, it is restricted to rotations in a plane.

We are working with Hamiltonian vector fields on the co-tangent bundle, which give the first derivatives of the co-tangent vectors at a point on the manifold,

$$\chi_H^\eta = -\frac{\delta H}{\delta R_\alpha} + ad^*_{\frac{\delta H}{\delta R_\alpha}}\eta, \qquad \chi_H^\nu = -\frac{\delta H}{\delta R_\beta}.$$

We wish to add the controls to these vectors such that at a point $P \in M$:

$$\dot{\eta}(P) = \chi_H^\eta(P) + \gamma - \mu \quad \dot{\nu}(P) = \chi_H^\nu(P) + \mu.$$

Now we also need to have the dreaded co-co-vectors of T^*T^*M, which is denoted by a p and subscript of their corresponding co-tangent elements. A cost per time given by $c_1 = 1 + \frac{a}{2}|\gamma|^2 + \frac{b}{2}|\mu|^2$. The Hamiltonian for the optimal control problem is

$$\mathcal{H} = -c + < \begin{pmatrix} e \\ n \\ \chi_H^\eta + \gamma - \mu \\ \chi_H^\nu + \mu \end{pmatrix}, \begin{pmatrix} p_\alpha \\ p_\beta \\ p_\eta \\ p_\nu \end{pmatrix} > .$$

The equations generated by this Hamiltonian are definitely complicated. The best way to handle it is to break it down in smaller problems, which can be solved or approximated separately. To begin consider the problem of the Hamiltonian system on the plane.

2.2 Equations of the planar problem

For the derivation of the equations for this problem we can use the representation of the problem in the complex plane. The generators for the rotational group, $U(1)$ are pure imaginary numbers, we also use imaginary numbers for the vertical axis. There is no need for the geometric algebra, it would be an equivalent representation, but everything commutes so we stick with the complex numbers. We use $\mathbb{I}_1, \mathbb{I}_2$ to denote the moment of inertias of the two limbs, m_1, m_2 to denote the mass of the limbs, and $\mathbf{cm}_1, \mathbf{cm}_2$ for the vectors pointing to the center of mass of the limbs. The Hamiltonian for the planar mechanics problem with no external forces is the total energy,

$$H = \frac{1}{2}\mathbb{I}_1|e|^2 + \frac{1}{2}\mathbb{I}_2|e+n|^2 + \frac{1}{2}m_2|\dot{\mathbf{cm}}_2|^2(R_\beta, e, n) + g < m_1\mathbf{cm}_1(R_\alpha) + m_2\mathbf{cm}_2(R_\alpha, R_\beta), \hat{e}_3 > . \tag{2.1}$$

By choosing the second rotation to be rotated with the first, the total angular velocity of the second limb is the addition of the two tangent vectors. While this may make these equations slightly less appealing, it simplifies the larger problem as the tangent vector to the rotation of the second limb still stays in one dimension with the constraints. The moment of inertias are different as the first vector is a rotation around the end of the limb and for the second we consider rotation around the center of mass. For simplicity we assume symmetry of the limbs such that $\mathbf{cm}_1 = R_\alpha(\frac{\mathbf{r_1}}{2})$ and the $\mathbf{r}$ vectors are of the initial orientation of the limb. The inertias can calculated standardly as the inertia of thin rods,

$$\mathbb{I}_1 = \frac{m_1}{3}|\mathbf{r}_1|^2, \quad \mathbb{I}_2 = \frac{m_2}{3}|\frac{\mathbf{r}_2}{2}|^2.$$

We have an inner product on the tangent space that gives the kinetic energy part of the Hamiltonian. The inner product is described by a symmetric, positive definite, metric

tensor, G. From here on we consider the tangent and co-tangent vectors to be imaginary numbers. This is the Lie algebra for rotations on a plane and we use a complex inner product, which since the vectors are pure imaginary contributes a negative sign.

$$KE = \frac{1}{2} < \begin{pmatrix} e \\ \\ n \end{pmatrix}, \begin{pmatrix} e \\ \\ n \end{pmatrix} >_G = -\frac{1}{2} \begin{pmatrix} e & n \end{pmatrix} G \begin{pmatrix} e \\ \\ n \end{pmatrix} = \frac{1}{2} < \begin{pmatrix} e \\ \\ n \end{pmatrix}, \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix} > .$$

The final pairing between the velocity and momentum corresponds to a complex inner product with the trivialization of the elements in a complex vector space. For the representation of the complex numbers to the space, we relate a vertical component of the vector with the imaginary part, and the horizontal with the real part. The potential energy term simply isolates the imaginary part of a rotated vector for the height. Rotations are the imaginary exponents, and the tangent vectors are pure imaginary numbers.

With this representation the initial condition is with both vectors in the direction of $-i$. We also use l_2 to be the length from the 'knee' joint to the center of mass of the second limb and l_1 the length from the 'hip' joint to the end of the first limb. The terms for the center of mass of the second limb and their derivatives can be expressed either through the rotation transformations, imaginary exponentials, or trigonometric functions. These are all equivalent; what is important is that it is expressed in terms of the rotations and their tangent vectors. The computation of G and V, the potential energy, is put off in the Appendix A.1.

The simple Poisson bracket for the abelian group is

$$\{f, g\} = \left\langle \begin{pmatrix} \frac{\delta f}{\delta R_\alpha} \\ \\ \frac{\delta f}{\delta R_\beta} \end{pmatrix}, \begin{pmatrix} \frac{\delta g}{\delta \eta} \\ \\ \frac{\delta g}{\delta \nu} \end{pmatrix} \right\rangle - \left\langle \begin{pmatrix} \frac{\delta g}{\delta R_\alpha} \\ \\ \frac{\delta g}{\delta R_\beta} \end{pmatrix}, \begin{pmatrix} \frac{\delta f}{\delta \eta} \\ \\ \frac{\delta f}{\delta \nu} \end{pmatrix} \right\rangle.$$

The partial derivatives are also trivial and could be expressed in terms of the angles,

$$< v, \frac{\delta H}{\delta R_\alpha} >= \overline{v}\frac{\delta H}{\delta R_\alpha} = \frac{d}{dt}H(e^{i\alpha+tv)})|_{t=0} = ve^{i\alpha}\frac{\partial H}{\partial e^{i\alpha}} = ve^{i\alpha}(-i)\overline{e^{i\alpha}}\frac{\partial H}{\partial \alpha}$$

$$\Rightarrow \frac{\delta H}{\delta R_\alpha} = i\frac{\partial H}{\partial \alpha}.$$

In this case it is just a change in notation, but represents a more general operation.

The full Hamiltonian for the planar problem can now be expressed as,

$$H = \frac{1}{2} < \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix}, \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix} >_{G^{-1}(R_\beta)} + V(R_\alpha, R_\beta), \tag{2.2}$$

where the metric matrix G is computed in the Appendix, A.1.

$$G = \begin{pmatrix} \mathbb{I}_1 + \mathbb{I}_2 + (l_1^2 + l_2^2)m_2 + l_1 l_2 m_2(R_\beta + \overline{R_\beta}) & \mathbb{I}_2 + l_2^2 m_2 + l_1 l_2 m_2 \frac{1}{2}(R_\beta + \overline{R_\beta}) \\ \mathbb{I}_2 + l_2^2 m_2 + l_1 l_2 m_2 \frac{1}{2}(R_\beta + \overline{R_\beta}) & \mathbb{I}_2 + l_2^2 m_2 \end{pmatrix}.$$

For computational purposes we wish to express the derivative with respect to R_β of the kinetic matrix as well as the derivative of its inverse in terms of the inverse and the derivative matrix. We need to understand some basic derivatives

$$< v, \frac{\delta(R_\beta + \overline{R_\beta})}{\delta R_b} >= \lim_{\epsilon \to 0}\frac{d}{d\epsilon}(e^{i\beta+v\epsilon} + e^{-i\beta-v\epsilon}) = -\overline{v}(R_\beta - \overline{R_\beta}).$$

This derivative takes real numbers to imaginary and imaginary to reals, $\cos \to -i\sin, \sin \to i\cos$. The derivative operator is also linear and passes term by term to G. For the inverse, which appears in the Hamiltonian, there is a formula for the derivative of a matrix inverse,

$$\frac{\delta G^{-1}}{\delta R_b} = -G^{-1}\frac{\delta G}{\delta R_b}G^{-1}.$$

Now the full equations of motion are $R_\alpha^{-1}R_\alpha = e, \quad R_\beta^{-1}R_\beta = n$ and $\begin{pmatrix} e \\ \\ n \end{pmatrix} = G^{-1}\begin{pmatrix} \eta \\ \\ \nu \end{pmatrix}$,

$$\begin{aligned}
\dot{\eta} &= \chi_H^{\eta}(R_\alpha, R_\beta) \\
&= -\frac{\delta H}{\delta R_\alpha} \\
&= -\frac{g}{2}((R_\alpha - \overline{R_\alpha})(\frac{l_1}{2}m_1 + m_2 l_1) + (R_\alpha R_\beta - \overline{R_\alpha R_\beta})m_2 l_2)) \\
\dot{\eta} &= -ig(\sin(\alpha)(\frac{l_1}{2}m_1 + m_2 l_1) + \sin(\alpha+\beta)m_2 l_2)), \qquad (2.3)
\end{aligned}$$

$$\begin{aligned}
\dot{\nu} &= \chi_H^{\eta}(R_\alpha, R_\beta, \eta, \nu) \\
&= -\frac{1}{2}\begin{pmatrix} \eta & \nu \end{pmatrix} G^{-1}\frac{\delta G}{\delta R_\beta}G^{-1}\begin{pmatrix} \eta \\ \nu \end{pmatrix} - \frac{g}{2}(R_\alpha R_\beta - \overline{R_\alpha R_\beta})m_2 l_2) \\
\dot{\nu} &= -\frac{1}{2}\begin{pmatrix} \eta & \nu \end{pmatrix} G^{-1}\frac{\delta G}{\delta R_\beta}G^{-1}\begin{pmatrix} \eta \\ \nu \end{pmatrix} - ig\sin(\alpha+\beta)m_2 l_2. \qquad (2.4)
\end{aligned}$$

2.2.1 Control

Adding the external torques our equations for motions become,

$$\dot{\nu} = \chi_H^{\nu} + \mu, \qquad \dot{\eta} = \chi_H^{\eta} + \gamma - \mu.$$

In this way the change in total angular momentum is due to the torque on the first joint. Again, the cost per time we use is simply $c_1 = 1 - \frac{1}{2}(a\gamma^2 + b\mu^2)$. Here the torques are once again imaginary as they are in the tangent space of the imaginary angular momentum. Our adjoint variables are, $p_\alpha, p_\beta, p_\eta, p_\nu$. The new control Hamiltonian is a real valued function on $T^*T^*M \times (i\mathbb{R})^2$, where the imaginary real lines are for the control space.

$$\mathcal{H} = -(1 - \frac{1}{2}(a\gamma^2 - b\mu^2)) + \langle \begin{pmatrix} p_\alpha \\ p_\beta \end{pmatrix}, G^{-1}\begin{pmatrix} \eta \\ \nu \end{pmatrix} \rangle$$

$$+\Big\langle \begin{pmatrix} p_\eta \\ p_\nu \end{pmatrix}, \begin{pmatrix} \chi_H^\eta(R_\alpha, R_\beta, \eta, \nu) + \gamma - \mu \\ \chi_H^\nu(R_\alpha, R_\beta, \eta, \nu) + \mu \end{pmatrix} \Big\rangle.$$

The inner products are complex, between purely imaginary variables. The complex nature comes in once more as their is a negative sign in front of all the partial derivatives with respect to any of these variables. While slightly confusing, this is necessary to continue with our treatment of the rotations and understanding the sign changes help in generalization to dual vectors for a more complex algebra. For an optimal trajectory we can use the extra freedoms to set this to be zero,

$$\mathcal{H}(R_\alpha, R_\beta, \eta, \nu, p_\alpha, p_\beta, p_\eta, p_\nu, \gamma^*, \nu^*) = 0.$$

For the exactly precise kick, we specify all but n at the terminal condition. This leaves an interesting relationship for η, ν and a transversality condition for p_η, p_ν,

$$\begin{pmatrix} \eta(t_1) \\ \nu(t_1) \end{pmatrix} = G|_{t_1} \begin{pmatrix} 0 \\ n(t_1) \end{pmatrix}, \quad \Big\langle \begin{pmatrix} p_\eta(t_1) \\ p_\nu(t_1) \end{pmatrix}, \begin{pmatrix} \eta(t_1) \\ \nu(t_1) \end{pmatrix} \Big\rangle = 0$$
$$\Rightarrow \begin{pmatrix} p_\eta(t_1) \\ p_\nu(t_1) \end{pmatrix} = G^{-T} \begin{pmatrix} x \\ 0 \end{pmatrix}. \tag{2.5}$$

To determine conditions of maximality for the instantaneous Hamiltonian, consider the partials with respect to the control torques:

$$\frac{\partial \mathcal{H}}{\partial \gamma} = a\gamma - p_\eta = 0,$$
$$\frac{\partial \mathcal{H}}{\partial \mu} = b\mu - p_\nu + p_\eta = 0.$$

The only critical point must be a maximum because of the quadratic nature, so we obtain the optimal forces directly from the co-state variables. Manipulating the end condition

yields

$$\mathcal{H}|_{t_1} = -1 - n(t_1)p_\beta(t_1) - (\chi_H^\eta(t_1) + \frac{\gamma(t_1)}{2} - \frac{\mu(t_1)}{2})p_\eta(t_1) - (\chi_H^\nu(t_1) + \frac{\mu(t_1)}{2})p_\nu(t_1) = 0.$$

It comes down to that we need to determine 3 unknowns, one decides $p_\eta(t_1), p_\nu(t_1)$, and two more for $p_\alpha(t_1), n(t_1)$. $p_\beta(t_1)$ is determined using zero for the Hamiltonian. We can make estimates, then we must calculate the trajectory and try to improve the estimate until we come up with an optimal trajectory that fits our initial as well as terminal trajectory. The difficulty comes from the need to work both backwards as well as forwards. If we knew the right initial adjoint boundary conditions, we could immediately compute the optimal trajectory, but all we have is very limited knowledge of the end condition. Here we rely on intuition in that we only need to consider the trajectories that lie on the important part of the manifold between the initial and end conditions.

The equations governing the evolution of the adjoint control variables are now expressed using the control Hamiltonian, with the more complex terms derived in A.2:

$$\dot{p_\alpha} = -\frac{\delta(\chi_H^\eta p_\eta + \chi_H^\nu p_\nu)}{\delta R_\alpha} = g < \begin{pmatrix} p_\eta \\ \\ p_\nu \end{pmatrix}, \begin{pmatrix} \cos(\alpha)(\frac{l_1}{2}m_1 + m_2 l_1) + \cos(\alpha+\beta)m_2 l_2) \\ \\ \cos(\alpha+\beta)m_2 l_2) \end{pmatrix} >, \tag{2.6}$$

$$\begin{aligned} \dot{p_\beta} = -\frac{\delta\mathcal{H}}{\delta R_\beta} = \quad & g < \begin{pmatrix} p_\eta \\ \\ p_\nu \end{pmatrix}, \begin{pmatrix} \cos(\alpha+\beta)m_2 l_2) \\ \\ \cos(\alpha+\beta)m_2 l_2) \end{pmatrix} > \\ & + \frac{1}{2} < \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix}, (-2G^{-1}\frac{\delta G}{\delta R_\beta}G^{-1}\frac{\delta G}{\delta R_\beta}G^{-1} + G^{-1}\frac{\delta^2 G}{\delta R_\beta^2}G^{-1}) \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix} > p_\nu \\ & - < \begin{pmatrix} p_\alpha \\ \\ p_\beta \end{pmatrix}, G^{-1}\frac{\delta G}{\delta R_\beta}G^{-1} \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix} >, \end{aligned} \tag{2.7}$$

$$\dot{p_\eta} = -\overline{\frac{\partial(ep_\alpha + np_\beta + \chi_H^\nu p_\nu)}{\partial \eta}} = \quad - < G^{-T} \begin{pmatrix} p_\alpha \\ \\ p_\beta \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 0 \end{pmatrix} >$$

$$- < \begin{pmatrix} p_\nu \\ \\ 0 \end{pmatrix}, G^{-1} \frac{\delta G}{\delta R_\beta} G^{-1} \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix} >, \qquad (2.8)$$

$$\dot{p_\nu} = -\overline{\frac{\partial(ep_\alpha + np_\beta + \chi_H^\nu p_\nu)}{\partial \nu}} = \quad - < G^{-T} \begin{pmatrix} p_\alpha \\ \\ p_\beta \end{pmatrix}, \begin{pmatrix} 0 \\ \\ 1 \end{pmatrix} >$$

$$- < \begin{pmatrix} 0 \\ \\ p_\nu \end{pmatrix}, G^{-1} \frac{\delta G}{\delta R_\beta} G^{-1} \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix} > . \qquad (2.9)$$

The first partial derivatives are straight forward to compute, as there is only linear dependance in the terms of the previous Hamiltonian vector fields. The second is much more complex, as there is dependance of R_β in the inverse metric tensor, necessary to express the tangent vectors. There is also only linear dependance for the momentum vector. As always, the adjoint system is linear with respect to the co-state variables. Note that the inner product appearing in these equations is just a vector inner product as a convenience of notation, the output in all of them is actually an purely imaginary number.

For an algorithm converging to an optimal trajectory, we need a distance function for when non-optimal trajectories terminate. They terminate if Im(e) < 0, or if $0 < \beta < \pi$, or if $\pi/2 < \alpha < 0$, or if n crosses zero twice. The distance function can be made of the form, $\cos(\alpha(t_0))^2 + \sin(\beta(t_0))^2 + |e(t_0)|^2 + |n(t_0)|^2$.

One more important piece of mechanics comes in again. We need to calculate the work done by the external torques to compare with the change in the Hamiltonian. This is quite simple because the energy added is just $\int -e\gamma dt$ if we only consider one force on the

first rotation. To show this, suppose ΔH is the change in total energy in the presence of external torques,

$$\Delta H = -\frac{1}{2}\begin{pmatrix} \eta+\Delta\gamma-\Delta\mu & \nu+\Delta\mu \end{pmatrix} G^{-1}\begin{pmatrix} \eta+\Delta\gamma-\Delta\mu \\ \nu+\Delta\mu \end{pmatrix}.$$

Collect the first order terms, and utilize the symmetry of G,

$$\Delta H = -\begin{pmatrix} \Delta\gamma-\Delta\mu & \Delta\mu \end{pmatrix} G^{-1}\begin{pmatrix} \eta \\ \nu \end{pmatrix} =< \begin{pmatrix} \Delta\gamma-\Delta\mu \\ \Delta\mu \end{pmatrix}, \begin{pmatrix} e \\ n \end{pmatrix} > .$$

Although it is not addressed in this problem we are solving, there is also a desire to solve the problem with an emphasis on the power delivered by the kick. To do this we would add a negative terminal cost, capturing the kinetic energy of the kick, $C_2 = -w\frac{1}{2}|\nu(t_1)|^2$. We would not want to use n just because to put it in terms of the co-tangent variables involves the inverse metric and both η, ν. To incorporate this factor we must differentiate and add it to the instantaneous cost with the term

$$c_2 = w\nu(\chi_H^\nu - \gamma + \mu).$$

This would not add much difficulty, but is not essential for understanding the problem.

An alternative approach to the controls is to not specify the end condition as a constraint, but to work it into the cost function. The terminal cost has the terms, $\frac{1}{2}d(\cos(\alpha)^2 + \sin(\beta)^2 + |e|^2)$, evaluated at the end time, t_1. In terms of the rotations, $d\frac{1}{8}((R_\alpha + \overline{R_\alpha})^2 + |R_\beta - \overline{R_\beta}|^2 + |e|^2)$. The time derivative adds to the instantaneous cost

$$c' = d\frac{1}{4}(e(R_\alpha - \overline{R_\alpha})(R_\alpha + \overline{R_\alpha}) - n(R_\beta + \overline{R_\beta})(R_\beta - \overline{R_\beta}) + \frac{d|e|^2}{dt}).$$

The last term is a little more complicated as it must be formulated in terms of the momentum variables and involves the inverse metric. Now, however, we can set the terminal condition

for all of the co-state variables to be zero. There is only one undetermined degree of freedom, which is the terminal condition for n. We should only have to find the value $n(t_1)$, for which the optimal trajectory with respect to the augmented cost passes through the initial condition. With this approach we can make d as large as we like to get a more accurate kick. We would need d to be relatively large because we are no longer guaranteed there is a solution hitting the initial condition as we have allowed the exact end state to vary. Making d too big could generate numerical problems in the approximation of the backwards evolution, but as noted these are trade-offs in complexities of the problem.

2.3 Full Roundhouse Kick Problem

For the most part, the notation should carry through from the previous problem, but now the variables we are dealing with are slightly different creatures. The tangent space to the rotations is now going to be the bi-vector algebra from the geometric algebra of 3D space. The rotors also form a subgroup found in the algebra, as discussed in the introduction, Section 1.1.4. The action of these rotations is no longer simple multiplication, but by the adjoint action or conjugation. The inertia tensors are now linear transformations on the algebra of bi-vectors. When the notation calls for a rotated inertia tensor, what is meant is that the frame of the tensor is rotated. When deriving the energy in terms of vectors, the vectors are simply incorporated as the vector elements of the geometric algebra. Due to the nature of the boundary conditions we must choose a basis. The vector basis is denoted by, $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$. The third vector is vertical, the direction of the second vector is the end condition for the kick, which ends with a counter clockwise rotation of the second limb if looking from above. The basis for the bi-vectors is $\{\hat{b}_3 = \hat{e}_1\hat{e}_2 = I\hat{e}_3, \hat{b}_1 = \hat{e}_2\hat{e}_3, \hat{b}_2 = \hat{e}_3\hat{e}_1\}$.

We always consider the second rotation to be in the plane of $\hat{b}_2$. The backwards ordering, is to reconnect these vectors with the plane of rotation and can also be reached by multiplying the basis vector with the pseudo-scalar, I. For the initial conditions, both rotations are at the identity. The reason that the initial plane of rotation for the second limb is perpendicular to the plane joining the initial and end conditions of the leg is an attempt to mimic the standard 'backward leaning stance' of Taekwondo from which this kick is generally thrown. In this stance the back, kicking leg, is actually point more then 90^o from the target. The final condition for the second rotation is back to the identity. The terminal condition first rotation can be constructed by considering the two rotations necessary to reach the end condition and multiplying the rotors on the left in the order that they are carried out,

$$R_\alpha(t_1) = e^{\frac{-\pi}{2}\hat{b}_2} e^{\frac{\pi}{4}\hat{b}_1} = -\frac{1}{\sqrt{2}}(\hat{b}_3 + \hat{b}_2).$$

The total energy can be expressed in terms of the rotational energy and the kinetic energy of the center of mass of the second limb,

$$\begin{aligned} H = \quad & \frac{1}{2} < e, R_\alpha(\mathbb{I}_1)(e) > + \frac{1}{2} < (e+n), R_\alpha(R_\beta(\mathbb{I}_2))(e+n) > + \frac{1}{2} m_2 |\dot{\mathbf{cm}}_2|^2 (R_\alpha, R_\beta, e, n) \\ & + g < m_1 \mathbf{cm}_1(R_\alpha) + m_2 \mathbf{cm}_2(R_\alpha, R_\beta), \hat{e}_3 > . \end{aligned} \tag{2.10}$$

For simplicity, we take the eigenvectors of the inertia tensor to be the x, y, z axis, and, moreover, to be symmetric for rotations around the x, y axis like a cylinder. The only difference between the first and second inertia tensors a factor of 4 for the rotations of the full length of the limb because the first limb is rotating around an end and the second is rotating around its center of mass,

$$\mathbb{I}(\hat{b}_3) = I_s \hat{b_3}, \quad \mathbb{I}(\hat{b}_2) = I_l \hat{b}_2, \quad \mathbb{I}(\hat{b}_1) = I_l(\hat{b}_1).$$

We would like to express this linear transformation with elements of the algebra

$$\mathbb{I}(B) =< B, \hat{b}_3 > I_s \hat{b}_3 + < B, \hat{b}_2 > I_l \hat{b}_2 + < B, \hat{b}_1 > I_l \hat{b}_1.$$

The inner products here are pairings with dual bi-vectors so using the inner product defined in Eq. (1.9) on the Clifford algebra these are actually, $B \cdot \overline{\hat{b}_3}$. As these are always pure bi-vectors, the conjugate here is simply the negative. The geometric algebra inner product of two bi-vectors pointing in the same directionally is negative, so this bracket inner product returns positive scalar output for bi-vectors of the same orientation. A rotation of this inertia tensor then corresponds to a rotation of the bi-vector basis and is naturally equivalent to a rotation of the end result with an inversely rotated input,

$$R(\mathbb{I})(B) = R(\mathbb{I}(R^{-1}(B))).$$

We seek to represent the kinetic energy as a metric, such that we can express the Hamiltonian with the inner product of the metric,

$$H = \frac{1}{2} \Big\langle \begin{pmatrix} e \\ \\ n \end{pmatrix}, \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix} \Big\rangle + V(R_\alpha, R_\beta),$$

where η, ν are generalized angular momentum bi-vectors in the co-tangent space, reached by a linear transformation G of e, n. G is the momentum map of the tangent bi-vectors, as well as the symmetric, positive definite kinetic energy tensor, if we wish to deal with the bi-vectors as four-vectors in the orthonormal basis for the problem. The equations are complicated and can be derived from the higher expression with the expression of the center of mass expressed. In the following notation consider the R, rotations in the group, as the actions on the space $\mathbb{R}_3$. The vectors, $\mathbf{r}_1, \mathbf{r}_2$, represent the two limbs, with the center of mass at half of the distance. The vectors initialize pointing down, in the $-\hat{e}_3$ direction. The

vector to the center of mass of the second limb is

$$\mathbf{cm}_2 = R_\alpha(\mathbf{r}_1 + R_\beta(\frac{\mathbf{r}_2}{2})).$$

The boldface $\mathbf{r}$ vectors are constants of the system, so the derivative involves a product rule only on the rotations. To do this we shall make use of a trivialization of the Lie group. The representation is the adjoint or conjugate action on the geometric algebra, $R_\alpha(\mathbf{v}) \to w_\alpha v w_\alpha^{-1}$, where w_α is a unitary element that is a scalar plus a bi-vector. This group is isomorphic to $SU(2)$, also the unit quaternions, which is a double cover of the rotations we want, $SO(3)$. The double cover does not pose any problem, we just recognize that $R_\alpha, -R_\alpha$ are equivalent rotations. The trivialization we are using for our variables yields

$$\dot{w_\alpha} = -\frac{1}{2} w_\alpha e, \quad \dot{w_\beta} = -\frac{1}{2} w_\beta n.$$

While these choices appear strange, this trivialization has a nice geometric significance and leads to some simplifications in calculation. The choice generates a family of one-parameter subgroups, $R_\alpha(t) = R_{0\alpha}(e^{\frac{-1}{2}et}) \to w_{0\alpha} e^{\frac{-1}{2}et}(\cdot) e^{\frac{1}{2}et} w_{0\alpha}^{-1}$. The derivative of the inverse is

$$\dot{w_\alpha^{-1}} = \frac{1}{2} e w_\alpha^{-1}.$$

The algebra element for the β rotation is in a single plane, so only needs a single dimension of bi-vector, $\hat{b}_2$, isomorphic to the complex numbers. The derivation for the matrix for the kinetic energy metric is in Eq. (A.3). The equations look the most neat when they are in matrix form,

$$H = -\frac{1}{2} \begin{pmatrix} \eta & \nu \end{pmatrix} G^{-1}_{R_\alpha, R_\beta}(\begin{pmatrix} \eta \\ \\ \nu \end{pmatrix}) + V(R_\alpha, R_\beta).$$

In this equation η is a three dimensional bi-vector, while ν is a bi-vector of a single dimension. The matrix multiplication can be considered as an inner product that is to be taken with the matrix positive definite matrix G^{-1} along with a flip in sign as one of the bi-vectors is the conjugate for the minus sign. The matrix G is symmetric and is composed of three components. G_e represents the linear transformation in the 3D bi-vector space, and G_n is the transformation in the 1D space. G_{ne} is a transformation from the 1D space to the 3D space, appearing as a column in the matrix

$$G = \begin{pmatrix} G_e & G_{ne} \\ G'_{ne} & G_n \end{pmatrix}.$$

For computational purposes we map this transformation onto a matrix of real vectors. This is necessary to calculate numerically the inverse, although there may be more efficient manner to find the inverse using the additional algebraic structure. This is a little bit more difficult because we have a matrix of transformations in the language of geometric algebra. While the components in the above equation can be expressed naturally as geometric algebra transformations, the inverse would be much more difficult. In coordinate free form the transformations come from the equations worked out before. The transformations that make of G are

$$\begin{aligned} G_e(e) &= R_\alpha(\mathbb{I}_1)(e) + R_\alpha(R_\beta(\mathbb{I}_2))(e) + m_2(r_1 \wedge [e, r_1] + R_\beta(r_2) \wedge [e, R_\beta(r_2)] + 2(r_1 \wedge [e, R_\beta(r_2)]), \\ G_n(n) &= R_\alpha(R_\beta(\mathbb{I}_2))(n) + m_2 r_2 \wedge [n, r_2], \\ G_{ne}(n) &= \frac{1}{2}(R_\alpha(R_\beta(\mathbb{I}_2))(n) + m_2(r_1 \wedge R_\beta([n, r_2]) + R_\beta(r_2) \wedge R_\beta([n, r_2])). \end{aligned} \tag{2.11}$$

To really get to the equations of motion, we need to take a variety of derivatives. When we are working with the representation as an adjoint action the partial derivatives,

$\frac{\delta}{\delta R_\beta}$ and $\frac{\delta}{\delta w_\beta}$, have equivalent meaning. $\frac{\delta}{\delta w_\beta}$ becomes an element in the co-tangent space of the planar rotation through the natural pairing,

$$< a, \frac{\delta f(w_\beta)}{\delta w_\beta} >= \frac{d}{dt} f(w_\beta e^{-\frac{1}{2}at})|_{t=0}.$$

For example for an arbitrary term in the kinetic energy,

$$< a, \frac{\delta}{\delta w_\beta}(< e, r_1 \wedge w_\beta [n, r_2] w_\beta^{-1} >) >= \frac{d}{dt}(< e, r_1 \wedge w_\beta e^{\frac{-1}{2}at}[n, r_2] e^{\frac{1}{2}at} w_\beta^{-1} >)$$

$$=< e, r_1 \wedge w_\beta (\frac{-1}{2} a[n, r_2] + [n, r_2]\frac{1}{2}a) w_\beta^{-1} >=< e, r_1 \wedge w_\beta ([[n, r_2], a]) w_\beta^{-1} > .$$

We can transpose the transformation to get $< [\overline{n}, r_2] \wedge w_\beta^{-1}[r_1, e] w_\beta, a >$, such that

$$\frac{\delta}{\delta w_\beta}(< e, r_1 \wedge w_\beta [n, r_2] w_\beta^{-1} >) = [r_2, n] \wedge w_\beta^{-1}[r_1, e] w_\beta.$$

We also need the potential energy,

$$\begin{aligned} V &= < w_\alpha (m_1 \frac{r_1}{2} + m_2(r_1 + w_\beta r_2 w_\beta^{-1})) w_\alpha^{-1}, e_3 >, \\ \frac{\delta V}{\delta w_\alpha} &= -(m_1 \frac{r_1}{2} + m_2(r_1 + w_\beta r_2 w_\beta^{-1})) \wedge w_\alpha^{-1} e_3 w_\alpha \in T^*M, \\ \frac{\delta V}{\delta w_\beta} &= -m_2 r_2 \wedge w_\beta^{-1} e_3 w_\beta \in T^*M. \end{aligned} \tag{2.12}$$

As the second rotation, R_β, is restricted to a plane, the derivative is restricted to a single dimension in the cotangent space. The derivatives on the linear transformation G have a slightly more complex flavor. For conciseness in notation, let π represent the vector with η and ν. We need to know

$$\frac{\delta < \pi, G^{-1}\pi >}{\delta w_\alpha} \in T^*M.$$

We could calculate it term by term, but there may be a way that it passes into the linear transformation. Some sample calculations are done in A.3. We can write out the term for

R_α, which is relatively simple,

$$\begin{aligned}\frac{\delta H}{\delta w_\alpha} &= \frac{1}{2}\frac{\delta}{\delta w_\alpha}\langle \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix}, G^{-1}\begin{pmatrix} \eta \\ \\ \nu \end{pmatrix}\rangle + \frac{\delta V}{\delta w_\beta} \\ &= -\begin{pmatrix} 1 & 1 \end{pmatrix}[R_\alpha^{-1}(G^{-T}\begin{pmatrix} \eta \\ \\ \nu \end{pmatrix}), R_\beta(\mathbb{I})(G^{-1}\begin{pmatrix} \eta \\ \\ \nu \end{pmatrix})] - m_2 r_2 \wedge w_\beta^{-1} e_3 w_\beta,\end{aligned}$$

where $\mathbb{I}$ is a more generalized moment of inertia, $\begin{pmatrix} \mathbb{I}_1(e) + \mathbb{I}_2(e+n) \\ \\ \mathbb{I}_2(n) \end{pmatrix}$. Of course in all the equations if there is a term with G^T this is nothing but G as it is symmetric, the transpose is only included to give an extra hint about where the term is originating. The term for R_β is much more complex as it appears in almost every term and in different ways. Since all the operations are linear, there is nothing that we cannot handle with this strategy. However, it is much more convenient to only consider this term as the abstract transformation. Since we know the transformation is linear and has a matrix transformation in terms of the basis, we can us all the machinery of linear transformations. The formula

$$\frac{\delta G^{-1}}{\delta R_\beta} = -G^{-1}(\frac{\delta G}{\delta R_\beta}(G^{-1}))$$

still holds. The transformations in terms of the geometric algebra terms are complicated as we would use the product rule to differentiate separately the terms of multiplication on the left and on the right, just as the transformations in terms of a basis are complicated.

The adjoint operator also appears in Hamilton's equations. Using the representation of the group with an adjoint action on the Clifford algebra, and right trivialization, the adjoint appears simply as operations in the algebra. The linearization of the adjoint leads

to a convenient expression

$$ad_n m = \frac{de^{-\frac{mt}{2}} n e^{\frac{mt}{2}}}{dt} = \frac{1}{2}(-mn + nm) = [n, m],$$

$$< m, ad^*_\gamma \eta >=< [\gamma, m], \eta >=< m, [\eta, \gamma] >,$$

$$ad^*_\gamma \eta = [\eta, \gamma]. \tag{2.13}$$

As promised there is no further complication from this term. We write out the final equations of motion from Eq. (1.1),

$$\begin{aligned} -2 \begin{pmatrix} w_\alpha \dot{w_\alpha} \\ \\ w_\beta \dot{w_\beta} \end{pmatrix} &= \begin{pmatrix} \frac{\delta H}{\delta \eta} \\ \\ \frac{\delta H}{\delta \nu} \end{pmatrix} = G^{-1} \begin{pmatrix} \eta \\ \\ \nu \end{pmatrix} = \begin{pmatrix} e \\ \\ n \end{pmatrix}, \\ \dot{\eta} &= -\frac{\delta H}{\delta w_\alpha} + [\eta, \frac{\delta H}{\delta w_\alpha}], \\ \dot{\nu} &= -\frac{\delta H}{\delta w_\beta}. \end{aligned} \tag{2.14}$$

2.3.1 Final Controls

The equations for an optimal trajectory for this full problem are nearly identical to the equations in Section 2.2.1. Some of the terms are more complicated, and for p_α, the co-adjoint term must also be added, but this comes simply from the commutator. The complicated terms from the R_β derivative are now more complicated, and they no longer dissapear for R_α. There is no added difficulty in solving for the optimal controls as the same partial derivative can be taken to get conditions for criticality, which is equivalent to maximality again in this case. It is hopefully apparent to the reader that the derivation of the equations for an optimal trajectory of the control problem is a straight forward although complex process. Nothing in the derivation of the equations could not be carried out by computers although I do not believe current computer algebra programs are very

advanced with manipulating equations of non-abelian algebras. I suggest that using the Clifford algebras is a good way to approach this problem as all the computations are done in the same framework.

However, this approach to optimal control is significantly more difficult for the full problem due to the extra dimensions. To find the correct terminal conditions for the adjoint system in the planar problem was difficult and it only required a search in a 3 dimensional space. Now that space would be 7 dimensions and there is less intuition to guide the search.

Approximation of the larger problem using the planar model may be useful. Coming up with an approximate answer will at least help, as it gives a starting point from where we can find an optimal solution. In Lie group theory the maximal torus of the manifold plays a large role, so we might presume the model restricted to this torus is also useful here. There is a very important result in Lie group theory, which can guide us for a simpler way to approximate the solution to the big problem. For any compact Lie group, every element in the group is conjugate to an element on a maximal torus of the group. We can use this because the planar model is the maximal torus for our 3D problem. Restricted to the single plane of motion we would have the same planar problem that we analyzed. Intuitively, the solution to the problem, if the end condition was also on on the plane, would stay on the plane as any extra motion would only increase the cost. We can also think that a smooth change in the terminal condition should lead to a smooth change in the optimal trajectory. Using these two facts, we must come up with an approximation architecture to map the optimal trajectory on the torus to the full 3-D problem. For the Lie group elements, we can do this by simply conjugating by a group element that varies smoothly and whose initial and end conditions satisfy the problems boundary conditions. The Lie algebra elements can

also be rotated on or off the torus. We should be able to get some further bounds as we expect the optimal motion to not have any extra deviations, i.e. few critical points for any of the derivatives of the conjugating element.

The Hamiltonian vector fields can also be projected onto the torus to become functions of the rotations on the torus and the co-tangent vectors, approximating the three dimensional dynamics. These vector fields will be part of the new input in the planar optimal control problem for approximating a full dimensional optimal control trajectory. From this projected approximation, we should be able to get an idea of how the momentum vectors transform through the conservation of energy and momentum. This is slightly complicated by the control torques, which disrupt this conservation, but we also get a new conservation from the control Hamiltonian. To handle this complication, we can use rescaling and sample trajectories to guide the approximation. In theory, a dense approximation architecture might give us a method to algorithmically converge to an optimal solution [2]. This method can be enhanced by computing the control adjoint vectors and control Hamiltonian to look for local optimality of solutions, as well as sampling trajectories and incorporating information from the costs in the approximation architecture; a method known as reinforcement learning. Finding a way to incorporate all the different approaches, I believe, has great potential in developing a robust approach to difficult non-linear control problems. However, as they all also pose difficulties in implementation this cannot be fully explored in this thesis.

3

Results

The equations obtained are very non-linear and are not eligible for an analytic solution. The analysis can be done through computer simulations of the dynamical system. The simulation of the planar system is straight forward. All of the programming is done using Matlab. In the algorithm used for approximating the integral we would like to maintain conservation of the Hamiltonian and any other conserved quantities. The algorithm used mixes a basic second order method with stochastic optimization to ensure conservation. The tangent vectors are computed at the point and then again at a partially evolved point. The two tangent vectors are then added together with a set of randomly generated weighting factors, and the vectors that best conserve the Hamiltonian are chosen. This methods incorporates the information from the immediate tangent vector, as well as tangent vector at a slightly displaced location, picking up some of the second order dynamics. As we do not know exactly where the evolution will take the system, this stochastic process of evolution allows us to on average get very close to the sub-manifold determined by the conservation laws. Figure 3.1, shows an uncontrolled trajectory on the torus manifold. Figure 3.2 shows

Uncontrolled Chaotic Trajectory on Torus

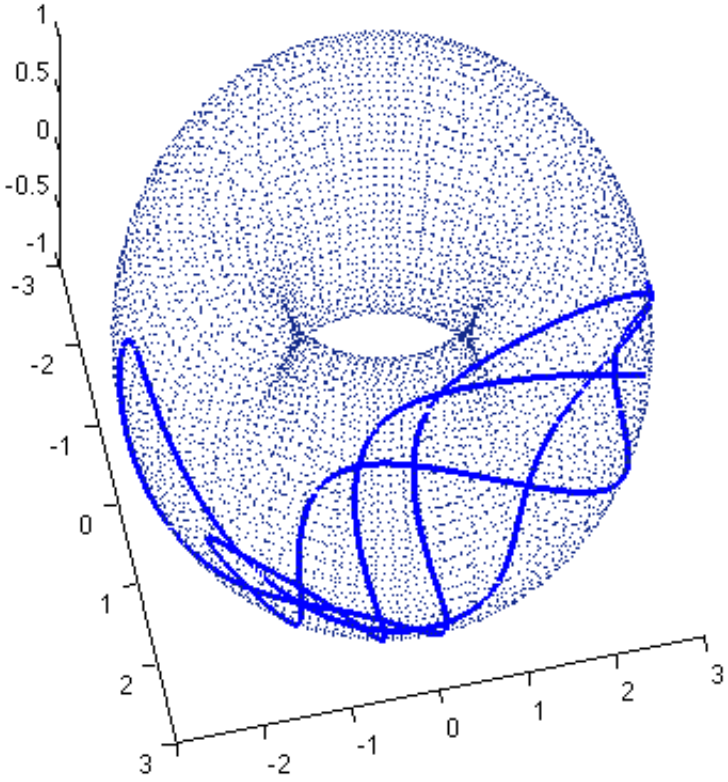

Figure 3.1: This is an uncontrolled trajectory initiated from a stationary horizontal position.

the conservation of the Hamiltonian. It is very near to zero, with blips on the order of 10^{-5} when the kinetic energy is at a max.

When the controls are added the computation is slightly more difficult. The method used is to calculate optimal trajectories backwards. We do this by calculating trajectories for a variety of terminal conditions for the co-state variables. To get optimal trajectories I limit the trajectory to a subset of the full manifold and tangent bundle and kill trajectories that go outside, avoiding unnecessary computations. A distance function is

Hamiltonian for trajectory in Figure 3.1

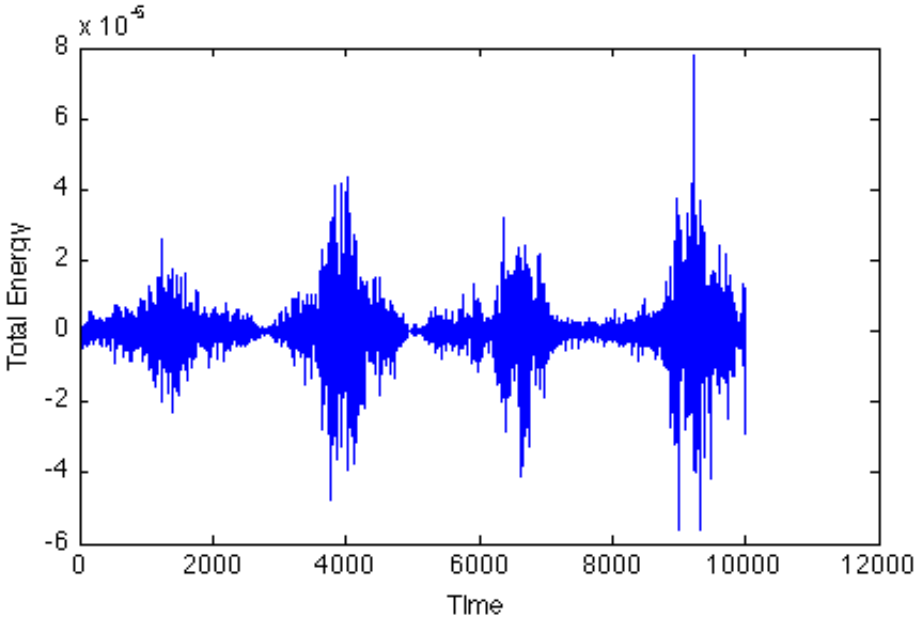

Figure 3.2: The Hamiltonian demonstrates that energy is very nearly conserved except for some numerical distortion.

used to check how far from our desired initial conditions the trajectories died. The minimal distance is used as a seed to randomly generate new terminal conditions to get closer to the correct initial conditions. An unexpected complication was that not all the end conditions that set the control Hamiltonian kept it at zero. Many conditions would immediately jump to a different value and stay near there. I am not sure what other constraint could be used to narrow those out, the sub-manifold of the good end conditions appears to be a 2-D non-linear surface. A plot of points near or on that surface is withheld as it is very difficult to recognize the surface in two dimensions. One possibility to recover the constraint onto this surface would be to set the time derivative of the Hamiltonian to zero, but this is an impractical calculation. Incorporating the possibility of non-conservation of the Hamiltonian into the distance function works to narrow out the end conditions off the surface. The search is

Data for trajectories in Figure 3.3

Color	End Condition	Cost	Distance
Blue	4.98, 0.84, -0.04	0.84	0.57
Green	5.23, 0.88, -0.10	0.80	1.00
Red	5.10, 0.85, -0.07	0.84	0.59
Teal	6.08, 0.99, -0.15	0.72	0.97
Pink	3.58, 0.61, 0.04	1.00	2.25

Table 3.1: The values of constants used for these trajectories were, $l_1 = 1, l_2 + 0.5, m_1 = m_2 = 1, g = 0.98, a = 1, b = 0.5$.

particularly difficult as the terminal condition space exhibits both sensitive dependance on small changes as well as near optimal conditions in a fairly wide range. Varying by as much as 0.01 could totally throw off the trajectory, where as close trajectories could be found at a distance of around 3.

The table provides the data for the five trajectories graphed. The end conditions that generated the condition are given. The first number is the end condition for the tangent vector, n. The second is for p_α, and the third gives the end conditions for p_η, p_ν according to the transversality condition.

It took a fair amount of tinkering to get the values close, but once it was close the distance function did its job to use the iteration to find the optimal solution. The distance function I use is

$$d = 25\sin(\alpha(0))^2 + 10\sin(\beta(0))^2 + 10(-\cos(\beta(0)))^2 + 0.1|\eta(0)|^2 + 0.1|\nu(0)|^2 + \mathcal{H}(0)^2,$$

where the $\cos(\beta)$ is only added if it is negative. The weighting factors only really mattered to relate all the terms on the same order of magnitude. Figure 3.3 shows a variety of trajectories that came up in the search for optimality. Really all of the trajectories solved for were optimal, simply with different starting positions.

From a Taekwondo perspective, some of the close trajectories had good technique,

Optimal Trajectories on the Torus

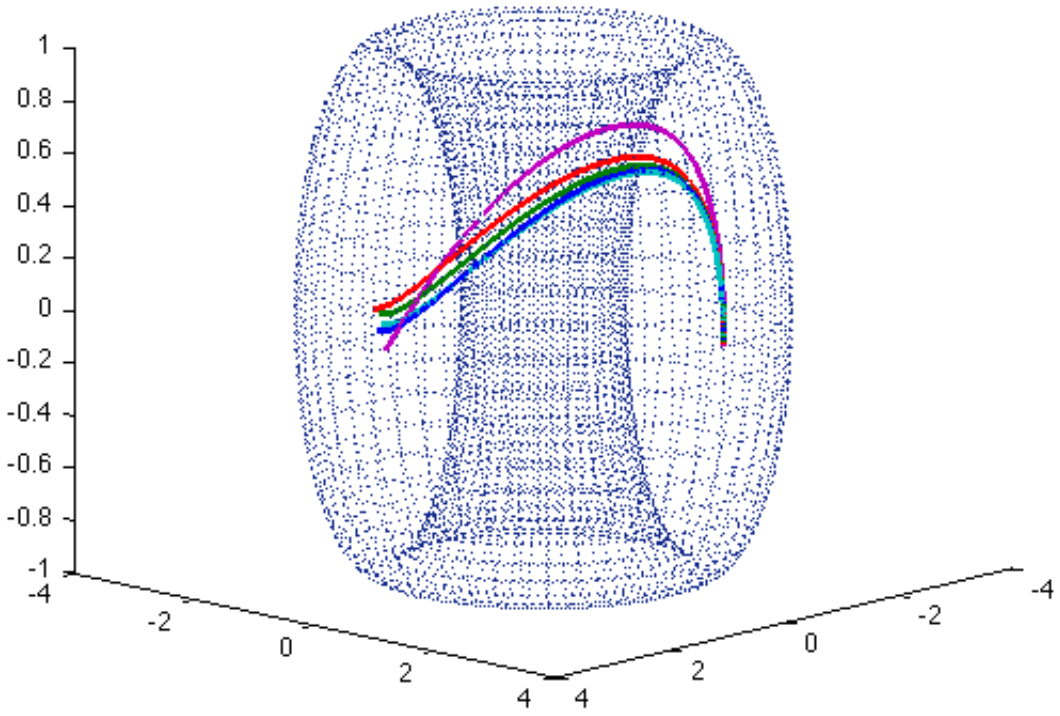

Figure 3.3: Five nearly-optimal trajectories. The big circle on the torus represents the rotation of the first, hip, joint. The trajectory goes from left to right, starting near zero and ending with only rotation of the second, knee, joint.

Diagram for a kick

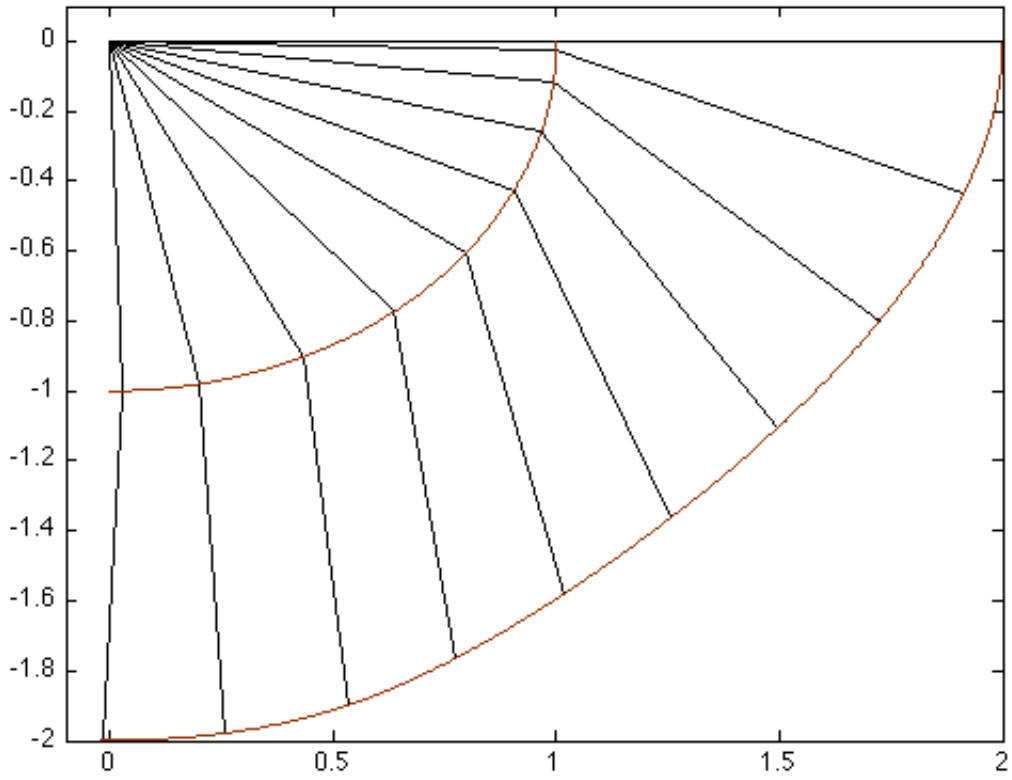

Figure 3.4: This graph is of the blue trajectory on the torus. The hip joint is at zero, and the knee joint traces an arc. The second arc shows the trajectory of the foot.

Torques for an optimal kick

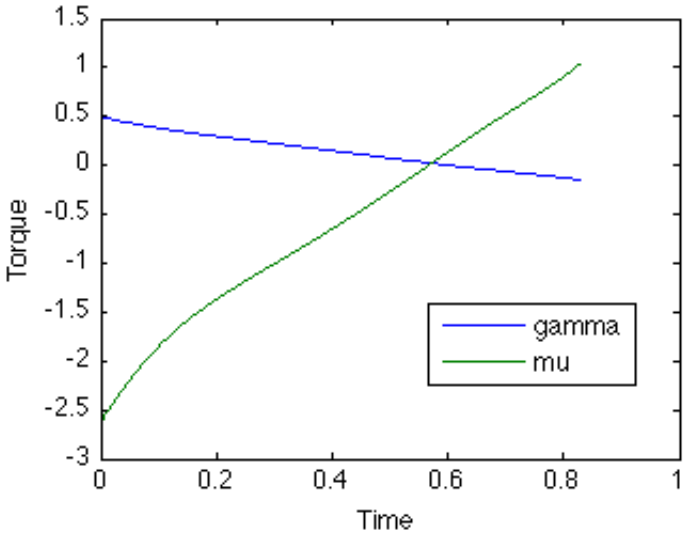

Figure 3.5: The torques, γ, μ, are plotted for the blue trajectory from the torus.

whereas others did not. Maybe adding the maximal power to the cost would help with the technique, as the optimal trajectories ended up converging to what looks like pretty weak kicks. It is not too much of a surprise that some of the near optimal trajectories look fairly different. In practice, much of getting the right kick is to be able to feel it, not necessarily know how to throw exactly the same kick every time.

For the full problem the only result is presented as an approximation with an optimal planar trajectory. To do the approximation we need a mapping from $R_\alpha \to w_\alpha$. We need the mapping to satisfy the beginning and end conditions, such that $w_\alpha(0) = 1, w_\alpha(t_1) = 0.5(1 - \hat{b}_1 + \hat{b}_2 + \hat{b}_3)$. The simplest mapping is

$$\begin{aligned} w_\alpha &= \cos(\alpha) - \sin(\alpha)\frac{1}{\sqrt{2}}(\hat{b}_3 + \hat{b}_2), \\ w_\beta &= \cos(\frac{\beta}{2}) - \sin(\frac{\beta}{2})\hat{b}_2. \end{aligned} \tag{3.1}$$

Three dimensional plot of an optimal planar kick.

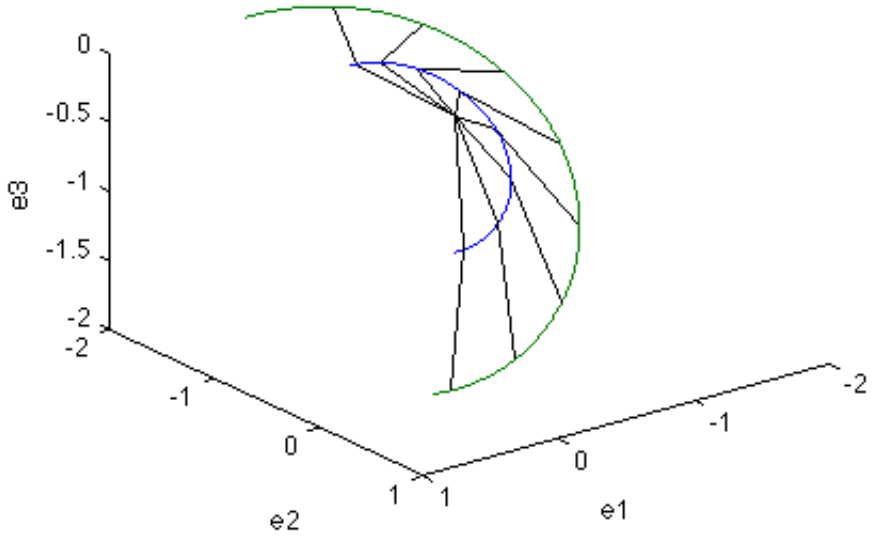

Figure 3.6: This is the plot of the blue trajectory from Figure 3.3, in 3-D space using the mapping of Eq. (3.1).

Using this mapping the blue trajectory from Figure 3.3 is plotted as a three dimensional kick in Figure 3.6. We expect this not to be optimal as we have not taken into account any of the new dynamics.

4

Conclusion

The overall project was a success in learning to apply the tools of Lie groups, Clifford algebras, and control theory to a difficult physical problem. Many of the computations, while straight forward are tedious. I believe an effective computer algebra problem would greatly help with these computations. The non-linear manipulations are, however, a little too complex for Mathematica and Matlab to handle neatly. One of the most promising uses of the Clifford algebra is how the manipulations could be handled efficiently and generally by a computer.

The overall calculation of optimal trajectories is still difficult and computationally intensive. While this approach can give a pretty good understanding of a well behaved system, it may not be well suited for real-time optimal control, or optimal control of systems with less well behaved constraints. The final step of achieving the optimal solution for the planar problem exemplified how this sort of modeling is as much an art as a science. Significantly different optimal trajectories were found near the optimal solution. These differences were small enough that the model was not accurate enough to tell them apart

in terms of real optimality. This also showcases the room there is in the martial arts for very different in styles and techniques to still be effective.

An area of further research is to introduce stochasticity. In Taekwondo, one is never exactly sure where the target is going to be in the next moment, so modeling the kick for a randomly moving target is of interest. There are also some interesting trade-offs in terms of computational ease when entering the stochastic case. Certain classes of control problems can be solved with path sampling, where superposition can be used even for non-linear problems.

The approach taken in this thesis may not really help our understanding of martial arts. These arts have developed over thousands of years and take into account many factors that would be nearly impossible to incorporate into a model. However, the problem is very much related to many practical problems for which such a deep understanding has not developed.

Appendix A

Derivation of Equations

A.1 Planar Problem Mechanics

We calculate the kinetic energy by adding the rotational energy and the translational energy of the center of mass for the second limb,

$$\begin{aligned}
\mathbf{cm}_2 &= R_\alpha(\mathbf{r}_1 + R_\beta(\frac{\mathbf{r}_2}{2})) = e^{i\alpha}(-il_1 + e^{i\beta}(-il_2)),\\
\dot{\mathbf{cm}}_2 &= -ie^{i\alpha}(el_1 + e^{i\beta}(el_2 + nl_2)) = -ie^{i\alpha}(el_1 + e^{i\beta}l_2(e+n)),\\
|\dot{\mathbf{cm}}_2|^2 &= -(e^2l_1^2 + (e+n)^2l_2^2 + l_1l_2e\overline{(e+n)}(e^{i\beta} + e^{-i\beta}))\\
|\dot{\mathbf{cm}}_2|^2 &= (l_1^2 + l_2^2 + 2l_1l_2cos(\beta))|e|^2 + l_2^2|n|^2 + 2(l_2^2 + l_1l_2cos(\beta))e\overline{n}.
\end{aligned}$$

Now the kinetic energy part of H can be expressed in matrix form taking into account that the tangent vectors are imaginary. Let $v = \begin{pmatrix} e \\ \\ n \end{pmatrix}$, the kinetic energy is

$$-\frac{1}{2}v'\begin{pmatrix} \mathbb{I}_1 + \mathbb{I}_2 + (l_1^2 + l_2^2)m_2 + l_1l_2m_2(R_\beta + \overline{R_\beta}) & \mathbb{I}_2 + l_2^2m_2 + l_1l_2m_2\frac{1}{2}(R_\beta + \overline{R_\beta}) \\ \mathbb{I}_2 + l_2^2m_2 + l_1l_2m_2\frac{1}{2}(R_\beta + \overline{R_\beta}) & \mathbb{I}_2 + l_2^2m_2 \end{pmatrix}v. \tag{A.1}$$

The potential energy is simply $m_1gh_1 + m_2gh_2$ where the heights are the imaginary part of the complex vector to the center of mass of the limbs.

$$\begin{aligned} V &= g\mathrm{Im}(\mathrm{m}_1\frac{\mathrm{l}_1}{2}\mathrm{e}^{\mathrm{i}\alpha}(-\mathrm{i}) + \mathrm{m}_2\mathrm{e}^{\mathrm{i}\alpha}(-\mathrm{il}_1 + \mathrm{e}^{\mathrm{i}\beta}(-\mathrm{il}_2)) \\ &= -g\frac{1}{2}((R_\alpha + \overline{R_\alpha})(m_1\frac{l_1}{2} + m_2l_1) + (R_\alpha R_\beta + \overline{R_\alpha R_\beta})m_2l_2 \\ &= -g(cos(\alpha)(m_1\frac{l_1}{2} + m_2l_1) + cos(\alpha+\beta)m_2l_2. \end{aligned}$$

The derivative of the metric matrix is easy to compute. Notice that every entry is imaginary,

$$-\frac{\delta G}{\delta R_\beta} = \begin{pmatrix} l_1l_2m_2(R_\beta - \overline{R_\beta}) & \frac{1}{2}l_1l_2m_2(R_\beta - \overline{R_\beta}) \\ \frac{1}{2}l_1l_2m_2(R_\beta - \overline{R_\beta}) & 0 \end{pmatrix}.$$

When computing the derivatives, the partial derivatives are straight forward to compute. The derivatives that we are actual after are the adjoint to these derivatives, since all of them are purely imaginary the complex conjugation simply switches signs.

A.2 Planar Problem Controls

Once again all the derivatives computed are imaginary, and are actually the negative partial derivatives when we take into account we want the output in the dual space to a purely imaginary variable. A few terms pose a little difficulty so there is brief discussion of the most complicated terms. A particularly tough term with appears in the equation for $\chi_{\mathcal{H}}^{p_\beta}$ is

$$-\frac{\delta < \chi_H^\nu, p_\nu >}{\delta R_\beta} = \frac{\delta}{\delta R_\beta}\left[\frac{1}{2}\begin{pmatrix} \eta & \nu \end{pmatrix}\frac{\delta G^{-1}}{R_\beta}\begin{pmatrix} \eta \\ \nu \end{pmatrix} - \frac{\delta V(R_\alpha, R_\beta)}{\delta R_\beta}\right]p_\nu.$$

The first term is particularly daunting, but is just a matter of using the product rule on the derivative of the inverse linear transformation,

$$\frac{1}{2}\begin{pmatrix} \eta & \nu \end{pmatrix}\frac{\delta^2 G^{-1}}{\delta R_\beta^2}\begin{pmatrix} \eta \\ \nu \end{pmatrix} = \begin{pmatrix} \eta & \nu \end{pmatrix}(2G^{-1}\frac{\delta G}{\delta R_\beta}G^{-1}\frac{\delta G}{\delta R_\beta}G^{-1} - G^{-1}\frac{\delta^2 G}{\delta R_\beta^2}G^{-1})\begin{pmatrix} \eta \\ \nu \end{pmatrix}.$$

Another novel term is from the inner product of the co-state variables and the tangent vectors. Appearing in the Hamiltonian vector field equations for p_η, p_ν is

$$\begin{pmatrix} \frac{\partial}{\partial \eta} \\ \frac{\partial}{\partial \nu} \end{pmatrix}\left\langle \begin{pmatrix} e \\ n \end{pmatrix}, \begin{pmatrix} p_\alpha \\ p_\beta \end{pmatrix} \right\rangle = -\begin{pmatrix} \eta & \nu \end{pmatrix} G^{-T}\begin{pmatrix} p_\alpha \\ p_\beta \end{pmatrix} = -G^{-T}\begin{pmatrix} p_\alpha \\ p_\beta \end{pmatrix}.$$

A.3 Full Problem Mechanics

The derivation of equations for the full problem proceeds in the same manner but involves non-commutative algebra. For the kinetic energy we must calculate the energy from the motion of the center of mass of the second joint as well as the rotational energy. The $[,]$ operator is half of the commutator of a vector and bi-vector and is also the Lie algebra product between two bi-vectors. When operating on a vector and a bi-vector the output is always a pure vector on the plane of the bi-vector, orthogonal to the original vector. When operating between two bi-vectors the commutator returns a mutually orthogonal bi-vector. We are really only concerned with the real number associated with the magnitude of the

derivative of the center of mass vector, which simplifies slightly,

$$\begin{aligned}
c\dot{m} &= \frac{d}{dt} w_\alpha (r_1 + w_\beta r_2 w_\beta^{-1}) w_\alpha^{-1} \\
c\dot{m} &= -\frac{1}{2} w_\alpha ((e r_1 - r_1 e) + (e(w_\beta r_2 w_\beta^{-1} - w_\beta r_2 w_\beta^{-1} e) + (w_\beta n r_2 w_\beta^{-1} - w_\beta r_2 n w_\beta^{-1}))) w_\alpha^{-1} \\
c\dot{m} &= w_\alpha ([r_1, e] + [w_\beta r_2 w_\beta^{-1}, e] + w_\beta [r_2, n] w_\beta^{-1}) w_\alpha^{-1} \\
c\dot{m}^2 &= [r_1, e]^2 + [w_\beta r_2 w_\beta^{-1}, e]^2 + [r_2, n]^2, \\
&\quad +2(< [r_1, e], [w_\beta r_2 w_\beta^{-1}, e] >_1 + < [r_1, e], w_\beta [r_2, n] w_\beta^{-1} >_1 \\
&\quad + < [w_\beta r_2 w_\beta^{-1}, e], w_\beta [r_2, n] w_\beta^{-1} >_1).
\end{aligned}$$

Note that this is now expressed in terms of the inner product on the vectors. For the purposes of getting a metric on the bi-vectors we wish to convert it to an inner product on bi-vectors. In the following manipulations take a,b to be vectors, and C,D to be bi-vectors. Inspired by the geometry of the products we see first that using I the pseudo-scalar, which commutes with all vectors and squares to -1. For the purposes of these calculations the subscripted inner products are the graded inner products of the algebra. When the inner product returns without a subscript it is understood one of the elements is to be conjugated (switch signs). The geometric product can switch over the inner product by reversal in the order of multiplication. The inner product relates to an inner product of bi-vectors by

$$< a, b >_1 = -\frac{1}{2}(aIbI + bIaI) = - < aI, bI >_2 .$$

The terms we have to deal with appear like

$$\begin{aligned}
< [a, C], [b, D] >_1 &= - < [a, C]I, [b, D]I >_2 = \frac{1}{4} < aCI - CaI, bDI - DbI > \\
&= \frac{1}{4} < C, bDa - abD + aDb - Dba > = < C, a \wedge [D, b] > .
\end{aligned}$$

The term on the right is now a single transformation of the bi-vector D. We would like to write the whole kinetic energy part of the Hamiltonian in terms of these inner products and a linear transformation,

$$\frac{1}{2}(< e, R_\alpha(\mathbb{I}_1)(e) > + < (e+n), R_\alpha(R_\beta(\mathbb{I}_2))(e+n) >$$

$$+m_2(< e, r_1 \wedge [e, r_1] > + < e, w_\beta r_2 w_\beta^{-1} \wedge [e, w_\beta r_2 w_\beta^{-1}] + < n, r_2 \wedge [n, r_2] >$$

$$+2(< e, r_1 \wedge [e, w_\beta r_2 w_\beta^{-1}] > + < e, r_1 \wedge w_\beta [n, r_2] w_\beta^{-1} >$$

$$+ < e, w_\beta r_2 w_\beta^{-1} \wedge w_\beta [n, r_2] w_\beta^{-1} >))$$

$\Rightarrow$

$$\frac{1}{2}(< e, R_\alpha(\mathbb{I}_1)(e) + R_\alpha(R_\beta(\mathbb{I}_2))(e+n) + m_2(r_1 \wedge [e, r_1]$$

$$+R_\beta(r_2) \wedge [e, R_\beta(r_2)] + 2(r_1 \wedge [e, R_\beta(r_2)] + r_1 \wedge R_\beta([n, r_2]) + R_\beta(r_2) \wedge R_\beta([n, r_2]))) >$$

$$+ < n, R_\alpha(R_\beta(\mathbb{I}_2))(e+n) + m_2 r_2 \wedge [n, r_2] >).$$

The transformations on the right hand side of the inner product are what appear in Equations (2.11), for the metric in the Hamiltonian.

To take the derivative of the kinetic energy part of the Hamiltonian we need to refer back to the definition, Eq. (1.2),

$$< a, \frac{\delta < \pi, G^{-1}(w_\alpha)\pi >}{\delta w_\alpha} >= \frac{d}{dt} < \pi, G^{-1}(w_\alpha e^{-\frac{at}{2}})\pi >= - < \pi, G^{-1} \frac{d}{dt} G(w_\alpha e^{-\frac{at}{2}}) G^{-1} \pi > .$$

The general form for a term in G can be expressed in terms of two linear transformations, t, s, that do not depend on R_α, like $G(\pi) = t(R_\alpha(b(\pi)))$. These terms may appear many times. As for the inertial tensors that are themselves rotated, we can consider it as just two applications of the rotation and the inverse rotation. Using the product rule we can deal with these separately. The time derivative appears as the commutator with a tangent

element and passes inside the first linear operator, $t(R_\alpha([b(\pi), a]))$. Here we continue trying to find the derivative for the inverse transformation appearing in the equations,

$$< G^{-T}\pi, t(R_\alpha([b(G^{-1}(\pi)), a])) >=< R_\alpha^{-1}(t^T(G^{-T}\pi)), [b(G^{-1}(\pi)), a] > .$$

We can cycle the inner product and the commutator to get,

$$< a, [R_\alpha^{-1}(t^T(G^{-T}\pi)), b(G^{-1}(\overline{\pi}))] > .$$

This is obviously a little more complicated then the case where this derivative is a scalar.

Bibliography

[1] Bullo F, Lewis A. Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Simple Mechanical Control Systems. Springer, 2005.

[2] Bertsekas D, Tsitsiklis J. Neuro-Dynamic Programming. Athena Scientific, 1996.

[3] Doran C, Lasenby A. Geometric Algebra for Physicists. Cambridge University Press, 2003.

[4] Hestenes D, Sobczyk G. Clifford Algebra to Geometric Calculus: A Unified Language for Mathematics and Physics. D. Reidel Publishing Company, 1984.

[5] Knowles G. An Introduction to Applied Optimal Control. Academic Press, Inc., 1981.

[6] Marsden J, Ratiu T. Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems. Springer-Verlag, 1994.

[7] Pontryagin L, Boltyanskii V, Gamkrelidze R, Mishchenko E. L.S. Pontryagin Selected Works, Volume 4: The Mathematical Theory of Optimal Processes. Gordon and Breach Science Publishers, 1986.

[8] Steeb W. Continuous Symmetries, Lie Algebras, Differential Equations and Computer Algebra. World Scientific, 1996.