INTERPOLATION IN RIEMANNIAN MANIFOLDS

Shreya Bhattarai

This thesis is presented for the degree of
Doctor of Philosophy
of The University of Western Australia
Department of Mathematics and Statistics
2013
Candidate’s statement of contribution

Chapter 3 in this thesis is a paper submitted for publication, which I jointly authored as primary author, with my supervisor W. Prof Lyle Noakes. All other chapters in the thesis are the result of my own work but with a lot of help from weekly discussions with Prof. Noakes.

Shreya Bhattarai (candidate) ________________________________

W Prof. Lyle Noakes (supervisor) ________________________________
Abstract

It is a very natural task to connect the dots between patterns of data we see. This fundamental procedure of connecting the dots, or interpolation, arises in many contexts including graphics, robot path planning, and medical imagining. Although the subject of interpolation has been thoroughly studied for Euclidean spaces, the more general setting of interpolating in a Riemannian manifold has only relatively recently received attention. The central aim of this thesis is to present results concerning the interpolation of data points in Riemannian manifolds, and in particular Lie groups.

Riemannian manifolds are smooth spaces where concepts such as distances and angles are defined using a metric. A dynamical system's state can often be represented as points in a Riemannian manifold and we can apply geometrical methods of inference to predict how the system would have behaved given a suitable model of the dynamics. This thesis in a broad sense looks at various ways to model a system's dynamical behaviour in order to interpolate effectively.

Chapter 3 concerns null Riemannian cubics in tension. Null Riemannian cubics in tension are a special sub-class of Riemannian cubics in tension and are the differential analogues of hyperbolic trigonometric functions, which are used to produce differentiable interpolants. Asymptotic behaviour of the curves are derived.

Chapter 4 concerns ways to incorporate dynamical systems into the interpolation scheme using the notion of a prior vector field. If a differential equation which attempts to model a dynamical system is specified by a vector field, then a natural choice of interpolant is one which minimises the error in the equation being satisfied. Several different frameworks, including second order and time varying systems, are considered and some necessary conditions concerning the solutions are described in the case of Lie groups.

Chapter 5 builds on the results of Chapters 3 and 4 by implementing an interpolation scheme based on a linearisation of the differential equation. We approximate natural conditional extremal splines and natural Riemannian cubics in tension splines using an analogue of the traditional B-spline methodology and analyse how effective the method is.

Chapter 6 focuses on discretisation of spaces of curves and outlines relationships between the continuous and discretised problem. An alternative method to produce interpolants using numerical optimization packages on (large) finite dimensional spaces is considered which allows for a much broader scope for applications. The results produced are compared to those of the previous chapter.
Acknowledgements

There are so many people who made it possible for me to finish my PhD. There is no possible way to thank everyone in the amount of words that I feel they each deserve.

First and foremost I would like to thank my supervisor, W Prof. Lyle Noakes who has shown me unimaginable amounts of care and patience. With your guidance and love for mathematics, you have made completing my PhD an extremely enjoyable experience and have taught me how to always bring out the best in other people.

I would also like to thank my examiners Prof. Anthony Bloch, Prof. Marin Kobilarov, and Dr. Cesare Tronci. Your valuable suggestions and feedback have proven invaluable and provide a wealth of future directions for study.

I would like to acknowledge and thank the University of Western Australia and the Australian government for allowing me to study under a University Postgraduate Scholarship and Australian Postgraduate Scholarship.

To Dad, Mum, and my brother Debesh, thanks for the unceasing encouragement and support.

I would also like to thank the many friends whom I have had the pleasure of sharing this time with. To the ‘amigos’ Brian Corr, Yohei Negi, Philip Schrader, Michael Pauley, and Ryan Mickler. Without each of you, life at university would be a major drag. To my past and present housemates Trent, Vlad, Chris, Steve and Matt. Thanks for the company and all the conversations. To Kenny, Tony, Sean, John, Adam, Bucky, Smriti, Jess, Anthony, Rohan(s), Sham, and all the others who I have missed, you made the last few years the best years of my life. To Hua, for being an excellent Go teacher and friend. To the ‘crossword gang’, Rob, Con and all the friendly staff at UWA who made lunch times quite memorable.

Lastly I would like to thank my beautiful girlfriend Rijbi Dixit who was always there to help me and clear any doubts I may ever had had.
Contents

0.1 Notation .......................................................... 7

1 Riemannian Geometry and Lie Groups ......................... 9
  1.1 Smooth Manifolds ............................................ 9
  1.2 Riemannian Geometry .................................... 13
  1.3 Calculus of Variations ...................................... 14
    1.3.1 First order Lagrangians and the Euler-Lagrange Equations 14
    1.3.2 Hamiltonian Formalism ................................ 15
  1.4 Lie algebras and Lie groups ................................ 16
    1.4.1 Matrix Lie groups ..................................... 20
    1.4.2 Lax constraints ....................................... 20

2 Literature Review ................................................ 23
  2.1 Interpolation in Manifolds .................................. 23
  2.2 Riemannian Cubics in Tension ................................ 26
  2.3 Prior fields .................................................. 27

3 Asymptotics of Null Riemannian Cubics in Tension in SO(3) ... 29
  3.1 Introduction .................................................. 29
    3.1.1 Null Riemannian Cubics in Tension and their connection to other families of curves ............................. 33
  3.2 Asymptotics of Null Lie Quadratics in Tension in so(3) .... 37
  3.3 Asymptotics of Null Riemannian Cubics in Tension in SO(3) .. 43
  3.4 Conclusion ................................................... 47

4 Prior fields ........................................................ 49
  4.1 Introduction .................................................. 49
  4.2 First order systems ......................................... 49
    4.2.1 Variable left-invariant priors ........................ 63
  4.3 Second order systems ....................................... 64
## CONTENTS

4.4 Time varying systems ........................................... 66
   4.4.1 Time varying prior fields for solvable $A(t)$ in a bi-invariant Lie group .............................. 71
   4.4.2 Time varying prior fields in SO(3) ....................... 73
   4.4.3 Second order time-variant systems ..................... 77
4.5 Conclusion ...................................................... 79

5 Linear interpolation schemes on bi-invariant Lie groups .... 81
   5.1 Introduction .................................................. 81
   5.2 Second order prior fields in SO(3) ....................... 82
      5.2.1 $A(t) = \text{constant}$ ................................ 84
   5.3 Approximating natural Riemannian cubics in tension splines ........................................... 97
   5.4 Conclusion ...................................................... 100

6 Numerical Interpolation Schemes for Second Order Lagrangians101
   6.1 Introduction .................................................. 101
   6.2 Approximations of derivatives .............................. 102
   6.3 Unconstrained optimisation ................................ 104
   6.4 Constrained optimisation ................................... 107
      6.4.1 Case study: $\mathcal{A}J_2$ on $S^3$ .................. 107
   6.5 Conclusion ...................................................... 110

7 Conclusions .......................................................... 113

Bibliography .......................................................... 117
0.1 Notation

Quick reference table for frequently used notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q$</td>
<td>Manifold</td>
</tr>
<tr>
<td>$U$</td>
<td>Neighbourhoods of $Q$</td>
</tr>
<tr>
<td>$q$</td>
<td>Points in $Q$</td>
</tr>
<tr>
<td>$v$</td>
<td>Vectors in $TQ$</td>
</tr>
<tr>
<td>$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$</td>
<td>Smooth vector fields on $Q$</td>
</tr>
<tr>
<td>$n$</td>
<td>Dimension of $Q$</td>
</tr>
<tr>
<td>$f^{(k)}$</td>
<td>The $k^{th}$ derivative of $f$</td>
</tr>
<tr>
<td>$\alpha, \beta$</td>
<td>Set index variables</td>
</tr>
<tr>
<td>$G$</td>
<td>Lie groups</td>
</tr>
<tr>
<td>$g, h$</td>
<td>Lie group elements (or sometimes functions)</td>
</tr>
<tr>
<td>$\mathfrak{g}$</td>
<td>Lie algebras</td>
</tr>
<tr>
<td>$A, B, C, D$</td>
<td>Lie algebra variables</td>
</tr>
<tr>
<td>$[A, B]$</td>
<td>Lie bracket of $A$ and $B$</td>
</tr>
<tr>
<td>$Ad_g : \mathfrak{g} \to \mathfrak{g}$</td>
<td>Adjoint action of $g$: $Ad_g A = dL_g \circ dR_{g^{-1}} A$</td>
</tr>
<tr>
<td>$ad_A : \mathfrak{g} \to \mathfrak{g}$</td>
<td>Adjoint action of $A$: $ad_A B = [A, B]$</td>
</tr>
<tr>
<td>$\langle v_0, v_1 \rangle$</td>
<td>Inner product of $v_0$ and $v_1$</td>
</tr>
<tr>
<td>$I$</td>
<td>The unit interval</td>
</tr>
<tr>
<td>$I_{k\times k}$</td>
<td>The $k \times k$ identity matrix</td>
</tr>
<tr>
<td>$e$</td>
<td>Lie group identity element</td>
</tr>
<tr>
<td>$C^{k}_{q_0, q_1}$</td>
<td>Space of $C^k$ curves connecting $q_0$ and $q_1$</td>
</tr>
<tr>
<td>$x(t)$</td>
<td>A curve $x$ in the manifold parameterised by $t$</td>
</tr>
<tr>
<td>$\gamma(t)$</td>
<td>Geodesics in $Q$</td>
</tr>
<tr>
<td>$X(t), W(t)$</td>
<td>Vector field along $x(t)$</td>
</tr>
<tr>
<td>$V(t)$</td>
<td>The left Lie reduction of $x^{(1)}(t)$</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of intervals to be interpolated</td>
</tr>
<tr>
<td>$\xi = (\xi_0, \ldots, \xi_N)$</td>
<td>Data set of points to be interpolated</td>
</tr>
<tr>
<td>$J_k$</td>
<td>The $k^{th}$ order energy functional</td>
</tr>
<tr>
<td>$J^\lambda_2$</td>
<td>Riemannian cubic in tension functional with tension $\lambda$</td>
</tr>
<tr>
<td>$A J_k$</td>
<td>The $k^{th}$ order prior field functional</td>
</tr>
</tbody>
</table>
Riemannian Geometry and Lie Groups

1.1 Smooth Manifolds

Manifolds are spaces which locally look like Euclidean space but globally may have different structure. Typically manifolds arise in physical systems satisfying nonlinear constraints such as conservation laws, and come equipped with additional structure, such as Riemannian metrics or symplectic forms. We will be interested in studying interpolation in manifolds and so we will only consider connected spaces.

Formally, an $n$-dimensional smooth manifold is a topological space $Q$ with an open cover $\{U_\alpha\}$ and a collection of homeomorphisms $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ such that each of the maps $\phi_\alpha \circ \phi_\beta^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ are smooth. The pairs $(U_\alpha, \phi_\alpha)$ are called charts. We say the standard Euclidean coordinates $(u_1, \ldots, u_n)$ are a local coordinate system on $U_\alpha$ where $(u_1, \ldots, u_n)$ represents $\phi(u_1, \ldots, u_n)$. A map between manifolds is considered smooth if the induced maps between all choices of coordinate charts are also smooth in the standard sense of maps between open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$.

Let $x_1, x_2 : (-1, 1) \to Q$ be two curves such that $x_1(0) = x_2(0) = q$. We say they are equivalent at $q$ if we have $(\phi_\alpha \circ x_1)'(0) = (\phi_\alpha \circ x_2)'(0)$. It can be shown that this definition does not depend on the chart. Equivalence classes of these curves are called tangent vectors at $q$ and the set of tangent vectors at $q$ can be made into a vector space by setting $[x_1(t)] + [x_2(t)] = [x_3(t)]$ when we have $(\phi_\alpha \circ x_1)'(0) + (\phi_\alpha \circ x_2)'(0) = (\phi_\alpha \circ x_3)'(0)$ and $r[x_1(t)] = [x_2(t)]$ when $r(\phi_\alpha \circ x_1)'(0) = (\phi_\alpha \circ x_2)'(0)$ for all $r \in \mathbb{R}$. We denote the set of tangent vectors at $q$ by $T_qQ$. Let $q = (u_1, \ldots, u_n)$ in a choice of local coordinates, then a basis for $T_qQ$ will be $\partial_i := [(u_1, \ldots, u_i + t, \ldots, u_n)]$ for $i = 1, \ldots, n$. Let $v \in T_qQ$ and $f : Q \to \mathbb{R}$, then define $vf := (f \circ \tilde{v})'(0)$ where $\tilde{v}$ is any representative curve of $v$.

The operation of $v$ is independent of choice of $\tilde{v}$, and so tangent vectors can be thought of as differential operators, and $\partial_i$ as partial derivatives in a coordinate system. Any tangent vector $v$ can be written with respect to local coordinates as
We can define the **tangent bundle** \( TQ \) of \( Q \) to be the disjoint union of tangent spaces \( T_qQ \), that is, \( TQ = \sqcup_{q \in Q} \{ q \} \times T_qQ \). There is a natural projection map \( \pi : TQ \to Q \) given by \( \pi(q, v) = q \). The tangent bundle has a natural topology and can be made into a manifold itself.

A **vector field** \( X \) on \( Q \) is a smooth section of the tangent bundle. That is to say, \( X : Q \to TQ \) is a smooth map and \( \pi \circ X = \text{id} \). A **derivation** \( D : C^\infty(Q) \to C^\infty(Q) \) is a map such that \( D(f + g) = Df + Dg \) and \( D(fg) = fDg + gDf \). The space of smooth vector fields form a vector space denoted \( \mathfrak{X}(Q) \) and they are in one to one correspondence with the space of derivations with the identification \( (\mathfrak{X}f)(q) = X(q)f \) (See [36, Addendum] for a proof). The **flow** \( \Phi \) of a vector field \( X \) is a map \( \Phi_t : Q \to Q \) defined by the differential equation \( \frac{d}{dt} \Phi_t(q_0) = X_{\Phi_t(q_0)} \) with \( \Phi_0(q_0) = q_0 \). The question of when \( \Phi_t \) is defined is answered in [36] which says that if \( Q \) is a Hausdorff manifold, there is a continuous function \( \epsilon : Q \to \mathbb{R}^+ \) such that \( \Phi_t(q_0) \) is defined when \( |t| < \epsilon(q_0) \). Moreover, \( \Phi_s \circ \Phi_t = \Phi_{s+t} \) when they are defined. Specific trajectories generated by a flow are also called integral curves of \( X \).

Given two vector fields \( \mathcal{X}, \mathcal{Y} \), we can define their **Lie bracket** \( [\mathcal{X}, \mathcal{Y}] \) by the formula \( [\mathcal{X}, \mathcal{Y}]f = \mathcal{X}(\mathcal{Y}f) - \mathcal{Y}(\mathcal{X}f) \) for all \( f \in C^\infty(Q) \). It can be shown that \([\mathcal{X}, \mathcal{Y}]\) is well defined and is in fact a vector field. With this operation, the space \( \mathfrak{X}(Q) \) is a Lie algebra (See section 1.4). The advantage of thinking of vector fields as derivations is that derivations have a natural Lie bracket operation.

The Lie bracket of two vector fields \( \mathcal{X} \) and \( \mathcal{Y} \) can be informally thought of as a measure of the failure of flows of \( \mathcal{X} \) and \( \mathcal{Y} \) to commute. More formally, let \( \Phi \) and \( \Psi \) be the flows of \( \mathcal{X} \) and \( \mathcal{Y} \) respectively. Then we have for any chart, 
\[
(\Psi_{-t} \circ \Phi_{-t} \circ \Psi_t \circ \Phi_t) x_0 = x_0 + t^2 [\mathcal{X}, \mathcal{Y}] + O(t^3).
\]

A more general notion than that of a tangent bundle is that of a vector bundle. A quadruple \( \mathbf{E} = (E, Q, \pi, F) \) is called a **smooth vector bundle** when both the total space \( E \) and the base space \( Q \) are smooth manifolds and \( \pi : E \to Q \) is a smooth surjection such that \( \pi^{-1}(q) \), the fiber over \( q \), has a vector space structure.
which is isomorphic to the fiber space $F$. Moreover we require that for any point in $Q$, there is a neighbourhood $U$ of that point and map $\phi$ such that the following diagram commutes:

$$
\begin{array}{ccc}
\pi^{-1}(U) & \xrightarrow{\phi} & U \times F \\
\downarrow & & \downarrow pr_1 \\
U & & 
\end{array}
$$

where $pr_1$ is the projection onto the first component. Moreover, for any point $q$ in $U$, the map $v \mapsto \phi^{-1}(q,v)$ is a linear isomorphism between $F$ and $\pi^{-1}(q)$. Such $(U, \phi)$ are called local trivializations. An example of a smooth vector bundle is $TQ = (TQ, Q, \pi, \mathbb{R}^n)$. We define a smooth section of a bundle to be a smooth map $s : Q \to E$ such that $\pi \circ s = id$. The space of all smooth sections on $E$ is denoted by $\Gamma(E)$, for example $\Gamma(TQ) = \mathfrak{X}(Q)$.

Let $\pi : E \to Q$ be a vector bundle and let $f : Q' \to Q$ be a continuous map. Then the pullback bundle is defined by $f^*E = \{(q, v) \in Q' \times E | f(q) = \pi(v)\}$ with the topology on $f^*E$ being defined by the subspace topology and the projection map $\pi' : f^*E \to Q'$ given by projection onto the first factor. Given two vector bundles $E_1, E_2$ over the same base space, we are able to construct the Whitney sum $E_1 \oplus E_2$ of the bundles and the tensor product $E_1 \otimes E_2$. The Whitney sum $E_1 \oplus E_2$ has fiber over $q$ equal to $\pi_1^{-1}(q) \oplus \pi_2^{-1}(q)$. Similarly the tensor product $E_1 \otimes E_2$ has fiber over $q$ equal to $\pi_1^{-1}(q) \otimes \pi_2^{-1}(q)$. We can also take the dual bundle, $E^*$, where fibers over $q$ equal $(\pi^{-1}(q))^*$. For more details regarding the construction of these spaces, see [45].

For any vector bundle $E$ and natural numbers $r,s$, define the tensor bundle of type $(r,s)$ as follows:

$$
T^r_s E = \underbrace{E^* \otimes \cdots \otimes E^*}_{r} \otimes \underbrace{E \otimes \cdots \otimes E}_{s}
$$

(1.1)
Sections of such a bundle are called tensor fields and we will refer to the space of all tensor fields $\Gamma^r_s E$. Such a tensor field is said to have $r$ contravariant indices and $s$ covariant indices.

A $\mathbb{R}$-linear mapping $\nabla : \Gamma E \rightarrow \Gamma(TQ^* \otimes E)$ is called a connection if it satisfies the Leibniz identity $\nabla fs = df \otimes s + f \nabla s$ for all $f \in C^\infty(Q)$ and $s \in \Gamma E$ where $d$ is the exterior derivative. That is to say, given a section $s$, $\nabla s$ is an element of $\Gamma(TQ^* \otimes E)$. This allows us to define the covariant derivative $\nabla_X s$ for any vector field $X$ and section $s$ by contracting $\nabla s$ with $X$, resulting in another section.

Given a connection on a manifold, we have a covariant derivative $\nabla : X(Q) \times X(Q) \rightarrow X(Q)$ which satisfies the following properties:

\begin{align*}
\nabla_{fX_1 + gX_2}Y &= f\nabla_{X_1}Y + g\nabla_{X_2}Y \quad \forall f, g \in C^\infty(Q) \\
\nabla_X (fY) &= f\nabla_X Y + (Xf)Y \quad \forall f \in C^\infty(Q) \\
\nabla_X (aY_1 + bY_2) &= a\nabla_X Y_1 + b\nabla_X Y_2 \quad \forall a, b \in \mathbb{R}
\end{align*}

Informally we can think of $\nabla_X Y$ as differentiating the vector field $Y$ in the direction of $X$. $(\nabla_X Y)(q)$ depends only on $X(q)$ and on $Y$ in a neighbourhood of $q$. In local coordinates, the connection defines $n^3$ smooth functions $\Gamma^k_{ij}$ where $\nabla_{\partial_i} \partial_j = \Gamma^k_{ij} \partial_k$. We say the connection is torsion-free (or symmetric) when

\begin{equation}
\nabla_X Y - \nabla_Y X = [X, Y] \quad \forall X, Y \in X(Q)
\end{equation}

Let $x$ be a curve in $Q$. We can then define the notion of vector fields over $x$. We say $X$ is a smooth vector field over or along $x$ when $X(t) \in T_{x(t)}Q$ and the map $t \mapsto X(t)f$ is smooth for any $f \in C^\infty(Q)$. Then we can define a new vector field along $x$ called the covariant derivative of $X$ along $x$ by $\nabla_{x(t)}X$, denoted $\nabla_t X$ for short. In local coordinates let $X(t) = X^i(t) \partial_i$. Then $\nabla_t X(t)$ is defined by

\begin{equation}
\nabla_t X(t) := (X^i)^{(1)}(t) \partial_i + X^i \nabla_{x(t)}(t) \partial_i.
\end{equation}

Covariant differentiation allows us to define the important notion of parallel trans-
porting a vector. If we have a vector field \( X(t) \) along \( x(t) \) such that \( \nabla_t X = 0 \), then \( X(t) \) is said to have been obtained by parallel transporting \( X(t_0) \) along \( x(t) \). We can write \( \Pi^t_0 X(t_0) = X(t) \).

### 1.2 Riemannian Geometry

Riemannian geometry is the study of manifolds for which we have a notion of inner products between vectors tangent to some point. A **Riemannian metric** is a smooth contravariant 2-tensor field \( g \) which is symmetric and positive definite on each tangent space. Such a metric defines an inner product at each tangent space \( T_q\mathbb{Q} \) given by
\[
\langle v, w \rangle = g_q(v, w)
\]
for \( v, w \in T_q\mathbb{Q} \). A smooth manifold together with a Riemannian metric \((\mathbb{Q}, g)\) is called a **Riemannian manifold**.

Given a Riemannian metric, we say a connection is **compatible** with the metric when
\[
X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle
\]
for all \( X, Y, Z \in \mathfrak{X}(\mathbb{Q}) \).

**Theorem 1.2.1** (Fundamental Theorem of Riemannian Geometry [18]). A Riemannian manifold has a unique torsion-free connection compatible with the metric

This torsion free and symmetric connection is known as the Levi-Civita connection.

For the Levi-Civita connection, if \( X \) and \( Y \) are parallelly transported vector fields along \( x(t) \), then compatibility of the metric implies that \( \langle X(t), Y(t) \rangle \) is fixed along \( x \). Let \( x \) be a small loop with \( x(t_0) = x(t_1) \). Now consider parallel transporting a vector \( v \) around this loop. Generally, \( v \) and the transported vector will differ. The curvature tensor is a measure of this effect to first order accuracy. Define the curvature \( R \) as a \((3, 1)\) tensor field:
\[
R(\mathcal{X}, \mathcal{Y})\mathcal{Z} = \nabla_\mathcal{X} \nabla_\mathcal{Y} \mathcal{Z} - \nabla_\mathcal{Y} \nabla_\mathcal{X} \mathcal{Z} - \nabla_{[\mathcal{X}, \mathcal{Y}]} \mathcal{Z}
\]

It is easily verified (See [18, Section 9]) that such a map depends on \( \mathcal{X}, \mathcal{Y}, \mathcal{Z} \) only at a point \( q \) and so it is indeed a tensor field. Choose a surface \( \sigma : [0, 1] \times [0, 1] \to \mathbb{Q} \) parameterised by \( r \) and \( s \) and define \( X = \frac{\partial \sigma}{\partial r} \) and \( Y = \frac{\partial \sigma}{\partial s} \). Define a curve
\( x_h : [0, 1] \to Q \) to be a traversing of the boundary of the image of \([0, h] \times [0, h]\) under \( \sigma \). Then the Ambrose-Singer theorem \([46]\) says that \( \frac{\partial^2}{\partial h^2} \Pi_x^1 v - v|_{r=0} = R(X, Y)v \) where \( \Pi_x^1 v \) is the parallel transport of \( v \) done piecewise along \( x_h(t) \).

1.3 Calculus of Variations

### 1.3.1 First order Lagrangians and the Euler-Lagrange Equations

In Lagrange’s formulation of classical mechanics, we are to consider paths that minimise the integral of a function \( L : TQ \times \mathbb{R} \to \mathbb{R} \) known as a (first order) Lagrangian. Classically the Lagrangian is given as the kinetic energy minus the potential energy. For example, for a particle moving in a potential field, \( L(q, v, t) = \frac{1}{2}m\|v\|^2 - V(q) \). We are interested in finding curves \( x : [0, 1] \to Q \) that minimise the following functional

\[
J(x) = \int_0^1 L(x(t), x^{(1)}(t), t) \, dt
\]  

(1.9)

over curves \( x(t) \) satisfying the boundary constraint \( x(0) = q_0 \) and \( x(1) = q_1 \).

Define a variation of a curve \( x \) as a smooth family of curves \( x_h \) parameterised by \( h \in (-\varepsilon, \varepsilon) \) where \( x_0 = x \). A necessary property of such a minimising curve \( x \) is that for any variation \( x_h \), the derivative of \( J \) with respect to \( h \) must vanish at \( h = 0 \). Moreover, any curve \( x \) that minimises \( J \) also minimises the functional that is given by the same integrand but on a subinterval \((t_0, t_1) \subseteq [0, 1]\). Let \( X(t) \) be a vector field along \( x(t) \) vanishing at \( t = t_0, t_1 \). Now define \( x_h \) to be any family of curves subject to the condition that \( \frac{d}{dh}|_{h=0} x_h(t) = X(t) \). Choose a local coordinate system containing the curve segment between \([t_0, t_1]\) where \( x_h(t) = u_h(t) \) and let \( L_q \) and \( L_v \) be the corresponding partial derivatives of \( L \).

\[
\frac{dJ(x_h)}{dh} = \frac{d}{dh} \int_{t_0}^{t_1} L(u_h, u_h^{(1)}, t) \, dt = \int_{t_0}^{t_1} \frac{d}{dh} L(u_h, u_h^{(1)}, t) \, dt
\]

\[
= \int_{t_0}^{t_1} \left( L_q(u_h, u_h^{(1)}, t) \frac{du_h}{dh} + L_v(u_h, u_h^{(1)}, t) \frac{du_h^{(1)}}{dh} \right) \, dt
\]
\[ = \int_{t_0}^{t_1} \left( L_q(u_h, u_h^{(1)}, t) - \frac{d}{dt} L_v(u_h, u_h^{(1)}, t) \right) \frac{d u_h}{d h} d t \]

The last step uses integration by parts and the fact that \( \frac{d u_h}{d h} \) vanishes at the endpoints of the variation. Evaluating at \( h = 0 \), we see

\[ \frac{d J}{d h} \bigg|_{h=0} = \int_{t_0}^{t_1} \left( L_x(u, u^{(1)}), t - \frac{d}{dt} L_v(u, u^{(1)}, t) \right) X(t) dt \quad (1.10) \]

Because this must vanish for all choices of \( X(t) \), the Fundamental Lemma of the Calculus of Variations [18] says that we must have the following theorem

**Theorem 1.3.1. Euler-Lagrange equations** For \( x(t) \) to be a critical point \( J \), it is a necessary condition that on any chart, \( x(t) = u(t) \) must satisfy

\[ L_q(u, u^{(1)}, t) - \frac{d}{dt} L_v(u, u^{(1)}, t) = 0 \quad (1.11) \]

Equation (1.11) is referred to as the *Euler-Lagrange equation*, which is a necessary but not sufficient local condition for minimality. It should be noted that the partial derivatives depend on the chart that we choose: more machinery is required to derive the equations in a coordinate free way.

### 1.3.2 Hamiltonian Formalism

We can reformulate the Euler-Lagrange equations in a way that is not dependent on charts. Let \( T^*Q \) be the cotangent bundle of \( Q \). The Liouville 1-form \( \Theta \) is defined on \( T^*Q \) as follows. Let \( (q, \omega) \in T^*_q Q \) and then let \( W \) be in \( T_{(q,\omega)}T^*Q \). Then \( \Theta_{(q,\omega)}(W) = \omega(d\pi^*(W)) \) where \( \pi^* : T^*Q \to Q \) is the projection map. The symplectic 2-form \( \Omega \) on \( T^*Q \) is then defined by taking the exterior derivative, \( \Omega = -d\Theta \).

Given a function \( H : T^*Q \to \mathbb{R} \) called the Hamiltonian, we define the Hamiltonian vector field \( X_H \) on the cotangent bundle as the vector field satisfying the property \( dH = \iota_{X_H} \Omega \) where \( \iota_X \Omega \langle Y \rangle = \Omega(X, Y) \). Flows under this vector field represent solutions to the classical Hamiltonian equations for \( H \).

The correspondence between the Hamiltonian formalism and the Lagrangian
formalism is given by the Legendre transform. Given a map \( L : TQ \to \mathbb{R} \), define the fiber derivative of \( L \), \( F_L : TQ \to T^*Q \), by
\[
F_L(v)(w) = \frac{d}{ds}\bigg|_{s=0} L(v + sw).
\]
We have the corresponding forms on \( TQ \) given by \( \Theta_L = (FL)^*\Theta \) and \( \Omega_L = (FL)^*\Omega \). \( \Theta_L \) is called the Lagrangian 1-form and \( \Omega_L \) the Lagrangian 2-form. When \( F_L \) is locally invertible, the Lagrangian is said to be regular. We will suppose that we have a regular Lagrangian for the remainder of this section.

Let \( X \) be a vector field on \( TQ \) such that \( \iota_X \Omega_L = dE \) where the energy \( E \) is defined by
\[
E(q,v) = F_L(v) \cdot v - L(v) = \Theta_L(X)(v) - L(v).
\]
The flows of \( X \) are solutions to the Euler Lagrange equations. To see this, suppose a curve \( (x(t),v(t)) \in TQ \) satisfies \( (x^{(1)}(t),v^{(1)}(t)) = X(x(t),v(t)) \). Then by examining local coordinates, we have \( v^i(t) = (u^i)^{(1)}(t) \) and that \( L_q(u,v) - \frac{d}{dt} L_v(u,v) = 0 \).

The second of these equations is the Euler-Lagrange equation.

### 1.4 Lie algebras and Lie groups

A **Lie algebra** (over \( \mathbb{R} \) or \( \mathbb{C} \)) is a vector space \( g \) together with a bilinear operation, the **Lie bracket** \([\cdot,\cdot] : g \times g \to g\), satisfying the following relationships

\[
[A, B] + [B, A] = 0 \quad \text{(Anti-commutativity)}
\]
\[
[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \quad \text{(Jacobi identity)}
\]
for all \( A, B, C \in g \).

An important example of a Lie algebra is the space of vector fields on a manifold under the Lie bracket operation. Another example is the space End(\( V \)) of endomorphisms of a vector space \( V \) with \([A, B] := AB - BA\).

A **Lie subalgebra** \( h \) of \( g \) is a vector subspace which is also closed under the Lie bracket operation. Many of the Lie algebras in this thesis will consider will be subalgebras of End(\( V \)). If \( h_1 \) and \( h_2 \) are two Lie subalgebras, we define \([h_1, h_2] := \{[A_1, A_2] : A_1 \in h_1, A_2 \in h_2\}\). We say \( h \) is an ideal of \( g \) when \([h, g] \subseteq h\). If \( g \) has no ideals apart from \( \{0\} \) and itself and if \([g, g] \neq \{0\} \), then \( g \) is called simple.

A **Lie algebra homomorphism** between \( g_1 \) and \( g_2 \) is a linear map \( \phi : g_1 \to g_2 \).
that satisfies $\phi[A, B] = [\phi A, \phi B]$. Let $A \in \mathfrak{g}$, then define the adjoint map $\text{ad}(A) : \mathfrak{g} \rightarrow \mathfrak{g}$ by $\text{ad}(A)B = [A, B]$. It can be seen that $(\text{ad}[A, B])C = [\text{ad} A, \text{ad} B]C$ and so the map $\text{ad} : \mathfrak{g} \rightarrow \text{End}(\mathfrak{g})$ is a Lie algebra homomorphism and the space $\text{ad} \mathfrak{g}$ is called the adjoint representation of $\mathfrak{g}$. A derivation is a linear map $d : \mathfrak{g} \rightarrow \mathfrak{g}$ which satisfies the rule that $d[A, B] = [dA, B] + [A, dB]$ for all $A, B \in \mathfrak{g}$. The Jacobi identity is equivalent to stating that $\text{ad} A$ is a derivation, that is

$$\text{ad}(A)[B, C] = [\text{ad}(A)B, C] + [B, \text{ad}(A)C]$$  \hspace{1cm} (1.12)

We say $Q$ is a Lie group when $Q$ has both a group structure and a smooth manifold structure and the group multiplication and inversion are both smooth (we will often denote $Q$ by $G$ when $Q$ is a Lie group). Define two diffeomorphisms of $G$, $L_g, R_g : G \rightarrow G$ by $L_g h = gh$ and $R_g h = hg$. A vector field is called left-invariant when $\mathcal{X}_{gh} = dL_g(\mathcal{X}_h)$ (Similarly right-invariant if $\mathcal{X}_{gh} = dR_h(\mathcal{X}_g)$). We will consider only left-invariant vector fields for now but the same ideas apply for right-invariance. The left-invariant vector fields form a vector space of finite dimension $\dim G$ and the Lie bracket of two left-invariant vector fields is also left-invariant and so the space of left-invariant vector fields form a Lie subalgebra of $\mathfrak{X}(G)$ called the Lie algebra of $G$. If $\mathcal{X}$ is left-invariant, then $\mathcal{X}_g = dL_g(\mathcal{X}_1)$ and so it is completely determined by the value at the identity. Using this fact, we can identify the Lie algebra of $G$ with the tangent space at the identity, $T_e G$. When we wish to think of elements of $\mathfrak{g}$ as vector fields, we will use the calligraphic variables $\mathcal{X}, \mathcal{Y}, \ldots$ and when we wish to think of them as as vectors, we will use the variables $A, B, C, \ldots$.

It should be noted that $L_g$ and $R_h$ are commuting operations and so their derivatives must commute too. If $f(t), g(t)$ are two curves in $G$, then we have the product rule:

$$\frac{d}{dt}(f(t)g(t)) = (dL_{f(t)})g(t)g^{(1)}(t) + (dR_{g(t)})f(t)f^{(1)}(t)$$ \hspace{1cm} (1.13)

A consequence of this and the chain rule gives us the formula for differentiating
an inverse function:

\[
\frac{d}{dt} (g(t)^{-1}) = -(dL_{g(t)^{-1}})(dR_{g(t)^{-1}})g^{(1)}(t) \tag{1.14}
\]

Define conjugation by \(g\) as the map \(h \mapsto ghg^{-1} = (R_{g^{-1}} \circ L_g)h\). The derivative of this map at the identity is called \(Ad(g) : \mathfrak{g} \rightarrow \mathfrak{g}\), namely \(Ad(g) := (dR_{g^{-1}})g \circ (dL_g)\). This map \(Ad : G \rightarrow GL(\mathfrak{g})\) is called the adjoint representation of \(G\).

Differentiating this map at the identity, we obtain \((dAd)_1 = ad\).

Suppose now our Lie group is also a Riemannian manifold. We say the metric is left (right) invariant when \(\langle X, Y \rangle\) is constant for any left (right) invariant vector fields \(X, Y\). When a metric is both left and right-invariant, it is said to be a bi-invariant metric. The following theorem about Riemannian manifolds gives a complete description about the Levi-Civita connection

**Theorem 1.4.1. Koszul's formula [18]**

\[
2\langle \nabla_X Y, Z \rangle = \langle X, \nabla_Y Z \rangle + \langle \nabla_X Z, Y \rangle - \langle Z, \nabla_X Y \rangle - \langle [X, Y], Z \rangle - \langle [Z, X], Y \rangle
\]

Suppose now that \(X, Y, Z\) in the above formula are left-invariant vector fields. Because every term on the right hand side is either zero or is not dependent on the point on the manifold, then \(\langle \nabla_X Y, Z \rangle\) must not depend on the point either. We can see that \(\langle (\nabla_X Y)_g, Z_g \rangle = \langle (\nabla_X Y)_1, Z_1 \rangle = \langle dL_g(\nabla_X Y)_1, Z_g \rangle\) and since \(Z_g\) is arbitrary and the metric non-degenerate, it follows that \(\langle \nabla_X Y \rangle_g = dL_g(\nabla_X Y)_1\) and so \(\nabla_X Y\) is left-invariant.

Now if the metric is bi-invariant, then since both \(dL_g\) and \(dR_{g^{-1}}\) preserve the metric, we can see that

\[
\langle Ad(g)X, Ad(g)Y \rangle = \langle X, Y \rangle
\]

Taking the derivative of this relationship at \(g = 1\), one obtains the formula

\[
\langle [Z, X], Y \rangle + \langle X, [Z, Y] \rangle = 0
\]
Using this together with Koszul’s formula, we see that for left-invariant vector fields with a bi-invariant metric, we have

\[ \nabla_X Y = \frac{1}{2}[X, Y] \]

A map \( \exp : \mathfrak{g} \to G \) is defined in the following way. Let \( \mathcal{X} \in \mathfrak{g} \) and let \( \gamma(t) \) be an integral curve of \( \mathcal{X} \) satisfying \( \gamma(0) = e \). By using symmetry arguments, \( \gamma(1) \) exists, and we denote it by \( \exp(\mathcal{X}) \). Now supposing that \( \gamma^{(1)}(t) = \mathcal{X}_{\gamma(t)} \), then \( \nabla_t \gamma^{(1)} = \nabla_{\mathcal{X}} \mathcal{X} = 0 \) and so \( \gamma(t) \) is a geodesic and the Lie group definition of the exponential map coincides with the definition using geodesics.

Define the Maurer-Cartan form \( \mathcal{L} : TG \to \mathfrak{g} \) by \( \mathcal{L}(g, v) = dL_{g^{-1}}(v) \). This form is used to identify a tangent vector as that of one belonging to a left-invariant vector field and then taking the value of the field at the identity. Suppose we have a differential equation for a variable \( x(t) \) that is invariant under the action of the group. The Maurer-Cartan form can be used to take such a differential equation and factor the equation into a first order differential equation on the Lie group together with a differential equation of lower order on the Lie algebra, which is a vector space and therefore easier to analyse. This process of reducing a curve on the Lie group to one on the Lie algebra using \( \mathcal{L} \) is called the Lie reduction and we often apply this to various quantities defined over a curve \( x(t) \). The equation \( x^{(1)}(t) = dL_{x(t)} V(t) \) is called the linking equation as it links the velocity of a curve \( x(t) \) to a vector in the Lie algebra. In this thesis, \( V(t) \) will represent the Lie reduction of \( x^{(1)}(t) \), unless otherwise stated.

**Example 1.4.2.** Consider the equation \( \nabla^2_{x^{(1)}(t)} x^{(1)}(t) = 0 \). Then letting \( V(t) = \mathcal{L}(x^{(1)}(t)) \), we can rewrite the equation as

\[
\begin{align*}
x^{(1)}(t) &= dL_{x(t)} V(t) \\
V^{(2)}(t) &= \frac{1}{2}[V^{(1)}(t), V(t)]
\end{align*}
\]
Solutions to this equation have been solved for various Lie groups such as SO(3) and SL(2, \mathbb{R}) [30].

### 1.4.1 Matrix Lie groups

Subgroups of GL(n, \mathbb{R}) (and GL(n, \mathbb{C})) form a large class of Lie groups known as matrix groups consisting of matrices with the usual multiplication. In this case, the derivatives $dL_g : T_hG \to T_{gh}G$ and $dR_g : T_hG \to T_{hg}G$ are given by matrix multiplication:

\begin{align*}
  dL_g(X) &= gX \\
  dR_g(X) &= Xg
\end{align*}

Under these operations, Equations (1.13) and (1.14) simplify to:

\begin{align*}
  \frac{d}{dt} (f(t)g(t)) &= f^{(1)}(t)g(t) + f(t)g^{(1)}(t) \\
  \frac{d}{dt} (g(t)^{-1}) &= -g^{-1}(t)g^{(1)}(t)g^{-1}(t)
\end{align*}

Identifying left-invariant vector fields as the tangent space to the identity (a subset of $M_{n\times n}(\mathbb{R})$), the Lie bracket operation $[A, B] = AB - BA$ agrees with the matrix commutator.

### 1.4.2 Lax constraints

Suppose $x^{(1)}(t) = dL_{x(t)}V(t)$ where $V(t) \in \mathfrak{g}$. Suppose further that there is a curve $Z(t) \in \mathfrak{g}$ which satisfies the Lax equation

\begin{equation}
  Z^{(1)}(t) = [Z(t), V(t)]
\end{equation}

Such a $Z(t)$ is called a Lax constraint on $V(t)$, and $(V(t), Z(t))$ are called a Lax pair. Noakes proves in [22] that if $\mathfrak{g}$ is a semisimple Lie algebra, there is an open dense subset $\mathfrak{g}^{(1)}$ of $\mathfrak{g}$ such that if $Z(0) \in \mathfrak{g}^{(1)}$, then the solutions for $x(t)$ can then be computed in quadrature in terms of $V(t)$ and $Z(t)$ using algebraic operations (where $(V(t), Z(t))$ is a Lax pair). Pauley [31, Chapter 3] shows that this is in fact true for any Lie group. This means that if we have a Lax constraint, we can (almost always) reduce the problem of finding $x(t)$ to that of finding its Lie
reduction $V(t)$ and a suitable Lax constraint $Z(t)$. In the case of $G = \text{SO}(3)$, it is shown in [29] that an alternative condition to $Z(0) \in \mathfrak{g}^{(t)}$ is the condition $Z(t) \neq 0$ for all $t$.

Throughout this thesis we are more interested in the general case and so we may say informally that a system can be solved in quadrature when strictly speaking, we mean that there is an open dense set of initial values for which the system can be solved. More often than not, we will be working in $\text{SO}(3)$ where we simply require $Z(t)$ to not be 0 anywhere.
Consider a situation where we have a sequence of data points\(^1\) (or nodes) \(\xi_0, \ldots, \xi_N\) in ordinary \(n\) dimensional Euclidean space, \(\mathbb{R}^n\). A natural question to ask is how we can find a well-behaved curve that also passes through these points. A first attempt at answering such a question would be to join the points up by \(N\) line segments, and indeed in some circumstances this piecewise-affine interpolant is sufficient for the task at hand. On the other hand, we may impose restrictions on what curves we may use to connect the points up with; We may prefer a curve without the kinks in the previous example and so could require the curve to be differentiable. In a broad sense, interpolation is the study of the different ways of filling in the gaps with certain conditions placed on just how we can do the filling.

A slightly specialised case of the interpolation problem is one where we also have prescribed a sequence of time values (knots) \(0 = t_0 < t_1 < \cdots < t_N = 1\) and we require the curve \(x: [0, 1] \to \mathbb{R}^n\) to satisfy the condition that \(x(t_i) = \xi_i\) for all \(i\). A polynomial curve of order \(r\) in \(\mathbb{R}^n\) is one that can be written as \(p(t) = \sum_{i=0}^{r-1} p_i t^i\) where each \(p_i \in \mathbb{R}^n\). If we require an interpolant to be \(C^k\), a \(k\)th differentiable curve, one solution to the interpolation problem can be achieved using curves defined piecewise by polynomial curves of order \(k + 2\) on each interval \((t_i, t_{i+1})\). These class of interpolants are known as polynomial splines of order \(k + 2\) which have been well studied (See \([8, 2, 41]\)). De Boor states in \([8, \text{Chapter IV}]\) that the most popular choice for interpolating using polynomial splines continues to be polynomial splines of order 4, also known as cubic splines, for \(C^2\) interpolation. There are \(2n\) degrees of freedom in a choice of a cubic spline interpolant and the natural interpolating cubic spline is defined as the unique interpolating cubic spline where the initial and final accelerations are zero.

Instead of simply finding interpolants that satisfy a certain level of differentiability, sometimes we are more interested in an interpolant that best fits a certain

\(^1\)The notation \(\xi\) is preferred for data points to distinguish them from other points
model of how we expect the curve should behave. The notion of best-fitting can be represented by the curve minimising some objective functional. Variational approaches to choosing interpolants began with a famous theorem by Holladay [2] that states for $C^2$ interpolants $x(t)$, the quantity $\int_0^1 \|x^{(1)}\|^2 dt$ is minimised by a natural cubic spline. It is for this reason Schoenberg [8, Chapter IV] coined their name due to the resemblance to the mechanical splines used by draftsmen [39].

Interpolation arises in many real life contexts including graphics, robot path planning, and medical imagining and a survey of the history of interpolation in various contexts can be seen in [17]. In some situations, the choice of data points and interpolants may naturally be described as belonging in a Riemannian manifold. Although the subject of interpolation has been thoroughly studied for Euclidean spaces, the more general setting of interpolating in a Riemannian manifold has only relatively recently received attention.

To motivate the study of interpolation in Riemannian manifolds, we will consider a simple example where we have a hypothetical video camera that can rotate in all directions and so its state can be thought of as a point in the manifold $\text{SO}(3)$. We wish to control the orientation of the camera subject to the condition that it must face specified orientations at fixed points in time. If this problem was in Euclidean space, one choice of interpolant could be to simply connect the camera positions up with straight lines. On the other hand, if we view $\text{SO}(3)$ as a subset of $\mathbb{M}^{3 \times 3}$, a straight line between any two data points would not lie in $\text{SO}(3)$. We might then consider what would happen if we simply projected an ordinary Euclidean space interpolant onto $\text{SO}(3)$. Consider the two data points given by:

\[
\xi_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \xi_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\]

(2.1)

The obvious choice for an interpolant in $\mathbb{M}^{3 \times 3}$ would be the straight line connecting $\xi_0$ and $\xi_1$. If we were to then project this onto $\text{SO}(3)$, the projection would simply be $\xi_0$ for the first half of the line, undefined at the midpoint of the line.
and $\xi_1$ for the second half. This is clearly an undesirable way of trying to fix the problem of the interpolant not remaining on SO(3). A different approach still is to use charts of SO(3) to carry out the interpolation. The problems with charts are numerous including unstable sensitivity to initial conditions, distortion of the geometry (such as choosing unnecessarily large paths) and a strong dependence on the choice of charts (See [19, 20]).

A better set of interpolants to choose from would be those defined within the manifold which do not depend on a choice of coordinates. One way of generating these is by finding solutions to chart-independent variational problems defined on the manifold, for example using minimal geodesics which minimise norm squared velocity. Connecting the data points up with minimal geodesics would solve the interpolation problem without any of the previous problems. If we required our curve to be differentiable, simply using geodesic splines would not be enough to guarantee this (even in Euclidean space).

In order to produce interpolants that are $C^2$, Gabriel and Kajiya [10] and Noakes et al. [20] independently defined a class of curves called Riemannian cubics, generalising cubic polynomial curves in Euclidean space to a Riemannian manifold. To see this, first consider the straight line in Euclidean space. We could define a line by various properties, perhaps the most obvious way is a curve with zero acceleration, $x^{(2)}(t) = 0$ and can be generalised to a Riemannian manifold by saying that $\nabla_t x^{(1)}(t) = 0$. This is a purely differential condition rather than a geometric condition and if our aim is to capture geometric properties of the Euclidean interpolants, this generalisation may not be suitable. If we choose instead to define a line by a curve that locally minimises arc-length between two points, we obtain a geometric description of a line although the choice of parametrisation of the line is ambiguous. Taking curves that locally minimise mean squared velocity (or mean kinetic energy) on the other hand produces the same curves with uniform speed. That is, a generalised line on a Riemannian manifold are critical points of the functional $J_1(x) = \int_0^1 \|x^{(1)}\|^2 dt$, these are just geodesics of the manifold. This geometric condition then imposes differential
conditions known as the Euler-Lagrange equations which in the case of lines is the same as the differential condition we had above.

Cubic polynomials in Euclidean space have a differential description that says they are those curves which satisfy their fourth derivative being zero, \( x^{(4)}(t) = 0 \). Holladay’s theorem on the other hand gives a geometric condition saying that they locally minimise mean squared acceleration. The differential condition for cubic polynomials generalises to the statement that \( \nabla^3_t x^{(1)}(t) = 0 \) whereas the geometric condition generalises cubics to critical points of \( J_2(x) = \int_0^1 \| \nabla_t x^{(1)}(t) \|^2 dt \). The Euler-Lagrange equations for the geometric description is given in [20]. The two alternative ways of characterising regular cubic polynomials leads to two starkly different generalisations to manifolds. In the case of cubics, the geometric point of view generalises to Riemannian cubics and the differential point of view generalises to Jupp and Kent cubics [14]. One goal of the existing research is to efficiently piece together Riemannian cubics to create the analogue of the \( C^2 \) Euclidean cubic splines used in interpolation [25]. Noakes in [27] approximates natural Riemannian cubic splines in the case the data is nearly geodesic. Chapter 5 focuses mainly around this principle and applies it to the case of second order conditional extrema on a Lie group.

### 2.2 Riemannian Cubics in Tension

Riemannian cubics in Tension were first introduced by Silva Leite et al. in [43, 44] as critical points of \( J_2^λ(x) = \int_0^1 \| \nabla_t x^{(1)} \|^2 + \lambda \| x^{(1)} \|^2 dt \) where it was also mentioned that these curves had practical applications in the Euclidean space by “tightening” the interpolants one would get if they had simply used Riemannian cubics. We prove in this thesis that this indeed is the right interpretation as these Riemannian cubics in tension arise as solutions to the constrained optimisation problem where we minimise \( J_2 \) subject to a constraint on the arc-length. Popiel and Noakes [34, 35] prove various results concerning extendibility of such curves, describe various constants of motion and describe their internal symmetry. Hussein and Bloch in [12] also study Riemannian cubics in tension (which they refer
to as $\tau$-elastic extremals) and prove some additional constraints regarding the motion, with applications to multiple spacecraft interferometric imaging. We are particularly interested in the case where our manifold is $\text{SO}(3)$ for which Popiel and Noakes also describe some long term asymptotic behaviour.

In this thesis, we first look at a specific class of Riemannian cubics in tension known as null Riemannian cubics in tension. While these are a special case of Riemannian cubics in tension which were originally defined variationally, they are exactly the differential analogues of hyperbolic trigonometric functions which have found use in interpolation applications [41, 51] in Euclidean space. Motivated by the results of Noakes in [24], Chapter 3 of this thesis provides significantly tighter asymptotic behaviour for null Riemannian cubics in tension in $\text{SO}(3)$ and suggests a method for which the asymptotic curves may be used for interpolation applications.

2.3 Prior fields

As stated before, interpolation should depend on what assumptions are made regarding the dynamics of the system which we are interpolating. Noakes [26] introduces the concept of a prior vector field $\mathcal{A}$ which describes the known dynamical behaviour of a system. The integral curves of $\mathcal{A}$ represent a model for the dynamical behaviour of $x(t)$, but in practice, one may find that the actual behaviour of the system only closely resembles this behaviour. If we are given a series of points $\xi_0, \ldots, \xi_N$ to interpolate which do not lie on any integral curve of $\mathcal{A}$, then we may attempt to find an interpolant that minimises $\mathcal{A}J_1(x) = \int_0^1 \|x^{(1)} - A_\varepsilon\|^2 dt$.

In the same paper, Noakes derives the Euler-Lagrange equations for the system and describes various solutions in special cases when the manifold and vector field display symmetry. In particular if $Q$ is a Lie group with a bi-invariant metric and $\mathcal{A}$ is a left-invariant vector field, then minimisers of $\mathcal{A}J_1$ are curves of the form $x(t) = \exp(At) \exp((V_0 - A)t)$ where $A = A_e$.

This thesis further continues the study of conditional extremals by studying their so called $\mathcal{A}$-Jacobi fields. An equation for the $\mathcal{A}$-Jacobi fields is derived
Chapter 2. Literature Review

for a Riemannian manifold. When the manifold is a bi-invariant Lie group and the prior field is left-invariant, a simpler equation to describe the $\mathcal{A}$-Jacobi fields exists. Explicit solutions are then derived as well as geometric properties of the solutions. In addition to the study of first order prior fields originally studied by Noakes, we introduce in this thesis second order and time varying prior fields.

We set up a framework and derive several exploratory results concerning the conditional extremals in these settings including a large class of solutions for affine time varying prior fields and extendibility results for second order prior fields in a bi-invariant Lie group. In Chapter 5, we consider an algorithm to efficiently solve for interpolation systems of such conditional extremals.

There are two differing points of view regarding interpolation in manifolds; these concern whether or not the interpolant is calculated extrinsically or intrinsically. The extrinsic point of view takes the manifold as an embedded submanifold of $\mathbb{R}^n$ and carries out the computations by incorporating the embedding. Work in this area, particularly for the space of quaternions due to its application to rigid body motion, include Barr et al. [3], Crouch et al. [7], and Pottman [37]. Intrinsic formulations appeared slightly before and key papers include Gabriel and Kajiya [10] and Noakes et al. [28] where minimizers of covariant acceleration are proposed, namely Riemannian cubics. The intrinsic formulation holds closer to the underlying philosophy of Riemannian geometry and we only consider in this thesis extrinsic interpolation schemes for comparison purposes. The thesis contributes to the active area of research on intrinsic methods of interpolation in Riemannian manifolds.
Asymptotics of Null Riemannian Cubics in Tension in 
\( \text{SO}(3) \)

3.1 Introduction

There are problems in mechanics, computer vision and engineering that reduce naturally to interpolation of points by curves in Riemannian manifolds.

Methods for interpolation in manifolds are rather undeveloped compared with what exists for Euclidean spaces. A popular class of interpolants in Euclidean spaces are natural cubic splines, which has a variational generalisation in Riemannian manifolds called natural Riemannian cubic splines. Piecewise Riemannian cubics in tension are curves more useful for interpolation than natural Riemannian cubic splines when more information is known about the length of the curve. Riemannian cubics in tension are still not well understood, but some geometric properties are known for the so-called null case. The asymptotic behaviour of such curves is useful for a general understanding and for designing interpolation algorithms. We first describe how Riemannian cubics in tension arise naturally in a constrained optimisation setting and then provide accurate asymptotics for null Riemannian cubics in tension in the Lie group \( \text{SO}(3) \) which is of special interest due to applications in computer graphics and rigid-body trajectory planning.

Example 3.1.1. Consider a mechanical system described by Lagrangian
\[ L(x(t), x'(t)) \]
where \( x(t) \) lies in some configuration space \( Q \). Sometimes \( Q \) is the Euclidean space \( E^n \) but more often \( Q \) is a subset of \( E^k \) for \( k \geq n \), usually a connected smooth manifold. Typically, \( L(x, q'(1)) = \langle x'(1), x'(1) \rangle - U(x) \) where the kinetic energy \( \langle x'(1), x'(1) \rangle \) is given by a Riemannian metric \( \langle \cdot, \cdot \rangle \) on \( Q \) and \( U : Q \to \mathbb{R} \) is a potential function. When \( U \) is constant, then \( x : \mathbb{R} \to Q \) is called a geodesic. In practice, \( U \) may be unknown and nearly but not exactly constant (in the \( L^2 \) sense) and the data set \( \xi_0, \ldots, \xi_n \in Q \) may be observed at times \( t_0, \ldots, t_n \). The problem is to estimate \( x(t) \) for all \( t_0 < t < t_n \) given this information. If we knew
Chapter 3. Asymptotics of Null Riemannian Cubics in Tension in SO(3)

We could attempt to solve the Euler-Lagrange equations

$$\nabla_t x^{(1)} = -\text{grad}(U)$$

where $\nabla$ is the Levi-Civita covariant derivative associated with $\langle , \rangle$ and the right hand side is proportional to the force exerted by the potential $U$. If $U$ is unknown but assumed nearly constant, then the force $\text{grad}(U)$ should be, on average, small.

We might choose an interpolant $x : [t_0, t_n] \to Q$ satisfying $x(t_i) = \xi_i$ for $i = 0, \ldots, n$ that minimises

$$J_2(x) = \int_{t_0}^{t_n} \|\text{grad}(x(t))\|^2 dt = \int_{t_0}^{t_n} \|\nabla_t x^{(1)}\|^2 dt. \quad (3.1)$$

Such interpolants are examples of natural Riemannian cubic splines [20, 44]. If in addition, an upper bound $K$ is known for the length of the curve $x$, then with the inequality

$$\int_{t_0}^{t_n} \|x^{(1)}\|^2 dt \leq \left( \int_{t_0}^{t_n} \|x^{(1)}\| dt \right)^2 = K^2.$$

in mind, a better choice of interpolant would then be to minimise $J_2$ subject to the additional condition $\int_{t_0}^{t_n} \|x^{(1)}\|^2 dt \leq K^2$. We will see later that such curves are piecewise Riemannian cubics in tension as first described by Popiel and Noakes [34], who have also described various other applications for these curves in interpolation problems.

For practical reasons, we will now consider a specialised situation where we have only two observed values and known initial and final velocities. For $i \in \{0, 1\}$, fix $q_i \in Q$ and $v_i \in T_{q_i}Q$ and let $\mathcal{C}_{v_0, v_1}$ be the space of smooth curves $x : [0, 1] \to Q$ satisfying $x(i) = q_i$ and $x^{(1)}(i) = v_i$. We say that a curve $x$ is a critical point of a functional $J : \mathcal{C}_{v_0, v_1} \to \mathbb{R}$ when for any smooth vector field $X$ whose value and derivative vanish at the endpoints of the curve, we have:

$$\lim_{s \to 0} \frac{J(\Phi_s(x)) - J(x)}{s} = 0 \quad (3.2)$$
where $\Phi$ is a local flow of $X$ taken pointwise on the curve $x(t)$. The geodesic energy $J_1$ of a curve in $C_{v_0,v_1}$ is given by $J_1(x) = \int_0^1 \|x^{(1)}(t)\|^2 dt$ and critical points of $J_1$ are called geodesics. Riemannian cubics on the other hand are critical points of the higher order energy functional $J_2$ defined by $J_2(x) := \int_0^1 \|\nabla_t x^{(1)}(t)\|^2 dt$. One way of trading off between geodesics and Riemannian cubics leads to a class of curves called Riemannian cubics in tension (abbreviated $RCT$) (See [44, 34]) which are critical points of the functional $J^\lambda_2 : C_{v_0,v_1} \rightarrow \mathbb{R}$ defined by:

$$J^\lambda_2(x) := \int_0^1 (\|\nabla_t x^{(1)}(t)\|^2 + \lambda\|x^{(1)}(t)\|^2) dt$$  \hspace{1cm} (3.3)

with $\lambda > 0$. The curves obtained when $\lambda < 0$ lie outside the scope of this thesis.

With Riemannian curvature defined by $R(X,Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]}Z$, we have the following theorem:

**Theorem 3.1.2** (Silva Leite et al. [44]). $x \in C_{v_0,v_1}$ is a critical point of $J^\lambda_2$ if and only if

$$\nabla^3_t x^{(1)}(t) + R(\nabla_t x^{(1)}(t), x^{(1)}(t))x^{(1)}(t) - \lambda \nabla_t x^{(1)}(t) = 0$$  \hspace{1cm} (3.4)

for all $t \in [0,1]$.

**Example 3.1.3.** Suppose $Q = E^n$, i.e. $\mathbb{R}^n$ with $\langle \cdot, \cdot \rangle$ the Euclidean inner product. Then (3.4) reads $x^{(1)}(t) - \lambda x^{(2)}(t) = 0$. So $x(t) = c_1 e^{\sqrt{\lambda}t} + c_2 e^{-\sqrt{\lambda}t} + c_3 t + c_4$ for some $c_1, c_2, c_3, c_4 \in E^n$.

Riemannian cubics in tension were originally conceived [33] as an ad hoc fix to lengthy piecewise Riemannian cubic spline interpolants by adding a penalty term corresponding to arc length. They also arise naturally in a constrained optimisation setting which we have seen before in Example 3.1.1 and is explained by the following Proposition.

**Proposition 3.1.4.** $C^\infty$ curves obtained by minimizing $J_2$ under the constraint that $J_1$ is bounded ($J_1(x) \leq c$) are Riemannian cubics in tension with $\lambda \geq 0$
**Proof.** The Sobolev space $H^2(I, Q)$ is a smooth Hilbert manifold and the maps $J_1, J_2 : H^2(I, Q) \to \mathbb{R}$ are smooth (See Eliasson [9]). In particular, there are bounded linear Fréchet derivatives $(dJ_1)_x, (dJ_2)_x$ at any $x \in H^2(I, Q)$ between the Hilbert spaces $H^2(I, \mathbb{R}^n)$ and $\mathbb{R}$. Let $Q'$ denote the closed submanifold with boundary $\{x \in H^2(I, Q) : J_1(x) \leq c\}$ for $c$ a regular value of $J_1$. Let $x$ be a $C^\infty$ critical point of $J_2$ in $Q'$ in the sense that $(dJ_2)_x(v') = 0$ for all $v' \in T_xQ'$. If $J_1(x) < c$, then $x$ is a critical point of $J_2$ in $Q$ and so it is a Riemannian cubic in tension ($\lambda = 0$). We will therefore assume that $J_1(x) = c$. We know that $T_x\partial Q' \subseteq \ker((dJ_1)_x, \ker((dJ_2)_x)$ and because $c$ is a regular value, $T_x\partial Q'$ has codimension 1 in $T_xH^2(I, Q)$. Therefore as $(dJ_1)_x \neq 0$, we must have that $(dJ_2)_x = -\lambda(dJ_1)_x$ for some $\lambda \in \mathbb{R}$ and so in particular, $d(J_2 + \lambda J_1)_x = 0$ and so $x$ is a critical point of (3.3). As we are minimising $J_2$, $\lambda$ must be non-negative. To see this, if $\lambda < 0$, then let $v$ be a vector in $T_xH^2(I, Q)$ such that $(dJ_1)_x(v) < 0$ meaning $v \in T_xQ'$, but $(dJ_2)_x(v) < 0$ contradicting the fact $J_2$ is minimised in $Q'$ at $x$.

If $c$ is not a regular value, then either $x$ is a geodesic (which by Theorem 3.1.2 is a Riemannian Cubic in Tension) or $x$ has a neighbourhood for which $c$ is a regular value, completing the argument. Because of the denseness of $C^\infty(I, Q)$ in $H^2(I, Q)$ [1] (relative to the Sobolev norm), it follows that critical points of smooth functionals (relative to the same norm) as defined originally are also critical points in the sense described above ($(dJ)_x = 0$).

This proposition is the familiar Karush-Kuhn-Tucker conditions on a Sobolev space (we could not find a suitable reference).

Now let $Q$ be a Lie group $G$ with bi-invariant Riemannian metric, identity $e$, Lie algebra $\mathfrak{g} := T_eG$ and Lie bracket $[\cdot, \cdot]$. Recall[18] that bi-invariance of a Riemannian metric $\langle \cdot, \cdot \rangle$ is equivalent to:

$$\langle [u, v], w \rangle = \langle u, [v, w] \rangle \quad \text{for all } u, v, w \in \mathfrak{g} \quad (3.5)$$

Let $L_g : G \to G$ be the left multiplication by $g \in G$. Let $I$ be an open interval. For a $C^\infty$ curve $x : I \to G$, define the Lie reduction $V : I \to \mathfrak{g}$ of $x(1)$ by the
3.1. Introduction

Linking equation

\[ V(t) := (dL_{x(t)^{-1}})_{x(t)}x^{(1)}(t) \] (3.6)

**Proposition 3.1.5** (Silva Leite et al. [44]). \( x : I \rightarrow G \) is a RCT if and only if

\[ V^{(3)}(t) = [V^{(2)}(t), V(t)] + \lambda V^{(1)}(t) \] (3.7)

for all \( t \in I \).

A solution \( V : I \rightarrow g \) of (3.7) is called a Lie quadratic in tension, abbreviated LQT. \( V(t) \) is considerably easier to analyse than \( x(t) \) as it satisfies only a quadratic second order ODE in a vector space rather than a non-linear fourth order ODE in a manifold. It is possible ([35], [23]) to recover \( x(t) \) from \( V(t) \) and so we will therefore focus on \( V(t) \).

**Corollary 3.1.6.** \( V : I \rightarrow g \) is a LQT if and only if

\[ V^{(2)}(t) = [V^{(1)}(t), V(t)] + \lambda V(t) + C \] (3.8)

for some constant \( C \in g \) and all \( t \in I \).

If \( C = 0 \), \( V(t) \) is said to be a null Lie quadratic in tension (abbreviated NLQT) with the corresponding cubic called a null Riemannian cubic in tension (abbreviated NRCT). Null and non-null LQTs are considerably different in character and require separate analysis. We will briefly discuss a relation between null Riemannian cubics in tension, Jupp and Kent quadratics as described in [30] and null Riemannian cubics.

### 3.1.1 Null Riemannian Cubics in Tension and their connection to other families of curves

Acting on differentiable functions \( f^{(1)} : \mathbb{R} \rightarrow \mathbb{R} \), we have a linear differential operator \( D^1 = \frac{d}{dt} \) whose eigenfunctions are solutions \( f^{(1)}(t) \) to the equation:

\[ f^{(2)}(t) - \lambda f^{(1)}(t) = 0 \] (3.9)
The reason for denoting the function by \( f^{(1)}(t) \) and not \( f(t) \) should soon become clear. The nullspace of \( D^1 \) is the space of constant functions and the remaining eigenfunctions are exponentials. If we look at the possible solutions for \( f(t) \), we see that we have either affine lines or affine line segments reparameterised by \( e^{\lambda t} \). Consider now the differential analogue of Equation (3.9) on a Riemannian manifold:

\[
\nabla_t x^{(1)}(t) - \lambda x^{(1)}(t) = 0
\]

Although an exact comparison of the \( \mathbb{R} \)-linear operator \( \nabla_t \) cannot be made as it may act on sections over different curves and thus different vector spaces, the solutions corresponding to \( \lambda = 0 \) have covariantly constant velocity, geodesics. The remaining solutions for \( x^{(1)}(t) \) are constant vectors scaled by an exponential function and so solutions for \( x(t) \) are reparameterised geodesics. That is to say, suppose \( \gamma(t) \) is a geodesic, then \( \tilde{\gamma}(t) = \gamma\left(\frac{t}{\lambda}\right) \) is a solution for \( \lambda \neq 0 \).

Consider now the operator \( D^2 = \frac{d^2}{dt^2} \). Here the eigenfunction equation for real valued functions is

\[
f^{(3)} - \lambda f^{(1)}(t) = 0
\]

Solutions for \( f^{(1)}(t) \) are affine when \( \lambda = 0 \) and \( f^{(1)}(t) = c_1 \cosh \sqrt{\lambda} t + c_2 \sinh \sqrt{\lambda} t \) when \( \lambda \neq 0 \). The solutions for \( f(t) \) are quadratic polynomials when \( \lambda = 0 \) however for \( \lambda \neq 0 \), they are functions of the form \( f(t) = d_1 \sinh \sqrt{\lambda} t + d_2 \cosh \sqrt{\lambda} t + d_3 \) which are quadratics reparameterised by solutions of \( f^{(1)} \) but for a different choice of \( \lambda \). Finally, let us consider the operator \( \nabla_t^2 \) and the eigenfunction equation:

\[
\nabla_t^2 x^{(1)}(t) - \lambda x^{(1)} = 0
\]

Solutions for \( x(t) \) to this equation when \( \lambda = 0 \) are known as Jupp and Kent quadratics [31, Chapter 4], or JK-quadratics for short. If we take the Lie reduction of Equation (3.12), we see that \( V^{(2)}(t) + \frac{1}{2}[V, V^{(1)}] - \lambda V = 0 \). Under the
transformation $\tilde{V} = \frac{1}{2}V(t)$, we see that $\tilde{V}^{(2)} + [\tilde{V}^{(1)}, \tilde{V}] - \lambda \tilde{V} = 0$. Solutions to this for $\lambda = 0$ are null Lie quadratics and the solutions when $\lambda > 0$ are null Lie quadratics in tension. One may ask whether null Riemannian cubics in tension are simply reparameterised null Riemannian cubics. Suppose $x(t) = y(f(t))$ where $x(t)$ is a NRCT and $y(t)$ is a NRC. Then

$$V_x(t) = x^{-1}(t)x^{(1)}(t)$$
$$= y(f(t))^{-1}y^{(1)}(f(t))f^{(1)}(t)$$
$$= V_y(f(t))f^{(1)}(t)$$

So suppose $V_x(t) = g(t)V_y(f(t))$. Then $V_x^{(1)} = g^{(1)}V_y(f) + gV_y'(f)f^{(1)}$ and $V_x^{(2)} = g^{(2)}V_y(f) + 2g^{(1)}V_y'(f)f^{(1)} + gV_y''(f)(f^{(1)})^2 + gV_y'(f)f^{(2)}$.

$$V_x^{(2)} + [V_x, V_x^{(1)}] - \lambda V_x = g^{(2)}V_y(f) + (2g^{(2)}f^{(1)} + gf^{(2)})V_y'(f) + gV_y''(f)(f^{(1)})^2$$
$$+ [gV_y(f), g^{(1)}V_y(f) + gV_y'(f)f^{(1)}] - \lambda gV_y(f)$$
$$= g(f^{(1)})^2V_y''(f) + g^{(2)}V_y(f) + (2g^{(1)}f^{(1)}$$
$$+ gf^{(2)})V_y'(f) + g^2f^{(1)}[V_y(f), V_y'(f)] - \lambda gV_y(f)$$

For the right hand side to be the same as the NLQ equation, we must have

$$g(f^{(1)})^2 - g^2 f^{(1)} = 0 \quad (3.13)$$
$$2g^{(1)}f^{(1)} + gf^{(2)} = 0 \quad (3.14)$$
$$g^{(2)} - \lambda g = 0 \quad (3.15)$$

In solving these equations, we see from (3.15) that $g = Ae^{\sqrt{\lambda}t} + Be^{-\sqrt{\lambda}t}$. (3.13) then implies $f = \frac{A}{\sqrt{\lambda}}e^{\sqrt{\lambda}t} - \frac{B}{\sqrt{\lambda}}e^{-\sqrt{\lambda}t} + C$. But (3.14) then says $3\sqrt{\lambda}(A^2e^{2\sqrt{\lambda}t} - B^2e^{-2\sqrt{\lambda}t}) = 0$ which would imply $\lambda = 0$ or $A = B = 0$. This shows that NRCs and NRCTs although closely related cannot just be expressed in terms of the other.
Authors such as Schumaker [41], Zhang et al. [51] make use of hyperbolic trigonometric functions in various interpolation applications. These motivate the study of NRCT in their own right as they represent a differential analogue to those of hyperbolic trigonometric functions which have already found use in interpolation. NRCT represent a third order differential equation on the manifold and moreover when the manifold is complete, solutions exist (Philip Schrader, private communication) for any pair of points and initial velocity, which means they are suitable curves to use for \( C^1 \) interpolation.

An important simplification for LQTs arise when we see that under the transformation \( V(t) \mapsto \sqrt{\lambda} V(\sqrt{\lambda}s) \), we can rearrange (3.8) to

\[
V^{(2)}(s) = [V^{(1)}(s), V(s)] + V(s) + C/\lambda^{3/2}.
\]

With the above discussion in mind, we will now focus on the null case in this chapter and assume without loss that \( \lambda = 1 \). The following propositions are proved in [34].

**Proposition 3.1.7** (Noakes, Popiel). Let \( V(t) \) be a NLQT. Then there is a unique solution to \( V(t) \) defined for all \( t \in \mathbb{R} \) for a given choice of \( V(t_0), V^{(1)}(t_0), \) and \( V^{(2)}(t_0) \). There also exists constants \( b, c, d_0, d_1 \in \mathbb{R} \) such that \( \forall t \in \mathbb{R} \),

\[
\frac{d^2}{dt^2} (\|V(t)\|) = 4\|V(t)\|^2 + 2b \quad (3.16)
\]

\[
\|V^{(1)}(t)\|^2 = \|V(t)\|^2 + b \quad (3.17)
\]

\[
\|V^{(2)}(t)\|^2 = \|V^{(1)}(t)\|^2 + c \quad (3.18)
\]

\[
\|V(t)\|^2 = d_0 e^{2t} + d_1 e^{-2t} - \frac{b}{2} \quad (3.19)
\]

\[
\|V^{(1)}(t)\|^2 = d_0 e^{2t} + d_1 e^{-2t} + \frac{b}{2} \quad (3.20)
\]

\[
\|V^{(2)}(t)\|^2 = d_0 e^{2t} + d_1 e^{-2t} + \frac{b}{2} + c \quad (3.21)
\]

Moreover, when \( G \) is SO(3), an additional choice for \( x(t_0) \) gives a unique solution to \( x(t) \).

We assume that \( V \) is nondegenerate which means that \( d_0, d_1 \) as defined in Corollary 3.1.7 are both not zero. For convenience we will define the functions \( F(t) := \|V(t)\|^2 \) and \( f(t) := \|V(t)\| \). The two different names for \( \|V\| \) and
\[\|V\|^2\] may seem superfluous however for calculations involving derivatives, will become handy later. An observation we will use throughout the chapter is that if \( g(t) = O(f(t)^\gamma) \) with \( \gamma < 0 \), then \( \int_{t}^{\infty} g(s)ds = O(f(t)^\gamma) \) because \( f(t) \) is of order \( e^t \).

The case where \( Q \) is the Lie group SO(3) of rotations of \( E^3 \) has practical applications in computer graphics and trajectory planning [20]. From now on we assume that \( G = SO(3) \). We identify its Lie algebra \( so(3) \) with \( E^3 \) using the Lie algebra isomorphism \( B : E^3 \to so(3) \) given by \( B(v)(w) = v \times w \).

When talking about asymptotics, we are referring to describing or approximating the behaviour of a function with respect to a parameter \( t \), usually as \( t \to \pm \infty \). In interpolation applications, asymptotics may be used within shooting methods for solving a boundary value problem [15] because the asymptotics of \( V(t) \) produce a reasonable estimate of \( V(t) \) for moderate values of \( t \) as well. Asymptotics turn out useful in other scenarios too: in [24] Noakes shows that there is a null Riemannian cubic underlying the motion of a naturally occurring mechanical system and the asymptotics computed can be used to approximate the state of the system at moderate to large time. In this chapter, we find accurate asymptotics for both null Lie quadratics in tension in \( so(3) \) and the null Riemannian cubics in tension in \( SO(3) \).

### 3.2 Asymptotics of Null Lie Quadratics in Tension in \( so(3) \)

Let \( U(t) = V/f \). Then it has been shown in [34] that we have:

\[
\|U^{(1)}(t)\|^2 = \frac{b + c}{F^2}.
\]  

Using Equation (3.22), it was also shown using the arc-length inequality \( \|U(a) - U(b)\| \leq \int_{a}^{b} \|U^{(1)}(s)\|ds \) that \( \lim_{t \to \infty} U(t) = \alpha \) and that \( \|V(t) - f(t)\alpha\| = O(1/f) \).

Here we can think of \( \alpha \) as a vector that describes the limit axis \( V(t) \) tends towards.

We aim in the rest of this chapter to achieve a much tighter asymptotic formula for such non-degenerate null Lie quadratics in tension in \( E^3 \). We will consider
only the asymptotics for \( t \geq 0 \) due to the internal symmetry property of null LQTs described in [34]
Lemma 3.2.1 and the fact that $\langle V, V^{(2)} - V \rangle = 0$ together imply that 

$$\| U''(t) \|^2 = F^3 / (b + c) + \left( \frac{b^2}{4} - 4d_0d_1 / (b + c) \right)^2 \text{ which means } 1 / \| U''(t) \|^2 = (b + c) / F^3 + O(1 / F^6).$$

The centre of curvature $\hat{U}(t)$ can now be computed using the formula:

$$\hat{U}(t) = U + \frac{U''(t)}{\| U''(t) \|^2} = U + \frac{V^{(2)} - V}{F^3 / 2} + \left( \frac{b^2}{4} - 4d_0d_1 \right) V + O \left( \frac{1}{F^{9/2}} \right). \quad (3.24)$$

As $\hat{U}(t)$ is rather complicated, we will simplify by taking $\hat{U}(t) := U + \frac{V^{(2)} - V}{F^3 / 2}$. Set $V(t) := f(t)\hat{U}(t)$.

**Proposition 3.2.2.** For $t \geq 0$, $\| \hat{V}(t) - f(t)\alpha \| = O(1 / f^2)$

**Proof.** For $t \geq 0$, differentiating $\hat{U}(t)$, we find $\frac{d}{dt} \hat{U}(t) = ((V^{(1)}F + VF^{(1)} + V^{(3)} - V^{(1)}F^3 / 2 - \frac{3}{2} F^{1/2} F^{(1)}(VF + V^{(2)} - V)) / F^3$. We can see that $V^{(3)} - V^{(1)} = V^{(2)} \times V = (V^{(1)} \times V + V) \times V = FV^{(1)} - \frac{1}{2} F^{(1)} V$ by Lagrange’s formula for triple cross products. Substituting this in, we find:

$$\frac{d}{dt} \hat{U}(t) = \frac{-3F^{(1)}(V^{(2)} - V)}{2F^{3/2}} \quad (3.25)$$

And so

$$\| \frac{d}{dt} \hat{U}(t) \|^2 = \frac{9(F^{(1)})^2(b + c)}{4F^5} \quad (3.26)$$

We then have that $\| \alpha - \hat{U}(s) \| \leq \int_s^\infty \| \frac{d}{dt} \hat{U}(t) \| dt = \int_s^\infty \frac{3F^{(1)}\sqrt{b + c}}{2F^{3/2}} dt$.

Because both $F$ and $F^{(1)}$ are of order $e^{2t}$, we have that $\| \alpha - \hat{U}(t) \| = O(1 / f^3)$. Multiplying both sides by $f(t)$ proves the result.

Because $\| \hat{V}(t) - V(t) \| = (\sqrt{b + c}) / F$, we also have that $\| V(t) - f(t)\alpha \| = O(1 / f^2)$. This also implies that $\langle V, V \rangle - 2f(t)\langle V, \alpha \rangle + f(t)^2 = O(1 / f^4)$ and so therefore $\langle V, \alpha \rangle = f + O(1 / f^5)$.

Recall that we have a Lie algebra isomorphism $B : E^3 \rightarrow so(3)$. Consider the orthogonal endomorphism $I_\alpha = B(\alpha)|_H : H \rightarrow H$ where $H = \{ v \in E^3 | \langle v, \alpha \rangle = 0 \}$ is the plane orthogonal to $\alpha$. By Equation (3.5), $I_\alpha$ is skew-adjoint and $I_\alpha^2 = -I_{3 \times 3}$. 

3.2. Asymptotics of Null Lie Quadratics in Tension in so(3)
So \( I_\alpha \) defines a complex structure on \( H \). We will denote the projection of a vector \( v \in E^3 \) onto \( H \) by \( v_\alpha := v - \langle v, \alpha \rangle \alpha \).

We would like to get a handle on the behaviour of \( V(t) \) in the directions perpendicular to \( \alpha \). Setting \( Z(t) := V^{(2)}(t) - V(t) = V^{(1)}(t) \times V(t) \), we also have \( Z^{(1)}(t) = Z(t) \times V(t) \). Notice that both \( Z(t) \) and \( Z^{(1)}(t) \) are close to being orthogonal with \( \alpha \).

**Lemma 3.2.3.** For \( t \geq 0 \), \( \langle \alpha, Z(t) \rangle = O(1/f^3) \)

**Proof.** \( \langle Z(t), \alpha \rangle - \langle Z(t), \tilde{U}(t) \rangle = \langle Z(t), \alpha - \tilde{U}(t) \rangle = O(1/f^3) \) by the Cauchy-Schwarz inequality but \( \langle Z(t), \tilde{U}(t) \rangle = (b+c)/f^3 \) so therefore \( \langle Z(t), \alpha \rangle = O(1/f^3) \).

The following theorem says that there is a vector \( \beta \), thought of as describing the plane orthogonal to it, such that the behaviour of \( Z(t) \) limits towards a known function in that plane. The theorem does not provide a way of computing what \( \beta \) is.

**Theorem 3.2.4.** For \( t \geq 0 \), and some unit vector \( \beta \in H \),

\[
Z(t) = (\sqrt{b+c}) \exp(-g(t)I_\alpha)\beta + O\left(\frac{1}{f(t)^2}\right) \tag{3.27}
\]

where \( g(t) := \int_0^t f(s)ds \).

**Proof.** We have seen in defining \( Z(t) \) that \( Z^{(1)}(t) + V(t) \times Z(t) = 0 \). Taking projections onto \( H \), we find that \( Z^{(1)}(t) + (V \times Z)_\alpha = 0 \). Using Proposition 3.2.2, we see that \( Z^{(1)}(t) + f(t)(\alpha \times Z)_\alpha = O(1/f^2) \) implying that \( Z^{(1)}(t) + f(\alpha \times Z)_\alpha = O(1/f^2) \) giving:

\[
W(t) := Z^{(1)}(t) + f(t)I_\alpha Z_\alpha(t) = O\left(\frac{1}{f(t)^2}\right) \tag{3.28}
\]

So \( Z_\alpha(t) \) is close to satisfying a simple first order differential equation in the complex plane that can be solved using an integrating factor. Since \( g^{(1)}(t) = f(t) \)
for \( t > 0 \), we have, for \( 0 < r < s \),

\[
\exp(g(s)I_\alpha)Z_\alpha(s) - \exp(g(r)I_\alpha)Z_\alpha(r) = \int_r^s \exp(g(t)I_\alpha)W(t)dt = O\left(\frac{1}{f(r)^2}\right).
\]

(3.29)

Therefore by the completeness of \( E^3 \), we know that

\[
\lim_{s \to \infty} \frac{\exp(g(s)I_\alpha)Z_\alpha(s)}{\sqrt{b+c}} \text{ must exist, call the limit } \beta.
\]

Taking limits as \( s \to \infty \), we have

\[
\exp(g(t)I_\alpha)Z_\alpha(t) = (\sqrt{b+c})\beta + O(1/f^2).
\]

Finally, we can conclude using Lemma 3.2.3 that

\[
Z(t) = Z_\alpha(t) + O(1/f^3)\alpha = (\sqrt{b+c})\exp(-g(t)I_\alpha)\beta + O(1/f^2).
\]

(3.30)

Now that asymptotics have been obtained for \( Z(t) \), asymptotics for \( V(t) \) and its derivatives can be found:

**Theorem 3.2.5.** For \( t \geq 0 \),

\[
V(t) = f(t)\alpha - \frac{\sqrt{b+c}}{f(t)^2} \exp(-g(t)I_\alpha)\beta + O\left(\frac{1}{f(t)^3}\right).
\]

(3.31)

**Proof.** By Equation 3.25 and Theorem 3.2.4, for \( 0 < r < s \),

\[
\hat{U}(s) - \hat{U}(r) = -\int_r^s -3F(1)(\sqrt{b+c})\exp(-g(t)I_\alpha)\beta dt + O\left(\frac{1}{f(r)^5}\right)
\]

\[
= -\frac{3}{2}(\sqrt{b+c})I_\alpha \int_r^s \left(\frac{F(1)}{F^3}\right) f(t)I_\alpha \exp(-g(t)I_\alpha)\beta dt + O\left(\frac{1}{f(r)^5}\right).
\]

Integration by parts gives the right hand side as

\[
= -\left(\int_s^r \left(\frac{F(2)F - 3(F(1))^2}{F^4}\right) \exp(-g(t)I_\alpha)\beta dt\right) + O\left(\frac{1}{f(r)^5}\right).
\]

All the terms are of at most order \( O(1/f^4) \) and so therefore, \( \hat{U}(s) - \hat{U}(r) = O(1/f(r)^4) \). Taking limits as \( s \to \infty \), we find that \( \hat{U}(t) = \alpha + O(1/f^4) \). This
means that \( V(t) = f(t)\alpha - Z/f(t)^2 + O(1/f^3) \). Substituting in for \( Z(t) \) using Theorem 3.2.4 proves the result.

**Corollary 3.2.6.**

\[
V^{(2)}(t) = Z + V = f(t)\alpha + (\sqrt{b + c}) \exp(-g(t)I_\alpha)\beta + O\left(\frac{1}{f^2}\right) \tag{3.32}
\]

**Corollary 3.2.7.** For \( t \geq 0 \),

\[
V^{(1)}(t) = f^{(1)}\alpha - \frac{(\sqrt{b + c})f^{(1)}}{f^3} \exp(-g(t)I_\alpha)\beta + \frac{\sqrt{b + c}}{f} I_\alpha \exp(-g(t)I_\alpha)\beta + O\left(\frac{1}{f^3}\right) \tag{3.33}
\]

**Proof.** We have seen that \( Z \times V = \frac{1}{2} F^{(1)} V - F V^{(1)} \) and so

\[
V^{(1)} = \frac{1}{2} F^{(1)} V - Z \times V \tag{3.34}
\]

Using Theorems 3.2.4 and 3.2.5 to substitute for \( Z(t) \) and \( V(t) \) respectively proves the result.

**Example 3.2.8.** Figures 3.2 and 3.3 plot two components of \( V(t) \) computed using Mathematica’s *NDSolve* together with an asymptotic approximation to \( V(t) \).

![Figure 3.2: One component of \( V(t) \) (dashed) with an asymptotic approximation](image)

It is interesting to notice that in Corollary 3.2.7, the approximation of \( V^{(1)}(t) \) is not the formal derivative of the approximation for \( V(t) \).
3.3. Asymptotics of Null Riemannian Cubics in Tension in SO(3)

Having determined asymptotic expressions for the NLQT $V(t)$ in so(3), we wish to find similar expressions for the corresponding NRCT $x(t)$ in SO(3). Recall that $Ad_x : so(3) \to so(3)$ is the derivative at the identity of the map $g \mapsto xgx^{-1}$, $g, x \in SO(3)$. The dual $V^*(t)$ of a NLQT $V(t)$ is defined by $V^*(t) := -Ad_{x(t)}V(t)$. We can assume without loss that $x(0) = 1$ because if $x(t)$ is a NRCT, then so is $x(0)^{-1}x(t)$. The following facts proved in [35] are important in reconstructing the NRCT:

\[
\begin{align*}
\frac{d^2}{dt^2} V^*(t) &= V^*(t) \\
\frac{d}{dt} V^*(t) &= -Ad_{x(t)} V^{(1)}(t) \\
Ad_{x(t)} Z(t) &= Z(0)
\end{align*}
\]

For $X \in so(3)$, define $X_E := B^{-1}(X)$ to be the inverse image of the Lie algebra isomorphism $B$. Then for $X, Y \in so(3)$ and $A \in SO(3)$, we have that $X = Ad_A Y$ if and only if $X_E = AY_E$. Using this and the facts above, we obtain:

\[
\begin{align*}
x(t)V_E(t) &= V_E(0) \cosh t + V_E^{(1)}(0) \sinh t \\
x(t)Z_E(t) &= Z_E(0) \\
x(t)(Z_E(t) \times V_E(t)) &= Z_E(0) \times \left( V_E(0) \cosh t + V_E^{(1)}(0) \sinh t \right)
\end{align*}
\]

Define $p_1(t), p_2(t)$ and $p_3(t)$ as the right hand sides of (3.38), (3.39), and (3.40)
respectively and define \( P(t) := \begin{bmatrix} p_1(t) & p_2(t) & p_3(t) \end{bmatrix} \). Because we know how \( x(t) \) acts on three orthogonal vectors in \( E^3 \), we can determine \( x(t) \).

\[
x(t) = P(t) \begin{bmatrix} V_E(t) & Z_E(t) & Z_E(t) \times V_E(t) \end{bmatrix}^{-1}
\] (3.41)

After inverting the matrix, we are left with:

\[
x(t) = P(t) \begin{bmatrix} \frac{1}{f(t)} V_E(t) & \frac{1}{b+c} Z_E(t) & \frac{1}{f(t)(b+c)} Z_E(t) \times V_E(t) \end{bmatrix}^T
\] (3.42)

**Theorem 3.3.1.** For \( t \geq 0 \),

\[
x(t) = \Pi_{SO(3)} \begin{bmatrix} P(t) \left[ \frac{1}{f(t)} \alpha^T \right] \\
\frac{1}{\sqrt{b+c}} (\exp(-I_a g(t)) \beta)^T \\
\frac{1}{f(t)(b+c)} (\exp(-I_a g(t))(\beta \times \alpha))^T \end{bmatrix} + O \left( \frac{1}{f(t)^2} \right)
\] (3.43)

**Proof.** Here \( \Pi_{SO(3)} \) is the distance minimising projection of a matrix in \( M_{3\times3}(\mathbb{R}) \) into \( SO(3) \) (a compact space) thought of as a subset of Euclidean space. Substituting in for \( V(t) \) and \( Z(t) \) in Equation (3.42) using Theorems 3.2.4 and 3.2.5 prove the result. Note that firstly we can replace \( (\exp(-I_a g(t)) \beta) \times \alpha \) by \( \exp(-I_a g(t))(\beta \times \alpha) \) since \( \exp(-I_a g(t)) \) is an orthogonal transformation fixing \( \alpha \). Secondly, taking projections does not change the asymptotics because the approximation becomes sufficiently close to \( SO(3) \) to make it appear flat, and a projection will therefore only decrease the error.

\( \square \)

**Example 3.3.2.** As \( f(t) \) is of exponential order, the approximations to \( x(t) \) converge incredibly fast. We will look at an example where the following initial conditions are used: \( x(0) = I \), \( V_E(0) = (2.97917, -0.35216, -0.02234) \) and \( V_E^{(1)}(0) = (1.61826, 0.24575, 1.16985) \) where the initial conditions were tweaked to produce an \( \alpha = (1, 0, 0) \). Although we have studied a tension parameter \( \lambda = 1 \), this example uses a parameter \( \lambda = 0.01 \) for aesthetic reasons.
As we can see in Figure 3.4, even for $t$ small, the approximation for $x(t)$ works quite well. We may wish to compute the values of the objective function $J_2$ (extended to other time intervals) to compare how well the asymptotic expansions do in that regard. Let $\hat{x}(t)$ be the asymptotic approximation for $x(t)$.

We can from Table 3.1 see that the $J_2$ value of the approximation $\hat{x}$ also rapidly approaches the $J_2$ value of $x$ in this example.

Let us now consider $J_2$ values more generally and for the more general interval $[0, t_0]$. Define $\hat{V}(t)$ as the asymptotic approximation of $V(t)$ using Equation (3.31).
Chapter 3. Asymptotics of Null Riemannian Cubics in Tension in SO(3)

Table 3.1: Comparison of $J_2$ values for NRCT $x(t)$ and asymptotic approximation $\hat{x}(t)$.

<table>
<thead>
<tr>
<th>Interval</th>
<th>$J_2(x(t))$</th>
<th>$J_2(\hat{x}(t))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 0$ to $t = 1$</td>
<td>4.47738</td>
<td>4.5532</td>
</tr>
<tr>
<td>$t = 1$ to $t = 2$</td>
<td>4.85373</td>
<td>4.86114</td>
</tr>
<tr>
<td>$t = 2$ to $t = 3$</td>
<td>5.42488</td>
<td>5.42613</td>
</tr>
<tr>
<td>$t = 3$ to $t = 4$</td>
<td>6.21374</td>
<td>6.21405</td>
</tr>
<tr>
<td>$t = 4$ to $t = 10$</td>
<td>71.4492</td>
<td>71.4494</td>
</tr>
</tbody>
</table>

Note that because our metric is left-invariant,

$$J_2^0(x) = \int_0^{t_0} \|V^{(1)}(t)\|^2 + \|V(t)\|^2 dt. \quad (3.44)$$

Rather than using our asymptotic approximation of $V^{(1)}(t)$, let us take the formal derivative of $\hat{V}(t)$ (denoted $\hat{V}^{(1)}(t)$) and examine the error.

$$\hat{V}^{(1)}(t) = f^{(1)}(t) - \frac{(\sqrt{b+c})f^{(1)}}{2f^3} \exp(-g(t)I_a)\beta + \frac{\sqrt{b+c}}{f} I_a \exp(-g(t)I_a)\beta \quad (3.45)$$

The difference between $\hat{V}^{(1)}(t)$ and the asymptotic expression for $V^{(1)}(t)$ in Equation (3.33) is of error $O\left(\frac{1}{f(t)^2}\right)$ and so we can say that $\hat{V}^{(1)}(t) = V^{(1)}(t) + O\left(\frac{1}{f(t)^2}\right)$ which proves

**Theorem 3.3.3.**

$$J_2^{\text{lo}}(x) = \omega(J_2)^{\text{lo}}(\hat{V}) + O\left(\frac{1}{f(t_0)}\right) \quad (3.46)$$

where $\omega(J_2)^{\text{lo}}$ is the functional obtained from the Lie reduction of the integrand of $J_2^{\text{lo}}$.

**Proof.** Combine Equation (3.44) together with the fact $\hat{V}(t)$ and $\hat{V}^{(1)}(t)$ are of order $O(f(t))$.

One unresolved issue is how well $\omega(J_2)^{\text{lo}}$ approximates the functional $J_2$ evaluated on the approximation of $x(t)$ from Equation (3.43). A second unresolved issue is that approximations are parametrized by constants $\alpha$ and $\beta$ for which
there is no known way of computing from the initial conditions or vice versa. If it could be shown the asymptotic formula produce a \( J_2 \) which closely approximates a real NRCT for small \( t \) too, then we may consider using them for interpolants rather than \( x(t) \).

When \( x(t_0) \) is known for some large \( t_0 \) as well as the initial conditions, we can estimate \( \alpha \) and \( \beta \) by solving (with a numerical optimization package) the non-linear over-constrained system of equations obtained by rearranging Equation (3.43):

\[
\exp(-I_\alpha g(t_0)) \begin{bmatrix} \alpha^T \\ \beta^T \\ (\beta \times \alpha)^T \end{bmatrix} \approx P(t_0)^{-1} \cdot D(t_0) \cdot x(t_0) \tag{3.47}
\]

where \( D(t) := \text{diag}(f(t), \sqrt{b+c}, f(t)\sqrt{b+c}) \).

3.4 Conclusion

In summary, we have provided several reasons to study Riemannian cubics in tension by drawing on a problem in constrained optimisation. We specialise to the null case due to tractability of analysis and a comparison of this case to well studied curves in Euclidean space. We focus on the case \( Q = \text{SO}(3) \) due to the abundance of practical applications and derive asymptotic formulae for the solutions to null Riemannian cubics in tension and the corresponding null Lie quadratics in tension. We mention how the functional values of the approximations can be compared as well as suggesting how asymptotic formulas may be used if the initial conditions and final endpoint are known.
Chapter 3. Asymptotics of Null Riemannian Cubics in Tension in SO(3)
Prior fields

4.1 Introduction

Let $Q$ be a complete Riemannian manifold that represents the possible configuration states of some dynamical system. We may take several observations of the system $\xi_0, \ldots, \xi_N$ at times $t_1, \ldots, t_N$ and we are interested in interpolating the behaviour of the system in between. The dynamics of the system are often approximated by a model such as a first order differential equation (representable by a prior vector field $A$)

$$x^{(1)}(t) \approx A(x(t))$$

or possibly higher order differentiable equation such as a particle in a potential field

$$\nabla_t x^{(1)} \approx A(x(t))$$

More often than not, models of dynamical systems make simplifying assumptions and so the observed data may not be representable exactly by a solution of the model. We overcome this obstacle by searching for interpolants that best fit the given model while agreeing with what is observed.

4.2 First order systems

For first order systems, Noakes in [26] introduces one such interpolant which is a $C^1$ curve $x(t)$ such that $x(t_i) = \xi_i$ for $i = 1, \ldots, N$ for which the functional

$$\mathcal{A}J_1(x) := \int_0^1 \|x^{(1)}(t) - A_{x(t)}\|^2 dt$$

is minimised. A critical point of $\mathcal{A}J_1$ is called a conditional extremal with prior field $A$. If we restrict our problem to the case where $N = 1$, we can analyse the constituent curves that make up such an interpolant. In the case that $A = 0$, critical points of $\mathcal{A}J_1$ would be the geodesics of the manifold and this framework would produce geodesic interpolation.
Example 4.2.1. Consider the vector field $A(q) = q$ in $\mathbb{R}^n$ and we observe $x(0) = \xi_0$ and $x(1) = \xi_1$. Then $A_1 J_1(x) = \int_0^1 \|x^{(1)}(t) - x(t)\|^2 dt$. The Euler-Lagrange equations for this system are $(x^{(2)}(t) - x^{(1)}(t)) + (x^{(1)}(t) - x(t)) = 0$. The conditional extremals have the form $x(t) = c_1 e^t + c_2 e^{-t}$ where $c_1 = \frac{\xi_1 - \xi_0}{e^2 - 1}$ and $c_2 = \xi_0 + \frac{\xi_0 - \xi_1}{e^2 - 1}$.

Suppose $x(t)$ is a conditional extremal and that we have a smooth one parameter family of variations $x_h(t)$ where $x_0(t) = x(t)$ and the endpoints are fixed. Set $W_h(t) := \frac{\partial x_h}{\partial h}$. We can compute the derivative of $A_1 J_1$ with respect to $h$:

\[
\frac{1}{2} \frac{\partial A_1 J_1(x_h)}{\partial h} = \int_0^1 \langle \nabla_h (x^{(1)} - A), x^{(1)} - A \rangle dt = \int_0^1 \langle \nabla_t W_h - \nabla_h A, x^{(1)} - A \rangle dt = \int_0^1 -\langle W_h, \nabla_t (x^{(1)} - A) \rangle - \langle \nabla_h A, x^{(1)} - A \rangle dt
\]

where the last step uses integration by parts and the fact $W_h(0) = W_h(1) = 0$. An important definition we will need to simplify this further is of the function $\Phi$ defined by

\[
\langle \Phi X, Y \rangle := \langle \nabla X Y, Z \rangle.
\]

This is well defined since the metric is non-degenerate and because $\nabla X Y$ is tensorial in $X$. Here we interpret $\nabla X Y$ as a linear operator with $\Phi X$ as its adjoint. This construction allows us to isolate $W(t)$ to one side of the inner product leaving:

\[
\frac{1}{2} \frac{\partial A_1 J_1(x_h)}{\partial h} = -\int_0^1 \langle W_h, \nabla_t (x^{(1)} - A) \rangle + \Phi_{x^{(1)} - A} A \rangle dt.
\]

Evaluating at $h = 0$, the fundamental lemma of calculus of variations gives us the Euler Lagrange equations for conditional extremals

\[
\nabla_t (x^{(1)} - A) + \Phi_{x^{(1)} - A} A = 0.
\]

For practical considerations, we are often interested in finding when the conditional extremals minimise $A_1 J_1$. An analysis requires higher order derivatives of $A_1 J_1$. For this, consider a two parameter family of variations of $x(t)$ given by $x_{h_1, h_2}(t)$ where $x_{h_1, h_2}(t) = x(t)$ when $h_1 = h_2 = 0$, and the endpoints are fixed.
4.2. First order systems

Defining $W_1$ and $W_2$ as the partial derivatives of $x_{h_1,h_2}$ with respect to $h_1$ and $h_2$ respectively, the Hessian of $\mathcal{A}J_1$ at the critical point is given by

$$
\frac{\partial \mathcal{A}J_1(x_{h_1,h_2})}{\partial h_2 \partial h_1} = -\int_0^1 \langle W_1, \nabla h_2 \left( \nabla t(x^{(1)} - \mathcal{A}) + \Phi_{x^{(1)}-\mathcal{A}} \mathcal{A} \right) \rangle dt \\
+ \left( \langle \nabla h_2 W_1, x^{(1)} - \mathcal{A} \rangle + \langle W_1, \nabla h_2 (x^{(1)} - \mathcal{A}) \rangle \right) \bigg|_{t=1}^{t=0}
$$

(4.4)

This is calculated in a similar way to the Euler-Lagrange equations but we now keep the boundary terms, differentiate a second time and substitute in the Euler-Lagrange equations. Using the right hand entry of the inner product inside the integral in (4.4), we define an $\mathcal{A}$-Jacobi field as a vector field $W(t)$ along a conditional extremal $x$ satisfying

$$
\nabla W \left( \nabla t(x^{(1)} - \mathcal{A}) + \Phi_{x^{(1)}-\mathcal{A}} \mathcal{A} \right) = 0.
$$

(4.5)

This says $W$ to first order accuracy preserves the Euler-Lagrange equations but only makes sense when we have a variation $x_h(t)$ already defined. We usually have no variation to begin with but instead would like to be able to find to first order accuracy a variation of $x(t)$ for which Euler-Lagrange equations hold. This means we need a differential equation for $W(t)$ which does not require differentiation of $x^{(1)}$ in the $W(t)$ direction because we no longer have a variation $x_h(t)$ in mind.

The following theorem can be taken as the definition of $\mathcal{A}$-Jacobi fields:

**Theorem 4.2.2.** The $\mathcal{A}$-Jacobi field equation can be re-expressed as

$$
\nabla^2_W W - \nabla_t \nabla W \mathcal{A} + R(W, x^{(1)})(x^{(1)} - \mathcal{A}) + R(\mathcal{A}, x^{(1)}) W + \Phi_{x^{(1)}-\mathcal{A}} \nabla W \mathcal{A} + \Phi_{\nabla W - \nabla W \mathcal{A}} \mathcal{A} - \Phi_{\Phi_{x^{(1)}-\mathcal{A}} \mathcal{A}} W = 0.
$$

(4.6)

When $\mathcal{A} = 0$, this results in the usual Jacobi field equations for geodesics:

$$
\nabla^2_W W + R(W, x^{(1)}) x^{(1)} = 0.
$$

**Proof.** To shorten the proof, we first make the substitution $Y = x^{(1)} - \mathcal{A}$. Then
we have

\[ \nabla_W (\nabla_t Y + \Phi_Y \mathcal{A}) = \nabla_t \nabla_W Y + R(W, x^{(1)}) Y + \nabla_W \Phi_Y \mathcal{A}. \]

The first term in this expression can be rewritten as \( \nabla^2_t W - \nabla_t \nabla_W \mathcal{A} \). To simplify the last term, we take its inner product with an arbitrary vector field \( \mathcal{X} \) and also extend \( W \) to a vector field \( \mathcal{W} \) defined over the whole manifold. Because the inner product will be tensorial in both \( \mathcal{W} \) and \( \mathcal{X} \), we may assume that \( [\mathcal{W}, \mathcal{X}] = 0 \).

\[
\langle \nabla_W \Phi_Y \mathcal{A}, \mathcal{X} \rangle = \mathcal{W} \langle \Phi_Y \mathcal{A}, \mathcal{X} \rangle - \langle \Phi_Y \mathcal{A}, \nabla_W \mathcal{X} \rangle \\
= \langle \nabla_W \nabla_X A, Y \rangle + \langle \nabla_X A, \nabla_W Y \rangle - \langle \Phi_Y \mathcal{A} \mathcal{W}, \mathcal{X} \rangle \\
= \langle \Phi_Y \nabla_W \mathcal{A} + R(\mathcal{A}, \mathcal{Y}) \mathcal{W} + \Phi_{\nabla_W Y} \mathcal{A} - \Phi_{\Phi_Y \mathcal{A}} \mathcal{W}, \mathcal{X} \rangle
\]

Since \( \mathcal{X} \) can be taken to have any value at a given point, we must have equality for the the left hand terms in the inner product. Finally, replacing \( \nabla_W Y \) by \( \nabla_t W - \nabla_W \mathcal{A} \) and using the fact that \( R(\mathcal{A}, \mathcal{A}) \mathcal{W} = 0 \) proves the result.

\( \Box \)

Two points \( x(t_0) \) and \( x(t_1) \) on a conditional extremal are said to be \textit{conjugate} if there is a non-trivial \( \mathcal{A} \)-Jacobi field \( W(t) \) along \( x(t) \) such that \( W(t_0) = W(t_1) = 0 \). Two general points \( q_0, q_1 \in Q \) are \( \mathcal{A} \)-conjugate if they are conjugate along some conditional extremal. Schrader and Noakes in [40] prove that \( \mathcal{A}J_1 \) satisfies the Palais-Smale condition on the space of absolutely continuous curves joining two submanifolds of \( Q \) to show the existence of extremals in any homotopy class and also prove that the conjugate points along any curve are isolated. Some stronger results can be obtained if we now restrict ourselves to the case that \( Q \) is a bi-invariant Lie group \( G \) and \( \mathcal{A} \) a left-invariant vector field\(^1\). Recall that a bi-invariant

---

\(^1\)The requirement \( \mathcal{A} \) be left-invariant is not necessary to do the Lie reduction but the equations are rather complicated even in somewhat simple cases and lie outside the scope of this thesis.
Lie group satisfies \( \langle [u, v], w \rangle = \langle u, [v, w] \rangle \) for all \( u, v, w \in T_g G \). Let \( v \in T_g G \), recall that the left Lie reduction \( \mathcal{L} \) of \( v \) is given by \( \mathcal{L}(v) := (dL_g^{-1})v \in T_e G \). Then for any vector field \( W(t) \) along a curve \( x(t) \), we can define a curve in the Lie algebra by \( \hat{W}(t) := \mathcal{L}(W(x(t))) \). In the case \( W(t) = x^{(1)}(t) \), we give the curve \( \hat{W}(t) = \mathcal{L}(x^{(1)}(t)) \) the name \( V(t) \).

**Example 4.2.3.** Suppose we have an object that we wish to rotate about its centre of mass at a given fixed angular velocity \( \Omega_0 \) for a specified amount of time. The state of a rigid body in absolute coordinates can be described by a curve \( x(t) \) in \( SO(3) \), a bi-invariant Lie group. The angular velocity is defined as \( \Omega(t) = x(t)^{-1}x^{(1)}(t) \). We are able to take some measurements of the body’s orientation in its motion and we find that the rotation of the object is not exactly as specified. We may like to interpolate the object’s orientation between the observations to display how it actually rotates over time. If it is the case that we expect on average the angular velocity to be \( \Omega_0 \), then it is sensible to choose an interpolant that minimises the mean square error between the body’s angular velocity and \( \Omega_0 \).

The rotating mechanism attempts to keep \( \Omega(t) \) a constant \( \Omega_0 \). This is to say that we try to keep \( x^{(1)} \) approximately \( \mathcal{A} = x(t)\Omega_0 \) which is a left-invariant vector field on \( SO(3) \). One such choice for an interpolating curve \( x(t) \) is one that minimises \( \mathcal{A}J_1 \). Trouble can arise when interpolating optimally as the observed orientations become further apart. As with interpolating using piecewise geodesics, if a conditional extremal interpolant contains an \( \mathcal{A} \)-conjugate point, it means we have not produced an optimal interpolant as the Hessian will not be positive definite [18].

We will employ the use of a Lie reduction to simplify the equations obtained for conditional extremals and their \( \mathcal{A} \)-Jacobi fields. In doing so, it is important to see how to calculate the Lie reductions for the various terms involved in these equations. Let \( \mathcal{E}_i \) be a basis for the left-invariant vector fields of \( G \) and let \( \mathcal{X}, \mathcal{Y} \) and \( \mathcal{Z} \) be vector fields on \( G \). Define smooth functions \( X, Y, Z : G \to \mathbb{R}^n \) such that
\[ X_g = X^i(g)(E_i)_g, \ Y_g = Y^j(g)(E_j)_g \] and \[ Z_g = Z^k(g)(E_k)_g. \] We can then identify \( X(g), Y(g), Z(g) \) as the Lie reduction represented in the chosen basis by seeing that \( \mathcal{L}(X_g) = X^i(g)(E_i)_e \) and so on. Then for covariant differentiation, omitting the auxiliary variable \( g \), we have:

\[
\mathcal{L}(\nabla X Y) = \mathcal{L}(\nabla X^i E_i^j)(Y^j E_j) = \mathcal{L}((E_i^j Y^j + X^i Y^j \nabla E_i^j)E_j) = dY \cdot X + \frac{1}{2} [X, Y].
\]

where \( dY \) and \( X \) are thought of as elements of \( \text{GL}(\mathfrak{g}) \) and \( \mathfrak{g} \) respectively represented in the chosen basis. In the case that \( \mathcal{X} \) is \( x^{(1)} \) along some curve \( x(t) \), we have

\[
\mathcal{L}(\nabla_t Y) = V^{(1)}(t) + \frac{1}{2} [V, Y]. \tag{4.7}
\]

Lie reduction of the curvature tensor is similarly calculated:

\[
\mathcal{L}(R(\mathcal{X}, \mathcal{Y}) \mathcal{Z}) = \mathcal{L}(R(X^i E_i, Y^j E_j)(Z^k E_k)) = \mathcal{L}(X^i Y^j Z^k R(E_i, E_j)E_k) = \frac{1}{4} [Z, [X, Y]]. \tag{4.8}
\]

Lastly, we also would like to be able to calculate the Lie reduction of the adjoint of the covariant differential, \( \Phi_X \mathcal{Y} \). To compute this, we must first take the inner product with an arbitrary vector field \( \mathcal{Z} \) and use the non-degeneracy of the metric.

\[
\langle \mathcal{L}(\Phi_X \mathcal{Y}), \mathcal{Z} \rangle = \mathcal{L}((\Phi_X \mathcal{Y}), \mathcal{Z}) = \mathcal{L}((\nabla Z \mathcal{Y}, \mathcal{X})) = \langle \nabla Z \mathcal{Y}, X \rangle = \langle dY \cdot Z + \frac{1}{2} [Z, Y], X \rangle = \langle dY^* \cdot X + \frac{1}{2} [Y, X], Z \rangle
\]

where \( dY^* \) represents the adjoint with respect to the inner product. Because \( \mathcal{Z} \) was arbitrary, this means that

\[
\mathcal{L}(\Phi_X \mathcal{Y}) = dY^* \cdot X - \frac{1}{2} [X, Y]. \tag{4.9}
\]
In the setting of a bi-invariant group with left-invariant vector field $\mathcal{A}$, Noakes in [26] uses the Lie reduction to reduce Equation (4.3) to a much simpler equation in the Lie algebra. Recall that since $\mathcal{A}$ is left-invariant, we can represent it by $A = A_e$.

**Theorem 4.2.4** (Noakes [26]). Let $x(t)$ be a conditional extremal. Then $V(t) = \mathcal{L}(x^{(1)}(t))$ satisfies

$$V^{(1)} = [V, A] \quad (4.10)$$

**Proof.**

$$\nabla_t(x^{(1)} - A) + \Phi_{x^{(1)} - A}A = 0 \iff \mathcal{L}(\nabla_t(x^{(1)} - A) + \Phi_{x^{(1)} - A}A) = 0$$

$$\iff V^{(1)} + \frac{1}{2}[V, V - A] + \frac{1}{2}[A, V - A] = 0 \iff V^{(1)} + [A, V] = 0$$

Using this new equation, Noakes[26] calculates the explicit solution of such $\mathcal{A}$-extremals given by the following theorem

**Theorem 4.2.5** (Noakes [26]). Let $x : [0, 1] \to G$ be a conditional extremal with $\mathcal{A}$ left-invariant (with $A = A_e$) and $x(0) = e$. Then $x(t)$ can be written explicitly in the form

$$x(t) = e^{(V(0) - A)t}e^{At}. \quad (4.11)$$

**Proof.** The proof we give is for matrix Lie groups although Noakes proves it more generally. Solving for (4.10) using an integration factor tells us that $V(t) = e^{-At}V(0)e^{At}$. Using the linking equation $x^{(1)}(t) = x(t)V(t)$, we see that

$$x^{(1)}(t)e^{-At} = x(t)e^{-At}V(0). \quad (4.12)$$

Using this and the product rule, we see that

$$\frac{d}{dt}(x(t)e^{-At}) = (x(t)e^{-At})(V(0) - A). \quad (4.13)$$
This is a differential equation for $x(t)e^{-At}$ which has the solution $x(t)e^{-At} = e^{(V(0)-A)t}$ and then multiplying both sides of the equation by $e^{At}$ proves the result.

In the study of conjugate points along $\mathcal{A}$-extremals, these extremals on Lie groups were the first place to consider looking for examples. The following theorem says that in bi-invariant Lie groups, solving conjugate points of prior fields for left-invariant $\mathcal{A}$ correspond precisely to solving conjugate points for geodesics.

**Theorem 4.2.6.** Let $x(t) : [0, 1] \to G$ be a conditional extremal with $\mathcal{A}$ left-invariant, $x(0) = e$. Then $x(1)$ is an $\mathcal{A}$-conjugate point if and only if $\exp_e V(0) - A$ is a regular conjugate point of the geodesic $\exp_e (V(0) - A)t$.

**Proof.** By Theorem 4.2.5, all $\mathcal{A}$-extremals take the form $e^{t(V(0)-A)}e^{tA}$ where $V(0)$ is variable. As $e^{tA}$ does not change over different $\mathcal{A}$-extremals, variations through $\mathcal{A}$-extremals correspond to variations through the left hand term $e^{t(V(0)-A)}$. These variations may fix the endpoints precisely when $e^{V(0)-A}$ is conjugate to $e$ along $e^{(V_0-A_0)t}$.

The proof of the above theorem means that for different choices of $A$ on a connected Lie group, we can have conjugate points placed anywhere. Figure 4.1 shows an example of conjugate points arbitrarily close together distance-wise in $S^3$.

Although we have reduced the problem of finding conjugate points to a simpler, more well studied problem, knowing about the $\mathcal{A}$-Jacobi fields along has applications in interpolation. Solutions for $\mathcal{A}$-Jacobi fields will let us answer questions such as how sensitive a change in the endpoint velocities will be to a perturbation of the initial conditions. We will now use the solutions for $x(t)$ to find a differential equation for the Lie reduction of an $\mathcal{A}$-Jacobi Field. In what will follow, we will assume $W(0) = \mathbf{0}$ so we are only dealing with $\mathcal{A}$-Jacobi fields that fix the initial point. If $V(0) = A$, then we have a geodesic so we will also assume that $V(0) \neq \mathbf{0}$.
Figure 4.1: Extremals on $S^3 \subset \mathbb{H}$ with $A = i\pi$ and $V(0) \approx 0$ (Stereographic projection into $\mathbb{R}^3$).

Proposition 4.2.7. The $A$-Jacobi fields $W(t)$ for which $x(0) = e$ is fixed are given by the series expression

$$W(t) = \left( \sum_{i=1}^{\infty} \frac{t^i}{i!} T^{i-1}(B) \right) x(t)$$

where $T = \text{ad}(V(0) - A)$ and $B = W^{(1)}(0)$.

Proof. All variations will have up to first order the form $x_h(t) = e^{(V(0) + hB - A)t}e^{At}$. Then we explicitly calculate the formula $W(t) = \frac{\partial x_h}{\partial h} \bigg|_{h=0}$ using the formula [49] for the derivative of the exponential map and it is easily verified that $B = W^{(1)}(0)$.

A more easily understood formula can be obtained if instead of directly solving variations of the solution, we study the effect of a variation on the underlying differential equation. In effect we could now set $A = 0$ however we will preserve it to see its effects in the process. In what will follow, we will effectively be computing the derivative of the exponential but in a much simpler way than the methods which exist in literature and has application to further chapters.

Theorem 4.2.8. Let $x(t)$ be a conditional extremal and suppose $\dot{W}$ is the Lie
reduction of an \( \mathcal{A} \)-Jacobi field. Then

\[
\dot{\mathcal{W}}^{(2)} + [V + A, \dot{\mathcal{W}}^{(1)}] + [V, [A, \dot{\mathcal{W}}]] = 0
\]  

(4.15)

**Proof.** Let \( x_h \) be a variation through conditional extremals and let \( V_h \) be the corresponding Lie reductions. Define \( X_h = \frac{dV}{dh} \). Differentiating the equation \( V_h^{(1)} + [A, V_h] = 0 \) we can differentiate this equation with respect to \( h \) and we have

\[
X_h^{(1)} + [A, X_h] = 0
\]  

(4.16)

The relationship between \( \dot{W} \) and \( V_h \) is that \( \dot{\mathcal{W}}^{(1)} + [\dot{\mathcal{W}}, V] = X \) (A proof of this fact is in Chapter 5, page 86). Substituting, we get

\[
\dot{\mathcal{W}}^{(2)} + [V + A, \dot{\mathcal{W}}^{(1)}] + [V^{(1)}, \dot{W}] + [A, [V, \dot{W}]] = 0
\]

Using the fact that \( V^{(1)} = [V, A] \) together with the Jacobi identity, we obtain

\[
\dot{\mathcal{W}}^{(2)} + [V + A, \dot{\mathcal{W}}^{(1)}] + [V, [A, \dot{\mathcal{W}}]] = 0
\]

Before we explicitly solve this equation, we make some observations. Define an operator which we will call the Lax operator defined by \( L_C = \frac{d}{dt} + \text{ad}_C \). So for example, \( L_C V := V^{(1)} + [C, V] \). Under this new definition, we see that the condition for \( V \) to be a conditional extremal can be written as \( L_A V = 0 \). The condition for a vector field to be an \( \mathcal{A} \)-Jacobi field was calculated by expanding \( L_V L_A \dot{W} = 0 \). The following proposition is an interesting property of conditional extremals.

**Proposition 4.2.9.** Let \( x(t) \) an \( \mathcal{A} \)-extremal. Then \( L_V L_A = L_A L_V \).

**Proof.** For any curve \( B \) in the Lie algebra, \( L_A L_V B = L_A (B^{(1)} + [V, B]) = B^{(2)} + [V^{(1)}, B] + [V, B^{(1)}] + [A, B^{(1)} + [V, B]] \). Using the fact \( V \) is an \( \mathcal{A} \)-extremal, this is equal to \( B^{(2)} + [[V, A], B] + [V + A, B^{(1)}] + [A, [V, B]] \). Applying the Jacobi identity,
we are left with $B(2) + [B(1), V + A] + [V, [A, B]]$ which we have already seen is $L_V L_A B$.

The following propositions list some identities relating to the geometry of $\tilde{W}$.

**Proposition 4.2.10.** There exists a constant $b_1 \in \mathbb{R}$ such that for all $t$,

$$\langle W^{(1)} + [V, \tilde{W}], A \rangle = b_1. \quad (4.17)$$

**Proof.** Taking inner products of (4.15) with $A$, we find:

$$\langle \tilde{W}^{(2)}, A \rangle + \langle [V, \tilde{W}^{(1)}], A \rangle + \langle [V^{(1)}, \tilde{W}], A \rangle = 0$$

which implies $\frac{d}{dt} \left( \langle \tilde{W}^{(1)}, A \rangle + \langle [V, \tilde{W}], A \rangle \right) = 0$. Integrating proves the result. \qed

**Proposition 4.2.11.** There exists a constant $b_2 \in \mathbb{R}$ such that for all $t$,

$$\langle \tilde{W}^{(1)} + [V, \tilde{W}], V \rangle = \langle \tilde{W}^{(1)}, V \rangle = b_2. \quad (4.18)$$

**Proof.** Taking inner products of (4.15) with $V$, we find:

$$\langle \tilde{W}^{(2)}, V \rangle + \langle [A, \tilde{W}^{(1)}], V \rangle = 0$$

which implies $\langle \tilde{W}^{(2)}, V \rangle + \langle \tilde{W}^{(1)}, V^{(1)} \rangle = 0$. \qed

**Proposition 4.2.12.** There exists a constant $b_3 \in \mathbb{R}$ such that for all $t$,

$$\langle \tilde{W}^{(1)} + [V, \tilde{W}], V^{(1)} \rangle = b_3 \quad (4.19)$$

**Proof.** Taking inner products of (4.15) with $V^{(1)}$, we find:

$$\langle \tilde{W}^{(2)}, V^{(1)} \rangle + \langle [V + A, \tilde{W}^{(1)}], V^{(1)} \rangle + \langle [A, [V, \tilde{W}]], V^{(1)} \rangle = 0$$

which implies

$$\langle \tilde{W}^{(2)}, V^{(1)} \rangle + \langle [V, \tilde{W}^{(1)}], V^{(1)} \rangle + \langle \tilde{W}^{(1)}, [V^{(1)}, A] \rangle + \langle [V, \tilde{W}], [V^{(1)}, A] \rangle = 0$$

and so

$$\langle \tilde{W}^{(2)} + [V, \tilde{W}^{(1)}], V^{(1)} \rangle + \langle \tilde{W}^{(1)} + [V, \tilde{W}], V^{(2)} \rangle = 0. \quad \square$$

**Proposition 4.2.13.** There exists a constant $b_4 \in \mathbb{R}$ such that for all $t$,

$$\langle \tilde{W}, V - A \rangle = (b_2 - b_1)t + b_4 \quad (4.20)$$
Proof.

\[ b_2 - b_1 = \langle \tilde{W}^{(1)}, V \rangle - \langle \tilde{W}^{(1)} + [V, \tilde{W}], A \rangle = \langle \tilde{W}^{(1)}, V - A \rangle + \langle \tilde{W}, V^{(1)} \rangle \]  

(4.21)

Integrating proves the result \( \square \)

**Proposition 4.2.14.** For conjugate points to exist, \( b_1 = b_2 \) and \( \langle \tilde{W}, V - A \rangle = 0 \)

**Proof.** In the case of Proposition 4.2.13, we choose our \( A \)-Jacobi fields to vanish initially so the constant \( b_4 = 0 \). This means for there to be conjugate points, we must have \( b_1 = b_2 \) otherwise \( \langle \tilde{W}, V - A \rangle \) would only be 0 at \( t = 0 \). \( \square \)

If \( X \) is a solution to \( L_A X = 0 \), then \( X \) has the form \( X(t) = \text{Ad}(e^{-tA})X(0) \).

We know \( V \) is also a solution to \( L_A V = 0 \) so \( V(t) = \text{Ad}(e^{-tA})V(0) \). Let us then solve equations of the form \( L_V \tilde{W} = X \) for \( \tilde{W} \). Since we choose \( \tilde{W}(0) = 0 \), we know that \( X(0) = W^{(1)}(0) \).

\[ \tilde{W}^{(1)} + [V, \tilde{W}] = \text{Ad}(e^{-tA})W^{(1)}(0) \]  

(4.22)

Applying \( \text{Ad}(e^{tA}) \) to both sides, we find due to \( \text{Ad} \) invariance of \( \text{ad} \) that

\[ \text{Ad}(e^{tA})\tilde{W}^{(1)} + [V(0), \text{Ad}(e^{tA})\tilde{W}] = \tilde{W}^{(1)}(0) \]  

(4.23)

Defining \( \tilde{Y}(t) = \text{Ad}(e^{tA})\tilde{W}(t) \), then \( \tilde{Y}^{(1)}(t) = [A, \text{Ad}(e^{tA})\tilde{W}] + \text{Ad}(e^{tA})\tilde{W}^{(1)} \) and

\[ \tilde{Y}^{(1)} + [\tilde{V}(0) - A, \tilde{Y}] = \tilde{W}^{(1)}(0) \]

This is just a constant coefficient linear system. Because \( G \) is bi-invariant, \( T = \text{ad}_{V(0)} - A \) is skew-adjoint and so \( \text{im} T \oplus \ker T = \mathfrak{g} \), and that \( T : \text{im} T \to \text{im} T \) is an isomorphism. By setting \( \tilde{Y} = \nu_1 + \nu_2 t \) with \( \nu_1 \in \text{im} T \) and \( \nu_2 \in \ker T \), there is a unique solution for \( \nu_1 \) and \( \nu_2 \) such that \( T(\nu_1) + \nu_2 = \tilde{W}^{(1)}(0) \). This leads to the general solution \( \tilde{Y} = \nu_1 + \nu_2 t - \text{Ad}(e^{-(V(0)-A)t})\nu_1 \) since \( \tilde{Y}(0) = 0 \) and implies the following theorem
Theorem 4.2.15. If \( x(t) \) is an \( A \)-extremal on a Lie group and \( \tilde{W}(t) \) is the Lie reduction of an \( A \)-Jacobi field \( W(t) \) which fixes the initial point \( x(0) = e \), then \( \tilde{W}(t) \) has the form:

\[
\tilde{W}(t) = \text{Ad}(e^{-tA})(\nu_2 t + (1 - \text{Ad}(e^{-(V(0) - A)t}))\nu_1)
\]

where \( W^{(1)}(0) = \nu_1 + \nu_2 \) and \( \nu_1 \in \text{im} \, T \) and \( \nu_2 \in \text{ker} \, T \). Moreover, \( W(t) = x(t)\tilde{W}(t) \) has the form:

\[
W(t) = e^{(V(0) - A)t}(\nu_2 t + (1 - \text{Ad}(e^{-(V(0) - A)t}))\nu_1)e^{At}
\]

Example 4.2.16. Let \( G \) be the algebra of unit quaternions in \( \mathbb{H} \) thought of as \( S^3 \). Let \( A_0 = (0, 1, 0, 0) \) and \( V(0) = (0, 4, 0, 4) \). The map \( T = ad_{V(0) - A} \) on the Lie algebra has \( \ker T = \text{span}(1, -1, 0) \) and \( \text{im} \, T = \ker T^\perp \). Then we see that \( \tilde{W}(t) = \text{Ad}(e^{-tA})(\nu_1 + \nu_2 t - \text{Ad}(e^{-(V(0) - A)t})\nu_1) \) which has a zero when \( \nu_1 + \nu_2 t - \text{Ad}(e^{-(V(0) - A)t})\nu_1 = 0 \). This example relates very closely to the motivating example of a rotating body, but we choose \( S^3 \) for a simpler description. \( S^3 \) is in fact just a double cover of \( SO(3) \) with isomorphic Lie algebras. To see that their Lie algebras are isomorphic, recall firstly that we already have a Lie algebra isomorphism between \( \text{so}(3) \) and \( E^3 \). Take the linear map between \( T_1 S^3 \) and \( E^3 \) given by the maps \( i \mapsto 2e_1, j \mapsto 2e_2, k \mapsto 2e_3 \). Then, as an example, \([\frac{1}{2}i, \frac{1}{2}j] = (\frac{1}{2}i) (\frac{1}{2}j) - (\frac{1}{2}j) (\frac{1}{2}i) = \frac{1}{2}k \) while we also have \( e_1 \times e_2 = e_3 \). This calculation can be repeated with a more general vector to prove that this map is an isomorphism.

In this example,

\[
x(t) = (\cos t \cos 5t \frac{3}{5} \sin t \sin 5t, \cos 5t \sin t + \frac{3}{5} \cos t \sin 5t, \frac{4}{5} \sin t \sin 5t, \frac{4}{5} \cos t \sin 5t)
\]

A possible Lie reduced \( A \)-Jacobi field would be

\[
\tilde{W}(t) = (-4 + 4 \cos 10t, 3 \sin 2t - \sin 8t - 4 \sin 12t, 3 \cos 2t + \cos 8t - 4 \cos 12t)
\]
If \( \nu_2 \neq 0 \), then this equation cannot have any non-trivial zeros as both \( \nu_1 \) and \( \text{Ad}(e^{-(V(0)-A)t})\nu_1 \) will be orthogonal to \( \nu_2 \). If \( \nu_2 = 0 \) on the other hand, the only non-trivial zeros occur when \( \text{Ad}(e^{-(V(0)-A)t}) = e \) which is when \( t = \frac{k\pi}{\|V(0)-A\|} \) for \( k \in \mathbb{N} \).

To see how such variations look, the first root of \( W(t) \) is \( t_0 = \frac{\pi}{5} \). Here \( \hat{V}^{(1)}(0) \) can be chosen anywhere in the space perpendicular to \( (3, 0, 4) \). We can see the conjugate points by taking small perturbations to \( V(0) \) of the form \( V(0) + \epsilon \hat{V}^{(1)}(0) \) for small \( \epsilon \).

**Figure 4.2:** Perturbations of the initial condition \( V(0) \) \( (V(0) \) represented by the red curve) in Example 4.2.16 for \( A \)-extremals with \( t \in [0, \frac{\pi}{5}] \). Stereographic projection of \( S^3 \) onto \( \mathbb{R}^3 \) is used to represent the curves.

**Example 4.2.17.** Let us now consider the Lie group \( \text{SL}(2, \mathbb{R}) \), the group of 2 by 2 real matrices with determinant 1. The Killing form [11] here is not positive definite so this example no longer relates to the minimisation problem posed at the beginning of the chapter however critical points can be calculated just the same and the same formula for \( \hat{W}(t) \) holds. Let \( A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \) and \( V_0 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \).
Then \( x(t) = \begin{pmatrix} e^t \cos t & -e^{-t} \sin t \\ e^t \sin t & e^{-t} \cos t \end{pmatrix} \). Suppose \( W^{(1)}(0) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \). Then \( \nu_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \) and \( \nu_2 = 0 \). We can then solve \( \dot{W}(t) = \begin{pmatrix} 2e^{-t} \sin^2 t & e^{-t} \sin 2t \\ e^t \sin 2t & -2e^t \sin^2 t \end{pmatrix} \).

By noticing the conjugate points occur when \( t \) is a multiple of \( 2k\pi \), we can find a point conjugate to the identity is \( \begin{pmatrix} e^{2\pi} \\ 0 \\ 0 \end{pmatrix} \).

### 4.2.1 Variable left-invariant priors

It is easy to imagine interpolation applications where the vector field may be modelled to be left-invariant but it is not known which vector field to use. Consider now the system where we have an unknown left-invariant prior field \( \mathcal{A} \) and suppose now we let \( t \) go from \([0, T]\). The resulting Euler Lagrange equations for the system will be the same of that of the fixed left-invariant prior field however an extra condition will come about when differentiating the functional \( \mathcal{A} J_1 \) with respect to \( \mathcal{A} \). Suppose \( \mathcal{A}_s = \mathcal{A} + s \mathcal{B} \) with \( \mathcal{B} \) also left-invariant (represented as \( A \) and \( B \) in \( g \)).

As we are dealing with a bi-invariant metric, we can reduce the functional to:

\[
\mathcal{A} J_1(x_h, s) = \int_0^T \langle V_r^{(1)} - \mathcal{A}_s, V_r^{(1)} - \mathcal{A}_s \rangle dt
\]  

(4.26)

Differentiating this expression with respect to \( s \), we get:

\[
\frac{1}{2} \frac{\partial}{\partial s} \mathcal{A} J_1(x_r, s) \big|_{s=0} = \int_0^T \langle \frac{d}{ds} \big|_{s=0} \mathcal{A}_s, V_r^{(1)} - \mathcal{A}_s \rangle dt

= \langle \mathcal{B}, V(T) - V(0) - T A \rangle
\]

If this is to equal \( 0 \) for all variations \( \mathcal{B} \) of \( \mathcal{A} \), then it is clear that \( \mathcal{A} = \frac{1}{T} (V(T) - V(0)) \). This shows that if \( \mathcal{A} \) is to be estimated from the data as well, the best choice to use is the average velocity between the points. The analysis then remains the same as when \( \mathcal{A} \) is known.
4.3 Second order systems

Usually in mechanics where forces and accelerations are involved, the dynamics are governed by a second order system. Let \( L : (TQ \oplus Q) \times \mathbb{R} \to \mathbb{R} \) be a smooth function and let \( x : [0, 1] \to Q \) be a \( C^2 \) curve with known endpoints \( x(0) = q_0 \) and \( x(1) = q_1 \). Let \( \nu \) be a lifting of the curve \( x \) given by \( \nu(x) = (x, \nabla_t x(1), x(1)) \) and define:

\[
J(x) = \int_{0}^{1} L(\nu(x)(t), t) dt
\]

Let \( \mathcal{A} = (A_1, A_2) \) be a section of \( TQ \oplus QTM \), and \( L(X, t) = \langle X - A_X, X - A_X \rangle \) where the inner product is given by a general orthogonal relationship

\[
\langle (A_1, A_2), (B_1, B_2) \rangle = \langle A_1, B_1 \rangle + \lambda \langle A_2, B_2 \rangle.
\]

Define

\[
\mathcal{A}J_2^\lambda(x) := \int_{0}^{1} \| \nabla_t x^{(1)} - A_1 \|^2 + \lambda \| x^{(1)} - A_2 \|^2 dt
\]

(4.27)

This functional generalises that of Riemannian cubics in tension. We can of course compute the Euler-Lagrange equations for \( \mathcal{A}J_2^\lambda \) although we may as well suppose \( \lambda = 0 \) since we have studied that case in the first section of this chapter and therefore we will write \( A_1 \) as \( \mathcal{A} \). Let \( x_h \) be a variation of \( x \) for which all of \( x^{(1)}(0) = x^{(1)}(1) = \nabla_t x^{(1)}(0) = \nabla_t x^{(1)}(1) = 0 \)

\[
\frac{1}{2} \frac{d}{dh} \mathcal{A}J_2^\lambda(x_h) = \int_{0}^{1} \langle \nabla_h (\nabla_t x^{(1)} - \mathcal{A}), \nabla_t x^{(1)} - \mathcal{A} \rangle dt
\]

\[
= \int_{0}^{1} \langle \nabla_t \nabla_h x^{(1)} + R(\frac{dx_h}{dh}, x^{(1)}), x^{(1)} - \nabla_h \mathcal{A}, \nabla_t x^{(1)} - \mathcal{A} \rangle dt
\]

\[
= \int_{0}^{1} \langle \nabla_t^2 \frac{dx_h}{dh} + R(\frac{dx_h}{dh}, x^{(1)}), x^{(1)} - \nabla_h \mathcal{A}, \nabla_t x^{(1)} - \mathcal{A} \rangle dt
\]

\[
= \int_{0}^{1} \langle \frac{dx_h}{dh}, \nabla_t^3 x^{(1)} - \nabla_t^2 \mathcal{A} + R(\nabla_t x^{(1)} - \mathcal{A}, x^{(1)}), x^{(1)} - \Phi_{\nabla_t x^{(1)} - \mathcal{A}} \rangle dt
\]

Theorem 4.3.1. Let \( x(t) \) be a critical point of the functional \( \mathcal{A}J_2 \) as defined in Equation (4.27) with \( \lambda = 0 \). Setting \( Y(t) = \nabla_t x^{(1)} - \mathcal{A} \), we see that \( x(t) \) satisfies

\[
\nabla_t^2 Y + R(Y, x^{(1)}), x^{(1)} - \Phi_Y \mathcal{A} = 0
\]

(4.28)
Compare this result with the equation for first order conditional extremals which when we set $Y(t) = x^{(1)} - A$, get the equation $\nabla_t Y + \Phi_Y A = 0$. Secondly, notice that every term in the expression involves $A$. If the manifold and vector field displays symmetry in the form of $Q$ being a bi-invariant Lie group and $\mathcal{A}$ being left-invariant, then an interesting phenomenon occurs. In what will follow, let $Q$ be a Lie group $G$ and $\mathcal{A}$ a left-invariant vector field. As for the first order conditional extremals, we can carry out the Lie reduction on the Euler Lagrange equations to examine the equation for $V(t)$.

**Theorem 4.3.2.** The Lie reduction $V(t)$ of an extremal of $A J_0^2$ defined in Equation (4.27) on a bi-invariant Lie group with $\mathcal{A}$ left-invariant satisfies

$$V^{(3)} + [V, V^{(2)}] = 0$$

**(Proof.** Setting $\tilde{Y}(t) = \mathcal{L}(Y(t))$, we have:

$$\mathcal{L} \left( \nabla_t^2 Y + R(Y, x^{(1)})x^{(1)} - \Phi_Y A \right)
= \left( \tilde{Y}^{(2)} + \frac{1}{2}[V^{(1)}, \tilde{Y}] + \frac{1}{2}[V, \tilde{Y}^{(1)}] + \frac{1}{2}[V, \tilde{Y}^{(1)} + [V, \tilde{Y}]] \right)
+ \left( \frac{1}{4}[V, [\tilde{Y}, V]] \right) + \left( \frac{1}{2}[\tilde{Y}, A] \right)
= \tilde{Y}^{(2)} + \frac{1}{2}[V^{(1)} - A, \tilde{Y}] + [V, \tilde{Y}^{(1)}] = \tilde{Y}^{(2)} + [V, \tilde{Y}^{(1)}]
= V^{(3)} + [V, V^{(2)}].$$

This theorem is rather intriguing because $\mathcal{A}$ does not show up in the equation for $V(t)$ at all showing that the left invariance of the vector field produces no local effect on the solutions. Interestingly, the Lie reduced equation obtained for second order systems is identical to that of Lie quadratics and so the second order conditional extremals on a Lie group are Riemannian cubics. The Euler Lagrange equations don’t show the entire picture however and removing constraints on the end-point velocities adds boundary conditions that are affected by $\mathcal{A}$. The effects
of such boundary conditions are discussed in Chapter 5.

4.4 Time varying systems

We wish to extend the constructions of the previous sections to those where the prior vector field $A$ may change with time. Many models of dynamical systems are time dependent and so we might like to model the systems behaviour by an equation such as $x^{(1)}(t) \approx A(x(t), t)$. As before, we would like to find critical points of the functional:

$$A(t)J_1(x) := \int_0^1 \|x^{(1)}(t) - A(x(t), t)\|^2dt$$

(4.30)

over $C^1$ curves $x(t)$. As $A$ is no longer strictly a vector field over the integration, we must first extend the domain to describe the problem in a form we are more familiar with. We do this by transforming the problem to the problem on the product manifold $Q \times [0, 1]$ and restricting variations of the curve to those which do not change the second parameter. In this manifold, $A(x, t) = (A(x, t), 1)$ is now a smooth vector field (where $A(x, t)$ is the original pseudo-vector field).

Let $x : [0, 1] \times (-\epsilon, \epsilon) \to Q \times [0, 1]$ define a variation of curves $x(t, h)$ such that $x(t, 0)$ is a critical point and moreover we will restrict ourselves to those variations for which $\pi_2x(h, t) = t$. Set $W_{h,t} = \frac{\partial x_{h,t}}{\partial h}$ and $X_{h,t} = \frac{\partial x_{h,t}}{\partial t}$ so as to avoid confusion regarding the $\nabla_t$ notation.

$$\frac{1}{2} \frac{\partial A(t)J_1(x_{h,t})}{\partial h} = \int_0^1 \langle \nabla_W(X - A), X - A \rangle dt$$

$$= \int_0^1 \langle \nabla_X W - \nabla_W A, X - A \rangle dt$$

$$= \int_0^1 -\langle W, \nabla_X(X - A) \rangle - (\nabla_W A, x^{(1)} - A) dt$$

$$= \int_0^1 -\langle W, \nabla_X(X - A) + \Phi_{X-A}A \rangle dt$$

Up to this point, nothing is different to time invariant systems. The difference arises because our $W$ is restricted purely to variations that do not change the
4.4. Time varying systems

second argument and we can only conclude that the projection of $\nabla_X (X - A) + \Phi_{X - A} A$ into $TQ$ is $0$. In order to calculate this projection, we must first compute the connection on a product manifold $Q_1 \times Q_2$.

Let $(u_1(x), \ldots, u_{n_1}(x), v_1(y), \ldots, v_{n_2}(y))$ be local coordinates for a chart on $Q_1 \times Q_2$ using an atlas of cartesian products of atlases on $Q_1$ and $Q_2$. Indices $i$ and $j$ are similar if they have the same sign, denoted $i \simeq j$. By definition of the metric, $g_{ij} = 0$ whenever $i \not\simeq j$. The formula for the Levi-Civita connection is given by $\bar{\nabla}_{\bar{U}} \bar{V} = (\bar{\nabla}_{\bar{U}} \bar{V}, \bar{V})$ where $\bar{\nabla}_{\bar{U}} \bar{V}$ is defined by $\frac{1}{2} g^{kr}(\partial_i g_{jr} + \partial_j g_{ir} - \partial_r g_{ij})$. When $i, j, k$ are similar, $\Gamma^k_{ij}$ is known since $g_{kr}$ is only ever non-zero for $k \simeq r$. Because of symmetry, there are only two other cases to examine: When $i \simeq j \not\simeq k$ and $i \simeq k \not\simeq j$. Because the only terms that contribute are when $r \simeq k$, it can be seen that $\Gamma^k_{ij} = 0$ in both the above cases. We can then define the covariant derivative as follows:

$$\nabla_{(U_1, U_2)}(V_1, V_2) = (\nabla_{U_1} V_1 + dV_1(U_2), \nabla_{U_2} V_2 + dV_2(U_1))$$

where $\nabla_{U_1}$ and $\nabla_{U_2}$ are connections on $Q_1$ and $Q_2$. The partial derivative terms $dV_1(U_2)$ and $dV_2(U_1)$ are defined because the tangent space remains fixed in the orthogonal direction. In the case of $Q \times \mathbb{R}$, we have the following for covariant derivatives along $x$:

$$\nabla_X A = (\nabla_X A + d_2 A, 0)$$
$$\nabla_X X = (\nabla_X x^{(1)}, 0)$$

where $d_2$ refers to differentiation with respect to the second coordinate of $A$.

Under this computation, we obtain the Euler-Lagrange equations with the original connection on $Q$.

**Theorem 4.4.1.** A solution $x(t)$ to a critical point of $A(t)J$ as defined in Equation (4.30) satisfies

$$\nabla_t (x^{(1)} - A) + \Phi_{x^{(1)} - A} A - \frac{\partial A}{\partial t} = 0.$$

(4.34)
If $Q$ is a bi-invariant Lie group $G$ and $A(g, t) = dL_g A(e, t) =: A(t)$, then we can compute the Lie reduction of the above equation. As the covariant derivatives are partial in the $X$ direction, the first two terms are treated the same as the Lie reduction for ordinary prior fields on a Lie group. The last term on the other hand is differentiated in the $t$ direction.

**Theorem 4.4.2.** A critical point of $J$ as defined in Equation (4.30) when there is time-dependent left-invariant prior field $A(t)$ on a bi-invariant Lie group satisfies:

$$V^{(1)}(t) - A^{(1)}(t) = [V(t), A(t)]$$

(4.35)

Making the substitution $Y(t) = V(t) - A(t)$, we obtain a nicer form of Theorem 4.4.2:

$$Y^{(1)}(t) = [Y(t), A(t)]$$

(4.36)

We have seen what solutions to this equation are when $A(t)$ is constant. The simplest case after $A(t)$ being constant is if $A(t)$ is linear. This is a useful practical case to study to make because we can approximate more complicated vector fields using line segments. If $A(t) = A_0 + A_1 t$, then various scenarios can occur. If $[A_0, A_1] = 0$, the solutions are $Y(t) = \text{Ad}(e^{-(A_0 t + \frac{1}{2} A_1 t^2)}) Y(0)$. These can be solved for $x(t)$ in the same manner as in the time-invariant case.

**Proposition 4.4.3.** Suppose $x(t)$ is a conditional extremal for the time-variant left-invariant prior field defined by $A(t) = A_0 + A_1 t$ where $[A_0, A_1] = 0$ and $x(0) = e$. Set $B(t) = A_0 t + \frac{1}{2} A_1 t^2$. Then the solutions for $x(t)$ are

$$x(t) = e^{(V(0) - A(0))t} e^{B(t)}$$

(4.37)

**Proof.** We have seen that $Y(t) = \text{Ad}(e^{-B(t)}) Y(0)$ and so $V(t) = \text{Ad}(e^{-B(t)})(V(0) - A(0)) + A(t) = \text{Ad}(e^{-B(t)})(V(0) - A(0) + A(t))$. Using the formula $x^{(1)}(t) = x(t)V(t)$ and the product rule, we can see that

$$(x(t)e^{-B(t)})^{(1)} = x^{(1)}(t)e^{-B(t)} - x(t)e^{-B(t)}A(t) = (x(t)e^{-B(t)}) (V(0) - A(0))$$
Therefore $x(t)e^{-B(t)} = e^{(V(0) - A(0))t}$ and so
\[ x(t) = e^{(V(0) - A(0))t}e^{B(t)} \]

We should note that this proof would work even if $A(t)$ was more general so long as $A(t)$ is contained in a Lie subalgebra $\mathfrak{h}$ with $[\mathfrak{h}, \mathfrak{h}] = 0$ in which case we would set $B(t) = \int A(t)dt$. As seen by the above theorem and the discussion in the proof of Theorem 4.2.6, the conjugate points of $x(t)$ depend only on $V(0)$ and the constant term of $A(t)$ which is quite curious. If on the other hand $[A_0, A_1] \neq 0$, then the solution becomes significantly more difficult to find. If we restrict ourselves to the manifold $\text{SO}(3)$, then solutions in this case can be solved up to a single quadrature.

**Theorem 4.4.4.** Solutions to the equation for $\nabla X(X - A) - \frac{dA}{dt} + \Phi X - A A = 0$ for $A = A_0 + A_1t$ with $[A_0, A_1] \neq 0$ and $X = x^{(1)}$ can be solved exactly in terms of a single quadrature when the Lie group is $\text{SO}(3)$.

**Proof.** We have the equation $Y^{(1)}(t) = [A_0 + A_1t, Y(t)] = (\text{ad} A_0 + t \text{ad} A_1)Y(t)$ where $Y(t) = V(t) - A(t)$. Suppose $\mathfrak{g}$ is a Lie Algebra such that $\text{ad} \mathfrak{g}$ belongs in the Lie algebra of some matrix subgroup $G$ of $\text{GL}(\text{dim} \mathfrak{g}, \mathbb{R})$. Then we can see that when extending $Y(t)$ to a fundamental matrix $\Psi(t)$, we have an equation of the form $\Psi^{(1)}(t) = (B_0 + tB_1)\Psi(t)$ where $\Psi(t) \in G$ and $B_0, B_1 \in T_eG$. Using the duality [21] property of null Riemannian cubics, we see that the same $\Psi(t)$ satisfies $\Psi^{(1)}(t) = \Psi(t)C(t)$ where $C(t)$ is a null Lie quadratic and so $\Psi(t)$ is a null Riemannian cubic. This shows that solutions to $Y(t)$ correspond to projections of a null Riemannian cubic in $G$. For example, consider $\mathfrak{g} = E^3$ and $\text{ad} \mathfrak{g}$ only contains $3 \times 3$ skew-symmetric matrices which are in the Lie algebra of $\text{SO}(3, \mathbb{R})$.

Pauley in [30] finds the solutions of null Riemannian cubics in $\text{SL}(2, \mathbb{C})$ implicitly in terms of special functions and then which can be transformed via a homomorphism into solutions for $\text{SO}(3)$. Also within the paper are solutions for null Lie quadratics and their first three derivatives. Let $Y(t)$ be a projection of a null Riemannian cubic $Y(t) = x_c(t) \cdot C_0$, where $C_0$ is chosen so that $x_c(0)C_0 = Y(0)$. Then $V(t) = x_c(t)C_0 - A_1t - A_0$ is known exactly.
For what is to follow, recall that we have a Lie algebra isomorphism $B : \mathfrak{so}(3) \to E^3$ (denoted by $B : D \mapsto D_E$) defined by $D \cdot v = D_E \times v$ where we also have the relationship $Ad_g D = g D E$ due to the properties of orthogonal matrices and the cross product. The result on duality [21] shows it is in fact possible to reconstruct the $A$-extremal if we know the quantity $Y(t)$. This is because we know that $\frac{d}{dt} Ad_{x(t)} Y(t) = Ad_{x(t)} (Y^{(1)}(t) + [V, Y(t)]) = Ad_{x(t)} (V^{(1)} - A^{(1)} + [V, -A]) = 0$ and so $Ad_{x(t)} Y(t) = D$. Considering the vectors $Y(t)$ and $D$ as elements of $E^3$, the equation can be rewritten $x(t)Y_E(t) = C_E$. We are then interested in the action of $x(t)$ in a vector orthogonal to $Y_E(t)$. $Y_E^{(1)}(t)$ is one such vector and it can be solved knowing $Y(t)$. Set $W_3(t) := Y_E(t)/\|Y_E(t)\|$, $W_1(t) := Y_E^{(1)}(t)/\|Y_E^{(1)}(t)\|$ and $W_2(t) := W_3(t) \times W_1(t)$ and define:

$$
\theta(t) := \int_0^t \langle W_1(s), W_2^{(1)}(s) + V_E(s) \times W_2(s) \rangle ds \tag{4.38}
$$

$$
U(t) := \begin{bmatrix}
W_1 \cos \theta(t) + W_2 \sin \theta(t) & W_2 \cos \theta(t) - W_1 \sin \theta(t) & W_3(t)
\end{bmatrix} \tag{4.39}
$$

Then $x(t) = U(0)U(t)^T$ is an $A(t)$-extremal. To prove this, we will show that $x^{(1)}(t) = x(t)V(t)$. We can see first of all that $W_3^{(1)} = W_3 \times V_E$ because $Y_E^{(1)} = Y_E \times V_E$ and as a consequence $\|Y_E\|$ is constant. So $U_3^{(1)} = -V_E \times U_3 = -VU_3$. Now, using the fact that $\theta^{(1)}(t) = \langle W_1(t), W_2^{(1)}(t) + V_E(t) \times W_2(t) \rangle$ and that $W_1, W_2$ and $W_3$ are orthonormally related, we obtain:

$$
\langle U_1^{(1)}, W_1 \rangle = -\langle W_1, V_E \times W_2 \rangle \sin \theta = -\langle VU_1, W_1 \rangle
$$

$$
\langle U_1^{(1)}, W_2 \rangle = \langle W_1, V_E \times W_2 \rangle \cos \theta = -\langle V_E \times W_1, W_2 \rangle \cos \theta = -\langle VU_1, W_2 \rangle
$$

$$
\langle U_1^{(1)}, W_3 \rangle = \langle U_1^{(1)}, U_3 \rangle = -\langle U_1, U_3^{(1)} \rangle = -\langle U_1, W_3 \times V_E \rangle = -\langle VU_1, W_3 \rangle
$$

Therefore $U_1^{(1)} = -VU_1$ as they act identically on a basis for $E^3$. Lastly, $U_2^{(1)} = -(V_E \times U_3) \times U_1 - U_3 \times (V_E \times U_1) = -V_E \times (U_3 \times U_1) = -VU_2$. It is then seen that $U^{(1)} = -VU$. Then $x^{(1)}(t) = (U(0)U(t)^T)^{(1)} = -U(0)U(t)^TV(t)^T = x(t)V(t)$.
since $V(t)$ is skew-symmetric.

4.4.1 Time varying prior fields for solvable $A(t)$ in a bi-invariant Lie group In some instances, $A(t)$ may be a vector field for which integral curves are already known. We now prove a remarkable fact that generalises many of the previous results. If we know the integral curves of $A(t)$, we can solve the equation $V^{(1)}(t) - A^{(1)}(t) = [V(t), A(t)]$ exactly and reconstruct the conditional extremal up to a single quadrature.

**Theorem 4.4.5.** Let $G$ be a bi-invariant Lie group and suppose that $z^{(1)}(t) = z(t)A(t)$ where $A(t)$ is a curve in the Lie algebra. Then we can solve $V(t)$ and $x(t)$ exactly in terms of $z(t)$, where $x^{(1)}(t) = x(t)V(t)$ and $V(t)$ satisfies the equation $V^{(1)}(t) - A^{(1)}(t) = [V(t), A(t)]$.

**Proof.** We prove the result using the notation of matrix groups although we note that the same line of proof holds for a general group. As before, set $Y(t) = V(t) - A(t)$ and without loss of generality, set $z(0) = x(0) = e$. The equation

$$\frac{d}{dt} \text{Ad}_{z(t)} Y(t) = \text{Ad}_{z(t)} (Y^{(1)}(t) + [A(t), Y(t)]) = 0$$

(4.40)

shows that $\text{Ad}_{z(t)} Y(t)$ is constant and so

$$Y(t) = \text{Ad}_{z(t)^{-1}} Y(0) = z(t)^{-1} Y(0) z(t).$$

(4.41)

Substituting the definition of $Y(t)$ gives us the following equation for $V(t)$:

$$V(t) = z(t)^{-1}(V(0) - A(0))z(t) + A(t).$$

(4.42)

Pre-multiplying by $x(t)$ and post-multiplying by $z^{-1}(t)$, we see that:

$$x^{(1)}z^{-1} = xz^{-1}(V(0) - A(0)) + xA z^{-1}$$

(4.43)
Applying the product rule on \((xz^{-1})^{(1)}\), we obtain:

\[
(xz^{-1})^{(1)} = xz^{-1}(V(0) - A(0)) + xAz^{-1} - xz^{-1}z^{(1)}z^{-1}
\]

\[
= xz^{-1}(V(0) - A(0)) + xAz^{-1} - xAz^{-1}
\]

\[
= (xz^{-1})(V(0) - A(0))
\]

(4.44)

and by solving the differential equation, we see that:

\[
x(t) = e^{(V(0)-A(0))t}z(t)
\]

(4.45)

This allows us to generate a whole family of solutions for when the prior vector field is integrable. There are many examples for which the solutions for \(z(t)\) are known.

- \(A(t) = At + B\) in \(\text{SL}(2, \mathbb{C})\) or \(\text{SO}(3)\) (Corresponding to \(z(t)\) being the inverse of a null Riemannian cubic)

- \(A(t)\) being a null Lie quadratic in \(\text{SL}(2, \mathbb{C})\) or \(\text{SO}(3)\) (Corresponding to \(z(t)\) being a null Riemannian cubic)

- \(A(t)\) is a multiple of a fixed vector. \(A(t) = f(t)A_0\) (\(z(t)\) is a reparameterised geodesic)

- \(A(t)\) is a Jupp and Kent quadratic, as seen in [31, Chapter 4]

Notice that we now have two alternate ways of solving for \(Y(t)\) in the equation \(Y^{(1)}(t) = [A_1t + A_0, Y(t)]\). One way relies on a special property of \(\text{so}(3)\) (where the adjoint map on \(E^3\) converts to regular multiplication in the matrix group) to convert the system into \(\Psi^{(1)}(t) = (\text{ad}(A_1)t + \text{ad}(A_0))\Psi(t)\). The second technique is more general and can reduce solving a linked equation of the form \(x^{(1)}(t) = x(t)V(t)\) and \(V^{(1)}(t) - A^{(1)}(t) = [V(t), A(t)]\) into solving for integral curves of \(A(t)\).
Suppose we now started with the equation \( V(1)(t) = [V(t), A(t)] \). If we know the solutions of \( z^{(1)}(t) = z(t)A(t) \), we would indeed be able to solve for \( V(t) \) using the same argument. We then however cannot easily construct a Lax pair and so we do not have a known way of integrating \( V(t) \) to find \( x(t) \). This shows that the form of the equation for \( V(t) \) is necessary for this theorem to work.

4.4.2 Time varying prior fields in SO(3) The method Pauley uses in [30] to solve null Riemannian cubics in \( SL(2, \mathbb{C}) \) are far more powerful than realised in the paper. The same methods can be used to solve a much wider class of differential equations. These additional solutions have applications in the case of time dependent prior fields. We will borrow heavily from Pauley’s methods to find solutions of interest.

What may well be the next simplest case after examining affine prior fields is the case when \( A(t) \) is of the form \( A(t) = f(t)A_1 + A_0 \) where \([A_1, A_0] \neq 0\). Recall the Lie reduction equation for time varying prior fields says \( Y^{(1)}(t) = \text{ad}(A(t))Y(t) \) and so we can write \( Y^{(1)}(t) = (f(t)\text{ad}(A_1) + \text{ad}(A_0))Y(t) \). Treating \( Y(t) \) as a curve in \( E^3 \), we extend the system to one involving the fundamental matrix \( \Psi(t) \), we have that \( \Psi^{(1)}(t) = (f(t)\text{ad}(A_1) + \text{ad}(A_0))\Psi(t) \) where \( \text{ad}(A_1), \text{ad}(A_0) \in \text{so}(3) \).

Because the factor of \( \Psi(t) \) is in \( T_{\Psi(t)} \text{SO}(3) \) and \( \Psi(0) = I_{3 \times 3} \in \text{SO}(3) \), it must be the case that \( \Psi(t) \in \text{SO}(3) \).

The equivalent problem in \( SL(2, \mathbb{C}) \) is \( \Psi^{(1)}(t) = (f(t)B_1 + B_0)\Psi^{(1)}(t) \) where \( B_1, B_0 \in \text{su}(2) \) are obtained by the usual Lie algebra isomorphism from \( \text{so}(3) \) into the Pauli matrices.

Because the matrix \( B_1 \) is skew-Hermitian, we can diagonalize it with a unitary matrix and the diagonal matrix will be skew-Hermitian too. Moreover since \( f(t) \) is arbitrary, we will assume \( B_1 \) is the diagonal Pauli matrix \( D(i, -i) \) and that the diagonal terms of \( B_2 \) are 0. We may assume the off diagonal terms of \( B_2 \) are non-zero otherwise we can solve the system with an integrating factor.

\[
\begin{pmatrix}
  p_1^{(1)} \\
  p_2^{(1)}
\end{pmatrix} = \begin{pmatrix} f(t) & i & 0 \\
  0 & 0 & -i 
\end{pmatrix} + \begin{pmatrix} 0 & z & 0 \\
  -\bar{z} & 0 & 0
\end{pmatrix} \begin{pmatrix}
  p_1 \\
  p_2
\end{pmatrix} \tag{4.46}
\]
Writing out these equations, we have

\begin{align}
  p_1^{(1)} &= if(t)p_1 + zp_2 \\
  p_2^{(1)} &= -if(t)p_2 - \bar{z}p_1
\end{align} \hspace{1cm} (4.47)

Differentiating (4.48) and substituting (4.47), we see that:

\begin{align}
  p_2^{(2)} &= -if^{(1)}(t)p_2 - if(t)p_2^{(1)} - \bar{z}(if(t)p_1 + zp_2) \\
  &= -if^{(1)}(t)p_2 - if(t)p_2^{(1)} + if(t)(if(t)p_2 + p_2^{(1)}) - \bar{z}zp_2 \hspace{1cm} (4.49)
\end{align}

Re-substituting (4.48) results in an equation with only \( p_2 \) terms

\begin{align}
  p_2^{(2)} &= -if^{(1)}(t)p_2 - if(t)p_2^{(1)} + if(t)(if(t)p_2 + p_2^{(1)}) - \bar{z}zp_2 \\
  &= -if^{(1)}(t)p_2 - if(t)p_2^{(1)} + if(t)(if(t)p_2 + p_2^{(1)}) - \bar{z}zp_2 \hspace{1cm} (4.50)
\end{align}

Simplified, this says, \( p_2^{(2)} + (\bar{z}z + f(t)^2 + if^{(1)}(t))p_2 = 0 \). By making the substitution \( g(t) = if(t), c^2 = z\bar{z} \geq 0 \), and setting \( \phi(t) := p_2(t) \), we are left with the task of solving the equation

\begin{align}
  \phi^{(2)}(t) + (c^2 - g(t)^2 + g^{(1)}(t))\phi(t) = 0 \hspace{1cm} (4.51)
\end{align}

Many possible functions for \( g(t) \) lead to this system having known solutions. Letting \( g(t) = \frac{1}{t+t_0} \), we obtain solutions \( \phi(t) = c_1(\frac{\cos(t)}{1+t} + \sin(t)) + c_2(\frac{\sin(t)}{1+t} - \cos(t)) \). This \( g(t) \) although making the ODE particularly easy to solve also could represent a situation where one vector fields effects are being damped over time. A more realistic situation with damping which may happen in the real world would be where we have a damping factor such as \( g(t) = e^{-\gamma t} \). While the solution is not straightforward, Mathematica shows it can be solved in terms of confluent hypergeometric functions of the second kind and generalised Laguerre polynomials.

Once \( p_2 \) is solved, \( p_1 \) is found from 4.48. The original \( f(t) \) is easily recoverable from the various simplifications that have been made and so we have completed what we set out to do.

Now that we have solved for \( \Psi(t) \), we can apply the homomorphism of SU(2)
into $\text{SO}(3)$ generated by the Lie algebra isomorphism and therefore produce a solution for $\Psi(t)$ in $\text{SO}(3)$. Recall that this represents the fundamental matrix of the solutions for $Y(t)$ and therefore $Y(t) = \Psi(t)Y(0)$ which means $V(t) = \Psi(t)Y(0) + A(t)$. Using the results on duality [21], we are then able to take this solution and produce a solution in quadrature for the time invariant conditional extremal $x(t)$.

We wish to investigate which functions $f(t)$ lead to a solvable system of equations. It is interesting to note there are two differential equations playing about in Equation 4.51 for $\phi(t)$.

\begin{align*}
\phi^{(2)}(t) + q(t)\phi(t) &= 0 \quad (4.52) \\
c^2 - g(t)^2 + g^{(1)}(t) &= q(t) \quad (4.53)
\end{align*}

The first of these is a second order linear ODE while the second is a first order non-linear ODE. When one tries to solve these equations in a computer package such as Mathematica, they find that solutions to (4.52) can be solved by the software when and only when solutions to (4.53) can be solved and moreover, the solutions are of a similar form. The family of equations given by (4.53) are a subclass of a more general family known as non-linear Riccati equations [13].

One interesting property of the Riccati equation ([38]) is that once we know a single solution, we can obtain the rest by quadrature.

Let $g_0(t)$ be a solution to (4.53), then the general class of solutions for (4.53) is:

\[ g(t) = y_0(t) + \Phi(t) \left( C - \int \Phi(t) dt \right)^{-1} \]

where $\Phi(t) = \exp \left( \int 2g_0(t) \right)$ and $C$ is an arbitrary constant.

If we make the substitution $\psi(t) = \exp \left( -\int g(t) \right)$, we see that

\[ \psi^{(2)}(t) + (q(t) - c^2)\psi(t) = 0 \]

This is a linear second order differential equation and so we see the system of
Chapter 4. Prior fields

The equations we are in fact solving for are

\[ \phi''(t) + q(t)\phi(t) = 0 \]
\[ \psi''(2) + q(t)\psi(t) = c^2\psi(t) \]

for some \( q(t) \). Another way of expressing this is that there is a differential operator \( \frac{d^2}{dt^2} + q(t) \) for which both 0 and \( c^2 \) are eigenvalues. By setting \( q(t) = -V(t) \) where \( V(t) \) is the negative of some potential function (the force), this operator is the negative of the well-known one-dimensional time-independent Schrödinger operator [42] (also called the Hamilton operator) \( H = -\nabla^2 + V(t) \) and if \( c^2 = -E \), we obtain the Schrödinger equations \( H\psi = E\psi \) and \( H\psi = 0 \). That is, we are after solutions to both a zero and a negative energy state. This also relates right back to the discussion in section 3.1.1 in which we described \( q(t) = 0 \).
4.4. Time varying systems

4.4.3 Second order time-variant systems A similar analysis can be made for the functional

\[ J_A^2(x) := \int_0^1 \| \nabla_t x^{(1)} - \mathcal{A}(x(t)) \|^2 dt \]  

(4.54)

The Euler Lagrange equations are given by

\[ \nabla_X^2 Z + R(Z, X) X - \phi_Z A \perp TQ \]  

(4.55)

where \( Z = \nabla_X X - A \). Recall that

\[ \nabla_X A = (\nabla_X^Q A + \frac{\partial}{\partial t} A, 0) \]  

(4.56)

\[ \nabla_X X = (\nabla_X^Q X, 0) \]  

(4.57)

Moreover, because of the curvature tensor of \( \mathbb{R} \) being 0 and the definition of the product metric, \( R(X, Y) Z = R^Q(\pi_1 X, \pi_1 Y) \pi_1 Z \) where \( \pi_1 \) is projection onto the first coordinate. To compute \( \phi_Z A \), we have seen that on the product manifold \( X \times Y \), that

\[ \nabla_{(U_X, U_Y)}(V_X, V_Y) = (\nabla_{U_X}^X V_X + dV_X(U_Y), \nabla_{U_Y}^Y V_Y + dV_Y(U_X)) \]

This implies that

\[ \phi_{(U_X, U_Y)}(V_X, V_Y) = (\phi_{U_X}^X V_X + dV_X^T(U_Y), \phi_{U_Y}^Y V_Y + dV_Y^T(U_X)) \]

Therefore \( \phi_Z A = (\phi_{Z_1} A, dA^T(Z_1)) \). Taking projections, we obtain the following Euler Lagrange equations on \( Q \):

**Theorem 4.4.6.** A critical point \( x(t) \) of the functional defined in Equation (4.54) satisfies:

\[ \nabla_X^2 Z - \frac{\partial}{\partial t}(\nabla_X A) - \nabla_X (\frac{\partial}{\partial t} A) - \frac{\partial^2}{\partial t^2} A + R(Z, X) X - \phi_Z A = 0 \]  

(4.58)
where \( X = x^{(1)}(t) \)

**Second order time varying systems on Lie groups**

Suppose now the manifold \( Q \) is a bi-invariant Lie group \( G \). In order to carry out the Lie reduction, we need to only focus on the three terms with the partial derivatives.

\[
\mathcal{L}(-\nabla_X A) = -\frac{1}{2} \left[ V, \frac{\partial A}{\partial t} \right]
\]

\[
\mathcal{L}\left(-\frac{\partial}{\partial t}(\nabla_X A)\right) = -\frac{1}{2} \left[ V, \frac{\partial A}{\partial t} \right]
\]

\[
\mathcal{L}\left(-\frac{\partial^2}{\partial t^2} A\right) = -\frac{\partial^2}{\partial t^2} A
\]

Combining this with the Lie reduction for the remaining equations to which we already know the value, we obtain the theorem:

**Theorem 4.4.7.** Let \( x(t) \) be a critical point of (4.54) in a bi-invariant Lie group with prior field \( A(t) \). Then the Lie reduction \( V(t) \) of \( x(t) \) satisfies

\[
V^{(3)}(t) - A^{(2)}(t) + [V(t), V^{(2)}(t) - A^{(1)}(t)] = 0 \quad (4.59)
\]

This then explains why the time invariant prior fields reduced to the Riemannian cubic equation. One substitution which can be made is if by setting \( Y(t) := V(t) - \int_0^t A(s)ds \), then \( Y(t) \) satisfies:

\[
Y^{(3)}(t) + \left[ Y(t) + \left( \int_0^t A(s)ds \right), Y^{(2)}(t) \right] = 0 \quad (4.60)
\]

If we set \( Z(t) = V^{(2)}(t) - A^{(1)}(t) \), then we see that \( Z^{(1)}(t) = [Z(t), V(t)] \) and so \((Z(t), V(t))\) is a Lax pair.

**Proposition 4.4.8.** The Lie reduction \( V(t) \) of a critical point of 4.54 satisfies:

\[
\|V^{(2)}(t) - A^{(1)}(t)\| = c \quad (4.61)
\]

\[
\langle V^{(3)}(t) - A^{(2)}(t), V(t) \rangle = 0 \quad (4.62)
\]
4.5. Conclusion

**Proof.** Take inner products of Equation (4.59) with $V^{(2)}(t) - A^{(1)}(t)$ and $V(t)$ respectively and use the bi-invariance of the metric.

**Corollary 4.4.9.** If $A(t)$ is a smooth function of $t$ defined on all of $\mathbb{R}$, then for a fixed $V(0)$, $V^{(1)}(0)$ and $V^{(2)}(0)$, $V(t)$ extends to a unique $C^\infty$ solution of 4.59 also defined on all of $\mathbb{R}$.

**Proof.** The constraint on $V^{(3)}(t)$ is smooth in all variables and so we have local existence and uniqueness by the Picard-Lindelöf theorem. A solution $V(t)$ can only fail to be indefinitely extended when any of $V(t)$, $V^{(1)}(t)$ and $V^{(2)}(t)$ diverge in finite time (See [47, Chapter 1]). Since $A$ is smooth, on any interval $[t_0, t_1]$, $A$ and its derivatives are bounded. Suppose then that $\|A^{(1)}(t)\| \leq 2k - c$. Then we have from Equation (4.61) that $\|V^{(2)}(t)\| \leq c + \|A^{(1)}(t)\| \leq 2k$. This also implies that $\|V^{(1)}(t)\| \leq 2kt + \|V^{(1)}(0)\|$ and $\|V(t)\| \leq kt^2 + \|V^{(1)}(0)\|t + \|V(0)\|$ implying none of $V(t)$, $V^{(1)}(t)$ and $V^{(2)}(t)$ diverge in finite time.

**Corollary 4.4.10.** If $G$ is compact, then there is a unique conditional extremal $x : \mathbb{R} \to G$ satisfying $x(0) = x_0$, $x^{(1)}(0) = v_0$, $\nabla_t x^{(1)}(0) = v_1$ and $\nabla^2_t x^{(1)}(0) = v_2$.

**Proof.** Corollary 4.4.9 guarantees a solution for $V(t)$ for all $t \in \mathbb{R}$. Lemma 2.4 in [18] says that a smooth vector field with compact support has flows defined for all $t$. Apply this lemma to the vector field $V(x, t) = xV(t)$ defined on $G \times [t_0, t_1]$ where $t_0$ and $t_1$ are arbitrary.

### 4.5 Conclusion

In summary, we have described $A$-Jacobi fields for first order prior fields and have derived solutions and properties for $A$-Jacobi fields in the special case bi-invariant Lie groups with a left-invariant prior field. We have shown that in the case of an unknown left-invariant prior fields, that the average of the initial and final velocities defines the optimal choice of prior field. For second order prior fields, the Euler-Lagrange equations were derived in the case of a general Riemannian manifold and the Lie reduced Euler-Lagrange equations in the case of a bi-invariant Lie group with a left-invariant prior field were obtained. In the
latter case, somewhat surprisingly, the prior field does not appear in the equations except as boundary conditions. We extend the concept of prior fields to those which may vary with time for both first and second order conditional extremals and extend the framework for handling this situation. Once specialising to the Lie group setting, we are able to calculate solutions in a variety of situations and mention an interesting connection to Riccati equations and the Schroedinger equation. For second order time varying prior fields on a compact Lie group, we prove the extendibility of solutions to all values of time.

We summarise the extremals considered so far and their corresponding functionals (symbol and formula) with a table. Time varying versions refer to the vector field $\mathcal{A}$ being a function of time as well as position in the manifold.

<table>
<thead>
<tr>
<th>Name of extremals</th>
<th>Symbol</th>
<th>Functional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geodesics</td>
<td>$J_1$</td>
<td>$\int_0^1 |x^{(1)}|^2 dt$</td>
</tr>
<tr>
<td>Riemannian cubics</td>
<td>$J_2$</td>
<td>$\int_0^1 |\nabla_t x^{(1)} - \mathcal{A}|^2 dt$</td>
</tr>
<tr>
<td>Riemannian cubics in tension</td>
<td>$J_2^\lambda$</td>
<td>$\int_0^1 |\nabla_t x^{(1)} - \mathcal{A}|^2 dt + \lambda J_1$</td>
</tr>
<tr>
<td>First order conditional extremals</td>
<td>$\mathcal{A}J_1$</td>
<td>$\int_0^1 |x^{(1)} - \mathcal{A}|^2 dt$</td>
</tr>
<tr>
<td>Second order conditional extremals</td>
<td>$\mathcal{A}J_2^\lambda$</td>
<td>$\int_0^1 |\nabla_t x^{(1)} - \mathcal{A}|^2 dt + \lambda (\mathcal{A}J_1)$</td>
</tr>
</tbody>
</table>

Table 4.1: List of various extremals with their corresponding functional’s symbol and expression.
Linear interpolation schemes on bi-invariant Lie groups

5.1 Introduction

In the previous two chapters, we concerned ourselves with Riemannian cubics in tension and conditional extrema in first and second order settings, as well as non-autonomous first order systems. What remains to be discussed are algorithms for interpolating efficiently and accurately. As we will see in this chapter, the problem of interpolation requires the solving of an inherently coupled system of boundary value problems. Even in the case for elastic curves in $\mathbb{E}^2$ where the solution is completely known [15], their use in interpolation applications is extremely limited due to the difficulty of satisfying nonlinear boundary value, interior value and smoothness conditions. Polynomial splines are effective because of their ability to reduce the system of equations to a linear system. Noakes in [27] introduces a technique for interpolating data that falls near a geodesic by using approximations to a natural Riemannian cubic spline. In this chapter, we give an exposition of these ideas and how they apply to the case of second order conditional extremals on a bi-invariant Lie group with $\mathcal{A}$ left-invariant. Moreover, we employ a technique analogous to those of the traditional $B$-splines in Euclidean space to reduce the problem of approximating a natural conditional extremal spline into a solving a tridiagonal linear system.

Let $Q$ be a smooth $n$-dimensional Riemannian manifold. A data list on $Q$ is an ordered $(N+1)$-tuple

$$D = \left((t_0, \xi_0), (t_1, \xi_1), \ldots, (t_N, \xi_N)\right)$$

where $0 = t_0 < t_1 < \cdots < t_N = 1$ and $\xi_i \in Q$ for $0 \leq i \leq N$.

Given a data list $D$, we say a curve $x(t)$ is $D$-feasible if $x(t_i) = \xi_i$ for all $0 \leq i \leq N$. Let $\mathcal{D} \subset ([0,1] \times M)^N$ be a non-empty set of data lists, then an interpolator is a $C^k$ function $I : \mathcal{D} \times [0,1] \to M$ such that $x_D(t) := I(D,t)$ is $D$ feasible for all $D \in \mathcal{D}$. We say $I$ is linear when given $D$, $I(D,t)$ can be computed
Chapter 5. Linear interpolation schemes on bi-invariant Lie groups

by solving a system of linear equations. We are interested in producing a linear interpolator which attempts to approximately minimise a variational problem.

**Definition 5.1.1.** Let \( R \) be an open neighbourhood of the 0-section in \( \Gamma(TQ) \). A \( C^\infty \) function \( E : R \rightarrow Q \), with all its derivatives bounded, is said to be quasi-exponential when, for any \( q \in Q \) and any \( v \in T_qQ \),

\[
E(q, 0) = q \quad \text{and} \quad \frac{d}{dh} E(q, hv) \bigg|_{h=0} = v \quad (5.1)
\]

The exponential map (on any complete Riemannian manifold) is quasiexponential although for the purpose of linear interpolation in SO(3), the following map is preferable

**Definition 5.1.2.** The Cayley transform \( T : TSO(3) \rightarrow SO(3) \) is given by the formula

\[
T(q, v) = q \left(1 + \frac{L(v)}{2}\right) \left(1 - \frac{L(v)}{2}\right)^{-1} \quad (5.2)
\]

where we recall that \( L(v) \) is the Lie reduction of \( v \). The reason we will use the Cayley transform is because calculating the inverse Cayley transform is performed by solving a linear system of equations.

### 5.2 Second order prior fields in SO(3)

We will begin by studying interpolants minimising the functional relating to second order time varying left-invariant prior fields in SO(3). First order prior fields are only required to be piecewise \( C^1 \) (for example, when \( A = 0 \), we have piecewise geodesic interpolation) and therefore the boundary value problems associated with these are considerably simpler. Consider the functional \( J \) restricted to \( D \)-feasible curves:

\[
A J_2(x) = \sum_{i=1}^{N-1} \int_{t_i}^{t_{i+1}} \| \nabla_t x^{(1)} - A(x(t)) \|^2 dt \quad (5.3)
\]

Minimizers of this functional are known as natural second order conditional splines. We are interested in constructing an interpolator which approximates
these splines and comes close to being a minimizer of $A J_2$.

Given a variation $x_h$ through piecewise extremals of $A J_2$ between any two
knots of this functional, if we differentiate with respect to $h$ and ignore the terms
corresponding to the Euler Lagrange equations (which are 0 since we vary through
piecewise extremals), we are left with:

$$
\left. \frac{d}{dh} \right|_{h=0} A J_2(x_h) = \sum_{i=1}^{N-1} \left[ \langle \nabla_t \frac{dx_h}{dh}, \nabla_t x_h^{(1)} - A \rangle - \langle \frac{dx_h}{dh}, \nabla_t^2 x_h^{(1)} - \nabla_t A \rangle \right]_{t_i}^{t_{i+1}} \tag{5.4}
$$

After applying a Lie reduction, we see that:

$$
\left. \frac{d}{dh} \right|_{h=0} A J_2(x_h) = \sum_{i=1}^{N-1} \left[ \langle W^{(1)} + \frac{1}{2} [V, W], V^{(1)} - A \rangle - \langle W, V^{(2)} + \frac{1}{2} [V, V^{(1)}] \rangle \right]_{t_i}^{t_{i+1}}
$$

Because we can set $W^{(1)}$ arbitrarily and in such a way that its support only
contains one $t_i$, we obtain the following equations:

$$
V^{(1)}(t_i) = A(t_i) \tag{5.5}
$$

$$
V^{(1)}(t_j^-) - A(t_j^-) = V^{(1)}(t_j^+) - A(t_j^+) \tag{5.6}
$$

for all $i \in \{0, N\}$ and $j \notin \{0, N\}$.

Although we previously required $x(t)$ be piecewise $C^2$, we can relax conditions
on differentiability at the interior knots in a way so that the integral is still defined.
In order for this to occur, discrete approximations to $\int_0^1 \| \nabla_t x^{(1)} - A \|^2$ must not
be too large and so we require that $x(t)$ be $C^1$ at the knots. Even under this
restriction, we find that equation (5.6) imposes the condition optimal solutions be
$C^2$.

Equations (5.6) and (5.5) together with the Euler-Lagrange Equations (4.58)
between any two knots give us all the necessary properties for critical points of
this interpolation problem and show that $V$ must be continuously differentiable.
To find a critical point, we must piece together solutions of the Euler Lagrange
equations in such a way that, at the first and last knots, we have $V^{(1)} = A$ and
moreover we require \( V^{(1)} \) to be continuous at the interior knots. While solving explicitly for \( V \) is difficult, we may approximate \( V \) by taking a simple extremal \( V_0 \) which satisfies \( V_0^{(1)}(t) = A(t) \) everywhere (not necessarily interpolating the interior data points), and then taking a Taylor expansion about this curve. If we are minimising the functional
\[
\int_{t_0}^{t_N} \| V^{(1)}(t) - A(t) \|^2 dt,
\]
then starting with the segment connecting \( \xi_0 \) and \( \xi_N \) with \( V_0^{(1)}(t) = A(t) \) is a reasonable first attempt.

Determining this curve requires solving a boundary value problem and is in general a difficult problem and we will restrict ourselves to cases for which it is possible. The easiest setting to be able to solve this is in the case that \( A \) is constant.

5.2.1 \( A(t) = \text{constant} \). We firstly require a curve \( x_0 : [0, 1] \to \text{SO}(3) \) satisfying \( V_0^{(1)}(t) = A \), \( x_0(0) = \xi_0 \), \( x_0(1) = \xi_N \) and define \( v_0 = V_0(0) \). It is shown in [21] that the solutions for such a system are inverses of null Riemannian cubics. The solutions to these curves have been solved exactly in terms of special functions [30] for the Lie group \( \text{SL}(2, \mathbb{C}) \). Although it is mentioned how to calculate null Riemannian cubics in \( \text{SO}(3) \) using the solutions of null Lie quadratics in \( \text{su}(2) \subseteq \text{sl}(2, \mathbb{C}) \) together with a Lie algebra homomorphism, we choose a slightly different method by directly applying a Lie group homomorphism to the null Riemannian cubics in \( \text{SU}(2) \subseteq \text{SL}(2, \mathbb{C}) \).

Define a homomorphism \( \phi : \text{SU}(2) \to \text{SO}(3) \) given by:

\[
\phi \left( \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix} \right) = \begin{bmatrix} \Re(a^2 - b^2) & \Im(a^2 + b^2) & -\Re(2ab) \\ -\Im(a^2 - b^2) & \Re(a^2 + b^2) & \Im(2ab) \\ ab + \bar{a}b & i(\bar{a}b - ab) & a\bar{a} - \bar{b}\bar{b} \end{bmatrix}
\]

where \( \Re \) and \( \Im \) are the real and imaginary components.

Let \( \bar{y} : [0, 1] \to \text{SU}(2) \) and suppose that \( \bar{y}^{(1)}(t) = (\bar{A}t + \bar{v}_0)\bar{y}(t) \) with \( \bar{y}(0) = I \). Because \( \bar{A} \) is skew-Hermitian, it is normal and therefore diagonalisable by a unitary matrix \( P \). Define \( A := P\bar{A}P^{-1} \), define \( y(t) := P\bar{y}(t)P^{-1} \) and \( v_0 := P\bar{v}_0P^{-1} \). Then \( y^{(1)}(t) = (At + v_0)y(t) \) with \( A \) diagonal. Moreover, conjugation by a unitary matrix preserves the property of being skew-Hermitian and therefore
y(t) ∈ SU(2) and At + v_0 ∈ su(2). Write At + v_0 = \[
\begin{pmatrix}
a(t) & -v_{21} \\
v_{21} & -a(t)
\end{pmatrix}.
\]

**Theorem 5.2.1** (Pauley [30]).

\[
y(t) = \begin{pmatrix}
\frac{1}{v_{21}}(\psi_1^{(1)} + a(t)\psi_1) \\
\frac{1}{v_{21}}(\psi_2^{(1)} + a(t)\psi_2)
\end{pmatrix}
\]  \hspace{1cm} (5.7)

where \(\psi_1, \psi_2\) are solutions of the initial value problem

\[
\psi^{(2)} + v(t)\psi(t) = 0 \quad \psi_1(0) = 0 \quad \psi_1^{(1)} = v_{21}
\]
\[
\psi_2(0) = 1 \quad \psi_2^{(1)} = -a(0)
\]

and \(v(t)\) is the quadratic function \(a^{(1)}(t) - a(t)^2 + \frac{v_{21}}{2}v_{21}\).

The solutions for \(\psi_1\) and \(\psi_2\) are parabolic cylinder functions which are thoroughly analysed in [4, Chapter 8]. Taking \(\phi(y(t))\) allows us to describe a map \(\Phi : so(3) → SO(3)\) that takes the initial velocity \(v_0\) and maps to \(\phi(y(1))\). The choice of the initial guess for \(v_0\) must be made with care because the observed data may have been sampled from a trajectory which loops SO(3) several times. As \(\phi\) and \(y\) are known, the non-linear system \(\Phi\) can be solved with a computer package such as \texttt{NSolve} in Mathematica which uses a variant of the Newton method [50]. Figure 5.1 shows an affine Riemannian cubic in its action (by multiplication) on the standard basis vectors in \(\mathbb{R}^3\) computed explicitly using parabolic cylinder functions. After producing the first guess to an interpolator, \(x_0(t)\), we then need a process to provide an interpolant that also satisfies the interior knot condition. We must be prepared to trade off accuracy in optimising the functional for efficiency in producing such a curve. Our strategy involves first approximating prior extremals \(x(t)\) by taking the derivative of variations of \(x_0\) through piecewise prior extremals. Let \(x_h(t)\) be a variation through prior extremals such that \(x_h(0) = x_0, x_h(1) = x_N\) and \(x_1(t) = x(t)\). We can define a vector field \(W_h(t)\) along \(x_h(t)\) defined by \(W_h(t) := \frac{\partial}{\partial h} x_h(t)\). The solutions for \(W_h(t)\) form a linear subspace of sections over \(x_h(t)\) because they will satisfy a linear differential equation. We
can approximate $x(t)$ by a curve $\hat{x}(t)$ defined by $\hat{x}(t) := T(x_0(t), W(t))$ where $W(t) := W_0(t)$. Our goal therefore is to choose $W(t)$ in such a way that $\hat{x}(t)$ is an interpolant (although not necessarily a conditional extremal) and see how well such an approximation minimises $J_A$.

One approach to computing $W(t)$ can be done by differentiating Euler Lagrange equations for $x_h(t)$ with respect to $h$ and then solving the resulting differential equation. This approach leads to rather complicated equations. What can be done instead is varying $V_h$, the Lie reduction of $x_h$.

We have seen that $V_h$ satisfies $V_h^{(3)} + [V_h, V_h^{(2)}] = 0$. If we differentiate with respect to $h$ and let $X_h := \frac{\partial}{\partial h} V_h$, we see that $X_h^{(3)} + [X_h, V_h^{(2)}] + [V_h, X_h^{(2)}] = 0$. Substituting $h = 0$ and setting $X := X_0$ gives us the formula

$$X^{(3)} + [V_0, X^{(2)}] + [X, V_0^{(2)}] = 0 \quad (5.8)$$

Now to see the relationship between $W$ and $X$, we see that:

$$X_h = \frac{\partial}{\partial h} V_h = \frac{\partial}{\partial h} (x_h^{-1} x_h^{(1)}) = -x_h^{-1} \left( \frac{\partial}{\partial h} x_h \right) x_h^{-1} x_h^{(1)} + x_h^{-1} \frac{\partial}{\partial h} x_h^{(1)}$$

$$= -\tilde{W}_h V_h + x_h^{-1} \frac{\partial}{\partial t} \left( \frac{\partial}{\partial h} x_h \right) = -\tilde{W}_h V_h + x_h^{-1} \frac{\partial}{\partial t} (x_h \tilde{W}_h)$$

$$= -\tilde{W}_h V_h + x_h^{-1} x_h^{(1)} \tilde{W}_h + \frac{\partial}{\partial t} \tilde{W}_h = -\tilde{W}_h V_h + V_h \tilde{W}_h + \frac{\partial}{\partial t} \tilde{W}_h$$
\[\frac{\partial}{\partial t} \tilde{W}_h + [V_h, \tilde{W}_h]\]

where \(\tilde{W}_h(t) := \mathcal{L}(W_h(t))\). This proves that the relationship between \(X\) and \(W\) is given by:

\[X = \tilde{W}^{(1)} + [V_0, \tilde{W}]\]  \hspace{1cm} (5.9)

Solving for \(W\) has therefore been reduced to solving the two linear differential equations 5.8 and 5.9 where we first solve for \(X\) and then for \(\tilde{W}\). \(W\) is then obtainable by calculating \(W(t) = x_0(t)\tilde{W}(t)\). We must now examine what conditions are placed upon a solution for \(W\).

We firstly require the endpoint positions to match up and so \(\hat{x}(0) = \mathcal{T}(x_0(0), W(0)) = x_0\) as well as \(\hat{x}(1) = \mathcal{T}(x_0(1), W(1)) = x_N\). This condition says that \(W(0) = W(1) = 0\) since \(x_0(0) = x_0\) and \(x_0(1) = x_N\). The second conditions we require are those imposed by Equation 5.5 which says that \(\hat{V}^{(1)}(0) = \hat{V}^{(1)}(1) = A\). These conditions are equivalent to the conditions

\[\nabla_t \hat{x}^{(1)}(i) = x(i)A\]  \hspace{1cm} (5.10)

for \(i = 0, 1\). Supposing now \(\hat{x}_h = \mathcal{T}(x_0, W_h) = \mathcal{T}(x_0, hW)\), we may differentiate Equation (5.10) with respect to \(h\) and obtain

\[\nabla_h \nabla_i \hat{x}^{(1)}(i) = \nabla^2 W_h(i) + R(W_h(i), x^{(1)}(i))x^{(1)}(i) = 0\]  \hspace{1cm} (5.11)

Setting \(h = 0\) and using the fact \(W(i) = 0\), we can take the Lie reduction of Equation (5.11) and obtain the following boundary condition on \(\tilde{W}\):

\[\tilde{W}^{(2)}(i) + [V_0(i), \tilde{W}^{(1)}(i)] = 0.\]  \hspace{1cm} (5.12)

The Equations (5.8) and (5.9) give \(X(t)\) and \(W(t)\) as solutions to a third and first order linear system respectively in a three dimensional vector space. We
would like to have a method for efficiently solving this linear system. Supposing that no two knot points \( x_0(t_i) \) and \( x_0(t_j) \) are conjugate with respect to the conditional extrema, for each \( i = 1, \ldots, N \), there will be 12 linearly independent solutions for \( W(t) \) with support on \([t_{i-1}, t_i]\) and together these will span a space of solutions \( K \) for which elements may not be continuous let alone \( C^2 \) at the knots. The condition that \( W(t) \) be \( C^2 \) on \([0, 1]\) is a linear condition imposed on \( K \) and thus forms a vector subspace \( L \). We wish to choose a suitable basis for \( L \) so that we may solve for a specific \( W(t) \). Imitating ideas from polynomial interpolation, we will construct an analogue of a \( B \)-spline basis (which we will hereon also refer to as \( B \)-splines) of variation fields along the affine cubic. If we restrict our “basis splines” to that of minimal support, we can solve the equations above efficiently. A \( B \)-spline basis of this form will lead to a block-tridiagonal system of equations.

\[ \text{Figure 5.2: Components of example B-Spline variation fields in } E^3 \text{ along } x_0. \]

*Constructing a basis of \( L \)*

To count dimensions of the basis, we notice that for each of the \( N \) segments, there are \( 4DN \) different variations where \( D \) is the dimension of the Lie algebra. There are however twice differentiability (because \( V \) is once differentiable) conditions on the interior knots giving \( 3D(N - 1) \) constraints. We will suppose for now that the exterior knots are free as we may wish to apply other methods of constructing the interpolant. This leaves us with \((N + 3)D\) basis elements, whose coefficients we will solve for.
The basic building block of the basis will consist of $B^i_j$ for $i = 2, \ldots, N - 2$ and $j = 1, \ldots, D$. This element will be such that $B^i_j(t_i) = E_j$ with support only on $(t_{i-2}, t_{i+2})$. We must take a dimension count to see that this indeed does uniquely specify $B^i_j$. For a given $B^i_j$, the support will consist of regions between the knot points $\{t_{i-2}, t_{i-1}, t_i, t_{i+1}, t_{i+2}\}$. The subspace of $K$ with support inside this region has dimension $4D$. Being $C^2$ at each knot adds $3D$ constraints and specifying $(B^i_j)^{(l)}(t_i) = \delta^i_k E_j$ adds $D$ constraints. Together, the $B^i_j$ will provide $(N - 3)D$ basis elements for $L$.

The remaining $6D$ basis elements can be made with support around the boundary and can be constructed in two ways. The first way is to use the elements $k B^i_j$ where $i = 1, N$, $j = 1, \ldots, D$ and $k = 0, 1, 2$. In this case we have $k B^i_j(t_i) = \delta^i_k E_j$ and the support will be along three segments. On the other hand, we can extend the curve $x_0(t)$ and add 6 additional knots $t_{-3}, t_{-2}, t_{-1}, t_{N+1}, t_{N+2}, t_{N+3}$ spaced with uniform distance $\frac{1}{N}$ from the knots $t_0$ and $t_N$ (or a suitable spacing to avoid conjugate points) and extend the basis set $B^i_j$. We will focus on the latter approach as it provides a more unified approach.

The way we construct such a basis $B^i_j$ of $L$ is by first producing a basis for $K$ by computing solutions for $\tilde{W}$. That is to say, between any two knots $(t_i, t_{i+1})$, we solve for $k \tilde{W}^i_j$ corresponding to each of the $4D$ initial values:

\[
(k \tilde{W}^i_j)^{(l)}(t_i) = \delta^i_k E_j \tag{5.13}
\]

where $i \in \{-3, -2, \ldots, N + 2, N + 3\}$, $j \in \{1, 2, 3\}$, and $k \in \{0, 1, 2, 3\}$.

Initial value problems are considerably less expensive to solve and it is important we solve them to produce the endpoint values $k W^i_j(t_{i+1})$. Once we have solved all such initial value problems, the task of putting together the B-splines reduces to solving $4ND$ small linear equations for coefficients $k \alpha^i_q$ where

\[
B^i_j = \sum_{q=-2}^{1} \sum_{j=1}^{3} \sum_{k=0}^{3} k \alpha^i_q \cdot k W^i_{j+q} \tag{5.14}
\]
Chapter 5. Linear interpolation schemes on bi-invariant Lie groups

The B-splines are useful for solving the boundary value problem efficiently. We have the important condition that \( V^{(1)}(0) = V^{(1)}(1) = A \). Because our \( V_0 \) already satisfies this condition, this imposes a straightforward condition on \( \tilde{W} \). The condition that we do not affect the covariant derivative at the endpoints as we vary \( W \) says that \( \nabla_W \nabla_{\xi} x^{(1)} = 0 \) at \( t = t_0, t_N \). This is precisely the condition any variation through geodesics must satisfy, and so \( W \) satisfies the Jacobi field condition on the endpoints. That is to say, \( \nabla^2_t W + R(W, x^{(1)}) x^{(1)} = 0 \) at \( t = t_0, t_N \).

In SO(3), the Jacobi field condition says that

\[
W^{(2)} + \frac{1}{2} [V, W^{(1)}] + \frac{1}{4} [V, [V, W]] + \frac{1}{4} [V, [V^{(1)}, V]] = W^{(2)} + \frac{1}{2} [V, W^{(1)}] = 0
\]

This imposes the final 2D linear conditions tying \( W, W^{(1)} \) and \( W^{(2)} \) together at the endpoints. The system one gets when solving for the coefficients of the B-spline basis to construct \( W \) can be arranged into a block tridiagonal form which has numerous numerical advantages when it comes to solving such systems for a large amount of points \[48\]. This is because the value of \( W(t_i) \) is determined completely by knowing only the coefficients of \( B_j^{i-1}, B_j^i \) and \( B_j^{i+1} \) where \( j \in \{1, 2, 3\} \).

\[
\begin{pmatrix}
    C_{j-1}^0 & C_j^0 & C_j^1 & 0 & 0 & \ldots & 0 & 0 & 0 \\
    B_j^{i-1} & I & B_j^i & 0 & 0 & \ldots & 0 & 0 & 0 \\
    0 & B_j^0 & I & B_j^2 & 0 & \ldots & 0 & 0 & 0 \\
    0 & 0 & B_j^1 & I & B_j^3 & \ldots & 0 & 0 & 0 \\
    0 & 0 & 0 & B_j^2 & I & \ldots & 0 & 0 & 0 \\
    \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
    0 & 0 & 0 & 0 & \ldots & I & B_j^N & 0 \\
    0 & 0 & 0 & 0 & \ldots & B_j^{N-1} & I & B_j^{N+1} \\
    0 & 0 & 0 & 0 & \ldots & C_j^{N-1} & C_j^N & C_j^{N+1} \\
\end{pmatrix}
\begin{pmatrix}
    c_{j-1}^{-1} \\
    c_j^0 \\
    c_j^1 \\
    \vdots \\
    c_j^{N-1} \\
    c_j^N \\
    c_j^{N+1}
\end{pmatrix}
= 
\begin{pmatrix}
    0 \\
    0 \\
    W(t_1) \\
    W(t_2) \\
    W(t_3) \\
    \vdots \\
    W(t_{N-1}) \\
    0 \\
\end{pmatrix}
\]

**Figure 5.3:** Tridiagonal matrix to compute B-spline coefficients.

Here, all matrices are evaluated at times given by their rows where the first and last rows also correspond to \( t_0 \) and \( t_N \) respectively. The matrices \( C_j^i \) correspond to the functions \( C_j^i = (B_j^i)^{(2)} + \frac{1}{2} [V_0, (B_j^i)^{(1)}] \) and are there to impose the boundary
Second order prior fields in SO(3)

conditions on $W$. This $D(N + 3)$ by $D(N + 3)$ system although large can be solved in relatively few operations. The Thomas algorithm for block tridiagonal matrices [48] with fixed size blocks can solve the system in $O(N)$ running time. Although the system is not quite in block tridiagonal form due to the presence of $C_j^1$ and $C_j^{N-1}$, a few simple block row operations fixes this.

Because $W(t_i)$ is computed by solving a linear equation when the Cayley transform is used, we therefore have a linear interpolation scheme for when $x_0$ and $x_N$ are fixed and $x_0(t)$ is known. We write our solution

$$\hat{x}(t) = T\left(x_0, \sum_{i=1}^{N+1} \sum_{j=1}^{3} c^j_i B^i_j\right)$$ (5.15)

To empirically test our algorithm against something potentially crude, we suppose that the values of $W(t_i)$ in the Lie algebra $E^3$ are interpolated using ordinary cubic splines rather than with $W(t)$. We feed an initial guess of $v_0 = A$ when solving for $x_0$ as that is the direction our physical model says we’re accelerating to. The data has been selected by taking a random perturbation of data which was generated to very roughly fit the model. Figures 5.4, 5.5, and 5.6 demonstrate the algorithm working in SO(3). These figures show the curve $x(t)$ acting by multiplication on two standard basis vectors in $E^3$ ($e_1$ and $e_2$). The blue dots are the data set, the black curves are the action of $x(t)$, the dashed red curve represents $W(t_i)$ interpolated with a basic Euclidean cubic spline in $E^3$, and the dashed green curve is the action of $x_0(t)$ on the basis vectors. The arrows drawn on the data set represent scaled push-forwards of $A$ at the given point on $S^2$. The $A J_2$ values are compared between our algorithm and using the Euclidean cubic spline interpolation of $W(t_i)$. 
Figure 5.4: Demonstration of algorithm for $N = 4$ for $A = (-3, 1, 4)$. It can be seen that acceleration in the direction of $A$ is preferred for $x(t)$ in the initial hook. The corresponding $\mathcal{A}J_2$ values are 47.71 and 268.82.
Figure 5.5: Algorithm for $N = 4$ for $A = (-3, 1, 4)$. The corresponding $A J_2$ values are 46.19 and 378.63.
Figure 5.6: Algorithm for $N = 10$ for $A = (5, 0, 5)$. The corresponding $A J_2$ values are 740.245 and 3976.53.
5.2. Second order prior fields in SO(3)

<table>
<thead>
<tr>
<th>N</th>
<th>time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.624</td>
</tr>
<tr>
<td>10</td>
<td>0.842</td>
</tr>
<tr>
<td>20</td>
<td>1.451</td>
</tr>
<tr>
<td>50</td>
<td>3.354</td>
</tr>
<tr>
<td>100</td>
<td>7.582</td>
</tr>
<tr>
<td>200</td>
<td>19.407</td>
</tr>
</tbody>
</table>

Table 5.1: Time taken to compute interpolant for different data sizes.

Table 5.1 is a list of times taken to solve the interpolation problem using Mathematica’s `Timing[]` on a Pentium i7 3.2 GHz computer. As we can see, the time taken appears to be following the pattern we expect which is linear growth in the number of knots.

**Analysis of algorithm** The question remains how accurately the interpolator approximates conditional extremals. The original motivation of such conditional extremal interpolators is that we wish to interpolate data that we expect to roughly follow the model \( \nabla_t x^{(1)}(t) \approx A \). Suppose that there is a curve \( x_0(t) \) such that \( \nabla_t x_0^{(1)}(t) = A \) and \( d(x_0(t_i), \xi_i) < \delta \). We are interested in what can be said about the interpolating curve \( \hat{x}(t) \) produced by the linear interpolation scheme that we have described. We closely follow the analysis of Noakes [27] and point out a few differences along the way. Contained in these proofs are expository ideas that explain how the error bounds are achieved.

Denote the left hand side of the Euler-Lagrange equation (4.28) by \( EL_A(x) \). Then the \( A \)-Jacobi field equations for the second order system about a curve \( x_0 \) are given by:

\[
\nabla_{\tilde{W}} (EL_A(x))|_{x=x_0} = 0 \tag{5.16}
\]

This equation is linear in the variable \( \tilde{W} \) and solutions for \( \tilde{W} \) form a \( 4D \) dimensional real vector space, where \( D \) is the dimension of \( G \). Let \( \tilde{Y} \) be any \( C^\infty \) variation of \( x_0 \) and define a curve \( x_{\tilde{Y}}(h, t) := E(x_0(t), h\tilde{Y}(t)) \) where \( E \) is a quasiexponential map. Define:

\[
\|\tilde{Y}\|_{k, \infty} := \max\{\|\nabla^j_t \tilde{Y}(t)\| : 0 \leq j \leq k, 0 \leq t \leq 1\} \tag{5.17}
\]
\[ H\tilde{Y} := \nabla_h x_\tilde{Y}(h,t)|_{h=0} = \nabla_h^2 \tilde{Y}(t) + R(\tilde{Y}(t), x_0^{(1)}(t)) x_0^{(1)}(t) \quad (5.18) \]

where \( \nabla_h^0 \) is the identity map.

**Lemma 5.2.2.** Let \( \tilde{Y} \) be \( C^\infty \) and suppose for \( H\tilde{Y}(t_*) = 0 \) for \( t_* = t_0, t_N \). Then for some \( C_{ext} > 0 \) depending on \( E \) and \( x_0 \), and for all \( h \in [0,1] \),

\[ \| \nabla_t x_\tilde{Y}^{(1)}(h,t_*) - \mathcal{A}(x_\tilde{Y}(h,t_*)) \| \leq C_{ext}\|Y\|_{2,\infty} \]

**Proof.** The proof follows that of [27, Lemma 3.5] but now \( \nabla_t x_\tilde{Y}(0,t_*) = \mathcal{A}(x_0(t_*)) \) and we use the fact that \( \nabla_h\mathcal{A} \) and \( \nabla_h^2\mathcal{A} \) are of order \( \|\mathcal{A}\|\|\tilde{Y}\|_{2,\infty}^2 \) since \( \mathcal{A} \) is left-invariant and \( E \) is quasiexponential. \( \square \)

**Lemma 5.2.3.** Let \( \tilde{W} \) be a \( \mathcal{A} \)-Jacobi field. Then for some \( C_{int} > 0 \) depending on \( E \) and \( x_0 \), for all \( h \in [0,1] \), and \( t \in [0,1] \),

\[ \| \text{EL}_\mathcal{A}(x_\tilde{W}) \| \leq C_{int}\|\tilde{W}\|_{3,\infty}^2 \]

**Proof.** The proof follows that of [27, Lemma 3.7] because we have already shown in Theorem 4.3.2 that for left-invariant \( \mathcal{A} \), the Euler-Lagrange equation of a second order conditional extremal is equivalent to that of a Riemannian cubic. \( \square \)

Given \( \mathcal{A} \)-Jacobi fields \( \tilde{W}^1, \ldots, \tilde{W}^N \) along \( x_0 \), their track-sum \( \tilde{W} \) is defined by \( \tilde{W}(t) := \tilde{W}^i(t) \) for \( t \in [t_{i-1}, t_i) \) and \( \tilde{W}(t_N) := W^N(t_N) \). Then \( \tilde{W} \) is a vector field defined along \( x_0 \), and is \( C^\infty \) except possibly at the knots \( t_1, \ldots, t_{N-1} \). The track-sum \( \tilde{W} \) is said to be an *infinitesimally natural conditionally extremal* when it is \( C^2 \) and \( (H\tilde{W}^1)(t_0) = (H\tilde{W}^N)(t_N) = 0 \). Lemmas 5.2.2 and 5.2.3 lead to the following Lemma:

**Lemma 5.2.4.** Let \( \tilde{W} \) be an infinitesimally natural conditional extremal vector field. Then \( x_\tilde{W} \) is \( C^2 \) and, on taking \( x = x_\tilde{W} \), we have:

\[
\max\{ \| \nabla_t x^{(1)}(t_0) - \mathcal{A}(x(t_0)) \|, \| \nabla_t x^{(1)}(t_N) - \mathcal{A}(x(t_N)) \|, \\
\| \text{EL}_\mathcal{A}(x(t)) \| : t \neq t_i \} \leq C\|\tilde{W}\|_{3,\infty}^2
\]
where \( C = \max\{C_{\text{int}}, C_{\text{ext}}\} \).

The algorithm in Section 5.2 produces, with the exception of knots being conjugate along \( x_0 \), an infinitesimally natural conditional extremal vector field \( \tilde{W} \) satisfying \( E(x_0(t_i), \tilde{W}(t_i)) = \xi_i \). Moreover, if \( \{\tilde{B}_j^i\} \) is a basis for the infinitesimally natural conditional extremal vector fields along \( x_0 \), then we have a 1-1 linear relationship between the values \( \tilde{W}(t_i) \) and coefficients \( c_j^i \) where \( \tilde{W} = c_j^i \tilde{B}_j^i \). For a data set \( D \), define \( \mu(D) := \max\{d(x_0(t_i), \xi_i)\} \) and define \( D_\delta \) as the set of all data sets \( D \) satisfying \( \mu(D) < \delta \). Then because \( E \) is quasiexponential, the implicit function theorem states, for some \( \delta > 0 \), that there is a unique \( C^\infty \) assignment

\[
D \in D_\delta \mapsto c(D) := c_j^i
\]

with \( c((x_0(t_0), t_0), \ldots, (x_0(t_N), t_N)) = 0 \). By setting \( \tilde{W}(D) \) to be the associated infinitesimally natural conditional extremal, we have the following theorem:

**Theorem 5.2.5** (Noakes [27]). Define \( I_\delta : D_\delta \times [t_0, t_N] \to Q \) by \( I_\delta(D, t) := x_{\tilde{W}(D)}(t) \). For \( \delta > 0 \) sufficiently small, \( I_\delta \) is linear and

\[
\max\{\|\nabla_t x^{(1)}(t_0) - A(x(t_0))\|, \|\nabla_t x^{(1)}(t_N) - A(x(t_N))\|, \|E_{L_A}(x(t))\| : t \neq t_i\} \leq b\mu(D)^2
\]

where \( b > 0 \) depends on the choice of basis \( \{\tilde{B}_j^i\} \).

Informally, this result says that if \( \mu(D) < \delta \), then the interpolant this algorithm outputs will almost satisfy the equations of a natural conditional extremal spline for the data to \( O(\mu(D)^2) \) accuracy.

**5.3 Approximating natural Riemannian cubics in tension splines**

Recall that a Riemmanian cubic in tension with tension parameter \( \lambda \) is a curve \( x(t) \) minimising the functional

\[
J_2^\lambda(x) := \int_0^1 (\|\nabla_t x^{(1)}(t)\|^2 + \lambda \|x^{(1)}(t)\|^2)dt
\]
Chapter 5. Linear interpolation schemes on bi-invariant Lie groups

and that according to Theorem 3.1.2, \( x(t) \) is a RCT if and only if it satisfies

\[
\nabla_t^3 x^{(1)}(t) + R(\nabla_t x^{(1)}(t), x^{(1)}(t)) x^{(1)}(t) - \lambda \nabla_t x^{(1)}(t) = 0
\]

Proposition 3.1.5 says that the Lie reduction \( V(t) \) satisfies the equation:

\[
V^{(3)}(t) + [V(t), V^{(2)}(t)] - \lambda V^{(1)}(t) = 0
\]

A simple but known solution to this equation is \( V_0(t) = V_0 \) is constant. If our data points are sufficiently close to a geodesic, we may follow the same approach as in Section 5.2 to produce a linear interpolation scheme using Riemannian cubics in tension. If we take a variation \( x_h \) about \( x_0 \) and take the derivative with respect to \( h \) about \( h = 0 \), we obtain the equations:

\[
X^{(3)}(t) + [V_0, X^{(2)}(t)] - \lambda X^{(1)}(t) = 0 \quad (5.19)
\]
\[
\dot{\tilde{W}}^{(1)}(t) + [V_0, \tilde{W}(t)] = X(t) \quad (5.20)
\]

where as before, \( X(t) = \frac{dV(t)}{dh} \bigg|_{h=0} \) and \( \tilde{W}(t) = \mathcal{L}(\frac{dx_h}{dh} \bigg|_{h=0}) \). In this case, Equation (5.20) can be solved directly given we know \( X(t) \) using an integration factor and therefore only Equation (5.19) needs to be solved. On the other hand, in \( E^3 \), Equation (5.19) can be rewritten in the coordinates parallel to \( V_0 \) and perpendicular to \( V_0 \) where we have the equations

\[
X^{(3)}_{||}(t) - \lambda X^{(1)}_{||}(t) = 0 \quad (5.21)
\]
\[
X^{(3)}_{\perp}(t) + iX^{(2)}_{\perp}(t) - \lambda X^{(1)}_{\perp}(t) = 0 \quad (5.22)
\]

where \( X_{||}(t) := \text{proj}_{V_0} X(t), X_{\perp}(t) := X(t) - X_{||}(t), \) and \( i \) represents the imaginary unit obtained by identifying \( V_0^\perp \) with \( \mathbb{C} \). Both these equations form a decoupled homogeneous constant coefficient third order linear system of ODEs which can readily be solved exactly and therefore we can solve for \( \tilde{W}(t) \) up to a single quadrature.
As before, given an interpolation problem with data points $\xi_0, \ldots, \xi_N$ and knots $t_0, \ldots, t_N$, we can take a variation of an optimal interpolant $x_h$ to determine the associated interior point conditions as well as boundary conditions. Differentiating $J_2^\lambda(x_h)$ with respect to $h$ gets us (omitting any terms that appear in the Euler Lagrange equations):

$$\frac{1}{2} \frac{d}{dh} J_2^\lambda(x_h) \bigg|_{h=0} = \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \langle \nabla_h \nabla_t x^{(1)}, \nabla_t x^{(1)} \rangle + \lambda \langle \nabla_h x^{(1)}, x^{(1)} \rangle dt \bigg|_{h=0}
$$

$$\equiv \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \langle \nabla^2_t W, \nabla_t x^{(1)} \rangle + \lambda \langle \nabla_t W, x^{(1)} \rangle dt
$$

$$\equiv \sum_{i=1}^N \left( \langle \nabla_t W, \nabla_t x^{(1)} \rangle + \lambda \langle W, x^{(1)} \rangle \right) \bigg|_{t_{i-1}}^{t_i} - \int_{t_{i-1}}^{t_i} \langle \nabla_t W, \nabla^2_t x^{(1)} \rangle dt
$$

$$\equiv \sum_{i=1}^N \left( \langle \nabla_t W, \nabla_t x^{(1)} \rangle + \lambda \langle W, x^{(1)} \rangle - \langle W, \nabla_t^2 x^{(1)} \rangle \right) \bigg|_{t_{i-1}}^{t_i}
$$

Because $x_h$ is an interpolant, $W(t_i) = 0$ for all $i$ and so we are left with the condition that for all variations $W$, we have:

$$\sum_{i=1}^N \langle \nabla_t W, \nabla_t x^{(1)} \rangle \bigg|_{t_{i-1}}^{t_i} = 0 \quad (5.23)$$

Since $\nabla_t W$ can be arbitrarily chosen, this implies that $\nabla_t x^{(1)}(t_i^-) = \nabla_t x^{(1)}(t_i^+)$ for all interior $t_i$ and that $\nabla_t x^{(1)}(t_0) = \nabla_t x^{(1)}(t_N) = 0$. The same argument as in the previous section can be used to show that critical points of $J_2^\lambda$ are $C^2$ at the knots.

These results allow us to construct a $B$-spline basis for solutions of $\tilde{W}$ to solve the associated boundary value problem in just the same way as in Section 5.2. We note in passing that the same analysis of the approximately natural Riemannian cubic splines in [27] would apply to approximately natural Riemannian cubics in tension splines with very little change and so we do not repeat it.

---

1The symbol $\equiv$ is used to denote that terms have been omitted
5.4 Conclusion

We have seen in this chapter that when the data we wish to interpolate behaves close to an existing dynamical model, for a slight trade-off in optimality, we can avoid the difficulty of solving a coupled system of boundary value problems. By only solving essentially 4D initial value problems in a vector space and splicing them up, we can solve a small set of equations to produce a $B$-spline basis and reduce the problem into solving a linear tridiagonal system of equations which can be solved in $O(DN)$ running time. By imitating the analysis of Noakes [27], we are able to show that this method produces interpolants that almost satisfy the requirements of a natural conditional extremal spline. We then show how the same method can be applied to data modelled on Riemannian cubics in tension lying near a geodesic.
CHAPTER 6

Numerical Interpolation Schemes for Second Order Lagrangians

6.1 Introduction

Much of what we have studied in this thesis has been theoretical results pertaining to the study of interpolation in often restricted settings. In practice, we may not always be able to work in specialised settings such as in a Lie group with a bi-invariant metric having left-invariant prior vector fields. Although there is much to be gained from the study of various extrema of functionals from a pure-mathematical point of view, the underlying problem we are set on dealing with is that of interpolation. Computer power now greatly exceeds the capacity that existed when the theory of interpolation in manifolds began. It may indeed be possible to minimise the functionals of interest for all practical concerns without needing to go into the theoretical aspects of such a minimiser. This chapter of the thesis looks at discretisation of the time domain to produce applicable interpolation algorithms and compares them with others mentioned in this thesis. As a passing remark, in the theoretical study of splines, the problem of interpolation of \( N \) points is considered a significantly more difficult task than the calculation of a single spline element connecting two points. For the linearisation algorithm in Chapter 5, the algorithm complexity grows linearly with the number of knots. In the setting of numerical optimisation on the other hand, solving the \( N \) point interpolation problem actually becomes slightly simpler due to the knot constraints reducing the order of the system that is to be minimised.

Numerical optimization broadly fits into two categories, direct search methods and gradient based methods. Direct search methods have no real application in any large scale optimization problem and so we only consider gradient based methods. The two broad class of gradient based methods fall into variants of gradient descent and quasi-Newton methods. Coleman et al. [6] provide an excellent overview of large scale numerical optimization. When problems become sufficiently large, gradient descent tends to converge quite slowly to a solution.

[5] including when the problem is constrained. For the purposes of this thesis, Mathematica’s [50] FindMinimum function, although memory intensive, serves our goal of being able to find local minima quite quickly for moderate sized searches (100-200 variables). We would like to use these results to make comparisons to the algorithm in Chapter 5.

6.2 Approximations of derivatives

Before we consider discretising a Lagrangian, we should decide how to take discrete approximations to the various derivative terms that appear in second order Lagrangians.

In charts,
\[ x(t + h) = x(t) + hx^{(1)}(t) + O(h^2). \]  

Similarly,
\[ x(t + h) = \exp_{x(t)}(\log_{x(t)} x(t + h)) = x(t) + \log_{x(t)} x(t + h) + O(h^2) \]

since \( d\exp \big|_{v=0} = \text{id}. \) From this we see that
\[ x^{(1)}(t) = \frac{1}{h} \log_{x(t)} x(t + h) + O(h). \]

Traditionally in numerical differentiation, the derivative is calculated centrally where we have
\[ \frac{f(x + h) - f(x - h)}{2h} = f^{(1)}(x) + O(h^2) \]

In this vein, let us consider instead how well \( \frac{1}{2h} (\log_{x(t)} x(t + h) + \log_{x(t)} x(t - h)) \) approximates \( x^{(1)} \). For this, we will need normal coordinates. Normal coordinates [36] give us a coordinate neighbourhood of a point \( p \) so that all geodesics through \( p \) are straight lines where \( \exp_p V = p + V \) in the coordinate system. With \( h \) sufficiently small, \( x(t + h) \) and \( x(t - h) \) will belong inside a normal neighbourhood of \( x(t) \). In this coordinate system, the Christoffel symbols at \( x(t) \) vanish. We
then have

\[ x(t + h) = \exp_{x(t)}(\log_{x(t)} x(t + h)) = x(t) + \log_{x(t)} x(t + h) \]  \hspace{1cm} (6.5)

\[ x(t - h) = \exp_{x(t)}(\log_{x(t)} x(t - h)) = x(t) + \log_{x(t)} x(t - h) \]  \hspace{1cm} (6.6)

and therefore

\[ x(t + h) - x(t - h) = \log_{x(t)} x(t + h) - \log_{x(t)} x(t - h). \]  \hspace{1cm} (6.7)

We also know that

\[ x(t + h) - x(t - h) = (x(t) + h x^{(1)}(t) + \frac{h^2}{2} x^{(2)}(t) + \frac{h^3}{6} x^{(3)}(t) + O(h^4)) \]

\[ - (x(t) - h x^{(1)}(t) + \frac{h^2}{2} x^{(2)}(t) - \frac{h^3}{6} x^{(3)}(t) + O(h^4)) \]

\[ = 2 x^{(1)} h + O(h^3). \]  \hspace{1cm} (6.10)

Therefore

\[ x^{(1)}(t) = \frac{1}{2h} \left( \log_{x(t)} x(t + h) - \log_{x(t)} x(t - h) \right) + O(h^2) \]  \hspace{1cm} (6.11)

On the other hand we can see

\[ x(t + h) + x(t - h) = 2x(t) + \log_{x(t)} x(t + h) + \log_{x(t)} x(t - h) \]

\[ = 2x(t) + h^2 x^{(2)} + O(h^4). \]  \hspace{1cm} (6.13)

But since the Christoffel symbols vanish at \( x(t) \) in this neighbourhood, we know that \( x^{(2)}(t) = \nabla_t x^{(1)}. \) Therefore

\[ \nabla_t x^{(1)}(t) = \frac{1}{h^2} \left( \log_{x(t)} x(t - h) + \log_{x(t)} x(t + h) \right) + O(h^2) \]  \hspace{1cm} (6.14)

We now have ourselves a way of approximating the covariant derivative in a chart. One observes that all terms that are involved are in the tangent space.
at \( x(t) \). Being in \( T_{x(t)}Q \) means the addition is a well defined chart-invariant operation. Moreover, both \( \log \) and \( \nabla \) are chart-invariant and so this equation aptly serves as a geometrically defined approximation to the covariant derivative because the error term must therefore be chart-invariant too.

### 6.3 Unconstrained optimisation

Consider a general second order Lagrangian function that we hope to discretise and minimise

\[
J(x) = \int_0^1 L(x, x^{(1)}, \nabla_t x^{(1)}) dt \quad (6.15)
\]

where \( x^{(1)}(0), \nabla_t x^{(1)}(0), \nabla_t^2 x^{(1)}(0), \) and \( \nabla_t^3 x^{(1)}(0) \) are known.

We begin the discretisation by approximating these initial conditions the best we can be setting values for the initial knots \( \{0, 1/k, 2/k, 3/k\} \).

Define \( \pi_k : C^2([0, 1], Q) \to Q^{k+1} \) by \( \pi_k(x) = (x(0), x(h), x(2h), \ldots, x(1)) \) where \( h = \frac{1}{k} \). We wish to approximate the functional \( J \) by a function \( J_k : Q^{k+1} \to \mathbb{R} \) such that \( J_k(\pi_k x) \) converges to \( J(x) \) as \( k \) increases. Moreover, we should have the property that if \( x \) minimises \( J \) and \( x_k \) minimises \( J_k \), then

\[
\sum_{i=0}^k d((\pi_k x)_i, (x_k)_i)^2 = O(1/k) \quad (6.16)
\]

where \( d \) is the distance defined by the Riemannian metric. Informally, this second requirement for \( J_k \) says that minimisers of \( J_k \) should sufficiently approximate the minimiser of the continuous functional as \( k \) gets large. We will see that this isn’t necessarily implied by the first requirement of \( J_k \).

We can choose to define \( J_k \) by first defining a discrete Lagrangian \( L_k : Q^3 \to \mathbb{R} \) and then defining

\[
J_k(q_0, \ldots, q_k) := \sum_{i=1}^{k-1} \frac{1}{k} L_k(q_{i-1}, q_i, q_{i+1}). \quad (6.17)
\]

where \( q_0, \ldots, q_3 \) are sampled from the minimiser \( x(t) \).

The topic of variational integrators refers to numerical integrators of La-
grangian systems where a discrete Lagrangian serves as an approximation to a continuous system. Traditionally when trying to solve for trajectories of mechanical systems, one would compute the Euler-Lagrange equations and then apply numerical ODE solvers to compute solutions. With variational integrators, the Lagrangian is discretised instead and the necessary conditions for optimality lead to an integration scheme. Such solutions preserve conserved quantities known as momentum maps by a discrete version of Noether’s theorem. Marsden and West provide an overview of the subject of variational integrators [16] for first order Lagrangians (those of the form $L(x, x^{(1)})$). Petrehus [32] extends some of the fundamental results for variational integrators of first order Lagrangians to higher order Lagrangians. One particular result of importance is the following corollary of Proposition 4:

**Proposition 6.3.1** (Petrehus). For a non-degenerate Lagrangian $L$, let

$$|S^E_k(v_0, v_1, \Delta) - S_k(v_0, v_1, \Delta)| = O(\Delta^{r+2})$$

Then

$$\|F^E_\Delta(v_0, v_1) - F_\Delta(v_0, v_1)\| = O(\Delta^{r+1})$$

Moreover, the points of the discrete trajectory given by $F^E_\Delta$ correspond to the to the exact trajectory given by the Euler-Lagrange equations.

To understand this proposition for the case of second order Lagrangians, we must understand the meaning behind some of the terms. If a second order Lagrangian is non-degenerate as defined in [32], then for any fixed initial condition $v \in TQ$, there exists a number $\Delta_v > 0$ and a neighbourhood $U_v \subseteq TQ$ of $v$ such that for any $u \in U_v$, there is a unique solution to the Euler-Lagrange equations for which $x^{(1)}(0) = v$ and $x^{(1)}(\Delta_v) = u$.

The functional $S^E_k(v_0, v_1, \Delta)$ is the action of the exact discrete Lagrangian and
is calculated by

\[ S_k^E(v_0, v_1, \Delta) = \int_0^\Delta L(y(t), y^{(1)}(t), \nabla_1 y^{(1)}(t)) dt \]  \hspace{1cm} (6.18)\

where \( y(t) \) is the solution to the Euler-Lagrange equations for \( J \) with initial and final velocities \( v_0 \) and \( v_1 \) respectively. \( S_k \) on the other hand is an approximation to this quantity given by the function \( L_k \):

\[ S_k(v_0, v_1, \Delta) = \int_0^\Delta L_k(v_0, v_1, \Delta). \]  \hspace{1cm} (6.19)\

The term \( F^E_\Delta \) corresponds to the exact evolution operator of the Lagrangian and represents the solution to the continuous Euler-Lagrange equations. That is, \( F^E_\Delta(v_0, v_1) := (v_1, v_2) \) where \( v_2 \) is given by solving the Euler-Lagrange equations with initial condition \( v_2 \) for time step \( \Delta \). \( F_\Delta(v_0, v_1) := (v_1, v_2) \) on the other hand is calculated implicitly by the fact \( v_0, v_1 \) and \( v_2 \) should be a minimiser of \( S_k(v_0, v_1, \Delta) + S_k(v_1, v_2, \Delta) \).

With repeated applications of the evolution operators, the proposition implies that \( \| (F^E_\Delta)^k(v_0, v_1) - F^k_\Delta(v_0, v_1) \| = O(\Delta^r) \) where \( k = O(1/\Delta) \). Because \( F^E_\Delta \) corresponds to sampling the continuous solution of the Euler-Lagrange equations, \( F^k_\Delta(v_0, v_1) \) has error \( O(\Delta^r) \). On the other hand, by taking a sum over each interval in the discretisation, \( |S^E_k(v_0, v_1, \Delta) - S_k(v_0, v_1, \Delta)| \) being \( O(\Delta^{r+2}) \) implies that \( |J - J_k| = O(\Delta^{r+1}) = O(1/k^{r+1}) \). Therefore, we require an approximation of our discrete Lagrangian to order \( O(1/k) \) in order to achieve convergent approximations to the minimum functional value but we need \( O(1/k^2) \) accuracy in order to achieve convergent approximations to the extremals.

Consider the functional \( \mathcal{A}J_2(x) = \int_0^1 \| \nabla_1 x^{(1)} - \mathcal{A} \|^2 dt \)

Define the discrete Lagrangian by

\[ L_k(v_0, v_1, \Delta) := \frac{1}{\Delta^2} \left( \log x_\ast h + \log x_\ast x(0) \right) - \mathcal{A}(x_\ast) \]  \hspace{1cm} (6.20)\

where \( x \) is the unique conditional extremal satisfying \( x^{(1)}(0) = v_0 \) and \( x^{(1)}(\Delta) = v_1 \).
and $x_* = x(\Delta/2)$. If we choose normal coordinates about $x_*$, because $A$ is left-
invariant, $A(x_* + u) = A(x_*) + O(u^2)$ by the formula for parallel translation. This fact together with Equation (6.14) produces the bound $|S_k - S_k^E| < C\Delta^3$ for some $C > 0$. Normally we would need to factor $\|\nabla_t x^{(1)} - A\|$ into the calculation for the error but because $x$ is smooth on a compact domain and $A$ is constant norm, we can assume this quantity is bounded. Therefore by Proposition 6.3.1, we can see that $\|d(\pi_k x), x_k\| = O(1/k)$ and $|J(x) - J_k(x_k)| = O(1/k^2)$.

6.4 Constrained optimisation

The general interpolation problem we are interested in solving is to minimise

$$J(x) = \int_0^1 L(x, x^{(1)}, \nabla_t x^{(1)}) dt \quad (6.21)$$

subject to the condition that $x(t_i) = \xi_i$ for $i = 0, \ldots, N$ over curves $x$ which are $C^2$ in any region $(t_i, t_{i+1})$. This is the situation we will be most interested in as it pertains to interpolation but the methods of the previous section fail due to the constraints changing the path of integration. In the case of $A J_2(x)$, we have the relation that $|J(x) - J_k(x_k)| = O(1/k^2)$ by the reasoning above. We may use this result to draw comparisons between the $J$ of Chapter 5 and the $J_k$.

6.4.1 Case study: $A J_2$ on $S^3$ Although we are typically interested in $SO(3)$, we will shift our focus temporarily to interpolation in $S^3$. This is because $S^3$ is easy to understand geometrically as a manifold embedded in $\mathbb{R}^4$. Moreover, $S^3$ is a double cover of $SO(3)$ with a similar metric and therefore results are easily comparable. We will be interested in comparing the time taken to interpolate various data points as well as comparing plots of the resulting interpolants. We compute the log in $S^3$ by:

$$\log_{q_1} q_2 = -\arccos(q_1 \cdot q_2) \frac{q_2 - (q_1 \cdot q_2)q_1}{\|q_2 - (q_1 \cdot q_2)q_1\|} \quad (6.22)$$

We minimise the discrete Lagrangian
Example 6.4.1. We have a system of 6 knots given by $t_i = i/5$ and the data points in $S^3$ given by the rows of:

$$
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0.995735 & -0.0479317 & -0.0745203 & 0.0256949 \\
0.994743 & 0.00655651 & -0.0799695 & 0.0636315 \\
0.988917 & 0.0793479 & -0.0570451 & 0.111774 \\
0.956889 & 0.157192 & -0.04585 & 0.239899 \\
0.889612 & 0.237285 & -0.0355599 & 0.388614
\end{bmatrix}
$$

with $A(1, 0, 0, 0) = (0, 1, 0, 1)$.

<table>
<thead>
<tr>
<th>Number of data points</th>
<th>time (secs)</th>
<th>$J_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.234</td>
<td>7.40921</td>
</tr>
<tr>
<td>11</td>
<td>6.552</td>
<td>8.53452</td>
</tr>
<tr>
<td>21</td>
<td>33.134</td>
<td>8.98314</td>
</tr>
<tr>
<td>41</td>
<td>170.602</td>
<td>9.11598</td>
</tr>
<tr>
<td>61</td>
<td>442.606</td>
<td>9.14182</td>
</tr>
</tbody>
</table>

Table 6.1: Computations for discretisations of various step sizes. The linearisation algorithm of Chapter 5 produces a solution in 0.780 seconds and $A J_2$ of 12.9485.

Figure 6.1: Discretization with 6 points and 6 knots.
6.4. Constrained optimisation

Figure 6.2: Discretization with 11 points and 6 knots.

Figure 6.3: Discretization with 21 points and 6 knots.

Figure 6.4: Discretization with 41 points and 6 knots.

Figure 6.5: Discretization with 61 points and 6 knots.
6.5 Conclusion

The problem of discretisation leads to interesting theory such as those of variational integrators and can show that the flows under a discrete Lagrangian remain quite close to the actual flows. In the case of interpolation which is constrained optimisation, other methods need to be employed to determine how well behaved the critical points of $J_k$ are. Regardless of the situation, in both the constrained and unconstrained situation, $J$ can be estimated quite well using discretisation and we use it as a guide to rate the interpolation algorithm produced in Chapter 5. The $J_k$ values take a very long time to compute and they aren’t too different from the $J$ calculated with nearly conditional extremal interpolation and so the methods in Chapter 5 look quite powerful. It should be highlighted the methods
of this chapter work in any Riemannian manifold and does not require the additional structure we required in analysis such as having a Lie group structure, a bi-invariant metric or a left-invariant vector field.
In this thesis, we originally studied null Riemannian cubics in tension. These are special cases of Riemannian cubics in tension but are also generalised hyperbolic trigonometric functions. We firstly prove that these curves arise in a constrained optimisation setting where we minimise $J_2$ subject to bounded length. We took the Lie reduction of these curves and showed that the Lie reduction satisfied certain tight asymptotic formula for large $t$. Using results about duality, we were able to then produce asymptotics for the null Riemannian cubics themselves. It still remains to be seen how to more efficiently compute or approximate the constants $\alpha$ and $\beta$ from the initial conditions as this would provide a lot of benefit in more specifically understanding the asymptotics.

In Chapter 4, we looked at a way of designing the interpolation scheme to be based upon a dynamical model of the system being interpolated for various situations including first and second order, autonomous and non-autonomous. For non-autonomous first order systems, we produced $\mathcal{A}$-Jacobi field equations for these curves in a general Riemannian manifold and specialised the equations using the Lie reduction to produce exact solutions of an interesting form. We show that if we are to model a system based on a first order left-invariant prior field on a Lie group, then the best choice for vector field is the average of the initial and final velocities. Secondly we look at second order fields and their Euler Lagrange equations in both the manifold and Lie group setting. Often in mechanics the dynamics of the system depend on time and so that motivated the introduction of prior fields which vary with time. Some interesting results were obtained by looking at the simplest non-trivial case $\mathcal{A}_0 + t\mathcal{A}_1$. Perhaps the most fascinating result in this chapter was the result that if $z(t)$ is an integral curve such that $z^{(1)}(t) = z(t)A(t)$, then solutions to $V^{(1)} - A^{(1)} = [V(t), A(t)]$ can be found exactly as an exponential multiplied by $z(t)$. One interesting turn of events was when we studied equations of the form $V^{(1)}$. We also look at time varying
second order prior fields and found the interesting observation that when $A$ is left-invariant, the equations reduce to the Lie quadratic equation for Riemannian cubics. Lastly we find some conserved quantities for these curves and prove that they are infinitely extendible for all time. There are many possible directions this research can go in. One possible direction to go moving forward is studying how to relax the assumption of the second order prior vector field being left-invariant in performing calculations and to study simple examples of this.

In Chapter 5, we use linearisation to try strike a balance between optimal interpolation and efficiency. We take data that was modelled with a second order left-invariant prior field but we found the observations didn’t quite meet with the data. We first describe a method for computing $x_0(t)$ which is an affine cubic satisfying $V^{(1)} = A_1 t + A_0$ where $x_0(t_0) = \xi_0$ and $x_0(t_N) = \xi_N$. This was done through a process of solving a small set of initial value problems. We extend the methods of Noakes [27] by employing the technique of B-splines to form a tridiagonal system of equations to solve. We apply the ideas of analysis introduced by Noakes and show that the interpolants are quite good when the initial guess $x_0$ is good. We then show how the method may also be applied to Riemannian cubics in tension. It would be really interesting to extend this basic approach to a more general set of data points to interpolate rather than requiring that the data is close to an affine Riemannian cubic.

In Chapter 6, we look at the problem of minimising a discretisation and its relationship to an extremal of a continuous Lagrangian. We use results of higher order Lagrangian variational integrators to show that if we have an unconstrained minimisation problem, we can quite accurately solve the extremal by minimising the discrete Lagrangian. In the case of interpolation or a constrained problem, we are happy to accept that the $J$ values will be close to optimal and so the minimisation can be used to decide how well an interpolation algorithm approximates an optimal one, such as in Chapter 5. It would be really interesting to see a result concerning the accuracy of approximation using discrete minimisation for the constrained system. Moreover, I am thankful to Prof. Kobilarov who points out that
it would be interesting to apply these numerical optimisation methods to more
general cases such as when we have a more general Riemannian manifold, since
the difficulty in their analysis is removed. The numerical methods presented in
this chapter would indeed extend as they are defined by using operations available
in the more general setting.
Chapter 7. Conclusions
Bibliography


