Abstract

OSBORNE, JASON, MATTHEW. On Geometric Control Design for Holonomic and Nonholonomic Mechanical Systems. (Under the direction of Dr. Dmitry Zenkov.)

Geometric mechanics as a field of study does not hesitate to draw upon foundational geometry for formulation, re-formulation and inspiration. Herein, we take *geometry* to mean not only Riemannian and differential geometry but also fibre bundle theory. Broadly and simply stated, an overarching and unifying theme for this document is that:

> Viewing mechanical systems through a geometric lens opens up an extensive set of tools that can be brought to bear upon energy, mass, and system–conscious control design for constrained mechanical systems.

To demonstrate this thesis we consider the dynamics and control for several mechanical systems.

The moving mass Chaplygin sleigh, a rigid platform and moving mass system with attached blade imposing nonholonomic or velocity constraints, when viewed through a geometric lens presents a nontrivial momentum equation. Analysis of this momentum equation reveals natural (uncontrolled) motions of the sleigh system that play a central role in our control design to steer the sleigh to any point in the plane using a moving mass. The details for this control problem can be found in Chapter 2.

In Chapter 3 we develop a geodesic–based proportional–derivative (PD) control logic for tracking on a class of Riemannian manifolds. As a specific application of this general control logic, we consider the double gimbal system (envision a telescope), a mechanical system comprised of a base, an outer gimbal attached to the base through

a revolute joint, and an inner gimbal also attached to the outer gimbal through a revolute joint. A Riemannian structure for the double gimbal system is derived from its kinetic energy tensor which is itself constructed from the gimbal mass properties. At this point, the double gimbal system is now in the control domain of our geodesic–based PD control logic. The free and minimal energy motions of the double gimbal (or the double gimbal geodesics) are the *natural* tendencies of the double gimbal system that take into account its mass distribution. Since double gimbal geodesics are central to our PD control design, we work with, rather than against, the gimbals natural tendencies.

The foundational geometries most often chosen to begin modeling mechanical systems are the tangent and cotangent bundles to a configuration manifold Q, denoted TQ and $T^{\star}Q$ respectively. That both $T^{\star}Q$ and TQ are associated bundles to the frame bundle (the bundle of linear frames) of a configuration manifold Q, denoted LQ seems to, at least initially, indicate that it is also a natural geometric model of mechanical systems. That the frame bundle is also an appropriate geometric model within which to begin the study of mechanical systems is strengthened upon realizing that LQ carries with it a generalized symplectic (Hamiltonian) structure. In this generalized setting, kinetic energy dynamics are formulated on the frame bundle of a mechanical systems configuration manifold. By adapting the frame bundle dynamics to the constraint distribution (that is, by appropriate choice of moving frame), a portion of the (constrained) generalized momenta dynamics are an n-symplectic version of the nonholonomic momentum equation, see Chapter 4. These general dynamics have been carried out for the simple examples of the vertical rolling hoop and a nonholonomic constrained particle. Since the n-symplectic theory allows for the introduction of potentials which appear at the generalized momenta level, preliminary work along the n-symplectic line of thought indicates the possiblity of potential shaping and momenta based control design for nonholonomic mechanical systems.

ON GEOMETRIC CONTROL DESIGN FOR HOLONOMIC AND NONHOLONOMIC MECHANICAL SYSTEMS

BY

JASON M. OSBORNE

A DISSERTATION SUBMITTED TO THE GRADUATE FACULTY OF

NORTH CAROLINA STATE UNIVERSITY

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

MATHEMATICS

RALEIGH, NORTH CAROLINA

MAY 2007

APPROVED BY:

Dmitry Zenkov

D.V. ZENKOV

CHAIR OF ADVISORY COMMITTEE

Larry K. Norris

L.K. NORRIS

R.O. Fulp

R.O. FULP

S.L. Campbell

S.L. CAMPBELL

Biography

Jason Osborne tried to nominate his first person to write his biography, but he was politely turned down. Despite pleas as to the obvious awkwardness of relaying information as a third person that only the first person can know, his first person was steadfast in the matter.

Jason Osborne was born in Fairbanks, Alaska in 1977; his father Larry was part of an Army meteorological team stationed there to gather high altitude wind data. Despite several attempts by his mother Margie to explain the series of transplantations leading them (and his sister Stacey) to Tennessee, the spacetime gel never quite set for him...

> ...he recalls long bus rides to school and the delightful satisfaction of shooting BB's at tin cans from his grandparents porch in x_0 =Oklahoma...$\Delta t = t_1$...

....he remembers wading through creeks, flipping over rocks in search of crawdads and the oppressing humidity (which, to this day and forever after, he despises) in x_1 =Lousiana...$\Delta t = t_2$...

> ....he recalls long car rides at night accompanied by NPR's Prairie Home Companion and All Things Considered (which he hated then but would grow up to enjoy very much) and a deliciously cold snocone on the day the space shuttle landed in x_2 =Whitesands, New Mexico....$\Delta t = t_3$...

...he remembers marbles, digging and camouflaging holes (forts) for his GI Joes, jumping out of swings, the centrifugal force of the merry–go–round and trying not to fly off, garbage pail kids and his first taste of freedom (walking to school through the woods [despite his mother's admonitions]) in x_3 =Fort Meade, Maryland...$\Delta t = t_4$...

> ...he recalls Betty the bus driver (he would come to love riding the bus and would view it as a wonderful time to think and collect his thoughts) and more freedom in the form of skateboards, bicycles, basketball, tennis, wallball and football played with all the neighborhood kids until dark each day throughout their domain around Seminole drive in x_4 =Johnson City, Tennessee.

Jason graduated from Science Hill High School where his interest in science (specifically biology) was encouraged by Mrs. Bowman. Since his bus arrived to school 30 minutes before the start of the school day, she would arrive to the classroom to find him waiting to ask questions about the previous nights readings on photosynthesis or whatever. To her credit, and a quality that he would find present in each of those professors most influential on him, she would patiently address each question in a way that would encourage more thought on his part.

Jason graduated from East Tennessee State University with a B.S. in mathematics and physics. His original plans were to major in biology and possibly biochemistry, but organic chemistry lab and research conducted at the Quillen College of Medicine convinced him that he was more of a theorist than a laboratory technician. His interest in the mathematics of physics was piqued in an astrophysics class wherein Dr. Don Luttermoser discussed the relationship between quarks and mathematical (Lie group) symmetry. He found this relationship very interesting (and puzzling as, in all honesty, most of his physics classes were). As he had been doing for much of his time at ETSU, he went to pester Dr. Jeff Knisley with questions along the symmetry lines. Fortunately for Jason, Dr. Knisley did not view these questions as annoying, but rather as an opportunity to teach Jason about Lie algebras in physics. Jason credits Dr. Knisley for his decision to not only major in mathematics but also to pursue mathematical physics in graduate school.

Thirty years after his birth, Jason graduated from North Carolina State University with a Ph.D. in mathematics. This third person cannot come close to relaying the wonderful times Jason had at graduate school...the experiences were too broad...`##overflow##`...and the growth too exponential...`##overflow##`....to be captured by a linear net. Suffice it to say that the long days were filled with studying, teaching, basketball, movies, traveling, sculpting, painting, and studying.

Jason fully expects the future to be as equally rich, eventful, and satisfying.

Acknowledgments

Please do not read any intention into any ordering of names or statements below. Since there are a lot of people to thank and since all of my organizational energies went into the dissertation itself, I feel justified and liberated to write on the fly.

To say that I am extremely grateful to my many teachers and professors is a gross understatement. Thank you Mrs. Bowman, Mrs. Durham, Mr. Mauldin, Mrs. Moody; Dr. Jeff Knisley, Dr. Don Luttermoser; Dr. Arkady Kheyfets, Dr. Ron Fulp, Dr. Larry Norris, Dr. Dmitry Zenkov.

I would like to single out Drs. Knisley, Fulp, and Norris for their thought and care in laying out a lucid big–mathematical–picture–to–interesting–physical–example curriculum. Each class came to represent a small piece in a large mysterious puzzle that, once started, was impossible for me to put down. Being slow at puzzles, I am truly grateful to you all for your patience and support.

I want to thank all of my committee members for giving me their valuable time.

I want to thank my advisor, Dr. Dmitry Zenkov, for his course on mechanics that added extra dimensionality to the mystery puzzle. Exploring these corners has provided many interesting problems and opportunities. For his support at each stage of the whole dissertation process and for the career development opportunities he (through the NSF) supported, I cannot be more thankful. For the headaches I know I caused him as he kindly dragged me into the world of multi–tasking, I am sorry.

For many valuable discussions and the fruitful collaboration at AFRL, I thank Dr. Rob Fuentes and Dr. Greg Hicks. Thanks to the ear piece attachment to my cellphone for making the hours long conference calls and “that” never–ending article possible.

For all sorts of random help too numerous to relay, my deepest thanks go to Brenda Currin, Di Bucklad, Denise Seabrooks, Carolyn Gunton, Charlene Wallace, Seyma Bennett–Shabbir, Frankie Stephenson and Janice Gaddy.

To Ali, Brandon and Chirag thanks for the Cup–A–Joe–India–Garden nights where ideas were let fly.

To the houseoffive (Brian Adams in particular), for the college dorm and roommate experience I never had.

To every coffee shop with free wireless in which I have spent countless hours, thank you.

To Mitch’s limeade, turkey sandwiches, black bean chili, and Lisa, thank you so much for providing a cool (temperature) and unique environment in which to study for qualifying exams.

My deepest gratitude goes to Catherine Daubert for making me aware of the NC State Craft Center; had you not suggested that first class I really have no concept of where I would be today. Thanks to the NC State Craft Center for providing the excellent studio space for me to work in while I recovered my sense of balance after my qualifiers. I would like to express my gratitude to Jock Gault for the excellent figure sculpture classes wherein he let me find my own path (with gentle nudges every once in awhile). To Jim Fatata, I thank you for the use of your studio space and the many helpful pointers.

To Amme Klose, your unedited truthfulness, spontaneity and constant friendship brings to mind

> a vinyl blue
> folding chair in the sun
> waiting for me to do nothing but
> sit and talk and spit cherry seeds

Thank you for everything.

To Susan Mahar, wherever in the world you are now, thank you for more than I can say. In my mind you conjure the only poem (*if i* by ee cummings) that ever really struck me to the point of remembering

...i say to hell
with that i
say that doesn't matter) but
if somebody
or you are beautiful or
deep or generous what
i say is
whistle that
sing that yell that spell
that out big (bigger than cosmic rays war earthquakes famine...

To Karey Markovich, wherever in the world you are now, attempting to cook noodles on the dash of your car using the sun, experiencing the profound silence of Haleakala and driving along the bumpy north side of Maui are experiences I will never forget. Thank you for your hospitality, friendship and generosity.

To Paddy, Frank, Dr. Griggs and the various incarnations of (the original) LaceEmUp, I would not trade the memories of lunchtime basketball, one–on–one-on–one games of 21 and our two championships for anything, not even for a theorem with my name on it. Thank you all.

To Jill Reese, for the (mis)adventures we have already had and the many more we have yet to experience, I want to say that "You are my biker".

To my parents. My work ethic stems from your examples. Thank you for your trust, understanding and giving me free reign.

Table of Contents

List of Tables

List of Figures

Notation

M (or Q)	n-Dimensional (configuration) manifold	31
$(q^{\mathcal{J}})$	$\mathcal{J} = 1..\dim(Q) = n$ Coordinates on Q	31
T_qQ	Tangent space to Q at q	31
$\hat{\Omega}^{\mathfrak{J}}$	n-Symplectic vector valued–two form defined by $d\hat{\theta}^{\mathfrak{J}}$	74
${}^{\Delta}\hat{\theta}^{\mathfrak{J}}$	Distribution adapted soldering form	76
$\Delta^{\star}$	(n-m)-dimensional co-distribution on Q	86
Δ	m-dimensional distribution on Q defining system constraints	87
$\mathfrak{X}(M)$	Set of smooth vector fields on M	105
$\mathfrak{F}(M)$	Set of smooth real-valued functions on M	105
X, Y	Typical elements of TQ or $\mathfrak{X}$(Q)	105
$S = Q/\mathfrak{G}$	Symmetry reduced (shape) manifold	137
(r^i)	$i = 1..\dim(Q/\mathfrak{G})$ Coordinates on $Q/\mathfrak{G}$	137
$\mathfrak{G}$	Lie group	138
(s^A)	$A = 1\ldots\dim(\mathfrak{G})$ coordinates on $\mathfrak{G}$	138
LQ	Frame bundle of Q	139
$u = (q, \underline{f})$	Reference frame on LQ	140
$u = (q, \underline{e})$	Arbitrary element of LQ	140
$\Pi^{\mathfrak{J}}_{\mathcal{J}}$	Canonical coordinates on LQ	141
$u = (q, \underline{\partial})$	Canonical reference frame on LQ	141
$\overset{\Delta}{\Pi}{}^{\mathfrak{J}}_{\mathcal{J}}$	Distribution Adapted Coordinates on LQ	141
TQ	Tangent bundle of Q; vector bundle associated to LQ	143
$P(Q, \mathfrak{G}, F)$	Principal fibre bundle P over Q w.r.t Lie group $\mathfrak{G}$ with fiber F	145

Chapter 1

Introduction

1.1 Selective Historical Excerpts

A relatively brief venture into the history of late nineteenth and early twentieth century mathematics, specifically the development of Lie groups [17], brought clearly and to the forefront of my attention that many great mathematicians were inspired by the advances in the physics of their time. Indeed, Hermann Weyl wrote (1949, [17, pg. 421])

> For myself I can say that the wish to understand what really is the mathematical substance behind the formal apparatus of relativity theory led me to the study of representations and invariants of groups...

where (general) relativity theory is the theory of gravitation developed by Albert Einstein in 1916. According to ([17, pg. 423]) and the references cited therein, Einstein began the 1916 paper by acknowledging a debt to the mathematician Hermann Minkowski:

> The generalization of the relativity theory was facilitated through the form that Minkowski had given to the special relativity theory. He was the first mathematician to clearly perceive the formal equivalence of the

> space and time coordinates; this made possible the construction of the [general] theory.

The form of special relativity theory, as mentioned above by Einstein, with roots in James Clerk Maxwell's electromagnetic theory of light, is based upon the equal treatment of both space and time. In Minkowski's own words (1907, [17, pg. 339])

> Out of the electromagnetic theory of light there recently seems to have come a complete transformation of our representations of space and time which must be of exceptional interest to the mathematician. The mathematician is also especially well prepared to pick up the new views because it is a question of acclimatization to conceptual schemes with which he has long been familiar. The physicist meanwhile must rediscover these concepts and must painfully cut way through a primeval forest of obscurities.

Embedded in the above quote is the call for mathematicians to look to physics for interesting problems that they are, by their training, well equipped to solve. The necessity of such a statement and why Minkowski would jestingly need to apologize in a letter to David Hilbert for being

> . . . thoroughly infected with physics . . .

gains meaning through an understanding of the mathematical atmosphere of the late 19^{th} century. As stated by the author in ([17, pg. 334]):

> . . . by 1890 mathematics had been professionalized to the point where university mathematicians rarely ventured outside the ever expanding domain of pure mathematics. By then, most mathematicians were no longer active participants in the science of their day as Euler, Lagrange and Gauss had been in their time—Poincare' being a notable exception to the general trend.

That Newtonian (foundational) mechanics was the science of the time for Euler and Lagrange no doubt influenced Hilbert's feelings that it would be

> very desirable that the discussion of the foundations of mechanics be taken up by mathematicians also. (1900, [17, pg. 334])

Given the advice of this great mathematician, one can feel the sense of bewilderment in the anonymous quote from the preface of [24] published in 1992!

> Many of the greatest mathematicians—Euler, Gauss, Lagrange, Riemann, Poincare', Hilbert, Birkoff, Atiyah, Arnold, Smale— were will versed in mechanics and many of the greatest advances in mathematics use ideas from mechanics in a fundamental way. Why is it no longer taught as a basic subject to mathematicians?

On a personal note, it has been my experience that the clear take home message from those historically minded individuals to those not-so historically interested is ubiquitously "learn your history so you are not doomed to repeat it". Regarding the mathematical study of mechanics, I have at least begun to learn my history for the following document reflects my efforts in the study of *geometric mechanics*.

In the next section we summarize more recent literature which play a central role to our contributions in geometric mechanics, specifically control design of holonomic and nonholonomic mechanical systems.

1.2 Literature Background and Statement of Contributions

This document largely addresses geometric control design for holonomic and nonholonomic mechanical systems and is centrally based upon the following foundational framework:

A: Symmetry reduction of Lagrangian mechanical systems with nonholonomic constraints as developed in BKMM [3], B: PD (proportional and derivative) feedback control on manifolds as introduced in Bullo et. al. [9] and further developed in Bullo [6],

C: n-Symplectic geometry (generalized cotangent bundle geometry) and n-symplectic dynamics (generalized Hamiltonian dynamics) as developed in [28].

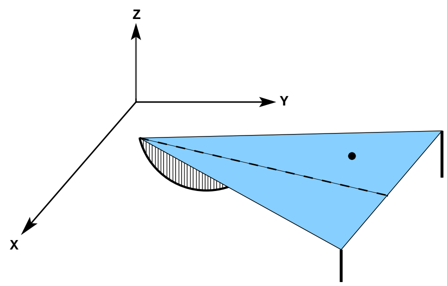
A zeroth order statement of this documents contributions are as follows:

A′: We show that the Chaplygin sleigh, a rigid platform with attached blade (imposing nonholonomic constraints, see Figure 1.1) can be steered using a moving mass, see Chapter 2 and Osborne and Zenkov [31]. The dynamics of the sleigh are a special case of those derived in [3].

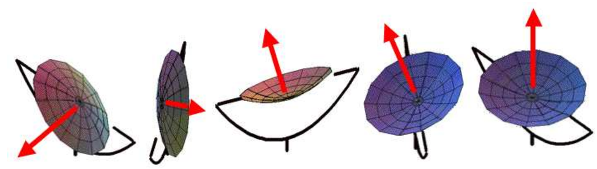
B′: Using geometric constructs (geodesic distance and parallel transport), we specialize the PD control theorem of [6] to a class of Riemannian manifolds, see Chapter 3 and Fuentes, Hicks and Osborne [14] and [15]. We specifically apply our controller to the double gimbal system, a coupled system of rigid bodies (see Figure 1.1).

C′: As an alternative to the tangent bundle (Lagrangian) and cotangent bundle (Hamiltonian or symplectic) formulations of mechanics, we demonstrate the frame bundle (n-symplectic/moving frame) approach to nonholonomic mechanical systems with symmetry. An n-symplectic equivalent of the nonholonomic momentum equation of [3] is obtained and the feasibility of momenta based control design is established.

The above foundational works are summarized in the following subsections for their contributions to the big picture in order that a more detailed statement of contributions may be given. Details of the foundational works will be given in the following chapters and in corresponding appendices.



(a)



(b)

Figure 1.1: Chaplygin Sleigh and Double Gimbal System

1.2.1 Short Summary of BKMM [3]

The dynamics of nonholonomic mechanical systems with symmetry are obtained in [3] using a constrained form of the *Lagrange–d'Alembert principle*, i.e. where both velocities and the variations are constrained. Utilizing also Lagrangian reduction techniques the constrained, reduced dynamics become the so called *(nonholonomic) momentum equation* which can be expressed as

$$\dot{p} = \langle \alpha(r)p, p \rangle + \langle \beta(r)p, \dot{r} \rangle + \langle \gamma(r)\dot{r}, \dot{r} \rangle \tag{1.1}$$

and the *shape equation* which can be expressed as

$$\ddot{r} = f(r, \dot{r}, p). \tag{1.2}$$

The shape equation is of Euler-Lagrange type on the reduced space with extra forcing terms. Together with the *reconstruction equation* or group dynamics equation expressed as

$$\dot{g} = g(J(r)p - A(r)\dot{r}), \tag{1.3}$$

these three equations are the full system of equations for nonholonomic mechanical systems with symmetry. See [3] for a derivation and the coordinate form of these dynamics or [16] for a coordinate free (bundle and connection oriented) approach.

An important feature of the dynamics in the above form is that while the shape and momentum dynamics are *coupled*, they are both *decoupled* from the group dynamics. The momentum equation characterizes a major difference between holonomic and nonholonomic systems. According to Noether's theorem, for a holonomic system with symmetry, the spatial components of momentum are conserved. On the contrary, for nonholonomic systems the momenta are, in general, dynamic variables with dynamics governed by a nontrivial momentum equation. The nonholonomic momentum equation and its *lack of conservation* are shown to be important in control, motion generation, and stability analysis of nonholonomic systems (see Bloch et al. [3], Zenkov et al. [41], and references therein).

That the nonholonomic momenta are *not*, in general, conserved and that the nonholonomic momenta dynamics are coupled to the shape dynamics allows us to restate our first contribution more precisely:

> Contribution A' restated: We design a steering control algorithm for the Chaplygin sleigh using a moving mass. The key to the design lies in certain control primitives that are discovered from an analysis of the momentum equation (2.2) for particular fixed placements of the mass. Assuming fully actuated control over the moving mass (the shape of the system) and utilizing the coupling between shape dynamics and nonholonomic momentum equation we are able to asymptotically steer the sleigh to any desired direction of motion and to any point in the plane.

The details of this contribution along with further supporting references are given in Chapter 2. The research for this chapter was supported by NSF grant DMS-0306017 and was conducted under the guidance of Dr. Dmitry Zenkov. Much of this chapter originally appeared in [31].

1.2.2 Short Summary of Bullo et. al. [9] and Bullo and Murray [6]

By way of two simple and yet illustrative examples we can gain a partial footing into the PD control results of Bullo and thusly to our specialized geometric PD control result.

Simply stated, to render the dynamics of a system in PD (proportional plus derivative) form is to add control forces to the systems free dynamics for which the controlled dynamics take the form of a damped, linear spring. For a point particle of mass m moving in 1D, the free/uncontrolled dynamics are $m\ddot{q} = 0$. Introduction of the control forces 1) the negative gradient of a positive definite quadratic potential $\frac{1}{2}kq^2$ about $q = 0$; the *proportional control forces* and 2) the damping constant times the velocity $c\dot{q}$; the *derivative control forces* renders the free point particle dynamics of the form $m\ddot{q} + c\dot{q} + kq = 0$. Adjusting or tuning the control parameters k, c, one can make the particles position asymptotically approach $q = 0$ with zero velocity $\dot{q} = 0$.

Extending the 1D damped spring example to mechanical systems with configuration manifold Q is the over arching idea behind [9] and [6]. A mechanical system of particular interest to us is the double gimbal system (see Figure 1.1). Simply stated, the double gimbal system is comprised of a base, an outer gimbal attached to the base through a revolute joint, and an inner gimbal attached to the outer gimbal also through a revolute joint. Since this system has two circular degrees of freedom, the spinning angles of both the inner and outer gimbals, the configuration manifold is the torus $T^2 = S^1 \times S^1$. Equipping the torus with the kinetic energy metric defined by the mass properties of the inner and outer gimbals, we obtain a Riemannian manifold that we call the *double gimbal torus*. Control on the double gimbal torus takes two flavors (a) *fixed point tracking* which corresponds to re-orientating a telescope pointed at fixed Star X to fixed Star Y and (b) *full tracking* which corresponds to a telescope tracking a comet.

Intuitively, the control forces of a *geodesic based fixed point tracking controller* on the double gimbal torus can be characterized as: (1) a force tangent to a minimal energy curve (a geodesic) connecting the points identified with Star X and Star Y and (2) a dissipation force opposing motion along a geodesic segment which will eventually stop the telescope when pointed at Star Y.

In Bullo et al. [9] the geodesic based PD *fixed point tracking* controller intuitively described above was introduced and implemented on the two-sphere. In the conclusions of Bullo et. al., it is stated that design of Lyapunov functions for tracking on not only the sphere and Lie groups but also a general Riemannian manifold are problems of future interest. Indeed, in Bullo and Murray [6] a general PD feedback control law was formulated for *tracking* on Riemmanian manifolds. The general feedback forces devised are analogous to those of the damped simple harmonic 1D oscillator example.

It should be emphasized at this point that to actually implement the general controller of Bullo and Murray on specific examples requires the explicit formulation of the feedback forces, which themselves are constructed from explicit formulations of a compatible pair of objects called i) the configuration error and ii) the transport map. Utilizing canonical cross product Lie algebra structure of $\mathbb{R}^3 \cong \mathrm{so}(3)$, the configuration and velocity errors are explicitly formulated in [6] for tracking on the two-sphere embedded in $\mathbb{R}^3$.

Contribution B' restated:

As a geometric, specifically geodesic, framework of [9] seemed to be the original impetus for [6] it would be natural and thus highly desirable, to return to a geodesic setting for tracking on Riemannian manifolds. Indeed, we achieved this geometric generalization in [14]. That is, we specialized the results of [6] to a class of Riemannian manifolds (one of which is the double gimbal torus) by explicitly using a quadratic geodesic based configuration error and compatible parallel transport map pair. Use of this pair, however, presented new problems. Namely, one need now find an appropriate region of, for example, the double gimbal torus for which any two points can be connected by a unique minimal geodesic. Using ideas from Riemannian and differential geometry ideas along with the software `Loki` [39], we are able to obtain estimates for the stability region on the double gimbal torus. With this knowledge we are able to confidently implement the controller.

The details of this contribution along with further supporting references are given in Chapter 3. This chapter was joint work with Dr. Robert Fuentes (then at, Boeing-SVS) and Dr. Gregory Hicks (then at, General Dynamics) and originally appeared in [14] and [15]. This authors work on the project was supported by the Air Force Research Labs (AFRL) Space Scholars program. During this time the author was also supported by NSF grant DMS-0306017 (PI: Dr. D. Zenkov).

1.2.3 Short Summary of LKN [28]

The geometries most often chosen to begin modeling mechanical systems are the tangent and cotangent bundles to a configuration manifold Q, denoted TQ and $T^\star Q$ respectively. For example, to obtain dynamics on $T^\star Q$ one needs only a geometric object, a real-valued one form, ϑ called the *canonical form* and a physical object, a real-valued function, f called an *observable or Hamiltonian function.* Specifically, given the canonical real-valued one-form, ϑ on $T^\star M$ and corresponding canonical non-degenerate real-valued two form $\Omega = d\vartheta$ on $T^\star M$ given in canonical coordinates (q^i, p^j) by

$$\Omega = dp_i \wedge dq^i \quad (\vartheta = p_i dq^i) \tag{1.4}$$

one can use the structure equation

$$df = -X_f \lrcorner\ \Omega \tag{1.5}$$

to uniquely determine a global Hamiltonian vector field X_f given a globally defined Hamiltonian function $f : T^\star M \to \mathbb{R}$. The integral curve equations of the global Hamiltonian vector field are exactly Hamilton's equations.

That both $T^\star Q$ and TQ are associated bundles to the frame bundle (the bundle of linear frames) of a configuration manifold Q, denoted LQ seems to, at least initially, indicate that LQ is also a natural geometric model of mechanical systems (see Appendix B for LQ details and Appendix C for associated bundle details). The appropriateness of LQ to the study of mechanical systems is strengthened by the discovery of the generalized symplectic (Hamiltonian) structure on LQ in [28]. Using the n-symplectic structure of LQ, this reference establishes the generalized momenta dynamics for arbitrary rank observables.

For example, following [28], one obtains dynamics of mechanical systems on the frame bundle by first lifting a rank-2 observable, the kinetic energy metric g, a (0,2) tensor field on Q, to an $\mathbb{R}^n \otimes \mathbb{R}^n$-valued tensorial function $\hat{g}$ on LQ and then solving the *n-symplectic structure equation*, a generalization of the structure equation on $T^\star Q$,

$$d\hat{g}^{ij} = -2\hat{X}_{\hat{g}}^{(i} \lrcorner \; \hat{\Omega}^{j)} \tag{1.6}$$

for n-vector fields $[\hat{X}_{\hat{g}}^i]$. Here $\hat{\Omega} = d\hat{\theta}$ where $\hat{\theta}$ is an $\mathbb{R}^n$-valued generalization of the canonical form θ on $T^\star Q$ called the *soldering form.* A critical difference in the n-symplectic versus the symplectic setting is that solutions $[X_{\hat{g}}^i]$ are equivalence classes of vector fields rather than a unique vector field. Equivalence here is characterized by an *n-symplectic gauge freedom* in the choice of a torsion free connection which, when specified, defines a unique solution vector field.

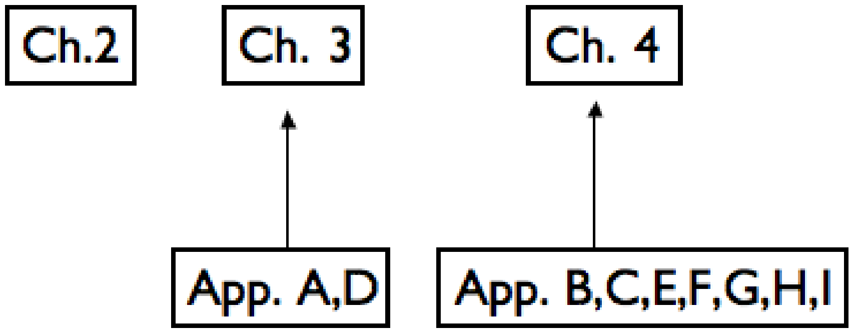
It is of general interest to the author to establish the applicability of n-symplectic geometry to mechanical systems and control. To the author's knowledge, no one has considered this approach. It is both interesting and encouraging to note that linear connections, the frame bundle and the distribution adapted frame bundle (used below) make an appearance in [21], a paper whose application of affine connections and distributions is to derive the nonholonomic momentum equation. To our reading and understanding, n-symplectic geometry played no role in the derivation.

The n-symplectic approach has been used to obtain results in field theory [26, 34, 35] but an analysis of the rank-2 observable (the kinetic energy) case with the control of mechanical systems perspective in mind has not been addressed.

> Contribution C' restated: For a distribution Δ modeling a mechanical systems nonholonomic constraints, we *adapt* the soldering form $\hat{\theta}$, the kinetic energy observable and the frame bundle to the distribution to obtain a distribution adapted structure equation
>
> $$d \overset{\vartriangle}{g}{}^{ij} = -2 \overset{\vartriangle}{X}{}^{(i}_{\hat{g}} \lrcorner \overset{\vartriangle}{\Omega}{}^{j)} \tag{1.7}$$
>
> specific solutions to which lead to constrained momenta dynamics. Some of the momenta dynamics are the nonholonomic momentum equation of [3]. That the n-symplectic theory allows for the introduction of potentials which appear at the generalized momenta level indicate the feasibility of potential shaping and momenta based control design for nonholonomic mechanical systems. We illustrate the above n-symplectic approach using two examples, the vertical rolling hoop and a nonholonomic constrained particle.

The details of this contribution are given in Chapter 4. Preliminary results relating to the use of the *n-symplectic gauge freedom* and the *n-symplectic scalar potentials* in control are addressed in Appendix I. The preliminary research in this chapter was conducted under the guidance and support of Dr. Larry Norris.

1.3 On the Structure of this Document



There are many appendices for this document that offer extra details and relevant background material for the readers' convenience. For Chapter 3, which utilizes a significant amount of geometry in developing a PD control logic on Riemmanian manifolds, we have added

- Appendix A which collects a detailed account of the requisite Riemmanian geometry objects and

- Appendix D which gives a detailed account of a fibre bundle needed to correctly formulate a *configuration error function* from our PD control design.

For Chapter 4, which utilizes n-symplectic (frame bundle, LQ) geometry to formulate the dynamics for constrained mechanical systems, we have added several appendices.

- Appendix B defines the frame bundle and its fibre bundle structure.

- Appendix C defines not only the *soldering form* (a generalization of the canonical one form on the cotangent bundle) but addresses the relationship between tensor fields (on the base manifold, Q) and tensorial vector–valued and matrix–valued functions on LQ. The soldering form and the tensor field/tensorial function relationship are key in formulating kinetic energy dynamics on the frame bundle.

- Appendix G, addresses the canonical formulation of generalized Hamiltonian mechanics on the frame bundle. Appendix F defines the (constraint) distribution adapted frame bundle, $L_\Delta Q$. The canonical dynamics and $L_\Delta Q$ are key to formulating the *constrained n-symplectic dynamics*.

- We have also added Appendix H which summarizes the nonholonomic momentum equation of BKMM [3] which will be compared to the generalized momenta dynamics on $L_\Delta Q$.

- Appendix I addresses preliminary results pertaining to the use of the n-symplectic gauge freedom and scalar potentials as control inputs to mechanical systems.

Chapter 2

The Moving Mass Chaplygin Sleigh

In this chapter we design a steering control algorithm for the Chaplygin sleigh with a moving mass. Our strategy is to only use the controlled dynamics to initiate short-time transitions between the various uncontrolled modes of the system in order to achieve the desired direction of motion.

2.1 Introduction

The objective of this chapter is to use a moving mass to control the direction of motion of the Chaplygin sleigh—a rigid body on a horizontal plane constrained by a blade. The blade limits the velocity of the body-plane contact point to a direction fixed in the body. This constraint is *nonholonomic* as it imposes a velocity restriction on the system which is not derivable from a position constraint. Rand and Ramani [33] and Ruina [36] point out that blade constraints similar to this have been used to model an underwater missile with fins.

In recent years much work has been done in using geometric structures to both formulate the equations and to address aspects of control of constrained mechanical systems. We summarize some such works. For a more complete list of references on the dynamics and control of nonholonomic systems see [4].

In the seminal paper by Bloch, Krishnaprasad, Marsden, and Murray [3] (hereafter referred to as BKMM), nonholonomic mechanical systems with symmetry are studied. For a system with symmetry, it is natural to split the configuration variables into the *group variables* g which describe the overall position (attitude) of the system and the *shape variables* r which describe the positions of the system's components relative to each other. In the case of the Chaplygin sleigh with a moving mass, the variable g is the element of the group of Euclidean transformations of the 2-dimensional plane and the variable r is the position of the moving mass relative to the contact point of the body and the plane.

The dynamics of a nonholonomic system with symmetry are governed by the system of equations

$$\ddot{r} = f(r, \dot{r}, p) + u, \tag{2.1}$$

$$\dot{p} = \langle \alpha(r)p, p\rangle + \langle \beta(r)p, \dot{r}\rangle + \langle \gamma(r)\dot{r}, \dot{r}\rangle, \tag{2.2}$$

$$\dot{g} = g(J(r)p - A(r)\dot{r}), \tag{2.3}$$

where p is the *nonholonomic momentum*, which in general is no longer conserved, and u represents control forces. We emphasize that the controls appear only in the shape equation. Note that equations (2.1) and (2.2) decouple from the the group dynamics (2.3). See [3] for details and formulae that define the various coefficients in equations (2.1)–(2.3). Equations (2.1), (2.2), and (2.3) are referred to as the *shape equation*, *momentum equation*, and *reconstruction equation*, respectively.

Utilizing the perturbation methods of [19] to study equations (2.1)–(2.3), Ostrowski [32] determines relations between the cyclic control inputs $u(t)$, resultant *momentum generation*, and ultimately, motion. These integral relations were critical in designing momentum generating and steering algorithms.

In Lewis and Murray [23], a *symmetric product* is introduced and used to formulate sufficient conditions for various types of configuration controllability of *simple* mechanical control systems. The general equations analyzed are of geodesic type with extra external force and control input terms.

Simple mechanical control systems *with constraints* are treated in Lewis [22]. The symmetric product was shown to address configuration controllability questions in this constraint setting also.

In Bullo, Leonard, and Lewis [7], simple mechanical control systems on a Lie group G are investigated. Here the dynamics of interest are of the form

$$\dot{p} = \langle \alpha\, p, p \rangle + u^a F_a, \tag{2.4}$$

$$\dot{g} = g(Jp), \tag{2.5}$$

where u^a are the control inputs and F_a are the directions in which they act. Implementing the perturbation approach of [19] and using the symmetric product technique, Bullo, Leonard, and Lewis [7] design steering control algorithms for equations (2.5) and (2.4).

The physical system of particular interest to us is the Chaplygin sleigh with a fully actuated moving mass. The dynamics of this system are of the form (2.1)–(2.3) (details are given in Section 2.2).

Unlike Ostrowski, in this paper we are not concerned with motion generation. We assume that the Chaplygin sleigh is already in motion and concentrate on *steering* the system using the movable mass. Since the mass is fully actuated, we can *assign* its position relative to the sleigh as a function of time. Therefore, the dynamics reduce to equations (2.2) and (2.3), where the moving mass position r relative to the contact point is interpreted as the control parameter. We emphasize that the dependence of the right-hand sides of (2.2) and (2.3) on r is *inherently nonlinear*, and thus the control design of Bullo, Leonard, and Lewis [7] is not applicable.

The perturbation control techniques and associated algorithm design of Ostrowski [32] and Bullo, Leonard, and Lewis [7] mentioned above are extremely useful and have a wide range of application. However, it is our philosophy that the dynamics of the *uncontrolled* system, which are not explicitly addressed in any of the above references, should play a critical role in the design of control algorithms.

Our control philosophy can be outlined as follows: We first study the variety of trajectories of equations (2.2) and (2.3) in the *uncontrolled* setting (*i.e.,* constant r). We then use the *controlled* dynamics (*i.e.,* equations (2.2) and (2.3) with non-constant r) to switch between the various types of *uncontrolled* dynamics which then lead to the goal configuration. This approach proved to be useful in various situations (see, e.g., [5]). We emphasize that the transfer is very short in duration and hence the system remains uncontrolled for most of the steering procedure.

On a technical note, we assume that, except for the short time that the actuators must implement the change in shape configuration, they are at rest. That is, the actuators are engineered to maintain the constancy of r when inactive. For example, in the Chaplygin sleigh, we can view the mass as sliding on a rod where the friction between the rod and sliding mass, not the actuator, is applied to keep the mass fixed.

The exposition is organized as follows: In Section 2.2 we summarize the properties of the uncontrolled dynamics of the Chaplygin sleigh. In particular, we list all possible types of trajectories of the contact point of the sleigh on the plane. In Section 2.3 we study "control primitives" which implement the transitions between the uncontrolled trajectories of the sleigh. These control primitives, when applied in the proper order, result in the desired reorientation of the system. Simulations are presented in Section 2.4.

2.2 The Dynamics of the Chaplygin Sleigh

2.2.1 The Configuration Variables

The Chaplygin sleigh is a rigid body sliding on a horizontal plane. The body is supported at three points, two of which slide freely without friction while the third is a knife edge, a constraint that allows no motion orthogonal to this edge. This mechanical system was introduced and studied in 1911 by Chaplygin [11] (although the work was actually finished in 1906).

The configuration space of this system is the group of Euclidean motions of the two-dimensional plane which we parameterize with coordinates (θ, x, y). As Figure 2.1 indicates, θ and (x, y) are the angular orientation of the blade (shown as the bold segment in the Figure) and position of the contact point of the blade on the plane, respectively. We view the sleigh as a platform whose center of mass is at the contact point. The mass and moment of inertia of the platform relative to the contact point are M and I, respectively. There is also a point mass m positioned at (a, b) relative to the platform, see Figure 2.1. In the classical Chaplygin sleigh this mass is motionless relative to the platform; in Section 2.3 we will control its position in order to steer the sleigh on the plane.

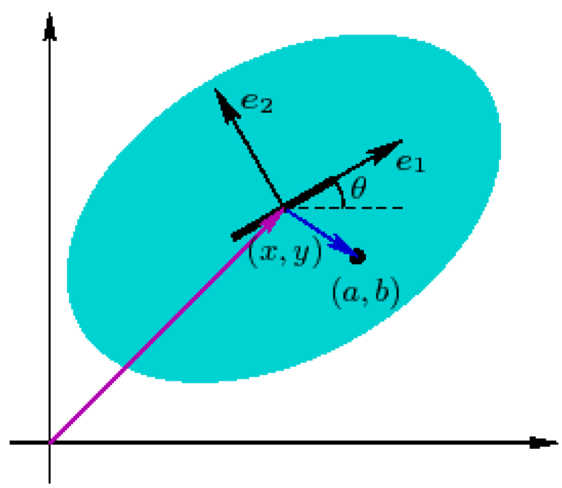


Figure 2.1: The Chaplygin sleigh (top view of Figure 1.1(a))

The constraint imposed by the blade reads

$$-\dot{x}\sin\theta + \dot{y}\sin\theta = 0. \tag{2.6}$$

This constraint is *nonholonomic*, whereby we mean it is not possible to derive the velocity constraint (2.6) from a position constraint $G(\theta, x, y) = 0$.

2.2.2 The Momentum Dynamics and Reconstruction

Let Ω^1 be the angular velocity of the platform and Ω^2, Ω^3 be the components of linear velocity of the contact point along and orthogonal to the blade, respectively. Constraint (2.6) implies $\Omega^3 = 0$.

Denote the nonholonomic momentum by (p_1, p_2). The components p_1 and p_2 satisfy the equations

$$\Omega^1 = \frac{(M+m)p_1 + mbp_2}{(M+m)(I+ma^2) + Mmb^2},$$
$$\Omega^2 = \frac{mbp_1 + (I+ma^2+mb^2)\,p_2}{(M+m)(I+ma^2) + Mmb^2},$$

see [3] and [42] for details and definitions. If $b = 0$, the components p_1 and p_2 equal the angular momentum of the sleigh relative to the contact point, and the projection of the linear momentum along the direction of the blade, respectively.

The dynamics of the Chaplygin sleigh is governed by the *momentum equations*

$$\dot{p}_1 = -ma\Omega^1\Omega^2, \qquad \dot{p}_2 = ma(\Omega^1)^2, \tag{2.7}$$

coupled with the *reconstruction equations*

$$\dot{\theta} = \Omega^1, \qquad \dot{x} = \Omega^2 \cos\theta, \qquad \dot{y} = \Omega^2 \sin\theta, \tag{2.8}$$

(see, e.g., [42]). This representation of the equations of motion allows one to first solve (2.7) and then find the trajectory of the sleigh by integrating equations (2.8). For the details on the derivation of these equations see BKMM [3] and Zenkov and Bloch [42].

The dynamics of the Chaplygin sleigh depends drastically on the value of a. This dependence is critical in the design of our control algorithm in Section 2.3.

If $a = 0$, the momentum components p_1 and p_2 are preserved. Equations (2.8) then imply that the trajectory of the contact point is either a circle or a straight line.

In both cases the contact point is moving at a constant rate. The existence of circular trajectories is very important for our steering control algorithm.

If $a \neq 0$, the trajectories in the momentum plane are either equilibria situated on the line $(M+m)p_1 + mbp_2 = 0$, or elliptic arcs, as shown in Figure 2.2. Assuming $a > 0$, the equilibria located in the upper half plane are asymptotically stable (filled dots in Figure 2.2) whereas the equilibria in the lower half plane are unstable (empty dots). The elliptic arcs form heteroclinic connections between the pairs of equilibria. The trajectories of the contact point that correspond to the momentum equilibria are

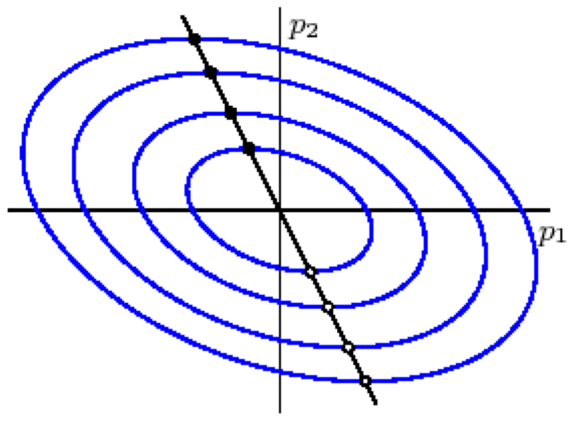
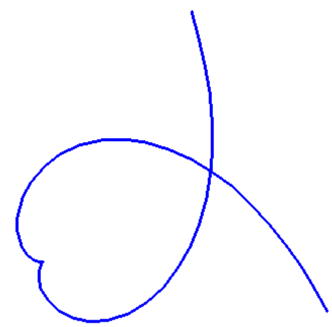


Figure 2.2: The momentum dynamics of the unbalanced sleigh.

straight lines in the xy-plane. They are stable if the mass m precedes the contact point and unstable otherwise.

The trajectories of the contact point reconstructed from the heteroclinic momentum trajectories should be regarded as the transfer solutions from an unstable straight line motion to a stable one. A typical transfer trajectory is shown in Figure 2.3. The shape of these transfer trajectories is predetermined by the inertia of the body and the position of the center of mass relative to the contact point, and is independent of the initial conditions. The angle between the asymptotic directions of a trajectory of the contact point in the xy-plane is evaluated in [18] for the case $b = 0$.[1]



(a)

Figure 2.3: A generic trajectory of the contact point.

2.3 Controllability of the Chaplygin Sleigh with a Moving Mass

2.3.1 The Reduced Controlled Dynamics

We now allow the point mass to change its position relative to the rigid body. That is, the quantities (a, b) are now dynamic variables. Assuming that the mass degrees of freedom are fully actuated, the system's dynamics are given by equations (2.2) and (2.3), where $r = (a, b)$ is viewed as the control parameter. Recall that the controller is active only when $\dot{a}^2 + \dot{b}^2 \neq 0$—see the discussion of the physical implementation of controllers in the Introduction.

In order to write equations (2.2) and (2.3) for the Chaplygin sleigh with a moving mass explicitly, let

$$\xi^1 = \frac{(M+m)(p_1 - ma\dot{b}) + mb(p_2 + M\dot{a})}{(M+m)(I+ma^2) + Mmb^2},$$

$$\xi^2 = \frac{m[b(p_1 - ma\dot{b}) - (I+ma^2)\dot{a}] + [I + m(a^2+b^2)]p_2}{(M+m)(I+ma^2) + Mmb^2},$$

and define η by

$$\frac{[Mmb^2 + I(M+m)]\dot{b} + a[(M+m)p_1 + mb(p_2 + M\dot{a})]}{(M+m)(I+ma^2) + Mmb^2}.$$

The momentum dynamics (2.2) for the Chaplygin sleigh with a moving mass is computed to be

$$\dot{p}_1 = -m\eta\xi^2, \qquad \dot{p}_2 = m\eta\xi^1. \tag{2.9}$$

Observe that for $(a, b) = \text{const}$, equations (2.9) reduce to (2.7).

After solving equations (2.9), the group configuration variables (θ, x, y) are obtained from the reconstruction equations

$$\dot{\theta} = \xi^1, \qquad \dot{x} = \xi^2 \cos\theta, \qquad \dot{y} = \xi^2 \sin\theta.$$

2.3.2 Controllability of Asymptotic Directions

Recall that if (p_1, p_2) is constant and $p_2 \neq 0$, there are three types of motions for the uncontrolled ($\dot{a}^2 + \dot{b}^2 = 0$) dynamics:

1. If $a \neq 0$ and $(M + m)p_1 + mbp_2 \neq 0$, then the system's trajectory is a curve that approaches straight-line motions as $t \to \pm\infty$ (see Figure 2.3).

2. If $a = 0$ and $(M + m)p_1 + mbp_2 \neq 0$, then the system moves along a circle in the xy-plane at a constant rate.

3. If $(M + m)p_1 + mbp_2 = 0$, the system moves along a straight line in the xy-plane at a constant speed.

We remark that the first type is *generic* (*i.e.,* observed with probability one when the initial conditions are randomly generated) whereas the second and the third are not.

The objective of this paper is: Assuming that the sleigh is sliding (that is, $\xi^2 \neq 0$), find the control inputs that put the system on a trajectory which asymptotically approaches a straight line with the desired direction in the xy-plane. In the theorems below we prove that it is possible to change the trajectory type by controlling parameters a and b. The existence of the desired steering control algorithm follows immediately from these theorems.

Theorem 2.3.1 Assume that the initial motion of the system is circular, i.e., $a = 0$, $b = \text{const}$, and $(M + m)p_1 + mbp_2 \neq 0$. Then there exist a continuously-differentiable function $a(t)$ and constants A, T_1, and T_2 with properties

- $a(t) = 0$ when $t \leq T_1$ and $a(t) = A$ when $t \geq T_2$,
- $a(t)$ is increasing when $T_1 < t < T_2$,

such that the trajectory of the system with $a = a(t)$, $b = \text{const}$ asymptotically approaches a straight line motion with a given direction in the xy-plane.

Without loss of generality assume that $\theta = 0$ at $t = 0$. Choose positive constants A and T and consider a continuously-differentiable function $f(t)$ such that $f(t) = 0$ for $t \leq 0$, $f(t) = A$ for $t \geq T$ and $f(t)$ is increasing on $0 < t < T$. The actual shape of $f(t)$ on the interval $0 < t < T$ is not important. For instance, one can define $f(t)$ as

$$F_{A,T}(t) = \begin{cases} 0 & \text{if} \quad t \leq 0 \\ \dfrac{2A}{T^2}\,t^2 & \text{if} \quad 0 < t \leq T/2 \\ A - \dfrac{2A}{T^2}\,(t-T)^2 & \text{if} \quad T/2 < t \leq T \\ A & \text{if} \quad t > T \end{cases}$$

(this will be our default choice).

Set $a = f(t)$. At the end of *transition interval* $0 < t < T$ the value of a becomes A. According to the classification of motions given above, the trajectory of the system for $t > T$ is either of type 1 or type 3. Let ϕ be the angle between the asymptotic direction of this trajectory as $t \to \infty$ (or the trajectory itself if it is a straight line) and the positive direction of the x-axis. Let ψ be the angle between the desired (asymptotic) direction of motion and the positive direction of the x-axis.

For the initial circular trajectory, let $T_1 \in \mathbb{R}$ be such that $\theta(T_1) = \psi - \phi$. Set $a(t)$ equal to $f(t - T_1)$. Then the trajectory of the system with $a = a(t)$ and $b = \text{const}$ satisfies the statement of the theorem. Indeed, this trajectory is obtained from the one corresponding to $a = f(t)$ by rotation about the center of the initial circular trajectory by the angle $\psi - \phi$. Therefore, the asymptotic direction of the trajectory forms the angle ψ with the positive direction of the x-axis.

Theorem 2.3.2 Assume that the system is moving along a trajectory of type 1, i.e., $a = A \neq 0$ and $(M+m)p_1 + mbp_2 \neq 0$.[2] Then there exist a continuously-differentiable function $a(t)$ and constants T_1 and T_2 with properties

- $a(t) = A$ when $t \leq T_1$ and $a(t) = 0$ when $t \geq T_2$,
- $a(t)$ is decreasing when $T_1 < t < T_2$,

such that the trajectory of the system with $a = a(t)$, $b = \text{const}$ becomes a circle for $t > T_2$.

Choose the values T_1, T_2 and set $T = T_2 - T_1$, $a(t) = A - f(t - T_1)$, where $f(t)$ is the function introduced in Theorem 2.3.1. Then at the end of the transition period the value of a equals 0. Therefore, the trajectory of the system for $t > T_2$ is either a circle, or a straight line. Adjusting the initial and terminal moments T_1 and T_2 of the transition period if necessary, it is possible to have $(M+m)p_1 + mbp_2 \neq 0$. Therefore, the trajectory becomes circular for $t > T_2$.

Theorem 2.3.3 Assume that the system is moving along a straight line, i.e., $b = B_1 = \text{const}$ and $(M+m)p_1 + mbp_2 = 0$. Assume that $a = A > 0$. Then there exist a continuously-differentiable function $b(t)$ and constants $B_2 \neq B_1$, T_1, and T_2 with properties

- $b(t) = B_1$ when $t \leq T_1$ and $b(t) = B_2$ when $t \geq T_2$,
- $b(t)$ is monotonic when $T_1 < t < T_2$,

such that the trajectory of the system with $a = A$, $b = b(t)$ becomes type 1 for $t > T_2$.

Define $b(t)$ by the formula $B_1 + F_{B_2-B_1,T_2-T_1}(t)$. By adjusting the values of T_1, T_2, and B_2, it is possible to satisfy the condition $(M+m)p_1(T_2) + mb(T_2)p_2(T_2) \neq 0$. Since the value of a has not changed, the trajectory of the system becomes type 1 for $t > T_2$.

Remark: The statement of the last theorem can be extended to the case of an initial straight line motion with $a = 0$. One just needs to change the value of a from 0 to A and then apply the algorithm of Theorem 2.3.3.

The reorientation algorithm can now be stated in the following steps:

1. Check if the trajectory of the sleigh is a straight line. If no, go to step 2. If yes, use Theorem 2.3.3 to transfer the sleigh to a generic trajectory and then go to step 2.

2. Check if the trajectory is circular. If yes, go to step 3. If no, use the control from Theorem 2.3.2 to transfer the sleigh to a circular trajectory and then go to step 3.

3. Using Theorem 2.3.1, exit the circular trajectory at an appropriate moment.

Remark: By Theorem 2.3.1, any point outside a circular trajectory in the plane belongs to an "exit" trajectory. It is now evident that the above three reorientation algorithm steps can be used to steer the Chaplygin sleigh through any point in the plane.

2.4 Simulations

In this section we illustrate the control primitives obtained in Theorems 2.3.1–2.3.3. We assume that the numerical values of the parameters of the system are $I = 10$, $M = 2$, and $m = 1$. In all simulations the initial value of b is set to 0.

Figure 2.4 illustrates the steering algorithm of Theorem 2.3.1. The value of a on the circular trajectory equals 0, and $f(t)$ is chosen to be $F_{1,4}(t)$.

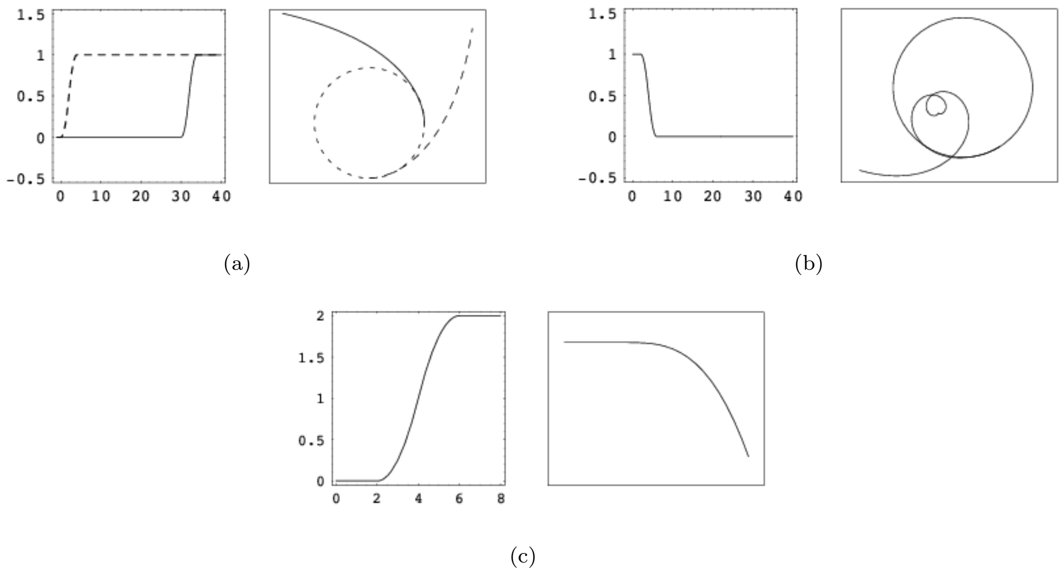


Figure 2.4: Control Primitives for Chaplygin Sleigh. (a) Transition form a circle to the trajectory with desired direction (b) Transition from a generic trajectory to a circle (c) Transition from a straight line to a generic trajectory.

If $a = f(t)$, the system's trajectory leaves the circle along the dashed curve in Figure 2.4 (a). The trajectory corresponding to $a = a(t) = f(t - 30)$ is the solid curve in Figure 2.4 (a).

Figure 2.4 (b) illustrates the control input that steers the system from a generic trajectory to a circular one. The initial value of a is 1, and $a(t)$ is set to $1 - F_{1,4}(t-2)$.

Figure 2.4 (c) illustrates the transfer from a straight line to a generic trajectory. The initial value of a is 0.1 and $b(t)$ equals $F_{2,4}(t)$.

2.5 Conclusions

In this chapter we have developed a dynamical system approach to controlling the asymptotic dynamics of the Chaplygin sleigh. The key feature of our algorithm is the use of the controlled dynamics only for switching to and from circular trajectories. As a consequence, the controller remains unpowered most of the time.

While our control algorithm design is problem specific (the uncontrolled dynamics change for each mechanical system chosen) as a philosophy it is a general principle. Whether it is applicable to a given situation depends of course on the nature of both the uncontrolled dynamics and controllers.

Chapter 3

Geodesic Based Proportional–Derivative (PD) Control on Riemannian Manifolds

In this chapter a geodesic-based formulation of geometric PD tracking control for fully actuated mechanical systems is derived from the Riemannian metric. The region of stability is determined directly from the size of the injectivity radius and, for a restricted set of control problems, the locus of cut points about a desired reference point in the manifold. Exponential stability is proven for controlled motion along a geodesic, yielding a particularly simple, yet elegant, methodology for control design. This chapter concludes with the geometry of the double gimbal system, a particular mechanical system on which the geodesic–based, PD control design can be implemented.

3.1 Introduction

The applicability of differential geometry and topology to the discipline of theoretical mechanics has been clear for some time [1]. Hence it is not a surprising development for one to conceive the notion that a more complete understanding of the *control* of mechanical systems, or simply, machines, may also reside within these frameworks. It so happens that many recent advances in control theory were obtained using the machinery of differential geometry and topology. Specifically regarding the theory of control design, it has clearly been demonstrated ([9, 10, 8]) that Riemannian geometry can be utilized to *naturally* approach a variety of control design problems and to obtain a proper perspective on commonly used control logic. Of particular interest to us is the notion of the geodesic spring introduced by Bullo, Murray, and Sarti in [9]. We expound upon this idea as it is the foundation of the control approach we shall take up herein.

Essentially, the philosophy of control system design is to create in the closed loop a well understood mechanical analogy. In more cases than not, the mass–spring–damper or the proportional derivative (PD) system is that to which the analogue system is compared. This is fitting, as its one of the few systems we understand quite well. The awkward part of this philosophy, in the context of mechanical systems, is the manner in which it is most often realized. By and large, typical control designs are brought to fruition and their subsequent analyses are carried out in $\mathbb{R}^n$. In some sense working in $\mathbb{R}^n$ is quite a natural thing to do and is easy enough to fall into; for it is typical to measure by some means the states of the physical system at hand, which then, either directly or indirectly, provide ordinates or n-tuples of real numbers from which a controls engineer elicits an input command. However, the configuration manifolds of ideal machines are typically not $\mathbb{R}^n$. As a result, the intrinsic nature of the machines one desires to understand and control is disregarded in typical control design.

In [9] a geodesic based, PD, set–point controller was introduced and "implemented" on the two-sphere. In the conclusions of this paper it was stated that the geometric design of controllers for tracking on, not only the sphere and Lie groups, but also a general Riemannian manifold, was a problem of future interest.

In [10], Bullo and Murray addressed this problem by introducing a fairly generic framework which extended the PD or generalized mass–spring–damper paradigm to the tracking problem on an arbitrary Riemannian manifold. This platform was hinged upon two mappings that generalized the notions of configuration error and velocity error. The latter object is constructed by means of a transport map that is "compatible" with the configuration error mapping. It should be emphasized that to actually specify the controller of [10] for specific machines requires the explicit presentation of both the configuration error and compatible transport mappings. Utilizing the algebraic structure of a Lie group, Bullo and Murray were able to construct a compatible pair and thus able to employ their controller on systems which evolve on SO(3) and SE(3).

Since the geodesic spring paradigm of [9] appears to have been the impetus for [10], it seems natural and highly desirable to return to an intrinsic geometric setting for tracking on Riemannian manifolds. Indeed, that is the purpose of this paper. That is, we are particularly interested in demonstrating herein, how a machine's intrinsic geometry (i.e. the geometry induced by the machine's kinetic energy tensor) can be exploited within the a generic framework similar to that of [10] in order to generalize [9]. In the intrinsic geometry setting there is an obvious choice of configuration error and transport map. Namely, they are the quadratic function of intrinsic distance and parallel transport along geodesic paths. However, control design based upon implicit geometry is technically difficult due to the non-trivial nature of Riemannian geometries. Nevertheless, we show that control is feasible on an "appropriate region" of the configuration manifold.

This paper is organized as follows. Sections 3.2 and 3.2.3 introduces those concepts from intrinsic geometry needed for the construction of a geodesic based controller. Should the reader already be familiar with the basic concepts of intrinsic geometry then they may better use their time by going directly to section 3.3 . The main objects, configuration error, compatible transport mapping and dissipation function, along with the main control result from Bullo et. al. are summarized in Section 3.3. In Section 3.4 we prove, as a corollary to the main control result from Section 3.3, a local exponential stability result for a specific class of control problems. Also in this section, it is explicitly argued that our control approach is a geodesic mass–spring–damper. In Section 3.5 we discuss our injectivity radius assumption and provide some examples.

3.2 Background Material

3.2.1 Some Conceptual and Notational Overhead

It is presumed that the reader has a knowledge of manifolds, bundles, and associated objects. We state here the notations associated to said notions that we shall use.

Q**:** A smooth connected manifold. Throughout this writing we shall make reference to pairs of points taken from Q. We will consistently use r and q to denote these points.

(U_α, ϕ_α)**:** An indexed coordinate chart compatible with the atlases comprising the differentiable structure of Q.

q_i**:** The ith coordinate function $x_i \circ \phi_\alpha : Q \to \mathbb{R}^n : q \mapsto (q_1(q), \ldots, q_n(q))$ associated to a chart (U_α, ϕ_α).

F**:** If f is a mapping between manifolds, the local representation thereof.

$f_x(y)$**:** If f is a function of $Q \times Q$ or any other product, then we use this convention to denote the restriction $f(x, y)|_{x\,\mathrm{fixed}}$. Similarly, $f_y(x) \triangleq f(x, y)|_{y\,\mathrm{fixed}}$.

T_qQ**:** The tangent space to Q at the point $q \in Q$.

v_q**:** An element of T_qQ.

T_q^*Q: The cotangent space to Q at the point $q \in Q$.

ω_q: An element of T_q^*Q.

(E, π, M): A fibre bundle with total space E, base space M, and projection π. We will often refer to the bundle by the total space.

TQ: The tangent bundle of Q.

v: An element of TQ.

q_i, v^j: The coordinate functions associated to a chart (U_α, ϕ_α) of TQ that reflects a local trivialization this bundle.

$(\vec{q}, \vec{v})$ If we desire to make use of coordinate vectors to describe the image of the q_i and v^k collectively, we will use this notation.

$\mathfrak{X}(Q)$: The set of smooth vector fields on Q. Typical representatives of this set will be denoted $X, Y, \ldots,$ etc. We use Y_r to indicate that the field Y has been evaluated at a particular point r and $Y(r)$ to denote a functional dependence upon a variable point r. The same conventions will be observed for tensor fields in general.

3.2.2 Basic Intrinsic Geometry

We begin with the purpose of providing context by introducing some of the basic elements of Riemannian geometry. Detailed expositions are abundant. Reference [13], for example, is decent and concise for the purpose of understanding this paper and with the exception of slight paraphrasing and notation, the definitions and results of this subsection and that to follow, are extracted therefrom.

Let Q denote a Riemannian manifold with metric tensor g. The metric structure is sufficient enough to describe a geometry on Q, for it emits the notion of lines and angles as well as length and angle measures. We will only briefly describe the lines of the geometry and the metric.

We focus first on the ruler. The restriction of a smooth curve $\sigma : \mathbb{R} \to Q$ to the compact interval $[a, b]$ shall be called a segment connecting points $r = \sigma(a)$ and $q = \sigma(b)$. Designating u as the curve "parameter" or coordinate on $\mathbb{R}$, the length of a segment is given by the line integral of its speed:

$$l^{\sigma}(r, q) = \int_a^b \sqrt{g\left(\frac{\mathrm{d}\sigma}{\mathrm{d}u}, \frac{\mathrm{d}\sigma}{\mathrm{d}u}\right)}\, du \ .$$

If $\mathcal{S}(r, q)$ denotes the set of all segments connecting r and q then the Riemannian manifold's intrinsic distance is defined as

$$\mathrm{dist}(r, q) = \inf_{\sigma \in\, \mathcal{S}(r,q)} l^{\sigma}(r, q) \ .$$

The intrinsic distance of a Riemannian manifold is a metric and its induced topology corresponds to that of the configuration manifold (a quite important fact indeed).

Now let us describe the lines of the geometry. The metric structure induces the Levi-Civita or Riemannian connection ∇, which itself effects a covariant derivative $\frac{D}{du}$ of vector fields on smooth curves. This covariant derivative then identifies the zero acceleration curves γ or the geodesics of the manifold (lines of the geometry). That is, $\frac{D\gamma'}{du} = 0$. We denote the parameterized geodesic with velocity v at $u = 0$ by $\gamma(u, v)$ and we simply use $\gamma(u)$ when the initial data is of little concern. Thus, components of the geodesic induced by a selection of chart satisfy the following second order system of ordinary differential equations (ODEs)

$$\frac{D}{du}\frac{d\gamma^k}{du} \quad = \quad \frac{d^2\gamma^k}{du^2} + \Gamma^k_{ij}\frac{d\gamma^i}{du}\frac{d\gamma^j}{du} = 0$$

where Γ^k_{ij} denote the Christoffel symbols and summation over repeated indices is implied. The classic existence and uniqueness theorems from the qualitative theory of ODEs ensures a flow box for this system.

3.2.3 The Exponential Map and Related Geometric Notions

Now we focus on the primary analytic tool associated to the Riemannian geometries, the *exponential mapping* $\exp\colon TQ \to Q$. This mapping allows us to gather the geodesic spray of points r in Q (all of the lines emanating from r) and is defined by

$$\exp(v) = \gamma(1, v)$$

for v sufficiently small.

For a fixed $r \in Q$, we also make use of the restriction $\exp_r\colon T_rQ \to Q$. The homogeneity property of geodesics ensures us that, for the given point r, the exponential mapping is defined on a sufficiently small neighborhood of $0 \in T_rQ$. Further, letting $B_\varepsilon(0)$ be a ball centered at the origin of $\mathbb{R}^n$ of radius ε, we have

Proposition 3.2.1 *Given $r \in Q$, $\exists \varepsilon > 0 \ni \exp_r\colon B_\varepsilon(0) \to Q$ is a diffeomorphism of $B_\varepsilon(0)$ onto an open subset of Q.*

If $\exp_r$ is a diffeomorphism of a neighborhood V of the origin in T_rQ, $\exp_r V = N$ is called a *normal neighborhood* of r in Q. If $B_\varepsilon(0)$ is such that $\overline{B_\varepsilon(0)} \subseteq V$, we call $B_\varepsilon(r) = \exp_r B_\varepsilon(0)$ the *normal ball* (or *geodesic ball*) with center r and radius ε.

Using this proposition, the exponential mapping allows us to create several special charts. By selecting an ordered orthonormal basis for T_rQ and inducing a coordinate system for this linear space w.r.t. this basis, we create by means of the local diffeomophism $\exp_r$ a system of coordinates on the image of the maximal normal neighborhood of r. Such a system of coordinates is called a normal coordinate system. Radial lines in this coordinate system get mapped onto the image (trace) of geodesics and $(l^\gamma)^2(r, q) = q_1^2 + \ldots + q_n^2$. We will make use of this special coordinate system. In the same manner, defining polar coordinates on tangent space we obtain the geodesic polar coordinate system in which the arclength actually plays the role of a coordinate.

Making use of these tools and other results the following proposition may be established:

Proposition 3.2.2 *Let $r \in N$, N a normal neighborhood of r, and $B \subseteq N$ a normal ball of center r. Let $\gamma : [0,1] \to B$ be a geodesic segment with $\gamma(0) = r$. If $\sigma : [0,1] \to Q$ is any piecewise differentiable curve joining $\gamma(0)$ to $\gamma(1)$, then $l^\gamma \leq l^\sigma$ and if equality holds then $\gamma([0,1]) = \sigma([0,1])$.*

Thus, geodesics are locally minimizing curves. That is, for q in a sufficiently small neighborhood, specifically, in a normal neighborhood of a point $r \in Q$,

$$\mathrm{dist}(r,q) = l^\gamma(r,q) \tag{3.1}$$

where γ is the unique (arc–length parameterized) geodesic connecting r to q. In fact, as we will illustrate at a later juncture, this relation is, under conditions to be discussed, smooth. *We shall utilize these facts extensively and equation (3.1) is the crux of the intrinsic view of the geodesic spring paradigm.* Consequently, we have need of knowing at what point this state of affairs ceases to persist.

Of course, this relation fails if a member fails to be defined. In an incomplete manifold, γ may not extend indefinitely and the spray of r may not cover Q. In this situation the right member becomes undefined for certain q and the relation must end. Thus, it may ease one's path to consider taking up the assumption that Q be geodesically complete. Under this assumption one may invoke the Hopf–Rinow theorem. This theorem asserts that the completeness assumption is equivalent to, amongst other things, the following:

1. $\exp_r$ is defined on all of T_rQ.
2. The closed and bounded sets of Q are compact.
3. Q is a complete metric space.

4. For any $q \in Q$ there exists a geodesic γ joining r to q with $\operatorname{dist}(r, q) = l^{\gamma}(r, q)$.

Equivalence 4 indicates to us that relation (3.1) becomes global once we alter the qualification "γ is the unique" to "γ is a". Though not a cure all, as $\operatorname{dist}(r, q)$ may not be a smooth everywhere, the situation is improved.

Should we have geodesic completeness and the exponential mapping be defined for all $v \in T_r Q$, then the smooth nature of $\operatorname{dist}(r, q)$, as we shall see, cannot be assured at those locations where the exponential mapping $\exp_r$ fails to be a diffeomorphism. This occurs at those points q such that $\exp_r$ or its tangent mapping become singular. As the proposition to follow asserts, singularity of the tangent mapping occurs at *conjugate points*, which are those points q at which nearby geodesics of a spray will *appear* to be destined to, and may, intersect.

Proposition 3.2.3 *Let $\gamma\colon\ [0, a] \to Q$ be a geodesic with $\gamma(0) = r$. The point $q = \gamma(u^*)$, $u^* \neq 0$ is* conjugate to r along γ *iff $v = u^* \frac{\mathrm{d}\gamma}{\mathrm{d}u}(0)$ is a critical point of exp_r.*

We will take up the assumption of geodesic completeness in order to ensure that these latter pitfalls are all we need worry with. **Working under this supposition we continue**.

Definition 3.2.1 *Let Q be a complete Riemannian manifold, and let $\gamma\colon\ [0, \infty) \to Q$ be a normalized geodesic with $\gamma(0) = r$. A* cut point of r along γ, *if it exists, is the last point $q = \gamma(s^*)$ along γ such that* $\operatorname{dist}(r, q) = s^*$. *The cut locus, denoted $C(r)$, is the union of cut points along all geodesics emanating from r.*

As our previous discussion suggests, the cut locus may be composed of two different types of points. The next proposition characterizes these points. Proof of this result can also be found in [13].

Proposition 3.2.4 *Let γ be defined as in definition 3.2.1 and suppose that $\gamma(s^*)$ is the cut point of $r = \gamma(0)$ along γ. Then*

1. *either $q = \gamma(s^*)$ is the first conjugate point of $\gamma(0) = r$*

2. *or there exists a geodesic* $\sigma \neq \gamma$ *from* r *to* $q = \gamma(s^*)$ *such that the arc lengths satisfy* $l^\sigma(r, q) = l^\gamma(r, q)$,

Conversely, if either 1 or 2 holds, then $\exists \tilde{s} \in (0, s^*]$ *such that* $\gamma(\tilde{s})$ *is the cut point of* r *along* γ.

Thus, at those points of $C(r)$ where the exponential mapping actually fails to be 1–1, two lines segments contained in the spray of r and of equal length intersect once more.

Apparently, the map $\exp_r$ along with its tangent mapping are injective on an open ball $B_\varepsilon(0)$ if the radius ε is less than the infimum of distances to $C(r)$. Hence the quantity

$$\mathrm{Inj}(r) = \inf_{q \in C(r)} \mathrm{dist}(r, q)$$

is called the *injectivity radius* of r. Defining

$$\mathrm{Inj}(Q) = \inf_r \mathrm{Inj}(r) \tag{3.2}$$

to be the *injectivity radius of the manifold* Q, we see that $\exp_r$ is a diffeomorphism on any open ball B_ε with $\varepsilon \leq \mathrm{Inj}(Q)$.

Of course, there is the degenerate case $\mathrm{Inj}(Q) = 0$. For purposes of formulating the control problem, it is important to know when the injectivity radius of a Riemannian manifold is non-degenerate. The answer to this question is always affirmative when the manifold in question is compact or the sectional curvature is non-positive [2, 13]. Interestingly enough, the radius can often times be estimated from knowledge of the sectional curvature and the shortest periodic geodesic. Section 3.5 highlights some results along this head and the applicability of these results to control design examples.

3.3 Dynamics and Control

Let us now describe the potential free (post compensated), fully actuated, simple, mechanical, control system or more simply, an *abstract machine*. An abstract machine is a mathematical construct of the form $[Q, g, \mathcal{F}]$, where

1. Q is a smooth manifold of dimension n called the configuration manifold,
2. g is a Riemannian metric on Q,
3. and $\mathcal{F}$ is a rank n control co-distribution.

Since an abstract machine comes equipped with a Riemannian manifold, it has an intrinsic geometry and carries with it all of the machinery described in section 3.2. In particular, the lines of the geometry indicate to us those motions of the machine that are locally the "easiest" (kinetic energy minimizing) to perform. The dynamics of the machine are given by the abstract form of the Lagrange-D'Alembert principle to be

$$g^\flat \frac{D\dot{\gamma}}{dt} = u_i \mathcal{F}^i \ . \tag{3.3}$$

where $g^\flat$ is that mapping which identifies vector fields with Pfaffians by means of the metric tensor g. Apparently, free motions are lines and we know only as much about the machine's tendencies as we do about the manifold's topology and geometry.

We now introduce those constructs from [10, 8] which are needed to describe the generalized PD controller, specifically 1) a tracking error function and 2) a transport mapping.

Definition 3.3.1 (Configuration or Tracking Error) *Let P be a connected submanifold of $Q \times Q$. A smooth function $\varphi\colon P \to \mathbb{R}$ is a* Tracking Error Function *if for each fixed $r \in Q$, $\varphi_r(q)$ is smooth, symmetric, proper, and bounded from below and φ satisfies*

1. $\varphi(r, r) = 0$
2. $d\varphi_r(q)|_{q=r} = 0$
3. $\operatorname{Hess} \varphi_r(q)|_{q=r}$ *is positive definite.*

The differential of the error function will serve us as geodesic "stiffness". That is, the tracking error function is a generalization of the quadratic Hookean potential.

In order to demonstrate, by the means provided us by Bullo and Murray, that we can obtain a generalized damped spring with the expected behavior, we shall also need the following definitions which are used specifically for proving local exponential stability.

Definition 3.3.2 *Let*

$$\begin{aligned} L_{\text{reg}}(\varphi_r(q), r) &\triangleq \sup\left\{L \in \mathbb{R} \mid q \in \varphi_r^{-1}([0, L]) \setminus \{r\} \Rightarrow d\varphi_r(q) \neq 0\right\} \\ L_{\text{reg}}(\varphi, Q) &\triangleq \inf\left\{L_{\text{reg}}(\varphi_r(q), r) \mid r \in Q\right\} \end{aligned}$$

A tracking error function φ is uniformly quadratic *if $L_{\text{reg}}(\varphi, Q) > 0$ and if, for all $L \in (0, L_{\text{reg}}(\varphi, Q))$ and $r \in Q$, there exist two constants $0 < a \leq b$ such that for all $q \in \varphi_r^{-1}(B_L(0)) \backslash \{r\}$, we have*

$$0 < a \; \|d\varphi_r(q)\|_g^2 \leq \varphi(r, q) \leq b \; \|d\varphi_r(q)\|_g^2$$

where for $\|\cdot\|_g$ denotes the induced or operator norm w.r.t. the norm associated to the inner product g.

Next we need to describe velocity error. This requires the concept of *transport map.* A *transport map* $\mathcal{T}$ is a smooth mapping on a connected submanifold P of $Q \times Q$ with point–wise behavior $(r, q) \mapsto GL(T_rQ, T_qQ)$ satisfying the property that $(r, r) \mapsto \text{id}$. In other words, a transport map is simply a construct through which one can compare velocities which reside in different fibers of the tangent bundle TQ. Thus, given a curve $(r(t), q(t)) \in P$ one may define a velocity error in T_qQ by

$$\dot{e} = \dot{q} - \mathcal{T}_{r \to q} \dot{r} \; .$$

Upon close inspection one may find the above description of the transport mapping to be a bit nebulous. To make this object concrete we recall bitensor fields (or alternatively, two-point tensor fields).

Remark 3.3.1 *In [40] (pg. 48) two-point tensors were introduced. As is common in the classic physics literature, this tensorial object was described in a coordinate context and was thought of as a multidimensional array which acts on like indexed objects. For example, in [40], we find* $T^{i'}_{j}$ *where* (D', C') *and* (D, C) *are the (domain, coordinate system) pair about two points* p' *and* p *of a manifold* Q*. The object* $T^{i'}_{j}$ *is viewed therein as a co-vector relative to coordinate transformations at* p *and a vector relative to coordinate transformations at* p'*. Viewed as such, each index could be covariantly differentiated according to the notation* $T^{i'}_{j;k}$ *or* $T^{i'}_{j;k'}$ *. Furthermore, transport of a vector* $v_p \in T_pQ$ *to a vector* $\bar{v}_{p'} \in T_{p'}Q$ *would have been written as* $\bar{v}^{i'} = T^{i'}_{j} v^{j}$*.*

Consider two finite dimensional vector spaces V and W of dimension n and m, respectively. The space of linear functionals on each of V and W, denoted V^* and W^*, can be used to generate the tensor product $V^* \otimes W^*$, the vector space of bilinear mappings on $V \times W$ of dimension nm. Selecting bases for V and W, the bilinear mappings of $V^* \otimes W^*$ can be represented as $n \times m$ matrices.

Now suppose that, for $p = (r, q) \in P \subseteq Q \times Q$, the vector spaces of concern are T_rQ and T^*_qQ. Then the mapping $T \in T^*_rQ \otimes T_qQ$ takes a vector at r, say v_r, and a covector at q, say ω_q, and maps them into $\mathbb{R}$. So, T is like a $(1, 1)$ tensor, but the arguments of the map work over spaces based at different points.

Take under examination the disjoint union

$$\mathrm{BT}^{0,1}_{1,0}P \triangleq \bigcup_{p \in P} T^*_rQ \otimes T_qQ.$$

$\mathrm{BT}^{0,1}_{1,0}P$ is a vector bundle over P. As such, we may talk about a section $\phi : P \to \mathrm{BT}^{0,1}_{1,0}P$ or a *bitensor field* (or alternatively, *two–point tensor field*). Such a field is *smooth* if the section $p \mapsto \phi_p$ is a smooth mapping. We note that given a point $p = (r, q) \in P$ and a vector v_r, or simply a pair $(v_r, q) \in TQ \times Q$, that $\phi_{(r,q)}(v_r, \cdot) = w_q$, a vector above q. So, it becomes clear that we may, as is commonly done with the typical $(1, 1)$ tensor, view the section ϕ as a mapping $\phi : p = (r, q) \mapsto gl(T_rQ, T_qQ)$, as previously discussed. i.e., a bitensor field provides a means of *transporting vectors*. Further, the notion of being smooth is made concrete in this context by insisting that $X(r, q) \triangleq \phi(r, q)(Y(r), \cdot)$ is a smooth section of the bundle $Q \times TQ$ whenever $Y \in \mathcal{X}(Q)$.

The above discussion motivates the following definition, which can be found to be completely equivalent to the vector bundle mapping approach to be found in [8].

Definition 3.3.3 (Transport Mapping) *A* Transport Mapping *is a smooth bitensor field $\mathcal{T}$ on a connected submanifold P of $Q \times Q$ satisfying the properties:*

1. $\mathcal{T}_{r\to q} \in GL(T_rQ, T_qQ)$.
2. $\mathcal{T}_{r\to r} = \mathrm{id}$.

where the notation $\mathcal{T}_{r\to q} \triangleq \mathcal{T}_{(r,q)}$ is used for clarity.

A given transport mapping is said to be *compatible with configuration error φ* if it satisfies the relation

$$d\varphi_q(r) = -\mathcal{T}^*_{r\to q}\left(d\varphi_r(q)\right) .$$

In order to understand conditions to be delineated in the statement of Bullo and Murray's theorem below, one needs a definition of the covariant derivative of a transport mapping, $\nabla\mathcal{T}$. As a bitensor, $\mathcal{T}(r, q)(v_r, \cdot)$, $v_r \in T_rQ$, is a smooth vector field over Q. As such, given a vector w_q, we can take the covariant derivative of this vector field, that is, $\nabla_{w_q}\mathcal{T}(r, q)(v_r, \cdot)$ makes sense. For understanding, it is helpful to consider the case of a parameterized curve $q(u)$. In this situation, the assignment $u \mapsto \mathcal{T}(r, q(u))(v_r, \cdot)$ describes a vector field, call it V, along the curve q, and $\frac{DV}{du} = \nabla_{q'}\mathcal{T}(r, q(u))(v_r, \cdot)$. So, we define $\nabla\mathcal{T} : TP \to TQ$ by the assignment $(v_r, w_q) \mapsto \nabla_{w_q}\mathcal{T}(r, q)(v_r, \cdot)$. As stated in [8], $\nabla\mathcal{T}$ is a vector bundle map.

Finally we have need of the dissipation function.

Definition 3.3.4 (Linear Rayleigh Dissipation Function (LRDF)) *A Linear Rayleigh Dissipation Function is a smooth, self-adjoint, positive-definite tensor field*

$$K_d(q) : T_qQ \to T_q^*Q \ .$$

We note immediately that $g^\flat$ is a LRDF. The LRDF along the velocity error allows us to abstract the dissipation term of the spring.

Finally, we conclude this section by stating the main theorem of [10]. It should be noted that in this paper, $P = Q \times Q$ was used as the domain of definition for the error function, etc.

Theorem 3.3.1 (Bullo and Murray) *Consider the control system in equation (3.3), and let $\{r(t), t \in \mathbb{R}_+\}$ be a twice differentiable reference trajectory. Let φ be an error function, $\mathcal{T}$ a transport map satisfying the compatibility condition and K_d be a dissipation function.*

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

$$\begin{aligned} F_{\mathrm{PD}} &= -d\varphi_r(q) - K_d\dot{e} \\ F_{\mathrm{FF}} &= g^\flat(q)\left((\nabla_{\dot{q}}\mathcal{T}_{r\to q})\dot{r} + \frac{d}{dt}\Big|_{q\ fixed}(\mathcal{T}_{r\to q}\dot{r})\right) \end{aligned}$$

then the curve $q(t) = r(t)$ is stable with Liapunov function

$$W_{total}(q, \dot{q}, r, \dot{r}) = \varphi(r, q) + \frac{1}{2}g(\dot{e}, \dot{e}) \ .$$

In addition, if the error function φ satisfies the quadratic assumption with a constant

L, and if the boundedness assumptions

$$d_1 \leq \inf_{q \in Q} \|K_d(q)\|_g \leq \sup_{q \in Q} \|K_d(q)\|_g \leq d_2$$

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

$$\sup_{(r,q) \in Q \times Q} \|\nabla d\varphi_r(q)\|_g < \infty$$

$$\sup_{t \in \mathbb{R}} \|\dot{r}\|_g < \infty$$

where $0 < d_1 \leq d_2$*, hold, then the curve* $q(t) = r(t)$ *is exponentially stable with Liapunov function* W_{total} *from all initial conditions* $(q(0), \dot{q}(0))$ *such that*

$$W_{total}(0) < L \ .$$

3.4 Intrinsic Tracking Control

Theorem 3.3.1 of Bullo and Murray does not tell one how to actually select the configuration error φ and a compatible transport mapping $\mathcal{T}$. As discussed in the introduction, if the configuration manifold is a Lie group, there are candidates which present themselves as obvious on account of the group structure. In fact, if one simply takes a smooth function f on G which is quadratic about the identity element e, then $\varphi(r, q) = f(r^{-1}q)$ offers itself immediately as a possibility for design. Using the machinery of generators associated with the Lie algebra $\mathfrak{g}$ corresponding to G one can also easily transport velocities. However, there is, in our view, a natural selection in all cases which in many instances provides the desired pair. Namely, we suggest, in the spirit of [9], that one use the constructs of intrinsic geometry. In this section we show, as a corollary to theorem 3.3.1, that the function

$$\varphi(r, q) = \frac{1}{2} \operatorname{dist}(r, q)^2 \tag{3.4}$$

and the traditional parallel transport along length minimizing geodesics form a compatible configuration error-transport mapping pair. As we shall see, in accord with our desire, the tracking design which results from (3.3.1) will tend always to push along the trace of the geodesic connecting r to q and with a strength proportional to the distance separating them.

To begin we must be careful to specify an appropriate domain $P \subseteq Q \times Q$ for φ. Namely, one in which it φ shall be smooth. Our stratagem is this. We will specify a submanifold of $Q \times Q$ in which (3.1) holds and such that $l^{\gamma}(r, q)$ depends smoothly on $p = (r, q) \in P$.

Assume that $\mathrm{Inj}(Q)$ is non-degenerate and consider the connected fibered submanifold

$$BQ_{\mathrm{Inj}(Q)} \quad = \quad \left\{ v \in TQ : \sqrt{g\,(v, v)} < \mathrm{Inj}(Q) \right\}.$$

of TQ (see Appendix D for the details of the bundle structure). The canonical charts inherited from TQ do not immediately provide local trivializations about points of $BQ_{\mathrm{Inj}(Q)}$, as the image of the coordinate neighborhoods have the form

$$\left\{ (q_i, v^j) \in \psi_\alpha(U_\alpha) \times \mathbb{R}^n : \sqrt{g_{ij} v^i v^j} < \mathrm{Inj}(Q) \right\}.$$

Nevertheless, using the coordinate neighborhoods (U_α, ψ_α) of Q and the canonical basis fields associated to these charts, such local trivializations can be constructed. Since the Riemannian metric is a smooth $(0, 2)$ tensor, we may act the Gram–Schmidt procedure upon the canonical basis field to obtain an orthonormal frame field. With respect to this frame field we may define coordinates on each vector space $\pi^{-1}(q)$ in the same manner that coordinates are defined on these fibres with respect to the canonical fields and thus we may define a chart mapping, say $\hat{\psi}_\alpha$, on TU_α. The image of $\pi^{-1}(U_\alpha)$ under $\hat{\psi}_\alpha$ has the desired form

$$\left\{ (q_i, w^j) \in \psi_\alpha(U_\alpha) \times \mathbb{R}^n | \sqrt{\delta_{ij} w^i w^j} < \mathrm{Inj}(Q) \right\} = \psi_\alpha(U_\alpha) \times B_{\mathrm{Inj}(Q)}(0).$$

Thus, $BQ_{\mathrm{Inj}(Q)}$ is, in fact, a fiber bundle in its own right and we refer to it as the ball–bundle.

We now construct an embedding of $BQ_{\mathrm{Inj}(Q)}$ into $Q \times Q$. The embedded space will be the appropriate domain for which the general configuration error defined in the next section will be smooth.

Proposition 3.4.1 *Let $f\colon BQ_{\mathrm{Inj}(Q)} \to Q \times Q$ be defined as*

$$f(v) \quad = \quad (\pi(v), exp(v))\,. \tag{3.5}$$

Then f is an embedding.

If $r = \pi(v) = \pi(v')$ then $v, v' \in T_rQ$. As a map on any fixed base point, $\exp_r$ is a diffeomorphism and hence $\exp_r(v_r) = \exp_r(v'_r)$ implies $v = v'$. Therefore, f is 1–1. All that remains is to show that f is differentiable with differentiable inverse in a neighborhood of each point.

Consider the map f in its local form

$$F(\vec{r}, \vec{v}) = (\vec{r}, \mathrm{Exp}(\vec{r}, \vec{v}))\,.$$

The exponential map exp is a differentiable function. Thus all partials of Exp exist and are smooth. The tangent mapping of F has the following local form:

$$TF = \begin{bmatrix} I & T\mathrm{Exp}_{\vec{v}} \\ 0 & T\mathrm{Exp}_{\vec{r}} \end{bmatrix}\,.$$

Given any fixed $r \in Q$, if $v \in B_{\mathrm{Inj}(Q)}(0)$ then $\exp_r(v)$ is not a conjugate point by proposition 3.2.4. Thus $T\mathrm{Exp}_{\vec{r}}(\vec{v})$ is injective by definition 3.2.3. That TF is injective at each point $(\vec{r}, \vec{v})$ follows as the diagonal blocks are injective. By the inverse function theorem, f is a local diffeomorphism. Since the local diffeomorphism maps local neighborhoods into neighborhoods of $Q \times Q$, f is an embedding.

Given a Riemannian manifold Q with non-zero injectivity radius, we define the set

$$EQ \triangleq f(BQ_{\mathrm{Inj}(Q)}) .$$

From the charts on $BQ_{\mathrm{Inj}(Q)}$ and the diffeomorphism f, we have a set of local trivializations for EQ of the form provided in equation (3.4). As such, EQ is indeed a fiber bundle with projection $\rho : EQ \to Q : (r, q) \mapsto r$. EQ is connected since f is a continuous map of $BQ_{\mathrm{Inj}(Q)}$.

From the above result we have the following needed result.

Corollary 3.4.1 *Let Q be a manifold with $\mathrm{Inj}(Q) > 0$. Then for $p = (r, q) \in EQ$, there exists a unique (unit speed), length minimizing geodesic γ between r and q that is differentiably dependent on r and q. As a consequence, equation (3.1) holds and depends smoothly on p.*

We have already discussed the existence and uniqueness of length minimizing segments. Since the geodesic vector field is smooth, geodesics γ are smooth functions of initial data $v \in BQ_{\mathrm{Inj}(Q)}$. Via the above bundle isomorphism, we see that they are differentiable functions of the boundary data p.

The set EQ is a differentiable manifold with fiber over each point r diffeomorphic to a normal neighborhood centered at r of radius $\mathrm{Inj}(Q)$. It is natural to think the trajectory of the mechanical system as "linked", through the fiber, to the reference trajectory path in the base space. The equilibrium of a controlled system will be considered to be the point of the "diagonal" of $Q \times Q$, that is $(r(t), r(t)) \in EQ$.

Consider a curve $p : \mathbb{R} \to EQ : u \mapsto (r(u), q(u))$ which does not touch the diagonal of $Q \times Q$. Composing this curve with l^γ provides a length function with domain, codomain $\mathbb{R}$ parameterized by u, which we shall also denote by l^γ. To which l^γ we refer should be clear by context. Being it the case that $l^\gamma : EQ \to \mathbb{R}$ is smooth and the curve p is smooth, $l^\gamma : \mathbb{R} \to \mathbb{R}$ is smooth and we may take its derivative with respect to the parameter u.

Let $\eta(s,u)$ be the family of length minimizing geodesics parameterized by arc length, s, between $r(u)$ and $q(u)$. Select some fixed parameter value and call it u^*. Taking $\gamma(s)$ to be the minimizing geodesic between $r = r(u)|_{u^*}$ and $q = q(u)|_{u^*}$ we see that $\eta(s,u)$ is a geodesic variation of $\gamma(s)$ with boundary values $\eta(s,u)|_{s=l^\gamma(u)} = q(u)$ and $\eta(s,u)|_{s=0} = r(u)$. That $\eta(s,u)$ depends smoothly on the boundary values $q(u)$ and $r(u)$ follows from proposition 3.4.1.

One finds $\frac{dl^\gamma}{du}$ to be related to the partials of η at u^* through the metric by the first variation formula, which is stated in the following lemma.

Lemma 3.4.1 (Gauss) *Let the curve p and the geodesic variation $\eta(s,u)$ be as described in the above discussion. Then*

$$\left.\frac{dl^\gamma}{du}\right|_{u^*} = g\left(\left.\frac{dq}{du}\right|_{u^*}, \left.\frac{\partial\eta}{\partial s}\right|_{(l^\gamma(u^*),u^*)}\right) - g\left(\left.\frac{dr}{du}\right|_{u^*}, \left.\frac{\partial\eta}{\partial s}\right|_{(0,u^*)}\right)$$

This lemma proves to be invaluable as it is the foundation of the following vital result.

We know that l^γ is smooth on EQ and thus the differentials dl^γ_q and dl^γ_r exist. Using the preceding lemma, we can calculate them.

Corollary 3.4.2 *Let the circumstances be those given in Lemma 3.4.1. Then the differential of the length function with respect to q and to r are, respectively*

$$dl^\gamma_r(q) = g^\flat(q)\frac{d\gamma}{ds}(l^\gamma(r,q)) \tag{3.6}$$

$$dl^\gamma_q(r) = -g^\flat(r)\frac{d\gamma}{ds}(0) \tag{3.7}$$

where, in the right members of these equations, γ is the unique geodesic segment connecting q to r.

Consider the vector v_q and let $\sigma(u) \in v_q$. Using the curve σ we may create, in accord with the above discussions, a curve $p(u) = (r, \sigma(u)) \in EQ$ and an associated geodesic variation $\eta(s,u)$ in which one endpoint, namely r, is fixed. It follows from Lemma

3.4.1 that

$$\begin{aligned}
[dl_r^\gamma](v_q) &= v_q(l_r^\gamma) \\
&= \frac{d}{du} l_r^\gamma(u)\Big|_{u=0} \\
&= g\left(v_q, \frac{\partial \eta}{\partial s}(l^\gamma(0), 0)\right) \\
&= g\left(v_q, \frac{d\gamma}{ds}(l^\gamma(r, q))\right) \\
&= \left[g^\flat(q)\, \frac{d\gamma}{ds}(l^\gamma(r, q))\right] v_q
\end{aligned}$$

Since such is the case for all $v_q \in T_q Q$ the result follows. Eq. (3.7) follows along exactly the same lines. We note the analogy between Corollary 3.4.2 and Lemma 1 of [9]. In another parallel with [9], we remark that $\frac{d\gamma}{ds}$ is the geodesic versor. In the case of the Euclidean sphere, the metric was induced from the ambient space and thus $g^\flat = \mathrm{id}$.

We now show that (3.4) is a local uniformly quadratic tracking error function.

Theorem 3.4.1 (Intrinsic Configuration Error) *Assuming* $\mathrm{Inj}(Q) > 0$*, the function*

$$\varphi(r, q) = \frac{1}{2}(l^\gamma(r, q))^2 \tag{3.8}$$

is a uniformly quadratic tracking error function with $L_{\mathrm{reg}} = \frac{1}{2}(\mathrm{Inj}(Q))^2$.

The proof that this mapping is tracking error function is indeed simple if we make use of a normal chart (U, ψ) centered at r. In such a chart we have the representation $\Phi_r(q_1, \ldots, q_n) = \frac{1}{2}(q_1^2 + \ldots + q_n^2)$. The local expression of the differential is then $(d\Phi_r)_{\psi(q)} = \begin{bmatrix} q_1 & \ldots & q_n \end{bmatrix}$. So $(d\Phi_r)_{\psi(r)=0} = 0$ and r is a critical point of φ_r. This is a geometric property and does not depend on the local representation. Since r is a critical point of φ_r, the Hessian of this function $\mathrm{Hess}\,\varphi_r$ is coordinate independent and is positive definite provided the matrix $\left[\frac{\partial^2 \Phi_r}{\partial q_i \partial q_j}\right]$ is positive definite for any chart whatsoever we should choose. Using normal coordinates we obtain the identity matrix which completes the discussion concerning the conjecture that φ_r is a tracking error function.

We now demonstrate the fact that φ is uniformly quadratic with $L_{\text{reg}} = \frac{1}{2}(\text{Inj}(Q))^2$. Let $r \in Q$. For a given geodesic γ emanating from r, φ is defined only within the injectivity radius and $d\varphi_r \neq 0$ for every interval $[0, L]$ with $L < \frac{1}{2}(\text{Inj}(Q))^2$. Hence $L_{\text{reg}}(\varphi_r(q), r) = \frac{1}{2}(\text{Inj}(r))^2$ for each $r \in Q$. Taking the infimum over all $r \in Q$, we have $L_{\text{reg}}(\varphi, Q) = \frac{1}{2}(\text{Inj}(Q))^2$.

Continuing we consider $\|d\varphi_r(q)\|_g$. Using equation (3.6) we discover by direct calculation that $\|d\varphi_r(q)\|_g^2 = \text{dist}^2(r, q)$. Using the Cauchy-Schwarz inequality, we have, suppressing the evaluation of $\frac{d\gamma}{ds}$ at $l^\gamma(r, q)$,

$$\begin{aligned} |d\varphi_r(q)v_q|^2 &= (l^\gamma)^2 g\left(\frac{d\gamma}{ds}, v_q\right)^2 \\ &\leq (l^\gamma)^2 g\left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds}\right) g\left(v_q, v_q\right) \\ &= (l^\gamma)^2 g\left(v_q, v_q\right). \end{aligned}$$

Taking the square root of both sides and then the supremum over $|v_q| = 1$, it is clear that $\|d\varphi_r(q)\|_g$ can be no greater than l^γ. If we take $v_q = \frac{d\gamma}{ds}(l^\gamma(r, q))$, the supremum is reached and thus the result. Taking $a = b = \frac{1}{2}$ in definition 3.3.2 completes the discussion.

Now we set up the compatible transport mapping. We begin with the proper definition of parallel transport. Let $v_r \in T_rQ$ and $\sigma \colon I \to Q$ be a curve in Q with $\sigma(0) = r$, $0 \in I$. The *parallel transport* of v_r from r to a point $\sigma(u^*) = q$ is the map $\sigma_{r\to q} \colon T_rQ \to T_qQ : v_r \mapsto \sigma_{r\to q}v_r \triangleq V(u^*)$, where $V(u)$ is the solution to the differential equation

$$\nabla_{\frac{d\sigma}{du}} V = 0, \qquad V(0) = v_r\,. \tag{3.9}$$

We now move to show that parallel transport defines a transport mapping on EQ. Our argument will need the aid of the following lemma, which is a standby of the science with which we are involved.

Lemma 3.4.2 *The parallel transport of two vectors $v_r, w_r \in T_rQ$ from a point r to* *another point q along a curve $\gamma \subset Q$ preserves the length of each vector and the inner product between the transport of v_r, w_r in the Riemannian metric $g(\cdot,\cdot)$. That is, $|v_r| = |\gamma_{r\to q} v_r|$ and $g(v_r, w_r) = g(\gamma_{r\to q} v_r, \gamma_{r\to q} w_r)$.*

Theorem 3.4.2 (Intrinsic Transport Map) *Parallel transport along length minimizing geodesics is a transport mapping on EQ as given by definition 3.3.3 .*

Writing the parallel transport equations out in coordinates we get the IVP

$$\frac{\mathrm{d}v^k}{\mathrm{d}u} + \Gamma^k_{ij} v^j \frac{\mathrm{d}\sigma_i}{\mathrm{d}u} = 0 \qquad k = 1, \ldots, n$$

with $v^k(0) = v^k_0$. Thus we get a linear (possibly nonautonomous) system and a unique solution with infinite continuation is assured. The solution at u^* is determined by the associated state transition matrix $\Phi(u^*, 0)$. So we see that, within a chart, parallel transport is, point–wise, a smooth linear mapping with local representation Φ. This conclusion, however, remains true even if we have need of involving multiple charts along the way from r to q. It is, therefore, clear that $\sigma_{r\to q}$ describes a bitensor field over EQ.

By the definition of parallel transport, $\sigma_{r\to r} = \mathrm{id}$. Further, Lemma 3.4.2 indicates that parallel transport preserves frames. Thus, we also know that $\sigma_{r\to q} \in GL(T_rQ, T_qQ)$. Should we show that $\sigma_{r\to q}$ is smooth, then we may conclude that it is a transport mapping. We now take the argument up.

The geodesic field on TQ is smooth. So, by the theory of ordinary differential equations, the geodesics depend smoothly on the initial data. Further, since f is a diffeomorphism of $BQ_{\mathrm{Inj}(Q)}$ onto EQ, the length minimizing geodesics depend smoothly on the boundary data $(r, q) \in EQ$. Thus the field described by equation (3.9) is smoothly parameterized by the boundary data $(r, q) \in EQ$. Hence, calling upon the theory of ordinary differential equations once more, we determine that the solutions $V(u; v_r, (r, q))$ depend smoothly on the initial condition v_r and the boundary data (r, q). Therefore, $\sigma_{r\to q}$ takes smooth fields to smooth sections of $Q \times TQ$. That is, for $Y \in \mathfrak{X}(Q)$, $V(1; Y(r), (r, q))$ is a smooth section of $Q \times TQ$.

It remains to show that this transport mapping is compatible with our selection of φ.

Theorem 3.4.3 (Compatability) *Parallel transport along length minimizing geodesics and the tracking error function (3.8) are* compatible *in the sense that*

$$d\varphi_q(r) = -\gamma^*_{r\to q}\left(d\varphi_r(q)\right) \ .$$

Since $d\varphi_r(q) = l^\gamma(r,q)dl^\gamma_r(q)$, $d\varphi_q(r) = l^\gamma(r,q)dl^\gamma_q(r)$, and $\gamma^*_{r\to q}$ is a linear operator, we need only show

$$d(l^\gamma_q(r)) = -\gamma^*_{r\to q}\left(d(l^\gamma_r(q))\right).$$

Consider an arbitrary vector $v_r \in T_rQ$. Computing directly:

$$\begin{aligned}
[dl^\gamma_q(r)](v_r) &\overset{eq.(3.7)}{=} \left[-g^\flat(r)\frac{d\gamma}{ds}(0)\right](v_r) \\
&= -g\left(\frac{d\gamma}{ds}(0), v_r\right) \\
&\overset{Lm.3.4.2}{=} -g\left(\gamma_{r\to q}\frac{d\gamma}{ds}(0), \gamma_{r\to q}v_r\right) \\
&= -g\left(\frac{d\gamma}{ds}(l^\gamma(r,q)), \gamma_{r\to q}v_r\right) \\
&= \left[-\gamma^*_{r\to q}\left(g^\flat(q)\left.\frac{d\gamma}{ds}\right|_{l^\gamma(r,q)}\right)\right](v_r) \\
&\overset{eq.(3.6)}{=} \left[-\gamma^*_{r\to q}\left(dl^\gamma_r(q)\right)\right](v_r).
\end{aligned}$$

This calculation holds for all $v_r \in T_rQ$ and so the result.

Theorems 3.4.1 and 3.4.3 ensure that the objective of this paper is obtained. Namely we have constructed a generalization of [9], or an instance of theorem 3.3.1, in which the intrinsic geometry of an abstract machine creates a closed-loop system analagous to a mass-spring–damper.

Corollary 3.4.3 *Suppose that one has an abstract machine with dynamics given by equation (3.3). Further assume that the machine's configuration manifold Q has a* *positive injectivity radius. Let the control force be defined as* $F = F_{\mathrm{FB}} + F_{\mathrm{FF}}$ *where*

$$\begin{aligned} F_{\mathrm{FB}} &= -\alpha\, d\varphi_r(q) - \beta\, g^\flat(q)\dot{e} \\ &= -\alpha\, g^\flat(q)\, \mathrm{grad}\, \varphi_r - \beta\, g^\flat(q)\dot{e} \\ &= -\alpha\ \mathrm{dist}(r,q)\, g^\flat(q)\, \frac{d\gamma}{ds} - \beta\, g^\flat(q)\dot{e} \\ \dot{e} &= \dot{q} - \gamma_{r\to q}\dot{r} \\ F_{\mathrm{FF}} &= g^\flat(q)\left(\frac{\mathrm{D}}{\mathrm{d}t}\gamma_{r\to q}\dot{r}\right) \end{aligned}$$

where $\varphi(r,q)$ *is the tracking error function given in equation (3.8) and* γ *is the unique minimizing geodesic between* $r(t)$ *and* $q(t)$ *at time* t*. Let*

$$V(t) = \frac{1}{2}\,\alpha\ \mathrm{dist}^2(r(t), q(t)) + \frac{1}{2}\, g\left(\dot{e}(t), \dot{e}(t)\right)$$

be the Lyapunov candidate. Then, $r(t)$ *is stable and should the boundedness conditions of theorem 3.3.1 be satisfied over* $P = EQ$ *the curve* $r(t)$ *is locally exponentially stable from all initial conditions satisfying*

$$\mathrm{dist}^2(r(0), q(0)) + \frac{1}{\alpha}\, g\left(\dot{e}(0), \dot{e}(0)\right) < \mathrm{Inj}(Q)^2\ . \tag{3.10}$$

In the next theorem, we show that this particular design induces the mass–spring–damper paradigm *explicitly* when the reference trajectory and mechanical system initial condition belongs to $T\mathrm{Im}(\gamma)$. We shall also see that corollary 3.4.3 has another advantage.

Theorem 3.4.4 *Let* $r(t)$ *be a smooth curve with image contained within the path of a geodesic* $\gamma : \mathbb{R} \to Q$*, where* Q *is a manifold with positive injectivity radius. For any initial condition of the mechanical system satisfying*

1. *Inequality 3.10,*

2. $q(0) \in \mathrm{Im}(\gamma)$*, and*

3. $\dot{q}(0) \in T_{q(0)}\mathrm{Im}(\gamma)$

the closed loop solution q to (3.3) under the control law of corollary (3.4.3) satisfies the following:

1. *$q(t)$ lies in the trace of γ for all $t > 0$*

2. *$q(t) = r(t)$ is exponentially stable with rate of convergence dictated by α and β.*

Choosing a base point $o \in \mathrm{Im}(\gamma)$ and a unit tangent vector $v_o \in T_o\mathrm{Im}(\gamma)$, parameterize the curve γ by the associated signed arc length measure s so that $\gamma(s) \triangleq \gamma(s, v_o)$. Let the reference trajectory be parameterized by, say $s_r(t)$, so that

$$r(t) = \gamma(s_r(t))\,.$$

Now, let us assume momentarily or "guess" that the signed arc length difference $\Delta s(t)$, where $\mathrm{dist}(r, q) = |\Delta s(t)|$, between the points $r(t)$ and $q(t)$ satisfies the mass–spring–damper differential equation:

$$\Delta\ddot{s} = -\alpha\Delta s - \beta\Delta\dot{s}\,. \tag{3.11}$$

We show that the solution to the closed loop dynamic system under the control action of corollary 3.4.3 is

$$q(t) = \gamma(s_r(t) + \Delta s(t))\,.$$

During our computations we will suppress the functional dependence on t.

Taking derivatives, we have the velocity vectors

$$\begin{aligned} \dot{q} &= \left.\frac{d\gamma}{ds}\right|_{(s_r+\Delta s)} (\dot{s}_r + \Delta\dot{s}) \\ \dot{r} &= \left.\frac{d\gamma}{ds}\right|_{s_r} \dot{s}_r. \end{aligned}$$

The parallel transport of $\dot{r}(t)$ to the base point $q(t)$ is given by

$$\begin{aligned}\gamma_{r\to q}\dot{r} &= \gamma_{r\to q}\frac{d\gamma}{ds}\Big|_{s_r}\dot{s}_r \\ &= \frac{d\gamma}{ds}\Big|_{(s_r+\Delta s)}\dot{s}_r\,.\end{aligned}$$

So,

$$\begin{aligned}\dot{e} &= \dot{q} - \gamma_{r\to q}\dot{r} \\ &= \frac{d\gamma}{ds}\Big|_{(s_r+\Delta s)}\Delta\dot{s}.\end{aligned}$$

Taking the covariant derivative of the velocity error vector, we have, suppressing the evaluation of $\frac{d\gamma}{ds}$ at $s_r + \Delta s$,

$$\begin{aligned}\frac{\mathrm{D}\dot{e}}{\mathrm{d}t} &= \frac{\mathrm{D}}{\mathrm{d}t}\left(\frac{d\gamma}{ds}\Delta\dot{s}\right) \\ &= \frac{\mathrm{D}}{\mathrm{d}t}\left(\frac{d\gamma}{ds}\right)\Delta\dot{s} + \frac{d\gamma}{ds}\Delta\ddot{s} \\ &= \frac{\mathrm{D}}{\mathrm{d}s}\left(\frac{d\gamma}{ds}\right)(\dot{s}_r+\Delta\dot{s})\,\Delta\dot{s} + \frac{d\gamma}{ds}\Delta\ddot{s} \\ &= \frac{d\gamma}{ds}\Delta\ddot{s}\end{aligned}$$

as $\frac{\mathrm{D}}{\mathrm{d}s}\frac{d\gamma}{ds} = 0$ due to the fact that γ is a geodesic. Substituting (3.11), we have

$$\begin{aligned}\frac{\mathrm{D}\dot{e}}{\mathrm{d}t} &= \frac{d\gamma}{ds}\Big|_{(s_r+\Delta s)}(-\alpha\Delta s - \beta\Delta\dot{s}) \\ &= -\alpha\,\mathrm{grad}\varphi_r(q) - \beta\,\dot{e}\end{aligned}$$

Under the operator $g^\flat(q)$ we arrive at the closed loop dynamics given by the control law of Corollary 3.4.3. By uniqueness of solutions we find that our "guess" of the mass–spring–damper dynamics in (3.11) was correct.

Furthermore, note that the Lyapunov function presented in Corollary 3.4.3 takes the form

$$\begin{aligned} V(t) &= \frac{\alpha}{2}\Delta s^2 + \frac{1}{2} g(\dot{e}, \dot{e}) \\ &= \frac{\alpha}{2}\Delta s^2 + \frac{1}{2}\Delta \dot{s}^2 \left| \frac{d\gamma}{ds} \right|^2 \\ &= \frac{\alpha}{2}\Delta s^2 + \frac{1}{2}\Delta \dot{s}^2 \end{aligned}$$

on the geodesic. This is a Lyapunov function we're used to seeing for the mass-spring damper.

For this restricted case, we can design the arc length and arc length rate to be over-damped, critically damped, or under-damped through judicious choice of positive constants α and β. Hence, this is a very "practical" design tool for those interested in tuning the performance of an abstract machine. Imbedded in this geometric control methodology is a second order linear, time-invariant ODE in kinetic energy distance. It is seen that the relative kinetic energy dissipates as the mechanical system's trajectory converges to the reference trajectory.

Another interesting consequence unfolds if we further restrict our considerations to the set point control problem with zero initial conditions.

Corollary 3.4.4 *Let $S(r) \triangleq Q \setminus C(r)$. Let $\dot{r}(t) = \dot{q}(0) = 0$. Then there exists a control law such that $q(t)$ exponentially converges to r for all $q(0) \in S(r)$.*

Choose α and β as to ensure the ODE in equation (3.11) is over-damped. The definitions for arc length and parallel transport can be extended along each geodesic up until its intersection with the cut locus. The arc length is strictly decreasing and the trajectory length can never leave the geodesic between r and $q(0)$. The convergence is exponential by the above argument.

3.5 The Assumption $\mathrm{Inj}(Q) > 0$

It behooves us to at least discuss that assumption which drives the use of intrinsic mechanisms for control and that assumption is that the injectivity radius of the abstract machine along its desired motion is nonzero. Being a bit more restrictive, let us, as we have done heretofore, consider only the case of the geodesically complete manifold and the condition that the injectivity radius of the machine is nonzero.

In several instances, which we shall delineate, certain other properties of the abstract machine's Riemannian structure ensures us that this focal assumption is satisfied. Let us begin with those which provide the greatest dividends w.r.t. this particular assumption. These are the suppositions which drive the theorem of Hadamard:

Theorem 3.5.1 (Hadamard; As Stated in [13]) *Let Q be a complete Riemannian manifold, simply connected, with sectional curvature $K(r,\sigma) \leq 0$ for all $r \in Q$ and for all $\sigma \subseteq T_rQ$. Then Q is diffeomorphic to $\mathbb{R}^n$, $n = \dim Q$; more precisely $\exp_r$ is a global diffeomorphism.*

Thus we see that under these assumptions none of the complications which we have previously discussed arise and $\mathrm{Inj}(Q) = \infty$. The concern is then shifted to the other assumption made in course to 3.4.3. Namely, the boundedness assumptions. Here we should say that is seems to be the case that these assumptions are gratuitous barring the desire to guarantee exponential stability. In fact, the control design works well for those machines which simply satisfy those requirements for Hadamard's theorem.

Example 3.5.1 (Tracking in the Poincarè Plane) *We take up for study here the Poincarè half-plane model of Hyperbolic geometry. Consider the upper half plane U as an open submanifold of $\mathbb{R}^2$. By restriction, the standard chart for $\mathbb{R}^2$, becomes a global chart for U and we make use of the corresponding coordinate system (u,v). We impose upon this smooth manifold the metric $ds^2 = \frac{1}{v^2}(du^2 + dv^2)$, which has components $g_{11} = g_{22} = \frac{1}{v^2}$ and $g_{12} = g_{21} = 0$. Thus, we acquire a Riemannian manifold $\mathbb{H}^2$ which is said to be conformal with ruler function $h = v$. By assuming the control co-distribution $\mathcal{F} = \{du, dv\}$, we complete the construction of an abstract machine.*

Now, what of the injectivity radius of this machine? $\mathbb{H}^2$ *is complete and clearly simply connected. Further, it is well known that hyperbolic space has constant sectional curvature* $K(\mathbb{H}^n, \sigma) = -1$ *for all tangent planes* σ*. Thus, this machine satisfies the hypotheses of Hadamard's theorem and* $\mathrm{Inj}(\mathbb{H}^2) = \infty$*. This tells us that our geometric control constructs are globally well defined as is the corresponding control law.*

In fact, everything is known of this machine. Free motions are vertical lines and semicircles orthogonal to the u–axis. For the pair $(R, Q) \in \mathbb{H}^2$ *defining the "line" (semicircle) with "endpoints"* (X, Y) *we have*

$$\mathrm{dist}(R, Q) = \ln\left(\frac{\overline{RX}\,\overline{QY}}{\overline{RY}\,\overline{XQ}}\right)$$

where juxtaposition with overbar indicates Euclidean distance. Likewise there is a closed form coordinate expression of the parallel transport operator for a given pair (R, Q)*, however, we will spare the detail and simply present the result of having put the controller to action. In Figure 3.1 we illustrate the result of tracking the circular reference* $r(t) = (u_r(t), v_r(t)) = (0.5\cos(t), 0.5\sin(t) + 1)$ *for the zero velocity initial position* $q(0) = (u_q(0), v_q(0)) = (1, 1)$*. Note that at two instances, including the initial, we indicate that line connecting the machine configuration with that of the reference.*

Another hypothesis with wide sweeping implications is that of compactness. If the abstract machine's configuration manifold is compact, then $\mathrm{Inj}(Q) > 0$ and we are guaranteed an exponential region of attraction, as the boundedness assumptions are also met. This covers a good number of actual machines, such as robot manipulators, whose configuration manifolds are the n-dim tori. Further, should one be able to get some controlling bounds on a compact manifold's sectional curvature, then more definite information can be obtained. This brings us to the following theorems, each of which provides definite estimates on $\mathrm{Inj}(Q)$.

Theorem 3.5.2 (Klingenberg; As Stated in [2]) *If* Q *is a compact Riemannian manifold with sectional curvature* $K \leq \Delta$ *everywhere then the injectivity radius of* Q

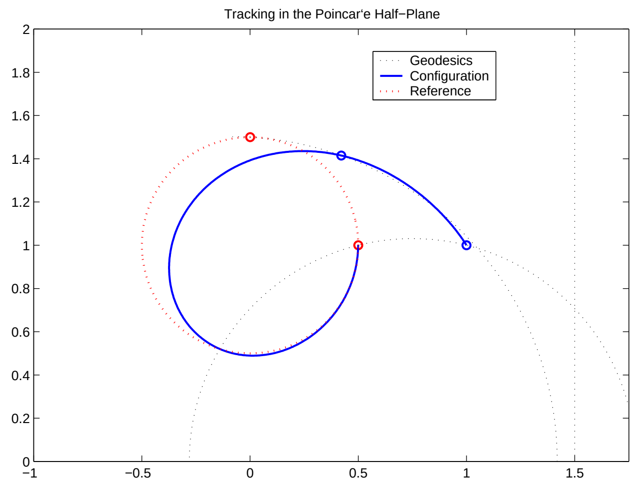


Figure 3.1: Tracking a circular trajectory in the Poincarè Plane

is not smaller than the lesser of the two numbers

1. $\frac{\pi}{\Delta}$

2. *half the length of the shortest periodic geodesic.*

The theorem of Cheeger aids in the assessment.

Theorem 3.5.3 (Cheeger; As Stated in [2]) *The length of the periodic geodesics on a compact manifold of given dimension can be controlled from below with the three following quantities:*

1. *a lower bound δ for the sectional curvature*

2. *a lower bound for the volume* $\mathrm{Vol}(Q)$ *and*

3. *an upper bound for the diameter* $\mathrm{diam}(Q)$.

Getting tighter we have

Theorem 3.5.4 (Klingenberg; As Stated in [2]) *Let Q be a compact, oriented, and simply connected manifold with $0 < K \leq \Delta$. If either*

1. *Q has odd dimension and satisfies the "pinching" hypothesis*

$$\delta \leq K \leq \Delta$$

with $\frac{\delta}{\Delta} > \frac{1}{4}$ or

2. *Q has even dimension*

then the injectivity radius of Q satisfies

$$\mathrm{Inj}(Q) \geq \frac{\pi}{\sqrt{\Delta}} .$$

As an application of this sections geodesic–based control design along with the corresponding cut locus and injectivity radius issues see Section 3.6 wherein we consider the double gimbal system.

3.6 The Geometry of the Double Gimbal

To undertake the general PD control design of Theorem 3.3.1 for the double gimbal system it is necessary to identify 1) a notion of intrinsic distance, 2) mass preferential directions, and 3) a consistent method for comparing velocities in this space. We use the mass distribution of the double gimbal system to induce a Riemannian geometry on T^2. It is this geometry that provides the extra structure needed to construct configuration and velocity errors for a geometric based PD control logic for the double gimbal system. That is, following Section 3.4, we use geodesic distance and parallel transport along geodesics to construct a compatible configuration error and transport map pair that are used in, for example, fixed point to fixed point tracking for the double gimbal system. The configuration space of the double gimbal system is $T^2 = S^1 \times S^1$. We realize this torus as a set of points in $\mathbb{R}^3$ defined in (θ, ϕ) coordinates by the parameterization

$$T(\theta,\phi) = \begin{bmatrix}(R + r\cos(\phi))\cos(\theta) & (R + r\cos(\phi))\sin(\theta) & r\sin(\phi)\end{bmatrix}^{\mathrm{T}} . \tag{3.12}$$

where θ and ϕ in the interval $[0, 2\pi)$ represent the spinning angles of the outer and inner gimbals, respectively (see Figure 3.2). By appropriate restriction of the parameterization in (3.12), T^2 is in fact a 2-dimensional manifold or a surface. Conspicuously absent from this manifold description of the double gimbal system are the mass properties for each of the gimbals. The remedy to this obvious shortcoming is to equip the torus T^2 with a *metric or energy tensor, g* which, simply stated, defines a mass and configuration dependent weighted dot product on T^2.

For the double gimbal system the energy tensor is given in coordinate form by

$$g(\mathbf{v},\mathbf{v}) = \mathbf{v}^{\mathrm{T}} M(\mathbf{q})\mathbf{v}\,, \tag{3.13}$$

where $\mathbf{v}$ is a velocity vector and $M(\mathbf{q})$ is the gimbal system's *generalized mass matrix* given by

$$M = \begin{bmatrix} m_{11} & m_{12} \\ m_{12} & m_{22}\end{bmatrix} = \begin{bmatrix} I_z + J_z\cos(\phi)^2 + J_x\sin(\phi)^2 & 0 \\ 0 & J_y\end{bmatrix} . \tag{3.14}$$

The inertias I and J denote, respectively, the body frame inertia scalars of the outer and inner gimbals with respect to their mass centers. The conspicuous feature of the energy tensor is the weights of the dot product defined are determined by the mass distribution of the double gimbal system in a particular configuration.

A manifold equipped with a metric tensor is called a Riemannian manifold. The manifold T^2 equipped with the energy tensor defined by the mass matrix in (3.14) is a Riemannian manifold we call the *double gimbal torus*.

With the energy tensor for the double gimbal system in hand an *intrinsic distance function* between two gimbal configurations can now be defined. First, the *arclength* *of a curve segment on* T^2, $\sigma(u)$ for u in the interval $[a,b]$ is given by

$$l_\sigma(q,r) = \int_{u=a}^{u=b} \sqrt{g\left(\frac{d\sigma}{du}, \frac{d\sigma}{du}\right)} du \tag{3.15}$$

$$= \int_{u=a}^{u=b} \sqrt{m_{11}\left(\frac{d\theta}{du}\right)^2 + 2m_{12}\frac{d\theta}{du}\frac{d\phi}{du} + m_{22}\left(\frac{d\phi}{du}\right)^2} du, \tag{3.16}$$

where $\sigma(u)$ is a curve on T^2 defined by $T(\gamma(u))$ for some curve $\gamma(u) = [\theta(u), \phi(u)]$ in the gimbal parameter space. Now, given the arclength function, the *intrinsic distance* between the gimbal configurations $q = \sigma(a)$ and $r = \sigma(b)$ is defined as

$$\text{dist}(q,r) = \inf_\sigma l_\sigma(q,r), \tag{3.17}$$

which gives the distance between q and r as the greatest lower bound on the lengths of all smooth curves σ connecting q, r. The intrinsic distance function of (3.17) is an immediate precursor to the definition of configuration error we use for inertia based PD control in the control section.

For all but those pathological Riemannian manifolds, there is at least one curve that yields the intrinsic distance. These special curves called *geodesics* are those curves, out of all possible curves, that minimize the energy functional (see [27]),

$$E(\alpha) = \int_{u=a}^{u=b} g\left(\frac{d\alpha}{du}, \frac{d\alpha}{du}\right) du. \tag{3.18}$$

Equivalently, geodesics are those curves which satisfy the *geodesic equations* (the analogue of Newton's F=ma for F=0 and m=1 on a Riemannian manifold) given in coordinates by

$$\sum_{j,k} \frac{d^2\alpha^i}{du^2} + \Gamma^i_{jk}\frac{d\alpha^j}{du}\frac{d\alpha^k}{du} = 0. \tag{3.19}$$

where Γ^i_{jk} are the *Christoffel symbols*, which for the double gimbal are those functions of the energy tensor and its partial derivatives, given by

$$\Gamma^i_{jk} = \sum_a \frac{1}{2} m^{ia}(m_{ak,j} + m_{ja,k} - m_{jk,a}), \tag{3.20}$$

where m^{ia} is the k^{th}-row and a^{th}-column entry of inverse of the mass matrix and $_{,i}$ denotes partial differentiation with respect to the i^{th} coordinate, see Table 3.7. From this point forward we use an implied summation on repeated upper and lower indices. The geodesic equations (or minimal energy dynamics) for the double gimbal system are

$$\frac{d^2\phi}{du^2} + \frac{\cos(\phi)\sin(\phi)(J_z - J_x)}{J_y}\left(\frac{d\theta}{du}\right)^2 = 0, \tag{3.21}$$

$$\frac{d^2\theta}{du^2} - \frac{2\cos(\phi)\sin(\phi)(J_z - J_x)}{I_z + J_z(\cos(\phi))^2 + J_x(\sin(\phi))^2}\left(\frac{d\theta}{du}\right)\left(\frac{d\phi}{du}\right) = 0. \tag{3.22}$$

Double gimbal geodesics between gimbal configurations p and q serve as the mass preferential directions of the PD control logic in the control section. With such a critical role to play, it is important to have as full a library as possible of double gimbal geodesics.

Analyzing *Clairaut's relation* (see [12] for details) given by

$$m_{11}\dot{\theta} = \text{ constant } = \sqrt{m_{11}}\cos(\psi), \tag{3.23}$$

where ψ is the angle between the geodesic and a parallel crossed by the geodesic leads to geodesics which are qualitatively a) *parallels*: $\phi = k_1$ (constant) and b) *meridians*: $\theta = k_2$ (constant) c) *bound geodesics*: those geodesics bound between two parallels and d) *asymptotic geodesics*: those geodesics which approach $\phi = k_3$ (constant). These solutions are summarized in Table 3.7 with plots given in Figure 3.3

The geodesic curves of Table 3.7 seen in Figure 3.3 are some of the minimal energy motions of the double gimbal system; motion here meaning the time evolution of the gimbal system absent of external forces. For instance, the geodesic curve which asymptotically approaches $\phi = \pi/2$ corresponds to the free/uncontrolled motion seen in Figure 3.4. That is, given the right initial conditions, the double gimbal system can point "due north" (without control). Since no geodesic asymptotically approaches all gimbal pointing direction, some form of control is necessary to reach these configurations.

We arrive at a simple model of the double gimbal that most engineers use in practice when we move beyond the Christoffel symbols to the *Riemann curvature tensor* (see Appendix A), the entries of which are constructed from the components of the mass matrix and their first and second partial derivatives by the formula

$$R_{\gamma\alpha\beta\cdot}^{\ \ \ \ \rho} \triangleq \Gamma^{\rho}_{\gamma\beta,\alpha} + \Gamma^{\Delta}_{\gamma\beta}\Gamma^{\rho}_{\Delta\alpha} - \Gamma^{\rho}_{\gamma\alpha,\beta} - \Gamma^{\Delta}_{\gamma\alpha}\Gamma^{\rho}_{\Delta\beta}. \tag{3.24}$$

A cursory analysis of the double gimbal data from Table 3.7 presents immediately for inspection $J_z = J_x$, since not only does the Gauss curvature vanish but so also can the geodesics equations be written as

$$\frac{d^2\phi}{du^2} = 0 \quad \text{and} \quad \frac{d^2\theta}{du^2} = 0. \tag{3.25}$$

We call this special gimbal system a *flat double gimbal.* The types of geodesics for the flat gimbal are summarized by the first row of Table 3.7. Apparently, an absence of curvature decouples the two gimbal motions while the presence of curvature acts to couple the gimbal motions.

To this point we have developed an intrinsic distance function that, as in Section 3.4, we use to define the gimbal PD control configuration error

$$\varphi(r, q) = \frac{1}{2}\,\mathrm{dist}(r, q)^2.$$

We have also introduced geodesics which play the role of mass preferential directions for the gimbal system.

Equally important to an inertia–based PD for the double gimbal is the notion of a *transport mapping*, $\mathcal{T}_{r\to q}$ which is, essentially, a smooth mapping that carries v_r, a velocity of gimbal in configuration r, to $\mathcal{T}_{r\to q}(v_r)$, a velocity of the gimbal in configuration q. A transport map is required since only velocities of the gimbal system in the same configuration can be subtracted. Working in a metric or inertia based geometry leads us to naturally choose parallel transport along the gimbal mass preferential directions (geodesics) as the transport map to define PD control law velocity error in the control section.

The coordinate free form of the parallel transport equations are

$$\frac{D}{du}V = 0, \tag{3.26}$$

where D/du is the *covariant derivative (with respect to the Levi-Civita connection) along a curve* $\sigma(u)$ and V is a vector field along $\sigma(u)$. Irrespective of the exact meaning of *covariant*, we deduce the solution vector field is, in some sense, constant along the curve $\sigma(u)$ (see Appendix A for the details of the Levi-Civita connection and a covariant derivative). In coordinates the parallel transport equation are the linear system of, possibly, non-autonomous ODE's given by

$$\frac{dV^i}{du} + \Gamma^i_{jk}\frac{d\gamma^j}{du}V^k = 0, \tag{3.27}$$

where Γ^i_{jk} are the Christoffel symbols from (A.32). For some examples of parallel transport see Appendix A.

Using the square of geodesic distance to define a configuration error function and parallel transport along geodesics defining to define a transport map, the tracking control design of Section 3.4 can be implemented. Since this control logic depends on finding a unique minimal geodesic between the desired configuration r and the actual machine configuration q and since there might be multiple geodesics between the two configurations, it is possible that the configuration error function is not well–defined. For this reason, we must take care in ascertaining the stability region of the double gimbal system.

Figure 3.6 illustrates that there are gimbal configurations connected by two minimal geodesics, for example, the red point and any point on the "antipodal" ring. As a result, the Hookean potential φ in (3.4) is not a configuration error function on the whole torus. Consequently the PD control law based on this potential cannot yield asymptotic stability for all reference configurations. Rather we must restrict the neighborhood of stability. The solution to this technical restriction issue is the subject of the rest of this section.

For the Hookean potential of (3.4) the problematic configurations q for each reference configuration r are collectively called the *cut locus* of r denoted by $C(r)$ (cf. Definition 3.2.1). By avoiding the cut locus we can then compute, as required by Theorem 3.3.1, the differential of φ to be

$$-\alpha\, d\varphi_r(q) = -\alpha\, g^\flat\,(\nabla\,\varphi_r) = -\alpha\,\mathrm{dist}(r,q)\; g^\flat\left(\frac{d\gamma}{ds}\right),$$

where γ is the unit speed geodesic segment connecting q to r and $g^\flat$ is the Riemannian metric thought of as a function acting on a vector.

To determine the extent of the cut locus for each point on the reference trajectory $r(t)$ is to know the stability region for the geodesic–based PD tracking control design of the previous section. The best possible circumstance is to find the maximum geodesic distance the actual configuration can be from any reference configuration without encountering a point of the cut locus. This distance denoted by $i(Q)$ is called the *injectivity radius of* Q (cf. equation 3.2). Simply stated, the injectivity radius defines the domain on which φ of (3.4) is in fact a configuration–error function.

That the configuration space $Q = \mathrm{T}^2$ of the double–gimbal system is compact is sufficient to guarantee that $i(\mathrm{T}^2) > 0$. Consequently, there exists some region of phase space determined by $i(\mathrm{T}^2)$ for which an energy– and distance–based control law provides local exponential stabilization for a given double–gimbal reference r. To determine the extent of the stability region is to ascertain the magnitude of $i(\mathrm{T}^2)$.

A result of Klingenberg (1959, see [2]), gives $i(Q)$ as the smaller of 1) half the length of the shortest closed geodesic and 2) π divided by the least upper bound on the sectional curvature (for the double–gimbal torus the sectional curvature is the Ricci curvature and is found from the Riemann curvature using $Ric_{ij} = \sum_a R_{aij\cdot}^{\cdot\cdot a}$). This result is double edged, since now we must discover all the closed geodesics on the double–gimbal torus. A general result of this nature is not known to the authors. We suspect, however, that in many applied cases the meridian geodesics (see Table 3.7) are the shortest closed geodesics. These geodesics correspond to the inner gimbal, which is usually the lightest of the two gimbals, spinning through 2π radians. Any other closed geodesic involves motion of the outer gimbal and thus requires more energy. On taking half the natural length of a purely inner gimbal motion and comparing to the sectional curvature bound, we conjecture half the meridian length to be a good estimate of the injectivity radius $i(\mathrm{T}^2)$ for many practical double gimbals, (see Figure 3.6). Future results in geometry may pave the way to tighter estimates.

Shifting our focus away from the general tracking problem to the fixed–point to fixed–point maneuverers, a little more can be said for the double–gimbal system. In the fixed point case the object of interest, as regards the region of stability, is the cut locus of the gimbal. According to [2], the geodesics of product manifolds are precisely those which project in each factor onto geodesics of the multiplicands. Since the configuration space of the double gimbal system is $\mathrm{T}^2 = \mathrm{S}^1 \times \mathrm{S}^1$, we can deduce almost immediately that the cut locus of a given point must consist of two rings.

To address specifically this problem, a software package called *Loki* has been written (see [38]) which computes the cut locus of an abstract, two–dimensional Riemannian manifold. The software requires only that the metric tensor coefficients g_{ij} or a parameterization for the two–dimensional manifold Q be given to compute the points of control indecision. For the double–gimbal torus, the cut locus of any point appears to be, as suspected, two rings seen for example in Figure 3.6. As a result, we conclude that fixed–point to fixed–point control is indeed possible so long as the two rings are avoided for each point on the geodesic segments connecting the two set–points (see Figure 3.7). Hence we can perform a series of fixed–point to fixed–point operations as illustrated in Figure 3.7 in an energy efficient manner.

3.7 Conclussion

We would like to conclude by commenting on the larger picture, as we see it. For any given fully-actuated abstract machine, Corollary 3.4.3 yields a control law specific to its geometry as a Riemannian manifold. This makes Theorem 3.3.1 applicable to a broader class of problems than was known before, but also introduces some additional complexity to the control design. For those problems whose explicit solutions are not known, numerical solutions to a geodesic boundary value problem need to be computed in practice. However, the advantage of such a design clearly lies in the result of Theorem 3.4.4 and Corollary 3.4.4. For those cases described, the closed loop characteristic behavior is revealed to be that of a spring-mass-damper in the geometry of the machine.

Due to our control design's dependence on intrinsic geometry, nearly every result providing insight into Riemannian geometry tells us something about the intrinsic control of abstract machines. For instance, as addressed in Section 3.5, the theorems of Klingenberg, Cheeger, and Buchner and Wall provide insight into just how the metric structure of the machine's manifold dictates controlling bounds over the stability region of our geodesic based controller.

This is no little thing when we consider further those facts which indicate that an abstract machine's topology alone dictates its capabilities. We speak, of course, of such results as the Morse inequalities, which dictate by means of a machine's Betti numbers the minimum number of equilibria under the influence of a potential. Viewing the configuration error as a potential we see our control limitations. The topology results, however, do not provide us insight into those configurations, for a given machine, at which the physical limitations are felt. Rather this is the role of the geometry (a.k.a. the metric). It doesn't seem a stretch for us to say that the cut loci indicates a machines physical limitations.

Therefore, the assumption on the injectivity radius is not a frivolous supposition. It is *necessary* in order to avoid those, so to speak, singular values of the machine's "natural" (intrinsically dictated by physical parameters) physical capabilities; where the machine "naturally" gets jammed up or confused. There is no more simple illustration of what we speak of than the tracking problem on the circle. Consider chasing someone about a circular corridor. The cut locus is the the antipodal point. It is near this location that the target desires to be in order to be avoid being caught. If the target and the tracker both start out in equilibrium, it is also that place at which there is no obvious direction in which to move and as a result the control law is undefined and the tracker is jammed.

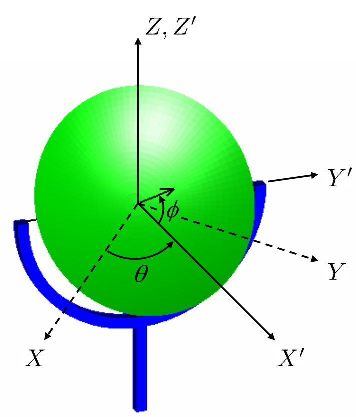


Figure 3.2: Coordinate system for the double gimbal. In a right handed coordinate system, positive θ (the inner gimbal rotation angle) is measured from the X–axis to the Y–axis while positive ϕ (the outer gimbal rotation angle) is measured from the X–axis to the Z–axis.

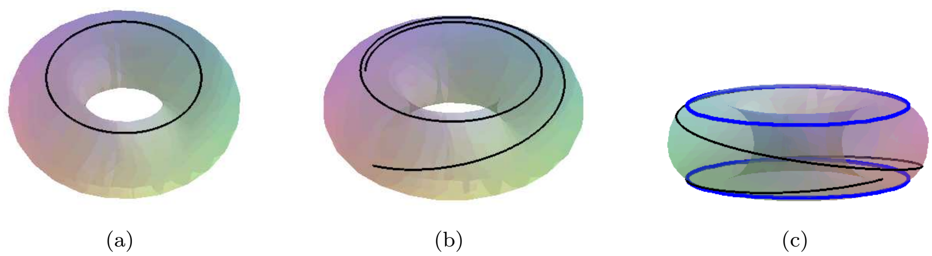


Figure 3.3: Some geodesics on the double gimbal torus. The black curves are the gimbal geodesics $\gamma(u) = (\theta(u), \phi(u))$ mapped under T given in (3.12) and plotted on the surface defined by the image of T. The *parallel, asymptotic and bound geodesics* seen here are typical of any inertias fitting the description of the second line of Table 3.7. In this particular case, the inertias are in the ratio $I_z = 2, J_z = 2, J_x = 1, J_y = 4$ with initial conditions $[\theta = 0, \phi = \pi/2, \dot{\theta} = .8164965808, \dot{\phi} = 0]$,$[\theta = 0, \phi = 0, \dot{\theta} = .6123724355, \dot{\phi} = .3535533910]$ and $[\theta = 0, \phi = 1, \dot{\theta} = .7794529900, \dot{\phi} = 0]$ for the parallel, asymptotic and bound geodesics, respectively. A geodesic on the double gimbal torus is a minimal energy *motion* (time evolution absent external forces) of the double gimbal system.

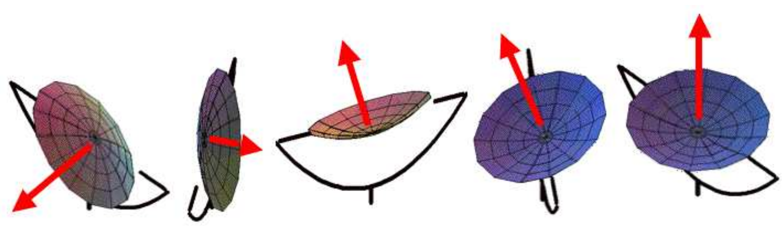


Figure 3.4: Asymptotic Motion. The motion of the double gimbal system approaches the "due north" configuration as indicated by the asymptotic gimbal geodesic from Figure 3.3. Since gimbal geodesics take into account the mass distribution of the system the corresponding motions are the *natural* tendencies of the double gimbal system. Since double gimbal geodesics are central to our inertia–based PD control design, we work with, rather than against, the gimbals natural tendencies.

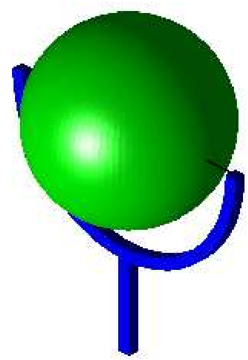


Figure 3.5: The flat double gimbal. An example of the "flat" double gimbal system is constructed by mounting a uniform, spherically symmetric body (denoted J) to the outer gimbal. Since the ratio J_z/J_x equals 1, the mass matrix of (3.14) becomes angle independent and thus the "flat" gimbal inherits its flatness from an angle independent kinetic energy metric. The angle independent metric of the flat gimbal has the effect of decoupling the inner and outer gimbal free motions (see Figure 3.7). Since it is geometrically straight forward to analyze, the flat double gimbal gives an intuitive and simple example into not only the behavior but also the control of more difficult (meaning "non-flat" or coupled) double gimbal systems.

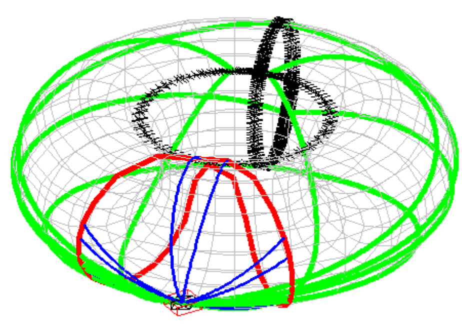


Figure 3.6: Klingenberg result and *Loki* computations on the double gimbal torus. The blue geodesics emanate from a point p for a distance $i(\mathrm{T}^2)$ given by the Klingenberg result which defines the red curve. The green geodesics emanate from point p until they reach the two black rings called the cut locus $C(p)$ as determined by the software *Loki*. For inertia–based tracking control of the double gimbal system, the region described by the red curve is a conservative estimate of the region of local exponential stability. For fixed–point to fixed–point maneuverers, the black rings, are the problematic reference configurations for local exponential stability of our inertia–based PD control logic.

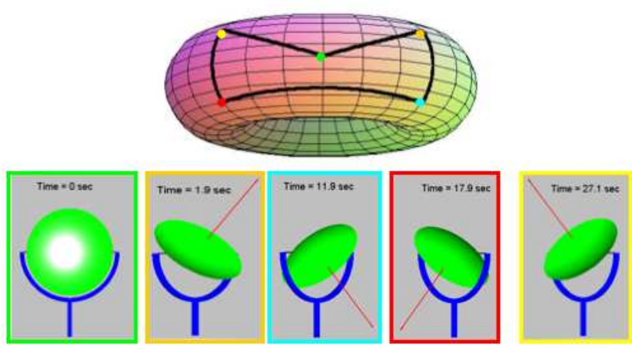


Figure 3.7: Fixed–point to fixed–point tracking on double gimbal torus. The geodesic segments given in black are the minimal energy motions between two successive double gimbal configurations. That the cut locus is avoided for each point along the geodesics indicates that our inertia–based PD control design is valid and proceeding in an energy efficient manner.

Table 3.1: Geometric data for the double gimbal torus. The mass matrix entries m_{ij} are substituted into eq. (A.32) to find the Christoffel symbols, Γ^i_{jk} which are substituted into eq. (A.16) to lead to the function R_{ijkl}. The Gauss curvature, K is then computed by the above equation.

	non-zero entries
m_{ij}	$m_{11} = I_z + J_z \cos(\phi)^2 + J_x \sin(\phi)^2, \quad m_{22} = J_y$
Γ^i_{jk}	$\Gamma^1_{12}=-\frac{\cos(\phi)\sin(\phi)(J_z-J_x)}{I_z+J_z(\cos(\phi))^2+J_x(\sin(\phi))^2} \quad , \Gamma^2_{11}=\frac{\cos(\phi)\sin(\phi)(J_z-J_x)}{J_y}$
R_{ijkl}	$R_{1212}=\frac{\left((J_z-J_x)(\cos(\phi))^4+(2I_z+2\,J_x)(\cos(\phi))^2-I_z-J_x\right)(J_z-J_x)}{I_z+J_z(\cos(\phi))^2+J_x(\sin(\phi))^2}$
K (Gauss Curvature)	$K = R_{1212}/(m_{11}m_{22})$

Table 3.2: Geodesic information for the double gimbal. For a double gimbal with inertia data given in the left column, the geodesic types (*parallel, meridian, bound or asymptotic*) are indicated.

Inertia Data	Geodesic type			
	Parallel $[\phi = k_1,\, \theta \in (0, 2\pi)]$	Meridian $[\theta = k_2,\, \phi \in (0, 2\pi)]$	Bound $[\phi_1 \leq \phi \leq \phi_2]$	Asymptotic [approach $\phi = k_3$]
$I_z, J_z, J_x, J_y > 0$ $J_z = J_x$	all	all	none	none
$I_z, J_z, J_x, J_y > 0$ $J_z > J_x$	$\phi = 0,\, \phi = \pi/2$ $\phi = \pi, \phi = 3\pi/2$	all	yes	$\phi = \pi/2, \phi = 3\pi/2$
$I_z, J_z, J_x, J_y > 0$ $J_z < J_x$	$\phi = 0,\, \phi = \pi/2$ $\phi = \pi, \phi = 3\pi/2$	all	yes	$\phi = 0, \phi = \pi$

Chapter 4

A Frame Bundle Approach to Nonholonomic Mechanical Systems

4.1 Introduction

In sections 4.2.1, 4.2.2, and 4.2.3 of this chapter we derive the constrained n-symplectic dynamics two ways. The first requires that the soldering form of Appendix C be adapted to the distribution. The second transforms the canonical dynamics of [28] (expressed in terms of canonical coordinates on LQ with details given in Appendix G) to the distribution adapted coordinates. We find that the two formulation are the same (as they should be) and take this as a double check that they are correct.

Given the correct mathematical formulation of the constrained n-symplectic dynamics we look to a couple of examples (the vertical rolling hoop and a nonholonomic constrained particle) to determine the applicability of these equations to mechanics and control. We find indeed that some useful information comes out, namely an n-symplectic formulation of the nonholonomic momentum equation of [3]. That there are extra momenta (specific to the n-symplectic formulation and thus unobtainable from the smaller geometric setting of TQ) and that the n-symplectic theory allows for potentials (the gradients of which occur in not only the nonholonomic momenta but also the the extra generalized momenta) indicate the feasibility of potential shaping and momenta based control design. For some preliminary thoughts along this line see Appendix I.

4.2 n-Symplectic Constrained Dynamics

We summarize here the main theorem of this section. Sections 4.2.1, 4.2.2, and 4.2.3 should be viewed as the proof of this theorem.

Theorem 4.2.1 *(N-Symplectic Constrained Dynamics) Given the n-symplectic rank 2 structure equation*

$$d\hat{g}^{(i_2 i_1)} = -2X_{\hat{g}}^{(i_2} \lrcorner\ \hat{\Omega}^{i_1)}$$

the integral curve equations for the equivalence class of solution vector fields $[X_{\hat{g}}^{i_2}]$ *adapted to a constraint distribution* Δ *(via the matrix* G*) are given by*

$$\begin{aligned} \dot{q}^{i_2 i_4} &= \overset{\vartriangle}{X}{}^{i_2 i_4} = \overset{\vartriangle}{g}{}^{i_7 i_5} \overset{\vartriangle}{\Pi}{}^{i_2}_{i_5}\, G^{i_4}_{i_7} + \mathrm{Extra}^{i_2 i_4} && \text{(4.1a)} \\ \dot{\overset{\vartriangle}{\Pi}}{}^{i_1 i_2}_{\cdot\cdot i_3} &= G^{i_6}_{i_3} \left[-\frac{1}{2} \overset{\vartriangle}{g}{}^{i_7 i_5}_{,i_6}\, \overset{\vartriangle}{\Pi}{}^{i_1}_{i_7} \overset{\vartriangle}{\Pi}{}^{i_2}_{i_5} + 2\, \overset{-1}{G}{}^{i_8}_{[i_4,i_6]} \overset{\vartriangle}{\Pi}{}^{(i_1}_{i_8}\, \overset{\vartriangle}{X}{}^{i_2) i_4} \right] + \mathrm{Extra}^{i_1 i_2 \cdot}_{\cdot\cdot i_3} && \text{(4.1b)} \end{aligned}$$

where $\overset{\vartriangle}{g}{}^{i_3 i_4} = g^{i_5 i_6}\, \overset{-1}{G}{}^{i_3}_{i_5} \overset{-1}{G}{}^{i_4}_{i_6}$ *and where*

$$\begin{aligned} \mathrm{Extra}^{i_2 i_4} &= B^{i_2 i_4} \\ \mathrm{Extra}^{i_1 i_2 \cdot}_{\cdot\cdot i_3} &= G^{i_7}_{i_3} C^{(i_1 i_2)}_{,i_7} + G^{i_7}_{i_3} T^{[i_1 i_2]}_{i_7} + \overset{-1}{G}{}^{i_5}_{i_8} \overset{\vartriangle}{\Pi}{}^{(i_1}_{i_5}\ B^{i_2) i_8}_{,i_7} G^{i_7}_{i_3} - B^{i_2 i_4}\ \overset{-1}{G}{}^{i_6}_{i_9,i_4}\ G^{i_9}_{i_3}\ \overset{\vartriangle}{\Pi}{}^{i_1}_{i_6} \\ &:= (\mathrm{Extra\_}C)^{(i_1 i_2)\cdot}_{\cdot\cdot i_3} + (\mathrm{Extra\_}T)^{[i_1 i_2]\cdot}_{\cdot\cdot i_3} + (\mathrm{Extra\_}VP_1)^{(i_1 i_2)\cdot}_{\cdot\cdot i_3} + (\mathrm{Extra\_}VP_2)^{i_1 i_2 \cdot}_{\cdot\cdot i_3} \end{aligned}$$

such that $(C(q))$ *are scalar potential terms (functions of the configuration manifold coordinates only),* $(B(q))$ *are vector potential terms (again functions of the configuration manifold coordinates only) and* (T) *are the terms characterizing the n-symplectic gauge freedom. We emphasize that to obtain an actual system of dynamics for a mechanical system one need specify*

- *A specific vector field, by specifying the* i_2 *index.*

- *A rank 2 observable by specifying the components g_{ij} of a kinetic energy metric tensor from which the components of the inverse energy tensor g^{ij} can be found.*
- *A matrix G defining how to transform from canonical coordinate vector field to the vector fields defining the constraint distribution (see Appendix F)*
- *Any extra scalar, vector and/or gauge terms.*

Applications and implementation of this theorem to the vertical rolling hoop and a nonholonomic constrained particle are addressed in Section 4.3.

4.2.1 Decompose Soldering Form wrt Distribution

Using the coordinated local reference frame field $\underline{f} = \underline{\partial}$, eq. (C.25) is given by

$$^{c}Id_p = dq^{\mathfrak{I}} \otimes \partial_{\mathfrak{I}}|_p \iff {}^{c}\hat{\theta}_u = \overset{c}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}}(u) dq^{\mathfrak{I}}|_u \otimes r_{\mathcal{J}} \tag{4.2}$$

where $\bar{q}^{\mathfrak{I}} = q^{\mathfrak{I}} \circ \pi$ and $\overset{c}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}}(u) = \overset{c}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}}((p, \underline{e}))$ are the canonical n-symplectic momenta from Remark B.2.1 defined by

$$\overset{c}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}}(u) = e^{\mathcal{J}}(\partial_{\mathfrak{I}}|_p) \quad (u = (p, \underline{e}) = \text{arbitrary frame at p})$$

Using the distribtution adapted local reference frame field $\overset{\triangle}{\underline{F}} = (\overset{\triangle}{f}_{\mathcal{J}}) = (\overset{\triangle}{R}_A, \overset{\triangle}{F}_i)$ with corresponding distribution adapted co-frame field $\overset{\triangle}{\underline{\omega}} = (\overset{\triangle}{\omega}{}^{\mathcal{J}}) = (\overset{\triangle}{\omega}{}^{A}, \overset{\triangle}{\omega}{}^{k})$ then

$$\begin{aligned} d\bar{q}^{\mathfrak{I}}(X_u) &= \left[dq^{\mathfrak{I}} \circ d\pi\right](X|_u) \\ &= \left[\overset{\triangle}{G}{}_{\mathcal{J}}^{\mathfrak{I}}(q)\, \overset{\triangle}{\omega}{}^{\mathcal{J}}|_q\right](\pi_\star X_u) \\ &= \left[\pi^\star(\overset{\triangle}{G}{}_{\mathcal{J}}^{\mathfrak{I}}(q)\, \overset{\triangle}{\omega}{}^{\mathcal{J}}|_q)\right](X_u) \end{aligned}$$

which implies

$$d\bar{q}^{\mathfrak{I}}|_u = (\overset{\triangle}{G} \circ \pi)^{\mathfrak{I}}_{\mathcal{J}}(u)\, \overset{\triangle}{\omega}{}^{\mathcal{J}}|_u$$

where $\omega^{\mathcal{J}}|_u \stackrel{\Delta}{=} \pi^\star(\omega^{\mathcal{J}}|_q)$ and thus, from the commutation of pull-back and differential,

$$d\omega^{\mathcal{J}}|_u = \pi^\star(d\omega^{\mathcal{J}}|_q)$$

The analog of eq. (4.2) is

$$^{\Delta}Id_p = \omega^{\mathcal{J}} \otimes e_{\mathcal{J}}|_p \iff {}^{\Delta}\hat{\theta}_u = \overset{\Delta}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}}(u)\, \overset{\Delta}{\omega}{}^{\mathfrak{I}}|_u \otimes r_{\mathcal{J}} \tag{4.3}$$

where

$$\overset{\Delta}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}}(u) = e^{\mathcal{J}}(\overset{\Delta}{f}_{\mathfrak{I}}|_p) \tag{4.4}$$

and $u = (p, \underline{e})$ is an arbitrary frame at p $\in \pi^{-1}(\tilde{U}) \subset LP$.

Note 4.2.1 *(Terminology) We call* $^{\Delta}\hat{\theta}_u$ *the distribution adapted soldering one-form.*

As in section B.2, $(q^k, \overset{\Delta}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}})$ and $(q^k, \overset{c}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}})$ define distribution adapted coordinates and canonical coordinates on $\pi^{-1}(\tilde{U}) \subset LP$, $\pi^{-1}(U) \subset LP$, respecitively.

A natural question is: *What is the relationship between the canonical soldering one form and the distribution adapted soldering one form?* . Let $u = (p, \underline{E}) \in \pi^{-1}(\hat{U}) = \pi^{-1}(U) \cap \pi^{-1}(\tilde{U})$ be a point in the domain of both coordinate systems. On this common domain, the two reference frame fields transform via some $g_p \in Gl(n)$ by

$$e_{\mathfrak{I}}|_p = g_{\mathfrak{I}}^{\mathcal{J}}(p)\partial_{\mathcal{J}}|_p.$$

As determined by eq. (F.8) we have $g_p = G_p$ and thus

$$\overset{\Delta}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}}(u) = E^{\mathcal{J}}(e_{\mathfrak{I}}|_p) = E^{\mathcal{J}}(G_{\mathfrak{I}}^{\mathcal{K}}(p)\partial_{\mathcal{K}}|_p) = (G \circ \lambda)_{\mathfrak{I}}^{\mathcal{K}}(u)\, \overset{c}{\Pi}{}_{\mathcal{K}}^{\mathcal{J}}(u).$$

It follows that

$$^{-1}(\overset{\Delta}{G} \circ \pi)_{\mathcal{K}}^{\mathfrak{I}}(u)\, \overset{\Delta}{\Pi}{}_{\mathfrak{I}}^{\mathcal{J}}(u) = \overset{c}{\Pi}{}_{\mathcal{K}}^{\mathcal{J}}(u). \tag{4.5}$$

and hence eq. (4.2) becomes

$$\begin{aligned} {}^{c}\hat{\theta}_u &= \overset{\triangle}{\Pi}{}^{\mathcal{J}}_{\mathcal{I}}(u)^{-1}G^{\mathcal{I}}_{\mathcal{K}}(p)d\bar{q}^{\mathcal{K}}|_u \otimes r_{\mathcal{J}} \\ &= \overset{\triangle}{\Pi}{}^{\mathcal{J}}_{\mathcal{I}}(u)\,\overset{\triangle}{\omega}{}^{\mathcal{I}}|_u \otimes r_{\mathcal{J}} = {}^{\triangle}\hat{\theta}_u \end{aligned} \tag{4.6}$$

where $\overset{\triangle}{\omega}{}^{\mathcal{J}}|_u = {}^{-1}(\overset{\triangle}{G}\circ\pi)^{\mathcal{J}}_{\mathcal{K}}(u)d\bar{q}^{\mathcal{K}}|_u$. That is, we have written the canonical soldering form in terms of the adapted momentum coordinates.

4.2.2 Via Constrained Soldering Form Structure Equation

The claim is the n-symplectic constrained dynamics are given by

$$d\hat{g}^{i_1 i_2} = -2\,\overset{\triangle}{X}{}^{(i_2}_{\hat{g}} \lrcorner\ {}^{\triangle}\hat{\Omega}^{i_1)} \tag{4.7}$$

where $\overset{\triangle}{X}$, ${}^{\triangle}\hat{\Omega} = d^{\triangle}\hat{\theta}$, and $\hat{g}$ are adapted to the constraints via

$$\begin{aligned} \overset{\triangle}{X}{}^{i_2} &= \overset{\triangle}{X}{}^{i_2 i_3}\frac{\partial}{\partial\,\overset{\triangle}{q}{}^{i_3}} + \overset{\triangle}{X}{}^{i_2 i_4}_{i_5}\frac{\partial}{\partial\,\overset{\triangle}{\Pi}{}^{i_4}_{i_5}} \\ {}^{\triangle}\hat{\theta}^{i_1} &= \overset{\triangle}{\Pi}{}^{i_1}_{i_3}\,\overset{-1}{G}{}^{i_3}_{i_4}\,dq^{i_4} \\ {}^{\triangle}\hat{\Omega}^{i_1} &= \overset{-1}{G}{}^{i_2}_{i_3}\left[d\,\overset{\triangle}{\Pi}{}^{i_1}_{i_2}\wedge dq^{i_3}\right] + \left(\overset{-1}{G}{}^{i_2}_{i_4,i_5}\overset{\triangle}{\Pi}{}^{i_1}_{i_2}\right)\left[dq^{i_5}\wedge dq^{i_4}\right] \\ \hat{g}^{i_1 i_2} &= \overset{\triangle}{g}{}^{i_3 i_4}\,\overset{\triangle}{\Pi}{}^{i_1}_{i_3}\overset{\triangle}{\Pi}{}^{i_2}_{i_4} \quad \text{where} \quad \overset{\triangle}{g}{}^{i_3 i_4} = g^{i_5 i_6}\,\overset{-1}{G}{}^{i_3}_{i_5}\overset{-1}{G}{}^{i_4}_{i_6} \end{aligned}$$

with $g^{i_3 i_4}$ the inverse of the metric on the configuration manifold Q and G and ${}^{-1}G$ really given by $G\circ\lambda$ and ${}^{-1}(G\circ\lambda)$.

Now computing the lhs and rhs of eq. (4.7) give

$$\overset{\triangle}{X}{}^{i_2} \lrcorner\ {}^{\triangle}\hat{\Omega}^{i_1} = -\overset{-1}{G}{}^{i_4}_{i_8}\,\delta^{i_1}_{i_5}\,\overset{\triangle}{X}{}^{i_2 i_8}\,d\,\overset{\triangle}{\Pi}{}^{i_5}_{i_4} + \left(2\,\overset{-1}{G}{}^{i_7}_{[i_5,i_8]}\overset{\triangle}{X}{}^{i_2 i_8}\,\overset{\triangle}{\Pi}{}^{i_1}_{i_7} + \overset{-1}{G}{}^{i_4}_{i_5}\overset{\triangle}{X}{}^{i_2 i_1}_{i_4}\right)dq^{i_5}$$

which implies

$$-2\,\overset{\triangle}{X}{}^{(i_2} \lrcorner\; {}^{\triangle}\hat{\Omega}^{i_1)} = -2\left(-\overset{-1}{G}{}^{i_4}_{i_8}\,\delta^{(i_2}_{i_5}\,\overset{\triangle}{X}{}^{i_1)i_8}\right) d\,\overset{\triangle}{\Pi}{}^{i_5}_{i_4} -2\left(2\,\overset{-1}{G}{}^{i_7}_{[i_5,i_8]}\overset{\triangle}{\Pi}{}^{(i_2}_{i_7}\,\overset{\triangle}{X}{}^{i_1)i_8} + \overset{-1}{G}{}^{i_4}_{i_5}\overset{\triangle}{X}{}^{(i_2 i_1)}_{i_4}\right) dq^{i_5}$$

and

$$d\hat{g}^{(i_2 i_1)} = \left(\overset{\triangle}{g}{}^{i_3 i_9}_{,i_5}\,\overset{\triangle}{\Pi}{}^{(i_2}_{i_3}\,\overset{\triangle}{\Pi}{}^{i_1)}_{i_9}\right) dq^{i_5} + \left(2\,\overset{\triangle}{g}{}^{i_4 i_3}\,\delta^{(i_2}_{i_5}\,\overset{\triangle}{\Pi}{}^{i_1)}_{i_3}\right) d\,\overset{\triangle}{\Pi}{}^{i_5}_{i_4}$$

Equating coefficients gives

$$\overset{\triangle}{g}{}^{i_3 i_9}_{,i_5}\,\overset{\triangle}{\Pi}{}^{(i_2}_{i_3}\,\overset{\triangle}{\Pi}{}^{i_1)}_{i_9} = -2^2\,\overset{-1}{G}{}^{i_7}_{[i_5,i_8]}\overset{\triangle}{\Pi}{}^{(i_2}_{i_7}\,\overset{\triangle}{X}{}^{i_1)i_8} - 2\,\overset{-1}{G}{}^{i_4}_{i_5}\overset{\triangle}{X}{}^{(i_2 i_1)}_{i_4} \tag{4.8}$$

$$2\,\overset{\triangle}{g}{}^{i_4 i_3}\,\delta^{(i_2}_{i_5}\,\overset{\triangle}{\Pi}{}^{i_1)}_{i_3} = 2\,\overset{-1}{G}{}^{i_4}_{i_8}\,\delta^{(i_2}_{i_5}\,\overset{\triangle}{X}{}^{i_1)i_8} \tag{4.9}$$

We first solve eq. (4.9): Unsymmetrizing we get

$$\overset{\triangle}{g}{}^{i_4 i_3}\left(\delta^{i_1}_{i_5}\,\overset{\triangle}{\Pi}{}^{i_2}_{i_3} + \delta^{i_2}_{i_5}\,\overset{\triangle}{\Pi}{}^{i_1}_{i_5}\right) = \overset{-1}{G}{}^{i_4}_{i_8}\left(\delta^{i_2}_{i_5}\,\overset{\triangle}{X}{}^{i_1 i_8} + \delta^{i_1}_{i_5}\,\overset{\triangle}{X}{}^{i_2 i_8}\right)$$

which for $i_1 = i_5$ reduces to

$$(n+1)\,\overset{\triangle}{g}{}^{i_4 i_3}\,\overset{\triangle}{\Pi}{}^{i_2}_{i_3} = (n+1)\,\overset{-1}{G}{}^{i_4}_{i_8}\overset{\triangle}{X}{}^{i_2 i_8}$$

and thus

$$\overset{\triangle}{X}{}^{i_2 i_8} = G^{i_8}_{i_4}\,\overset{\triangle}{g}{}^{i_4 i_3}\,\overset{\triangle}{\Pi}{}^{i_2}_{i_3} \tag{4.10}$$

Now solving eq.(4.8) for the symmetrized coefficients $\overset{\triangle}{X}{}^{(i_2 i_1)}_{i_4}$ gives

$$\begin{aligned}
\overset{\triangle}{X}{}^{(i_2 i_1)}_{i_4} &= G^{i_5}_{i_4}\left[-\frac{1}{2}\,\overset{\triangle}{g}{}^{i_3 i_9}_{,i_5}\,\overset{\triangle}{\Pi}{}^{i_1}_{i_3}\overset{\triangle}{\Pi}{}^{i_2}_{i_9} - 2\,\overset{-1}{G}{}^{i_7}_{[i_5,i_8]}\overset{\triangle}{\Pi}{}^{(i_1}_{i_7}\,\overset{\triangle}{X}{}^{i_2)i_8}\right] \\
&= G^{i_5}_{i_4}\left[-\frac{1}{2}\,\overset{\triangle}{g}{}^{i_3 i_9}_{,i_5}\,\overset{\triangle}{\Pi}{}^{i_1}_{i_3}\overset{\triangle}{\Pi}{}^{i_2}_{i_9} - \overset{-1}{G}{}^{i_7}_{[i_5,i_8]}\left(\overset{\triangle}{\Pi}{}^{i_1}_{i_7}\overset{\triangle}{X}{}^{i_2 i_8} + \overset{\triangle}{\Pi}{}^{i_2}_{i_7}\overset{\triangle}{X}{}^{i_1 i_8}\right)\right]
\end{aligned} \tag{4.11}$$

We now have the equivalence class of vector fields $[\overset{\triangle}{X}_{\hat{g}}]$ given by

$$
\begin{aligned}
[\overset{\triangle}{X}{}^{i_2}_{\hat{g}}] &= \overset{\triangle}{X}{}^{i_2 i_8} \frac{\partial}{\partial \overset{\triangle}{q}{}^{i_8}} + \left(\overset{\triangle}{X}{}^{i_2 i_1}_{i_4} + T^{[i_2 i_1]}_{i_4} \right) \frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_4}} \\
&= \overset{\triangle}{X}{}^{i_2 i_8} \frac{\partial}{\partial \overset{\triangle}{q}{}^{i_8}} + \left(\overset{\triangle}{X}{}^{(i_2 i_1)}_{i_4} + \overset{\triangle}{X}{}^{[i_2 i_1]}_{i_4} + T^{[i_2 i_1]}_{i_4} \right) \frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_4}}
\end{aligned}
$$

with and unique vector field $\overset{\triangle}{X}_{\hat{g}}$ corresponding to $T^{[i_2 i_1]}_{i_4} = - \overset{\triangle}{X}{}^{[i_2 i_1]}_{i_4}$ given by

$$
\begin{aligned}
\overset{\triangle}{X}{}^{i_2}_{\hat{g}} &= \overset{\triangle}{X}{}^{i_2 i_8} \frac{\partial}{\partial \overset{\triangle}{q}{}^{i_8}} + \overset{\triangle}{X}{}^{(i_2 i_1)}_{i_4} \frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_4}} \qquad (4.12) \\
&= \left(G^{i_8}_{i_4} \overset{\triangle}{g}{}^{i_4 i_3} \overset{\triangle}{\Pi}{}^{i_2}_{i_3} \right) \frac{\partial}{\partial \overset{\triangle}{q}{}^{i_8}} + \left(G^{i_5}_{i_4} \left[-\frac{1}{2} \overset{\triangle}{g}{}^{i_3 i_9}_{,i_5} \overset{\triangle}{\Pi}{}^{(i_2}_{i_3} \overset{\triangle}{\Pi}{}^{i_1)}_{i_9} - \overset{-1}{G}{}^{i_7}_{[i_5,i_8]} \left(\overset{\triangle}{\Pi}{}^{i_1}_{i_7} \overset{\triangle}{X}{}^{i_2 i_8} + \overset{\triangle}{\Pi}{}^{i_2}_{i_7} \overset{\triangle}{X}{}^{i_1 i_8} \right) \right] \right) \frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_4}} \\
&= \left(G^{i_8}_{i_4} \overset{\triangle}{g}{}^{i_4 i_3} \overset{\triangle}{\Pi}{}^{i_2}_{i_3} \right) \frac{\partial}{\partial \overset{\triangle}{q}{}^{i_8}} + \left(G^{i_5}_{i_4} \left[-\frac{1}{2} \overset{\triangle}{g}{}^{i_3 i_9}_{,i_5} \overset{\triangle}{\Pi}{}^{(i_2}_{i_3} \overset{\triangle}{\Pi}{}^{i_1)}_{i_9} - 2 \overset{-1}{G}{}^{i_7}_{[i_5,i_8]} \overset{\triangle}{\Pi}{}^{(i_2}_{i_7} \overset{\triangle}{X}{}^{i_1) i_8} \right] \right) \frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_4}}
\end{aligned}
$$

4.2.3 Via Transform of Canonical Rank Two Solution

As addressed in Appendix G the equivalence class of solutions to

$$
d\hat{g}^{(i_2 i_1)} = -2 X^{(i_2}_{\hat{g}} \lrcorner\ \hat{\Omega}^{i_1)}
$$

for the observable expressed as $\hat{g} = g^{-1} \cdot \Pi \cdot \Pi + B \cdot \Pi + C$ or in coordinates

$$
\hat{g}^{i_2 i_1} = g^{i_5 i_6} \Pi^{(i_2}_{i_5} \Pi^{i_1)}_{i_6} + \Pi^{(i_2}_{i_8} B^{i_1) i_8} + C^{(i_2 i_1)}
$$

where $C = C(q)$ are scalar potentials and $B = B(q)$ are vector potentials is given by

$$
[X^{i_2}_{\hat{g}}] = \left(g^{i_3 i_4} \Pi^{i_2}_{i_3} + B^{i_2 i_4} \right) \frac{\partial}{\partial q^{i_4}} - \frac{1}{2} \left(g^{i_5 i_6}_{,i_7} \Pi^{(i_2}_{i_5} \Pi^{i_1)}_{i_6} + 2 \Pi^{(i_2}_{i_8} B^{i_1) i_8}_{,i_7} + 2 C^{(i_2 i_1)}_{,i_7} + T^{[i_2 i_1]}_{i_7} \right) \frac{\partial}{\partial \Pi^{i_1}_{i_7}} \qquad (4.13)
$$

where $C_,$ and $B_,$ are the derivatives of scalar and vector potential terms and T is the gauge term characterizing the equivalence class. The following transformation relations will be critical in transforming $[X^{i_2}_{\hat{g}}]$ to $[\tilde{X}^{i_2}_{\hat{g}}]$ (which will then be compared to $\overset{\vartriangle}{X}{}^{i_2}_{\hat{g}}$ in eq. (4.12)).

$$\begin{aligned}
\overset{\vartriangle}{q}{}^{i_3} &= q^{i_3} && (4.14)\\
\overset{\vartriangle}{\Pi}{}^{i_3}_{i_4} &= G^{i_5}_{i_4}\Pi^{i_3}_{i_5} && (4.15)\\
\Pi^{i_5}_{i_6} &= \overset{-1}{G}{}^{i_8}_{i_6}\overset{\vartriangle}{\Pi}{}^{i_5}_{i_8} && (4.16)\\
\overset{\vartriangle}{g}{}^{i_3 i_4} &= g^{i_5 i_6}\,\overset{-1}{G}{}^{i_3}_{i_5}\overset{-1}{G}{}^{i_4}_{i_6} && (4.17)\\
\frac{\partial}{\partial \Pi^{i_1}_{i_7}} &= G^{i_7}_{i_8}\frac{\partial}{\partial\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8}} && (4.18)\\
\frac{\partial}{\partial q^{i_4}} &= \frac{\partial}{\partial\,\overset{\vartriangle}{q}{}^{i_4}} - \overset{-1}{G}{}^{i_6}_{i_3,i_4}\,G^{i_3}_{i_8}\,\overset{\vartriangle}{\Pi}{}^{i_7}_{i_6}\,\frac{\partial}{\partial\,\overset{\vartriangle}{\Pi}{}^{i_7}_{i_8}} && (4.19)\\
\delta^{i_3}_{i_4} &= \overset{-1}{G}{}^{i_5}_{i_4}\,G^{i_3}_{i_5} && (4.20)
\end{aligned}$$

Note 4.2.2 *Transformation equation (4.19) follows from*

$$\begin{aligned}
\frac{\partial}{\partial q^k} &= B^l_k\frac{\partial}{\partial\,\overset{\vartriangle}{q}{}^l} + A^i_{kj}\frac{\partial}{\partial\,\overset{\vartriangle}{\Pi}{}^i_j}\\
\Longrightarrow\quad & B^l_k = \frac{\partial\,\overset{\vartriangle}{q}{}^l}{\partial q^k} = \delta^l_k \quad \textit{and} \quad A^i_{kj} = \frac{\partial\,\overset{\vartriangle}{\Pi}{}^i_j}{\partial q^k} = \frac{\partial(G^n_j\Pi^i_n)}{\partial q^k} = G^n_{j,k}\Pi^i_n
\end{aligned}$$

wherby eq. (4.26) $G_,$ becomes $-\overset{-1}{G}_,\, G\cdot G$ and the result is obtained. Eq. (4.18) follows similarily.

We first transform the $\partial/\partial q^{i_4}$ term:

$$g^{i_3 i_4}\Pi^{i_2}_{i_3} = g^{i_5 i_6}\delta^{i_3}_{i_5}\delta^{i_4}_{i_6}\Pi^{i_2}_{i_3} = g^{i_5 i_6}\,\overset{-1}{G}{}^{i_7}_{i_5}\,G^{i_3}_{i_7}\,\overset{-1}{G}{}^{i_8}_{i_6}\,G^{i_4}_{i_8}\Pi^{i_2}_{i_3} = \overset{\vartriangle}{g}{}^{i_7 i_8}\,\overset{\vartriangle}{\Pi}{}^{i_2}_{i_7}\,G^{i_4}_{i_8}$$

which gives

$$\left(g^{i_3i_4}\Pi^{i_2}_{i_3} + B^{i_2i_4}\right)\frac{\partial}{\partial q^{i_4}} = \left(\overset{\triangle}{g}{}^{i_7i_8}\,\overset{\triangle}{\Pi}{}^{i_2}_{i_7}\,G^{i_4}_{i_8} + B^{i_2i_4}\right)\frac{\partial}{\partial q^{i_4}} := (\overset{\triangle}{X}{}^{i_2i_4} + B^{i_2i_4})\frac{\partial}{\partial q^{i_4}} \tag{4.21}$$

$$= (\overset{\triangle}{X}{}^{i_2i_4} + B^{i_2i_4})\frac{\partial}{\partial \overset{\triangle}{q}{}^{i_4}} - G^{i_7}_{i_8}\left(\overset{-1}{G}{}^{i_5}_{i_7,i_6}\overset{\triangle}{\Pi}{}^{i_1}_{i_5}\left[\overset{\triangle}{X}{}^{i_2i_6} + B^{i_2i_6}\right]\right)\frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_8}} \tag{4.22}$$

We now transform the $\partial/\partial\Pi^{i_1}_{i_7}$ term:

$$-\frac{1}{2}\left(g^{i_5i_6}_{,i_7}\Pi^{i_2}_{i_5}\Pi^{i_1}_{i_6} + 2\Pi^{(i_2}_{i_8}B^{i_1)i_8}_{,i_7} + 2C^{i_2i_1}_{,i_7} + T^{i_2i_1}_{i_7}\right)\frac{\partial}{\partial\Pi^{i_1}_{i_7}} \tag{4.23}$$

$$= -\frac{1}{2}G^{i_7}_{i_8}\left(g^{i_5i_6}_{,i_7}\Pi^{i_2}_{i_5}\Pi^{i_1}_{i_6} + 2\Pi^{(i_2}_{i_8}B^{i_1)i_8}_{,i_7} + 2C^{i_2i_1}_{,i_7} + T^{i_2i_1}_{i_7}\right)\frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_8}} \tag{4.24}$$

where, in particular, the $g_{,} \cdot \Pi \cdot \Pi$ term becomes

$$\begin{aligned} g^{i_5i_6}_{,i_7}\Pi^{i_2}_{i_5}\Pi^{i_1}_{i_6} &= \left(\overset{\triangle}{g}{}^{i_3i_4}\,G^{i_5}_{i_3}G^{i_6}_{i_4}\right)_{,i_7}\overset{-1}{G}{}^{i_8}_{i_5}\overset{-1}{G}{}^{i_9}_{i_6}\overset{\triangle}{\Pi}{}^{i_1}_{i_8}\overset{\triangle}{\Pi}{}^{i_2}_{i_9} \\ &= \overset{\triangle}{g}{}^{i_3i_4}_{,i_7}\,\overset{\triangle}{\Pi}{}^{i_1}_{i_3}\overset{\triangle}{\Pi}{}^{i_2}_{i_4} + \overset{\triangle}{g}{}^{i_3i_4}\,\left(G^{i_5}_{i_3}G^{i_6}_{i_4}\right)_{,i_7}\overset{-1}{G}{}^{i_8}_{i_5}\overset{-1}{G}{}^{i_9}_{i_6}\overset{\triangle}{\Pi}{}^{i_1}_{i_8}\overset{\triangle}{\Pi}{}^{i_2}_{i_9}\,. \end{aligned}$$

Now the product rule again gives

$$\begin{aligned} \overset{\triangle}{g}{}^{i_3i_4}\,\left(G^{i_5}_{i_3}G^{i_6}_{i_4}\right)_{,i_7}\overset{-1}{G}{}^{i_8}_{i_5}\overset{-1}{G}{}^{i_9}_{i_6}\overset{\triangle}{\Pi}{}^{i_1}_{i_8}\overset{\triangle}{\Pi}{}^{i_2}_{i_9} &= \overset{\triangle}{g}{}^{i_3i_4}\,\left[G^{i_5}_{i_3,i_7}G^{i_6}_{i_4} + G^{i_5}_{i_3}G^{i_6}_{i_4,i_7}\right]\overset{-1}{G}{}^{i_8}_{i_5}\overset{-1}{G}{}^{i_9}_{i_6}\overset{\triangle}{\Pi}{}^{i_1}_{i_8}\overset{\triangle}{\Pi}{}^{i_2}_{i_9} \\ &= \overset{\triangle}{g}{}^{i_3i_4}\,G^{i_5}_{i_3,i_7}\,\overset{-1}{G}{}^{i_8}_{i_5}\overset{\triangle}{\Pi}{}^{i_2}_{i_4}\overset{\triangle}{\Pi}{}^{i_1}_{i_8} + \overset{\triangle}{g}{}^{i_3i_4}\,G^{i_6}_{i_4,i_7}\,\overset{-1}{G}{}^{i_9}_{i_6}\overset{\triangle}{\Pi}{}^{i_1}_{i_3}\overset{\triangle}{\Pi}{}^{2}_{i_9} \end{aligned} \tag{4.25}$$

From $0 = \partial(\delta) = \partial(\overset{-1}{G} \cdot G)$ one obtains

$$G^{i_5}_{i_3,i_7} = -\overset{-1}{G}{}^{i_9}_{i_6,i_7}\,G^{i_5}_{i_9}G^{i_6}_{i_3} \;\; and \;\; G^{i_6}_{i_4,i_7} = -\overset{-1}{G}{}^{i_8}_{i_5,i_7}\,G^{i_6}_{i_8}G^{i_5}_{i_4} \tag{4.26}$$

which can be substituted into eq. (4.25) to give

$$
\begin{aligned}
& \overset{\vartriangle}{g}{}^{i_3i_4}\left[-\overset{-1}{G}{}^{i_9}_{i_6,i_7}\, G^{i_5}_{i_9}G^{i_6}_{i_3}\right]\overset{-1}{G}{}^{i_8}_{i_5}\overset{\vartriangle}{\Pi}{}^{i_2}_{i_4}\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8} + \overset{\vartriangle}{g}{}^{i_3i_4}\left[-\overset{-1}{G}{}^{i_8}_{i_5,i_7}\, G^{i_6}_{i_8}G^{i_5}_{i_4}\right]\overset{-1}{G}{}^{i_9}_{i_6}\overset{\vartriangle}{\Pi}{}^{i_1}_{i_3}\overset{\vartriangle}{\Pi}{}^{i_2}_{i_9} \\
= \;& -\overset{\vartriangle}{g}{}^{i_3i_4}\, G^{i_6}_{i_3}\,\overset{\vartriangle}{\Pi}{}^{i_2}_{i_4}\overset{-1}{G}{}^{i_8}_{i_6,i_7}\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8} - \overset{\vartriangle}{g}{}^{i_3i_4}\, G^{i_5}_{i_4}\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_3}\overset{-1}{G}{}^{i_9}_{i_5,i_7}\overset{\vartriangle}{\Pi}{}^{i_2}_{i_9} \\
= \;& -\overset{\vartriangle}{X}{}^{i_6i_2}\,\overset{-1}{G}{}^{i_8}_{i_6,i_7}\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8} - \overset{\vartriangle}{X}{}^{i_5i_1}\,\overset{-1}{G}{}^{i_9}_{i_5,i_7}\overset{\vartriangle}{\Pi}{}^{i_2}_{i_9} \\
= \;& -\overset{-1}{G}{}^{i_8}_{i_6,i_7}\left(\overset{\vartriangle}{X}{}^{i_6i_2}\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8} + \overset{\vartriangle}{X}{}^{i_6i_1}\,\overset{\vartriangle}{\Pi}{}^{i_2}_{i_8}\right) \\
= \;& -2\,\overset{-1}{G}{}^{i_8}_{i_6,i_7}\overset{\vartriangle}{\Pi}{}^{(i_2}_{i_8}\,\overset{\vartriangle}{X}{}^{i_1)i_6} \qquad (4.27)
\end{aligned}
$$

It follows now that $[\tilde{X}^{i_2}_{\hat{g}}]$ (i.e. $[X^{i_2}_{\hat{g}}]$ transformed to distribution variables) is given by

$$
\begin{aligned}
[\tilde{X}^{i_2}_{\hat{g}}] \;=\;& (\overset{\vartriangle}{X}{}^{i_2i_4} + B^{i_2i_4})\frac{\partial}{\partial\,\overset{\vartriangle}{q}{}^{i_4}} \qquad (4.28) \\
& + G^{i_7}_{i_8}\left[-\frac{1}{2}\,\overset{\vartriangle}{g}{}^{i_3i_4}_{,i_7}\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_3}\overset{\vartriangle}{\Pi}{}^{i_2}_{i_4} + \underbrace{\overset{-1}{G}{}^{i_5}_{i_6,i_7}\overset{\vartriangle}{\Pi}{}^{(i_2}_{i_5}\,\overset{\vartriangle}{X}{}^{i_1)i_6}}_{from\ eq.(4.27)} \underbrace{-\, \overset{-1}{G}{}^{i_5}_{i_7,i_6}\overset{\vartriangle}{\Pi}{}^{i_1}_{i_5}\left[\overset{\vartriangle}{X}{}^{i_2i_6} + B^{i_2i_6}\right]}_{from\ eq.(4.22)}\right]\frac{\partial}{\partial\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8}} \qquad (4.29) \\
& + G^{i_7}_{i_8}\left[-\Pi^{(i_2}_{i_8}B^{i_1)i_8}_{,i_7} - C^{i_2i_1}_{,i_7} - \frac{1}{2}T^{i_2i_1}_{i_7}\right]\frac{\partial}{\partial\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8}} \\
=\;& (\overset{\vartriangle}{X}{}^{i_2i_4} + B^{i_2i_4})\frac{\partial}{\partial\,\overset{\vartriangle}{q}{}^{i_4}} \qquad (4.30) \\
& + G^{i_7}_{i_8}\left[-\frac{1}{2}\,\overset{\vartriangle}{g}{}^{i_3i_4}_{,i_7}\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_3}\overset{\vartriangle}{\Pi}{}^{i_2}_{i_4} + \overset{-1}{G}{}^{i_5}_{i_6,i_7}\overset{\vartriangle}{\Pi}{}^{(i_2}_{i_5}\,\overset{\vartriangle}{X}{}^{i_1)i_6} - \left(\overset{-1}{G}{}^{i_5}_{i_7,i_6}\overset{\vartriangle}{\Pi}{}^{(i_2}_{i_5}\,\overset{\vartriangle}{X}{}^{i_1)i_6} + \overset{-1}{G}{}^{i_5}_{i_7,i_6}\overset{\vartriangle}{\Pi}{}^{[i_2}_{i_5}\,\overset{\vartriangle}{X}{}^{i_1]i_6}\right)\right]\frac{\partial}{\partial\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8}} \\
& + G^{i_7}_{i_8}\left[-\,\overset{-1}{G}{}^{i_5}_{i_7,i_6}\overset{\vartriangle}{\Pi}{}^{i_1}_{i_5}\, B^{i_2i_6} - \Pi^{(i_2}_{i_8}B^{i_1)i_8}_{,i_7} - C^{i_2i_1}_{,i_7} - \frac{1}{2}T^{i_21}_{i_7}\right]\frac{\partial}{\partial\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8}} \qquad (4.31)
\end{aligned}
$$

where $\overset{\vartriangle}{X}{}^{i_2i_4} = \overset{\vartriangle}{g}{}^{i_7i_8}\,\overset{\vartriangle}{\Pi}{}^{i_2}_{i_7}\, G^{i_4}_{i_8}$. Now for $B = C = 0$ and $T^{i_2i_1}_{i_7} = -2\,\overset{-1}{G}{}^{i_5}_{i_7,i_6}\overset{\vartriangle}{\Pi}{}^{[i_2}_{i_5}\,\overset{\vartriangle}{X}{}^{i_1]i_6}$ a unique vector $\tilde{X}^{i_2}_{\hat{g}}$ is obtained,

$$
\tilde{X}^{i_2}_{\hat{g}} \;=\; \overset{\vartriangle}{X}{}^{i_2i_4}\,\frac{\partial}{\partial\,\overset{\vartriangle}{q}{}^{i_4}} + G^{i_7}_{i_8}\left[-\frac{1}{2}\,\overset{\vartriangle}{g}{}^{i_3i_4}_{,i_7}\,\overset{\vartriangle}{\Pi}{}^{1}_{i_3}\overset{\vartriangle}{\Pi}{}^{i_2}_{i_4} + 2\,\overset{-1}{G}{}^{i_5}_{[i_6,i_7]}\overset{\vartriangle}{\Pi}{}^{(i_2}_{i_5}\,\overset{\vartriangle}{X}{}^{i_1)i_6}\right]\frac{\partial}{\partial\,\overset{\vartriangle}{\Pi}{}^{i_1}_{i_8}} \qquad (4.32)
$$

4.2.4 Integral curve Equations

From eq. (4.12) we have

$$\begin{aligned}
\overset{\triangle}{X}{}^{i_2}_{\hat g} &= \overset{\triangle}{X}{}^{i_2 i_8} \frac{\partial}{\partial \overset{\triangle}{q}{}^{i_8}} + G^{i_5}_{i_4}\left[-\frac{1}{2}\,\overset{\triangle}{g}{}^{i_3 i_9}_{,i_5}\,\overset{\triangle}{\Pi}{}^{(i_2}_{i_3}\,\overset{\triangle}{\Pi}{}^{i_1)}_{i_9} -2\,\overset{-1}{G}{}^{i_7}_{[i_5,i_8]}\overset{\triangle}{\Pi}{}^{(i_2}_{i_7}\,\overset{\triangle}{X}{}^{i_1) i_8}\right]\frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_4}} \\
&= \overset{\triangle}{X}{}^{i_2 i_8} \frac{\partial}{\partial \overset{\triangle}{q}{}^{i_8}} + G^{i_5}_{i_4}\left[-\frac{1}{2}\,\overset{\triangle}{g}{}^{i_3 i_9}_{,i_5}\,\overset{\triangle}{\Pi}{}^{(i_2}_{i_3}\,\overset{\triangle}{\Pi}{}^{i_1)}_{i_9} +2\,\overset{-1}{G}{}^{i_7}_{[i_8,i_5]}\overset{\triangle}{\Pi}{}^{(i_2}_{i_7}\,\overset{\triangle}{X}{}^{i_1) i_8}\right]\frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_4}}
\end{aligned}$$

and from eq. (4.32) we have

$$\tilde{X}^{i_2}_{\hat g} = \overset{\triangle}{X}{}^{i_2 i_8} \frac{\partial}{\partial \overset{\triangle}{q}{}^{i_8}} + G^{i_5}_{i_4}\left[-\frac{1}{2}\,\overset{\triangle}{g}{}^{i_3 i_9}_{,i_5}\,\overset{\triangle}{\Pi}{}^{(i_2}_{i_3}\,\overset{\triangle}{\Pi}{}^{i_1)}_{i_9} +2\,\overset{-1}{G}{}^{i_7}_{[i_8,i_5]}\overset{\triangle}{\Pi}{}^{(i_2}_{i_7}\,\overset{\triangle}{X}{}^{i_1) i_8}\right]\frac{\partial}{\partial \overset{\triangle}{\Pi}{}^{i_1}_{i_4}}$$

where $\overset{\triangle}{X}{}^{i_2 i_8} = G^{i_8}_{i_4}\,\overset{\triangle}{g}{}^{i_4 i_3}\,\overset{\triangle}{\Pi}{}^{i_2}_{i_3}$. The above vector fields are equal whereby we take this to mean that they been double checked to be correct.

The integral curve equations for the vector fields $\overset{\triangle}{X}{}^{i_2}_{\hat g} = \tilde{X}^{i_2}_{\hat g}$ are

$$\begin{aligned}
\dot{q}^{i_2 i_4} &= \overset{\triangle}{X}{}^{i_2 i_4} = \overset{\triangle}{g}{}^{i_7 i_5}\,\overset{\triangle}{\Pi}{}^{i_2}_{i_5}\,G^{i_4}_{i_7} \\
\dot{\overset{\triangle}{\Pi}}{}^{i_1 i_2}_{\cdot\cdot i_3} &= G^{i_6}_{i_3}\left[-\frac{1}{2}\,\overset{\triangle}{g}{}^{i_7 i_5}_{,i_6}\,\overset{\triangle}{\Pi}{}^{i_1}_{i_7}\overset{\triangle}{\Pi}{}^{i_2}_{i_5} +2\,\overset{-1}{G}{}^{i_8}_{[i_4,i_6]}\overset{\triangle}{\Pi}{}^{(i_1}_{i_8}\,\overset{\triangle}{X}{}^{i_2) i_4}\right]
\end{aligned} \tag{4.33}$$

Remark 4.2.1 *If we want to include scalar potential terms ($C(q)$), vector potential terms ($B(q)$) and extra gauge terms (T) we would add to the above equations (following eq. (4.31))*

$$\begin{aligned}
\text{Extra}^{i_2 i_4} &= B^{i_2 i_4} \\
\text{Extra}^{i_1 i_2 \cdot}_{\cdot\cdot i_3} &= G^{i_7}_{i_3} C^{(i_1 i_2)}_{,i_7} + G^{i_7}_{i_3} T^{[i_1 i_2]}_{i_7} + \overset{-1}{G}{}^{i_5}_{i_8}\overset{\triangle}{\Pi}{}^{(i_1}_{i_5}\; B^{i_2) i_8}_{,i_7} G^{i_7}_{i_3} - B^{i_2 i_4}\,\overset{-1}{G}{}^{i_6}_{i_9,i_4}\,G^{i_9}_{i_3}\,\overset{\triangle}{\Pi}{}^{i_1}_{i_6} \\
&:= (\text{Extra_}C)^{(i_1 i_2)\cdot}_{\cdot\cdot i_3} + (\text{Extra_}T)^{[i_1 i_2]\cdot}_{\cdot\cdot i_3} + (\text{Extra_}VP_1)^{(i_1 i_2)\cdot}_{\cdot\cdot i_3} + (\text{Extra_}VP_2)^{i_1 i_2 \cdot}_{\cdot\cdot i_3}
\end{aligned}$$

One can, for example, choose $i_2 = (A = 1)$ which selects the $\overset{\triangle}{X}{}^{i_2=(A=1)}$ vector field to find the integral curves for. This choice leads to , after splitting eq. (4.33) into $i_1 = (i, 1, \bar{A})$ pieces, the equations

$$\dot{q}^{i_4} = \overset{\triangle}{g}^{i_7 i_5} \overset{\triangle}{\Pi}{}^{A=1}_{i_5} G^{i_4}_{i_7} \tag{4.34a}$$

$$\dot{\overset{\triangle}{\Pi}}{}^{A=1}_{i_3} = G^{i_6}_{i_3}\left[-\frac{1}{2}\overset{\triangle}{g}{}^{i_7 i_5}_{,i_6} \overset{\triangle}{\Pi}{}^{A=1}_{i_7} \overset{\triangle}{\Pi}{}^{A=1}_{i_5} + 2\overset{-1}{G}{}^{i_8}_{[i_4,i_6]} \overset{\triangle}{\Pi}{}^{A=1}_{i_8} \dot{q}^{i_4}\right] \tag{4.34b}$$

$$\dot{\overset{\triangle}{\Pi}}{}^{\bar{A}}_{i_3} = G^{i_6}_{i_3}\left[-\frac{1}{2}\overset{\triangle}{g}{}^{i_7 i_5}_{,i_6} \overset{\triangle}{\Pi}{}^{\bar{A}}_{i_7} \overset{\triangle}{\Pi}{}^{A=1}_{i_5} + \overset{-1}{G}{}^{i_8}_{[i_4,i_6]}\left(\overset{\triangle}{\Pi}{}^{\bar{A}}_{i_8} \dot{q}^{i_4} + \overset{\triangle}{\Pi}{}^{A=1}_{i_8} \overset{\triangle}{X}{}^{\bar{A} i_4}\right)\right] \tag{4.34c}$$

$$\dot{\overset{\triangle}{\Pi}}{}^{i}_{i_3} = G^{i_6}_{i_3}\left[-\frac{1}{2}\overset{\triangle}{g}{}^{i_7 i_5}_{,i_6} \overset{\triangle}{\Pi}{}^{i}_{i_7} \overset{\triangle}{\Pi}{}^{A=1}_{i_5} + \overset{-1}{G}{}^{i_8}_{[i_4,i_6]}\left(\overset{\triangle}{\Pi}{}^{i}_{i_8} \dot{q}^{i_4} + \overset{\triangle}{\Pi}{}^{A=1}_{i_8} \overset{\triangle}{X}{}^{i i_4}\right)\right] \tag{4.34d}$$

where $\overset{\triangle}{X}{}^{i i_4} = \overset{\triangle}{g}{}^{i_7 i_5} \overset{\triangle}{\Pi}{}^{i}_{i_5} G^{i_4}_{i_7}$ and $\overset{\triangle}{X}{}^{\bar{A} i_4} = \overset{\triangle}{g}{}^{i_7 i_5} \overset{\triangle}{\Pi}{}^{\bar{A}}_{i_5} G^{i_4}_{i_7}$.

Remark 4.2.2 *The extra terms above are split amongst the above momentum equations as*

$$\mathrm{Extra}^{A=1 i_4} = B^{A=1 i_4} \tag{4.35a}$$

$$\begin{aligned}\mathrm{Extra}^{A=1A=1\cdot}_{\cdot\cdot i_3} &= (\mathrm{Extra\_}C)^{A=1A=1\cdot}_{\cdot\cdot i_3} + \overbrace{(\mathrm{Extra\_}T)^{[A=1A=1]\cdot}_{\cdot\cdot i_3}}^{=0} \\ &\quad + (\mathrm{Extra\_}VP_1)^{A=1A=1\cdot}_{\cdot\cdot i_3} + (\mathrm{Extra\_}VP_2)^{A=1A=1\cdot}_{\cdot\cdot i_3}\end{aligned} \tag{4.35b}$$

$$\begin{aligned}\mathrm{Extra}^{\bar{A}A=1\cdot}_{\cdot\cdot i_3} &= (\mathrm{Extra\_}C)^{\bar{A}A=1\cdot}_{\cdot\cdot i_3} + (\mathrm{Extra\_}T)^{\bar{A}A=1\cdot}_{\cdot\cdot i_3} \\ &\quad + (\mathrm{Extra\_}VP_1)^{\bar{A}A=1\cdot}_{\cdot\cdot i_3} + (\mathrm{Extra\_}VP_2)^{\bar{A}A=1\cdot}_{\cdot\cdot i_3}\end{aligned} \tag{4.35c}$$

$$\begin{aligned}\mathrm{Extra}^{iA=1\cdot}_{\cdot\cdot i_3} &= (\mathrm{Extra\_}C)^{iA=1\cdot}_{\cdot\cdot i_3} + (\mathrm{Extra\_}T)^{iA=1\cdot}_{\cdot\cdot i_3} \\ &\quad + (\mathrm{Extra\_}VP_1)^{iA=1\cdot}_{\cdot\cdot i_3} + (\mathrm{Extra\_}VP_2)^{iA=1\cdot}_{\cdot\cdot i_3}\end{aligned} \tag{4.35d}$$

4.3 n-Symplectic Dynamics Examples

4.3.1 Vertical Rolling Hoop

Section Summary 4.3.1 *(N-symplectic Form of the Dynamics of a Vertical Rolling Hoop) For the choices of potentials $C^{43}(t, x, y) = 0$ and C^{33}, C^{13} and C^{23} given by* *equations*

$$
\begin{aligned}
C^{33}(t,x,y) &= K\left(-\cos\left(\Omega t+\phi_0\right)y+\sin\left(\Omega t+\phi_0\right)x\right) \text{ where } K=-mR\omega\Omega \\
C^{13}(t,x,y) &= \left(-\overset{\triangle}{\Pi}{}^{1}_{2}(0)x+\overset{\triangle}{\Pi}{}^{1}_{1}(0)y\right)\frac{\Omega}{2}\cos\left(\frac{\Omega}{2}t+\phi_0\right) \\
&\qquad +\left(-\overset{\triangle}{\Pi}{}^{1}_{1}(0)x-\overset{\triangle}{\Pi}{}^{1}_{2}(0)y\right)\frac{\Omega}{2}\sin\left(\frac{\Omega}{2}t+\phi_0\right) \\
C^{23}(t,x,y) &= \left(-\overset{\triangle}{\Pi}{}^{2}_{2}(0)x+\overset{\triangle}{\Pi}{}^{2}_{1}(0)y\right)\frac{\Omega}{2}\cos\left(\frac{\Omega}{2}t+\phi_0\right) \\
&\qquad +\left(-\overset{\triangle}{\Pi}{}^{2}_{1}(0)x-\overset{\triangle}{\Pi}{}^{2}_{2}(0)y\right)\frac{\Omega}{2}\sin\left(\frac{\Omega}{2}t+\phi_0\right)
\end{aligned}
$$

and for the specific n-symplectic momenta $\overset{\triangle}{\Pi}{}^{i}_{j}$ *given by*

$$
\begin{aligned}
\overset{\triangle}{\Pi}{}^{1}_{1}(t) &= \overset{\triangle}{\Pi}{}^{1}_{1}(0)\cos\left(\frac{\Omega}{2}t+\phi_0\right)-\overset{\triangle}{\Pi}{}^{1}_{2}(0)\sin\left(\frac{\Omega}{2}t+\phi_0\right) \\
\overset{\triangle}{\Pi}{}^{1}_{2}(t) &= \overset{\triangle}{\Pi}{}^{1}_{2}(0)\cos\left(\frac{\Omega}{2}t+\phi_0\right)+\overset{\triangle}{\Pi}{}^{1}_{1}(0)\sin\left(\frac{\Omega}{2}t+\phi_0\right) \\
\overset{\triangle}{\Pi}{}^{2}_{1}(t) &= \overset{\triangle}{\Pi}{}^{2}_{1}(0)\cos\left(\frac{\Omega}{2}t+\phi_0\right)-\overset{\triangle}{\Pi}{}^{2}_{2}(0)\sin\left(\frac{\Omega}{2}t+\phi_0\right) \\
\overset{\triangle}{\Pi}{}^{2}_{2}(t) &= \overset{\triangle}{\Pi}{}^{2}_{2}(0)\cos\left(\frac{\Omega}{2}t+\phi_0\right)+\overset{\triangle}{\Pi}{}^{2}_{1}(0)\sin\left(\frac{\Omega}{2}t+\phi_0\right)
\end{aligned}
$$

where $\overset{\triangle}{\Pi}{}^{i}_{j}(0)$ *are the initial momenta satisfying* $\overset{\triangle}{\Pi}{}^{1}_{1}(0)\overset{\triangle}{\Pi}{}^{2}_{2}(0)-\overset{\triangle}{\Pi}{}^{1}_{2}(0)\overset{\triangle}{\Pi}{}^{2}_{1}(0)\neq 0$*, the n-symplectic momenta along with the consistent system of n-symplectic momenta dynamics are given by*

$$
\left[\overset{\triangle}{\Pi}{}^{\mathfrak{I}}_{\mathfrak{J}}(t)\right]=\begin{pmatrix}
\overset{\triangle}{\Pi}{}^{1}_{1}(t) & \overset{\triangle}{\Pi}{}^{1}_{2}(t) & 0 & 0 \\
\overset{\triangle}{\Pi}{}^{2}_{1}(t) & \overset{\triangle}{\Pi}{}^{2}_{2}(t) & 0 & 0 \\
mR\cos(\phi)\dot{\theta} & mR\sin(\phi)\dot{\theta} & (m+M_1R^2)\dot{\theta} & M_2\dot{\phi} \\
\overset{\triangle}{\Pi}{}^{4}_{1}(t) & \overset{\triangle}{\Pi}{}^{4}_{2}(t) & \overset{\triangle}{\Pi}{}^{4}_{3}(t) & \overset{\triangle}{\Pi}{}^{4}_{4}(t)
\end{pmatrix}
$$

and

$$\left[\frac{d}{dt}\overset{\triangle}{\Pi}{}^{\mathcal{J}}_{\mathcal{J}}(t)\right] = \begin{pmatrix} C^{13}_{,3} & C^{13}_{,4} & 0 & 0 \\ C^{23}_{,3} & C^{23}_{,4} & 0 & 0 \\ C^{33}_{,3} & C^{33}_{,4} & 0 & 0 \\ 0 & 0 & \frac{-R}{2}\dot{\phi}\left(\overset{\triangle}{\Pi}{}^{4}_{1}(t)\sin(\phi) - \overset{\triangle}{\Pi}{}^{4}_{2}(t)\cos(\phi)\right) & 0 \end{pmatrix}$$

where $\dot{\theta}\dot{\phi} = \omega\Omega$ *which implies* $\theta = \theta(t) = \omega t + \theta_0$, $\phi = \phi(t) = \Omega t + \phi_0$ *and*

$$\dot{x} - R\cos(\phi)\dot{\theta} = 0 \quad \textit{and} \quad \dot{y} - R\sin(\phi)\dot{\theta} = 0.$$

Remark 4.3.1 *It is not unexpected that the n-symplectic dynamics (on* $L_\triangle Q$*) are related to the nonholonomic momentum dynamics of [3] (see Appendix H)*

$$\frac{d}{dt}J_A(q) - \Gamma^C_{A\mathcal{J}}J_C(q)\dot{q}^{\mathcal{J}} = 0$$

since the n-symplectic generalized momentum dynamics (cf. eq. 4.1) are

$$\dot{\overset{\triangle}{\Pi}}{}^{i_1 i_2}_{\cdot\cdot i_3} - 2G^{i_6}_{i_3}\overset{-1}{G}{}^{i_8}_{[i_4,i_6]}\overset{\triangle}{\Pi}{}^{(i_1}_{i_8}\,\dot{q}^{i_2)i_4} + \frac{1}{2}G^{i_6}_{i_3}\overset{\triangle}{g}{}^{i_7 i_5}_{,i_6}\overset{\triangle}{\Pi}{}^{i_1}_{i_7}\overset{\triangle}{\Pi}{}^{i_2}_{i_5} = 0$$

The two sets of equations are structurally *similar with* $\overset{\triangle}{\Pi}$ *playing the role of* J *and with* $G \cdot \partial G$ *playing the role of* Γ*.*

The dynamics for a vertical rolling hoop are addressed from both the Lagrange-d'-Alembert and the nonholonomic momenta (both spatial and reduced) perspectives in [3]. One finds that for the hoop Lagrangian

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\phi}^2 \tag{4.36}$$

and the constraint 1-forms (i.e. the basis vectors of the constraints codistribution

$\Delta^\star$)

$$
\begin{aligned}
\omega^1 &= dx - R\cos(\phi)d\theta \\
\omega^2 &= dy - R\sin(\phi)d\theta
\end{aligned}
$$

which define the (rolling) constraint vectors (i.e. the basis vectors of the constraint distribution Δ)

$$
\begin{aligned}
\overset{\vartriangle}{F}_1 &= \partial_\theta + R\cos(\phi)\partial_x + R\sin(\phi)\partial_y \\
\overset{\vartriangle}{F}_2 &= \partial_\phi
\end{aligned}
$$

the dynamics from the (constrained) Lagrange-d'-Alembert principle are

$$
\begin{aligned}
(I + mR^2)\ddot{\theta} &= 0 \\
J\ddot{\phi} &= 0 \\
\dot{x} &= R\cos(\phi)\dot{\theta} \\
\dot{y} &= R\sin(\phi)\dot{\theta}.
\end{aligned}
$$

while from the non-holonomic momentum (spatial) perspective (see also Appendix H) the dynamics are equivalently

$$
\begin{aligned}
\dot{x} &= R\cos(\phi)\dot{\theta} \\
\dot{y} &= R\sin(\phi)\dot{\theta} \\
\dot{J}_1 &= 0 \\
\dot{J}_3 &= 0 \\
J_1 &= (I + mR^2)\dot{\theta} \\
J_3 &= J\dot{\phi}
\end{aligned}
$$

We now approach the rolling hoop from an n-symplectic perspective by utilizing equations (4.34). As stated in Theorem 4.2.1, the inputs to the general equations (4.1) required to obtain a specific set of dynamics equations are

- A specific vector field X^{i_2}: In this case $i_2 = 3 = (A = 1)$ where the variables are indexed as

$$[\{q^{i_1}\}] = [\{r^i\}, \{s^A\}] = [r^1, r^2, s^1, s^2] = [\theta, \phi, x, y]$$

  A justification for this choice of vector field is given in Section 4.3.3.

- Mass Matrix:

$$g = \begin{pmatrix} M_1 & 0 & 0 & 0 \\ 0 & M_2 & 0 & 0 \\ 0 & 0 & m & 0 \\ 0 & 0 & 0 & m \end{pmatrix}$$

- Transformation from Canonical Basis $\underline{\partial}$ to Disbribution Basis $\underline{\overset{\vartriangle}{F}}$ (cf. Appendix F equation (F.8)):

$$\overset{\vartriangle}{G} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & R\cos(\phi) & 0 \\ 0 & 1 & R\sin(\phi) & 0 \end{pmatrix}$$

- Both g^{-1} and $\overset{\vartriangle}{G}^{-1}$ can be found from the previous two points.

One can but is not required to utilize the full scalar potential (C), the full vector potential (B) or the full gauge terms (T) which will add terms of the form in eq. (4.35). To obtain the vertical rolling hoop equations we will need select scalar potential terms. For the time being we disregard the vector potential and gauge terms, though it is a source of future work to consider the contribution of these terms to the dynamics (with momenta based control specifically in mind). With these required inputs and the extra scalar potential functions, the dynamic of eqns.(4.34) can be written as

$$\left(\overset{\triangle}{\Pi}{}^{\mathcal{J}}_{\mathcal{J}}(t)\right) = \begin{pmatrix} \overset{\triangle}{\Pi}{}^{1}_{1} & \overset{\triangle}{\Pi}{}^{1}_{2} & 0 & 0 \\ \overset{\triangle}{\Pi}{}^{2}_{1} & \overset{\triangle}{\Pi}{}^{2}_{2} & 0 & 0 \\ mR\cos(\phi)\dot{\theta} & mR\sin(\phi)\dot{\theta} & (m+M_1R^2)\dot{\theta} & M_2\dot{\phi} \\ \overset{\triangle}{\Pi}{}^{4}_{1} & \overset{\triangle}{\Pi}{}^{4}_{2} & \overset{\triangle}{\Pi}{}^{4}_{3} & \overset{\triangle}{\Pi}{}^{4}_{4} \end{pmatrix} \tag{4.37}$$

and

$$\left(\frac{d}{dt}\overset{\triangle}{\Pi}{}^{\mathcal{J}}_{\mathcal{J}}(t)\right) = \tag{4.38}$$

$$\begin{pmatrix} C^{13}_{,x} & C^{13}_{,y} & \frac{R}{2M_2}\overset{\triangle}{\Pi}{}^{3}_{4}\left[\left(-\overset{\triangle}{\Pi}{}^{1}_{1}+\frac{2M_2}{\overset{\triangle}{\Pi}{}^{3}_{4}}C^{13}_{,y}\right)\sin(\phi)+\left(\overset{\triangle}{\Pi}{}^{1}_{2}+\frac{2M_2}{\overset{\triangle}{\Pi}{}^{3}_{4}}C^{13}_{,x}\right)\cos(\phi)\right]+C^{13}_{,\theta} & C^{13}_{,\phi} \\ C^{23}_{,x} & C^{23}_{,y} & \frac{R}{2M_2}\overset{\triangle}{\Pi}{}^{3}_{4}\left[\left(-\overset{\triangle}{\Pi}{}^{2}_{1}+\frac{2M_2}{\overset{\triangle}{\Pi}{}^{3}_{4}}C^{23}_{,y}\right)\sin(\phi)+\left(\overset{\triangle}{\Pi}{}^{2}_{2}+\frac{2M_2}{\overset{\triangle}{\Pi}{}^{3}_{4}}C^{23}_{,x}\right)\cos(\phi)\right]+C^{23}_{,\theta} & C^{23}_{,\phi} \\ C^{33}_{,x} & C^{33}_{,y} & C^{33}_{,x}\cdot R\cos(\phi)+C^{33}_{,y}\cdot R\sin(\phi)+C^{33}_{,\theta} & C^{33}_{,\phi} \\ C^{43}_{,x} & C^{43}_{,y} & \frac{R}{2M_2}\overset{\triangle}{\Pi}{}^{3}_{4}\left[\left(-\overset{\triangle}{\Pi}{}^{4}_{1}+\frac{2M_2}{\overset{\triangle}{\Pi}{}^{3}_{4}}C^{43}_{,y}\right)\sin(\phi)+\left(\overset{\triangle}{\Pi}{}^{4}_{2}+\frac{2M_2}{\overset{\triangle}{\Pi}{}^{3}_{4}}C^{43}_{,x}\right)\cos(\phi)\right]+C^{43}_{,\theta} & C^{43}_{,\phi} \end{pmatrix}$$

where $C^{33}_{,x}$ and $C^{33}_{,y}$ must satisfy the consistency equations

$$\begin{aligned} -mR\sin(\phi)\dot{\theta}\dot{\phi}+mR\cos(\phi)\ddot{\theta}-C^{33}_{,x} &= 0 \\ mR\cos(\phi)\dot{\theta}\dot{\phi}+mR\sin(\phi)\ddot{\theta}-C^{33}_{,y} &= 0 \end{aligned}$$

and the two constraint equations are

$$\begin{aligned} \dot{x}-R\cos(\phi)\dot{\theta} &= 0 \\ \dot{y}-R\sin(\phi)\dot{\theta} &= 0. \end{aligned}$$

One can see that $\overset{\triangle}{\Pi}{}^{3}_{3}=(M_1+mR^2)\dot{\theta}$ is the first non-holonomic momentum and that $\overset{\triangle}{\Pi}{}^{3}_{4}=M_2\dot{\phi}$ is the second non-holonomic momentum.

Assumption 4.3.1 *($C^{\partial 3}$ functional restriction) We assume that*

$$C^{13}_{,\theta} = C^{13}_{,\phi} = C^{23}_{,\theta} = C^{23}_{,\phi} = C^{33}_{,\theta} = C^{33}_{,\phi} = C^{43}_{,\theta} = C^{43}_{,\phi} = 0 \tag{4.39}$$

and thus we assume the function dependence

$$C^{\partial 3}(t, x, y). \tag{4.40}$$

From assumption 4.3.1, $\frac{d}{dt} \overset{\vartriangle}{\Pi}{}^{3}_{4} = 0$ which indicates conservation of the second nonholonomic momentum $\overset{\vartriangle}{\Pi}{}^{3}_{4}$ and that

$$\phi(t) = \Omega t + \phi_0.$$

The presence of potential terms seem to indicate that the first nonholonomic momentum $\overset{\vartriangle}{\Pi}{}^{3}_{3}$ is *not* conserved. However, one can solve the PDE

$$R\left(\left(\frac{\partial}{\partial x} C^{33}\,(x,y)\right)\cos\left(\Omega t + \phi_0\right) + \left(\frac{\partial}{\partial y} C^{33}\,(x,y)\right)\sin\left(\Omega t + \phi_0\right)\right) = 0$$

to find

$$C^{33}(t,x,y) = K\left(-y\cos\left(\Omega t + \phi_0\right) + \sin\left(\Omega t + \phi_0\right)x\right) \tag{4.41}$$

where K is any constant. That is, for $C^{33}(t,x,y)$ as given the nonholonomic momentum $\overset{\vartriangle}{\Pi}{}^{3}_{3}$ is conserved. A natural question is why introduce scalar potentials at all for without them $\overset{\vartriangle}{\Pi}{}^{3}_{3}$ would have been immediately conserved? The lack of scalar potential terms cause problems at the consistency equation level as follows: With scalar potentials as above or without them, that $\overset{\vartriangle}{\Pi}{}^{3}_{1}$ is conserved implies that $\ddot{\theta}$ is 0 which lead to the consistency equations

$$\begin{aligned} -mR\sin(\phi)\dot{\theta}\dot{\phi} &= C^{33}_{,x} \\ mR\cos(\phi)\dot{\theta}\dot{\phi} &= C^{33}_{,y} \end{aligned} \tag{4.42}$$

Were $C^{33}(t,x,y)=0$ (i.e. no scalar potentials present) then the only solution to the consistency equation given by eq. (4.42) would be $\dot{\phi}=0$ (as $\dot{\theta}\neq 0$ or else there would be no *rolling*). That is, the hoop motion would be only in a straight line. And while this is possible it is not the full story. So scalar potential terms need to be present to recover all of the physics of the rolling hoop. To maintain conservation of one of the nonholonomic momenta, $C^{33}(t,x,y)$ must be given by eq. (4.41) and which, upon substitution into eq. (4.42), give

$$\dot{\theta}\dot{\phi}=\frac{K}{-mR}$$

where from conservation of the two momenta, $\dot{\theta}$ and $\dot{\phi}$ are constants denoted say ω and Ω. As K can be any constant then let it be $K=-mR\omega\Omega$. We have thus recovered in the larger context of n-symplectic dynamics the dynamics of the vertical rolling hoop (an example of a non-holonomic mechanical system with symmetry).

There is, however, another problem arising from the n-symplectic formulation, namely that $\frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{\mathcal{I}}_{\mathcal{J}}$ does not have the [1,3] and [2,3] entries 0 and so the $L_{\vartriangle}Q$ condition which ensures that the dynamics satisfy the constraints for all t is not met (see Appendix F for $L_{\vartriangle}Q$ details). The $L_{\vartriangle}Q$ condition can be met can by solving the equations

$$-\overset{\vartriangle}{\Pi}{}^{1}_{1}+\frac{2M_2}{\overset{\vartriangle}{\Pi}{}^{3}_{4}}C^{13}_{,y}=\lambda\cos(\Omega t+\phi_0)\quad\text{and}\quad \overset{\vartriangle}{\Pi}{}^{1}_{2}+\frac{2M_2}{\overset{\vartriangle}{\Pi}{}^{3}_{4}}C^{13}_{,x}=-\lambda\sin(\Omega t+\phi_0)\tag{4.43}$$

$$-\overset{\vartriangle}{\Pi}{}^{1}_{2}+\frac{2M_2}{\overset{\vartriangle}{\Pi}{}^{3}_{4}}C^{23}_{,y}=\lambda\cos(\Omega t+\phi_0)\quad\text{and}\quad \overset{\vartriangle}{\Pi}{}^{2}_{2}+\frac{2M_2}{\overset{\vartriangle}{\Pi}{}^{3}_{4}}C^{23}_{,x}=-\lambda\sin(\Omega t+\phi_0)\tag{4.44}$$

for $\overset{\vartriangle}{\Pi}{}^{\mathcal{I}}_{\mathcal{J}}(t)$ where these solutions must also satisfy the consistency equations

$$\begin{aligned}\frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{1}_{1}&=C^{13}_{,x}\quad\text{and}\quad \frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{1}_{2}=C^{13}_{,y}\\ \frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{2}_{1}&=C^{23}_{,x}\quad\text{and}\quad \frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{2}_{2}=C^{23}_{,y}\end{aligned}\tag{4.45}$$

Since $\frac{2M_2}{\overset{\triangle}{\Pi}{}^3_4} = \frac{2}{\dot{\phi}}$ where $\dot{\phi} = \Omega = const.$ then we need to solve the linear system

$$\begin{pmatrix} \frac{d}{dt}\overset{\triangle}{\Pi}{}^1_1 \\ \frac{d}{dt}\overset{\triangle}{\Pi}{}^1_2 \\ \frac{d}{dt}\overset{\triangle}{\Pi}{}^2_1 \\ \frac{d}{dt}\overset{\triangle}{\Pi}{}^2_2 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{\Omega}{2} & 0 & 0 \\ \frac{\Omega}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{\Omega}{2} \\ 0 & 0 & \frac{\Omega}{2} & 0 \end{pmatrix} \begin{pmatrix} \overset{\triangle}{\Pi}{}^1_1 \\ \overset{\triangle}{\Pi}{}^1_2 \\ \overset{\triangle}{\Pi}{}^2_1 \\ \overset{\triangle}{\Pi}{}^2_2 \end{pmatrix} + \begin{pmatrix} -\lambda\frac{\Omega}{2}\sin(\Omega t + \phi_0) \\ \lambda\frac{\Omega}{2}\cos(\Omega t + \phi_0) \\ -\lambda\frac{\Omega}{2}\sin(\Omega t + \phi_0) \\ \lambda\frac{\Omega}{2}\cos(\Omega t + \phi_0) \end{pmatrix} \tag{4.46}$$

We focus on the case $\lambda = 0$. The solutions then are

$$\begin{aligned}
\overset{\triangle}{\Pi}{}^1_1(t) &= \overset{\triangle}{\Pi}{}^1_1(0)\cos\left(\frac{\Omega}{2}t + \phi_0\right) - \overset{\triangle}{\Pi}{}^1_2(0)\sin\left(\frac{\Omega}{2}t + \phi_0\right) \\
\overset{\triangle}{\Pi}{}^1_2(t) &= \overset{\triangle}{\Pi}{}^1_2(0)\cos\left(\frac{\Omega}{2}t + \phi_0\right) + \overset{\triangle}{\Pi}{}^1_1(0)\sin\left(\frac{\Omega}{2}t + \phi_0\right) \\
\overset{\triangle}{\Pi}{}^2_1(t) &= \overset{\triangle}{\Pi}{}^2_1(0)\cos\left(\frac{\Omega}{2}t + \phi_0\right) - \overset{\triangle}{\Pi}{}^2_2(0)\sin\left(\frac{\Omega}{2}t + \phi_0\right) \\
\overset{\triangle}{\Pi}{}^2_2(t) &= \overset{\triangle}{\Pi}{}^2_2(0)\cos\left(\frac{\Omega}{2}t + \phi_0\right) + \overset{\triangle}{\Pi}{}^2_1(0)\sin\left(\frac{\Omega}{2}t + \phi_0\right)
\end{aligned} \tag{4.47}$$

where $\overset{\triangle}{\Pi}{}^{\mathcal{J}}_{\mathcal{J}}(0)$ are the intial momenta and are required to satisfy

$$\overset{\triangle}{\Pi}{}^1_1(0)\,\overset{\triangle}{\Pi}{}^2_2(0) - \overset{\triangle}{\Pi}{}^1_2(0)\,\overset{\triangle}{\Pi}{}^2_1(0) \neq 0$$

The consistency equations given by eq. (4.45) are satisfied by choosing

$$\begin{aligned}
C^{13}(t,x,y) = &\left(-\overset{\triangle}{\Pi}{}^1_2(0)x + \overset{\triangle}{\Pi}{}^1_1(0)y\right)\frac{\Omega}{2}\cos\left(\frac{\Omega}{2}t + \phi_0\right) \\
&+ \left(-\overset{\triangle}{\Pi}{}^1_1(0)x - \overset{\triangle}{\Pi}{}^1_2(0)y\right)\frac{\Omega}{2}\sin\left(\frac{\Omega}{2}t + \phi_0\right)
\end{aligned} \tag{4.48}$$

$$\begin{aligned}
C^{23}(t,x,y) = &\left(-\overset{\triangle}{\Pi}{}^2_2(0)x + \overset{\triangle}{\Pi}{}^2_1(0)y\right)\frac{\Omega}{2}\cos\left(\frac{\Omega}{2}t + \phi_0\right) \\
&+ \left(-\overset{\triangle}{\Pi}{}^2_1(0)x - \overset{\triangle}{\Pi}{}^2_2(0)y\right)\frac{\Omega}{2}\sin\left(\frac{\Omega}{2}t + \phi_0\right)
\end{aligned} \tag{4.49}$$

It is a matter of future work to determine a meaning and use of extra scalar potential

C, vector potential B and gauge T terms in, for example, potential shaping and energy conserving control of the hoop.

4.3.2 A Nonholonomic Constrained Particle

Section Summary 4.3.2 *(N-symplectic Form of the Dynamics of a Nonholonomic Particle) For the choices of potentials $C^{13} = C^{23} = 0$ and for*

$$\begin{aligned} C^{33}(t,x,z) &= \frac{Q \cdot A\left[z - x \cdot (At+B)\right]}{\left[1+(At+B)^2\right]^{3/2}} \\ C^{23}(t,y) &= -\frac{y}{2}\left[\frac{(At+B)\cdot Q \cdot K}{\sqrt{1+(At+B)^2}}\right] \end{aligned}$$

where $A, B,$ and Q are arbitrary constants and $K = \overset{\triangle}{\Pi}{}^2_2(0) \neq 0$ (and $\overset{\triangle}{\Pi}{}^2_1(0) = 0$), the n-symplectic momenta along with a consistent system of n-symplectic momenta dynamics are given by

$$\left(\overset{\triangle}{\Pi}{}^{\mathfrak{I}}_{\mathfrak{J}}(t)\right) = \begin{pmatrix} \overset{\triangle}{\Pi}{}^1_1 & \overset{\triangle}{\Pi}{}^1_2 & 0 \\ \overset{\triangle}{\Pi}{}^2_1 & \overset{\triangle}{\Pi}{}^2_2 & 0 \\ y\dot{x} & \dot{x}(1+y^2) & \dot{y} \end{pmatrix}$$

and

$$\left(\frac{d}{dt}\overset{\triangle}{\Pi}{}^{\mathfrak{I}}_{\mathfrak{J}}(t)\right) = \begin{pmatrix} 0 & \frac{1}{2}\dot{y}\,\overset{\triangle}{\Pi}{}^1_1 & 0 \\ 0 & 0 & 0 \\ \frac{Q\cdot A}{[1+(At+B)^2]^{3/2}} & y\dot{x}\dot{y} & 0 \end{pmatrix}$$

where the constraint equation is given by

$$\dot{z} - y\dot{x} = 0.$$

These first order dynamics not only build the second order dynamics

$$\ddot{x} = \frac{-y\dot{x}\dot{y}}{1+y^2} \quad \textit{and} \quad \ddot{y} = 0$$

but also contain the consistency equation

$$\dot{x} = \frac{Q}{\sqrt{1+y^2}}$$

along the solution to $\ddot{y} = 0$.

We now consider the details for the dynamics of a particle of mass $m = 1$ moving in three-space subject to the nonholonomic constraints $\dot{z} = y\dot{x}$ (see also [3] and [21]). As in the rolling hoop example from section 4.3.1, the n-symplectic dynamics program requires the input of vector field to start (take again $i_2 = 3$ with the coordinates ordered as $[\{r^i\}, \{s^A\}] = [r^1, r^2, s^1] = [x, y, z]$, see Section 4.3.3) along with a mass matrix, g and the transformation matrix, $\overset{\triangle}{G}$ (cf. Appendix F equation (F.8))

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & y & 0 \end{pmatrix}.$$

It will be necessary to again include certain terms of the scalar potential matrix, C (we assume both the gauge terms T and the vector potential terms B are zero). The dynamics of (4.34) can be written as

$$\left(\overset{\triangle}{\Pi}{}^{\mathcal{I}}_{\mathcal{J}}(t)\right) = \begin{pmatrix} \overset{\triangle}{\Pi}{}^1_1 & \overset{\triangle}{\Pi}{}^1_2 & 0 \\ \overset{\triangle}{\Pi}{}^2_1 & \overset{\triangle}{\Pi}{}^2_2 & 0 \\ y\dot{x} & \dot{x}(1+y^2) & \dot{y} \end{pmatrix}$$

and

$$\left(\frac{d}{dt}\,\overset{\triangle}{\Pi}{}^{\mathcal{I}}_{\mathcal{J}}(t)\right) = \begin{pmatrix} C^{13}_{,z} & \frac{1}{2}\,\overset{\triangle}{\Pi}{}^3_3\overset{\triangle}{\Pi}{}^1_1 + C^{13}_{,x} + yC^{13}_{,z} & C^{13}_{,y} \\ C^{23}_{,z} & \frac{1}{2}\,\overset{\triangle}{\Pi}{}^2_1\left(\overset{\triangle}{\Pi}{}^3_3 + y\dot{x}\right) + C^{23}_{,x} + yC^{23}_{,z} & \frac{1}{2}y\dot{x}\,\overset{\triangle}{\Pi}{}^2_2 + C^{23}_{,y} \\ C^{33}_{,z} & y\dot{x}\,\overset{\triangle}{\Pi}{}^3_3 + C^{33}_{,x} + yC^{33}_{,z} & C^{33}_{,y} \end{pmatrix}$$

where the constraint equation is given by

$$\dot{z} - y\dot{x} = 0.$$

Assumption 4.3.2 *(Scalar Potential C restriction) Let $C^{13} = 0$ and let*

$$\begin{aligned} C^{33}(t,x,z) &= \frac{Q \cdot A\,[z - x \cdot (At+B)]}{[1 + (At+B)^2]^{3/2}} && (4.50) \\ C^{23}(t,y) &= -\frac{y}{2}\left[\frac{(At+B)\cdot Q \cdot K}{\sqrt{1+(At+B)^2}}\right] && (4.51) \end{aligned}$$

where $A, B,$ and Q are arbitrary constants and $K = \overset{\vartriangle}{\Pi}{}^{2}_{2}(0)$.

For the scalar potentials restricted as above, $\frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{3}_{3} = 0$ where $\overset{\vartriangle}{\Pi}{}^{3}_{3} = \dot{y}$ and thus $\ddot{y} = 0$ which gives $y(t) = At + B$. Consequently, at least along solution to $\ddot{y} = 0$, the generalized momentum dynamics reduce to

$$\left(\frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{\mathcal{I}}_{\mathcal{J}}(t)\right) = \begin{pmatrix} 0 & \frac{1}{2}\dot{y}\,\overset{\vartriangle}{\Pi}{}^{1}_{1} & 0 \\ 0 & \frac{1}{2}\,\overset{\vartriangle}{\Pi}{}^{2}_{1}\,(\dot{y} + y\dot{x}) & \frac{1}{2}y\dot{x}\,\overset{\vartriangle}{\Pi}{}^{2}_{2} + C^{23}_{,y} \\ \frac{Q\cdot A}{[1+(At+B)^2]^{3/2}} & y\dot{x}\dot{y} & 0 \end{pmatrix} \tag{4.52}$$

Note that we have, as in the rolling hoop example, recovered the nonholonomic momenta and the nonholonomic momentum equation

$$\overset{\vartriangle}{\Pi}{}^{3}_{2} = (1+y^2)\dot{x} = \dot{x} + y\dot{z} \;\text{ and }\; \frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{3}_{2} = y\dot{x}\dot{y} = \dot{z}\dot{y}$$

Differentiating the [3,1] and the [3,2] entries of the matrix (4.3.2) we obtain

$$\frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{3}_{1} = \dot{y}\dot{x} + y\ddot{x} \;\text{ and }\; \frac{d}{dt}\overset{\vartriangle}{\Pi}{}^{3}_{2} = \ddot{x}(1+y^2) + 2\dot{x}\dot{y}.$$

which are also equal to the corresponding entries of the generalized momenta dynamics matrix in eq. (4.52). Equating the two sets of terms, we obtain the two equations

$$\dot{x}\dot{y} + y\ddot{x} = \frac{Q \cdot A}{[1 + (At + B)^2]^{3/2}} \tag{4.53}$$

$$\ddot{x} = \frac{-y\dot{x}\dot{y}}{1 + y^2} \tag{4.54}$$

Equation (4.54), $\ddot{y} = 0$ and the constraint equation $\dot{z} = y\dot{x}$ complete the dynamics for this nonholonomic particle. Equation (4.53) is a consistency equation that arises from our n-symplectic formulation. Substituting eq. (4.54) into eq. (4.53) gives

$$\dot{x} = \frac{Q}{\sqrt{1 + (At + B)^2}} \tag{4.55}$$

As pointed out in BKMM [3] and references cited therein, eq. (4.54) is the total derivative of

$$\dot{x} = \frac{Q}{\sqrt{1 + y^2}}. \tag{4.56}$$

So our consistency equation (4.55) is eq. (4.56) along the solutions to $\ddot{y} = 0$. To ensure that the $L_\triangle Q$ condition is met (namely that $\overset{\triangle}{\Pi}{}^1_3$ and $\overset{\triangle}{\Pi}{}^2_3$, which we have constrained to be 0, stay 0 under time evolution). That is, we must show that $\frac{d}{dt} \overset{\triangle}{\Pi}{}^2_3 = 0$. Since $\frac{d}{dt} \overset{\triangle}{\Pi}{}^2_1 (t) = 0 \ \forall t$ then setting the initial momenta $\overset{\triangle}{\Pi}{}^2_1 (0) = 0$ gives $\overset{\triangle}{\Pi}{}^2_1 (t) = 0 \ \forall t$. Consequently, $\frac{d}{dt} \overset{\triangle}{\Pi}{}^2_2 (t) = 0 \ \forall t$ and so by setting $\overset{\triangle}{\Pi}{}^2_2 (0) = K \neq 0$ then $\overset{\triangle}{\Pi}{}^2_2 (t) = K \ \forall t$. So the equation we need to satisfy, along the solutions to $\ddot{y} = 0$ and $\ddot{x} = -y\dot{x}\dot{y}/\sqrt{1 + y^2}$ is

$$\frac{1}{2} \frac{(At + b) \cdot K\dot{Q}}{\sqrt{1 + (At + B)^2}} + C^{23}_{,y} = 0.$$

For the potential in eq. (4.51) this equation is indeed satisfied.

Finally, our n-symplectic formulation of this nonholonomic particle is consistent with the dynamics derived by BKMM [3]. We would like to point out here that A.D. Lewis in [21] considers this same nonholonomic particle example but from the perspective of restricting connections on the configuration space to the constraint distribution. That is, using a new connection (and not n-symplectic dynamics), the nonholonomic momentum for this particle are derived. It is encouraging and interesting to note that the frame bundle, the distribution adapted frame bundle and the soldering form are known objects to this author but the generalized symplectic structure, which we have used to derive the dynamics for this particle, appears to be unknown. The relationship between the two formulations would be interesting to explore.

We have recovered in the larger context of n-symplectic dynamics the equations for a particle subject to the nonholonomic constraints $\dot{z} = x\dot{y}$. To obtain these dynamics we were required to impose scalar potentials on the larger space, LQ to obtain the dynamics [3] formulate in terms of TQ. The next step and a source of future work is to use more potentials and gauge terms to address potential shaping and energy conserving control strategies. For some preliminary results along these lines see Appendix I.

4.3.3 Why Choose the $i_2 = 3 = (A = 1)$ Vector Field?

As stated in Theorem 4.2.1, to obtain a specific system of constrained n-symplectic dynamics it is necessary to choose a specific vector field $\overset{\triangle}{X}{}^{i_2}$. Which vector field (or linear combination of vector fields) should be chosen and by what mathematical means is this vector field actually specified? Mathematically, we specify a vector field using an element α from $\mathbb{R}^{n\star}$ [represented in coordinates as $(\alpha_1, \alpha_2 \ldots, \alpha_n)$] which can then be multiplied (traced) with the n vector fields to obtain $\alpha_{i_2} X^{i_2}$. The selection of α essentially choses a "slice" of the full n-symplectic dynamics. Finding the correct "slice" is our next task. To the n-symplectic way of thinking $\overset{\triangle}{g}{}^{ab}\overset{\triangle}{\Pi}{}^i_a\overset{\triangle}{\Pi}{}^j_b$ is the rank-2 (i.e. matrix valued) kinetic energy observable adapted to the constraint distribution. This n-symplectic matrix should in some way be related to the constrained kinetic energy function (Lagrangian) for the specific mechanical system. Tracing out $\overset{\triangle}{g} \cdot \overset{\triangle}{\Pi} \cdot \overset{\triangle}{\Pi}$ with the unknown but desired α according to $\alpha_i \alpha_j (\overset{\triangle}{g}{}^{ab}\overset{\triangle}{\Pi}{}^i_a\overset{\triangle}{\Pi}{}^j_b)$ should give the constrained Lagrangian for the mechanical system.

Indeed in the n-symplectic vertical rolling hoop example, for $\alpha = (0, 0, 1, 0)$, we obtain

$$\alpha_i \alpha_j (\overset{\triangle}{g}^{ab} \overset{\triangle}{\Pi}{}^i_a \overset{\triangle}{\Pi}{}^j_b) = \frac{1}{2}(I + mR^2)\dot{\theta}^2 + \frac{1}{2} J \dot{\phi}^2 \tag{4.57}$$

which is the constrained Lagrangian for the vertical rolling hoop. For this choice of α, $\alpha_{i_2} \overset{\triangle}{X}{}^{i_2} = \overset{\triangle}{X}{}^{3}$ which justifies our assignment of $i_2 = 3$ (the $A = 1$ or 1^{st} group variable index).

In the nonholonomic particle example, we chose $\alpha = (0, 0, 1)$ since

$$\alpha_i \alpha_j (\overset{\triangle}{g}^{ab} \overset{\triangle}{\Pi}{}^i_a \overset{\triangle}{\Pi}{}^j_b) = (1 + y^2)\dot{x}^2 + \dot{y}^2 \tag{4.58}$$

which is the constrained Lagrangian for the nonholonomic constrained particle.

References

[1] R. Abraham and J.E. Marsden. *Foundations of mechanics.* Westview Press, Cambridge, 1978.

[2] M. Berger. *A panoramic view of Riemannian geometry.* Springer, New York, 2003.

[3] A. Bloch, P.S. Krishnaprasad, J.E. Marsden, and R.M. Murray. Nonholonomic Mechanical Systems with Symmetry. *Arch. Rational Mech. Anal.*, 136:21–99, 1996.

[4] A.M. Bloch. *Nonholonomic Mechanics and Control*, volume 24. Springer-Verlag, 2003.

[5] A.M. Bloch and J.E. Marsden. Controlling homoclinic orbits. *Theoretical and Computation Fluid Dynamics*, 1:179–190, 1989.

[6] F. Bullo. *Nonlinear control of mechanical systems: a Riemannian geometry approach.* Ph.d. dissertation, California Institute of Technology, Pasadena California, 1999.

[7] F. Bullo, N.E. Leonard, and A.D. Lewis. Controllability and motion algorithms for underacutated lagrangian systems on lie groups. *Transactions on Automatic Control*, 45:1437–1454, 2000.

[8] F. Bullo and A.D. Lewis. *Geometric control of mechanical systems.* Springer, New York, 2005.

[9] F. Bullo, R. M. Murray, and A. Sarti. Control on the sphere and reduced attitude stabilization. In *Symposium on Nonlinear Control Systems*, volume 2, pages 495–501, Tahoe City, CA, June 1995.

[10] F. Bullo and R.M. Murray. Tracking for fully actuated mechanical systems: a geometric framework. 35:17–34, 1999.

[11] S.A. Chaplygin. On the thoery of motion of nonholonomic systems. the theorem on the reducing multiplier. *Math. Sbornik (in Russian)*, XXVIII:303–314, 1911.

[12] M.P. DoCarmo. *Differential geometry of curves and surfaces.* Prentice-Hall, New Jersey, 1976.

[13] M.P. DoCarmo. *Riemannian geometry.* Birkhäuser, Boston, 1992.

[14] R. Fuentes, G. Hicks, and J. Osborne. Intrinsic Control Designs for Abstract Machines. In *Mathematical Theory of Networks and Systems*, pages 984–995, 2006.

[15] R. Fuentes, G. Hicks, and J. Osborne. Intrinsic control designs for abstract machines. In Springer-Verlag, editor, *Lecture Notes in Control and Information Sciences*, 2007.

[16] J.E. Marsden H. Cendra and T.S. Ratiu. Lagrangian Reduction by Stages. *Memoirs of the American Mathematical Society*, 152, 2001.

[17] Thomas Hawkins. *Emergence of the Theory of Lie Groups: An Essay in the History of Mathematics 1869-1926.* Springer-Verlag, New York, 2000.

[18] Neimark. Ju. I. and N.A. Fufaev. *Dynamics of Nonholonomic Systems.* AMS (Translation of Methematical Monographs), Providence, 1972.

[19] H.K. Khalil. *Nonlinear Systems.* Prentice Hall, 2000.

[20] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry, Vol. 1.* John Wiley, New York, 1963.

[21] A.D. Lewis. Affine Connections and Distributions with Applications to Nonholonomic Mechanics. *Reports on Mathematical Physics*, 42:135–164, 1998.

[22] A.D. Lewis. Simple mechanical control systems with constraints. *Transactions on Automatic Control*, 45:1420–1436, 2000.

[23] A.D. Lewis and R.M. Murray. Configuration controllability of simple mechanical control systems. *SIAM Journal of Control and Optimization*, 35:766–790, 1997.

[24] Jerrold Marsden. *Lectures on Mechanics*. Cambridge University Press, Cambridge, 1992.

[25] J. McCleary. *Geometry from a differentiable viewpoint*. Cambridge University Press, Cambridge, 1994.

[26] M. McLean and L.K. Norris. Covariant Field Theory on Frame Bundles of Fibered Manifolds. *Journal of Mathematical Physics*, 40:6808–6823, 2000.

[27] J. Milnor. *Morse theory*. Princeton University Press, Princeton, New Jersey, 1963.

[28] L.K. Norris. Generalized Symplectic Geometry on the Frame Bundle of a Manifold. *Proceeding of Symposia in Pure Mathematics*, 54:435–465, 1993.

[29] L.K. Norris. Schouten–Nijenhuis Brackets. *Journal of Mathematical Physics*, 38:2694–2709, 1997.

[30] B. O'Neill. *Semi-Riemannian geometry with applications to relativity*. Academic Press, San Diego, 1983.

[31] J. Osborne and D.V. Zenkov. Steering the Chaplygin Sleigh with a Moving Mass. In *Proceedings of Conference on Decision and Control and European Conference and Control*, 2005.

[32] J. Ostrowski. Steering for a class of dynamics nonholonomic systems. *Transaction on Automatic Control*, 45:1492–1498, 2000.

[33] Rand. R.H. and D.V. Ramani. Relaxing nonholonomic constraints. In *Proceedings of First International Symposium on Impact and Friction of Solids, Structures and Intelligent Machines, Conference on Decision and Control and European Conference and Control*, pages 113–116, 2000.

[34] J.K. Lawson R.O. Fulp and L.K. Norris. Geometric prequantization on the spin bundle based on n-symplectic geometry: The dirac equation. *International Journal of Theoretical Physics*, 33:1011–1028, 1994.

[35] J.K. Lawson R.O. Fulp and L.K. Norris. Generalized symplectic geometry as a covering theory for the hamiltonian theoreis of classical particles and fields. *Journal of Geometry and Physics*, 20:195–206, 1996.

[36] A. Ruina. Non-holonomic stability aspects of piecewise-holonomic systems. *Rep. Math. Phys.*, 42:91–100, 1998.

[37] J.A. Schouten. *Ricci–Calculus*. Springer-Verlag, Berlin, Göttingen, Heidelberg, 1954.

[38] R. Sinclair and M. Tanaka. Loki: software for computing cut loci. *Experimental Mathematics*, 11(1):25 pages, 2002.

[39] R. Sinclar and M. Tanaka. Loki: Software for Computing Cut Loci. *Experimental Mathematics*, 11, 2000.

[40] J.L. Synge. *Relativity: the general theory*. North-Holland Publishing Company, Amsterdam, 1971.

[41] D. Zenkov, A. Bloch, and J. Marsden. The Energy-Momentum Method for the Stability of Nonholonomic Systems. *Dynamics and Stability of Systems*, 13:123–166, 1998.

[42] D.V. Zenkov and A.M. Bloch. Invariant measures of nonholonomic flows with internal degrees of freedom. *Nonlinearity*, pages 1793–1807, 2003.

Appendices

Appendix A

Constructs from Riemannian Geometry

A.1 Introduction

In this appendix the details behind the connection/covariant derivative, specifically the Levi-Civita connection, and the Riemmann curvature are given. With the Levi-Civita connection defined we then move on to Parallel Transport and Geodesics. These geometric constructs are needed for the geodesic based PD control on Riemannian manifolds and the geometry of the double gimbal chapters. The structure of this appendix follows [30] and is as follows:

- Section A.1.1: Notation
- Section A.2: Connections and Statement of Levi-Civita Connection
- Section A.2.1: Tensor Derivations
- Section A.2.2: Components of Covariant Derivative (specifically of DX, DΘ, Dg)
- Section A.2.3 Derivation/Proof of Levi-Civita Connection

- Section A.3 Riemman Curvature Tensor, Ricci Tensor, Ricci Scalar
- Section A.4 Parallel Transport and Geodesics

A.1.1 Notation

We will use the following notation:

- M-smooth manifold
- p-point in M, x=$(x^1, \ldots, x^n)$ is a chart on U (open set about p in M)
- $\mathfrak{X}$(M)-Set of all (tangent) vector fields on M . In terms of tensor fields we say that $\mathfrak{X}$(M) is a (1,0) tensor field. Notation: $\mathfrak{X}$(M)=$\mathfrak{T}^1_0$.
- $\mathfrak{X}^\star$(M)-Set of all covector fields on M (i.e. the dual space to $\mathfrak{X}$(M), i.e. the space of all $\mathfrak{F}$(M) linear maps from $\mathfrak{X}$(M) into $\mathfrak{F}$(M)). In terms of tensor fields we say that $\mathfrak{X}^\star$(M) is a (0,1) tensor field. Notation: $\mathfrak{X}^\star$(M)=$\mathfrak{T}^0_1$.
- $\mathfrak{F}$(M)-Set of all smooth functions from M into $\mathbb{R}$, f $\in \mathfrak{F}$(M)
- X=$X^i\partial_i$,Y=$Y^j\partial_j \in \mathfrak{X}$(M)-vector fields where $X^i, Y^j \in \mathfrak{F}$(M) and $\partial_i = \frac{\partial}{\partial x^i}, \partial_j = \frac{\partial}{\partial x^j} \in \mathfrak{X}$(M)
- $\Theta = \Theta_k dx^k$-covector field where $\Theta_k \in \mathfrak{F}(M)$ and $dx^k \in \mathfrak{X}^\star$(M).
- g=$g_{ij}dx^i \otimes dx^j \in \mathfrak{T}^0_2$-metric tensor where $g_{ij} \in \mathfrak{F}$(M). A (0,2) tensor field that is symmetric and non-degenerate is said to be a metric tensor field.
- (M,g)-semi-Riemannian manifold, g a metric tensor field.

A short explantation of the notation is as follows.

1. We are using the Einstein convention. So whenever there is a repeated upper and lower index we are summing over that index. That is,

$$X^i\partial_i = \sum_{i=1}^{m} X^i\partial_i$$

where m is the dimension of the manifold M.

2. A tensor field is an $\mathfrak{F}(M)$-multilinear map into $\mathfrak{F}(M)$. Our main examples will be (0,2),(1,0) and (0,1) tensor fields:

   - The metric tensor g, is a (0,2) tensor field which means g is an $\mathfrak{F}(M)$-bilinear map from $\mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{F}(M)$. That is, $g(fX, Y) = fg(X, Y)$ and $g(X + Y, Z) = g(X, Z) + g(Y, Z)$. Ditto for the second slot.
   - A vector field X, is a (1,0) tensor field which means that X is an $\mathfrak{F}$(M)-linear map from $\mathfrak{X}^\star$(M) $\to \mathfrak{F}(M)$. That is, $X(f\Theta) = fX(\Theta)$ and $X(\Theta + \Psi) = X(\Theta) + X(\Psi)$ [*Note:* We are using here that $\mathfrak{X}$(M)$\cong$ $\mathfrak{X}^{\star\star}$(M) via the isomorphism ρ defined by $\rho(X)(\Theta) = \Theta(X)$]
   - A covector field Θ, is a (0,1) tensor field which means that Θ is an $\mathfrak{F}$(M)-linear map from $\mathfrak{X}$(M) $\to \mathfrak{F}(M)$. That is, $\Theta(fX) = f\Theta(X)$ and $\Theta(X + Y) = \Theta(X) + \Theta(Y)$.

3. Covector fields act on vector fields to give smooth functions. That is,

$$\begin{aligned}\Theta(X) &= \Theta_j dx^j(X^i\partial_i) \\ &= \Theta_j X^i dx^j(\partial_i) \\ &= \Theta_j X^i \delta_i^j \\ &= \Theta_j X^j \in \mathfrak{F}(M)\end{aligned}$$

4. Vector fields act on smooth functions to give smooth functions. That is,

$$X(f) = X^i\partial_i(f) = X^i f_i \in \mathfrak{F}(M)$$

5. Vector fields act on points in M to give vectors at a point. That is,

$$X(p) = X|_p = X^i(p)\partial_i|_p$$

where $X^i(p) \in \mathbb{R}$ and $\partial_i|p \in T_pM$ (tangent space to M at p)

6. g a metric tensor acts on two vector fields to return a smooth function. That is,

$$\begin{aligned} g(X,Y) &= g_{ij}dx^i \otimes dx^j(X^a\partial_a, Y^b\partial_b) \\ &= g_{ij}X^aY^b dx^i \otimes dx^j(\partial_a, \partial_b) \\ &:= g_{ij}X^aY^b dx^i(\partial_a)dx^j(\partial_b) \\ &= g_{ij}X^aY^b\delta^i_j\delta^j_b \\ &= g_{ij}X^iY^j \in \mathfrak{F}(M) \end{aligned}$$

7. The coordinates of g given by g_{ij} are obtained by acting g on the vector fields ∂_a, ∂_b. That is,

$$\begin{aligned} g(\partial_a, \partial_b) &= g_{ij}dx^i \otimes dx^j(\partial_a, \partial_b) \\ &= g_{ij}dx^i(\partial_a)dx^j(\partial_b) \\ &= g_{ij}\delta^i_a\delta^j_b \\ &= g_{ab} \end{aligned}$$

This is how we find the coordinates of any tensor field. For more on tensor fields (that is, more details than just saying a tensor field is an $\mathfrak{F}$(M)-multilinear function into $\mathfrak{F}$(M)) see [[30], Chapter 2].

A.2 Connection/Covariant Derivative

In this section we define the connection and prove a theorem about a special connection called the Levi-Civita connection.

Definition A.2.1 *([30], pg. 59) A* connection *D on (M,g) is a map D:*$\mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ *such that*

1. *$D(X,Y){:=}D_X Y$ is $\mathfrak{F}(M)$-linear in the first slot (i.e the X slot)*

2. *$D(X,Y){:=}D_X Y$ is $\mathbb{R}$-linear in the second slot (i.e. the Y slot)*

3. $D_X(fY) := D_X(f)Y + fD_X(Y) := (Xf)Y + f(D_X Y)$

D_X is called the *covariant derivative with respect to X for the connection D* and $D_X Y$ is called the *covariant derivative of Y with respect to X for the connection D*. We can see that 3. is saying that D_X acts as a *derivation* or that D_X obeys a *product rule* and hence the connection *is not* a tensor.

Now we see from the definition that given two vector fields X,Y the covariant derivative of Y with respect to X is another vector field, say

$$D_X Y = Z = Z^k \partial_k. \tag{A.1}$$

We can think of D:$\mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ in another way, that is, as a map from D:$\mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}^\star(M) \to \mathfrak{F}(M)$. Hence,

$$D(X, Y, \Theta) := (D_X Y)(\Theta)$$

where D is $\mathfrak{F}(M)$-linear in the X slot, $\mathbb{R}$-linear in the Y slot and $\mathfrak{F}(M)$-linear in the Θ slot. Hence we can think of writing D as

$$D = \Gamma^k_{ij} dx^i \otimes dx^j \otimes \partial_k$$

where $\Gamma^k_{ij} \in \mathfrak{F}(M)$ are the component functions of D (by 3. we know that they are *not* tensor components). To find the components we do exactly as we did in 7. from Section 1. That is let X=∂_a and Y=∂_b to obtain

$$\begin{aligned} D(X, Y, \cdot) &= \Gamma^k_{ij} dx^i \otimes dx^j \otimes \partial_k(\partial_a, \partial_b, \cdot) \\ &= \Gamma^k_{ij} dx^i(\partial_a) dx^j(\partial_b) \partial_k(\cdot) \\ &= \Gamma^k_{ab} \partial_k \end{aligned} \tag{A.2}$$

where the dot indicates that we are to leave that spot open. So we can equate eq. 1 and eq. 2 to see that $Z^k = \Gamma^k_{ab}$. Hence, we can just go straight to the result that

$$D_{\partial_i}\partial_j = \Gamma^k_{ij}\partial_k.$$

The Γ^k_{ij} are called the *connection coefficients* and we say again are *not* the components of a tensor.

Theorem A.2.1 *If (M,g) is a semi-Riemannian manifold then there is a unique connection D called the Levi-Civita connection such that*

4. $[X,Y] = D_X Y - D_Y X$ *(torsion free condition)*
5. *X(g(Y,Z))=*$g(D_X Y, Z) + g(Y, D_X Z)$ *(non-metricity condition)*

We will soon get to the proof but first we examine why 4. is called the torsion free condition and 5. the non-metricity condition. Let us work out 4. in local coordinates. The left side works out when we define [·,·] (the commutator bracket of vector fields) and switch some indices:

$$\begin{aligned}
[X,Y] &:= X(Y^i)\partial_i - Y(X^j)\partial_j \\
&= X^j\partial_j((Y^i)\partial_i) - Y^i\partial_i((X^j)\partial_j) \\
&= X^jY^i_{,j}\partial_i - Y^iX^j_{,i}\partial_j \\
&= (X^jY^k_{,j} - Y^iX^k_{,i})\partial_k \\
&= (X^jY^k_{,j} - Y^jX^k_{,j})\partial_k
\end{aligned}$$

The right side works out when we use the derivation property of D_{∂_i} and D_{∂_i}:

$$\begin{aligned} D_X Y - D_Y X &= X^i D_{\partial_i}(Y^j \partial_j) - YjD_{\partial_j}(X^i \partial_i) \\ &= X^i(Y^j_{,i}\partial_j + Y^j D_{\partial_i}\partial_j) - Y^j(X^i_{,j}\partial_i + X^i D_{\partial_j}\partial_i) \\ &= X^i Y^j_{,i}\partial_j - Y^j X^i_{,j}\partial_i + X^i Y^j \Gamma^k_{ij} - Y^j X^i \Gamma^k_{ji} \\ &= (X^i Y^k_{,i} - Y^j X^k_{,j})\partial_k + X^i Y^j(\Gamma^k_{ij} - \Gamma^k_{ji}) \\ &= [X,Y] + X^i Y^j T^k_{ij} \end{aligned}$$

T^k_{ij} is called the torsion tensor (it is indeed a tensor because basically the non-tensor components of the connection coefficients are subtracted out; more on this in a minute). So we have T_{ij}=0 $\implies$ $[X,Y] = D_X Y - D_Y X$. To see that $[X,Y] = D_X Y - D_Y X$=0 $\implies$ $T_{ij} = 0$ follows immediately from $[\partial_i, \partial_j]$=0. That is,

$$0 = [\partial_i, \partial_j] = D_{\partial_i}(\partial_j) - D_{\partial_j}(\partial_i) = \Gamma^k_{ij} - \Gamma^k_{ji} = T^k_{ij}$$

A connection that satisfies 4. is called a *symmetric connection* since $T^k_{ij} = 0$ implies $\Gamma^k_{ij} = \Gamma^k_{ji}$ and hence the connection coefficients are *symmetric* in the lower indices.

To understand the non-metricity condition we must consider the concept of a *tensor derivation.*

A.2.1 Tensor Derivations

Definition A.2.2 ([30], pg. 43) *Let* $A \in \mathfrak{T}^r_s$ *and* $B \in \mathfrak{T}^m_n$. *A* tensor derivation $\mathcal{D}^p_q$ *is* <u>*set*</u> *of* $\mathbb{R}$*-linear maps*

$$\mathcal{D}^p_q : \mathfrak{T}^p_q \to \mathfrak{T}^p_q \, where \, p = r + m, q = s + n$$

such that

1. $\mathcal{D}^p_q(A \otimes B) = \mathcal{D}^r_s A \otimes B + A \otimes \mathcal{D}^m_n B$

2. $\mathcal{D}$ commutes with any contraction, that is $\mathcal{D}^{r-1}_{s-1}(CA) = C(\mathcal{D}^r_s A)$

Note: For the concept of a contraction see [[30], pg. 40]. But as an example of a contraction in coordinates, the contraction of a (1-2) tensor field A is given by $C^1_2(A^i_{jk}) = A^i_{ik}$. That is, the C^1_2 contraction just sums out the first contravariant index with the first covariant index. By extension, C^1_3 would sum out the first contravariant index with the second covariant index to give $C^1_3(A^i_{jk}) = A^i_{ji}$

In [[30], pg. 44] a product rule formula for tensor derivations given a tensor field of type (p,q) is derived. This formula requires the use of both 1. and 2. from above. The jist of this formula can be seen in the special cases of a vector field X, a covector field Θ, a (1,1) tensor field A and a (0,2) tensor field g (thinking of the metric tensor field). That is,

$$\mathcal{D}^0_0(A(\Theta, X)) = \mathcal{D}^1_1 A(\Theta, X) + A(\mathcal{D}^0_1\Theta, X) + A(\Theta, \mathcal{D}^1_0 X) \tag{A.3}$$

and

$$\mathcal{D}^0_0(\Theta(X)) = \mathcal{D}^0_1\Theta(X) + \Theta(\mathcal{D}^1_0 X) \tag{A.4}$$

and

$$\mathcal{D}^0_0(X(\Theta)) = \mathcal{D}^1_0 X(\Theta) + X(\mathcal{D}^0_1\Theta) \tag{A.5}$$

and

$$\mathcal{D}^0_0(g(X, Y)) = \mathcal{D}^0_2 g(X, Y) + g(\mathcal{D}^1_0 X, Y) + g(X, \mathcal{D}^1_0 Y) \tag{A.6}$$

There are three points to be made by these examples

1. From eq. 3 and eq. 6 we can solve for $\mathcal{D}^1_1$ and $\mathcal{D}^0_2$ to obtain

$$\begin{aligned}
\mathcal{D}^1_1 A(\Theta, X) &= \mathcal{D}^0_0(A(\Theta, X)) - A(\mathcal{D}^0_1\Theta, X) - A(\Theta, \mathcal{D}^1_0 X) \\
\mathcal{D}^0_2 g(X, Y) &= \mathcal{D}^0_0(g(X, Y)) - g(\mathcal{D}^1_0 X, Y) - g(X, \mathcal{D}^1_0 Y).
\end{aligned}$$

   That is, all we really need to know to find tensor derivations of higher order tensor fields are $\mathcal{D}^0_0(f)$, $\mathcal{D}^1_0(X)$ and $\mathcal{D}^0_1(\Theta)$.

2. But from eq. 4 we can find $\mathfrak{D}_1^0$ in terms of $\mathfrak{D}_0^0$ and $\mathfrak{D}_0^1$. That is,

$$\mathfrak{D}_1^0\Theta(X) = \mathfrak{D}_0^0(\Theta(X)) - \Theta(\mathfrak{D}_0^1 X)$$

3. So all we really need to find tensor derivation of higher order tensor fields are $\mathfrak{D}_0^1$ and $\mathfrak{D}_0^0$.

The point of the argument in [[30], pg. 12 (the bottom of the page)] is that

$$\mathfrak{D}_0^0(f) = Xf\ \forall f \in \mathfrak{F}(M)$$

and the point of the theorem in [[30], pg. 45] (where $\delta = D_X$ since $D_X(fY) := D_X(f)Y + fD_X(Y) := (Xf)Y + f(D_XY)$ by property 3. of the connection D) is that

$$\mathfrak{D}_0^1(X) = D_Z X\ \ for some Z \in \mathfrak{X}(M)$$

Taking into account these formulas we can write eq. 6 as (where we have switched X to Y and Y to Z and denoted $\mathfrak{D}_2^0 = \mathfrak{D}$)

$$X(g(Y,Z)) = \mathfrak{D}g(X,Y) + g(D_XY,Z) + g(Y,D_XZ). \tag{A.7}$$

Comparing eq. 7 and 5. (i.e. $X(g(Y,Z)) = g(D_XY,Z) + g(Y,D_XZ)$) from the theorem on pg. 4 we can see that

$$X(g(Y,Z)) = g(D_XY,Z) + g(Y,D_XZ) \iff \mathfrak{D}g = 0$$

where we mean

$$\mathfrak{D}g = D_Zg = Dg(\cdot,\cdot,Z) = 0\ \forall Z$$

and hence the terminology non-metricity condition makes sense. A connection that satisfies 5. is said to be a *metric* connection. A metric, symmetric connection is called the *Levi-Civita* connection. That is, the connection D which satisfies $\mathfrak{D}g = D_Zg =$

$0 \ \forall Z$ is the Levi-Civita connection.

Before we proof the theorem we must introduce one more concept, that is the components of the covariant derivative of a tensor field.

A.2.2 Components of DT, T $\in \mathfrak{T}^r_s$

Given a tensor field of type (r,s) T (with coordinates $T^{i_1,\dots,i_r}_{j_1,\dots j_s}$ then $D : \mathfrak{T}^r_s \to \mathfrak{T}^r_{s+1}$ is given by

$$DT(\Theta_1, \dots, \Theta_r, X_1, \dots, X_s, X) := D_X(\Theta_1, \dots, \Theta_r, X_1, \dots, X_s)$$

where the components of the new tensor are given by $DT^{i_1,\dots,i_r}_{j_1,\dots,j_s;k}$. The term covariant derivative is fitting in the sense that a new lower index (a covariant index), namely ;k, is introduced. To describe the components of the new tensor DT, we will consider smaller examples as these will sufficiently illustrate the general.

- If T=X=$X^a\partial_a$ then what are the $X^a_{;k}$ of DX=$X^a_{;k}\partial_a \otimes dx^k$? We present the following computations:

$$\begin{aligned} DX(\cdot, Y) &= X^a_{;k}\partial_a \otimes dx^k(\cdot, Y^j\partial_j) \\ &= Y^j X^a_{;k}\delta^k_j\partial_a \\ &= Y^k X^a_{;k}\partial_a \\ &= Y^k X^n_{;k}\partial_n \text{ (switch a to n)} \end{aligned}$$

but by definition of DX we have

$$\begin{aligned} DX(\cdot, Y) &:= [D_Y X](\cdot) \\ &= [Y^k D\partial_k (X^a \partial_a)](\cdot) \\ &= [Y^k (X^a_{,k}\partial_a + X^a D_{\partial_k}(\partial_a))](\cdot) \text{ (tensor derivation property of} D_{\partial_k}) \\ &= [Y^k (X^n_{,k} + X^a \Gamma^n_{ka})\partial_n](\cdot) \\ &= Y^k (X^n_{,k} + X^a \Gamma^n_{ka})\partial_n. \end{aligned}$$

Equating the two computations gives $X^n_{;k} = X^n_{,k} + X^a\Gamma^n_{ka}$, the components of DX.

- If T=$\Theta = \Theta_a dx^a$ then what are the $\Theta_{a;k}$ of $D\Theta = \Theta_{a;k} dx^a \otimes dx^k$? We present the following computations:

$$\begin{aligned} D\Theta(\cdot, Y) &= \Theta_{a;k} dx^a \otimes dx^k(\cdot, Y^j\partial_j) \\ &= Y^j \Theta_{a;k} \partial^k_j dx^a \\ &= Y^j \Theta_{a;j} dx^a \\ &= Y^j \Theta_{k;j} dx^k \text{ (switch a to k)} \end{aligned}$$

but by defintion of DΘ we have

$$\begin{aligned} D\Theta(Y, \cdot) &:= [D_Y \Theta](\cdot) \\ &= [Y^j D_{\partial_j}(\Theta_a dx^a)](\cdot) \\ &= [Y^j(\Theta_{a,j} dx^a + \Theta_a D_{\partial_j}(dx^a))](\cdot) \text{ (using tensor derivation property)} \\ &= [Y^j(\Theta_{a,j} dx^a - \Theta_a \Gamma^a_{jk} dx^k)](\cdot) \text{ (using} D_{\partial_j}(dx^a) = -\Gamma^a_{jk} dx^k) \\ &= [Y^j(\Theta_{k,j} - \Theta_a \Gamma^a_{jk}) dx^k](\cdot) \\ &= Y^j(\Theta_{k,j} - \Theta_a \Gamma^a_{jk}) dx^k. \end{aligned}$$

Equating the two computations gives $\Theta_{k;j} = \Theta_{k,j} - \Theta_a\Gamma^a_{jk}$, the components of

DΘ.

- Following the same method as above we find that for g $\in \mathfrak{T}^0_2$ then

$$g_{ij;k} = g_{ij,k} - g_{aj}\Gamma^a_{ki} - g_{ia}\Gamma^a_{kj} \tag{A.8}$$

A.2.3 Levi-Civita Connection

We now prove the

Theorem A.2.2 *If (M,g) is a semi-Riemannian manifold then there is a unique connection D called the Levi-Civita connection such that*

4. $[X, Y] = D_X Y - D_Y X \iff T^k_{ij} = 0$

5. X(g(Y,Z))=g$(D_X Y, Z) + g(Y, D_X Z) \iff \mathfrak{D}g = 0 \iff D_X g = 0\ \forall X \iff g_{ij;k} = 0\ \forall i, j, k$

Proof: Let D be a connection. Consider $N_{abi} := -g_{ab;i} + g_{ib;a} + g_{ai;b}$ where a,b,i are fixed but arbitrary indices. Expanding out N_{abi} using eq. 8 three times and $g_{ij} = g_{ji}$ gives

$$\begin{aligned}
N_{abi} &= -g_{ab,i} + g_{kb}\Gamma^k_{ia} + g_{ak}\Gamma^k_{ib} \\
&\quad\ g_{ib,a} - g_{kb}\Gamma^k_{ai} - g_{ik}\Gamma^k_{ab} \\
&\quad\ g_{ai,b} - g_{ki}\Gamma^k_{ba} - g_{ak}\Gamma^k_{bi} \\
&= -g_{ab,i} + g_{ib,a} + g_{ai,b} + g_{kb}(\Gamma^k_{ia} - \Gamma^k_{ai}) + g_{ak}(\Gamma^k_{ib} - \Gamma^k_{bi}) - g_{ki}(\Gamma^k_{ab} + \Gamma^k_{ba})
\end{aligned}$$

Now notice that Γ^k_{ij} which is not a tensor can be “factored” into a symmetric object and a tensor via

$$\begin{aligned}
\Gamma^k_{ij} &= \frac{1}{2}(\Gamma^k_{ij} + \Gamma^k_{ji}) + \frac{1}{2}(\Gamma^k_{ij} - \Gamma^k_{ji}) \\
&= \Gamma^k_{(ij)} + \frac{1}{2}T^k_{ij}.
\end{aligned} \tag{A.9}$$

Using eq. 9 in can easily be seen, for example, that

$$\begin{aligned} \Gamma^k_{ia} - \Gamma^k_{ai} &= T^k_{ia} \\ \Gamma^k_{ab} + \Gamma^k_{ba} &= 2\Gamma^k_{(ab)} \end{aligned}$$

which implies

$$N_{abi} = -g_{ab,i} + g_{ib,a} + g_{ai,b} + g_{kb}T^k_{ia} + g_{ak}T^k_{ib} - 2g_{ki}\Gamma^k_{(ab)}.$$

We may solve the above for $2g_{ki}\Gamma^k_{(ab)}$ and use eq. 9 to obtain a formula that we can use to solve for Γ^k_{ab}. That is,

$$2g_{ki}\Gamma^k_{(ab)} = 2g_{ki}(\Gamma^k_{ab} - \frac{1}{2}T^k_{ab}) = 2g_{ki}\Gamma^k_{ab} - g_{ki}T^k_{ab} \tag{A.10}$$

and thus

$$2g_{ki}\Gamma^k_{ab} = -g_{ab,i} + g_{ib,a} + g_{ai,b} + g_{kb}T^k_{ia} + g_{ak}T^k_{ib} + g_{ki}T^k_{ij} - N_{abi})$$

where by applying g^{ji} to both sides we obtain the desired result,

$$\delta^j_k\Gamma^k_{ab} = \Gamma^j_{ab} = \frac{1}{2}g^{ji}(-g_{ab,i} + g_{ib,a} + g_{ai,b} + g_{kb}T^k_{ia} + g_{ak}T^k_{ib} + g_{ki}T^k_{ij} - N_{abi}) \tag{A.11}$$

In words, eq. 11 expresses the components of a general connection in terms of the metric tensor components, the torsion tensor components, and the covariant derivative of the metric tensor components. We can see that by applying 4. kills off all torsion terms in eq. 11 and 5. kills off the N_{abi} leaving only

$$\Gamma^j_{ab} = \frac{1}{2}g^{ji}(g_{ib,a} + g_{ai,b} - g_{ab,i}) \tag{A.12}$$

called the *Christoffel symbols* or the *components of the Levi-Civita connection* or the *components of a symmetric, metric connection*. One could invent new notation for the Christoffel symbols to distinguish them from arbitrary connection coefficients, say

$$\left\{ {j \atop ab} \right\} \text{ or } \omega^j_{ab}.$$

A.3 Curvatures

A.3.1 Riemann Curvature

The Riemann curvature tensor has an algebraic definition. It is given in [30, pg. 74] as

Definition: Given (M,g) and Levi-Civita connection D the mapping
R:$\mathfrak{X}$(M)$\times\mathfrak{X}$(M)$\times\mathfrak{X}$(M) $\to \mathfrak{X}$(M) given by

$$R_{(X,Y)}Z = R_{[XY]}Z - [D_X, D_Y]Z$$

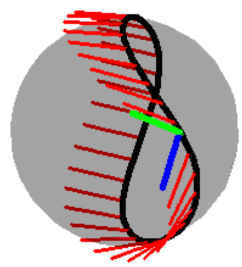
is called the *Riemann curvature tensor.* In terms of coordinates,

$$\begin{aligned} R^a_{\cdot ibj} &= \Gamma^a_{ij,b} - \Gamma^a_{ib,j} + \Gamma^a_{bn}\Gamma^n_{ji} - \Gamma^a_{jn}\Gamma^n_{ai} \\ &:= \Gamma^a_{[|i|j,b]} + \Gamma^a_{[b|n|}\Gamma^n_{j]i} \end{aligned}$$

where $[a|c|b]$ means acb-bca.

In its most straight-forward incarnation, the *Riemann curvature of an n-dimensional Riemannian manifold* (M, g) is, simply stated, that notion which measures the local difference between a subset (U, g) of (M, g) and Euclidean $\mathbb{R}^n$. What follows will be a diagrammatic illustration of curvature based upon parallel transport followed by an argument and explicit formula for the *Riemannian curvature tensor.* We conclude with a formula of the *Gauss curvature of a surface.*

From a pictorial standpoint, a manifold has curvature if one can find a closed curve which does not traverse a 2-dimensional topological hole for which the parallel transport (see section A.4) of an initial vector V_0 around this curve returns a final vector, $V_f \neq V_0$. That is, the difference between an initial vector V_0 and the parallel transported vector, V_f is a measure of curvature. One can say that if $V_0 - V_f = 0$ then (U, g) has zero curvature or that M is locally *flat* or even that M is locally Euclidean $\mathbb{R}^n$. Moreover, if $V_0 - V_f \neq 0$ then one can say that the subset $(U, g) \subset (M, g)$ has non-zero curvature or that M is locally *non-flat* or even that M is locally non-Euclidean.



(a)

Figure A.1: Parallel Transport on Euclidean S^2. Parallel transport of initial vector, V_0 (in blue) around closed curve (in black) to a final vector, V_f (in green)

A good example is given by Figure A.2. There the closed curve (in black) does not traverse a hole (there are none since the Betti number $B_1(S^2) = 0$) and the blue and green vectors are obviously different. We have pictorially discovered that the sphere is curved, as expected. To address the question of "How curved is the manifold?" will require an explicit curvature formula.

To obtain a curvature formula one can ask the following question: Given a coordinate system (Ψ^α) for the open set U of M with metric tensor expressed in this coordinate system $g_{\alpha\beta}(p) \neq \delta_{\alpha\beta}$ for every p belonging to U, does there exist a coordinate system (Φ^k) on U such that the metric tensor in this new system $g_{ij}(p)$ takes the form of the identity tensor δ_{ij}? In other words, if a Riemannian manifold does not appear Euclidean in one coordinate system then is it Euclidean in another? The answer to the existence question can be broken into three steps (one can find further detail in [25]):

[Step 1: $g_{ij} = \delta_{ij}$ if and only if $\Gamma^k_{ij} = 0$] This result follows from the formula for the Christoffel symbols, Γ (the components of the Levi-Civita connection) and that the covariant derivative (w.r.t to the Levi-Civita connection) of the metric tensor, denoted $g_{;}$, is 0:

$$\Gamma^k_{ij} = \frac{1}{2} g^{ka}(g_{aj,i} + g_{ia,j} - g_{ij,a}) \tag{A.13}$$

$$g_{ij;l} = g_{ij,l} - g_{aj}\Gamma^a_{li} - g_{ib}\Gamma^b_{lj} = 0 \tag{A.14}$$

Now if $g_{ij} = \delta_{ij}$, then by equation A.14, $\Gamma^k_{ij} = 0$. If $\Gamma^k_{ij} = 0$, then by equation (A.14), $g_{ij,l} = 0$ which have a particular solution $g_{ij} = \delta_{ij}$.

[Step 2: $\Gamma^k_{ij} = 0$ is equivalent to the statement: Φ^k satisfies the system of PDE's given by equation (A.15)] This result follows from the formula for the non-tensor transformation law for the Christoffel symbols

$$\Gamma^\sigma_{\alpha\beta}\Phi^k_{,\sigma} = \Phi^k_{,\alpha\beta} + \Phi^i_{,\alpha}\Phi^j_{,\beta}\Gamma^k_{ij}$$

where $\Phi^k = \Phi^k(\Psi^\alpha)$ are the change of coordinate functions with $_{,\alpha}$ partial differentiation wrt the Ψ^α coordinate and $_{,\alpha\beta}$ is partial differentiation with respect to Ψ^α then with respect to Ψ^β. Now for $\Gamma^k_{ij} = 0$, the desired system of PDE's is

$$\Gamma^\sigma_{\alpha\beta}\Phi^k_{,\sigma} = \Phi^k_{,\alpha\beta} \tag{A.15}$$

That is, the desired coordinates (Φ^k) for which $\Gamma^k_{ij} = 0$ need to satisfy equation (A.15).

[Step 3: Riemann curvature tensor = 0 implies Φ^k satisfy equation (A.15)] This result follows from an analysis of the integrability conditions (a.k.a the necessary conditions for Φ^k to be a solution) of equation (A.15). The integrability conditions, namely that mixed partials commute, lead to

$$R_{\gamma\alpha\beta\cdot}^{\ \ \ \ \rho} := \Gamma^{\rho}_{\gamma\beta,\alpha} + \Gamma^{\Delta}_{\gamma\beta}\Gamma^{\rho}_{\Delta\alpha} - \Gamma^{\rho}_{\gamma\alpha,\beta} - \Gamma^{\Delta}_{\gamma\alpha}\Gamma^{\rho}_{\Delta\beta} = 0. \tag{A.16}$$

For an n dimensional manifold, the n^4 functions $R_{\gamma\alpha\beta\cdot}^{\ \ \ \ \rho}$ define the *Riemann curvature tensor* and are determined from the metric tensor $g_{\alpha\beta}$ by first computing $\Gamma^{\sigma}_{\rho\eta}$ by equation (A.32) and then using equation (A.16).

The above steps are summarized as,

Curvature result: If the Riemann curvature tensor in the original coordinate system is zero then there exists a coordinate system (Φ^k) *such that* $g_{ij} = \delta_{ij}$ *and hence* M *would locally appear to be Euclidean* $\mathbb{R}^n$ *(i.e. flat). However, a Riemann curvature with non-zero components indicates a manifolds local deviation from flatness.*

With the Riemann curvature tensor in hand, the *Gaussian curvature of a surface or 2-D manifold* is given by

$$K := \frac{R_{1212}}{g_{11}g_{22} - (g_{12})^2} \tag{A.17}$$

where R_{1212} is the [1212]-entry of $R_{jkli} := g_{ia}R_{jkl\cdot}^{\ \ \ \ a}$

Some examples using the curvature result are

Example 2a. The cylinder

Choose a parametrization (coordinate system) of (on) the cylinder given by

$$(\theta, z) \to [\cos(\theta), \sin(\theta), z]$$

which will give the components of the metric tensor

$$\begin{aligned} g_{11} &= \ = [-\sin(\theta), \cos(\theta), 0] \cdot [-\sin(\theta), \cos(\theta), 0] = 1 \\ g_{22} &= [0, 0, 1] \cdot [0, 0, 1] = 1 \\ g_{12} &= 0 \end{aligned}$$

and hence $\Gamma = 0$ (as the metric tensor components are constant on an open set U given by, e.g. $\theta = 0..2\pi$ and z=0..1).

Example 2b. The cone

Choose a parametrization (coordinate system) of (on) the cone given by

$$(u, v) \rightarrow [u\cos(v), u\sin(v), u]$$

which will give the components of the metric tensor

$$g_{11} = 2, g_{22} = u^2, g_{12} = 0$$

which are not constant on an open set but still the curvature tensor is

$$R^{\rho}_{\gamma\alpha\beta} = 0.$$

Example 2c. The mobius strip

Choose a parametrization (coordinate system) of (on) the mobius strip given by

$$(u, \theta) \rightarrow [(4 - u\sin(\theta/2))\cos(\theta), (4 - u\sin(\theta/2))\sin(\theta), u\cos(\theta/2)]$$

which will define the components of the metric tensor as

$$g_{11} = 1, g_{22} = 16 - 8u\sin(\theta/2) + 5/4u^2 - u^2\cos^2(\theta/2), g_{12} = 0$$

which are not constant on an open set and the curvature tensor is

$$R_{\rho\gamma\alpha\beta} = R_{1212} = \frac{16}{32u\sin(\theta/2) - 5u^2 + 4u^2\cos^2(\theta/2) - 64}$$

which is not zero on an open set.

There is also another physical interpretation of the Riemann tensor. The Riemann tensor makes an appearance in the Jacobi equation, which for geodesic variations, is an equation describing how "nearby" geodesics (free-falling observers) attract or repel each other, see [30].

Back to the curvature tensor itself.

Remark A.3.1 *(***Prelude To Riemann Differences***) One can also write $R_{(X,Y)}Z = R(Z, X, Y)$, the order here is important. This is the order Oneill uses (he calls it the classical/historical order) while Milnor uses $\mathcal{R}_{(X,Y)}Z = \mathcal{R}(X, Y, Z)$ (he calls it a more convenient order). At the end of Section A.3.3 we will examine the differences between R and $\mathcal{R}$.*

Using the classical order

$$R(\partial_k, \partial_i, \partial_j) = R_{\partial_i \partial_j}\partial_k = R^a_{\cdot kij}\partial_a$$

the first symmetry of the Riemann tensor (pg. 75, [30]), $R_{XY} = -R_{YX}$ means $R(Z, X, Y) = -R(Z, Y, X)$ and hence at the coordinate level

$$R^a_{\cdot kij} = -R^a_{\cdot kji} \tag{A.18}$$

Consider now the object $g(W, R(Z, X, Y))$. In coordinates, using the classical ordering, we obtain

$$\begin{aligned} g(\partial_n, R(\partial_k, \partial_i, \partial_j)) &= g(\partial_n, \cdot)[R(\partial_k, \partial_i, \partial_j)] \\ &= g_{nb}dx^b[R^a_{\cdot kij}\partial_a] \\ &= g_{nb}R^b_{\cdot kij} \\ &:= R_{nkij} \end{aligned}$$

The first step used

$$g(\partial_n, X) = g_{nb}X^b = g_{nb}dx^b(X) \implies g(\partial_n, \cdot) = g_{nb}dx^b.$$

The last step above is an example of lowering a contra-variant index to a covariant index using the metric. The meaning of the $\cdot$ in the symbol $R^a_{\cdot kij}$ is used to indicate that the a contravariant index is to be lowered to that spot.

The second symmetry of the Riemann tensor [30, pg. 75], $g(W, R(Z, X, Y)) = -g(Z, R(W, X, Y))$ can now be expressed at the coordinate level, using the classical ordering, by

$$R_{nkij} = -R_{knij} \tag{A.19}$$

Remark A.3.2 *(***Raising and Lowering***) Raising and lowering indices (using a metric) creates new tensors from an original tensor. The new tensors are said to be* metrically equivalent *to the original. For instance, as above, let $R = R^1_3$ be the (1,3) Riemman tensor field given in coordinates by $R^a_{\cdot kij}$. Define the (0,4) tensor $\downarrow^1_1 R^1_3 = R^0_4$ by*

$$R^0_4(W, Z, X, Y) = g(W, R^1_3(Z, X, Y))$$

where the coordinates of R^0_4 are given by

$$R_{nkij}$$

In coordinates, we see the significance of the subscripts and superscripts of $\downarrow^1_1$; the metric g will lower the first contra-variant index to the first covariant index of the new tensor. Other metric equivalent tensor can be obtained. For instance $\downarrow^1_2 (R^a_{kij}) = R_{kaij}$. Lowering via $\downarrow^1_1$ will be the convention for the Riemann tensor as indicated by the notation $R^a_{\cdot kij}$ where $\cdot$ says "lower the index to this spot".

One can raise covariant indices to contra-variant indices using the inverse of the metric. For example given the original (1,3) Riemann tensor $R = R^a_{\cdot kij}$ the new tensor

$\uparrow_1^2 R$ is a (2,2) tensor defined by

$$R_{ij}^{an} = g^{nk} R_{.kij}^{a}$$

Now that we can raise indices we can contract tensors we couldn't before (i.e to contract the Ricci tensor to the Ricci scalar)

A.3.2 Contraction

We recall first contraction of a contra-variant and co-variant index. Contraction is a map which drops the rank of the tensor field being contracted by 2. In the case of a (1,1)-tensor field $T = \theta \otimes X$ contraction is defined by

$$C(T) := T(dx^i, \partial_i) = (\theta \otimes X)(dx^i, \partial_i) = \theta_i X^i = \theta(X). \qquad \text{(A.20)}$$

which is a function on M or 0-tensor field. This could be expressed as $T_b^a = \theta_b X^a \xrightarrow{C} T_i^i = \theta_i X^i$. One could also think of the contraction in the following way:

$$C(\theta \otimes X) = \theta(X) = g(Y, X) \qquad \text{(A.21)}$$

where Y is the vector metrically equivalent to θ. In coordinates this would be

$$g_{ij} Y^j dx^i (X^k \partial_k) = g_{ij} Y^j X^i = \theta_i X^i$$

As this example has only one contravariant and one covariant index there was no choice as to the contraction to be performed. But contraction can be extended to higher order tensor fields. For example consider the (1,2)-tensor field A (meaning $A(\theta, X, Y)$) given in coordinates by A_{jk}^i. Now there are two lower indices that can be used to contract out the upper index, i.e there are two maps $C = C_1^1$ and $C = C_2^1$

defined, following eq. (A.20),

$$\begin{aligned} C^1_1(A) &= A(dx^i, \partial_i, Y) \quad , Y\, fixed \\ C^1_2(A) &= A(dx^k, X, \partial_k) \quad , X\, fixed \end{aligned}$$

with respective coordinate descriptions

$$\begin{aligned} C^1_1(A^i_{jk}) &= A^i_{ik} \\ C^1_2(A^i_{jk}) &= A^k_{ik} \end{aligned}$$

So we have illustrated how to contract upper (contravariant) indices with lower (covariant) indices of tensor fields (i.e. $\mathcal{F}(M)$-valued multi–linear functions).

However, the Riemann curvature tensor is a $\mathfrak{X}(M)$-valued multilinear function, i.e. vector valued function given by $(X, Y, Z) \longrightarrow R(X, Y, Z)$. To make this a "true" tensor field we need to make this a $\mathcal{F}(M)$-valued multi-linear function as follows: Given $R(X, Y, Z) \in \mathfrak{X}(M)$ define

$$\bar{R}(\theta, X, Y, Z) = \theta(R(X, Y, Z)) \in \mathcal{F}(M) \tag{A.22}$$

where a coordinate description of $\bar{R}$ is

$$\begin{aligned} \bar{R}^b_{ijk} = \bar{R}(dx^b, \partial_i, \partial_j, \partial_k) &= dx^b(R(\partial_i, \partial_j, \partial_k)) \\ &= dx^b(R^a_{\cdot ijk}\partial_a) \\ &= R^b_{\cdot ijk} \end{aligned} \tag{A.23}$$

It is in the sense of eqs. (A.22) and (A.23) that we say the vector valued object R can be interpreted as a (1,3)-tensor field $\bar{R}$. To contract a vector valued object like R is to contract its corresponding tensor field via the higher rank analogue of (A.20) given by

$$(C^1_3\bar{R})(X, Y) = \bar{R}(dx^k, X, Y, \partial_k) = dx^k(R(X, Y, \partial_k)) \tag{A.24}$$

or in coordinates by

$$C^1_3(\bar{R}^a_{\cdot ijk}) = R^k_{\cdot ijk} \tag{A.25}$$

A different contraction of the Riemann tensor could be

$$(C^1_2\bar{R})(X,Z) = \bar{R}(dx^k, X, \partial_k, Z) = dx^k(R(X,\partial_k,Z)) \tag{A.26}$$

with coordinate description

$$C^1_2(\bar{R}^a_{\cdot ijk}) = R^k_{\cdot ikj} = -R^k_{\cdot ijk} = -C^1_3(\bar{R}^a_{\cdot ijk}) \tag{A.27}$$

So $C^1_2 = -C^1_3$. These are the only non-zero contractions as

$$C^1_1(R^a_{\cdot ijk}) = R^i_{\cdot ijk} = g^{ia}R_{aijk} = -g^{ia}R_{iajk} = -R^i_{\cdot ijk}$$

and so $C^1_1(\bar{R}) = 0$.

Remark A.3.3 *(**Tensor Interpretations**) The distinction between $\bar{R}$ and R will be dropped from now on. This is fine, since, as seen above, the components of multilinear function, $\bar{R}$, are exactly the components of the vector valued object R. So without further thought, (1,3)-contraction of R is given by $R^k_{\cdot ijk}$. Just for reference, given say a (1,3)-tensor field $(\theta, X, Y, Z) \longrightarrow \bar{T}(\theta, X, Y, Z)$ one can define*

$$T(X,Y,Z) = \bar{T}(\cdot, X, Y, Z)$$

which is an equation interpreting $\bar{T}$ as a vector valued object T.

So we can contract the upper and lower indices of a tensor field and we can perform a contraction of a vector valued object. It will be necessary to contract two lower or two upper indices of a tensor field. We won't get into much detail about this except to do a couple of specific examples in coordinates.

For instance, consider the (0,2) tensor field T_{ij} and the object $C_{12}(T_{ij})$ which we know should be some sort of function but the exact function requires a definition of

C_{12}. The meaning of $C_{12}(T_{ij})$ is:

1. Find the metric equivalent tensor of T_{ij}. This is, T_j^k where $g_{ki}T_j^k = T_{ij}$.

2. Now we have an upper and lower index, so contract in the usual way, i.e.

$$T_j^k \xrightarrow{C} T_j^j$$

3. Now $C_{12}(T_{ij})$ will be the function of T_{ij} for which making the substitution $g_{ki}T_j^k = T_{ij}$ will be exactly T_j^j. This is exactly,

4. $C_{12}(T_{ij}) = g^{ij}T_{ij}$

Now consider the (0,4) tensor field R_{aijk} and the object $C_{14}(R_{aijk})$. The meaning of $C_{14}(R_{aijk})$ is:

1. Find a metric equivalent tensor of R_{aijk}. Now we have a choice here, which index do we raise and to which spot? Say that we raise the first index to the first spot, so a metric equivalent tensor to R_{aijk} is $R^b_{\cdot ijk}$ where $g_{ba}R^b_{\cdot ijk} = R_{aijk}$.

2. Now that we have an upper and lower index, we can contract. Again, there is a choice; which lower index to contract with the upper index? Say that we contract the first and third, i.e.

$$R^b_{\cdot ijk} \xrightarrow{C_3^1} R^k_{\cdot ijk}$$

3. Now, as 3) above, using $R^k_{\cdot ijk} = g^{ak}R_{aijk}$ leads to,

4. $C_{14}(R_{aijk}) = g^{ak}R_{aijk}$

Remark A.3.4 *(***Prelude To Contraction Differences***) In the $C_{14}(R_{aijk})$ example, there were two choices that had to be made: i) which index to raise (here we raised the first) and then ii) which contraction to take (here we took C_3^1). As we will address at the end of Section A.3.3, different books (and Maple !!) can make different choices.*

A.3.3 Ricci Curvature Tensor, Ricci Scalar

The Ricci tensor, $Ric = R_{ij}$ is the (1,3) contraction of the Riemann tensor as defined in eqs. (A.24) and (A.25) . That is,

$$\begin{aligned} Ric(X,Y) = (C_3^1 \bar{R})(X,Y) &= \bar{R}(dx^j, X, Y, \partial_j) \\ &= dx^j(R(X,Y,\partial_j)) \end{aligned}$$

or in coordinates as

$$R_{ki} = C_3^1(R^a_{\cdot kij}) = R^j_{\cdot kij}$$

The Ricci scalar, S is a contraction of the Ricci tensor. That is,

$$S = C(Ric)$$

which in coordinates is

$$S = R^i_i = g^{ki} R_{ki}$$

Remark A.3.5 *(***Oneill, Milnor and Maple***) We now point out the difference between the Oneill and Milnor definitions of curvature. We also bring in how Maple is computing Ricci and Ricci scalar. From what I gather from footnote at the bottom of Milnor pg. 51 the definitions* $\mathcal{R}(X,Y,Z) = \mathcal{R}_{(X,Y)}Z$ *and* $(\downarrow^1_4 \mathcal{R})(X,Y,Z,W) = g(\mathcal{R}(X,Y,Z),W)$ *are made since*

$$\begin{aligned} g(\mathcal{R}(\partial_i,\partial_j,\partial_k),\partial_n) &= g(\mathcal{R}_{ijk\cdot}{}^a \partial_a, \partial_n) \\ &= g_{bn} \mathcal{R}_{ijk\cdot}{}^b \\ &= \mathcal{R}_{ijkn} \end{aligned}$$

As shown on pg. 53 of Milnor, the curvature tensor $\mathcal{R}$ *satisfies the same symmetry relations as R. That is,* $\mathcal{R}_{ijkn} = -\mathcal{R}_{jikn}$ *and* $\mathcal{R}_{ijkn} = -\mathcal{R}_{ijnk}$*. Contrast this with Oneills classical definition* $R(Z,X,Y) = R_{(X,Y)}Z$ *and* $(\downarrow^1_1 R)(W,Z,X,Y) = g(W,R(Z,X,Y))$

which gave

$$\begin{aligned} g(\partial_n, R(\partial_k, \partial_i, \partial_j)) &= g(\partial_n, \cdot)[R(\partial_k, \partial_i, \partial_j)] \\ &= g_{nb} dx^b [R^a_{\cdot kij} \partial_a] \\ &= g_{nb} R^b_{\cdot kij} \\ &:= R_{nkij} \end{aligned}$$

We have seen that for Oneill, that the Ricci tensor is $R_{ij} = R^a_{\cdot ija}$. *From what I can gather from Milnor pg. 104, the Ricci tensor is given by (cf. eq. (A.24))*

$$(C^1_2 \bar{\mathcal{R}})(X, Y) = \bar{\mathcal{R}}(dx^k, X, \partial_k, Y) = dx^k(\mathcal{R}(X, \partial_k, Y))$$

which coordinate description

$$\mathcal{R}_{ij} = \mathcal{R}_{iaj\cdot}{}^{a}$$

To understand what Maple is computing we consider the example, Euclidean S^2 *parametrized by* $sphere(\theta, \phi) = (\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi))$ *where* $\theta = 0..2\pi$, $\phi = 0..\pi/2$. *Maple computes for* S^2

$$\begin{aligned} &g_{22} = \sin^2(\phi), \quad g_{11} = 1 \\ &g^{22} = 1/\sin^2(\phi), \quad g^{11} = 1 \\ &RMN = R_{ijkl} = R_{1212} = \sin^2(\phi) \\ &Ricci = Ric = R_{ij} \;\; where \;\; R_{11} = -\sin^2(\phi), \; R_{22} = -1 \\ &RS = S = R^i_i = -2 \end{aligned}$$

For all these quantities to make sense Maple has to be using for Riemann

$$\mathfrak{R}_{ijk\cdot}{}^{a} \tag{A.28}$$

as in Milnor while for Ricci

$$\mathfrak{R}_{ij} = \mathfrak{R}_{kij\cdot}{}^{k} = g^{ka}\mathfrak{R}_{kija} \tag{A.29}$$

which is the negative of Milnor. We justify these via computations: Raising the fourth index of $\mathfrak{R}_{ijka}$ *(of which the only non-zero entry is* $\mathfrak{R}_{1212} = sin^2(\phi)$*) using* g^{-1} *gives the non-zero entries*

$$\begin{aligned} \mathfrak{R}_{121\cdot}{}^{2} = &\quad sin^2(\phi) &= -\mathfrak{R}_{211\cdot}{}^{2} \\ \mathfrak{R}_{122\cdot}{}^{1} = &\quad -1 &= -\mathfrak{R}_{212\cdot}{}^{1} \end{aligned}$$

and thus computing the (1,1) contraction of $\mathfrak{R}_{ijk\cdot}{}^{a}$ *gives* $\mathfrak{R}_{ij}$ *of which the only non-zero entries are*

$$\begin{aligned} \mathfrak{R}_{22} = \mathfrak{R}_{211\cdot}{}^{2} &= -sin^2(\phi) \\ \mathfrak{R}_{11} = \mathfrak{R}_{122\cdot}{}^{1} &= -1 \end{aligned}$$

which is what Maple computed. Note: the (1,1) contraction here is first up index with first down index. Maple actually labels the first up index as the 4 total index and the first down index as first total index and thus (1,1)-contraction is at (1,4) contraction.

Summary A.3.1 *We summarize the Riemman and Ricci curvatures with the table:*

author	*Riemann*	*Ricci*
Oneill	$R^{a}_{\cdot ijk}$	$R_{ij} = R^{a}_{\cdot ija}$
Milnor	$\mathcal{R}_{ijk\cdot}{}^{a}$	$\mathcal{R}_{ik} = \mathcal{R}_{iak\cdot}{}^{a}$
Maple	$\mathfrak{R}_{ijk\cdot}{}^{a}$	$\mathfrak{R}_{ij} = \mathfrak{R}_{aij\cdot}{}^{a}$

where the symmetries of each are

$$R^a_{\cdot ijk} = -R^a_{\cdot ikj} \quad \textit{and} \quad R_{nijk} = -R_{injk}$$
$$\mathcal{R}_{ijk\cdot}{}^a = -\mathcal{R}_{jik\cdot}{}^a \quad \textit{and} \quad \mathcal{R}_{ijkn} = -\mathcal{R}_{ijnk}$$
$$\mathfrak{R}_{ijk\cdot}{}^a = -\mathfrak{R}_{jik\cdot}{}^a \quad \textit{and} \quad \mathfrak{R}_{ijkn} = -\mathfrak{R}_{ijnk}$$

and thus

$$\mathcal{R}_{ik} = \mathcal{R}_{iak\cdot}{}^a = -\mathcal{R}_{aik\cdot}{}^a = -\mathfrak{R}_{ik}.$$

A.3.4 Sectional Curvature

Given two independent vector fields on M, say X and Y, then they define a tangent plane at each point of M, say Π_p. That sectional curvature of Π is

$$K(X,Y) = g(R_{(X,Y)}X,Y)/Q(X,Y)$$

where R is the Riemann tensor using Oneill's convention and $Q(X,Y) = g(X,X)g(Y,Y) - g(X,Y)^2$. If the vector fields are orthonormal then we get

$$K(X,Y) = g(R_{XY}X,Y).$$

We move to higher dimensions, Let $E_i \;\; i = 1 \ldots dim(TM)$ be an orthonormal basis for TM. Fix the leg E_1 of the orthonormal frame and form dim(TM)-1 tangent planes using the other legs of the frame and find the sectional curvature of each. That is,

$$\sum_i K(E_1,E_i) = \sum_i g(R_{E_1E_i}E_1,E_i) = Ric(E_1,E_1)$$

It can now be seen that the n-th diagonal entry of the Ricci tensor is the sum of all the sectional curvatures of the planes containing E_n leg of an orthonormal frame.

As pointed out in [Remark 55, pg 89, [30]] to convert from their classical notation based on R to more convenient notation based on $\mathcal{R}$ one need only switch R in Oneill's definitions to $-\mathcal{R}$ to find the new definitions. He gives the example of the sectional curvature defined through the curvature tensor:

$$\begin{aligned} K(X,Y) &= g(R_{XY}X,Y)/Q(X,Y) \\ &= -g(\mathcal{R}_{XY}X,Y)/Q(X,Y) \\ &= g(\mathcal{R}_{XY}Y,X)/Q(X,Y). \end{aligned}$$

That is, $K(X,Y) = g(R_{XY}X,Y)/Q(X,Y)$ is the definition of sectional curvature defined through R. In order to keep the sign of the sectional curvature the same when using $\mathcal{R}$ one must use the definition $K(X,Y) = g(\mathcal{R}_{XY}Y,X)/Q(X,Y)$.

A.4 Parallel Transport and Geodesics

It is necessary in building a geometry-based controller for the double gimbal that one has the notion of a *transport mapping*, $\mathcal{T}_{r\to q}$ which is, essentially, a smooth mapping that carries v_r, a tangent vector at r in the Riemannian manifold to another tangent vector $\mathcal{T}_{r\to q}(v_r)$ at the point q. The utility of said map lies in the fact that only vectors at the same point can be added or subtracted. Parallel transport along geodesic curves is the transport map we use in defining an intrinsic *velocity error* required for an intrinsic geometry based control law, see Chapter 3. So we now address parallel transport and geodesics and give some examples.

The coordinate free form of the parallel transport equations are

$$\frac{D}{du}V = 0, \tag{A.30}$$

where D/du is the *covariant derivative (w.r.t the Levi-Civita connection) along a curve* $\sigma(u)$ and V is a vector field along $\sigma(u)$. The solution vector field is that which is *covariatn* constant along the curve $\sigma(u)$. In coordinates the parallel transport equation are the linear system of, possibly, non-autonomous ODE's given by

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{dV^i}{du}+\Gamma^i_{jk}\frac{d\gamma^j}{du}V^k=0, \tag{A.31}$$

where Γ^i_{jk} are the *Christoffel symbols*, namely those functions of the metric tensor and its derivatives, given by

$$\Gamma^i_{jk}=\sum_{a=1}^{n}\frac{1}{2}g^{ia}(g_{ak,j}+g_{ja,k}-g_{jk,a}), \tag{A.32}$$

where g^{ia} is the k^{th}-row and a^{th}-column entry of $(g)^{-1}$ and $_{,i}$ denotes partial differentiation wrt the i^{th} coordinate. From this point on we will consistently use an implied summation over repeated upper and lower indices.

A *geodesic curve* on M is that curve, $\sigma(u)$ whose tangent vector field $d\sigma/du$ satisfies the equation

$$\frac{D}{du}\frac{d}{du}\sigma(u)=0. \tag{A.33}$$

Equation A.33 is called the geodesic equation and is that equation requiring the tangent vector field to a geodesic curve be parallel transported. The geodesic equation is Newton's equation $F=ma$ on a Riemannian manifold with $F=0$ and $m=1$. Alternatively, geodesics are those curves $\sigma(u)$ on (M,g), out of all possible curves $\alpha(u)$ on (M,g), which minimize the energy functional (see [27]),

$$E(\alpha)=\int_{u=a}^{u=b}g\left(\frac{d\alpha}{du},\frac{d\alpha}{du}\right)du.$$

The coordinate form of equation (A.33) is obtained by substitution of $V^k=d\alpha^k/du$ into equation (A.31):

$$\frac{d^2\alpha^i}{du^2}+\Gamma^i_{jk}\frac{d\alpha^j}{du}\frac{d\alpha^k}{du}=0. \tag{A.34}$$

We now give some examples of parallel transport and geodesics. To illustrate the parallel transport equations consider the curve in Euclidean $\mathbb{R}^3$ given by $\sigma(u)=$

$I(\boldsymbol{\gamma}(u))$ where $\boldsymbol{\gamma}(u) = \left[\, \gamma^1(u)\ \gamma^2(u)\ \gamma^3(u) \,\right]^{\mathrm{T}} = \left[\, x(u)\ y(u)\ z(u) \,\right]^{\mathrm{T}}$ and $\Gamma^i_{jk} = 0$ (which follows from the derivatives of the constant entries of $G_{\mathbb{R}^3} = I$ being 0). In this case the parallel transport equations are

$$\frac{d}{du}V^i(\boldsymbol{\gamma}(u)) = 0$$

That is, in the Euclidean $\mathbb{R}^3$ case, the covariant derivative can be replaced by a regular time derivative. For the initial conditions $[V^i(\boldsymbol{\gamma}(u_0))] := V_0$ the solution vector field, V along $\boldsymbol{\gamma}(u)$ is that with constant components V_0, see Figure A.2. We say that $V(\boldsymbol{\gamma}(u_f)) := V_f$ is the transport of V_0 along $\boldsymbol{\gamma}(u)$.

As another illustration consider the curve on S^2 $(r = 1)$ given by $\sigma(u) = \mathbf{S}(\boldsymbol{\gamma}(u))$ with $\boldsymbol{\gamma}(u) = [\theta(u), \phi(u)]^{\mathrm{T}}$. We equip S^2 with the standard metric, the components of which can be used in equation (A.32) to find the Christoffel symbols, see Table A.4.

Table A.1: Metric tensor, g and Christoffel symbol, Γ data for standard S^2 $(r = 1)$

	non-zero entries
g_{ij}	$g_{11} = \sin^2(\phi) \quad g_{22} = 1$
Γ^i_{jk}	$\Gamma^1_{12} = \Gamma^1_{21} = \cot(\phi), \Gamma^2_{11} = -\sin(\phi)\cos(\phi)$

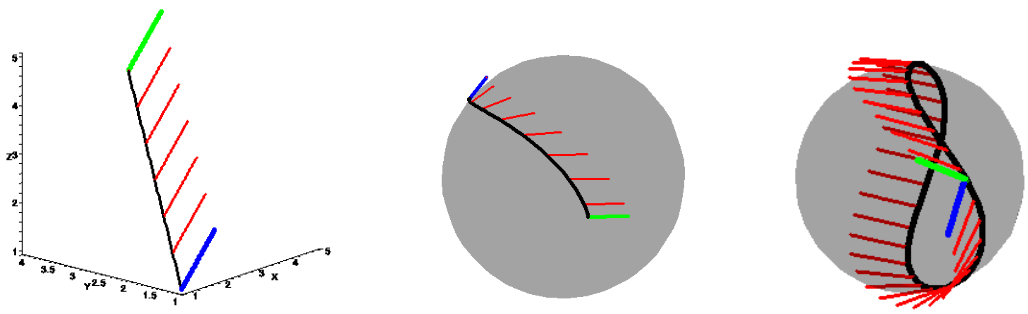
With the data from Table A.4 the parallel transport equations in equation (A.31) become

$$\begin{aligned} \frac{dV^1}{du} + \cot(\phi(u))\frac{d\phi}{du}V^1 + \cot(\phi(u))\frac{d\theta}{du}V^2 &= 0 \qquad \text{(A.35)} \\ \frac{dV^2}{du} - \cos(\phi(u))\sin(\phi(u))\frac{d\theta}{du}V^1 &= 0 \qquad \text{(A.36)} \end{aligned}$$

which, for example, along the curve $\gamma(u) = [u, u]$ are

$$\begin{aligned} \frac{dV^1}{du} + \cot(u)V^1 + \cot(u)V^2 &= 0 \\ \frac{dV^2}{du} - \cos(u)\sin(u)V^1 &= 0. \end{aligned}$$

These equations can be solved for $\mathbf{V} = [V^1, V^2]^{\mathrm{T}}$ and mapped under $D\mathbf{S}$ to the curve $\sigma(u)$, see Figure A.2.



(a) (b) (c)

Figure A.2: Parallel Transport on Euclidean $\mathbb{R}^3$ and standard S^2. a) (left) For $\mathbb{R}^3$ the curve is $\gamma(u) = [x(u), y(u), z(u)] = [\sin(u), \cos(u), u]$ is given in black with initial vector $V_0 = [0, 0, 1]$ (time $u_0 = 1$) given in blue and the transported vector $V_f = [0, 0, 1]$ (time $u_f = 7$) given in green. b) (middle) For S^2 the curve $\sigma(u) = \mathbf{S}(u, u)$ is given in black with initial vector $V_0 = [1, 1]$ (time $u_0 = 1$) mapped to the curve $\sigma(u)$ under $D\mathbf{S}$ given in blue and the transported vector given in green (time $u_f = 3$). c) (right) Parallel transport of initial vector, V_0 (in blue) around closed curve (in black) to a final vector, V_f (in green)

For Euclidean $\mathbb{R}^3$ the geodesics equations are

$$\frac{d^2x}{du^2} = 0 \quad \text{and} \quad \frac{d^2y}{du^2} = 0 \quad \text{and} \quad \frac{d^2z}{du^2} = 0. \tag{A.37}$$

The solutions, $\gamma(u)$ to equation (A.37) are easily seen by inspection to be

$$\gamma(u) \quad = \quad [x(u), y(u), z(u)] = [x_0 + v_{x0}u, y_0 + v_{y0}u, z_0 + v_{z0}u]$$

where the geodesics $\sigma(u) = I(\gamma(u))$ (on Euclidean $\mathbb{R}^3$) are straight lines through the point $p = [x_0, y_0, z_0]$ with initial velocity $\mathbf{v} = [v_{x0}, v_{y0}, v_{z0}]$.

For standard S^2 the geodesic equations are (A.38)

$$\frac{d^2\theta}{du^2} + 2\cot(\phi(u))\frac{d\theta}{du}\frac{d\phi}{du} = 0 \;\; \text{and} \;\; \frac{d^2\phi}{du^2} - \cos(\phi(u))\sin(\phi(u))\left(\frac{d\theta}{du}\right)^2 = 0. \tag{A.38}$$

Some solutions, $\gamma(u)$ to equation (A.38) easily seen by inspection are

$$\gamma(u) = [\theta(u), \phi(u)] = [\theta_{\text{const.}}, u] \;\text{ or }\; \gamma(u) = [u, 0]$$

with corresponding geodesics on S^2, $\sigma(u) = \mathbf{S}(\gamma(u))$, given by

$$\sigma(u) = \text{lines of longitude } \text{ or } \; \sigma(u) = \text{equator.}$$

Appendix B

Coordinates on $Q \to Q/G$ and LQ

B.1 Adapted Coordinates On $Q \to Q/G$

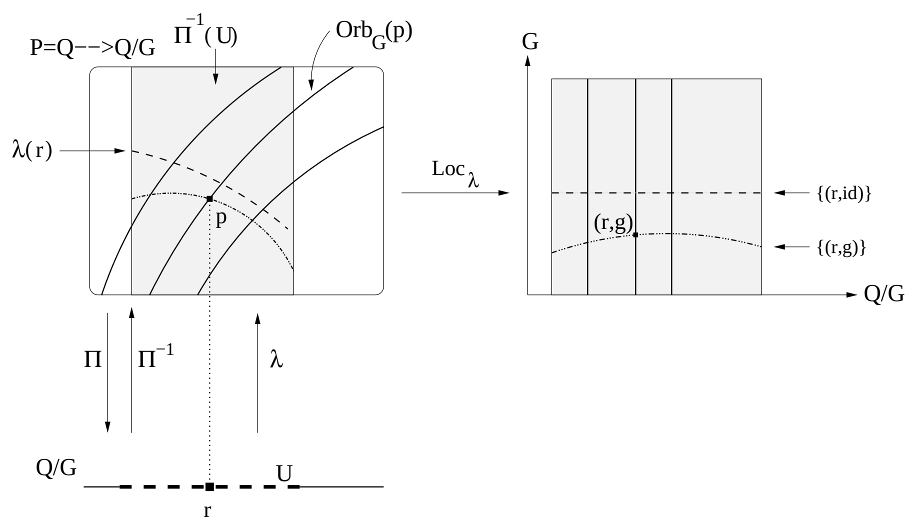


Figure B.1: Adapted Coordinates on Q from a Local Trivialization of a Local Section

Let $Q \overset{\pi_Q}{\to} Q/\mathfrak{G}$ be a principal fiber bundle which can be be viewed as union of group orbits. Assume $Q/\mathfrak{G}$ is a manifold (called *shape space*) with coordinates (r^i) defined on $U \subset Q/\mathfrak{G}$ about $r \in Q/\mathfrak{G}$. Now let σ_Q be a local section of π_Q, which is to say that $\sigma_Q(U)$ is a submanifold of Q for which

$$\pi_Q(\sigma_Q(r)) = r \quad \forall \;\, r \in U$$

One can define coordinates on $Q \to Q/\mathfrak{G}$ using a *local trivialization of a local section of* π_Q, Loc_{σ_Q}. Let $Loc_{\sigma_Q} : \pi_Q^{-1}(U) \to U \times \mathfrak{G} \subset Q/\mathfrak{G} \times \mathfrak{G}$ be defined by

1. $Loc_{\sigma_Q}(\sigma_Q(r)) = (r, id) \; \forall \; r \in U$

2. $Loc_{\sigma_Q}(\sigma_Q(r) \cdot g) = (r, g) \; \forall \; r \in U$

Clarification of 2: Let $q \in \pi_Q^{-1}(U) \subset Q$ be any point. As Q is a union of group orbits, q has to lie on some group orbit over some r. To figure out which compute

$$\pi_Q(q) = r.$$

The section can then be used to find

$$\sigma_Q(r)$$

Since $\pi_Q(\sigma_Q(r)) = r$ then $\sigma_Q(r)$ is on the same fiber as q and hence there must exist a $g \in \mathfrak{G}$ for which

$$q = R_g(\sigma_Q(r)) = \sigma_Q(r) \cdot g$$

Impose the condition (a "commutivity" condition requiring shifting then trivializing to be equal to trivializing then shifting)

$$Loc_{\sigma_Q}(\sigma_Q(r) \cdot g) = Loc_{\sigma_Q}(\sigma_Q(r)) \cdot g = (r, id) \cdot g := (r, g)$$

Then for any $q \in \pi^{-1}(U)$, $Loc_{\sigma_Q}(q) = (r, g) \in U \times \mathfrak{G}$. Now let (s^A) be coordinates on $\mathfrak{G}$ which can be used to define coordinates $(q^{\mathfrak{I}})$ on $\pi_Q^{-1}(U) \subset P$ (using a local trivialization of the section π_Q) by

$$q^{\mathcal{I}}(q) = q^{\mathcal{I}}(Loc_{\sigma_Q}(q)) = q^{\mathcal{I}}(r,g) := (r^i, s^A)(r,g) := (r^i(r), s^A(g)) \tag{B.1}$$

We emphasize that the index convention being used is

$$\mathcal{I}, \mathcal{J}, \mathcal{K} \text{ (upper case script)} = 1 \ldots \dim(Q) = n \tag{B.2}$$

$$i, j, k \text{ (lower case roman)} = 1 \ldots \dim(Q/\mathfrak{G}) \tag{B.3}$$

$$A, B, C \text{ (upper case roman)} = 1 \ldots \dim(\mathfrak{G}) \tag{B.4}$$

B.2 Coordinates On LQ

We now consider the principle fibre bundle $LQ \stackrel{\pi_{LQ}}{\to} Q$, the bundle of linear frames to an n-dimensional manifold Q which is itself a principal fiber bundle $Q \stackrel{\pi_Q}{\to} Q/\mathfrak{G}$. Notationally, LQ is given by

$$LQ = \{(p, \underline{e})\} \tag{B.5}$$

where $\underline{e} = (e_{\mathcal{J}})$ is a basis of TQ . That LQ can be made into a manifold is now addressed. Let $(q^{\mathcal{I}})$ be coordinates on $U \subset Q$ about $q \in Q$ and let σ_{LQ} be a local section of π_{LQ} be given by $\sigma_{LQ}(U) = \{(q, \underline{e}) | \, \underline{e} = \{e_{\mathcal{J}}\}$ a basis of $T_q Q\} \subset \pi^{-1}(U)$ where

$$\pi_{LQ}(\sigma_{LQ}(q)) = q \quad \forall \;\; q \in U \subset Q$$

That is, a local section is an assignment to each point $q \in U \subset Q$ a particular frame. Hence a local section of LQ can be called a *local frame field*. Define a *local trivialization of* $\pi_{LQ}^{-1}(U)$ *given a section* σ_{LQ}, denoted $Loc_{\sigma_{LQ}} : \pi_{LQ}^{-1}(U) \to U \times Gl(n) \subset P \times Gl(n)$ by

1. $Loc_{\sigma_{LQ}}(\sigma_{LQ}(q)) = (q, id) \; \forall \; q \in U$

2. $Loc_{\sigma_{LQ}}(\sigma_{LQ}(q) \cdot g) = (q, g) \; \forall \; q \in U$

Clarification of 2: Let $u \in \pi_{LQ}^{-1}(U) \subset LQ$ be any point. As LQ can be viewed as union of group orbits, u has to lie on some group orbit over some q. To figure out which compute

$$\pi_{LQ}(u) = q.$$

The section can then be used to find $\sigma_{LQ}(q)$. Since $\pi_{LQ}(\sigma_{LQ}(q)) = q$ then $\sigma_{LQ}(q)$ is on the same fiber as u. So there must exist a $g_q \in \mathfrak{G}$ (the $_q$ indicating that the group element needed to shift the section to another frame field will vary from point to point) for which

$$u = R_{g_q}(\sigma_{LQ}(q)) = \sigma_{LQ}(p) \cdot g_q$$

We then impose the condition (a "commutivity" condition requiring shifting then trivializing to be equal to trivializing then shifting)

$$Loc_{\sigma_{LQ}}(\sigma_{LQ}(q) \cdot g_q) = Loc_{\sigma_{LQ}}(\sigma_{LQ}(q)) \cdot g_q = (q, id) \cdot g_q := (q, g_q)$$

From now on we assume that g's q-dependence is implicit, $g_q = g$.

We can restate the above in terms of local reference frame fields. Assume that a local frame field of LQ is given as $\sigma_{LQ}(q) = (q, \underline{f})$. Via the local trivialization map, $Loc_{\sigma_{LQ}}(q, \underline{f}) = (q, id)$. Now let u be an arbitrary point in $\pi^{-1}(U) \subset LQ$. Using the section σ_{LQ} we identified that $\sigma_{LQ}(q)$ is on the same fiber as u and hence there exists $g \in Gl(n)$ such that $u = \sigma_{LQ}(q) \cdot g = (q, \underline{e}) \cdot g$. Denote this arbitrary point $u \in \pi_{LQ}^{-1}(U) \subset LQ$ by $u = (q, \underline{e})$ where $\underline{e} = \underline{f} \cdot g$. Using the local trivialization map, $Loc_{\sigma_{LQ}}(q, \underline{e}) = (q, g)$. We can think of $(q, \underline{e})$ as another frame field related to a given (reference) local frame field $(q, \underline{f})$ by $(q, \underline{e}) = (q, \underline{f} \cdot g)$. That is, the group element obtained from the local trivialization tells us how any frame field is related to the reference frame field.

We can now define coordinates on $\pi_{LQ}^{-1}(U) \subset LQ$ given a (reference) local frame field $(q, \underline{f})$. Let $u = (q, \underline{e})$ be an arbitrary element of $\pi_{LQ}^{-1}(U)$. Define (momentum) coordinates $\bar{q}^{\mathcal{K}}$ and $\Pi^{\mathcal{I}}_{\mathcal{J}}$ by

$$\begin{aligned}\bar{q}^{\mathcal{K}}(q,\underline{e}) &:= q^{\mathcal{K}}(q)\\ \Pi^{\mathcal{I}}_{\mathcal{J}}(q,\underline{e}) &:= e^{\mathcal{I}}(f_{\mathcal{J}}(q))\end{aligned} \tag{B.6}$$

where $f_{\mathcal{J}}(q)$ is the $\mathcal{J}-th$ vector from the reference frame field, $e^{\mathcal{I}}$ is the $\mathcal{I}-th$ co-vector of the arbitrary $\mathcal{I}$-th vector from the arbitrary frame field and $q^{\mathcal{K}}$ are defined in eq. (B.1). One can also define (velocity) coordinates $\bar{q}^{\mathcal{K}}$ and $V^{\mathcal{I}}_{\mathcal{J}}$ by

$$\begin{aligned}\bar{q}^{\mathcal{K}}(q,\underline{e}) &= q^{\mathcal{K}}(q)\\ V^{\mathcal{I}}_{\mathcal{J}}(q,\underline{e}) &= f^{\mathcal{I}}(e_{\mathcal{J}}(q))\end{aligned} \tag{B.7}$$

It follows from the definitions that

$$e^{\mathcal{J}} = \Pi^{\mathcal{J}}_{\mathcal{I}}(u) f^{\mathcal{I}} \;\; [\text{or invariantly } \bar{e} = \Pi \cdot \bar{f}] \tag{B.8}$$

$$e_{\mathcal{I}} = V^{\mathcal{J}}_{\mathcal{I}}(u) f_{\mathcal{J}} \;\; [\text{or invariantly } \underline{e} = \underline{f} \cdot V] \tag{B.9}$$

which are equations describing the $\Pi's$ and $V's$ role in relating the reference co-frame, $\bar{f}$ and reference frame, $\underline{f}$ to an arbitrary co-frame, $\bar{e}$ and arbitrary frame, $\underline{e}$ respectively. It also follows that

$$\begin{aligned}\Pi^{\mathcal{J}}_{\mathcal{K}}(u) V^{\mathcal{K}}_{\mathcal{I}}(u) &= \delta^{\mathcal{J}}_{\mathcal{I}}(u)\\ V^{\mathcal{K}}_{\mathcal{J}}(u) \Pi^{\mathcal{J}}_{\mathcal{I}}(u) &= \delta^{\mathcal{K}}_{\mathcal{I}}(u)\end{aligned} \tag{B.10}$$

and hence the $\Pi's$ and $V's$ are inverses of each other.

Remark B.2.1 *We note that the definitions given in eqs. (B.6) and (B.7) are slightly more general than those originally used by LKN [28], in that we allow for a reference frame field* $(p,\underline{e})$ *not necessarily equal to* $(q,\underline{\partial})$ *where* $\underline{\partial} = (\partial/\partial q^{\mathcal{I}}) = (\partial_{\mathcal{I}})$*. That is,* *LKN [28] uses*

$$\Pi^{\mathcal{I}}_{\mathcal{J}}(p, \underline{e}) = e^{\mathcal{I}}(\partial_{\mathcal{J}}(q))$$

which we will denote as $\overset{c}{\Pi}{}^{\mathcal{I}}_{\mathcal{J}}(q, \underline{e})$ *and call the canonical n-sympectic momenta. The generality will allow us to take as a reference frame field one which is in some sense adapted to the constraints and adapted to the fundamental vertical vectors, see Appendix F.*

Appendix C

More on LQ

C.1 TP and $T^{\star}P$, Associated Bundles of LP

This section can be summarized by the following two diagrams.

Summary C.1.1 *The diagram*

$$\begin{tikzcd}
u \in LP \arrow[dr, "\pi"'] & & v \in \mathbb{R}^n \arrow[r, shift left, "\phi_u"] & TP \cong E = LP \times_{Gl(n)} \mathbb{R}^n \ni [u,v] \arrow[l, shift left, "\phi_u^{-1}"] \arrow[dll, "\tau"] \\
& P & &
\end{tikzcd}$$

where ϕ_u is the well-defined mapping

$$\phi_u(v) = [u,v] \quad \ni \tau[u,v] = p$$

such that $[u,v] \in E$ is an equivalence class summarizes how TP can be viewed as a vector bundle associated to LQ. It follows that a section of τ, f is a smooth assignment of a vector X_p to each $p \in P$. That is f is a vector field on P.

Summary C.1.2 *The diagram*

$$
\begin{tikzcd}
u \in LP \arrow[dr, "\pi"'] & \alpha \in \mathbb{R}^{n\star} \arrow[r, shift left, "\psi_u"] & T^\star P \cong E = (LP \times_{GL(n)} \mathbb{R}^{n\star}) \ni [u,\alpha] \arrow[l, shift left, "\psi_u^{-1}"] \arrow[dl, "\sigma"] \\
& P &
\end{tikzcd}
$$

where ψ_u is the well-defined mapping

$$\psi_u(\alpha) = [u,\alpha] \quad \ni \sigma[u,\alpha] = p$$

such that $[u,\alpha]$ is an equivalence class summarizes how $T^\star P$ can be viewed as a vector bundle associated to LQ.. A section of σ is a co-vector field on P.

In section B.2 we have seen the sense in which LQ is considered to be a principle $\mathfrak{G}$=Gl(n)-bundle (right action of Gl(n) on LQ). A left action of Gl(n) on $V = \mathbb{R}^n$ can be defined by

$$L_g(v) := g^{-1} \cdot v \;\; \forall v \in \mathbb{R}^n$$

Remark C.1.1 *The above is a left action since*

$$L_{gh}(v) = (gh)^{-1} \cdot v = h^{-1}(g^{-1} \cdot v) = L_h(L_g v)$$

Remark C.1.2 *Coordinate formulas will periodically be presented which illustate various invariant definitions. For example,*

$$
\begin{aligned}
R_g(u) &:= (p, \underline{e} \cdot g) = (p, (f_\mathfrak{I} g^\mathfrak{I}_\mathfrak{J})) \; \forall g \in Gl(n), u \in LP \\
L_g(v) &:= g^{-1} \cdot v = ({}^{-1}g^\mathfrak{I}_\mathfrak{J} v^\mathfrak{J}) \; \forall g \in Gl(n), v \in \mathbb{R}^n
\end{aligned}
$$

where the row vector-matrix multiplication $f_\mathfrak{I} g^\mathfrak{I}_\mathfrak{J}$ gives $e_\mathfrak{J}$, the $\mathfrak{J}$-th frame vector of the frame $\underline{f} \cdot g := \underline{e}$ and the matrix-column vector multiplication ${}^{-1}g^\mathfrak{I}_\mathfrak{J} v^\mathfrak{J}$ gives $\tilde{v}^\mathfrak{I}$, the $\mathfrak{I}$-th *coordinate of the vector* $g^{-1} \cdot v := \tilde{v}$ *where we are using the index convention*

$$1 \leq \mathcal{I}, \mathcal{J}, \mathcal{K} \leq dim(P)=n$$

Comparing the above to equation (B.9) we see that

$$f_{\mathcal{I}} g^{\mathcal{I}}_{\mathcal{J}} = e_{\mathcal{J}}$$

is an equation relating the reference frame vector $f_{\mathcal{I}}$ *at* p *to an arbitrary frame vector* $e_{\mathcal{J}}$ *at* p *where*

$$g^{\mathcal{I}}_{\mathcal{J}}(p) = V^{\mathcal{I}}_{\mathcal{J}}((p, \underline{e})) \tag{C.1}$$

In [20, pg. 54] it is shown how given a PFB P(M,G) [*right action of* $mathfrakG$ *on P*] one can construct a vector bundle E(P,M,$\mathfrak{G}$,F=V) called *the associated bundle of P with standard fiber V* [*left action of G on V*]. As addressed above we have the ingredients to construct an associated bundle of LP with standard fiber $V = \mathbb{R}^n$. Following [20] we construct

$$E = (LP \times \mathbb{R}^n)/GL(n) = \{[u, v] | u \in LP \text{ and } v \in \mathbb{R}^n\}$$

where $[u, v], [\tilde{u}, \tilde{v}] \in E$ are equal if

1. p=$\tilde{p}$

2. there exists $g \in GL(n)$ such that $\tilde{u} = u \cdot g$ and $\tilde{v} = {}^{-1}g \cdot v \; \forall u \in LP$ and $v \in \mathbb{R}^n$.

That is, E is the collection of equivalence classes {[u,v]} where $[u, v] = \{(u \cdot g, g^{-1} \cdot v) \forall\ u \in LP, v \in \mathbb{R}^n, g \in Gl(n)\}$.

Remark C.1.3 *To get a feeling for the above equivalence, recall that two vector fields* $X = X^{\mathcal{I}} e_{\mathcal{I}}$ *and* $Y = Y^{\mathcal{J}} f_{\mathcal{J}} \in TP$ *are equal if they are related by the coordinate transform (the basis transform via the right group action and the coordinates transform* *via the left group action) given by*

$$X = X^{\mathfrak{I}} e_{\mathfrak{I}} = X^{\mathfrak{I}} \delta^{\mathfrak{J}}_{\mathfrak{I}} e_{\mathfrak{J}} = X^{\mathfrak{I}} ({}^{-1}g^{\mathfrak{K}}_{\mathfrak{I}} g^{\mathfrak{J}}_{\mathfrak{K}}) e_{\mathfrak{J}}$$
$$= {}^{-1}g^{\mathfrak{K}}_{\mathfrak{I}} X^{\mathfrak{I}} e_{\mathfrak{J}} g^{\mathfrak{J}}_{\mathfrak{K}} = Y^{\mathfrak{K}} f_{\mathfrak{K}} = Y \ \ \forall p$$

Substituting eq. (C.1) and using the relationship (B.10) implies

$$ {}^{-1}g^{\mathfrak{J}}_{\mathfrak{I}}(p) = \Pi^{\mathfrak{J}}_{\mathfrak{I}}(p, \underline{f}) \tag{C.2}$$

We will soon see the relationship between E and TM.

A critical formula can be derived using the equivalence relation defined on E. We ask for what $\tilde{v}$ the following will hold, $[g \cdot u, v] = [u, \tilde{v}]$. This implies there exists $h \in G$ for which $h \cdot (g \cdot u) = u$ and $h^{-1} \cdot v = \tilde{v} \implies h = g^{-1}$ and thus $\tilde{v} = g \cdot v$. We now have the critical formula (*for right actions of $\mathfrak{G}$ on P and left actions of $\mathfrak{G}$ on V*)

$$[u \cdot g, v] = [u, g \cdot v] \tag{C.3}$$

Now define $\tau : E \to P$ by

$$\tau([u, v]) := \pi(u) = p$$

As is always the case when defining a function on equivalence classes we must show *well-definedness*. That is, if $[u, v] = [\tilde{u}, \tilde{v}]$ then we must show that $\tau([u, v]) = \tau([\tilde{u} = u \cdot g, \tilde{v} = g^{-1} \cdot v])$. This follows from $\tau([u, v]) = \pi(u) = p$ and $\tau([\tilde{u}, \tilde{\xi}]) = \pi(u \cdot g) = p$ since R_g keeps u on the fiber over p. We consider the implications of $\tau([u, v]) := \pi(u) = p$. Let τ^{-1} be the fiber of $E \xrightarrow{\tau} P$ defined by

$$\tau^{-1}(p) = \{[u, v] | \tau([u, v]) = p\} \subset E.$$

We show that τ_p^{-1} is a vector space. Define addition of $[u,v],[\tilde{u},\tilde{v}] \in \tau^{-1}(p)$ by

$$\begin{aligned}[u,v]+[\tilde{u},\tilde{v}] &= [u,v]+[u\cdot g,\tilde{v}] \\ &= [u,v]+[u,g\cdot\tilde{v}] \text{ (by (C.3))} \\ &:= [u,v+g\cdot\tilde{v}]\end{aligned}$$

Now $[u,v+g\cdot\tilde{v}] \in \tau^{-1}(p)$ and hence τ_p^{-1} is a vector space. We now show that the fibers $\tau^{-1}(p)$ are isomorphic to $\mathbb{R}^n$ by *identifying a frame with a mapping*. That is, to each $u \in LM$ define $\phi_u : \mathbb{R}^n \to \pi^{-1}(p) \subset E$ by

$$\phi_u(v) = [u,v] \quad \text{where } \tau[u,v] = p \tag{C.4}$$

Define $\phi_u^{-1} : \tau^{-1}(p) \to \mathbb{R}^n$ by

$$\phi_u^{-1}([u,v]) = v \tag{C.5}$$

where $[u,v]$ is any point on the fiber $\tau^{-1}(p)$. Again, as this is a function defined on an equivalence class we must show that it is well-defined. That is, we have to show that if $[u,v]$ and $[\tilde{u},\tilde{v}]$ are on the same fiber $\tau^{-1}(p)$ then $\phi^{-1}([u,v]) = \phi^{-1}([\tilde{u},\tilde{v}]) = v$. By the equivalence relation and the crucial formula, $\phi^{-1}([\tilde{u},\tilde{v}]) = \phi^{-1}([u\cdot g, g^{-1}\cdot v]) = \phi^{-1}([u, gg^{-1}\cdot v]) = v$. Clearly ϕ_u is 1-1 and onto and now invertible which implies that $\tau^{-1}(p) \cong \mathbb{R}^n$. So we have shown that $E = (LP \times \mathbb{R}^n)/Gl(n)$ is a vector bundle over P with fiber $\tau^{-1}(p) \cong \mathbb{R}^n$.

That $E \cong TP$ follows using

$$[u,v] = [(p,\underline{e}),v] \cong X = v^{\mathfrak{I}} e_{\mathfrak{I}} \ \ \forall p \tag{C.6}$$

which says that the vector field X is formed using the components of $v^{\mathfrak{I}}$ of $v = v^{\mathfrak{I}} r_{\mathfrak{I}} \in V = \mathbb{R}^n$ summed out with the $\mathfrak{I}$-th frame vector $e_{\mathfrak{I}}$.

Remark C.1.4 *We note that the vector v and frame u used in eq. (C.6) are those of the represenative $(u,v) = ((p,\underline{e}),v) \in [u,v]$. Were any other $(\tilde{u},\tilde{v}) \in [u,v]$ chosen* *it would have defined a vector field Y equal (in the sense of Remark C.1.3) to X by*

$$X = v^{\mathfrak{I}} e_{\mathfrak{I}} \cong [u, v] = [u \cdot g, g^{-1} \cdot v] \cong e_{\mathfrak{I}} g^{\mathfrak{I}}_{\mathfrak{J}} ({}^{-1}g^{\mathfrak{J}}_{\mathfrak{I}} v^{\mathfrak{I}}) = \tilde{v}^{\mathfrak{J}} f_{\mathfrak{J}} = Y \ \ \forall p$$

and thus the mapping $[u, v] \to X$ is well-defined. A specific vector X_p will be obtained once a specific frame at p and specific $v \in V$ are chosen.

Other coordinate formulas are

$$\phi_u^{-1}([u, v]) = v \to \phi_{f_{\mathfrak{J}}}^{-1}(v^{\mathfrak{I}} f_{\mathfrak{I}}) = v^{\mathfrak{J}} r_{\mathfrak{J}} \tag{C.7}$$

$$\phi_u(v) = [u, v] \to \phi_{f_{\mathfrak{J}}}(v^{\mathfrak{I}} r_{\mathfrak{I}}) = v^{\mathfrak{J}} f_{\mathfrak{J}} \tag{C.8}$$

$$\phi_{u \cdot g}^{-1}([u, v]) = g^{-1} \cdot v \to \phi_{e_{\mathfrak{J}}}^{-1}(v^{\mathfrak{I}} f_{\mathfrak{I}}) = {}^{-1}g^{\mathfrak{J}}_{\mathfrak{I}} v^{\mathfrak{I}} r_{\mathfrak{J}} \tag{C.9}$$

$$\phi_{u \cdot g}(v) = [u, g^{-1} \cdot v] \to \phi_{e_{\mathfrak{J}}}(v^{\mathfrak{I}} r_{\mathfrak{I}}) = {}^{-1}g^{\mathfrak{J}}_{\mathfrak{I}} v^{\mathfrak{I}} f_{\mathfrak{J}} \tag{C.10}$$

where $\{r_{\mathfrak{I}}\}$ is a basis for $\mathbb{R}^n$ and $\underline{f} = (f_{\mathfrak{J}})$ is the (reference) frame field and $\underline{e} = \{e_{\mathfrak{I}}\}$ is another arbitrary frame field.

As expected $T^{\star}Q$ can also be constructed as an associated bundle of LQ. The construction is

$$E = (LP \times \mathbb{R}^{n\star})/GL(n) \quad (\alpha \in \mathbb{R}^{n\star}).$$

where equivalence is given by $[u, \alpha] = [\tilde{u}, \tilde{\alpha}] = [u \cdot g, \alpha \cdot g]$ (such that $L_g \alpha = \alpha \cdot g$ is a left action since $L_{gh}(\alpha) = \alpha \cdot gh = L_h(L_g \alpha)$). This notion of equivalence leads to the critical formula (cf. arguement leading to eq. (C.3)) $[u \cdot g, \alpha] = [u, \alpha \cdot g^{-1}]$. Analagous to the TP case one defines $\sigma([u, \alpha]) := \pi(u) = p$ and $\psi_u : \sigma^{-1}(p) \to \mathbb{R}^{n\star}$ is given by $\psi_u(\alpha) = [u, \alpha]$ with inverse $\psi_u^{-1}([u, \alpha]) = \alpha$ which gives $\psi^{-1}(p) \cong \mathbb{R}^{n\star}$. Analagous to (C.9) and (C.10) we have that

$$\psi_{u \cdot g}^{-1}([u, \alpha]) = \alpha \cdot g \to \psi_{e_{\mathfrak{J}}}^{-1}(\alpha_{\mathfrak{I}} f^{\mathfrak{I}}) = g^{\mathfrak{I}}_{\mathfrak{J}} \alpha_{\mathfrak{I}} r^{\mathfrak{J}} \tag{C.11}$$

$$\psi_{u \cdot g}(\alpha) = [u, \alpha \cdot g] \to \psi_{e_{\mathfrak{J}}}(\alpha_{\mathfrak{I}} r^{\mathfrak{I}}) = g^{\mathfrak{I}}_{\mathfrak{J}} \alpha_{\mathfrak{I}} f^{\mathfrak{J}} \tag{C.12}$$

That $E \cong T^\star P$ follows using

$$[u, \alpha] \cong \Theta = \alpha_{\mathfrak{I}} e^{\mathfrak{I}}$$

which says that co-vector Θ is formed using the components of $\alpha_{\mathfrak{I}}$ of $\alpha = \alpha_{\mathfrak{I}} r^{\mathfrak{I}} \in V = \mathbb{R}^{n\star}$ and summed out with the $f^{\mathfrak{I}}$ the co-frame of the (reference) frame defined by $u = (p, \underline{f} = (f_{\mathfrak{I}})) \in LP$. A smooth section of λ is a co-vector field on P.

Remark C.1.5 *In a very natual way one can put the diagrams for TP and $T^\star P$ together into a single diagram for a new associated vector bundle $\tilde{E}$*

$$\begin{tikzcd}
LP \arrow[dr, "\pi"'] & \mathbb{R}^n \otimes \mathbb{R}^{n\star} \arrow[r, shift left, "\xi_u"] & \tilde{E} = TP \otimes T^\star P \arrow[l, shift left, "\xi_u^{-1}"] \arrow[dl, "\lambda"] \\
& P &
\end{tikzcd}$$

where $\xi_u(v \otimes \alpha) := (\phi_u \otimes \psi_u)(v \otimes \alpha) = \phi_u(v) \otimes \psi_u(\alpha) = [u, v] \otimes [u, \alpha] \cong X \otimes \Theta$ and $\lambda(X \otimes \Theta) := (\tau \otimes \sigma)(X \otimes \Theta) = \tau(X) \otimes \sigma(\Theta) = p \otimes p \sim p$. A smooth section of λ is a 1-1 tensor field on P.

Remark C.1.6 *Extrapolating Remark C.1.5 to its full generality gives the diagram*

$$\begin{tikzcd}
LP \arrow[dr, "\pi"'] & \mathbb{R}^n \otimes_s^r \mathbb{R}^{n\star} \arrow[r, shift left] & \bar{E} \cong TP \otimes_s^r T^\star P \arrow[l, shift left] \arrow[dl, "\gamma"] \\
& P &
\end{tikzcd}$$

where a section of γ is an (r,s) tensor field on P.

C.2 Tensorial p-forms on LP

We begin with the general definition of a *pseduo-tensorial p-form* on P(M,G) as given [20, pg. 75]

Definition C.2.1 *Let P(M,$\mathfrak{G}$) be principle $\mathfrak{G}$-bundle [right action of $\mathfrak{G}$ on P] and let ρ be a representation of $\mathfrak{G}$ on V. A* pseudo-tensorial p-form on P of type (ρ,V) *is a V-valued p-form ψ such that*

$$R_g^\star \psi = \rho(g^{-1}) \cdot \psi \ \forall g \in \mathfrak{G}$$

The p-form ψ is said to be tensorial *if $\psi(X_1, \ldots, X_p) = 0$ whenever $d\pi(X_i) = 0$ for some $1 \leq i \leq dim(P)$.*

We focus on tensorial 0-forms and tensorial 1-forms in the case where LP(P,Gl(n)) and ρ is the standard representation of Gl(n) on $V = \mathbb{R}^n$. That is,

$$\rho(g^{-1}) \cdot \xi = g^{-1} \cdot \xi$$

C.2.1 Tensorial 0-Forms on LP

First we discuss $\hat{f} : LM \to \mathbb{R}^n$ a tensorial 0-form. Note that $\hat{f}$ fills in the gap of a previous diagram where we define f to be a (smooth) section of τ (i.e. f (smooth) defined pointwise by $p \to X_p$ such that $\tau(X_p) = p$). That is, the section f of τ smoothly assigns to each p a specific vector X_p and hence can be considered a vector field.

$$
\begin{tikzcd}
LP \arrow[rr, "\hat{f}"] \arrow[drr, "\pi"'] & & \mathbb{R}^n \arrow[r, shift left, "\phi_u"] & TP \cong E \arrow[l, shift left, "\phi_u^{-1}"] \\
& & P \arrow[ur, "f(section)"'] &
\end{tikzcd}
$$

Recall from section C.1 that $\tau^{-1} : P \to E \cong TP$, $\phi_u : \mathbb{R}^n \to \tau^{-1}(p) = f(p)$ and $\phi_u^{-1} : f(p) \to \mathbb{R}^n$ were defined pointwise by

$$
\begin{aligned}
\tau^{-1}(p) &= [u, v] \text{ such that } \tau[u, v] = p \\
\phi_u^{-1}([u, v]) &= v \\
\phi_u(v) &= [u, v] \cong X
\end{aligned}
$$

where $E \cong TM$ via $[(p, \underline{e}, v] \cong v^{\mathfrak{I}} e_{\mathfrak{I}} = X_p$. Specifying the frame u at a given p and a vector v determines X_p a specific element of T_pP. The relationship between a section f and a tensorial function, $\hat{f}$ is given by the

Theorem C.2.1 *Sections of τ and tensorial 0-forms are in one-to-one correspondence.*

Proof: *First let $f : M \to TM$ be a section. Define (pointwise), $\hat{f} : LM \to \mathbb{R}^n$ by*

$$\hat{f}_u := \phi_u^{-1}(f_p) \tag{C.13}$$

where f_p is a point on the same fiber of $E \cong TM$ over p as $[u, v] \in \tau^{-1}(p)$. That is f_p can be expressed as $f_p = [\bar{u}, \bar{v}]$ where $\tau(f_p) = \tau([\bar{u}, \bar{v}]) = p$. As $[\bar{u}, \bar{v}]$ is on the same fiber as $[u, v]$ there exists an $h \in Gl(n)$ such that $f_p = [u \cdot h, \bar{v}] = [u, h \cdot \bar{v}] \implies \hat{f}_u = \phi_u^{-1}(f_p) = h \cdot \bar{v}$. To show $\hat{f}$ is tensorial we must show (again pointwise)

$$(R_g^\star \hat{f})|_u = g^{-1} \cdot \hat{f}_u.$$

or, since the pull-back of a function f *through* R_g *is* $f \circ R_g$

$$\hat{f}_{u \cdot g} = g^{-1} \cdot \hat{f}_u$$

By the above definition,

$$\begin{aligned}
\hat{f}_{u \cdot g} &= \phi^{-1}_{u \cdot g}(f_p) = \phi^{-1}_{u \cdot g}([\bar{u}, \bar{v}]) \\
&= \phi^{-1}_{u \cdot g}([u, h \cdot \bar{v}]) \\
&= \phi^{-1}_{u \cdot g}[u \cdot gg^{-1}, h \cdot \bar{v}] \\
&= \phi^{-1}_{u \cdot g}[u \cdot g, g^{-1} \cdot (h \cdot \bar{v})] \\
&= g^{-1} \cdot (h \cdot \bar{v}) = g^{-1} \cdot \hat{f}_u
\end{aligned}$$

Thus we have shown that a section f *defines a tensorial function* $\hat{f}$*. Now given* $\hat{f} : LM \to \mathbb{R}^n$ *(tensorial) define* $f : P \to TP$ *(pointwise) by*

$$f_p := \phi_u(\hat{f}_u). \tag{C.14}$$

We just need to assure that two frames which are equivalent get mapped to the same vector. That is, we must show

$$\phi_{\tilde{u}}(\hat{f}_{\tilde{u}}) = \phi_u(\hat{f}_u) \ \ \textit{where} \ \ \tilde{u} = u \cdot g$$

Now

$$\begin{aligned}
\phi_{\tilde{u}}(\hat{f}_{\tilde{u}}) = \phi_{u \cdot g}(\hat{f}_{u \cdot g}) &= \phi_{u \cdot g}(g^{-1} \cdot \hat{f}_u) \ \ \textit{(using } \hat{f} \textit{ tensorial)} \\
&= [u \cdot g, g^{-1} \cdot \hat{f}_u] \\
&= [u, gg^{-1} \cdot \hat{f}_u] \\
&= \phi_u(\hat{f}_u) = [u, \hat{f}_u]
\end{aligned}$$

Thus we have shown that a tensorial function $\hat{f}$ *gives a* well defined *section* f *which* *completes the one-to-one correspondence of sections of τ with $\mathbb{R}^n$-valued tensorial functions on LP.* $\square$

Remark C.2.1 *It is readily apparent that one can generalize the above theorem to arbitrary tensor fields. That is, sections of γ (cf. Summary (i.e. (r,s)-tensor fields) are in one-to-one correspondence with $\mathbb{R}^n \otimes_s^r \mathbb{R}^{n\star}$ valued tensorial functions on LP. This is illustrated in the (1,1) tensor field case by the diagram (cf. Remark C.1.5)*

$$\begin{tikzcd} LP \arrow[r, "\hat{s}"] \arrow[dr, "\pi"'] & \mathbb{R}^n \otimes \mathbb{R}^{n\star} \arrow[r, shift left, "\xi_u"] & \tilde{E} = TP \otimes T^\star P \arrow[l, shift left, "\xi_u^{-1}"] \\ & P \arrow[ur, "s"'] & \end{tikzcd}$$

where a section of λ, s , is a (1,1) tensor field on P given by the formula

$$s_p = \xi_u(\hat{s}_u) \tag{C.15}$$

Given a (1,1) tensor field s_p there corresponds a $\mathbb{R}^n \otimes \mathbb{R}^{n\star}$-valued tensorial function on LP given by the formula

$$\hat{s}_u = \xi_u^{-1}(s_p). \tag{C.16}$$

C.2.2 Tensorial One Forms on LP

Theorem C.2.2 *Tensorial one forms on LP are in one-to-one correspondence with (1,1)-tensor fields on TP.*

Proof: *First, let $\hat{\psi} : T(LP) \to \mathbb{R}^n$ be a tensorial one-form on LP given point-wise by*

$$\hat{\psi}_u \hat{X}_u \in \mathbb{R}^n$$

where we assume

$$\hat{X}_u \in T_u LP \quad s.t. \quad d\pi(\hat{X}_u) = X_p \in T_p P \tag{C.17}$$

Now define $\psi : P \to T_1^1(TP)$ *given* $\hat{\psi}$ *and the map* ϕ_u *defined in (C.4) by*

$$\psi_p X_p := \phi_u(\hat{\psi}_u \hat{X}_u) = [u, \hat{\psi}_u \hat{X}_u] \in T_p P. \tag{C.18}$$

As ψ_p *maps into an equivalence class (via* $TP \cong E$*) we must show it is a well-defined function. That is, were another tangent vector to* LP *which projects to* X_p *given would we obtain the same value for* $\psi_p X_p$*. Notice that* $\hat{X}_{u\cdot g} = dR_g \hat{X}_u \in T_{u\cdot g} LP$ *is an arbitrary vector which projects to* X_p *as*

$$d\pi(\hat{X}_{u\cdot g}) = d\pi(dR_g \hat{X}_u) = d(\pi \circ R_g)\hat{X}_u = \hat{X}_{(\pi\circ R_g)(u)} = \hat{X}_p$$

it follows that

$$\begin{aligned}
\phi_{u\cdot g}(\hat{\psi}_{u\cdot g}\hat{X}_{u\cdot g}) &= \phi_{u\cdot g}(g^{-1}\cdot \hat{\psi}_u \hat{X}_u) \;\; (\textit{using } \hat{\psi} \textit{ tensorial}) \\
&= [u\cdot g, g^{-1}\cdot \hat{\psi}_u \hat{X}_u] \\
&= [u, g\cdot g^{-1}\cdot \hat{\psi}_u \hat{X}_u] \;\; (\textit{using (C.3)}) \\
&= [u, \hat{\psi}_u \hat{X}_u] = \psi_p X_p
\end{aligned}$$

so the map $\psi_p : T_p P \to T_p P$ *defined through the given map* $\hat{\psi}_u : T_u LP \to \mathbb{R}^n$ *is well-defined for each* p *and hence* $\psi : P \to T_1^1(TP)$ *is well defined. Now let* $\psi : P \to T_1^1(TP)$ *be given pointwise by*

$$\psi_p X_p \cong [\bar{u}, \bar{v}] = [u, h\cdot \bar{v}] \in T_p P \quad \textit{for some } h \in Gl(n)$$

Given this map ψ*, the map* ϕ^{-1} *from (C.5) and again assuming that* $d\pi(\hat{X}_u) = X_p$ *we define the function* $\hat{\psi} : T(LP) \to \mathbb{R}^n$ *pointwise by*

$$\hat{\psi}_u \hat{X}_u = \phi_u^{-1}(\psi_p(d\pi \hat{X}_u)) \tag{C.19}$$

We now show that $\hat{\psi}_u$ is indeed tensorial. We must show that

$$\hat{\psi}_{u\cdot g}(\hat{X}_{u\cdot g}) = g^{-1} \cdot \hat{\psi}_u \hat{X}_u \tag{C.20}$$

The right side of (C.20) works out to (since $\hat{X}_u$ projects to X_p)

$$g^{-1} \cdot \phi_u^{-1}([u, h \cdot \bar{v}]) = g^{-1} \cdot (h \cdot \bar{v}).$$

The left hand side of (C.20) works out to (since $\hat{X}_{u\cdot g}$ projects to X_p)

$$\begin{aligned} \phi_{u\cdot g}^{-1}(\psi_p X_p) &= \phi_{u\cdot g}^{-1}([u, h \cdot \bar{v}])) \\ &= \phi_{u\cdot g}^{-1}([u \cdot g, g^{-1}(h \cdot \bar{v})]) \\ &= g^{-1} \cdot (h \cdot \bar{v}) \end{aligned}$$

and thus $\hat{\psi}_p : T_pLP \to \mathbb{R}^n$ is a tensorial function for each p and hence $\hat{\psi} : T(LP) \to \mathbb{R}^n$ is a tensorial function. We have now shown that tensorial one-forms on LP are in one-to-one correspondence with (1,1)-tensor fields on TP. $\square$

Summary C.2.1 We summarize tensorial 1-forms on LP with the diagram

$$\begin{tikzcd}
T(LP) \arrow[dr, "\hat{\psi}"] & & & \\
LM \arrow[r, "\hat{f}"] \arrow[dr, "\pi"'] & \mathbb{R}^n \arrow[r, shift left, "\phi_u"] & TP \arrow[l, shift left, "\phi_u^{-1}"] \arrow[r, "\psi"] \arrow[dl, "\tau"] & TP \\
& P & &
\end{tikzcd}$$

C.3 Soldering Form on LP

The soldering form is an $\mathbb{R}^p$-valued one form on LP defined by

$$\hat{\theta}_u(\hat{X}_u) \triangleq \phi_u^{-1}(d\pi(\hat{X}_u)) \tag{C.21}$$

where (as in C.2.2) we require $d\pi X_u = X_p \cong ([(q, \underline{f}), v]$ (i.e. X_p is expressed relative to the reference frame) and that $\tau([u, v]) = p)\ \forall p$.

Note C.3.1 *The canonical 1-form on $T^\star Q$ is given by*

$$\vartheta_{[q,\alpha]}(X_{[q,\alpha]}) = \alpha(d\pi X_{[q,\alpha]}) = \alpha(X_q)$$

which in canonical local coordinates on $T^\star Q$ given by

$$\bar{q}^{\mathfrak{I}}([q, \alpha]) = (q^{\mathfrak{I}} \circ \pi)([q, \alpha]) = q^{\mathfrak{I}}(q) \quad \textit{and} \;\; p_{\mathfrak{I}}([q, \alpha]) = \alpha_q \left(\frac{\partial}{\partial x^{\mathfrak{I}}} | q \right)$$

can be written as

$$\begin{aligned} \alpha_q(X^{\mathfrak{I}}(q)\partial_{\mathfrak{I}}) &= p_{\mathfrak{I}}([q, \alpha])X^{\mathfrak{I}}(q) \\ &= p_{\mathfrak{I}}([q, \alpha])dq^{\mathfrak{I}}(X_q) \\ &= p_{\mathfrak{I}}([q, \alpha])d\bar{q}^{\mathfrak{I}}(X_{[q,\alpha]}) \\ \Longrightarrow &\quad \vartheta_{[q,\alpha]} = p_{\mathfrak{I}}([q, \alpha])d\bar{q}^{\mathfrak{I}}|_{[q,\alpha]} \end{aligned}$$

which implies ϑ as an $\mathbb{R}$-valued one-form on $T^\star Q$ can be written

$$\vartheta = p_{\mathfrak{I}} \wedge d\bar{q}^{\mathfrak{I}}.$$

That $\hat{\theta}_u$ is a tensorial one-form on LP follows as in the proof from C.2.2. The corresponding (1,1)-tensor field on TP is (using eq. (C.18))

$$\theta_p X_p = \phi_u(\hat{\theta}_u \hat{X}_u) = \phi_u(\phi_u^{-1} X_p) = X_p \tag{C.22}$$

which implies that the soldering form $\hat{\theta}_u$ corresponds to the identity (1,1) tensor, Id_p on T_pP for each p. Hence $\hat{\theta}$ corresponds to the identity tensor field on TP. Using the coordinate formulation in eq. (C.9) and (C.2) we obtain a coordinate formulation of (C.21) given by

$$\hat{\theta}_{(p,e_{\mathcal{J}})}(v^{\mathfrak{I}} f_{\mathfrak{I}}) = \Pi^{\mathcal{J}}_{\mathfrak{I}}((p,\underline{e}))v^{\mathfrak{I}} r_{\mathcal{J}} \in \mathbb{R}^n$$

or as an $\mathbb{R}^n$-valued one form on LP by

$$\hat{\theta}_{(p,\underline{e})} = \Pi^{\mathcal{J}}_{\mathfrak{I}}((p,\underline{e}))f^{\mathfrak{I}} \otimes r_{\mathcal{J}} \in T^{\star}_p(LP). \tag{C.23}$$

Relative to the canonical reference frame $\underline{f} = \underline{\partial}$, eq. (C.23) is

$$\hat{\theta}_u = \Pi^{\mathcal{J}}_{\mathfrak{I}}(u)dq^{\mathfrak{I}} \otimes r_{\mathcal{J}}$$

whereupon removing the u dependence and the range index gives

$$\hat{\theta}^{\mathfrak{I}} = \Pi^{\mathfrak{I}}_{\mathcal{J}} \wedge dq^{\mathcal{J}} \tag{C.24}$$

which should be compared with the above $T^{\star}Q$ version. A coordinate formulation of θ_p is $Id_p = e^{\mathfrak{I}} \otimes e_{\mathfrak{I}}$ and so we have the correspondence

$$Id_p = f^{\mathfrak{I}} \otimes f_{\mathfrak{I}}|_p \iff \hat{\theta}_u = \Pi^{\mathcal{J}}_{\mathfrak{I}}(u)f^{\mathfrak{I}}|_u \otimes r_{\mathcal{J}} \tag{C.25}$$

C.4 More LQ Coordinate Formulations

We now show the coordinate relationship between $\mathbb{R}^n \otimes^r_s \mathbb{R}^{n\star}$–valued tensorial functions on LP and sections of $TP \otimes^r_s T^{\star}P$ (cf. Summary C.2.1) using the example of sections of $TP \otimes T^{\star}P$ (cf. Remark C.1.5).

Using the definition of the map $\xi_u^{-1} := \phi_u^{-1} \otimes \psi_u^{-1} : T^1_1 M \to \mathbb{R}^{n\star} \otimes \mathbb{R}^n$ from Remark

C.1.5 and eqs. (C.12), (C.10) it follows that

$$\begin{aligned}
\xi^{-1}_{u\cdot g}([u,v]\otimes[u,\alpha]) &:= (\phi^{-1}_{u\cdot g}\otimes\psi^{-1}_{u\cdot g})([u,v]\otimes[u,\alpha]) \\
&= \phi^{-1}_{u\cdot g}([u,v])\otimes\psi^{-1}_{u\cdot g}([u,\alpha])) \\
&= g^{-1}\cdot v\otimes\alpha\cdot g \\
&= \Pi\cdot v\otimes\alpha\cdot V
\end{aligned}$$

and thus given a 1-1 tensor $T(p) = T^{\mathcal{I}}_{\mathcal{J}}(p)(e_{\mathcal{I}}\otimes e^{\mathcal{J}})$ on T_pP we have that $\hat{T}(u) = \xi^{-1}_u(T(p))$ can be computed in coordinates as

$$\begin{aligned}
\xi^{-1}_u(T^{\mathcal{I}}_{\mathcal{J}}(p)(e_{\mathcal{I}}\otimes e^{\mathcal{J}})) &= T^{\mathcal{I}}_{\mathcal{J}}(p)\phi^{-1}_{f_{\mathcal{N}}}(e_{\mathcal{I}})\psi^{-1}_{f_{\mathcal{M}}}(e^{\mathcal{J}}) \\
&= T^{\mathcal{I}}_{\mathcal{J}}(p)\Pi^{\mathcal{N}}_{\mathcal{I}}(u)r_{\mathcal{N}}V^{\mathcal{J}}_{\mathcal{M}}(u)r^{\mathcal{M}} \\
&= T^{\mathcal{I}}_{\mathcal{J}}(p)\Pi^{\mathcal{N}}_{\mathcal{I}}(u)V^{\mathcal{J}}_{\mathcal{M}}(u)r_{\mathcal{N}}\otimes r^{\mathcal{M}} \\
&:= \hat{T}^{\mathcal{N}}_{\mathcal{M}}(u)r_{\mathcal{N}}\otimes r^{\mathcal{M}}
\end{aligned}$$

We now have the coordinate correspondence of (1-1) tensor fields of TP with $\mathbb{R}^n\otimes^r_s\mathbb{R}^{n\star}$–valued tensorial functions on LP:

$$T^{\mathcal{I}}_{\mathcal{J}}(p)\longrightarrow T^{\mathcal{I}}_{\mathcal{J}}(p)\Pi^{\mathcal{N}}_{\mathcal{I}}(u)V^{\mathcal{J}}_{\mathcal{M}}(u) := \hat{T}^{\mathcal{N}}_{\mathcal{M}}(u) \tag{C.26}$$

It follows by extension that (0,2) tensor fields on TP correspond to $\mathbb{R}^{n\star}\otimes\mathbb{R}^{n\star}$–valued tensorial functions on LP given by

$$g_{\mathcal{I}\mathcal{J}}(p)\longrightarrow\hat{g}_{\mathcal{N}\mathcal{M}}(u) = g_{\mathcal{I}\mathcal{J}}(p)V^{\mathcal{I}}_{\mathcal{N}}(u)V^{\mathcal{J}}_{\mathcal{M}}(u) \tag{C.27}$$

and (2,0) tensor fields correspond to $\mathbb{R}^n\otimes\mathbb{R}^n$-valued tensorial function on LP given by

$$g^{\mathcal{I}\mathcal{J}}(p)\longrightarrow\hat{g}^{\mathcal{N}\mathcal{M}}(u) = g^{\mathcal{I}\mathcal{J}}(p)\Pi^{\mathcal{N}}_{\mathcal{I}}(u)\Pi^{\mathcal{M}}_{\mathcal{J}}(u) \tag{C.28}$$

Remark C.4.1 *The equation (C.27) is very important in the following sense: take* *the metric $g_{\mathcal{I}\mathcal{J}}(p)$ on P to be the kinetic energy metric then $\hat{g}_{\mathcal{N}\mathcal{M}}(u)$ can be viewed as the corresponding kinetic energy metric on LP. In Chapter 4.2 and Appendix G we derive dynamic equations on LP with $\hat{g}_{\mathcal{N}\mathcal{M}}(u)$ just as one would derive dynamic equations with $g_{\mathcal{I}\mathcal{J}}$ on P.*

Remark C.4.2 *We can effectively say that $\hat{g}$ is a metric tensor on LP and hence we can use it to* raise *and* lower *the indices of tensorial functions on LP. For example, $\hat{g}_{\mathcal{N}\mathcal{M}}\hat{T}^{\mathcal{N}}_{\mathcal{P}} = \hat{T}_{\mathcal{M}\mathcal{P}}$ since*

$$\begin{aligned}
\hat{g}_{\mathcal{N}\mathcal{M}}\hat{T}^{\mathcal{N}}_{\mathcal{P}} &= g_{\mathcal{I}\mathcal{J}}V^{\mathcal{I}}_{\mathcal{N}}V^{\mathcal{J}}_{\mathcal{M}}T^{\mathcal{A}}_{\mathcal{B}}\Pi^{\mathcal{N}}_{\mathcal{A}}V^{\mathcal{B}}_{\mathcal{P}} \\
&= g_{\mathcal{I}\mathcal{J}}T^{\mathcal{A}}_{\mathcal{B}}V^{\mathcal{I}}_{\mathcal{N}}V^{\mathcal{J}}_{\mathcal{M}}\Pi^{\mathcal{N}}_{\mathcal{A}}V^{\mathcal{B}}_{\mathcal{P}} \\
&= g_{\mathcal{I}\mathcal{J}}T^{\mathcal{A}}_{\mathcal{B}}\delta^{\mathcal{I}}_{\mathcal{A}}V^{\mathcal{J}}_{\mathcal{M}}V^{\mathcal{B}}_{\mathcal{P}} \\
&= g_{\mathcal{I}\mathcal{J}}T^{\mathcal{I}}_{\mathcal{B}}V^{\mathcal{J}}_{\mathcal{M}}V^{\mathcal{B}}_{\mathcal{P}} \\
&= T_{\mathcal{J}\mathcal{B}}V^{\mathcal{J}}_{\mathcal{M}}V^{\mathcal{B}}_{\mathcal{P}} \\
&= \hat{T}_{\mathcal{M}\mathcal{P}}
\end{aligned}$$

Similarily $\hat{g}^{\mathcal{N}\mathcal{M}}$ will raise the index of a covariant tensorial function on LP.

Appendix D

On $BM_{inj(M)}$

D.1 Summary of the Tangent Bundle as an Associated (Vector) Bundle of the (Principle Fibre) Bundle of Linear Frames

In [20, pg. 54] it is shown that given a principle fibre bundle $P(M, G)$ [*right action of G on M*] one can construct a vector bundle $E(P, V) = (P \times V)/G$ called *the associated bundle of P with standard fiber V* [*left action of G on V*]. Specifically the tangent bundle TM is vector bundle over M with fiber isomorphic to $\mathbb{R}^n$ associated to the principle fiber bundle LM, the bundle of linear frames. The bundle maps can be summarized by the diagram

$$\begin{tikzcd}
u \in LM \arrow[dr, "\pi"'] & v_p \in \mathbb{R}^n_p \arrow[r, shift left, "\phi_u"] & \tau^{-1}(p) \subset E = (LM \times \mathbb{R}^n)/Gl(n) \cong TM \ni X \arrow[l, shift left, "\phi_u^{-1}"] \arrow[dl, "\tau"] \\
 & p \in M &
\end{tikzcd}$$

A summary of all objects and maps are as follows:

- (The Bundle of Linear Frames) $LM = \{u = (p, \underline{f})\}$ where $\underline{f} = \{f_i\}_{i=1}^{dim(M)}$

is any basis of T_pM. That is, LM is the collection of all the frames of M for all points of M.

- (Projection of LM onto M) $\pi(u) = p \;\; \forall u \in LM$
- (The Associated Bundle of LM) $E = (LM \times \mathbb{R}^n)/GL(n) = \{[u,v] | u \in LM \ni X$ and $v \in \mathbb{R}^n\}$ (i.e. a set of Gl(n) equivalence classes) where $[u,v], [\tilde{u},\tilde{v}] \in E$ are equal if

  1. p=$\tilde{p}$
  2. there exists $g \in GL(n)$ such that $\tilde{u} = u \cdot g$ and $\tilde{v} = {}^{-1}g \cdot v \; \forall u \in LP$ and $v \in \mathbb{R}^n$ *(Note the right action of G on LM and the left action of G on* $\mathbb{R}^n$)

  A most critical formula can be derived from 2. (see eq. (C.3))

  $$[u \cdot g, v] = [u, g \cdot v] \tag{D.1}$$

- (Projection of E onto M) $\tau([u,v]) = \pi(u) = p$ is well-defined.
- (Standard Fibres of E) $\tau^{-1}(p) = \{[u,v] | \tau([u,v]) = p$ is a vector space
- (Identifying a Frame with a Mapping) $\phi_u(v) = [u,v] \in \tau^{-1}(p)$ and $\phi_u^{-1}([u,v]) = v \in \mathbb{R}^n_p$ (i.e. a copy of $\mathbb{R}^n$ for each $p \in M$. (this is well defined)
- (Identification of E with TM) $E \cong TM$ follows using

  $$[u,v] = [(p, \underline{e}), v] \cong X = v^i e_i \;\; \forall \;\; p \in M$$

  which says that the vector field X_p is formed using the components of $v^i(p)$ of $v_p = v^i(p) r_i|_p \in V = \mathbb{R}^n$ ($\{r_i|_p\}$ basis of $\mathbb{R}^n_p$) summed out with the frame vectors $e_i|_p$. The mapping $[u,v] \to v^i e_i$ is well-defined.

For the details of the above points (e.g. coordinates on LM and the well-definedness proofs) see Appendices B and C.

D.2 $BM|_{<Inj(M)}$ a Fibre Bundle over M

Theorem D.2.1 *Let (M,g) be a Riemannian manifold and let* $OM = \{u = (p, \underline{f}) | g(f_i, f_j) = \delta_{ij} \;\; \forall \;\; \underline{f} = \{f_i\} \in T_pM \;\; and \;\; \forall \;\; p \in M\}$ *be the orthonormal frame bundle where OM is the collection of all orthonormal frames of M for all points in M. OM is a subbundle of LM with structure group O(n), the set of orthogonal matrices as shown in [KNvol1, Example 5.7 pg. 60]. Define*

$$BM|_{<Inj(M)} := \{X \in TM | \sqrt{g(X,X)} = |X| < Inj(M)]\}.$$

$BM|_{<Inj(M)}$ *is a fibre bundle over M with fiber isomorphic to the submanifold* $B_{Inj(M)}(0)$*. We define the ball with respect to the standard (Euclidean) inner product in* $\mathbb{R}^n$*. Thus* $B_{Inj(M)}(0)$ *is a (n-1)-sphere of radius Inj(M) centered at* $0 \in \mathbb{R}^n$*. The bundle maps are summarized by the diagram:*

$$\begin{tikzcd} u \in OM \arrow[dr, "\pi"'] & B_{Inj(M)}(0) \arrow[r, shift left, "\phi_u"] & \tau^{-1}(p) \subset E = (OM \times B_{Inj(M)}(0))/O(n) \cong BM|_{<Inj(M)} \ni X_{<Inj(M)} \arrow[l, shift left, "\phi_u^{-1}"] \arrow[dl, "\tau"] \\ & p \in M & \end{tikzcd}$$

Proof: The proof is essentially an explanation of the modifications made to the diagram from Section D.1. Essentially Gl(n) is replaced by O(n) to guarantee that vectors in $BM|_{Inj(M)}(0)$ stay in $BM|_{Inj(M)}(0)$ which follows from O(n), by definition, preserving the standard Euclidean inner product. That we have a Euclidean inner product on each $BM|_{<Inj(M)}$ when we restrict to only orthonormal frames will complete the proof. To be specific:

- $E = \{[u, v] \mid u \in OM$ and $v \in B_{Inj(M)}(0)\}$ with equivalence classes given by $[u, v] = \{u \cdot O, O^t \cdot v \;\; \forall O \in O(n)\}$. That O is an orthogonal transformation guarantees that the O(n) orbits of (u,v), [u,v] remain of the form (OM,$B_{Inj(M)}(0)$).

- $E \cong BM|_{<Inj(M)}$ via

$$[u, v] = [(p, \underline{e}), v] \cong X_{<Inj(M)} = v^i e_i \cong v^i r_i \ \forall \ p \in M$$

meaning that the vector field $X_{<Inj(M)}|_p$ is formed using the components of v^i of $v = v^i r_i \in BM|_{Inj(M)}(0)$ (where $\{r_i\} = \{(0 \ldots \underbrace{1}_{i^{th} spot}, \ldots 0)\}$ is the standard orthonormal basis in $\mathbb{R}^n$) summed out with arbitrary orthonormal frame vectors $e_i|_p$ (which are in 1-1 correspondence with r_i). As in Section D.1 the map $[u, v] \to v^i e_i$ is well-defined. It now follows that for each $p \in M$

$$g(X_{<Inj(M)}, X_{<Inj(M)}) = \delta_{ij} v^i v^j = v^i v_i \implies |X_{<Inj(M)}| < Inj(M)$$

- $\tau([u, v]) = \pi(u) = p$ and $\tau^{-1}(p) = \{[u, v] | \tau([u, v] = p\}$

- $\phi_u : B_{Inj(M)}(0) \to \tau^{-1}(p)$ defined by $\phi_u(v) = [u, v]$ with $\phi_u^{-1} : \tau^{-1}(p) \to B_{Inj(M)}(0)$ well-defined by $\phi_u^{-1}([u, v]) = v$

The previous points explain the diagram and proof that $E \cong BM|_{<Inj(M)}$ is a fibre bundle with fibers $\tau^{-1}(p)$ isomorphic to the manifold $B_{Inj(M)}(0)$. $\square$

Appendix E

On Constraints

E.1 Canonical Bases of Distribution, Δ and Co-Distribution, $\Delta^\star$

Section Summary E.1.1 *The canonical bases of both Δ and $\Delta^\star$, derived in the following section, can be summarized by the matrix equations:*

$$\begin{pmatrix} \overset{\vartriangle}{R}_A & \overset{\vartriangle}{F}_i \end{pmatrix} = \begin{pmatrix} \partial_k & \partial_B \end{pmatrix} \begin{pmatrix} -e_i^k \omega_A^i & e_i^k \\ \delta_A^B + e_i^k \omega_A^i A_k^B & -A_k^B e_i^k \end{pmatrix} \tag{E.1}$$

and

$$\begin{pmatrix} \overset{\vartriangle}{\omega}{}^A \\ \overset{\vartriangle}{\omega}{}^i \end{pmatrix} = \begin{pmatrix} A_k^A & \delta_B^A \\ \tilde{e}_k^i + \omega_B^i A_k^B & \omega_B^i \end{pmatrix} \begin{pmatrix} dr^k \\ ds^B \end{pmatrix} \tag{E.2}$$

where we have freedom to choose

$$e_i^k \text{ (invertible with inverse } \tilde{e}_k^i) \text{ and } \omega_A^i. \tag{E.3}$$

In this section we show how to decompose an arbitrary TQ-valued one form $\omega = \omega^{\mathfrak{I}} \otimes e_{\mathfrak{I}}$ (called an *Ehressman connection* on TQ) into ω^v and ω^h, projections onto ver(TQ) and hor(TQ) respecitvely.

Let ω be a TQ-valued 1-form on $Q \to Q/G$. Let Δ be a m-dimensional distribution spanned by $\overset{\triangle}{F}_i$ and ${}^\star\Delta$ be the (q-m)-dimensional co-distribution spanned by ω^A. We fill out the frames and co-frames on Q using the riggings $\overset{\triangle}{R}_A$ and ω^i. That is, we split ω into

$$\begin{aligned}\omega = \omega^{\mathcal{J}} \otimes e_{\mathcal{J}} &= \overbrace{(\omega^i_k dr^k + \omega^i_B ds^B)}^{=\omega^i \text{ co-frame rigging}} \otimes \overset{\triangle}{F}_i + \overbrace{(\omega^A_k dr^k + \omega^A_B ds^A)}^{=\omega^A \text{ basis of } {}^\star\Delta} \otimes \overset{\triangle}{R}_A \\ &:= \omega^h + \omega^v\end{aligned}$$

and

$$e_{\mathcal{J}} = \underbrace{e^n_i \partial_n + e^B_i \partial_B}_{=\overset{\triangle}{F}_i \text{ span } \Delta} + \underbrace{e^m_A \partial_m + e^C_A \partial_C}_{=\overset{\triangle}{R}_A \text{ frame rigging}}$$

A theorem (to be found in Spivak) showed that the canonical basis for the co-distribution could be written as

$$\omega^A = ds^A + A^A_i(r^j, s^B) dr^i. \tag{E.4}$$

Requiring that ω^v be a projection into the space spanned by the $\overset{\triangle}{R}_A$ and that ω^v annihilate the space spanned by the $\overset{\triangle}{F}_i$ implies that

$$\begin{aligned}\omega^v(\overset{\triangle}{R}_B) &= (\omega^A \otimes \overset{\triangle}{R}_A)(\overset{\triangle}{R}_B) = \delta^A_B \overset{\triangle}{R}_A = \overset{\triangle}{R}_B \\ \omega^v(\overset{\triangle}{F}_i) &= (\omega^A \otimes \overset{\triangle}{R}_A)(\overset{\triangle}{F}_i) = 0.\end{aligned}$$

Solving the above for $\overset{\triangle}{F}_i, \overset{\triangle}{R}_A$, the basis for the distribution is found to be

$$\overset{\triangle}{F}_i = (\partial_k - A^A_k(r^j, s^B)\partial_A) e^k_i \tag{E.5}$$

while the rest of the frame is

$$\overset{\triangle}{R}_A = \partial_A + e^k_A(\partial_k - A^B_k(r^j, s^B)\partial_B). \tag{E.6}$$

Note E.1.1 The matrices e_i^k, e_A^k represent some freedom in the choice of the frame. Set $\overset{\vartriangle}{F}_i = t_k e_i^k$ where $t_k := \partial_k - A_k^A \partial_A$ are k linearly independent vectors. Since $\{\overset{\vartriangle}{F}_i\}$ are also k linearly independent vectors it must be the case that e_i^k is invertible. Denote the inverse of e_i^k as $\tilde{e}_i^k$

Requiring that ω^h be a projection into the space spanned by the $\overset{\vartriangle}{F}_i$ and that ω^h annihilate the space spanned by the $\overset{\vartriangle}{R}_A$ implies that

$$\begin{aligned} \omega^h(\overset{\vartriangle}{F}_i) &= (\omega^j \otimes \overset{\vartriangle}{F}_j)(\overset{\vartriangle}{F}_i) = \delta_i^j \overset{\vartriangle}{F}_j = \overset{\vartriangle}{F}_i \\ \omega^h(\overset{\vartriangle}{R}_A) &= (\omega^j \otimes \overset{\vartriangle}{F}_j)(\overset{\vartriangle}{R}_A) = 0 \end{aligned}$$

for ω^i. Solving the above give

$$\omega^i = \tilde{e}_k^i dr^k + \omega_B^i (ds^B + A_k^B dr^k). \tag{E.7}$$

Note E.1.2 Again there is some freedom in the choice of due to the matrix ω_B^i. The apparent freedom to choose e_A^k was ficticious as it can be shown that it is defined by the relation $e_A^k = -e_i^k \omega_A^i$. So the only freedom is in the choice of e_i^k and ω_A^i.

Remark E.1.1 (On Terminology) Using the above constructed canonical basis of TQ $\{\overset{\vartriangle}{F}_i, \overset{\vartriangle}{R}_A\}$ we have

$$\begin{aligned} TQ &= span\{\overset{\vartriangle}{R}_A\} \oplus span\{\overset{\vartriangle}{F}_i\} \\ &:= ver(TQ) \oplus hor(TQ) \end{aligned}$$

where ver(TQ) and hor(TQ) are the vertical and horizontal space of the tangent space, respectively.

By constuction, given an arbitrary vector $X \in TQ$ we have

$$\begin{aligned} \omega^v(X) &= \omega^v(X^A \overset{\Delta}{R}_A + X^i \overset{\Delta}{F}_i) \\ &= (\omega^B \otimes \overset{\Delta}{R}_B)(X^A \overset{\Delta}{R}_A + X^i \overset{\Delta}{F}_i) \\ &= X^B \overset{\Delta}{R}_B \in ver(TQ) \end{aligned}$$

or more intrinsically

$$\begin{aligned} \omega^v(X) &= \omega^v(X^v + X^h) = X^v \\ \omega^h(X) &= \omega^h(X^v + X^h) = X^h. \end{aligned}$$

We can now represent the identity tensor field on TQ (decomposed wrt to the disbribtution, Δ into horizontal (the distribution directions) and vertical (the rigging directions) pieces)

$$^{\Delta}Id = \omega^v + \omega^h = Id^v + Id^h. \tag{E.8}$$

The above basis of horizontal vectors, $\{\overset{\Delta}{F}_i\}$ (wrt to an *Ehressman connection,* ω) also called the distribution directions or the basis of the distribution Δ, the rigging vectors, $\overset{\Delta}{R}_A$ (which fill out the basis on TQ), and the basis of the co-distribution $\Delta^\star$ are summarized in eqns. (E.1) and (E.2).

Appendix F

$L_\pi Q$ and $L_\Delta Q$

F.1 $L_\pi Q$

Section Summary F.1.1 *When passing from LQ to $L_\pi Q$ the general $Gl(n)$ matrices $\overset{v}{\Pi}{}^{\mathcal{J}}_{\mathcal{J}}(u)$ and $\overset{v}{V}{}^{\mathcal{J}}_{\mathcal{J}}(u)$ become, relative to the coordinated section $\sigma(q) = (q, \underline{\partial}) = (q, \{\partial_i, \partial_A\})$,*

$$\overset{v}{\Pi}{}^{\mathcal{J}}_{\mathcal{J}}(u) = \begin{pmatrix} \overset{v}{\Pi}{}^{i}_{j}(u) & 0 \\ \overset{v}{\Pi}{}^{B}_{j}(u) & \overset{v}{\Pi}{}^{B}_{A}(u) \end{pmatrix} \quad \textit{and} \quad \overset{v}{V}{}^{\mathcal{J}}_{\mathcal{J}}(u) = \begin{pmatrix} \overset{v}{\mathbb{V}}{}^{i}_{j}(u) & 0 \\ -\overset{v}{\mathbb{V}}{}^{B}_{A}(u)\, \overset{v}{\mathbb{V}}{}^{i}_{j}(u)\, \overset{v}{\Pi}{}^{A}_{i}(u) & \overset{v}{\mathbb{V}}{}^{C}_{A}(u) \end{pmatrix}$$

where $\overset{v}{\Pi}{}^{i}_{j}(u)$ and $\overset{v}{\Pi}{}^{B}_{A}(u)$ are invertible with inverses $\overset{v}{\mathbb{V}}{}^{i}_{j}(u)$ and $\overset{v}{\mathbb{V}}{}^{B}_{A}(u)$, respectively and $\overset{v}{\Pi}{}^{B}_{j}(u)$ are freely specifiable. Notice that $\overset{v}{\Pi}{}^{\mathcal{J}}_{\mathcal{J}}, \overset{v}{V}{}^{\mathcal{J}}_{\mathcal{J}} \colon L_\pi Q \to G \subset Gl(n)$ where G is a subgroup of $Gl(n)$. The basis relations as matrix equations are

$$\begin{pmatrix} \overset{v}{R}{}^{i} \\ \overset{v}{E}{}^{B} \end{pmatrix} = \begin{pmatrix} \overset{v}{\Pi}{}^{i}_{j}(u) & 0 \\ \overset{v}{\Pi}{}^{B}_{j}(u) & \overset{v}{\Pi}{}^{B}_{A}(u) \end{pmatrix} \begin{pmatrix} dx^j \\ ds^A \end{pmatrix} \tag{F.1}$$

and

$$\begin{pmatrix} \overset{v}{R}_{j} & \overset{v}{E}_{A} \end{pmatrix} = \begin{pmatrix} \partial_i & \partial_B \end{pmatrix} \begin{pmatrix} \mathbb{V}^i_j(u) & 0 \\ -\mathbb{V}^B_A(u)\mathbb{V}^i_j(u)\, \overset{v}{\Pi}{}^{A}_{i}(u) & \mathbb{V}^B_A(u) \end{pmatrix} \tag{F.2}$$

where $\overset{v}{\Pi}{}^{A}_{\mathcal{J}}(u)$ *are freely specifiable.*

Let Q be a configuration manifold with coordinates $\{q^{\mathfrak{I}}\}$. Let $LQ \xrightarrow{\pi_{LQ}} Q$ be the frame bundle of Q with coordinates $\{(q^{\mathfrak{I}}, \Pi^{\mathfrak{I}}_{\mathcal{J}})\}$. Let σ_{LQ} be an arbitrary local section of π_{LQ}, say $\sigma_{LQ}(q) = (q, \underline{f} = (f_{\mathfrak{I}}))$, then $\Pi^{\mathfrak{I}}_{\mathcal{J}} : LQ \to Gl(n)$ can be defined by

$$\Pi^{\mathfrak{I}}_{\mathcal{J}}(u) = \Pi^{\mathfrak{I}}_{\mathcal{J}}(q, \underline{e}) = e^{\mathfrak{I}}(f_{\mathcal{J}}|_q)$$

where $\bar{e} = \{e^{\mathfrak{I}}\}$ is the dual frame to $\underline{e} = \{e_{\mathcal{J}}\}$. One can view the local section $\sigma(q) = (q, \underline{f})$ as a local reference frame with $\Pi^{\mathfrak{I}}_{\mathcal{J}}(u)$ the Gl(n) matrix relating the local reference co-frame $f^{\mathcal{J}}|_q$ to an arbitrary co-frame $e^{\mathfrak{I}}|_q$. That is,

$$e^{\mathfrak{I}}|_q = \Pi^{\mathfrak{I}}_{\mathcal{J}}(u) f^{\mathcal{J}}|_q$$

Denoting $V^{\mathfrak{I}}_{\mathcal{J}}(u)$ the inverse of $\Pi^{\mathfrak{I}}_{\mathcal{J}}(u)$ then

$$e_{\mathcal{J}}|_q = V^{\mathfrak{I}}_{\mathcal{J}}(u) f_{\mathfrak{I}}|_q$$

is an equation relating the local reference frame $\underline{f}$ to an arbitrary frame $\underline{e}$. The canonical choice of local section σ is the coordinated section $\sigma(q) = (q, \underline{\partial}) = (q, (\partial_{\mathfrak{I}})) = \left(q, \left(\frac{\partial}{\partial q^{\mathfrak{I}}}\right)\right)$ and thus

$$\begin{aligned} e^{\mathfrak{I}}|_q &= \overset{c}{\Pi}{}^{\mathfrak{I}}_{\mathcal{J}}(u) dx^{\mathcal{J}}|_q \\ e_{\mathcal{J}}|_q &= \overset{c}{V}{}^{\mathfrak{I}}_{\mathcal{J}}(u) \partial_{\mathfrak{I}}|_q. \end{aligned}$$

We emphasis that there is nothing unique about this choice of section.

Now we assume that Q has a principle fiber bundle structure $Q \xrightarrow{\pi_Q} Q/\mathfrak{G}$ where the action of $\mathfrak{G}$ on Q is free and proper so that $Q/\mathfrak{G}$ is a smooth manifold. Coordinates on $Q \xrightarrow{\pi_Q} Q/\mathfrak{G}$ can be obtained via a local trivialization of a local section σ_Q of π_Q. Denote these coordinates as $\{q^{\mathfrak{I}}\} = \{r^i, s^A\}$ where $\{r^i\}$ are coordinates on $Q/\mathfrak{G}$ and

$\{s^A\}$ are coordinates on $\mathfrak{G}$. The split in coordinates is manifest at the frame bundle level by

$$\begin{pmatrix} e^i \\ e^A \end{pmatrix} = \begin{pmatrix} \Pi^i_j(u) & \Pi^i_A(u) \\ \Pi^B_j(u) & \Pi^B_A(u) \end{pmatrix} \begin{pmatrix} dx^j \\ dx^A \end{pmatrix}$$

and

$$\begin{pmatrix} e_j & e_A \end{pmatrix} = \begin{pmatrix} \partial_i & \partial_B \end{pmatrix} \begin{pmatrix} V^i_j(u) & V^i_A(u) \\ V^B_j(u) & V^B_A(u) \end{pmatrix}$$

where we do not have any additional information about the sub-matrices other than they build an invertible matrices $\Pi^{\jmath}_{\jmath}(u)$ and $V^{\jmath}_{\jmath}(u)$.

Now we place restrictions on the frames $\underline{e} = \{e_{\jmath}\}$. Define

$$L_\pi Q = \{(p, \underline{e}) = (p, \{\underline{\overset{v}{R}}, \underline{\overset{v}{E}}\})\} \tag{F.3}$$

where $\underline{\overset{v}{E}} = \{\overset{v}{E}_A\}$ are independent vertical vectors and $\underline{\overset{v}{R}} = \{\overset{v}{R}_i\}$ is a rigging (i.e. fills out the basis). It follows that $\overset{v}{E}_A$ can be linear combinations of the coordinated basis of vertical vectors, i.e.

$$\overset{v}{E}_A\,|_q = \overset{v}{V}{}^B_A(u)\partial_B|_q$$

where $\overset{v}{V}{}^B_A(u)$ is invertible, denote this inverse $\mathbb{P}^B_A(u)$. Let $(p, \{\bar{R}, \bar{E}\})$ be the co-frame of $(p, \{\underline{R}, \underline{E}\})$. The basis condition $\overset{v}{R}{}^j\,(\overset{v}{E}_A) = 0$ holds if and only if

$$\overset{v}{\Pi}{}^j_A(u) = 0 \tag{F.4}$$

and thus

$$\overset{v}{R}{}^j\,|_q = \overset{v}{\Pi}{}^j_k(u)dr^k|_q$$

Again, both $\overset{v}{R}{}^j$ and dr^k are bases at q and so $\overset{v}{\Pi}{}^j_k(u)$ is invertible, denote the inverse $\overset{v}{\mathbb{V}}{}^j_k(u)$. It remains to determine the coefficients of

$$\begin{aligned} \overset{v}{R}_i &= \overset{v}{V}{}^i_{\jmath}(u)\partial_{\jmath} = \overset{v}{V}{}^i_j(u)\partial_j + \overset{v}{V}{}^i_A(u)\partial_A \\ \overset{v}{E}{}^B &= \overset{v}{\Pi}{}^B_{\jmath}(u)dq^{\jmath} = \overset{v}{\Pi}{}^B_i(u)dr^i + \overset{v}{\Pi}{}^B_A(u)ds^A \end{aligned}$$

such that $\overset{v}{R}{}^{j}(\overset{v}{R}_i) = \delta_i^j$, $\overset{v}{E}{}^{B}(\overset{v}{R}_i) = 0$ and $\overset{v}{E}{}^{B}(\overset{v}{E}_A) = \delta_A^B$.

First some relations following from $\overset{v}{\Pi}{}^{\mathcal{J}}_{\mathcal{I}}(u)\,\overset{v}{V}{}^{\mathcal{I}}_{\mathcal{K}}(u) = \delta^{\mathcal{J}}_{\mathcal{K}}$:

$$
\begin{array}{ll}
\overset{v}{\Pi}{}^{A}_{\mathcal{I}}(u)\,\overset{v}{V}{}^{\mathcal{I}}_{i}(u) = 0 & \Longrightarrow \overset{v}{V}{}^{C}_{i}(u) = -\,\overset{v}{\mathbb{V}}{}^{C}_{A}(u)\,\overset{v}{V}{}^{j}_{i}(u)\,\overset{v}{\Pi}{}^{A}_{j}(u) \\
\overset{v}{\Pi}{}^{i}_{\mathcal{I}}(u)\,\overset{v}{V}{}^{\mathcal{I}}_{A}(u) = 0 & \Longrightarrow \overset{v}{V}{}^{k}_{A}(u) = 0 \;\; (using \;\; \overset{v}{\Pi}{}^{i}_{B}(u) = 0) \\
\overset{v}{\Pi}{}^{A}_{\mathcal{I}}(u)\,\overset{v}{V}{}^{\mathcal{I}}_{B}(u) = \delta^{A}_{B} & \Longrightarrow \overset{v}{V}{}^{D}_{B}(u) = \overset{v}{\mathbb{V}}{}^{D}_{B}(u) \;\; (using \;\; \overset{v}{V}{}^{i}_{B}(u) = 0) \\
\overset{v}{\Pi}{}^{i}_{\mathcal{I}}(u)\,\overset{v}{V}{}^{\mathcal{I}}_{j}(u) = \delta^{i}_{j} & \Longrightarrow \overset{v}{V}{}^{k}_{j}(u) = \overset{v}{\mathbb{V}}{}^{k}_{j}(u) \;\; (using \;\; \overset{v}{\Pi}{}^{i}_{B}(u) = 0)
\end{array}
$$

We now know that

$$\overset{v}{R}_i = \overset{v}{\mathbb{V}}{}^{k}_{i}(u)\partial_k - (\overset{v}{\mathbb{V}}{}^{D}_{A}\overset{v}{\mathbb{V}}{}^{j}_{i}(u)\,\overset{v}{\Pi}{}^{A}_{j}(u))\partial_D$$

To find the coefficients of E^B one need analyze the equations $\overset{v}{E}{}^{B}(\overset{v}{R}_i) = 0$ and $\overset{v}{E}{}^{B}(\overset{v}{E}_A) = \delta^B_A$. The second equations yields

$$\overset{v}{\Pi}{}^{B}_{D}(u)\,\overset{v}{\mathbb{V}}{}^{D}_{A}(u) = \delta^{B}_{A} \Longrightarrow \overset{v}{\Pi}{}^{B}_{D}(u) = \mathbb{P}^{B}_{D}(u)$$

while the first equations gives

$$\overset{v}{\Pi}{}^{B}_{j}(u)(\overset{v}{\mathbb{V}}{}^{j}_{i}(u) - \overset{v}{\mathbb{V}}{}^{j}_{i}(u)) = 0 \Longrightarrow \overset{v}{\Pi}{}^{B}_{j}(u) \text{ freely specifiable}$$

Equations (F.1) and (F.2) summarize the above formulas.

F.2 $L_\Delta Q$

We again place restrictions on the frames $\underline{e} = (e_{\mathcal{J}})$. Define

$$L_\Delta Q = \{(p, \overset{\Delta}{\underline{e}}) = (p, (\overset{\Delta}{\underline{R}}, \overset{\Delta}{\underline{E}}))\} \tag{F.5}$$

where $\overset{\Delta}{\underline{E}} = \{\overset{\Delta}{\underline{E}}_i\}$ are a collection of independent vectors in Δ and $\overset{\Delta}{\underline{R}} = \{\overset{\Delta}{R}_A\}$ is a rigging (i.e. fills out the basis of TQ). From Appendix E, section E.1 we know that the section of LQ

$$\sigma(q) = (q, \underline{\overset{\vartriangle}{f}}) = (q, \{\overset{\vartriangle}{f}_{\jmath}\}) = (q, (\{\overset{\vartriangle}{R}_A, \overset{\vartriangle}{F}_i\}))$$

with $(\overset{\vartriangle}{F}_i)$ and $(\overset{\vartriangle}{R}_A)$ summarized by the matrix equation (E.1) is actually a section of $L_\Delta Q$. We also saw that the co-distribution basis is $\{\omega^B\}$ with rigging $\{\omega^i\}$ summarized by matrix equation (E.2) satisfied

$$\overset{\vartriangle}{\omega}^A (\overset{\vartriangle}{F}_i) = 0 \quad \overset{\vartriangle}{\omega}^A (\overset{\vartriangle}{R}_B) = \delta^A_B \quad \overset{\vartriangle}{\omega}^i (\overset{\vartriangle}{R}_A) = 0 \quad \overset{\vartriangle}{\omega}^i (\overset{\vartriangle}{F}_j) = \delta^i_j.$$

Using the language of local (reference) frame field from section B.2 (cf. Remark B.2.1) we interpret equations (E.1) and (E.2) the following way: If we set the local (reference) frame field as the coordinated basis $\underline{\partial} = (\partial_k, \partial_B)$ as is the standard in [28], then a new reference field, this time adapted to the distribution, $\underline{\overset{\vartriangle}{f}} = (\overset{\vartriangle}{R}_A, \overset{\vartriangle}{F}_i)$, just another point on the fiber of LQ over p, is related to the coordinated reference frame via the matrix

$$\overset{\vartriangle}{G} = \begin{pmatrix} -e^k_i \omega^i_A & e^k_i \\ \delta^B_A + e^k_i \omega^i_A A^B_k & -A^B_k e^k_i \end{pmatrix} \tag{F.6}$$

while the local co-frame field adapted to the distribution $\overset{\vartriangle}{\omega} := (\overset{\vartriangle}{\omega}^i, \overset{\vartriangle}{\omega}^A)$ is related to the reference co-frame via the matrix

$$\overset{\vartriangle}{G}^{-1} = \begin{pmatrix} A^A_k & \delta^A_B \\ \tilde{e}^i_k + \omega^i_B A^B_k & \omega^i_B \end{pmatrix}. \tag{F.7}$$

where $(\tilde{e}^i_k) = (e^k_i)^{-1}$. Using the freedom to choose $\omega^i_A = 0 \ \forall i, A$ and $e^k_i = \delta^k_i$ eqs. (F.6) and (F.7) reduce to

$$\overset{\vartriangle}{G} = \begin{pmatrix} 0 & \delta^k_i \\ \delta^B_A & -A^B_i \end{pmatrix} \qquad \overset{\vartriangle}{G}^{-1} = \begin{pmatrix} A^A_k & \delta^A_B \\ \delta^i_k & 0 \end{pmatrix}. \tag{F.8}$$

and thus

$$\overset{\triangle}{\underline{F}} = \underline{\partial}\cdot \overset{\triangle}{G} \qquad \bar{\omega} = \overset{\triangle}{G}^{-1} \cdot\bar{\partial} \tag{F.9}$$

where $\bar{\partial} = (dq^{\mathfrak{I}}) = (dr^k, ds^A)$ and $\underline{\partial} = (\partial/\partial q^{\mathfrak{I}}) = (\partial_k, \partial_A)$

Remark F.2.1 *(Terminology and Key Idea) We will call* $\overset{\triangle}{\underline{F}}$ *the local frame field adapted to the distribution and* $\overset{\Delta}{\bar{\omega}}$ *the local co-frame field adapted to the distribution. We shall take distribtution adapted frame field,* $\underline{f} = \overset{\triangle}{\underline{F}}$ *as the local reference frame rather than the coordinated frame field* $\underline{\partial}$ *as is standard in [28]*

Following the methods of section F.1, the matrix group element relating any other frame at q, whose last vectors are in the distribution, to the reference section $\overset{\Delta}{\sigma}(q) = (q, \{\overset{\triangle}{R}_A, \overset{\triangle}{F}_i\})$ will be given as

$$\overset{\triangle}{\Pi}{}^{\mathfrak{I}}_{\mathfrak{J}}(u) = \begin{pmatrix} \overset{\triangle}{\Pi}{}^{i}_{j}(u) & 0 \\ \overset{\triangle}{\Pi}{}^{A}_{j}(u) & \overset{\triangle}{\Pi}{}^{A}_{B}(u) \end{pmatrix}$$

where $\overset{\triangle}{\Pi}{}^{i}_{j}(u)$ and $\overset{\triangle}{\Pi}{}^{A}_{B}(u)$ are invertible and $\overset{\triangle}{\Pi}{}^{A}_{j}(u)$ is an arbitrary matrix.

The utility of taking $\overset{\triangle}{\underline{F}}$ as the reference frame field will be apparent in the n-symplectic dynamics section I.5(a).

Appendix G

n-Symplectic Dynamics Background

G.1 LM Dynamics Background

In this appendix we derive the canonical n-symplectic dynamics of both rank 1 and rank 2 observables. The dynamics for general rank observables first appeared in [28]. For completeness of this document we focus on these two subcases. The rank 2, canonical dynamics (using for example, the kinetic energy of a mechanical system as the rank 2 observable) derived here serve as the *foundation* for our derivation of the constrained rank 2 n–symplectic dynamics in Chapter 4. By foundation, we mean to obtain the constrained n-symplectic dynamics from an appropriate change of coordinates on the canonical dynamics derived here.

The outline for this section is to first give a foundation for n–symplectic (*generalized* Hamiltonian) dynamics by summarizing Hamiltonian dynamics. In Section G.1.2, we consider the rank 1 case in full detail. Our ultimate goal, the rank 2 dynamics, are addressed in the G.1.3 section. An important feature of the n–symplectic dynamics, namely the presence of gauge freedom, is addressed in Section G.1.4.

While still in the preliminary stages, it is our hope that the *gauge* terms might be thought of as control forces and thus useful in control design. An examples along this preliminary line of thought is given in Appendix I. Since these gauge/control forces carry over to the (nonholonomic) constrained dynamics setting, it is a goal to use these terms in control of nonholonomic mechanical systems.

G.1.1 $T^{\star}M$ Dynamics

Given the canonical real-valued one-form, ϑ on $T^{\star}M$ and corresponding canonical non-degenerate real-valued two form $\Omega = d\theta$ on $T^{\star}M$ given in canonical coordinates (q^i, p^j) by

$$\Omega = dp_i \wedge dq^i \quad (\vartheta = p_i dq^i) \tag{G.1}$$

one can use the equation

$$df = -X_f \lrcorner\ \Omega \tag{G.2}$$

to uniquely determine a global Hamiltonian vector field X_f given a globally defined Hamiltonian function $f : T^{\star}M \to M$. The integral curve equations of the global Hamiltonian vector field are exactly Hamilton's equations. The integrability conditions (i.e the necessary condition for existence of solutions) for eq. (G.2) are

$$d(X_f \lrcorner\ \Omega) = 0 \tag{G.3}$$

To give insight into eq. (G.3) we recall that Y is a *local Hamiltonian vector field* if

$$L_Y \Omega = 0 \tag{G.4}$$

Using the formula for the Lie derivative of a two-form, that Ω closed, that Y is locally Hamiltonian and the converse to Poincare's lemma we obtain

$$0 = L_Y \Omega = Y \lrcorner\ d\Omega + d(Y \lrcorner\ \Omega) = d(Y \lrcorner\ \Omega) \tag{G.5}$$

which implies

$$Y \lrcorner\ \Omega = dF$$

where $F : U \subset T^{\star}M \to M$, i.e. F locally defined function called a *local Hamiltonian function*. It is now apparent that the integrability condition (G.3) is equivalent to X_f being a local Hamiltonian vector field. Which makes sense: f a global function defines a global Hamiltonian vector field only if f locally defines a local Hamiltonian vector field.

G.1.2 LM Dynamics for $\mathbb{R}^n$-valued Tensorial Functions

Section Summary G.1.1 *(Rank One Case)*

- *(**Dynamic Equation**) The dynamic equation for $\hat{f}$ a $\mathbb{R}^n$-valued tensorial functions on LM $\sim$ vector field on M (called the rank one structure equation) is*

$$d\hat{f}^i = -X_{\hat{f}} \lrcorner\ \hat{\Omega}^i$$

which is a system of PDE's describing how to determine $X_{\hat{f}}$ given $\hat{f}$ or $\hat{f}$ given $X_{\hat{f}}$. The uniqueness follows from the non-degeneracy of $\hat{\Omega}^i$.

- *(**Canonical Vector Field**) Given the canonical $\hat{f}^i(u) = \Pi^i_j(u) f^j(p)$ where $f = f^j \partial_j$ is a vector field on M and $\hat{f}$ the corresponding vector valued tensorial function on LM one finds*

$$X_{\hat{f}} = f^j(p)\frac{\partial}{\partial x^j} - \frac{\partial f^j(p)}{\partial x^n}\Pi^i_j(u)\frac{\partial}{\partial \Pi^i_n}$$

- *(**More General Vector Field**) Analyzing the integrability conditions $d(-X_{\hat{f}} \lrcorner\ \hat{\Omega}^i) = 0$ one finds that a more general X is available:*

$$X_{\hat{f}} = C^k(x)\frac{\partial}{\partial x^k} - \left(\frac{\partial C^i(x)}{\partial x^b}\Pi^b_j + \frac{\partial \xi^i(x)}{\partial x^j}\right)\frac{\partial}{\partial \Pi^i_j}$$

where $C^k(x)$ and $\xi^i(x)$ are both n arbitrary functions constant in Π. The tensorial vector-valued function on LM corresponding to this more general vector *field on LM is*

$$\hat{f} = \hat{f}^i r_i = (C^k(x)\Pi^i_k + \xi^i(x))r_i$$

- **(Gauge Freedom)** *Though not addressed previously, we now point out that there may be an even more general vector field X and corresponding function $\hat{f}$ due to a* gauge freedom. *This* gauge freedom *can be described as follows: Given $X_{\hat{f}}$ with unique solution $\hat{f}$ to $d\hat{f}^i = -X_{\hat{f}} \rfloor \hat{\Omega}^i$ then is there a vector field, say T, for which the unique solution $\hat{F}$ to $d\hat{F}^i = -(X_{\hat{f}} + T) \rfloor \hat{\Omega}^i$ is equal to $\hat{f}$? If such T exist then solutions to $d\hat{f}^i = -X_{\hat{f}} \rfloor \hat{\Omega}^i$ are* equivalence classes $[X_{\hat{f}}] = \{X_{\hat{f}} + T\}$. *To find such T notice that*

  $$d\hat{F}^i = -(X_{\hat{f}} + T) \rfloor \hat{\Omega}^i = -X_{\hat{f}} \rfloor \hat{\Omega}^i - T \rfloor \hat{\Omega}^i$$

  which for

  $$-T \rfloor \hat{\Omega}^i = 0 \tag{G.6}$$

  will have solution $\hat{f}$. So the gauge freedom *is characterized by those T which satisfy eq. (G.6) . In the rank one case, the only solutions to eq. (G.6) are T=0 so there is no gauge freedom and solutions to the structure equation are uniquely determined.*

We now give the details to the above summary. The canonical $\mathbb{R}^n$-valued one form, $\hat{\theta}$ on LM and non-degenerate $\mathbb{R}^n$-valued two form, $\hat{\Omega}$ on LM are defined in canonical coordinates $(q^i, {}^c\Pi^i_j = \Pi^i_j)$ by

$$\hat{\Omega}^i = d\Pi^i_j \wedge dx^j \quad (\hat{\theta}^i = \Pi^i_j dx^j) \tag{G.7}$$

There is a nature correspondence between $\mathbb{R}^n$-valued tensorial functions, $\hat{f} = \hat{f}^i r_i$ on LM and vector fields, $f = f^i \partial_i$ on M given by

$$\hat{f}^i(u) = \Pi^i_j(u) f^j(p) \tag{G.8}$$

One can ask for the LM equivalent of eq. (G.2). One would guess that it is

$$d\hat{f}^i = -X_{\hat{f}} \lrcorner\ \hat{\Omega}^i \tag{G.9}$$

where $\hat{f}$ is some (global)tensorial function on LM, not necessarily of the form of eq.(G.8).

As pointed out in [28] eq. (G.9) has a "geometric derivation": Start with a torsion free connection ω on LM

$$\Theta^i = \hat{\Omega}^i + \omega^i_j \wedge \hat{\theta}^j = 0.$$

Now let H_i be horizontal vector fields on LM satisfying $\hat{\theta}^j(H_i) = \delta^j_i$ and $\omega(H_i) = 0$ then

$$\hat{\Omega}^i(H_k) = (-\omega^i_j \wedge \hat{\theta}^j)(H_k) = \omega^i_j \hat{\theta}^j(H_k)$$

and thus

$$\omega^i_k = H_k \lrcorner\ \hat{\Omega}^i \tag{G.10}$$

Now we use the connection from (G.10) in the covariant derivative formula for an arbitrary tensorial function

$$D\hat{f}^i = d\hat{f}^i + \omega^i_j \cdot \hat{f}^j = d\hat{f}^i + (H_j \lrcorner\ \hat{\Omega}^i)\hat{f}^j = d\hat{f}^i + (\hat{f}^j H_j \lrcorner\ \hat{\Omega}^i).$$

And thus if $\hat{f}$ is covariant constant then

$$d\hat{f}^i = -(\hat{f}^j H_j \lrcorner\ \hat{\Omega}^i) = -(X_{\hat{f}} \lrcorner\ \hat{\Omega}^i) \tag{G.11}$$

where $X_{\hat{f}}(u) := \hat{f}^j(u)H_j|_u$ is the (horizontal) vector field on LM defined by $\hat{f}$.

Remark G.1.1 *The above derivation of eq.(G.9) (where the solution vectors are horizontal) can be disregarded leaving only the*

$$Axiom:\ d\hat{f}^i = -X_{\hat{f}} \lrcorner\ \hat{\Omega}^i$$

are the dynamic equations for an $\mathbb{R}^n$-valued tensorial function on LM. The integrability conditions are

$$d(-X_{\hat{f}} \lrcorner\ \hat{\Omega}^i) = 0. \tag{G.12}$$

(Given the specific $\mathbb{R}^n$-valued tensorial function $\hat{f}^i(u) = f^j(p)\Pi^i_j(u)$ one can solve eq. (G.9) to obtain a locally Hamiltonian vector field

$$X_{\hat{f}} = f^j(p)\partial_j|_u - \frac{\partial f^j(p)}{\partial x^n}\Pi^i_j(u)\frac{\partial}{\partial \Pi^i_n}. \tag{G.13}$$

which has both a horizontal and vertical piece. The solution (G.13) follows from

$$\begin{aligned} d\hat{f}^i &= \frac{\partial}{\partial x^n}\left(f^j_p \Pi^i_j(u)\right) \wedge dx^n + \frac{\partial}{\partial \Pi^a_b}\left(f^j \Pi^i_j(u)\right) \wedge d\Pi^a_b \\ &= \frac{\partial f^j_p}{\partial x^n}\Pi^i_j(u) \wedge dx^n + f^j \wedge d\Pi^i_j \end{aligned}$$

and

$$\begin{aligned} -X_{\hat{f}} \lrcorner\ \hat{\Omega}^i &= (d\Pi^i_j(u) \wedge dx^j|_p)\left(-X^a\partial_a - X^n_m \frac{\partial}{\partial \Pi^n_m}\right) \\ &= -X^i_j dx^j + X^j d\Pi^i_j \end{aligned} \tag{G.14}$$

The question/objective seems to me to be: *Are the solutions (G.13) the most general solutions to the eq. (G.9)? If they are not, what are the extra pieces? How do these extra pieces influence the corresponding tensorial function?*. We must analyze the integrability conditions $(G.12)$. Using eq. (G.14) where $X^i_j = X^i_j(x, \Pi)$ and

$X^j = X^j(x, \Pi)$ we obtain

$$\begin{aligned}
0 &= d(-X_{\hat{f}} \lrcorner\ \hat{\Omega}^i) \\
&= d(-X^i_j \wedge dx^j + X^k \wedge d\Pi^i_k) \\
&= -dX^i_j \wedge dx^j + dX^k \wedge d\Pi^i_k \\
&= -\frac{\partial X^i_j}{\partial \Pi^n_m} d\Pi^n_m \wedge dx^j - \frac{\partial X^i_j}{\partial x^a} dx^a \wedge dx^j \\
&\quad + \frac{\partial X^k}{\partial \Pi^e_f} d\Pi^e_f \wedge d\Pi^i_k + \frac{\partial X^k}{\partial x^c} dx^c \wedge d\Pi^i_k \\
&= \left(\delta^i_n \frac{\partial X^k}{\partial x^j} + \frac{\partial X^i_j}{\partial \Pi^n_k} \right) dx^j \wedge d\Pi^n_k \\
&\quad + \left(-\frac{\partial X^i_j}{\partial x^a} \right) dx^a \wedge dx^j + \left(\delta^i_a \frac{\partial X^j}{\partial \Pi^e_f} \right) d\Pi^e_f \wedge d\Pi^a_j
\end{aligned}$$

whereby skew-symmetrizing we obtain

$$\delta^i_n \frac{\partial X^k}{\partial x^j} + \frac{\partial X^i_j}{\partial \Pi^n_k} = 0 \tag{G.15}$$

$$\frac{\partial X^i_a}{\partial x^j} - \frac{\partial X^i_j}{\partial x^a} = 0 \tag{G.16}$$

$$\delta^i_a \frac{\partial X^k}{\partial \Pi^e_f} - \delta^i_e \frac{\partial X^f}{\partial \Pi^a_k} = 0. \tag{G.17}$$

Setting i=a and i=e in (G.17) one obtains, respectively,

$$\begin{aligned}
n \frac{\partial X^k}{\partial \Pi^e_f} &= \frac{\partial X^f}{\partial \Pi^e_k} \\
n \frac{\partial X^f}{\partial \Pi^a_k} &= -\frac{\partial X^k}{\partial \Pi^a_f} \stackrel{a \to e}{\Longrightarrow} n \frac{\partial X^f}{\partial \Pi^e_k} = -\frac{\partial X^k}{\partial \Pi^e_f}
\end{aligned}$$

which collectively give

$$\left(n - \frac{1}{n} \right) \frac{\partial X^k}{\partial \Pi^e_f} = 0 \stackrel{n \neq 1}{\Longrightarrow} \frac{\partial X^k}{\partial \Pi^e_f} = 0 \implies X^k = C^k(x) \tag{G.18}$$

where $C^k(x)$ are n functions constant in Π. We use this constancy to find that X^i_j are *linear* in Π. Specifically, applying $\partial/\partial\Pi^a_b$ to (G.15), using the symmetry of mixed partials and that X^k is constant in Π we obtain

$$\frac{\partial^2 X^i_j}{\partial\Pi^a_b \partial\Pi^n_k} = 0 \implies X^i_j = A^{i\alpha}_{j\beta}(x)\Pi^\beta_\alpha + B^i_j(x) \tag{G.19}$$

where A^i_j and B^i_j are n^2 functions constant in Π. Substituting (G.19) and (G.18) into the integrability condition (G.15) we obtain

$$\delta^i_n \frac{\partial C^k(x)}{\partial x^j} + A^{ik}_{jn}(x) = 0$$

which, upon fixing i=n and absorbing the constant $-n$ under the derivative yields

$$\frac{\partial C^k(x)}{\partial x^j} = A^{ik}_{ji}(x) \tag{G.20}$$

Eq. (G.20) says that, at this moment, we know only a "trace" of the full components of A. Substituting (G.19) into the last integrability condition (G.16) we obtain

$$(A^{i\alpha}_{a\beta,j} - A^{i\alpha}_{j\beta,a})\Pi^\beta_\alpha + (B^i_{a,j} - B^i_{j,a}) = 0$$

Since it is not possible for the terms with coordinates Π to cancel with terms with no Π coordinates (recall that A and B are functions of only the coordinates x) we must have that

$$\frac{\partial B^i_a(x)}{\partial x^j} - \frac{\partial B^i_j(x)}{\partial x^a} = 0 \tag{G.21}$$

$$\left(A^{i\alpha}_{a\beta,j} - A^{i\alpha}_{j\beta,a}\right)\Pi^\beta_\alpha = 0 \tag{G.22}$$

Since eq. (G.22) is being summed over the index β we split it into two pieces; when $\beta = i$ and when $\beta \neq i$ (denoted by $\bar{i}$) and then use eq. (G.20) to obtain

$$\begin{aligned}
\left(A^{i\alpha}_{ai,j} - A^{i\alpha}_{ji,a}\right)\Pi^{i}_{\alpha} + \left(A^{i\alpha}_{a\bar{i},j} - A^{i\alpha}_{j\bar{i},a}\right)\Pi^{\bar{i}}_{\alpha} &= 0 \\
\left(C^{\alpha}_{,aj} - C^{\alpha}_{,ja}\right)\Pi^{i}_{\alpha} + \left(A^{i\alpha}_{a\bar{i},j} - A^{i\alpha}_{j\bar{i},a}\right)\Pi^{\bar{i}}_{\alpha} &= 0 \\
0 + \left(A^{i\alpha}_{a\bar{i},j} - A^{i\alpha}_{j\bar{i},a}\right)\Pi^{\bar{i}}_{\alpha} &= 0
\end{aligned}$$

This last system of PDE's has at least the trivial solution $A^{i\alpha}_{j\bar{i}} = 0 \; \forall i, j, \bar{i}, \alpha$ which indicates that we need only concern ourselves with the non-zero, trace terms of eq. (G.20). To understand the implications of eq. (G.21), we view the n^2 functions $B^i_j(x)$ as the coefficients of a (1,1) tensor field $B = B^i_j(x) dx^j \otimes \partial_i$ which we can view as n vector–valued one forms $B^i = B^i_j(x) dx^j$. Taking the exterior derivative of the B^i yield

$$dB^i = \frac{\partial B^i_j(x)}{\partial x^a} dx^a \wedge dx^j = \frac{1}{2}\left(\frac{\partial B^i_j(x)}{\partial x^a} - \frac{\partial B^i_a(x)}{\partial x^j}\right) dx^i \wedge dx^a$$

and thus if $\frac{\partial B^i_j(x)}{\partial x^a} - \frac{\partial B^i_a(x)}{\partial x^j} = 0$ then $dB^i = 0$ which by the converse of Poincare's lemma gives

$$B^i = -d\xi^i = -\frac{\partial \xi^i(x)}{\partial x^j} dx^j \quad \text{(the minus for convenience)}$$

where ξ^i are locally defined functions. Thus, at least locally,

$$B^i_j(x) = -\frac{\partial \xi^i(x)}{\partial x^j}.$$

We can now conclude that the most general vector field X which satisfies $d(X \rfloor\ \Omega^i) = 0$ can be written in local coordinates as

$$
\begin{aligned}
X &= X^k \frac{\partial}{\partial x^k} + X^i_j \frac{\partial}{\partial \Pi^i_j} \\
&= C^k(x) \frac{\partial}{\partial x^k} - \left(\frac{\partial C^i(x)}{\partial x^b} \Pi^b_j + \frac{\partial \xi^i(x)}{\partial x^j} \right) \frac{\partial}{\partial \Pi^i_j}
\end{aligned}
\tag{G.23}
$$

Substituting eq. (G.23) into the structure equation $d\hat{f}^i = -X \rfloor\ \Omega^i$ one obtains a system of PDE's for the most general $\mathbb{R}^n$-valued function corresponding to X:

$$
\begin{aligned}
\frac{\partial \hat{f}^i}{\partial x^k} &= -X^i_k = \frac{\partial C^i(x)}{\partial x^b} \Pi^b_k + \frac{\partial \xi^i(x)}{\partial x^k} \\
\frac{\partial \hat{f}^i}{\partial \Pi^i_k} &= X^k = C^k(x)
\end{aligned}
$$

which have solution

$$
\hat{f}^i = C^k(x) \Pi^i_k + \xi^i(x) \tag{G.24}
$$

G.1.3 LM Dynamics for $(\mathbb{R}^n \otimes \mathbb{R}^n)$-valued Tensorial Functions

---

Section Summary G.1.2 *(Rank 2 Case)*

- *(***Dynamic Equation***)The dynamic equation for $\hat{g}$ an $\mathbb{R}^n \otimes \mathbb{R}^n$-valued tensorial function on LM $\sim$ a (0,2) tensor field on M (called say the second rank structure equation) is*

$$
d\hat{g}^{ij} = -2X^{(i}_{\hat{g}} \rfloor\ \hat{\Omega}^{j)}.
$$

  As we will see in the fourth point below, given $\hat{g}$ the corresponding $X_{\hat{g}}$ is not uniquely defined but rather is an equivalence class of vector fields.

- **(Canonical Vector Field)** *Given the canonical $\hat{g}^{ij}(u) = g^{ab}(p)\Pi^i_a(u)\Pi^j_b(u)$ where $g = g_{ab}dx^a dx^b$ is the (0,2)-tensor field on M and $\hat{g}$ the corresponding* *tensorial function one finds using the second rank structure equation that*

$$X^i_{\hat{g}} = g_p^{ab}\Pi^i_b(u)\frac{\partial}{\partial x^a} - \frac{1}{2}\frac{\partial g_p^{ab}}{\partial x^n}\Pi^i_a(u)\Pi^j_b(u)\frac{\partial}{\partial \Pi^j_n}.$$

- **(More General Vector Field)** *Analyzing the integrability conditions* $d(X^{(i}_{\hat{g}} \lrcorner\ \hat{\Omega}^{j)}) = 0$ *one finds that a more general* X *is available:*

$$\begin{aligned} X^i_{\hat{g}} &= (A^{ab}(x)\Pi^i_b + B^{ia}(x))\frac{\partial}{\partial x^a} \\ -\frac{1}{2}&\left(\frac{\partial A^{ab}(x)}{\partial x^n}\Pi^i_a(u)\Pi^j_b(u) + \frac{\partial B^{ja}(x)}{\partial x^n}\Pi^i_a(u) + \frac{\partial C^{ij}(x)}{\partial x^n}\right)\frac{\partial}{\partial \Pi^j_n} \end{aligned}$$

- **(Gauge Freedom)** *Unlike the rank one case, the rank two case exhibits some* gauge freedom *characterized by the equation*

$$T^{(\alpha} \lrcorner\ \hat{\Omega}^{i)} = 0 \implies T^\alpha = T_b^{\alpha a}\frac{\partial}{\partial \Pi^a_b} \quad where \ \ T_b^{(\alpha a)} = 0$$

and thus the equivalence class of vector fields is given by

$$\begin{aligned} [X^i_{\hat{g}}] &= (A^{ab}(x)\Pi^i_b + B^{ia}(x))\frac{\partial}{\partial x^a} \\ -\frac{1}{2}&\left(\frac{\partial A^{ab}(x)}{\partial x^n}\Pi^i_a(u)\Pi^j_b(u) + \frac{\partial B^{ja}(x)}{\partial x^n}\Pi^i_a(u) + \frac{\partial C^{ij}(x)}{\partial x^n} + T^{ij}_n\right)\frac{\partial}{\partial \Pi^j_n} \end{aligned}$$

We now give the details to the above summary. The analog of eq. (G.9) for $\hat{g} = \hat{g}^{ij} r_i \otimes r_j$ follows from assuming $\hat{g}$ is covariant constant

$$0 = D\hat{g}^{ij} = d\hat{g}^{ij} + \omega^i_a\hat{g}^{aj} + \omega^j_b\hat{g}^{ib}$$

and using using again that the torsion of the connection is 0 to obtain

$$\omega^i_a = H_a \lrcorner\ \hat{\Omega}^i.$$

Both of these equation then lead to

$$\begin{aligned} d\hat{g}^{ij} &= -(\hat{g}^{aj} H_a \lrcorner\ \hat{\Omega}^i) - (\hat{g}^{ib} H_b \lrcorner\ \hat{\Omega}^j) \\ &:= -(X^j_{\hat{g}} \lrcorner\ \hat{\Omega}^i) - (X^i_{\hat{g}} \lrcorner\ \hat{\Omega}^j) \\ &= -2X^{(i}_{\hat{g}} \lrcorner\ \hat{\Omega}^{j)}. \end{aligned}$$

Remark G.1.2 *As in Remark G.1.1, independent of the derivation, we can take as*

$$Axiom : d\hat{g}^{ij} = -2X^{(i}_{\hat{g}} \lrcorner\ \hat{\Omega}^{j)} \tag{G.25}$$

are the dynamic equations for an $(\mathbb{R}^n \otimes \mathbb{R}^n)$-valued tensorial function on LM. The integrability conditions of eq. (G.25) are

$$d(X^{(i}_{\hat{g}} \lrcorner\ \hat{\Omega}^{j)}) = 0 \tag{G.26}$$

We now parallel section G.1.2.

Given the canonical $\mathbb{R}^n \otimes \mathbb{R}^n$-valued tensorial function on LM given by $\hat{g}^{ij}(u) = g^{ab}(p)\Pi^i_a(u)\Pi^j_b(u)$ we ask for the corresponding solution vector fields $X_{\hat{g}}$. The solution will follow from

$$\begin{aligned} d\hat{g}^{ij} &= \frac{\partial}{\partial x^n}\left(g^{ab}_p \Pi^i_a(u)\Pi^j_b(u)\right) \wedge dx^n + \frac{\partial}{\partial \Pi^c_d}\left(g^{ab}\Pi^i_a(u)\Pi^j_b(u)\right) \wedge d\Pi^c_d \\ &= \left(\frac{\partial g^{ab}_p}{\partial x^n}\Pi^i_a(u)\Pi^j_b(u)\right) \wedge dx^n + 2g^{ba}_p \Pi^{(i}_b(u) \wedge d\Pi^{j)}_a \end{aligned} \tag{G.27}$$

and

$$\begin{aligned} -X^i_{\hat{g}} \lrcorner\ \hat{\Omega}^j &= (d\Pi^j_n \wedge dx^n)\left(-X^{ik}\partial_k - X^{ia}_b \frac{\partial}{\partial \Pi^a_b}\right) \\ &= -X^{ij}_n dx^n + X^{in} d\Pi^j_n \\ &= -X^{ij}_n dx^n + X^{ia} d\Pi^j_a. \end{aligned}$$

since $d\Pi^j_a = d(\delta^j_i \Pi^i_a) = \delta^j_i d\Pi^i_a$ as δ^i_j constant. Similarly

$$\begin{aligned} -X^j_{\hat{g}} \lrcorner \ \hat{\Omega}^i &= (d\Pi^i_n \wedge dx^n)\left(-X^{jk}\partial_k - X^{ja}_b \frac{\partial}{\partial \Pi^a_b}\right) \\ &= -X^{ji}_n dx^n + X^{jn} d\Pi^i_n \\ &= -X^{ji}_n dx^n + X^{ja} d\Pi^i_a \end{aligned}$$

and hence

$$-X^{(i}_{\hat{g}} \lrcorner \ \hat{\Omega}^{j)} = -2X^{(ij)}_n \wedge dx^n + 2X^{(ia} \wedge d\Pi^{j)}_a. \tag{G.28}$$

Equating eq. (G.27) and (G.28) yields

$$\begin{aligned} -2X^{(ij)}_n &= \frac{\partial g^{ab}_p}{\partial x^n} \Pi^i_a(u) \Pi^j_b(u) \\ X^{ia} &= g^{ba}(p) \Pi^i_b(u). \end{aligned}$$

Since the right hand side of the first equation is symmetric in i and j then $X^{(ij)} = X^{ij}$ and hence

$$X^i_{\hat{g}} = g^{ab}_p \Pi^i_b(u) \frac{\partial}{\partial x^a} - \frac{1}{2}\frac{\partial g^{ab}_p}{\partial x^n} \Pi^i_a(u) \Pi^j_b(u) \frac{\partial}{\partial \Pi^j_n}. \tag{G.29}$$

To determine more general solutions one would have to analyze the integrability conditions $d(X^{(i}_{\hat{g}} \lrcorner \ \hat{\Omega}^{j)}) = 0$ using the same techniques as from section G.1.2. The result for arbitrary rank tensorial functions is given in [28]. For the rank two condition one obtains

$$\begin{aligned} X^i_{\hat{g}} &= (A^{ab}(x)\Pi^i_b + B^{ia}(x)) \frac{\partial}{\partial x^a} \\ -\frac{1}{2}&\left(\frac{\partial A^{ab}(x)}{\partial x^n} \Pi^i_a(u) \Pi^j_b(u) + \frac{\partial B^{ja}(x)}{\partial x^n} \Pi^i_a(u) + \frac{\partial C^{ij}(x)}{\partial x^n}\right) \frac{\partial}{\partial \Pi^j_n} \end{aligned}$$

Unlike the rank 1 case there is some *gauge freedom* characterized by finding n vector fields on LM denoted $T^\alpha = T^{\alpha n}\frac{\partial}{\partial x^n} + T^{\alpha a}_b \frac{\partial}{\partial \Pi^a_b}$ which satisfy

$$T^{(\alpha} \lrcorner \ \hat{\Omega}^{i)} = 0$$

Upon working this out one finds that

$$T^{\alpha} = T_b^{\alpha a}\frac{\partial}{\partial \Pi_b^a} \quad where \;\; T_b^{(\alpha a)} = 0 \tag{G.30}$$

and thus the most general solutions to $d\hat{g}^{ij} = -2X_{\hat{g}}^{(i} \lrcorner \; \hat{\Omega}^{j)}$ are the equivalence classes

$$\begin{aligned}
[X_{\hat{g}}^i] \;\; &= \;\; (A^{ab}(x)\Pi_b^i + B^{ia}(x))\frac{\partial}{\partial x^a} \\
-\frac{1}{2}&\left(\frac{\partial A^{ab}(x)}{\partial x^n}\Pi_a^i(u)\Pi_b^j(u) + \frac{\partial B^{ja}(x)}{\partial x^n}\Pi_a^i(u) + \frac{\partial C^{ij}(x)}{\partial x^n} + T_n^{ij}\right)\frac{\partial}{\partial \Pi_n^j}
\end{aligned}$$

G.1.4 Gauge Freedom Breaks Equivalence Classes

Summary G.1.1 *(Rank 2 Gauge Freedom Case)*

- *(***Dynamic Equation***)The rank two structure equation for $\hat{g}$ an $\mathbb{R}^n \otimes \mathbb{R}^n$-valued tensorial function on $LM \sim$ a (0,2) tensor field on M is*

$$d\hat{g}^{ij} = -2X_{\hat{g}}^{(i} \lrcorner \; \hat{\Omega}^{j)}.$$

- **(Canonical Vector Field with Gauge Freedom)** *Given the canonical $\hat{g}^{ij}(u) = g^{ab}(p)\Pi_a^i(u)\Pi_b^j(u)$ where $g = g_{ab}dx^a dx^b$ is the (0,2)-tensor field on M and $\hat{g}$ the corresponding tensorial function one finds the equivalence class of solutions to be*

$$[X_{\hat{g}}^i] = g_p^{ab}\Pi_b^i(u)\frac{\partial}{\partial x^a} - \left(\frac{1}{2}\frac{\partial g_p^{ab}}{\partial x^n}\Pi_a^i(u)\Pi_b^j(u) + T_n^{ij}\right)\frac{\partial}{\partial \Pi_n^j}.$$

- **(Two Choices of Gauge Freedom)**
  - *An obvious choice of gauge freedom is $T_n^{ij} = 0$ has the corresponding unique rank two solution*

$$X_{\hat{g}}^i = g_p^{ab}\Pi_b^i(u)\frac{\partial}{\partial x^a} - \frac{1}{2}\frac{\partial g_p^{ab}}{\partial x^n}\Pi_a^i(u)\Pi_b^j(u)\frac{\partial}{\partial \Pi_n^j}. \tag{G.31}$$

- *Choosing* $T^{ij}_n = \frac{1}{2} g^{kj}\Gamma^i_{nk} - \frac{1}{2} g^{kj}\Gamma^i_{nk}$ *(the* Γ*'s the coefficients for a symmetric connection not necessarily the Levi–Civita connection) gives the corresponding unique rank two solution*

$$X^i_{\hat{g}} = \hat{g}^{i\alpha} H_\alpha + \frac{1}{2}\hat{Q}^{\cdot in}_{\rho\cdot\cdot} E^{\rho\star}_n \tag{G.32}$$

where vertical and horizontal are given respectively by

$$E^{\rho\star}_n = -\Pi^\rho_\beta \frac{\partial}{\partial \Pi^n_\beta} \quad \text{and} \quad H_\alpha = \overset{-1}{\Pi}{}^i_\alpha \left(\frac{\partial}{\partial x^i} + \Pi^l_j \Gamma^j_{ik} \frac{\partial}{\partial \Pi^l_k} \right) \tag{G.33}$$

We now give the details to the above summary. In this section we ask for the form of the gauge freedom

$$T^{ij}_k \frac{\partial}{\partial \Pi^j_k} \quad \text{where} \quad T^{(ij)}_k = 0 \tag{G.34}$$

which transforms the purely kinetic energy solution to the n-symplectic rank two structure equation written in canonical coordinates

$$X^i_{\hat{g}} = g^{ab}\Pi^i_b \frac{\partial}{\partial x^a} - \frac{1}{2}\frac{\partial g^{ab}}{\partial x^n}\Pi^i_a \Pi^j_b \frac{\partial}{\partial \Pi^j_n} + T^{ij}_k \frac{\partial}{\partial \Pi^j_k} \tag{G.35}$$

into

$$X^i_{\hat{g}} = \hat{g}^{i\alpha} H_\alpha + \frac{1}{2}[H_\rho \lrcorner\ D\hat{g}^{in}] E^{\rho\star}_n \tag{G.36}$$

where vertical and horizontal are given respectively by

$$E^{\rho\star}_n = -\Pi^\rho_\beta \frac{\partial}{\partial \Pi^n_\beta} \quad \text{and} \quad H_\alpha = \overset{-1}{\Pi}{}^i_\alpha \left(\frac{\partial}{\partial x^i} + \Pi^l_j \Gamma^j_{ik} \frac{\partial}{\partial \Pi^l_k} \right) \tag{G.37}$$

and $D\hat{g}$ is the covariant derivative given by

$$D\hat{g}^{in} = d\hat{g}^{in} + \omega^i_d \hat{g}^{dn} + \omega^n_c \hat{g}^{ic}. \tag{G.38}$$

This problem was considered and solved in [29] but the exact form of the gauge terms and some details were not expressly written. In this section we supply all the details.

First we deal with the horizontal part by inserting a Kronecker delta and then adding and subtracting an appropriate vertical piece:

$$
\begin{aligned}
g^{ab}\Pi^i_b \frac{\partial}{\partial x^a} &= g^{ab}\Pi^i_b \delta^k_a \frac{\partial}{\partial x^k} \\
&= g^{ab}\Pi^i_b \overset{-1}{\Pi}{}^k_\alpha \Pi^\alpha_a \frac{\partial}{\partial x^k} \\
&= \hat{g}^{i\alpha} \overset{-1}{\Pi}{}^k_\alpha \frac{\partial}{\partial x^k} + 0 \\
&= \hat{g}^{i\alpha} \overset{-1}{\Pi}{}^k_\alpha \frac{\partial}{\partial x^k} + \hat{g}^{i\alpha} \overset{-1}{\Pi}{}^k_\alpha \Pi^\gamma_\xi \Gamma^\xi_{k\beta} \frac{\partial}{\partial \Pi^\gamma_\beta} \\
&\qquad - \hat{g}^{i\alpha} \overset{-1}{\Pi}{}^k_\alpha \Pi^\gamma_\xi \Gamma^\xi_{k\beta} \frac{\partial}{\partial \Pi^\gamma_\beta} \\
&= \hat{g}^{i\alpha} \overset{-1}{\Pi}{}^k_\alpha \left(\frac{\partial}{\partial x^k} + \Pi^\gamma_\xi \Gamma^\xi_{k\beta} \frac{\partial}{\partial \Pi^\gamma_\beta} \right) \\
&\qquad - \hat{g}^{i\alpha} \overset{-1}{\Pi}{}^k_\alpha \Pi^\gamma_\xi \Gamma^\xi_{k\beta} \frac{\partial}{\partial \Pi^\gamma_\beta} \\
&= \hat{g}^{i\alpha} H_\alpha - \hat{g}^{i\alpha} \overset{-1}{\Pi}{}^k_\alpha \Pi^\gamma_\xi \Gamma^\xi_{k\beta} \frac{\partial}{\partial \Pi^\gamma_\beta}
\end{aligned}
$$

So we have the horizontal piece we want plus an extra vertical piece which can be written

$$
-\hat{g}^{i\alpha} \overset{-1}{\Pi}{}^k_\alpha \Pi^\gamma_\xi \Gamma^\xi_{k\beta} \frac{\partial}{\partial \Pi^\gamma_\beta} = -g^{ak}\Pi^i_a \Pi^\gamma_\xi \Gamma^\xi_{k\beta} \frac{\partial}{\partial \Pi^\gamma_\beta} \tag{G.39}
$$

and collected with the other vertical pieces of eq. (G.35) to obtain

$$
-\left(\frac{1}{2} g^{ab}_{,\beta} \Pi^i_a \Pi^\gamma_b + g^{ak}\Pi^i_a \Pi^\gamma_\xi \Gamma^\xi_{k\beta} + T^{i\gamma}_\beta \right) \frac{\partial}{\partial \Pi^\gamma_\beta} \tag{G.40}
$$

The canonical Π directions can be written in terms of the vertical directions according to

$$
-\frac{\partial}{\partial \Pi^\gamma_\beta} = -\delta^n_\gamma \delta^\beta_m \frac{\partial}{\partial \Pi^n_m} = -\delta^n_\gamma (\overset{-1}{\Pi}{}^\beta_\rho \Pi^\rho_m) \frac{\partial}{\partial \Pi^n_m} = \delta^n_\gamma \overset{-1}{\Pi}{}^\beta_\rho E^{\rho\star}_n \tag{G.41}
$$

and thus eq.(G.40) can be written as

$$\overset{-1}{\Pi}{}^{\beta}_{\rho}\left(\left[\frac{1}{2}g^{ab}_{,\beta}+g^{ak}\Gamma^{b}_{k\beta}+T^{ab}_{\beta}\right]\Pi^{i}_{a}\Pi^{n}_{b}\right)E^{\rho\star}_{n} \tag{G.42}$$

The question now is for what choice of T^{ab}_{β} can we make eq. (G.42) take the form

$$\frac{1}{2}(H_{\rho}\lrcorner\ D\hat{g}^{in})E^{\rho\star}_{n}=\frac{1}{2}[H_{\rho}\lrcorner\ (d\hat{g}^{in}+\omega^{i}_{d}\hat{g}^{dn}+\omega^{n}_{c}\hat{g}^{ic})]E^{\rho\star}_{n} \tag{G.43}$$

We attack this by first computing the above and then match terms with eq.(G.42); whatever doesn't match with be what we will choose for the T terms.

First, as we are using a torsion free connection characterized by $\Gamma^{\xi}_{\beta k}=\Gamma^{\xi}_{k\beta}$, then

$$H_{\rho}\lrcorner\ \omega^{n}_{c}\hat{g}^{ic}=\hat{g}^{ci}\,\overset{-1}{\Pi}{}^{k}_{c}\,\overset{-1}{\Pi}{}^{\beta}_{\rho}\,\Pi^{n}_{\xi}\left(\Gamma^{\xi}_{\beta k}-\Gamma^{\xi}_{k\beta}\right)=0 \tag{G.44}$$

and similarly $H_{\rho}\lrcorner\ \omega^{i}_{d}\hat{g}^{dn}=0$. So for a torsion free connection

$$H_{\rho}\lrcorner\ D\hat{g}^{ij}=H_{\rho}\lrcorner\ d\hat{g}^{ij}$$

It can be shown that

$$\begin{aligned}\hat{Q}^{\cdot in}_{\rho\cdot\cdot}:=\frac{1}{2}d_{\rho}\hat{g}^{in} &:= \frac{1}{2}H_{\rho}\lrcorner\ d\hat{g}^{in}\\ &= \frac{1}{2}[g^{ab}_{,\beta}+g^{ka}\Gamma^{b}_{\beta k}+g^{kb}\Gamma^{a}_{\beta k}]\Pi^{i}_{a}\Pi^{n}_{b}\ \overset{-1}{\Pi}{}^{\beta}_{\rho}\\ &= \frac{1}{2}\,\overset{-1}{\Pi}{}^{\beta}_{\rho}\ g^{ab}_{;\beta}\Pi^{i}_{a}\Pi^{n}_{b}\quad (g^{ab}_{;\beta}:=Q^{ab\cdot}_{\cdot\cdot\beta})\\ &:= \frac{1}{2}\widehat{g^{in}_{;\rho}}\end{aligned} \tag{G.45}$$

using

$$d\hat{g}^{in}=g^{ab}_{,c}\Pi^{i}_{a}\Pi^{n}_{b}dx^{c}+g^{ed}\Pi^{i}_{d}d\Pi^{n}_{e}+g^{ed}\Pi^{n}_{d}d\Pi^{i}_{e}. \tag{G.46}$$

It now follows that we can make eq.(G.42) look like eq. (G.45) by adding in and subtracting to eq. 39 the term

$$\frac{1}{2} g^{ka} \Gamma^{b}_{\beta k} \Pi^{i}_{a} \Pi^{n}_{b} \overset{-1}{\Pi}{}^{\beta}_{\rho}$$

which defines the T^{ab}_{β} as

$$T^{ab}_{\beta} = \frac{1}{2} g^{kb} \Gamma^{a}_{\beta k} - \frac{1}{2} g^{ka} \Gamma^{b}_{\beta k} \tag{G.47}$$

which is totally skew-symmetric and thus a valid n-symplectic gauge freedom.

Appendix H

Nonholonomic Momentum Map and Equation (spatial form)

This appendix is a summary of the nonholonomic momentum equation of BKMM with particular application to the vertical rolling hoop.

H.1 Notation and Objects

- (Configuration Manifold Q a PFB)
  $Q \to Q/\mathfrak{G}$ a PFB (structure group $\mathfrak{G}$ with Lie algebra $\mathfrak{g}$) with local coordinates $q^{\mathfrak{I}} = (r^i, s^A)$

- (Constraints defined by a Distribution)
  Δ a distribution defined by those vectors killed by constraint 1-forms

- (Intersection of Orbits with Constraints)

$$S_q = \Delta_q \cap T_q Orb_{\mathfrak{G}}(q)$$

- (Distribution Part of Lie Algebra)

$$\xi^q \in \mathfrak{g}^{\Delta} = \{\xi \in \mathfrak{g} | \xi_q^\star \in S_q\}$$

- (Fixed (Canonical) Basis of Lie Algebra)

  $f_A = (0 \dots \underbrace{1}_{A^{th}\ spot} \dots 0)$ canonical basis of $\mathfrak{g}$ with

$$f_A^\star|_q = T_A^{\mathfrak{I}}(q)\partial_{\mathfrak{I}}|_q \tag{H.1}$$

  where $\partial_{\mathfrak{I}} = (\partial_i, \partial_A) = (\partial/\partial r^i, \partial/\partial s^A)$

- (Moving Basis Criteria)

  Define a *moving basis of* $\mathfrak{g}$ by

$$e_B(q) = \Psi_A^B(q) f_B \tag{H.2}$$

  where Ψ is characterized by

$$\xi^A = \xi^C(q) \overset{-1}{\Psi}{}_C^A(q) \ \ (constant)$$

  That is there are two representations of $\xi \in \mathfrak{g}$ given by

$$\xi = \underbrace{\xi^A(q) f_A}_{\text{fixed basis rep.}} = \xi^C(q) \overset{-1}{\Psi}{}_C^A(q) \Psi_A^B(q) f_B = \underbrace{\xi^B e_B(q)}_{\text{moving basis rep.}}$$

  where in the fixed basis rep. the coordinates are q dependent while in the moving basis rep. (i.e. q dependent basis) the coordinates are fixed. The choice of Ψ will depend on what the constant ξ^A are desired to be.

For example, in the case A=3, one could take

$$\xi = \begin{pmatrix} f_1 & f_2 & f_3 \end{pmatrix} \begin{pmatrix} \xi^1(q) \\ \xi^2(q) \\ \xi^3(q) \end{pmatrix} = \begin{pmatrix} f_1 & f_2 & f_3 \end{pmatrix} \underbrace{\begin{pmatrix} (\xi^1(q) & 0 & 0 \\ 0 & \xi^2(q) & 0 \\ 0 & 0 & \xi^3(q) \end{pmatrix}}_{:=K} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{H.3}$$

which gives $\xi^A = (1,1,1)^t$ and $e_B(q) = f_A K^A_B(q)$

H.2 Non-Holonomic Momentum Map

BKMM define the non-holonomic momentum map J^{nhc} as

$$J^{nhc}(\xi^q) = \left\langle \frac{\partial L}{\partial \dot{q}}, (\xi^q)^\star \right\rangle \tag{H.4}$$

where $\xi^q \in S_q$. The momentum map, J in Lagrangian form is

$$J(\xi) = \left\langle \frac{\partial L}{\partial \dot{q}}, \xi^\star_q \right\rangle \tag{H.5}$$

and thus the non-holonomic momentum map, J^{nhc} is the canonical momentum map on TQ restricted to those Lie algebra elements g^q for which $(g^q)^\star \in \Delta_q$.

H.2.1 Non-Holonomic Momentum Map in Fixed Basis

The coordinate form of eq. (H.4) in the fixed (canonical) basis f_A of $\mathfrak{g}$ is

$$J^{nhc}(\xi^q) = J^{nhc}(\xi^A(q) f_A) = J^{nhc}(f_A)\xi^A(q) := J_A \xi^A(q)$$

where $J_A := J^{nhc}(f_A)$ is given by

$$J_A = \left\langle \frac{\partial L}{\partial \dot{q}}, (f_A)^\star \right\rangle = \left\langle \frac{\partial L}{\partial \dot{q}^\jmath} dq^\jmath, T^\jmath_A(q)\partial_\jmath \right\rangle = \frac{\partial L}{\partial \dot{q}^\jmath} T^\jmath_A(q). \tag{H.6}$$

H.2.2 Non-Holonomic Momentum Map in Ψ Moving Basis

The coordinate form of eq. (H.4) in the moving basis $e_A(q)$ defined by (H.2) is

$$J^{nhc}(\xi^q) = J^{nhc}(\xi^A e_A(q)) = J^{nhc}(e_A(q))\xi^A := J_A(q)\xi^A$$

where $\xi^A = \xi^C(q) \overset{-1}{\Psi}{}^A_C(q)$ (constant) and $J_A(q)$ is given by

$$J_A(q) = \left\langle \frac{\partial L}{\partial \dot{q}}, (e_A(q))^\star \right\rangle = \left\langle \frac{\partial L}{\partial \dot{q}}, (f_B)^\star \right\rangle \Psi^B_A(q) = \frac{\partial L}{\partial \dot{q}^\eth} T^\eth_B(q) \Psi^B_A(q) \tag{H.7}$$

Note H.2.1 $J^{nhc}(\xi^q)$ *in the moving or fixed basis of* $\mathfrak{g}$ *are equal because*

$$J^{nhc}(\xi^q) = J_A \xi^A(q) = J_A \Psi^A_B(q)\xi^B = J_B(q)\xi^B$$

but J_A *and* $J_A(q)$ *are not;* J_A *is the component of* J^{nhc} *in the fixed* f_A *direction while* $J_A(q)$ *is the component of* J^{nhc} *is a configuration dependent direction* $e_A(q)$*.*

H.3 Non-Holonomic Momentum Equation

BKMM proof that the derivative of the non-holonomic momentum map is given by

$$\frac{d}{dt}\left[J^{nhc}(\xi^{q(t)})\right] = \left\langle \frac{\partial L}{\partial \dot{q}}, \left(\frac{d}{dt}\xi^{q(t)}\right)^\star \right\rangle$$

where $\xi^{q(t)}$ is a curve in $\mathfrak{g}^\Delta \subset \mathfrak{g}$, i.e. $(\xi^{q(t)})^\star \in S_{q(t)}$.

H.3.1 Non-Holonomic Momentum Equation in Fixed Basis

$$\frac{d}{dt}\left[J^{nhc}(\xi^{q(t)})\right] = \left\langle \frac{\partial L}{\partial \dot{q}}, \left(\frac{d}{dt}\xi^{q(t)}\right)^\star \right\rangle = \left\langle \frac{\partial L}{\partial \dot{q}}, f^\star_A \right\rangle \left[\frac{d}{dt}\xi^A(q(t))\right] = \frac{\partial L}{\partial \dot{q}^\eth} T^\eth_B(q) \left[\frac{d}{dt}\xi^A(q)\right] \tag{H.8}$$

H.3.2 Non-Holonomic Momentum Equation in Moving Basis

In a moving basis given by $e_A(q)$ the non-holonomic momentum equation becomes

$$\frac{d}{dt}\left[J^{nhc}(\xi^q)\right] = \left\langle \frac{\partial L}{\partial \dot{q}}, \left(\frac{d}{dt}\xi^q\right)^\star \right\rangle = \left\langle \frac{\partial L}{\partial \dot{q}}, \left(\frac{d}{dt}e_A(q)\right)^\star \right\rangle \xi^A \tag{H.9}$$

whereby using

$$\frac{d}{dt}e_A(q) = \frac{d}{dt}(\Psi^B_A(q) f_B) = \Psi^B_{A,\mathcal{J}}(q)\dot{q}^{\mathcal{J}} f_B = \Psi^B_{A,\mathcal{J}}(q) \overset{-1}{\Psi}{}^C_B(q)\dot{q}^{\mathcal{J}} e_C(q) := \Gamma^C_{A\mathcal{J}}(q)\dot{q}^{\mathcal{J}} e_C(q) \tag{H.10}$$

gives

$$\frac{d}{dt}\left[J^{nhc}(\xi^q)\right] = \left\langle \frac{\partial L}{\partial \dot{q}}, e_C(q)^\star \right\rangle \Gamma^C_{A\mathcal{J}}(q)\dot{q}^{\mathcal{J}}\xi^A. \tag{H.11}$$

Upon re-introduction of the notation

$$\left\langle \frac{\partial L}{\partial \dot{q}}, e^\star_A(q) \right\rangle := J_A(q)$$

one obtains for eq. (H.9)

$$\left[\frac{d}{dt}J_A(q) - \Gamma^C_{A;\mathcal{J}} J_C(q)\dot{q}^{\mathcal{J}}\right]\xi^A = 0 \tag{H.12}$$

Remark H.3.1 *Again, the coordinate representations of* $\frac{d}{dt}\left[J^{nhc}(\xi^q)\right]$ *given by eqs. (H.8) and (H.12) are equal. The question at this state is can a moving basis* $e_A(q) = \Psi^A_B(q) f_A$ *be chosen through judicious choice of* Ψ *to make make* $\xi^A = (0\ldots \underbrace{1}_{N^{th}\ spot} \ldots 0)$.
In this moving frame, only the N^{th}*-component of the non-holonomic momentum will survive and will be given by*

$$\frac{d}{dt}J_N(q) = \Gamma^C_{N\mathcal{J}} J_C(q)\dot{q}^{\mathcal{J}}. \tag{H.13}$$

The end result being that $J_N(q)$*, the N-th component of the momentum in the moving* *frame $e_A(q)$ is not in general a conserved quantity. So the key is in the existence and the choice of Ψ.*

H.4 Rolling Hoop–see also BKMM

The configuration manifold of the Rolling Hoop is $Q = \mathbb{R}^2 \times S^1 \times S^1$ with coordinates (x, y, θ, ϕ) and Lagrangian

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}\dot{\phi}^2$$

with constraint 1-forms

$$\begin{aligned}
\omega^1 = dx - R\cos(\phi)d\theta &\implies 0 = \omega^1(\dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\phi}\partial_\phi) \\
&\implies \dot{x} = R\cos(\phi)\dot{\theta} \ \ \text{(rolling constraint)} \\
\omega^2 = dy - R\sin(\phi)d\theta &\implies 0 = \omega^2(\dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\phi}\partial_\phi) \\
&\implies \dot{y} = R\sin(\phi)\dot{\theta} \ \ \text{(rolling constraint)}
\end{aligned}$$

The constrained Lagrangian is thus

$$L_c = \frac{1}{2}(I + mR^2)\dot{\theta}^2 + \frac{1}{2}J\dot{\phi}^2$$

with constrained dynamics (see eq. 2.1.6 in BKMM, a straightforward calculation)

$$\begin{aligned}
(I + mR^2)\ddot{\theta} &= 0 \\
J\ddot{\phi} &= 0.
\end{aligned}$$

The constraint distribution Δ_q follows from find those X, Y for which $\omega^1(X) = 0$ and $\omega^2(Y) = 0$:

$$\Delta_q = span\{\partial_\phi, \partial_\theta + R\cos(\phi)\partial_x + R\sin(\phi)\partial_y\}$$

The *coupled* constrained dynamics for the rolling hoop are summarized as

$$(I + mR^2)\ddot{\theta} = 0 \tag{H.14}$$

$$J\ddot{\phi} = 0 \tag{H.15}$$

$$\dot{x} = R\cos(\phi)\dot{\theta} \tag{H.16}$$

$$\dot{y} = R\sin(\phi)\dot{\theta} \tag{H.17}$$

H.4.1 Non-Holonomic Momentum Equation for SE(2) action on $Q = \mathbb{R}^2 \times S^1 \times S^1$

Following BKMM we show how the non-holonomic momentum equation (for an SE(2)) action relates to eq. (H.15).

Define the action of SE(2) on $\mathbb{R}^2 \times S^1 \times S^1$ with coordinates

$$g = \{s^A\} = (x, y, \phi)$$

are the group variables and $r = \{r^i\} = (\theta)$ is the shape variable by

$$\begin{aligned} \alpha(t) &= L_{h(t)}(r, g) = (\Phi(t), X(t), Y(t)) \cdot (\theta, \phi, x, y) \\ &:= [\theta, \phi + \Phi(t), x\cos\Phi(t) - y\sin\Phi(t) + X(t), x\sin\Phi(t)) + y\cos\Phi(t) + Y(t)] \end{aligned}$$

where

$$h(t) = (\Phi(t), X(t), Y(t))$$

is a curve in SE(2) such that $h(0) = Id.$ and $\alpha(0) = h(0) \cdot (r, g) = (r, g)$. That is, $\alpha(t)$ is a curve in P through (r,g) at t=0 and h(t) is a curve in SE(2) through the identity at t=0. It follows that

$$L_{h^{-1}(t)}(r,g) = [\theta, \phi - \Phi(t), x\cos\Phi(t) + y\sin\Phi(t) - X(t)\cos\Phi(t) - Y(t)\sin\Phi(t),$$
$$-x\sin\Phi(t) + y\cos\Phi(t) + X(t)\sin\Phi(t) - Y(t)\cos\Phi(t)]$$

Recall that the push-forward of $X_g = \dot{s}^A \partial_A|_g$ to a vector $(L_{h^{-1}})_\star X_g$ is given by

$$\dot{s}^A[(L_{h^{-1}})_\star \partial_A|_g] = \dot{s}^A Z_A^B(g,r)\partial_B|_{h^{-1}g} := \xi^B(g,r)|_{h^{-1}g}\partial_B|_{h^{-1}g}$$

where $Z_A^B(g,r)$ are given by

$$(L_{h^{-1}})_\star(\partial_A|_g)(s^B) = \partial_A|_g[L_{h^{-1}}s^B] = \partial_A|_g[L_{h^{-1}}g]^B$$

So, if $h(t) = g(t) = (x(t), y(t), \phi(t))$ then

$$\xi^B(r,g)|_{id} = \dot{s}^A(\partial_A|_g[L_{g^{-1}(t)}g]^B) \tag{H.18}$$

or

$$\xi(r,g) = \dot{g} \cdot (Jacobian_{\text{wrt group coord}} L_{g^{-1}(t)}g) \tag{H.19}$$

where $\xi(r,g)|_{id}$ are the components of ξ thought of as a Lie algebra elements in the basis $\partial_A|_{id}$. That is,

$$\xi = \xi(g,r)^B|_{id}\partial_B|_{id}$$

Using $L_{g^{-1}(t)}g$ on the previous page we get

$$\xi^2(r,g) = \dot{x}\cos\phi(t) - \dot{y}\sin\phi(t) \tag{H.20}$$

$$\xi^3(r,g) = \dot{x}\sin\phi(t) + \dot{y}\cos\phi(t) \tag{H.21}$$

$$\xi^1(r,g) = \dot{\theta} \tag{H.22}$$

Using the notation $\xi = (\xi^1, \xi^2, \xi^3) = \xi^1 f_1 + \xi^2 f_2 + \xi^3 f_3$ [thinking of the matrices

$$f_1 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} f_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} f_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

as the vectors, respectively,

$$f_1 = (0, 0, 1) = \partial_\phi|_{id}, \quad f_2 = (1, 0, 0) = \partial_x|_{id}, \quad f_3 = (0, 1, 0) = \partial_y|_{id}$$

the tangents to the group orbits can be found by computing $\dot{\alpha}(0)$. In coordinates we obtain

$$\begin{aligned} \frac{d}{dt}(\alpha^A(0))\partial_A &= \frac{d}{dt}\left(L_{g(t)}(r, g)\right)|_{t=0} \\ &= (-y\xi^1 + \xi^2)\partial_x + (x\xi^1 + \xi^3)\partial_y + (\xi^1)\partial_\theta \\ &= -y\xi^1\partial_x + x\xi^1\partial_y + \xi^1\partial_\phi + \xi^2\partial_x + \xi^3\partial_y. \end{aligned}$$

where $\dot{x}(0) = \xi^2, \dot{y}(0) = \xi^3$ and $\dot{\phi}(0) = \xi^1$. In the $\{f_A\}$-basis, notice that the curves $g_2(t) = (0, x(t), 0)$, $g_3(t) = (0, 0, y(t))$ and $g_1(t) = (\phi(t), 0, 0)$ give

$$\begin{aligned} \alpha_1(t) = g_1(t) \cdot (r, g) &\rightarrow \dot{\alpha}_3(0) = \xi^1(-y\partial_x + x\partial_y + \partial_\phi) \\ \alpha_2(t) = g_2(t) \cdot (r, g) &\rightarrow \dot{\alpha}_2(0) = \xi^2\partial_x \\ \alpha_3(t) = g_3(t) \cdot (r, g) &\rightarrow \dot{\alpha}_2(0) = \xi^3\partial_y \end{aligned}$$

which suggests the definition of the map $\star : \mathfrak{g} \rightarrow T_pOrb$ by

$$(f_1)^\star(q) = -y\partial_x + x\partial_y + \partial_\phi, \quad (f_2)^\star(q) = \partial_x, \quad (f_3)^\star(q) = \partial_y$$

It follows that

$$T_(g, r)Orb_{SE(2)}(g, r) = span\{\partial_x, \partial_y, \partial_\phi\}$$

which implies

$$S_q = span\{\partial_\phi\}.$$

Thus, following the notation of sections H.2 and H.3,

$$\xi^q = yf_2 - xf_3 + 1f_1 \quad since \;\; (\xi^q)^\star = yf_2^\star - xf_3^\star + 1f_1^\star = \partial_\phi$$

with

$$(\xi^A(q)) = (1, y, -x) \implies \left(\frac{d}{dt}\xi^A(q)\right) = (0, \dot{y}, -\dot{x})$$

and

$$T = (T_A^\partial(q)) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ -y & 1 & 0 \\ x & 0 & 1 \end{bmatrix}$$

feed into eq. (H.6) to give the components of J^{nhc} is the fixed basis f_A

$$(J_A) := (J^{nhc}(f_A)) = \begin{pmatrix} J\dot{\phi} - my\dot{x} + mx\dot{y} & m\dot{x} & m\dot{y} \end{pmatrix}$$

while the non-holonomic momentum map given by (H.8) is

$$J^{nhc}(\xi^q) \quad = \quad \frac{\partial L}{\partial \dot{q}^\partial} T_A^\partial(q)\xi^A(q) = \begin{pmatrix} I\dot{\theta} & J\dot{\phi} & m\dot{x} & m\dot{y} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = J\dot{\phi}.$$

The non-holonomic momentum equation given by eq. (H.9) becomes

$$\frac{d}{dt}\left[J^{nhc}(\xi^q)\right] \;=\; \frac{\partial L}{\partial \dot{q}^{\mathfrak{d}}} T^{\mathfrak{d}}_A(q)\left[\frac{d}{dt}\xi^A(q)\right] = \begin{pmatrix} I\dot{\theta} & J\dot{\phi} & m\dot{x} & m\dot{y} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \dot{y} \\ -\dot{x} \end{pmatrix} = 0$$

The goal now is to look at a particular moving basis determined by $\Psi(q)$ (defined in eq. (H.2)) for which there is a single non-zero component of the non-holonomic momenta . An obvious choice presents itself by simply noticing that

$$\xi^q = yf_2 - xf_3 + 1f_1 = \xi^1 e_1(q) = (1)e_1(q) + (0)f_2(q) + (0)f_3(q)$$

with $e_1(q) := yf_2 - xf_3 + 1f_1$, $e_2(q) = f_2$ and $e_3(q) = f_3$ and thus

$$(\Psi^A_B(q)) = \begin{pmatrix} 1 & 0 & 0 \\ y & 1 & 0 \\ -x & 0 & 1 \end{pmatrix}$$

which gives the non-holonomic momenta in the moving basis $\{e_A(q)\}$ of $\mathfrak{g}$ (cf. eq. (H.7))

$$(J_A(q)) := (J^{nhc}(e_A(q))) = \frac{\partial L}{\partial \dot{q}^{\mathfrak{d}}} T^{\mathfrak{d}}_B(q)\Psi^B_A(q) = \begin{pmatrix} J\dot{\phi} & m\dot{x} & m\dot{y} \end{pmatrix} \tag{H.23}$$

which because ξ^q has coordinate form $\xi^A = (1,0,0)$ in the moving basis $J\dot{\phi}$ is the first component of J^{nhc}, $J_1(q)$ in the moving basis, $\{e_A(q)\}$. One readily finds that $\dot{J}_2 = 0$ from which the equation $\ddot{J}_2(q) = 0$ is recovered.

Note H.4.1 *Notice that* $e_2^\star(q) = \partial_\phi$ *and so*

$$\begin{pmatrix} e_2^\star(q) \end{pmatrix} = \begin{pmatrix} \partial_\theta & \partial_\phi & \partial_x & \partial_y \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{H.24}$$

H.4.2 Non-Holonomic Momentum Equation for $\mathbb{R}^2 \times S^1$ action on $Q = \mathbb{R}^2 \times S^1 \times S^1$

Following BKMM we show how the non-holonomic momentum equation (for an $\mathbb{R}^2 \times S^1$) action relates to eq. (H.14).

Define the action of $\mathbb{R}^2 \times S^1$ on $\mathbb{R}^2 \times S^1 \times S^1$ by (thinking of (x, y, θ) as the group variables and ϕ as the shape variable)

$$\begin{aligned} \alpha(t) &= L_{h(t)}(r, g) = (\Theta(t), X(t), Y(t)) \cdot (\theta, \phi, x, y) \\ &:= [\theta + \Theta(t), \phi, x + X(t), y + Y(t)] \end{aligned}$$

where $h(t) = (X(t), Y(t), \Theta(t))$ is a curve in $\mathbb{R}^2 \times S^1$ such that $h(0) = Id$ and $\alpha(0) = h(0) \cdot (r, g) = (r, g)$. That is, $\alpha(t)$ is a curve in P through (r,g) at t=0 and h(t) is a curve in SE(2) through the identity at t=0. It follows that

$$L_{h^{-1}(t)}(g, r) = [x - X(t), y - Y(t), \theta - \Theta(t), \phi]$$

which, from (H.19), one obtains

$$\xi = [\dot{\theta}, \dot{x}, \dot{y}, \dot{\theta}]$$

From $\frac{d}{dt}(\alpha(t))|_{t=0}$ one finds

$$\begin{aligned} f_2 = (1,0,0) &\quad\rightarrow\quad f_1^\star = \partial_x \\ f_3 = (0,1,0) &\quad\rightarrow\quad f_2^\star = \partial_y \\ f_1 = (0,0,1) &\quad\rightarrow\quad f_3^\star = \partial_\theta \end{aligned}$$

and so

$$T_(r,g)Orb_{\mathbb{R}^2\times S^1}(r,g) = span\{\partial_x, \partial_y, \partial_\theta\}$$

which gives

$$S_q = span\{\partial_\theta + R\cos(\phi)\partial_x + R\sin(\phi)\partial_y\}$$

and thus

$$\begin{aligned} \xi^q &= 1f_1 + R\cos(\phi)f_2 + R\sin(\phi)f_3 = (1)e_1(q) + (0)e_2(q) + (0)e_3 \\ \implies &\quad (\xi^A(q)) = (R\cos(\phi), R\sin(\phi), 1) or (\xi^A) = (1,0,0) \\ \implies &\quad (\frac{d}{dt}\xi^A(q)) = (-R\sin(\phi)\dot{\phi}, R\cos(\phi)\dot{\phi}, 0) \end{aligned}$$

which feed into (H.6) to give

$$\begin{aligned} J^{nhc}(\xi^q) &= \frac{\partial L}{\partial \dot{q}^\eth} T_A^\eth(q)\xi^A(q) \\ &= \begin{pmatrix} m\dot{x} & m\dot{y} & J\dot{\phi} & I\dot{\theta} \end{pmatrix} \begin{pmatrix} R\cos(\phi) \\ R\sin(\phi) \\ 1 \\ 0 \end{pmatrix} \\ &= (I + mR^2)\dot{\theta} \,\text{(upon substitution of constraints)} \end{aligned}$$

and into (H.9) to give (again using the constraints)

$$\frac{d}{dt}\left[J^{nhc}(\xi^q)\right] = \frac{\partial L}{\partial \dot{q}^\eth} T_A^\eth(q) \left[\frac{d}{dt}\xi^A(q)\right] = 0$$

which recovers $(I+mR^2)\ddot{\theta} = 0$. Introduction of a moving basis (as in the SE(2)-action section) defined by $e_1(q) = 1f_1 + R\cos(\phi)f_2 + R\sin(\phi)f_3$, $e_2(q) = f_2$ and $e_3(q) = f_3$ gives

$$\Psi(q) = \begin{bmatrix} 1 & 0 & 0 \\ R\cos(\phi) & 1 & 0 \\ R\sin(\phi) & 0 & 1 \end{bmatrix} \implies \xi^A = (1,0,0)^t$$

gives

$$J_1(q) = (I+mR^2)\dot{\theta} \;\; and \;\; \dot{J}_1(q) = 0 \implies (I+mR^2)\ddot{\theta} = 0$$

Note H.4.2 *Notice that* $e_1^\star(q) = \partial_\theta + R\cos(\phi)\partial_x + R\sin(\phi)\partial_y$ *and so*

$$\begin{pmatrix} e_1^\star(q) \end{pmatrix} = \begin{pmatrix} \partial_\theta & \partial_\phi & \partial_x & \partial_y \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ R\cos(\phi) \\ R\sin(\phi) \end{pmatrix} \tag{H.25}$$

which together with eq. (H.24) give

$$\begin{pmatrix} e_1^\star(q) & e_2^\star(q) \end{pmatrix} = \begin{pmatrix} \partial_\theta & \partial_\phi & \partial_x & \partial_y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ R\cos(\phi) & 0 \\ R\sin(\phi) & 0 \end{pmatrix} \tag{H.26}$$

Viewing $e_1^\star(q), e_2^\star(q)$ *as part of a basis of* T_qQ *which we can fill out (rig) to a full basis via the addition of* $R_1 = \partial_x$ *and* $R_2 = \partial_y$ *we obtain*

$$\begin{pmatrix} R_1 & R_2 & e_1^\star(q) & e_2^\star(q) \end{pmatrix} = \begin{pmatrix} \partial_\theta & \partial_\phi & \partial_x & \partial_y \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & R\cos(\phi) & 0 \\ 0 & 1 & R\sin(\phi) & 0 \end{pmatrix} \tag{H.27}$$

This matrix is critical in the n-symplectic derivation of the non-holonomic momentum map and momentum equations, see section I.5(a).

At this point we can summarize the *coupled* dynamics for the rolling hoop as two constraint equations (eqs. (H.28) and (H.29)), two (non-holonomic) conservation laws (eq. (H.30) and (H.31)) and two momentum equations (eqs. (H.32) and (H.33))

$$\dot{x} = R\cos(\phi)\dot{\theta} \tag{H.28}$$

$$\dot{y} = R\sin(\phi)\dot{\theta} \tag{H.29}$$

$$\dot{J_1(q)} = 0 \tag{H.30}$$

$$\dot{J_2(q)} = 0 \tag{H.31}$$

$$J_1(q) = (I + mR^2)\dot{\theta} \tag{H.32}$$

$$J_2(q) = J\dot{\phi} \tag{H.33}$$

Appendix I

A Glimpse at the Role of n-Symplectic Gauge Freedom and Scalar Potentials in Control of Mechanical Systems

I.1 Canonical Rank 2 Dynamics with Gauge Freedom

The goal for this subsection is to briefly explore the affect various gauge selections have on control of a moving frame attached to a particles geodesic motion. First we show that the canonical rank–2 n-symplectic dynamics can be written as a geodesic equation and a modified form of the parallel transport of the generalized momenta. The modification is manifest in the form of extra forces due to a torsion–free connection. In the last part of this subsection we address how to control a particular leg of moving frame along the geodesic motion using a *semi-metric* form of a torsion free connection.

Note I.1.1 *(On index notation) For this section we use the greek indices* $\alpha, \beta, \eta, \mu, \xi =$

$1 \ldots n = dim(Q)$ *as the configuration space coordinate indices.*

It follows from eq. (G.32) that the dynamic equations for the n-vector fields $X^{\gamma}_{\hat{g}}$ for non-trivial gauge freedom T given by eq. (G.47) are

$$\begin{aligned} \dot{x}^{\beta} &= g^{\alpha\beta}\Pi^{\gamma}_{\alpha} \\ \dot{\Pi}^{\eta}_{\beta} &= -\frac{1}{2}\left(Q^{\alpha\xi\cdot}_{\cdot\cdot\beta} - 2\Gamma^{\xi}_{\mu\beta}g^{\alpha\mu}\right)\Pi^{\gamma}_{\alpha}\Pi^{\eta}_{\xi} \end{aligned} \tag{I.1}$$

Equation (I.1) follow from a straight-forward substitution of eq. (G.33) and eq. (G.45) into eq. (G.32). Again, for the $\gamma = 1$ vector field and $\eta = (1, \bar{\eta})$ where $\bar{\eta} = 2..n$ eq. (I.1) becomes

$$\dot{x}^{\beta} = g^{\alpha\beta}p_{\alpha} \tag{I.2}$$

$$\dot{p}_{\beta} = -\frac{1}{2}\left(Q^{\alpha\xi\cdot}_{\cdot\cdot\beta} - 2\Gamma^{\xi}_{\mu\beta}g^{\alpha\mu}\right)p_{\alpha}p_{\xi} \tag{I.3}$$

$$\dot{\Pi}^{\bar{\eta}}_{\beta} = -\frac{1}{2}\left(Q^{\alpha\xi\cdot}_{\cdot\cdot\beta} - 2\Gamma^{\xi}_{\mu\beta}g^{\alpha\mu}\right)p_{\alpha}\Pi^{\bar{\eta}}_{\xi} \tag{I.4}$$

where we define $p_{\alpha} := \Pi^{1}_{\alpha}$. We recall here that a general torsion free connection Γ can be expressed as (c.f. the proof of the Levi-Civita connection in Appendix A which itself follows the proof in [37, pg. 132])

$$\begin{aligned} \Gamma^{j}_{ab} &= \left\{{}^{j}_{ab}\right\} + \frac{1}{2}g^{ji}(g_{ab;i} - g_{bi;a} - g_{ia;b}) \\ &= \left\{{}^{j}_{ab}\right\} + \frac{1}{2}g^{ji}(-Q_{abi} + Q_{bia} + Q_{iab}) \end{aligned}$$

$$= \left\{{}^{j}_{ab}\right\} + \frac{1}{2}\left(-Q^{\cdot\cdot j}_{ab\cdot} + Q^{\cdot j\cdot}_{b\cdot a} + Q^{j\cdot\cdot}_{\cdot ab}\right) \tag{I.5}$$

$$= \left\{{}^{j}_{ab}\right\} + \frac{1}{2}\left(-Q^{\cdot\cdot j}_{ba\cdot} + Q^{\cdot j\cdot}_{b\cdot a} + Q^{j\cdot\cdot}_{\cdot ab}\right) \tag{I.6}$$

where, for example, $-Q_{abi} := g_{ab;i}$ and the last two equations equal as the metric g is symmetric. That $-Q_{abi} := g_{ab;i}$ follows from

$$\begin{aligned}
Q^{ab\cdot}_{\cdot\cdot i} &:= g^{ab}_{;i} \\
\Longrightarrow &\qquad Q^{\cdot b\cdot}_{c\cdot i} = g_{ac}Q^{ab\cdot}_{\cdot\cdot i} = g_{ac}(g^{ab}_{;i}) = (g_{ac}g^{ab})_{;i} - g^{ab}g_{ac;i} \\
\Longrightarrow &\qquad Q_{cdi} = Q_{dci} = -g_{dc;i}
\end{aligned}$$

Note I.1.2 *(On equivalent notation) Herein we put the covariant index last while in [37] it is put first. That is, we use $-Q_{abi} = g_{ab;i}$ while [37] use $-Q_{iab} = \nabla_i g_{ab}$. To test the equivalence of decomposition (I.6) with the decomposition of [37] write the above decomposition as*

$$\begin{aligned}
\Gamma^j_{ab} &= \left\{{}^{\,j}_{ab}\right\} + \frac{1}{2}g^{ji}\left(\nabla_i g_{ab} - \nabla_a g_{bi} - \nabla_b g_{ia}\right) \\
&= \left\{{}^{\,j}_{ab}\right\} + \frac{1}{2}g^{ji}(-Q_{iab} + Q_{abi} + Q_{bia}) \\
&= \left\{{}^{\,j}_{ab}\right\} + \frac{1}{2}\left(-Q^{j\cdot\cdot}_{\cdot ab} + Q^{\cdot\cdot j}_{ab\cdot} + Q^{\cdot j\cdot}_{b\cdot a}\right) \\
&\qquad \textit{use } j \to \chi \;\; a \to \mu \;\; b \to \lambda \\
\Gamma^\chi_{\mu\lambda} &= \left\{{}^{\,\chi}_{\mu\lambda}\right\} + \frac{1}{2}\left(-Q^{\chi\cdot\cdot}_{\cdot\mu\lambda} + Q^{\cdot\cdot\chi}_{\mu\lambda\cdot} + Q^{\cdot\chi\cdot}_{\lambda\cdot\mu}\right) \qquad \text{(I.7)}
\end{aligned}$$

which is eq. 3.5 in [37, pg. 132], indicating that the decomposition (I.6) and Schouten's (I.7) are equivalent.

Using decomposition (I.6) in eq. (I.3) (after symmetrizing on α and ξ) yields

$$\begin{aligned}
\dot{p}_\beta &= -\frac{1}{2}\left(Q^{\alpha\xi\cdot}_{\cdot\cdot\beta}\right)p_\alpha p_\xi + \frac{1}{2}\left[\Gamma^{\xi}_{\mu\beta}g^{\alpha\mu} + \Gamma^{\alpha}_{\mu\beta}g^{\xi\mu}\right]p_\alpha p_\xi \\
&= -\frac{1}{2}\left(Q^{\alpha\xi\cdot}_{\cdot\cdot\beta}\right)p_\alpha p_\xi \\
&\quad +\frac{1}{2}\left[\left(\{{}^{\xi}_{\mu\beta}\} + \frac{1}{2}\left(-Q^{\cdot\cdot\xi}_{\beta\mu\cdot} + Q^{\cdot\xi\cdot}_{\beta\cdot\mu} + Q^{\xi\cdot\cdot}_{\cdot\mu\beta}\right)\right)g^{\alpha\mu}\right] \\
&\quad +\left[\left(\{{}^{\alpha}_{\mu\beta}\} + \frac{1}{2}\left(-Q^{\cdot\cdot\alpha}_{\beta\mu\cdot} + Q^{\cdot\alpha\cdot}_{\beta\cdot\mu} + Q^{\alpha\cdot\cdot}_{\cdot\mu\beta}\right)\right)g^{\xi\mu}\right]p_\alpha p_\xi \\
&= -\frac{1}{2}\left(Q^{\alpha\xi\cdot}_{\cdot\cdot\beta}\right)p_\alpha p_\xi + \frac{1}{2}\left[\{{}^{\xi}_{\mu\beta}\}g^{\alpha\mu} + \{{}^{\alpha}_{\mu\beta}\}g^{\xi\mu}\right]p_\alpha p_\xi \\
&+\frac{1}{2}\left[-\frac{1}{2}Q^{\cdot\alpha\xi}_{\beta\cdot\cdot} + \frac{1}{2}Q^{\cdot\xi\alpha}_{\beta\cdot\cdot} + \frac{1}{2}Q^{\xi\alpha\cdot}_{\cdot\cdot\beta} - \frac{1}{2}Q^{\cdot\xi\alpha}_{\beta\cdot\cdot} + \frac{1}{2}Q^{\cdot\alpha\xi}_{\beta\cdot\cdot} + \frac{1}{2}Q^{\alpha\xi\cdot}_{\cdot\cdot\beta}\right]p_\alpha p_\xi \\
&\quad \text{1 and 5; 2 and 4 cancel and using symmetry } Q^{\alpha\xi\cdot}_{\cdot\cdot\beta} = Q^{\xi\alpha\cdot}_{\cdot\cdot\beta} \\
&= -\frac{1}{2}\left(Q^{\alpha\xi\cdot}_{\cdot\cdot\beta}\right)p_\alpha p_\xi + \frac{1}{2}Q^{\xi\alpha\cdot}_{\cdot\cdot\beta}p_\alpha p_\xi + \frac{1}{2}\left[\{{}^{\xi}_{\mu\beta}\}g^{\alpha\mu} + \{{}^{\alpha}_{\mu\beta}\}g^{\xi\mu}\right]p_\alpha p_\xi
\end{aligned}$$

Again using $Q^{\alpha\xi\cdot}_{\cdot\cdot\beta} = g^{\alpha\xi}_{;\beta}$ is symmetric in α and ξ and un-symmetrizing the Levi-Civita terms the last equation above reduces to

$$\begin{aligned}
\dot{p}_\beta &= \{{}^{\xi}_{\mu\beta}\}g^{\alpha\mu}p_\alpha p_\xi \\
&= g_{\Delta\xi}\{{}^{\xi}_{\mu\beta}\}\dot{x}^\mu\dot{x}^\Delta
\end{aligned}$$

Using eq. (I.2) it is clear that $p_\beta = g_{\eta\beta}\dot{x}^\eta$ and thus $\dot{p}_\beta = g_{\eta\beta,\sigma}\dot{x}^\sigma\dot{x}^\eta + g_{\eta\beta}\ddot{x}^\eta$. The partial derivative can be replaced with a covariant derivative by adding in and subtracting the appropriate terms; that is

$$\begin{aligned}
g_{\eta\beta,\sigma} &= g_{\eta\beta,\sigma} - g_{n\beta}\{{}^{n}_{\sigma\eta}\} - g_{\eta m}\{{}^{m}_{\sigma\beta}\} + g_{n\beta}\{{}^{n}_{\sigma\eta}\} + g_{\eta m}\{{}^{m}_{\sigma\beta}\} \\
&= g_{\eta\beta;\sigma} + g_{n\beta}\{{}^{n}_{\sigma\eta}\} + g_{\eta m}\{{}^{m}_{\sigma\beta}\}.
\end{aligned}$$

As g is covariant constant (wrt to the Levi-Civita connection $\{\cdot\}$) the momentum dynamics $\dot{p}_\beta$ become

$$\begin{aligned} & g_{\eta\beta}\ddot{x}^\eta + \left[g_{a\beta}\{{}^{a}_{\sigma\eta}\} + g_{\eta\beta}\{{}^{\xi}_{\sigma\beta}\}\right]\dot{x}^\sigma\dot{x}^\eta = g_{\eta\xi}\{{}^{\xi}_{\sigma\beta}\}\dot{x}^\sigma\dot{x}^\eta \\ \Longrightarrow\quad & \ddot{x}^\Delta + \{{}^{\Delta}_{\sigma\eta}\}\dot{x}^\sigma\dot{x}^\eta = 0 \end{aligned} \tag{I.8}$$

which are the geodesic equations for the Levi-Civita connection. Again, using the decomposition formula (I.5) in eq. (I.1) and the symmetry $Q^{\alpha\xi\cdot}_{\cdot\cdot\beta} = Q^{\xi\alpha\cdot}_{\cdot\cdot\beta}$ one obtains

$$\begin{aligned} \dot{\Pi}^{\bar{\eta}}_{\beta} &= \left[-\frac{1}{2}Q^{\alpha\xi\cdot}_{\cdot\cdot\beta} + \left(\{{}^{\xi}_{\mu\beta}\} + \frac{1}{2}\left(-Q^{\cdot\cdot\xi}_{\mu\beta\cdot} + Q^{\cdot\xi\cdot}_{\beta\cdot\mu} + Q^{\xi\cdot\cdot}_{\cdot\mu\beta}\right)\right)g^{\alpha\mu}\right]p_\alpha\Pi^{\bar{\eta}}_{\xi} \\ &= \left[\{{}^{\xi}_{\mu\beta}\} + \frac{1}{2}\left(Q^{\cdot\xi\cdot}_{\beta\cdot\mu} - Q^{\cdot\cdot\xi}_{\mu\beta\cdot}\right)\right]g^{\alpha\mu}p_\alpha\Pi^{\bar{\eta}}_{\xi} \end{aligned}$$

which can be written as

$$\frac{D}{dt}\Pi^{\bar{\eta}}_{\beta} = \frac{1}{2}\left(Q^{\cdot\xi\cdot}_{\beta\cdot\mu} - Q^{\cdot\cdot\xi}_{\mu\beta\cdot}\right)\dot{x}^\mu\Pi^{\bar{\eta}}_{\xi} \tag{I.9}$$

where $\frac{D}{dt}\Pi^{\bar{\eta}}_{\beta} := \dot{\Pi}^{\bar{\eta}}_{\beta} - \{{}^{\xi}_{\mu\beta}\}\dot{x}^\mu\Pi^{\bar{\eta}}_{\xi}$.

Remark I.1.1 *Equations (I.8) and (I.9) are given in [29] modulo the detail we have provided. For Q=0 the geodesic equation (I.8) and the parallel transport of the spatial triad (or laboratory frame) equation (I.9) collectively represent, in relativity language, the dynamics of a free falling, non-rotating observer in spacetime. One would not be remise to envision a spacecraft hurtling through a region of "empty" space with the center of mass tracing a "straight" line and the outer hull lifelessly devoid of motion about the center of mass. A natural question is "Can the attitude of the spacecraft be controlled to, say, align the cockpit of the spacecraft with the overall (geodesic) direction of motion using the Q terms as "external" control forces?"*

Following Schouten [37, pg. 133], a natural choice is to use a *semi-metric* connection

$$
\begin{aligned}
Q^{ab\cdot}_{\cdot\cdot i} &= g^{ab} A_i \\
Q_{dci} &= -g_{dc} A_i \\
Q^{\cdot b\cdot}_{c\cdot i} &= -\delta^b_c A_i \\
g^{ai} Q_{dci} = Q^{\cdot\cdot a}_{dc\cdot} &= -g_{dc} g^{ai} A_i
\end{aligned}
$$

where the second to last line follows from $Q^{\cdot b\cdot}_{c\cdot i} = g^{ab} Q_{cai} = -g^{ab} g_{ca} A_i$. At least at this stage of the thought process and in keeping with the spacecraft analogy, we view the vector of control inputs $[A_i]$ as a thrust vector. The steps to formulating and solving the resulting controlled system of equations are

- Start with a given metric $g_{\alpha\beta}$ (and the inverse $g^{\alpha\beta}$) from which the Christoffel symbols $\{{}^{\xi}_{\sigma\delta}\}$ can be found.
- Choose the vector of control inputs $[A_\zeta]$.
- The dynamics for $Q^{ab\cdot}_{\cdot\cdot i} = g^{ab} A_i \neq 0$ are

$$
\begin{aligned}
&\ddot{x}^{\Delta} = -\{{}^{\Delta}_{\sigma\eta}\} \dot{x}^{\sigma} \dot{x}^{\eta} \\
&\dot{\Pi}^{\bar{\eta}}_{\beta} = \left[\{{}^{\xi}_{\mu\beta}\} + \frac{1}{2} \left(-\delta^{\xi}_{\beta} A_{\mu} - -g_{\mu\beta} g^{\xi\zeta} A_{\zeta} \right) \right] \dot{x}^{\mu} \Pi^{\bar{\eta}}_{\xi}
\end{aligned}
\tag{I.10}
$$

- Solve the above coupled system for $\dot{x}^{\alpha} := v^{\alpha}$ and $\Pi^{\bar{\eta}}_{\xi}$
- Solve for p_β using $p_\beta = g_{\alpha\eta} v^{\alpha}$
- Form the matrix denoted Π with first row p_β then 2 through n rows $\Pi^2_\beta, \Pi^3_\beta \ldots \Pi^n_\beta$. This is the matrix describing how the coframe moves relative to the fixed spatial (coordinated) coframe.
- Invert the matrix Π and call it V. The matrix V describes how the frame moves along the geodesic relative to a fixed spatial (coordinated) frame.

As a special case consider $g = \delta \implies \{\cdot\} = 0$. That is, we envision body whose center of mass tracks an actual straight line in $\mathbb{R}^3$ where the "skeleton" of the body is defined by the spatial triad attached to the center of mass. The time evolution or motion of this body is determined by the position of the center of mass along the line and the orientation of the triad for each point along the line.

The dynamics (I.10) in this case reduce to

$$\begin{aligned}
\dot{v}^{\Delta} &= 0 \implies v^{\Delta} = const. \\
\dot{x}^{\Delta} &= v^{\Delta} \implies x^{\Delta} = linear \\
\dot{\Pi}^{\bar{\eta}}_{\beta} &= -\frac{1}{2}A_{\mu}v^{\mu}\Pi^{\bar{\eta}}_{\beta} + \frac{1}{2}v_{\beta}A^{\xi}\Pi^{\bar{\eta}}_{\xi}
\end{aligned}$$

where the components of the row vector A_μ equal the components of the column vector A^μ as the metric is the identity (similarly for v_β and v^β). For this special case, we are interested in finding the control inputs A which aligns a leg of the triad along the straight line motion in $\mathbb{R}^3$ of center of mass.

The frame dynamics of equation (I.11) can be written in expanded form as

$$\begin{aligned}
\frac{d}{dt}\Pi_{21} + \frac{1}{2}\left(A_2 v_2 + A_3 v_3\right)\Pi_{21} - \frac{1}{2}\left(A_2\Pi_{22} + A_3\Pi_{23}\right)v_1 = 0 \\
\frac{d}{dt}\Pi_{22} + \frac{1}{2}\left(A_1 v_1 + A_3 v_3\right)\Pi_{22} - \frac{1}{2}\left(A_1\Pi_{21} + A_3\Pi_{23}\right)v_2 = 0 \\
\frac{d}{dt}\Pi_{23} + \frac{1}{2}\left(A_1 v_1 + A_2 v_2\right)\Pi_{23} - \frac{1}{2}\left(A_1\Pi_{21} + A_2\Pi_{22}\right)v_3 = 0 \\
\frac{d}{dt}\Pi_{31} + \frac{1}{2}\left(A_2 v_2 + A_3 v_3\right)\Pi_{31} - \frac{1}{2}\left(A_2\Pi_{32} + A_3\Pi_{33}\right)v_1 = 0 \\
\frac{d}{dt}\Pi_{32} + \frac{1}{2}\left(A_1 v_1 + A_3 v_3\right)\Pi_{32} - \frac{1}{2}\left(A_1\Pi_{31} + A_3\Pi_{33}\right)v_2 = 0 \\
\frac{d}{dt}\Pi_{33} + \frac{1}{2}\left(A_1 v_1 + A_2 v_2\right)\Pi_{33} - \frac{1}{2}\left(A_1\Pi_{31} + A_2\Pi_{32}\right)v_3 = 0.
\end{aligned}$$

Along with $\dot{p}_\alpha = \dot{\Pi}^1_\alpha = 0$, the full momentum dynamics can be expressed as as linear system of the form

$$\dot{\vec{\Pi}} = M(v, A)\vec{\Pi} \tag{I.11}$$

where the vector $\vec{\Pi}$ is given by $[p_1, p_2, p_3, \Pi_{21}, \ldots, \Pi_{33}]$ and

$$M(v, A) = \mathrm{diag}(0, block_1, block_1)$$

such that *0* is a 3×3 matrix of zeros and

$$block_1 = \begin{bmatrix} -\frac{1}{2}\,\mathrm{A}_2\,\mathrm{v}_2 - \frac{1}{2}\,\mathrm{A}_3\,\mathrm{v}_3 & \frac{1}{2}\,\mathrm{A}_2\,\mathrm{v}_1 & \frac{1}{2}\,\mathrm{A}_3\,\mathrm{v}_1 \\ \frac{1}{2}\,\mathrm{A}_1\,\mathrm{v}_2 & -\frac{1}{2}\,\mathrm{A}_1\,\mathrm{v}_1 - 1/2\,\mathrm{A}_3\,\mathrm{v}_3 & \frac{1}{2}\,\mathrm{A}_3\,\mathrm{v}_2 \\ \frac{1}{2}\,\mathrm{A}_1\,\mathrm{v}_3 & \frac{1}{2}\,\mathrm{A}_2\,\mathrm{v}_3 & -\frac{1}{2}\,\mathrm{A}_1\,\mathrm{v}_1 - \frac{1}{2}\,\mathrm{A}_2\,\mathrm{v}_2 \end{bmatrix}.$$

Assumption I.1.1 *For a general metric g, the vector $v = [v_1, v_2, v_3]$ will not be constant in time. For $g = \delta$, however, v is constant in time. Making the further assumption that the vector of control inputs $A = [A_1, A_2, A_3]$ is also constant in time then the $M(v, A)$ is a constant matrix and the resulting linear system can be solved by the matrix exponential.*

The solution to the linear system (I.11) is given by

$$\vec{\Pi}(t) = \vec{\Pi}_0 \Phi(t) \tag{I.12}$$

where $\vec{\Pi}_0$ is the vector of initial momenta and $\Phi(t)$ is $\exp(M \cdot t)$ given by $\mathrm{diag}[I, block_2, block_2]$ where $block_2 = \exp(block_1)$ is given by

$$\begin{bmatrix} \frac{\mathrm{A}_1\,\mathrm{v}_1 + \mathrm{A}_2\,\mathrm{v}_2\,e^{-K \cdot t} + \mathrm{A}_3\,\mathrm{v}_3\,e^{-K \cdot t}}{2 \cdot K} & -\frac{\mathrm{A}_2\,\mathrm{v}_1\,\left(e^{-K \cdot t} - 1\right)}{2 \cdot K} & -\frac{\mathrm{A}_3\,\mathrm{v}_1\,\left(e^{-K \cdot t} - 1\right)}{2 \cdot K} \\ -\frac{\mathrm{A}_1\,\mathrm{v}_2\,\left(e^{-K \cdot t} - 1\right)}{2 \cdot K} & \frac{\mathrm{A}_2\,\mathrm{v}_2 + \mathrm{A}_1\,\mathrm{v}_1\,e^{-K \cdot t} + \mathrm{A}_3\,\mathrm{v}_3\,e^{-K \cdot t}}{2 \cdot K} & -\frac{\mathrm{A}_3\,\mathrm{v}_2\,\left(e^{-K \cdot t} - 1\right)}{2 \cdot K} \\ -\frac{\mathrm{A}_1\,\mathrm{v}_3\,\left(e^{-K \cdot t} - 1\right)}{2 \cdot K} & -\frac{\mathrm{A}_2\,\mathrm{v}_3\,\left(e^{-K \cdot t} - 1\right)}{2 \cdot K} & \frac{\mathrm{A}_3\,\mathrm{v}_3 + \mathrm{A}_1\,\mathrm{v}_1\,e^{-K \cdot t} + \mathrm{A}_2\,\mathrm{v}_2\,e^{-K \cdot t}}{2 \cdot K} \end{bmatrix} \tag{I.13}$$

such that $K = \frac{1}{2}(A_1 + A_2 + A_3)$. The 9 vector solution to equation (I.11) can be written as a 3×3 matrix $\Pi(t)$ and inverted to give the matrix $V(t)$. The matrix $V(t)$ is the transformation matrix describing the time evolution of a fixed spatial frame along a geodesic in $\mathbb{R}^3$ subject to the control inputs $A = [A_1, A_2, A_3]$. Of particular interest to us is the matrix $V(0)$ which describes the initial triad and the matrix $V(\infty) := \lim_{t\to\infty} V(t)$ which describes the asymptotic orientation, if any, of the controlled spatial triad.

Assumption I.1.2 *It can be shown that the matrix* $V(0)$ *takes the form* $\begin{pmatrix} \star & \star & \star \\ \star & \star & \star \\ 0 & 0 & \star \end{pmatrix}$ *when* $\Pi_{23}(0) = \Pi_{31}(0) = \Pi_{32}(0) = 0$ *and* $\Pi_{21}(0), \Pi_{22}(0), \Pi_{33}(0)$ *remain free to choose so long as* $\Pi_{21}(0) \neq \Pi_{22}(0)$. *The limit* $V(\infty)$ *exists when* $K < 0$.

For initial momenta, velocities and control inputs satisfying the assumptions in I.1.2 we obtain

$$V(0) = \begin{bmatrix} -\frac{\Pi_{22}(0)}{-v_1(0)\,\Pi_{22}(0)+\Pi_{21}(0)\,v_2(0)} & \frac{v_2(0)}{-v_1(0)\,\Pi_{22}(0)+\Pi_{21}(0)\,v_2(0)} & \frac{v_3(0)\,\Pi_{22}(0)}{-v_1(0)\,\Pi_{22}(0)\,\Pi_{33}(0)+\Pi_{21}(0)\,v_2(0)\,\Pi_{33}(0)} \\ \frac{\Pi_{21}(0)}{-v_1(0)\,\Pi_{22}(0)+\Pi_{21}(0)\,v_2(0)} & -\frac{v_1(0)}{-v_1(0)\,\Pi_{22}(0)+\Pi_{21}(0)\,v_2(0)} & -\frac{v_3(0)\,\Pi_{21}(0)}{-v_1(0)\,\Pi_{22}(0)\,\Pi_{33}(0)+\Pi_{21}(0)\,v_2(0)\,\Pi_{33}(0)} \\ 0 & 0 & \Pi_{33}(0)^{-1} \end{bmatrix}$$

$$V(\infty) = \begin{bmatrix} \frac{A_1}{v_2(0)\,A_2+A_1\,v_1(0)+v_3(0)\,A_3} & 0 & 0 \\ \frac{A_2}{v_2(0)\,A_2+A_1\,v_1(0)+v_3(0)\,A_3} & 0 & 0 \\ \frac{A_3}{v_2(0)\,A_2+A_1\,v_1(0)+v_3(0)\,A_3} & 0 & 0 \end{bmatrix}$$

It follows from the form of $V(\infty)$ that, independent of the initial configuration, the x–leg of the spatial triad asymptotically approaches the vector

$$\left[\frac{A_1}{v_1(0)A_1 + v_2(0)A_2 + v_3(0)A_3}, \frac{A_2}{v_1(0)A_1 + v_2(0)A_2 + v_3(0)A_3}, \frac{A_3}{v_1(0)A_1 + v_2(0)A_2 + v_3(0)A_3}\right]$$

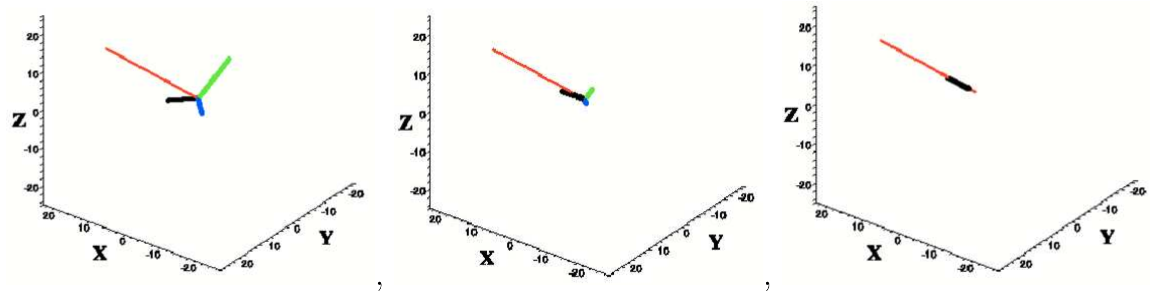
while the y–leg and z–leg tend to $[0,0,0]$. Given an initial velocity $v(0) = [v_1(0), v_2(0), v_3(0)]$ a natural question is to ask for the control inputs that asymptotically align the x–leg of the spatial frame with a desired direction $dir = [x_d, y_d, z_d]$. That is, given the initial velocity $v(0)$ and desired direction dir can the system of equations

$$
\begin{aligned}
\frac{A_1}{v_1(0)A_1 + v_2(0)A_2 + v_3(0)A_3} &= x_d \\
\frac{A_2}{v_1(0)A_1 + v_2(0)A_2 + v_3(0)A_3} &= y_d \\
\frac{A_3}{v_1(0)A_1 + v_2(0)A_2 + v_3(0)A_3} &= z_d
\end{aligned}
$$

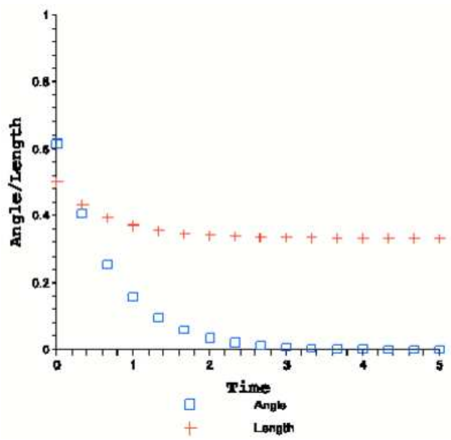
be solved for the control inputs A? For example, when *dir* is the geodesic direction $[v_1(0), v_2(0), v_3(0)]$ the control inputs are $A = [A_2, A_2, -A_2]$ where $A_2 < 0$, see Figure I.1. As another example, for $v = [1, 1, 0]$ and $A = [-1, -1, A_3]$ then the x–leg approaches $[1/2, 1/2, -A_3/2]$ which is a vector aligned with the geodesic x,y direction but elevated (for negative A_3) or de-elevated (for positive A_3) by the angle

$$\theta = \arccos\left(\frac{\sqrt{2}}{\sqrt{2 + (A_3)^2}}\right).$$

In this case it is clear that for a desired θ_d one can find the necessary input A_3. For example, for $\theta_d = 1.5$ (nearly vertical), $A_3 = \pm 19.94$ stretches and drives the x–leg to line up with the geodesic x,y direction with the z–component nearly vertical.



(a) Moving triad at t=0, t=0.9, and t=3



(b) Angle between x–leg of moving triad (black) and geodesic direction (red); Length of x–leg of moving triad (blue)

Figure I.1: Control of moving frame along the geodesic using the control input vector $A = [-1, -1, 1]$.

I.2 Constrained Rank 2 Dynamics with Scalar Potentials

The goal for this subsection is to briefly explore the affect various scalar potentials have on control of a nonholonomic constrained particle.

Recall from Section 4.3.2 that the second order dynamics of a particle of mass $m = 1$ moving in three space [x,y,z] with kinetic energy $\frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$ subject to the nonholonomic constraint $\dot{z} = y\dot{x}$ are

$$\ddot{x} = \frac{-y\dot{x}\dot{y}}{1+y^2} \quad \text{and} \quad \ddot{y} = 0.$$

These equations in first order form can be written as

$$\begin{aligned} \dot{v} &= \frac{-yvw}{1+y^2} \\ \dot{w} &= 0 \\ \dot{x} &= v \\ \dot{y} &= w \\ \dot{z} &= yv. \end{aligned} \tag{I.14}$$

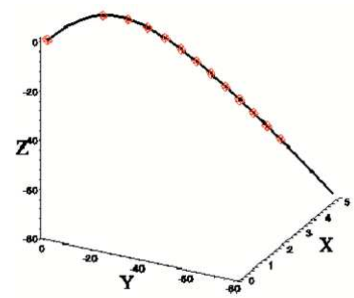
In Section 4.3.2 these equations were written equivalently (relative to a moving basis adapted to the constraint distribution) in first order form as

$$\begin{aligned} \dot{p}_1 &= \frac{yp_1p_2}{1+y^2} \\ \dot{p}_2 &= 0 \\ \dot{x} &= \frac{p_1}{1+y^2} \\ \dot{y} &= p_2 \\ \dot{z} &= \frac{yp_1}{1+y^2} \end{aligned} \tag{I.15}$$

where $p_1 = \dot{x}(1+y^2)$ and $p_2 = \dot{y}$. For initial conditions $IC_1 = [x_0, y_0, z_0, v_0, w_0] = [0, 0, 0, 1, -1]$ for the first system and $IC_2 = [x_0, y_0, z_0, p_{10}, p_{20}] = [x_0, y_0, z_0, v_0(1+y_0^2), w_0]$ for the second system, a typical solution is seen in Figure I.2.

As indicated in Section 4.3.2, the n–symplectic theory allows for the introduction of scalar potential forces in the generalized momenta dynamics.

Assumption I.2.1 *(Quadratic Scalar Potential) We focus specifically on a scalar potential of the form* $C^{33}(x,y,z,t) = C_x(x-x_{ref}(t))^2 + C_y(y-y_{ref}(t))^2 + C_z(z-z_{ref}(t))^2$ *where* $C_x, C_y,$ *and* C_z *are constants and* $[x_{ref}(t), y_{ref}(t), z_{ref}(t)]$ *is a possibly moving*



(a) Black solid line solution to I.14 and red points solution to I.15

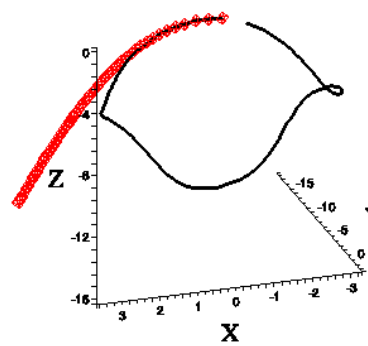
Figure I.2: Typical trajectory of particle whose z–velocity is nonholonomically constrained by $\dot{z} = y\dot{x}$

reference trajectory. That is, C^{33} is assumed to be a quadratic potential centered at a moving reference point.

The new controlled dynamics for the nonholonomic particle can now be written as

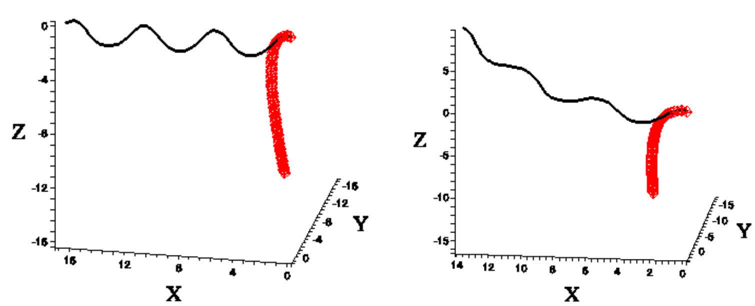
$$\begin{aligned}
\dot{p}_1 &= \frac{yp_1p_2}{1+y^2} + C_x(x - x_{ref}) + C_z y(z - z_{ref}) \\
\dot{p}_2 &= C_y(y - y_{ref}) \\
\dot{x} &= \frac{p_1}{1+y^2} \\
\dot{y} &= p_2 \\
\dot{z} &= \frac{yp_1}{1+y^2}.
\end{aligned} \tag{I.16}$$

In Figure I.3–I.5 some numerical simulations are shown which illustrate the potential force affects on the constrained particle dynamics. It is a matter of future work to put these ideas together to a general control logic which can move the nonholonomic particle to any point in space from any starting point.



(a) The black curve is the controlled trajectory while the red curve is the uncontrolled trajectory.

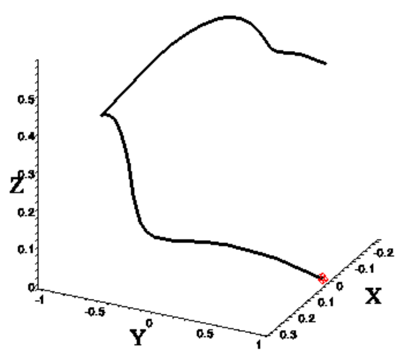
Figure I.3: The particles motion towards negative infinity can be arrested and we can return the particle to its starting position. The necessary input data is $C_x = -0.1, C_y = -0.1, C_z = 0$, $x_{ref} = y_{ref} = z_{ref} = 0$, IC: $[x_0, y_0, z_0, \dot{x}_0, \dot{y}_0] = [0, 0, 0, 1, -1]$.



,

(a) The black curves are the controlled trajectories while the red curves are the uncontrolled trajectories.

Figure I.4: The particles motion towards negative infinity can be arrested and we can redirect the particle in a "perpendicular" and "elevated" direction. The necessary input data is $C_x = 0, C_y = -1, C_z = 0$, $x_{ref} = y_{ref} = z_{ref} = 0$ (left), $x_{ref} = y_{ref} = z_{ref} = 0.1t$ (right), IC: $[x_0, y_0, z_0, \dot{x}_0, \dot{y}_0] = [0, 0, 0, 1, -1]$.



(a) The black curve is the controlled trajectory while the red dot is the equilibrium.

Figure I.5: Using a cyclic reference point, the particle can be given momentum and thus move away from equilibrium. The necessary data input data is $C_x = -0.1, C_y = -0.1, C_z = 0$, $y_{ref} = z_{ref} = 0$, $x_{ref} = \sin(t)$, IC: $[x_0, y_0, z_0, \dot{x}_0, \dot{y}_0] = [0, y_0 \neq 0, 0, 0, 0]$.