LIPSCHITZ ESTIMATES FOR GEODESICS IN THE HEISENBERG GROUP

by

Robert D. Berry

B.S. in Mathematics and Physics

Utah State University, 2003

M.S. in Mathematics

Brigham Young University, 2004

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Robert D. Berry

It was defended on

April 9, 2009

and approved by

J. Manfredi, Ph. D., Professor

P. Hajlasz, Ph. D., Associate Professor

F. Beatrous, Ph. D., Professor

A. Vainchtein, Ph. D., Associate Professor

D. Boyanovsky, Ph. D., Professor

Dissertation Director: J. Manfredi, Ph. D., Professor

LIPSCHITZ ESTIMATES FOR GEODESICS IN THE HEISENBERG GROUP

Robert D. Berry, PhD

University of Pittsburgh, 2009

In many modern approaches to solving Monge's mass transport problem (that is, optimal transport with respect to linear costs) in various metric spaces, one attempts to reduce the problem to one dimension by decomposing the measures along so-called transport (geodesic) rays. Certain key Lipschitz estimates on geodesics are needed in order provide such a decomposition. Herein these estimates for the (three dimensional, sub-Riemannian) Heisenberg Group are provided as a step towards solving Monge's problem in this metric space.

Keywords: Monge-Kantorovich, optimal mass transportation, Heisenberg group, Carnot group, horizontal curve, subRiemannian geodesics.

TABLE OF CONTENTS

LIST OF FIGURES

PREFACE

This dissertation is dedicated to my loving wife Sarah and charming daughter Olivia. Their love and support have blessed and strengthened me. I also extend my gratitude to my family, my parents Dan R. and Kathleen S. Berry, and my parents-in-law Phillip F. and Sallie K. Drinkwater. Each of them has extended a good deal of support and love to my little family in our difficult years as students.

My thanks and appreciation to Juan Manfredi for his support and the time he has invested with me. His expertise has been invaluable in showing me how to find success in this profession. I am grateful for his patient tutelage during a time when he filled many other important roles in the Department and College.

I thank the members of my committee for their kind suggestions and instruction which have improved this work.

Thank you to the many friends we have made, in school and church, along the way.

I also thank the Mathematics Department at Princeton University for hosting me for a period over which much of this work was completed.

1.0 INTRODUCTION

Gaspard Monge is credited with first posing the optimal mass transport problem in his 1781 memoir. He asks: in what way should one utilize a given embankment of earth to fill in a certain ditch in order to minimize the total distance of transporting earth? This problem is difficult enough that it remained unsolved in $\mathbb{R}^n$ until the independent papers [3, 4, 6], published in 1999-2002.

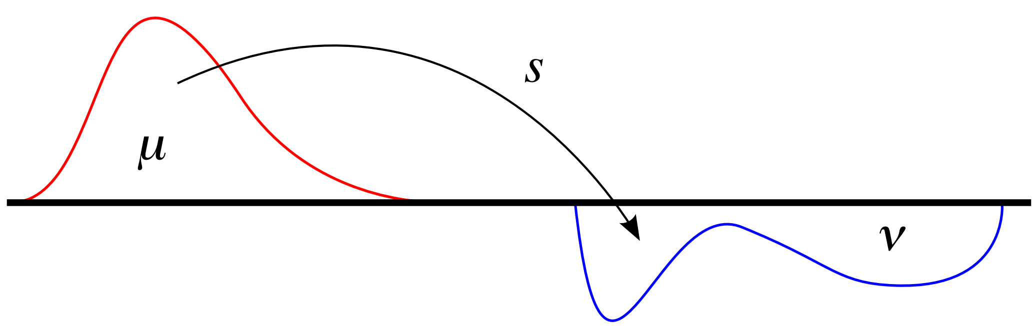

Figure 1: Monge's Problem

The study of Monge's problem lends insight into the study of several partial differential equations. If the mapping $s(x)$ pushes a source measure μ with $d\mu = f(x)\,dx$ into a destination measure μ with $d\mu = g(x)\,dx$, and is realized as the gradient of some potential function $s(x) = \nabla u(x)$, then u must satisfy the Monge-Ampère equation:

$$\det D^2 u(x) = \frac{f(x)}{g(\nabla u(x))} = F(x, u(x), \nabla u(x)). \tag{1.0.1}$$

See, e.g., [7, 8]. The optimal mapping s can be constructed by considering the sequence of potential functions u_p satisfying the $p-$Laplace equation:

$$-\text{div}(|Du_p|^{p-2}Du_p(x)) = f(x) - g(x) \tag{1.0.2}$$

in the limit as $p \to \infty$. In this construction, one can show $u_p \to u$ where the latter satisfies

$$|Du| \leq 1, -\text{div}(aDu_p(x)) = f(x) - g(x) \tag{1.0.3}$$

for $a \geq 0$, termed the transport density. See, e.g., [4]. Also, [3, 5] discuss this differential equations setting.

1.1 MONGE'S PROBLEM IN ONE DIMENSION

Monge's problem is easily solved in one dimension. Consider the following example: find a function which transports a mass uniformly distributed on the interval $[0, 2]$ into a mass uniformly distributed on $[10, 12]$. One can show this problem has a solution, but the solution is not unique; indeed, each of the functions

$$p(x) = x + 10,$$
$$q(x) = 12 - x$$

are optimal.

For simplicity, let us consider probability distributions (total mass 1) so that we need not be concerned with matching total mass in initial and goal distributions. Let f and g be probability distribution functions on $\mathbb{R}$ with cumulative distribution functions F, G respectively. Furthermore, suppose that f is supported on $[a, b]$ and g is supported on $[c, d]$ where $a < b < c < d$. One can show (under proper assumptions) that the mappings given by

$$p(x) = G^{-1}(F(x)),$$
$$q(x) = G^{-1}(1 - F(x))$$

are optimal in pushing f forward into g. Each represents a monotonic translation of the mass represented by f.

Two questions are vital to transportation: (1) in which direction and (2) how far? In this simple example there is no question about the direction the mass given by f must travel: it must go to the right. The content of this example then addresses the magnitude of displacement.

In other metric spaces the issue of direction is not trivial. The next two sections will discuss these issues in $\mathbb{R}^n$ and in Riemannian manifolds, respectively. The content of this dissertation is to extend these ideas and results to the subRiemannian Heisenberg group.

1.2 MONGE'S PROBLEM IN N-DIMENSIONS

Even in $\mathbb{R}^2$ one is forced to consider both aspects of mass transfer: direction and magnitude. Sudakov showed that in $\mathbb{R}^2$, the mass transfer occurs along straight lines connecting points in the support of f and the support of g. As such, he was able to build an optimal mapping by decomposing the measures on this family of lines and executing the 1-dimensional transfer problem along each.

In $\mathbb{R}^n$ for $n > 2$ another technicality arises: the direction of these transport lines (or rays) must vary Lipschitz continuously (this is automatically true in $\mathbb{R}^2$) in order to decompose the mass along the family of transport rays. This Lipschitz control is obtained in [3, 4, 6], published in 1999-2002.

In [4] a differential equations approach is taken, where the transportation distance is obtained via an ODE along the transport rays.

1.3 MONGE'S PROBLEM IN RIEMANNIAN MANIFOLDS

In [5], the authors approach Monge's problem on Riemannian manifolds in a method analogous to the paper [3] about mass transportation on $\mathbb{R}^n$. The authors rely on the smoothness of the distance function, the Hopf-Rinow theorem, and the existence of convex geodetic balls to bring the key Lipschitz controls from $\mathbb{R}^n$ to the manifold, giving a Lipschitz control on the directions of geodesic transport rays.

1.4 MASS TRANSFER IN THE HEISENBERG GROUP

Parts of the approach to Monge's Problem in a Riemannian manifold (in [5]) must be modified in several non-trivial ways. In the Heisenberg group, the distance is not locally differentiable, and no non-trivial subset is geodetically convex. Furthermore, the Hopf-Rinow theorem doesn't apply. Since we can write the formulas for the geodesics explicitly, some of the Lipschitz estimates can be obtained by direct calculation. The key Lipschitz estimate will rely on a rigidity result about horizontal curves in the Heisenberg group. These results are necessary to the decomposition of the initial and goal measures along the geodesic transport rays.

In [1], a solution for mass transportation, optimal with respect to distance-squared, is provided. As with other metric spaces, the L^2 optimality comes by convexity arguments; while the distance (d) is not convex in this group, d^2 is. In [3, 6], optimality of the transfer map is established via (strictly) convex approximations to the distance function. Such an approach is not possible in the Heisenberg group: d^p is not convex for any $p < 2$.

2.0 MONGE'S PROBLEM

Gaspard Monge's transport problem (in the modern treatment) is to find a map pushing one (probability) distribution into another, minimizing the average distance transported. More specifically, for a complete separable metric space (M, d):

Problem 2.0.1. *Find a mapping $s : M \to M$ which minimizes the functional*

$$I[s] = \int_M d(x, s(x))d\mu(x)$$

amongst all Borel maps $s \in \mathcal{A}(\mu, \nu)$ which push the Borel measure μ forward to the Borel measure ν. All such Borel maps satisfy

$$\int_M \phi(s(x))d\mu(x) = \int_M \phi(y)d\nu(y),$$

for any continuous function $\phi \in C(M)$.

2.1 RELAXATION AND DUALITY

We will denote by Ω_μ, Ω_ν the supports of the measures μ, ν, respectively. It is not clear on the outset that $\mathcal{A}(\mu, \nu)$ is nonempty. Kantorovich posed the following problem which is a relaxation of Monge's. Let $\mathcal{B}(\mu, \nu)$ be the set of (probability) measures on $\Omega_\mu \times \Omega_\nu$ with marginals (or projections) μ and ν. In other words, if $\eta \in \mathcal{B}(\mu, \nu)$ then for any measurable $A \subset \Omega_\mu$ and $B \subset \Omega_\nu$,

$$\mu(A) = \eta(A \times \Omega_\nu) \text{ and } \nu(B) = \eta(\Omega_\mu \times B).$$

Problem 2.1.1. *(Relaxation). Find a probability measure $\eta \in \mathcal{B}(\mu,\nu)$ for which the functional*

$$J[\eta] = \int_{\Omega_\mu \times \Omega_\nu} d(x,y) d\eta(x,y).$$

is minimum.

This problem already has the advantage of minimizing over a nonempty collection; indeed, the product measure $\mu \times \nu \in \mathcal{B}(\mu,\nu)$. Furthermore, since J is linear in η and $\Omega_\mu \times \Omega_\nu$ is compact, this relaxation has at least one solution:

$$J[\sigma] = \min_{\eta \in \mathcal{B}} J[\eta].$$

Note also that if Problem 2.0.1 has a solution s then η_s, the measure pushed forward onto the graph of s, is in $\mathcal{B}(\mu,\nu)$. Thus

$$J[\zeta] = \min_{\eta \in \mathcal{B}} J[\eta] \leq \inf_{t \in \mathcal{A}} I[t] = I[s].$$

Kantorovich also noted this relaxation has a dual maximization problem as follows. Let

$$\mathrm{Lip}_1(M,d) = \big\{ u : M \to \mathbb{R} : \text{ for all } x, y \in M, |u(x) - u(y)| \leq d(x,y) \big\}.$$

Problem 2.1.2. *(Kantorovich). Maximize*

$$K[v] = \int_M v(d\mu - d\nu).$$

for $v \in Lip_1(M,d)$.

Kantorovich's optimality principle states

$$\min_{\eta \in \mathcal{B}} J[\eta] = \max_{v \in \mathrm{Lip}_1} K[v].$$

Another useful statement (which is shown in Proposition 2.1.1 to be equivalent to Problem 2.1.2) is as follows.

Problem 2.1.3. *(Dual Formulation). Let μ and ν be the measures in Problem 2.0.1. Find a pair of functions (φ_0, ψ_0) maximizing the functional*

$$\hat{K}(\varphi, \psi) = \int_{\Omega_\mu} \varphi(x) d\mu(x) + \int_{\Omega_\nu} \psi(y) d\nu(y)$$

amongst all pairs of continuous functions $(\varphi, \psi) \in \mathcal{D}(\mu, \nu)$, where each of the pairs in this class satisfy

$$\varphi(x) + \psi(y) \leq d(x, y) \text{ for all } x \in \Omega_\mu, y \in \Omega_\nu.$$

The duality $\inf I = \sup \hat{K}$, will be discussed in Proposition 2.1.2. The equivalence of Problems 2.1.2 and 2.1.3 is established by the following Proposition.

Proposition 2.1.1. (Lipschitz Maximizer). Let (M, d) be a complete separable metric space with two finite Borel measures μ, ν having compact support $\Omega_\mu, \Omega_\nu \subset M$ and the same mass (e.g. 1). Then a maximizing pair (φ_0, ψ_0) exists for Problem 2.1.3 satisfying

$$\varphi_0 = u \text{ on } \Omega_\mu, \ \psi_0 = -u \text{ on } \Omega_\nu,$$

with $u \in \mathrm{Lip}_1(M, d)$. Moreover,

$$\varphi_0(x) = \inf_{y \in \Omega_\nu} (d(x, y) - \psi_0(y)) \text{ for any } x \in \Omega_m u$$

$$\psi_0(y) = \inf_{x \in \Omega_\mu} (d(x, y) - \phi_0(x)) \text{ for any } y \in \Omega_n u.$$

In particular, Problem 2.1.2 and Problem 2.1.3 are equivalent.

Proof. See [5] or [7]. The last statement about equivalence can be seen in the calculation

$$K[u] = \hat{K}(u, -u) = \max \hat{K}(\varphi, \psi) \geq \max \hat{K}(v, -v) = \max K[v] = K[u].$$

□

The duality argument for Monge's Problem 2.0.1 is summed up in the following Theorem. It is important to note that the existence of an $s \in \mathcal{A}(\mu, \nu)$ is assumed.

For $p \in M$, let

$$R_p = \{q \in M : d(p, q) = |u(p) - u(q)|\}.$$

Proposition 2.1.2. (Duality). Fix $u \in \mathrm{Lip}_1(M, d)$ and let $s \in \mathcal{A}(\mu, \nu)$. If $u(p) \in R_p$ for $\mu-$a.e. $p \in \Omega_\mu$, then:

I. s is an optimal map in Problem 2.0.1, minimizing the functional I.

II. u is a Kantorovich potential maximizing the functional K in Problem 2.1.2.

III. The infimum in Problem 2.0.1 is equal to the supremum in Problem 2.1.2

IV. Every optimal map $\hat{s}$ and Kantorovich potential $\hat{u}$ also satisfy

$$\hat{u}(p) - \hat{u}(\hat{s}(p)) = d(p, \hat{s}(p)).$$

for $\mu-$a.e. $p \in \Omega_\mu$,

Proof. See [4] or [5].

Recall that they hypothesis $s \in \mathcal{A}(\mu, \nu)$ means

$$\int_{\Omega_\mu} (h \circ u) d\mu = \int_{\Omega_\nu} h \ d\nu \tag{2.1.1}$$

for each $h \in C(\Omega_\nu)$, and the hypothesis $u(p) \in R_p$ means

$$u(p) - u(s(p)) = d(p, s(p)) \text{ for } \mu - \text{a.e. } p \in \Omega_\mu, \tag{2.1.2}$$

It follows that

$$\begin{aligned}
\int_{\Omega_\mu} d(p, s(p)) d\mu(p) &= \int_{\Omega_\mu} [u(p) - u(s(p))] \, d\mu(p) && (2.1.3)\\
&= \int_{\Omega_\mu} u(p) \ d\mu(p) - \int_{\Omega_\nu} u(q) \ d\nu(q) && (2.1.4)\\
&= \int_{\Omega} u \ d(\mu - \nu) && (2.1.5)\\
&\leq \max_{w \in \mathrm{Lip}_1} \int_{\Omega} w \ d(\mu - \nu) && (2.1.6)\\
&= \min_{\eta \in \mathcal{B}} \int_{\Omega} d(p, q) \ d\eta && (2.1.7)\\
&\leq \inf_{r \in \mathcal{A}} \int_{\Omega_\mu} d(p, r(p)) d\mu(p). && (2.1.8)
\end{aligned}$$

Since $s \in \mathcal{A}$, we have equality. □

2.2 TRANSPORT INFRASTRUCTURE

Much of the literature addressing Monge's Problem proceeds by obtaining a Kantorovich potential u from Proposition 2.1.1 and studying the properties of mappings s satisfying (2.1.2). The final step in solving Monge's problem in some (M, d) is then to show that one of these s is in $\mathcal{A}(\mu, \nu)$.

Definition 2.2.1. For $p \in M$, the *transport ray* through p is the set

$$R_p = \{q \in M : d(p, q) = |u(p) - u(q)|\}.$$

Call T, the union of all transport rays of positive length, the *transport set.* A measurable set $E \subseteq T$ is a *transport subset* if

$$p \in E \text{ implies } R_p \subseteq E.$$

Finally, we call a function in

$$\mathcal{T}(\mu, \nu) = \{f : T \to T : f(p) \in R_p\}$$

an *optimal mapping.*

Proposition 2.1.2 then guarantees that if there is an $s \in \mathcal{A} \cap \mathcal{T}$ (a mass transport mapping which is also an optimal mapping) then s solves Monge's Problem 2.0.1 (is an optimal mass transport mapping).

The important property of mass balance on transport subsets is true in this general setting, as established in the following Lemma.

Lemma 2.2.2. *(Mass Balance). Let F be a transport subset. Then*

$$\mu(F) = \nu(F).$$

Proof. (See [4].) Suppose $E \subseteq F$ is a closed transport subset. Let $h = \chi_E$ and put

$$u_\varepsilon(p) = u(p) + \varepsilon h(z),$$
$$v_\varepsilon(q) = \min_p\{d(p,q) - u_\varepsilon(p)\}.$$

Note that

$$v(q) = \min_p\{d(p,q) - u(p)\} = -u(q).$$

Furthermore,

$$u_\varepsilon(p) + v_\varepsilon(q) \leq d(p,q)$$

and so, by Proposition 2.1.2

$$\int_{\Omega_\mu} u_\varepsilon d\mu + \int_{\Omega_\nu} v_\varepsilon d\nu \leq \int_{\Omega_\mu} u d\mu + \int_{\Omega_\nu} v d\nu.$$

Thus

$$\int_{\Omega_\mu} \chi_E d\mu = \int_{\Omega_\mu} \left(\frac{u_\varepsilon - u}{\varepsilon}\right) d\mu \leq \int_{\Omega_\nu} \left(\frac{v - v_\varepsilon}{\varepsilon}\right) d\nu.$$

CASE 1: If $q \in E \cap \Omega_\nu$,

$$\begin{aligned} v_\varepsilon(q) &= \min_p\{d(p,q) - u(p) - \varepsilon\chi_E(p)\} \\ &= -u(q) - \varepsilon \\ &= v(q) - \varepsilon\chi_E(q), \end{aligned}$$

so that

$$\frac{v(q) - v_\varepsilon(q)}{\varepsilon} = 1 = \chi_E(q).$$

CASE 2: If $q \in \Omega_\nu \setminus E$, since E is a closed transport subset,

$$\min_p\{d(p,q) + u(p) - u(q)\} = \theta > 0$$

for some constant $\theta = \theta(q)$. Therefore

$$\min_p\{d(p,q) - u(p) - \varepsilon\chi_E(p)\} = \theta - \varepsilon - u(q).$$

On the other hand,

$$\inf_{p\in E^c}\{d(p,q) - u(p) - \varepsilon\chi_E(p)\} = -u(q).$$

Thus if $\varepsilon = \varepsilon(y) > 0$ is small so that $\varepsilon < \theta$, then

$$v_\varepsilon(q) = -u(q) = v(q),$$

and so

$$\frac{v(q) - v_\varepsilon(q)}{\varepsilon} = 0 = \chi_E(q).$$

Hence

$$\lim_{\varepsilon \to 0^+} \frac{v - v_\varepsilon}{\varepsilon} = \chi_E.$$

Since $0 \leq \frac{v - v_\varepsilon}{\varepsilon} \leq 1$ a.e.,

$$\mu(E) \leq \nu(E).$$

By a symmetric argument,

$$\nu(E) \leq \mu(E),$$

so that

$$\mu(E) = \nu(E)$$

for every closed set $E \subseteq F$. Taking the supremum over all such E gives

$$\mu(F) = \nu(F).$$

□

What remains is to find an $s \in \mathcal{T}(\mu, \nu)$ which is also in $\mathcal{A}(\mu, \nu)$. The information that is missing is: how far along the transport ray should a point be mapped? Two methods in answering this question are prominent: either decompose the measures μ, ν along transport rays, reducing the problem to a single dimension (see e.g. [3], [5], [6]) or alternatively, to build an ODE along the transport rays, for which s is the time-1 mapping (see [4]). The main difference is when the idea of a "transport density" is introduced. In the former method, this density is hidden (couched in the 1-dimensional solution). In the latter method, this density is front-and-center from the beginning. These methods both rely on certain Lipschitz estimates which essentially gives a second-order estimate on the structure of the "transport infrastructure," whose elements are introduced in Definition 2.2.1.

3.0 THE HEISENBERG GROUP

We model the Heisenberg Group as $\mathbb{H} = (\mathbb{R}^3, *)$ where the multiplication of $x, a \in \mathbb{H}$ is given by

$$(x_1, x_2, x_3) * (a_1, a_2, a_3) = \left(x_1 + a_1, x_2 + a_2, x_3 + a_3 + \frac{1}{2}(x_1 a_2 - x_2 a_1)\right).$$

Note that the identity in $\mathbb{H}$ is 0, $x^{-1} = -x$, and that the multiplication is not commutative. The center of $\mathbb{H}$ is $\mathcal{Z} = \{(0, 0, z) : z \in \mathbb{R}\}$.

3.1 HORIZONTAL AND VERTICAL BUNDLES

We calculate the left invariant vector fields in $\mathbb{H}$ as those which are the left push-forward of the standard basis at the origin. Differentiating the multiplication law with respect to each of the coordinates of a at the point (x, y, z) gives:

$$\begin{aligned} a_1 &: \quad \left(1, 0, -\frac{y}{2}\right), \quad && X = \partial_x - \frac{y}{2}\partial_z \\ a_2 &: \quad \left(0, 1, \frac{x}{2}\right), \quad && Y = \partial_y + \frac{x}{2}\partial_z \\ a_3 &: \quad \left(0, 0, 1\right), \quad && Z = \partial_z. \end{aligned}$$

These vector fields are everywhere linearly independent, and the only non-zero Lie bracket is:

$$[X, Y] = X\left(\frac{x}{2}\right)\partial_z - Y\left(-\frac{y}{2}\right)\partial_z = \partial_z = Z.$$

Thus by Frobenius' theorem, the distribution $\operatorname{span}\{X, Y\}$ is not integrable. Furthermore,

$$\operatorname{span}\{X, Y, [X, Y]\}_p = T_p\mathbb{H}$$

so this distribution is said to be step 2 (everywhere).

For the dual space $T_p^*\mathbb{H}$ we take the standard basis dual to $\{X, Y, Z\}$. Specifically, let

$$\theta_1 = dx,$$
$$\theta_2 = dy,$$
$$\theta_3 = \frac{y}{2}dx - \frac{x}{2}dy + dz,$$

so that

$$\operatorname{span}\{\theta_1, \theta_2, \theta_3\}_p = T_p^*\mathbb{H}.$$

Definition 3.1.1. At a point $p \in \mathbb{H}$, the *horizontal space* $H_p\mathbb{H} \subset T_p\mathbb{H}$ is spanned by X_p and Y_p. The *vertical space* $V_p\mathbb{H} \subset T_p\mathbb{H}$ is spanned by Z_p. Furthermore, the disjoint unions $H\mathbb{H} = \bigcup_{p\in\mathbb{H}} H_p\mathbb{H}$ and $V\mathbb{H} = \bigcup_{p\in\mathbb{H}} V_p\mathbb{H}$ are the *horizontal bundle* and *vertical bundle,* respectively. Similarly, we denote $H_p^*\mathbb{H} = \operatorname{span}\{\theta_1, \theta_2\}_p$, $V_p^*\mathbb{H} = \operatorname{span}\{\theta_3\}$, $H^*\mathbb{H} = \bigcup_p H_p^*\mathbb{H}$, and $V^*\mathbb{H} = \bigcup_p V_p^*\mathbb{H}$.

Also, θ_3 is the annihilator of the horizontal space since $\theta_3(X) = \theta_3(Y) = 0$, by definition of dual basis.

3.2 HORIZONTAL CURVES

Definition 3.2.1. For an interval $I \subset \mathbb{R}$, a piecewise C^1 curve $p : I \to \mathbb{H}$ is a *horizontal curve* if $\dot{p}(t) \in H_{p(t)}\mathbb{H}$, whenever it exists.

Theorem 3.2.2 (Chow)**.** *If M is a manifold and the distribution $HM \subset TM$ is bracket-generating, then the set of points that can be connected to $p \in M$ by a horizontal curve is the connected component of M containing p. (see, e.g., [2]).*

Notation 3.2.3. Denote the canonical projection $[\cdot] : \mathbb{H} \to \mathbb{H}/\mathcal{Z}$, which is the projection into the $xy-$plane: $[(x, y, z)] = (x, y)$.

A key new ingredient, employed in the analysis of geodesics which follows, is the following lemma.

Lemma 3.2.4 (Rigidity of Horizontal Curves in $\mathbb{H}$). *Let $p, q : (-1,1) \to \mathbb{H}$ be horizontal curves. If $[p(t)] = [q(t)]$ for all $t \in (-1,1)$, then there is a constant $c \in \mathbb{R}$ such that $p_3 = q_3 + c$.*

Proof. We suppose that

$$(p_1(t), p_2(t)) = [p(t)] = [q(t)] = (q_1(t), q_2(t)).$$

These horizontal curves are differentiable. In particular,

$$p_1' = q_1' \text{ and } p_2' = q_2'.$$

Since the curves are horizontal,

$$\begin{aligned} p_3' &= \left\{ -\frac{p_2}{2}p_1' + \frac{p_1}{2}p_2' \right\} \\ q_3' &= \left\{ -\frac{q_2}{2}q_1' + \frac{q_1}{2}q_2' \right\}. \end{aligned}$$

Let $c(t) = p_3(t) - q_3(t)$. Finally, we calculate:

$$\begin{aligned} c' = p_3' - q_3' &= \left\{ -\frac{p_2}{2}p_1' + \frac{p_1}{2}p_2' \right\} - \left\{ -\frac{q_2}{2}q_1' + \frac{q_1}{2}q_2' \right\} \\ &= \frac{1}{2}\Big((q_2 - p_2)p_1' + (p_1 - q_1)p_2' \Big) \equiv 0. \end{aligned}$$

Therefore there exists a constant $c \in \mathbb{R}$ such that $p_3 = q_3 + c$. □

In a Riemannian (or Euclidean) setting, this point is trivial because the center in the tangent space is trivial.

4.0 GEODESICS IN $\mathbb{H}$

In order to introduce geodesics in $\mathbb{H}$, we first discuss a canonical distribution in $T(T^*\mathbb{H})$ (a connection) which gives rise to parallel covectors. Then we introduce a subRiemannian metric and Hamiltonian vector fields, giving rise to geodesic equations.

Let $x = (x_i)$ give coordinates in $\mathbb{H}$ and $\lambda = (\lambda_i)$ give coordinates in $T_p^*\mathbb{H}$ with respect to the basis θ_i. Let $X_1 = X, X_2 = Y, X_3 = Z$.

4.1 CONNECTION AND PARALLEL TRANSPORT OF COVECTORS

Notation 4.1.1. Let $\pi : T^*\mathbb{H} \to \mathbb{H}$ be the canonical projection so that $\pi_* : T(T^*\mathbb{H}) \to T\mathbb{H}$. Let $\mathcal{V} = \ker(\pi_*) \subset T(T^*\mathbb{H})$, a smooth distribution.

In coordinates, we write $\mathcal{V} = \text{span}\{\partial_{\lambda_i}\}$

Definition 4.1.2. The *canonical 1-form* $\eta \in T^*(T^*\mathbb{H})$ is

$$\eta_\alpha(h) = \alpha(\pi_* h)$$

where $\alpha \in T^*\mathbb{H}$, $h_\alpha \in T(T^*\mathbb{H})$.

So, for $h = \sum(h_i X_i + v_i \partial_{\lambda_i})$ at $\alpha = \sum \lambda_i(\alpha)\theta_i$,

$$\begin{aligned}\eta_\alpha(h) = \alpha(\pi_* h) &= \sum_i \lambda_i(\alpha)\theta_i\left(\sum_j h_j X_j\right)\\ &= \sum_i \lambda_i(\alpha) h_i.\end{aligned}$$

Thus

$$\eta_\alpha = \sum_i \lambda_i(\alpha)\theta_i$$

where the θ_i are thought of as 1-forms on $T(T^*\mathbb{H})$.

Definition 4.1.3. The *canonical symplectic form* is $\omega = d\eta$.

In coordinates this is:

$$\begin{aligned}
\omega = d\eta &= d\big(\sum_i \lambda_i\theta_i\big) \\
&= \sum_i \big(d\lambda_i \wedge \theta_i + \lambda_i d\theta_i\big) \\
&= \sum_i (d\lambda_i \wedge \theta_i) - \lambda_3(\theta_1 \wedge \theta_2)
\end{aligned}$$

since

$$\begin{aligned}
d\theta_1 &= ddx = 0, \\
d\theta_2 &= ddy = 0, \\
d\theta_3 &= d\big(\frac{y}{2}dx - \frac{x}{2}dy + dz\big) = -dx \wedge dy = -\theta_1 \wedge \theta_2.
\end{aligned}$$

Definition 4.1.4. A *connection* $\mathcal{H}$ is a smooth distribution $\mathcal{H}_\alpha \subseteq T_\alpha(T^*\mathbb{H})$ such that $\mathcal{H}_\alpha \cap \mathcal{V}_\alpha = 0$ and $\mathcal{V} \oplus \mathcal{H} = T(T^*\mathbb{H})$. The *canonical connection* satisfies: $h_\alpha \in \mathcal{H}_\alpha$ implies $\tilde{\omega}_\alpha(h, \cdot) \equiv 0$ where $\tilde{\omega}$ is the restriction of the canonical symplectic form ω to $T^*\mathbb{H}$.

In coordinates, let $h_\alpha = h_1X + h_2Y + h_3Z + \beta_1\partial_{\lambda_1} + \beta_2\partial_{\lambda_2} + \beta_3\partial_{\lambda_3} \in T_\alpha(T^*\mathbb{H})$, $\alpha \in T_p^*\mathbb{H}$. Then

$$\begin{aligned}
\omega(h, \cdot) &= \Big(\sum_i (d\lambda_i \wedge \theta_i) - \lambda_3(\theta_1 \wedge \theta_2)\Big)(h, \cdot) \\
&= \beta_1\theta_1 - h_1 d\lambda_1 + \beta_2\theta_2 - h_2 d\lambda_2 + \beta_3\theta_3 - h_3 d\lambda_3 - \lambda_3(h_1\theta_2 - h_2\theta_1) \\
&= (\beta_1 + \lambda_3 h_2)\theta_1 + (\beta_2 - \lambda_3 h_1)\theta_2 + \beta_3\theta_3 - h_1 d\lambda_1 - h_2 d\lambda_2 - h_3 d\lambda_3 \qquad (4.1.1)
\end{aligned}$$

so that

$$\tilde{\omega}(h, \cdot) = (\beta_1 + \lambda_3 h_2)\theta_1 + (\beta_2 - \lambda_3 h_1)\theta_2 + \beta_3\theta_3. \qquad (4.1.2)$$

Since the θ_i are linearly independent, $\tilde{\omega}(h, \cdot) \equiv 0$ implies

$$\begin{aligned} \beta_1 &= -\lambda_3 h_2 \\ \beta_2 &= \lambda_3 h_1 \\ \beta_3 &= 0. \end{aligned} \tag{4.1.3}$$

Therefore, $h \in \mathcal{H}$ implies

$$\begin{aligned} h &= h_1 X + h_2 Y + h_3 Z - \lambda_3 h_2 \partial_{\lambda_1} + \lambda_3 h_1 \partial_{\lambda_2} \\ &= h_1(X + \lambda_3 \partial_{\lambda_2}) + h_2(Y - \lambda_3 \partial_{\lambda_1}) + h_3 Z. \end{aligned}$$

We observe the vector field basis for $\mathcal{H}$:

$$\mathcal{H} = \operatorname{span}\{X + \lambda_3 \partial_{\lambda_2}, Y - \lambda_3 \partial_{\lambda_1}, Z\}.$$

Remark 4.1.5. Note that $T^*\mathbb{H}$ is the 2-step free group whose Lie algebra is generated by the three vectors $\{X + \lambda_3 \partial_{\lambda_2}, Y - \lambda_3 \partial_{\lambda_1}, \partial_{\lambda_3}\}$. Explicitly:

$$\begin{aligned} [X + \lambda_3 \partial_{\lambda_2}, Y - \lambda_3 \partial_{\lambda_1}] &= Z, \\ [X + \lambda_3 \partial_{\lambda_2}, \partial_{\lambda_3}] &= -\partial_{\lambda_2}, \\ [Y - \lambda_3 \partial_{\lambda_1}, \partial_{\lambda_3}] &= \partial_{\lambda_1}, \end{aligned}$$

and all other brackets are zero.

The spaces $\mathcal{H}_{(\alpha_p)}$ and $\mathcal{V}_{(\alpha_p)}$ may be identified in a canonical way with $T_p\mathbb{H}$ and $T_p^*\mathbb{H}$ respectively via the linear mappings

$$\pi_* : \mathcal{H}_{(\alpha_p)} \to T_p\mathbb{H}, \qquad K : \mathcal{V}_{(\alpha_p)} \to T_p^*\mathbb{H}$$

where $K(\partial_{\lambda_i}) = \theta_i$ for $i = 1, 2, 3$.

Definition 4.1.6. A curve in the cotangent bundle $\alpha_p : I \to T^*\mathbb{H}$ is *parallel* along the curve $\pi \circ \alpha_p = p : I \to \mathbb{H}$ if

$$\partial_t(\alpha_p) \in \mathcal{H}_{(\alpha_p)}.$$

4.2 SUBRIEMANNIAN METRIC AND HAMILTONIAN FORMALISM

We endow $\mathbb{H}$ with a subRiemannian metric, defined on the horizontal spaces, so that X and Y are everywhere orthonormal. Let $\|\cdot\|$ be the norm induced by this inner product. Let $\beta : I \to \mathbb{H}$ be a horizontal curve. Then $\|\dot{\beta}(t)\|$ is defined, and let

$$\text{length}(\beta) = \int_I \|\dot{\beta}(t)\| \, dt.$$

The Carnot-Carathéodory distance between two points $p, q \in \mathbb{H}$ is:

$$d_c(p,q) = \min\{\text{length}(\alpha) : \alpha \text{ connects } p \text{ and } q, \text{ and is horizontal}\}.$$

In the minimization we will not only have a Hamiltonian system, but we must also respect the annihilator of the horizontal space. As such, geodesics may be of two types: Normal or Abnormal.

Definition 4.2.1. A vector field $\alpha \mapsto h_\alpha \in T(T^*\mathbb{H})$ is a *Hamiltonian vector field* for the function $H \in C^1(T^*\mathbb{H})$ (called the *Hamiltonian*) if it satisfies:

$$dH(Y) = -\omega(h, Y)$$

for all Y.

Definition 4.2.2. A *normal curve* is a horizontal curve which is the projection (into $\mathbb{H}$) of a curve $(p, \alpha) : I \to T^*\mathbb{H}$ whose derivative is a Hamilton vector field with respect to the Hamiltonian

$$H(x, \lambda) = \frac{1}{2}(\lambda_1^2 + \lambda_2^2). \tag{4.2.1}$$

That is,

$$dH_{\alpha_p}(Y) = -\omega\big((\dot{p}, \dot{\alpha}), Y\big) \tag{4.2.2}$$

for all $Y \in T_{\alpha_p}(T^*\mathbb{H})$.

Definition 4.2.3. An *abnormal curve* is a horizontal curve which is the projection (into $\mathbb{H}$) of a curve $(p, \alpha) : I \to V^*\mathbb{H}$ which does not intersect the zero section, and whose derivative, whenever it exists, is the kernel of the canonical symplectic form restricted to $V^*\mathbb{H}$. That is,

$$\omega\big((\dot{p}, \dot{\alpha}), Y\big) = 0$$

for all $Y \in T_{\alpha_p}(V^*\mathbb{H})$.

For either type of extremal, we must calculate $\omega\big((\dot{p}, \dot{\alpha}), \cdot\big)$. We proceed then by letting p be a horizontal curve, and (p, α) be a lifting of that curve to a curve in $T^*\mathbb{H}$. Write

$$\begin{aligned} \dot{p} &= h_1 X + h_2 Y \text{ and} \\ \dot{\alpha} &= \sum_i \dot{\alpha}_i \partial_{\lambda_i}. \end{aligned}$$

Then from (4.1.1)

$$\omega_{(p,\alpha)}\big((\dot{p}, \dot{\alpha}), \cdot\big) = \big(\dot{\alpha}_1 + \alpha_3 h_2\big)\theta_1 + \big(\dot{\alpha}_2 - \alpha_3 h_1\big)\theta_2 + \dot{\alpha}_3\theta_3 - \sum_{j=1}^{2} h_j d\lambda_j$$

Consider now two cases:

Case 1. Suppose that (p, α) projects to an abnormal curve. Then $\alpha_1 = \alpha_2 = 0$ and then the restriction of ω to $V^*\mathbb{H}$ eliminates the $d\lambda_1$ and $d\lambda_2$ terms. Then we would have

$$\omega_{(p,\alpha)}\big((\dot{p}, \dot{\alpha}), \cdot\big) = \alpha_3 h_2 \theta_1 - \alpha_3 h_1 \theta_2 + \dot{\alpha}_3 \theta_3 = 0.$$

Since the θ_i are a basis, we have:

$$\begin{aligned} \alpha_3 h_2 &= 0, \\ \alpha_3 h_1 &= 0, \\ \dot{\alpha}_3 &= 0, \text{ and} \\ \alpha_3 &\neq 0. \end{aligned}$$

There are no non-trivial solutions to this system. Thus in $\mathbb{H}$ there are no abnormal curves.

Case 2. Suppose that (p, α) projects to an normal curve. Now we analyze the Hamiltonian given by (4.2.1). Since

$$\begin{aligned} dH_{\alpha_p} &= \big(\lambda_1 d\lambda_1 + \lambda_2 d\lambda_2\big)_{\alpha_p} \\ &= \alpha_1 d\lambda_1 + \alpha_2 d\lambda_2, \end{aligned}$$

$(\dot{p}, \dot{\alpha})$ is a Hamiltonian vector field with respect to H if (4.2.2) holds. Substituting from the foregoing calculations:

$$\begin{aligned} dH_{\alpha_p} &= -\omega\big((\dot{p}, \dot{\alpha}), \cdot\big) \\ \alpha_1 d\lambda_1 + \alpha_2 d\lambda_2 &= -\big(\dot{\alpha}_1 + \alpha_3 h_2\big)\theta_1 - \big(\dot{\alpha}_2 - \alpha_3 h_1\big)\theta_2 - \dot{\alpha}_3\theta_3 + \sum_{j=1}^{2} h_j d\lambda_j \end{aligned}$$

or

$$\big(\dot{\alpha}_1 + \alpha_3 h_2\big)\theta_1 + \big(\dot{\alpha}_2 - \alpha_3 h_1\big)\theta_2 + \dot{\alpha}_3\theta_3 + \big(\alpha_1 - h_1\big)d\lambda_1 + \big(\alpha_2 - h_2\big)d\lambda_2 = 0.$$

Since $\{\theta_1, \theta_2, \theta_3, d\lambda_1, d\lambda_2\}$ are linearly independent, we have

$$\begin{aligned} \alpha_1 &= h_1, \\ \alpha_2 &= h_2, \\ \dot{\alpha}_1 &= -\alpha_3 h_2, \\ \dot{\alpha}_2 &= \alpha_3 h_1, \\ \dot{\alpha}_3 &= 0. \end{aligned}$$

Thus α_p is parallel, and the Hamiltonian contributes the duality between tangent and cotangent (horizontal) vectors. Solutions to this system for $p(t) = \big(x(t), y(t), z(t)\big)$, with initial conditions $x(0) = 0, y(0) = 0, z(0) = 0, \alpha_1(0) = v_x, \alpha_2(0) = v_y, \alpha_3(0) = a$ are:

$$\begin{aligned}
&\text{for } a = 0: && x(t) = v_x t,\\
& && y(t) = v_y t,\\
& && z(t) = 0;\\
&\text{and for } a \neq 0: && x(t) = \frac{v_x \sin at - v_y(1-\cos at)}{a},\\
& && y(t) = \frac{v_y \sin at + v_x(1-\cos at)}{a},\\
& && z(t) = \frac{v_x^2 + v_y^2}{2a^2}(at - \sin at).
\end{aligned}$$

It is sufficient to consider normal curves through the identity by left invariance. We will see these normal curves are geodesics if $0 \leq at \leq 2\pi$.

4.3 GEODETIC EXPONENTIAL FUNCTION

Definition 4.3.1. We may define a geodetic *exponential function* $\exp : T^*\mathbb{H} \to \mathbb{H}$ as:

$$\exp(v_x\theta_1 + v_y\theta_2 + a\theta_3)_p) = p * (x(1), y(1), z(1)).$$

We shall see that this function is bijective on $(\mathbb{R}^2 \setminus 0) \times (-2\pi, 2\pi)$.

Remark 4.3.2. Note that this exponential function is defined on the cotangent bundle of $\mathbb{H}$, as opposed to the Riemannian geodesic exponential which is often defined on the tangent bundle of a manifold. A Riemannian metric naturally gives rise to a bijection between the tangent and cotangent bundles. Via this bijection, the Riemannian exponential function may equivalently be defined on the manifold's cotangent bundle. Indeed, when developing the connection (as above) we see the exponential function is most naturally defined on the cotangent bundle. Similarly, from the duality of horizontal vectors and covectors coming from the subRiemannian metric, we may equivalently consider the domain of this exponential function on $\mathbb{H}$ to be $H\mathbb{H} \times V^*\mathbb{H}$, or by the vector space duality for the basis in $V^*\mathbb{H}$, the domain may be considered to be $T\mathbb{H}$. This may be more familiar, but is not natural in the subRiemannian setting; this consideration relies on both the (subRiemannian) metric duality on the horizontal spaces and vector space duality in the vertical spaces.

Dynamic Interpretation. This definition is equivalent to the exponential given by Ambrosio and Rigot. However, while they give a "spherical coordinates" interpretation, we interpret the third parameter a as the initial horizontal acceleration of the geodesic. This is indeed natural since the acceleration

$$\ddot{p} = \dot{h}_1 X + \dot{h}_2 Y = -ah_2 X + ah_1 Y$$

is determined (beyond velocity) by a, and for a unit speed curve,

$$\|\ddot{p}\| = |a|.$$

Vectors are parallel transported along a curve if they are horizontal in $T(T^*\mathbb{H})$. Transportation of a vector $v_x X + v_y Y + v_z Z$ along a curve parameterized as $p : [0, T] \to \mathbb{H}$, with coordinate representation $p(t) = (x(t), y(t), z(t)$ and tangent vector $\dot{p}(t) = h_x(t)X + h_y(t)Y + h_z(t)Z$, satisfies

$$\begin{aligned} \dot{v}_x &= -v_z h_y, \\ \dot{v}_y &= v_z h_x, \\ \dot{v}_z &= 0. \end{aligned}$$

Thus,

$$V(t) = \{-v_z(y(t) - y(0)) + v_x\}X + \{v_z(x(t) - x(0)) + v_y\}Y + v_z Z.$$

In order to ensure that these normal curves are geodesics, we investigate the Jacobi Fields.

Definition 4.3.3. Given a 1-parameter family of normal curves $p_b(t)$, a *Jacobi Field* along $p_0(t)$ is:

$$J(t) = \left.\frac{\partial p_b(t)}{\partial b}\right|_{b=0}.$$

Definition 4.3.4. Points $p, q \in \mathbb{H}$ are *conjugate* if along a normal curve connecting them, there exists a non-zero Jacobi field which vanishes at p and q.

In order for a normal curve to be a geodesic, its interior must not contain two points which are conjugate.

Let $J(t) \in T_{p_0(t)}\mathbb{H}$, $J(t) = J_1(t)X + J_2(t)Y$ be a Jacobi field. Then:

$$\dot{J}(t) = \left.\frac{\partial \dot{p}_b(t)}{\partial b}\right|_{b=0},$$

so that

$$\dot{J}_i(t) = \left.\frac{\partial \alpha_i(t)}{\partial b}\right|_{b=0},$$
$$\ddot{J}_i(t) = \left.\frac{\partial \dot{\alpha}_i(t)}{\partial b}\right|_{b=0}.$$

Thus:

$$\ddot{J}_1 = \partial_b \dot{\alpha}_1 = \partial_b(-\alpha_2\alpha_3) = -\dot{J}_2\alpha_3 - \alpha_2\dot{J}_3,$$
$$\ddot{J}_2 = \partial_b \dot{\alpha}_2 = \partial_b(\alpha_1\alpha_3) = \dot{J}_1\alpha_3 - \alpha_1\dot{J}_3,$$
$$\ddot{J}_3 = \partial_b \dot{\alpha}_3 = 0.$$

Along a normal curve $p : [0,1] \to \mathbb{H}$ with acceleration parameter $\alpha_3 = a$, consider a Jacobi field with initial conditions $J(0) = 0$, $\dot{J}_1(0) = \phi_1$, $\dot{J}_2(0) = \phi_2$, $\dot{J}_3(0) = 0$. Then:

$$\ddot{J}_1 = -a\dot{J}_2,$$
$$\ddot{J}_2 = a\dot{J}_1,$$
$$\ddot{J}_3 = 0,$$

so that $\dot{J}(t)$ is parallel-transported along $p(t)$. Furthermore, $J(t)$ shares the periodicity of $[p(t)]$ (in fact, they satisfy the same differential equations). Thus each point in $p((0,1]) \cap [p(0)]$ are conjugate with $p(0)$. Therefore we have established the following characterization of geodesics in $\mathbb{H}$.

Lemma 4.3.5. *A horizontal curve $p : [0,1] \to \mathbb{H}$ is a geodesic if and only if there exist $v_x, v_y \in \mathbb{R} \setminus 0$ and $a \in [-2\pi, 2\pi]$ such that $p(t) = p(0) * (x(t), y(t), z(t))$ where*

$$\begin{aligned} \text{for } a = 0\text{:} \qquad & x(t) = v_x t, \\ & y(t) = v_y t, \\ & z(t) = 0; \\ \text{and for } a \neq 0: \qquad & x(t) = \frac{v_x \sin at - v_y(1-\cos at)}{a}, \\ & y(t) = \frac{v_y \sin at + v_x(1-\cos at)}{a}, \\ & z(t) = \frac{v_x^2 + v_y^2}{2a^2}(at - \sin at). \end{aligned}$$

Furthermore, the geodetic exponential function

$$\exp((v_x, v_y, a)_q) = q * (x(1), y(1), z(1))$$

is bijective on $(\mathbb{R}^2 \setminus 0) \times (-2\pi, 2\pi) \cup \{0\} \subset \mathbb{R}^3$ for every $q \in \mathbb{H}$.

This exponential function is equivalent (outside of group multiplication convention) to one defined in [1], for $A + Bi \in \mathbb{C}$ and $w \in [-\pi/2, \pi/2]$, related via:

$$\exp_{AR}(A + Bi, w) = \exp(AX + BY + 4wZ).$$

Define $\mathcal{D}_p = \mathbb{R}^2 \times (-2\pi, 2\pi) \subset T_p\mathbb{H}$.

Lemma 4.3.6. *Let $\alpha : [0,1] \to \mathbb{H}$ be a geodesic connecting $p = \alpha(0)$ and $q = \alpha(1)$. Suppose $a, b \in \alpha([0,1])$ $(a \neq b)$ such that $[a] = [b]$. Then (as sets)*

$$\{a, b\} = \{p, q\}.$$

Proof. If $a, b \in \alpha$ and $[a] = [b]$, then a, b are conjugate. In order to satisfy the criterion for absolutely minimizing distance, a, b must be the endpoints of α. □

The Carnot-Carathéodory distance from the identity: $d_{cc}(0,\cdot):\mathbb{H}\to\mathbb{R}$ is thus differentiable everywhere but the center $Z\leq\mathbb{H}$. Staying away from Z, we apply the inverse function theorem for

$$\exp_0\{(d\cos\theta)X+(d\sin\theta)Y+aZ\}=\begin{pmatrix}\frac{d}{a}(\sin(a+\theta)-\sin\theta)\\ \frac{d}{a}(\cos(a+\theta)-\cos\theta)\\ \frac{d^2}{2a^2}(a-\sin a)\end{pmatrix}$$

to obtain the X,Y,Z derivatives of (d,θ,a). First,

$$\frac{\partial(x,y,z)}{\partial(d,\theta,a)}=\begin{pmatrix}\frac{1}{a}(\sin(a+\theta)-\sin\theta) & \frac{d}{a}(\cos(a+\theta)-\cos\theta) & \frac{d}{a^2}(a\cos(a+\theta)-\sin(a+\theta)+\sin\theta)\\ \frac{1}{a}(\cos(a+\theta)-\cos\theta) & -\frac{d}{a}(\sin(a+\theta)-\sin\theta) & \frac{d}{a^2}(-a\sin(a+\theta)+\cos(a+\theta)-\cos\theta)\\ \frac{d}{a^2}(a-\sin a) & 0 & -\frac{d^2}{2a^3}(a-2\sin a+a\cos a)\end{pmatrix}.$$

When this matrix is invertible, we have:

$$\frac{\partial(d,\theta,a)}{\partial(x,y,z)}=\begin{pmatrix}\frac{1}{2}(\cos(a+\theta)+\cos\theta) & -\frac{1}{2}(\sin(a+\theta)+\sin\theta) & \frac{a}{d}\\ * & * & \frac{a^2}{d^2}\frac{\sin a}{a-2\sin a+a\cos a}\\ * & * & -\frac{2a^2}{d^2}\frac{\sin a}{a-2\sin a+a\cos a}\end{pmatrix}.$$

(NOTE: some entries are replaced with $*$ for convenience.) In particular:

$$\begin{aligned}Xd&=\partial_x d-\frac{y}{2}\partial_z d=\cos(a+\theta),\\ Yd&=\partial_y d+\frac{x}{2}\partial_z d=\sin(a+\theta),\\ Zd&=\partial_z d=\frac{a}{d}\end{aligned}\tag{4.3.1}$$

4.4 GEODESICS IN THE TANGENT BUNDLE

Recall that the spaces $\mathcal{H}_{(\alpha_p)}$ and $\mathcal{V}_{(\alpha_p)}$ may be identified in a canonical way with $T_p\mathbb{H}$ and $T_p^*\mathbb{H}$ respectively via the linear mappings

$$\pi_* : \mathcal{H}_{(\alpha_p)} \to T_p\mathbb{H}, \qquad K : \mathcal{V}_{(\alpha_p)} \to T_p^*\mathbb{H}$$

where $K(\partial_{\lambda_i}) = \theta_i$ for $i = 1, 2, 3$. The inner product on $H\mathbb{H}$ induces a subRiemannian inner product on $T(T^*\mathbb{H})$ as follows: Let $g, h \in T_{(\alpha_p)}(T^*\mathbb{H})$. Then g, h are uniquely represented as

$$g = g_{\mathcal{V}} + g_{\mathcal{H}}, \qquad h = h_{\mathcal{V}} + h_{\mathcal{H}}, \qquad \text{where} \qquad g_{\mathcal{V}}, h_{\mathcal{V}} \in \mathcal{V}_{(\alpha_p)} \qquad \text{and} \qquad g_{\mathcal{H}}, h_{\mathcal{H}} \in \mathcal{H}_{(\alpha_p)}$$

If $\pi_* g_{\mathcal{H}}$ and $\pi_* h_{\mathcal{H}}$ are in $H_p\mathbb{H}$ then we may define the inner product $\langle \cdot, \cdot \rangle_{(\alpha_p)}^{T^*\mathbb{H}}$ on $T_{(\alpha_p)}T^*\mathbb{H}$ by

$$\langle g, h \rangle_{(\alpha_p)}^{T^*\mathbb{H}} = \langle \pi_* g_{\mathcal{H}}, \pi_* h_{\mathcal{H}} \rangle_p + \ll K g_{\mathcal{V}}, K h_{\mathcal{V}} \gg_p$$

where $\ll \cdot, \cdot \gg_p$ is the induced cometric. Explicitly, let

$$\begin{aligned} g &= g_1 X + g_2 Y g_3 Z + \tilde{g}_1 \partial_{\lambda_1} + \tilde{g}_2 \partial_{\lambda_2} + \tilde{g}_3 \partial_{\lambda_3} \\ &= g_1 (X + \lambda_3 \partial_{\lambda_2}) + g_2 (Y - \lambda_3 \partial_{\lambda_1}) + g_3 Z + (\tilde{g}_1 + g_2 \lambda_3) \partial_{\lambda_1} + (\tilde{g}_2 - g_1 \lambda_3) \partial_{\lambda_2} + \tilde{g}_3 \partial_{\lambda_3}, \text{ and} \\ h &= h_1 X + h_2 Y h_3 Z + \tilde{h}_1 \partial_{\lambda_1} + \tilde{h}_2 \partial_{\lambda_2} + \tilde{h}_3 \partial_{\lambda_3} \\ &= h_1 (X + \lambda_3 \partial_{\lambda_2}) + h_2 (Y - \lambda_3 \partial_{\lambda_1}) + h_3 Z + (\tilde{h}_1 + h_2 \lambda_3) \partial_{\lambda_1} + (\tilde{h}_2 - h_1 \lambda_3) \partial_{\lambda_2} + \tilde{h}_3 \partial_{\lambda_3}. \end{aligned}$$

Then, if $g_3 = h_3 = 0$,

$$\langle g, h \rangle_{(\alpha_p)}^{T^*\mathbb{H}} = g_1 h_1 + g_2 h_2 + (\tilde{g}_1 + g_2 \lambda_3)(\tilde{h}_1 + h_2 \lambda_3) + (\tilde{g}_2 - g_1 \lambda_3)(\tilde{h}_2 - h_1 \lambda_3)$$

Let $(x, y, z, \lambda_1, \lambda_2, \lambda_3, \gamma_1, \ldots, \gamma_6)$ give coordinates in $T^*T^*\mathbb{H}$. The connection in $T^*T^*\mathbb{H}$ is then

$$\begin{aligned} \text{span}\{ & X + \lambda_3 \partial_{\lambda_2} + \gamma_4 \partial_{\gamma_2} + \gamma_6 \partial_{\gamma_3}, \\ & Y - \lambda_3 \partial_{\lambda_1} - \gamma_4 \partial_{\gamma_1} + \gamma_5 \partial_{\gamma_3}, \\ & Z - \gamma_6 \partial_{\gamma_1} - \gamma_5 \partial_{\gamma_2}, \\ & \partial_{\gamma_1}, \partial_{\gamma_2}, \partial_{\gamma_3} \}, \end{aligned}$$

and the geodesic equations are:

$$\dot{\gamma}_1 = -\gamma_4\gamma_2 - \gamma_6\gamma_3$$
$$\dot{\gamma}_2 = \gamma_4\gamma_1 - \gamma_5\gamma_3$$
$$\dot{\gamma}_3 = \gamma_5\gamma_2 + \gamma_6\gamma_1$$
$$\dot{\gamma}_4 = \dot{\gamma}_5 = \dot{\gamma}_6 = 0.$$

The constant vector $(\gamma_4, \gamma_5, \gamma_6)$ can be interpreted as an acceleration coefficient as we re-write the system:

$$\begin{pmatrix} \dot{\gamma}_1 \\ \dot{\gamma}_2 \\ \dot{\gamma}_3 \end{pmatrix} = \begin{pmatrix} 0 & -\gamma_4 & -\gamma_6 \\ \gamma_4 & 0 & -\gamma_5 \\ \gamma_6 & \gamma_5 & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}. \tag{4.4.1}$$

Furthermore, since there are no non-trivial solutions to

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & -\gamma_4 & -\gamma_6 \\ \gamma_4 & 0 & -\gamma_5 \\ \gamma_6 & \gamma_5 & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix},$$

for $(\gamma_4, \gamma_5, \gamma_6) \neq (0, 0, 0)$, there are no abnormal geodesics. Integrating (4.4.1), letting A be the constant matrix on the r.h.s. and $\gamma = (\gamma_1, \gamma_2, \gamma_3)$, we have

$$\gamma(t) = e^{tA}\gamma(0).$$

Geodesics are then only unique for $0 \leq t\sqrt{\gamma_4^2 + \gamma_5^2 + \gamma_6^2} < 2\pi$.

5.0 TRANSPORT SETS

We now set out to construct an admissible map which satisfies (2.1.2) for $(\mathbb{H}, d_c)$. By Proposition 2.1.2, such a map would solve Monge's transport problem. We start with a solution $u \in \mathrm{Lip}(\mathbb{H}, d_c)$ of the Kantorovich dual problem 2.1.2, which exists by Proposition 2.1.1.

5.1 TRANSPORT RAYS

Suppose now that $p \in \Omega_\mu, q \in \Omega_\nu$ satisfy

$$u(p) - u(q) = d_c(p, q).$$

Since u is 1-Lipschitz, along any geodesic $\sigma : [0,1] \to \mathbb{H}$ connecting $p = \sigma(0)$ and $q = \sigma(1)$, if $r \in \sigma([0,1])$ then $u(r) = u(q) + d_c(r, q)$. Indeed,

$$\begin{aligned} d_c(r, q) \geq u(r) - u(q) &= (u(p) - u(q)) - (u(p) - u(r)) \\ &\geq d_c(p, q) - d_c(p, r) = d_c(r, q). \end{aligned}$$

This motivates the following definition.

Definition 5.1.1. A *transport ray* R with *upper end* $a \in \Omega_\mu$ and *lower end* $b \in \Omega_\nu$ is the image of a geodesic $\sigma : [0,1] \to \mathbb{H}$ with $\sigma(0) = a, \sigma(1) = b$, such that

1. $u(a) - u(b) = d_c(a, b)$ and

2. R is maximal in the following sense: either $[a] = [b]$, or for all admissible $t > 1$ such that $b_t := \exp_p(t \exp_p^{-1}(b)) \in \Omega_\nu$ and $a_t := \exp_b(t \exp_b^{-1}(a)) \in \Omega_\mu$ we have

$$\begin{aligned} |u(a_t) - u(b)| &< d_c(a_t, b) \text{ and} \\ |u(b_t) - u(a)| &< d_c(b_t, a). \end{aligned}$$

Again, if follows that for any $r \in R$,

$$u(r) = u(b) + d_c(r, b) = u(a) - d(a, r).$$

We will also denote by R_b^a a transport ray with upper end a and lower end b.

Definition 5.1.2. Let $\mathcal{R}$ be the set of all transport rays. Denote by $T = \{p \in R \in \mathcal{R}\}$, the union of all transport rays. Define the *rays of length zero*:

$$T_0 = \{w \in \Omega_\mu \cap \Omega_\nu : \text{ for all } \tilde{w} \in \Omega_\mu \cup \Omega_\nu, |u(w) - u(\tilde{w})| < d_c(w, \tilde{w})\}$$

Also, denote $A = \{a : R_b^a \in \mathcal{R}\}$ and $B = \{b : R_b^a \in \mathcal{R}\}$ the sets of upper endpoints and lower endpoints, respectively.

Lemma 5.1.3. $\Omega_\mu \cup \Omega_\nu \subseteq T_0 \cup T$.

Proof. For $z \in \Omega_\mu \cup \Omega_\nu$, since u is 1-Lipschitz, either there exists a $z' \in \Omega_\mu \cup \Omega_\nu$, $z \neq z'$, such that $|u(z) - u(z')| = d(z, z')$, or no such z' exists. In the case where we have this z, then $z \in T$ because it is on a transport ray. Otherwise, $z \in T_0$ by definition. Thus $z \in T_0 \cup T$, so $\Omega_\mu \cup \Omega_\nu \subseteq T_0 \cup T$. □

Lemma 5.1.4. *$T_0 \cup T$ is compact.*

Proof. Since $\mathbb{H}$ and the Euclidean space $\mathbb{R}^3$ are topologically equivalent, it is sufficient to show $T_0 \cup T$ is closed and bounded.

Let $v : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ be given by

$$v(p, q) = u(p) - u(q).$$

This function is continuous, and so obtains a maximum $L < \infty$ on the compact set $\Omega_\mu \times \Omega_\nu$. Furthermore, $L \geq 0$ since, in the case $\Omega_\mu \cap \Omega_\nu$ is nonempty, then for any $x \in \Omega_\mu \cap \Omega_\nu$, $v(x, x) = 0$. In the case that $\Omega_\mu \cap \Omega_\nu$ is empty, then T is nonempty, and so there exists a transport ray R_b^a, with $v(a, b) = d(a, b) > 0$.

If $A = (T_0 \cup T) \setminus (\Omega_\mu \cup \Omega_n u)$ is nonempty, then any $z \in A$ lies on a transport ray R_z. Let a, b be the upper and lower ends of R_z, so that

$$d(a, z) + d(z, b) = d(a, b) = v(a, b) \leq L.$$

This implies that A lies in an $L-$neighborhood of Ω_μ and Ω_ν, and thus $T_0 \cup T$ is bounded.

Now, let $z_n \in T_0 \cup T_1$ be a sequence that converges to $z \in \mathbb{H}$. If an infinite subsequence of $\{z_n\}$ lies in $\Omega_\mu \cup \Omega_\nu$, then $z \in \Omega_\mu \cup \Omega_\nu$ by the compactness assumption. We may therefore assume that the $z_n \in A$, and as such, there exists a sequence of transport rays $\{R_n\}$ such that $z_n \in (R_n)_{b_n}^{a_n}$. Taking a subsequence if necessary,

$$a_{n_j} \to a \in \Omega_\mu,$$
$$b_{n_j} \to b \in \Omega_\nu,$$

and for each n_j,

$$d(a_{n_j}, z_{n_j}) + d(z_{n_j}, b_{n_j}) = d(a_{n_j}, b_{n_j}) = u(a_{n_j}) - u(b_{n_j}).$$

As such, we obtain by continuity that

$$d(a, z) + d(z, b) = d(a, b) = u(a) - u(b)$$

so that either z is on a transport ray, or $z = a = b \in T_0$. Either way, $z \in T_0 \cup T$, which is therefore closed, and thus compact. □

Lemma 5.1.5. $A \cap B = \emptyset$.

Proof. Arguing by contradiction, suppose there was a $p \in A \cap B$. Then there would exist transport rays R_p^a, R_b^p (with $p \neq a$ and $p \neq b$). On the other hand,

$$
\begin{aligned}
d(a,b) &\leq d(a,p) + d(p,b) \\
&= u(a) - u(p) + u(p) - u(b) \\
&= u(a) - u(b) \leq d(a,b).
\end{aligned}
$$

We therefore have the equality $u(a) - u(b) = d(a,b)$, indicating that p is in the interior of the transport ray R_b^a, contradicting the assumption that p is in the set of endpoints of transport rays. □

Lemma 5.1.6 (Transport Rays Are Disjoint)**.** *Let R_1, R_2 be two distinct transport rays such that $R_1 \cap R_2 \neq \emptyset$. Then either*

i. $R_1 \cap R_2 = \{c\}$ and c is either the upper end of both rays, or the lower end of both rays; or else

ii. $R_1 \cap R_2 = \{c_u, c_l\}$ where c_u is the upper end of both rays and c_l is the lower end of both rays.

In particular, an interior point of a transport ray cannot lie on any other transport ray:

$$
\big(R_1\big)^\circ \cap \big(R_2\big)^\circ = \emptyset.
$$

Proof. First we note that if $d(x,y) + d(y,z) = d(x,z)$ then y is in the minimizing geodesic connecting x and z.

Since $R_i, i = 1,2$ are transport rays, it follows that each R_i is a minimizing geodesic $\sigma_i : [0,1] \to \mathbb{H}$, where we chose the parameterization so that the functions $\tau \mapsto u(\sigma_i(\tau))$ are decreasing on $[0,1]$.

Suppose $R_1 \neq R_2$ share a point c, and let $c = \sigma_1(\tau_1) = \sigma_2(\tau_2)$, where $\tau_i \in [0,1]$. Then the vectors $v_1, v_2 \in T_c\mathbb{H}$ so that $\sigma_i(t) = \exp_c[(t - \tau_i)v_i]$ are not collinear. Indeed, since $R_i \subset \{\exp_c tv_i : t \in \mathbb{R}\}$ and each R_i are maximal, if the v_i were collinear then $R_1 = R_2$.

Suppose $R_1 \cap R_2$ contains at least two distinct points $\{c_u, c_l\}$ where $u(c_u) > u(c_l)$. Then $u(c_u) = u(c_l) + d(c_u, c_l)$. Since tangent vectors to R_1 and R_2 at c_l are not collinear, R_1 and R_2 between c_u and c_l do not coincide. Then c_u lies in the cut locus of c_l since the R_i are minimizing geodesics. Thus c_l and c_u are the endpoints of each R_i, and $R_1 \cap R_2 = \{c_l, c_u\}$.

Suppose now that R_1, R_2 have only one point in common: $R_1 \cap R_2 = \{c\}$. Let $a_i = \sigma_i(0)$, and $b_i = \sigma_i(1)$ for $i = 1, 2$. These are the upper and lower endpoints of R_i.

We will assume $c \neq b_2$ and show this forces $c = a_1$. Since R_1 has non-zero length, it follows that $c \neq b_1$ which, by symmetry, forces $c = a_2$, completing the proof. Similarly, assuming $c \neq a_2$ forces $c = b_1 = b_2$.

First, assume $c \neq b_2$. This means that $b_2 \notin R_1$. By definition (5.1.1 part 1),

$$u(c) = u(b_2) + d(c, b_2), \qquad\qquad u(c) = u(a_1) - d(a_1, c),$$

so that

$$u(a_1) - u(b_2) = d(a_1, c) + d(c, b_2) \geq d(a_1, b_2).$$

Strict inequality would violate the Lipschitz condition on u so equality must hold. This implies that c lies on a minimizing geodesic γ from a_1 to b_2. Since γ and σ_1 both minimize distance from c to a_1, either they coincide or c lies in the cut locus of a_1. If c is in the cut locus of a_1, then $c = b_2$, contradicting our assumption. The curve segments must then coincide between a_1 and c. It then follows that $\gamma \subset \{\exp_c tv_1 : t \in \mathbb{R}\}$. Since R_1 is maximal, $\gamma \subset R_1$, and thus $b_2 \in R_1$, contradicting our assumption. Thus $c = a_1$. $\square$

Lemma 5.1.7 (Differentiability of u along rays)**.** *If z_0 lies in the relative interior of some transport ray R, then u is differentiable at z_0. If R has upper endpoint p and is parametrized as $\sigma : [0,1] \to \mathbb{H}$ such that $\sigma(t) = \exp_p tV$ for $V \in \mathcal{D}_p$, then the function $t \mapsto u(\sigma(t))$ is decreasing on $[0,1]$. Suppose $z_0 = \sigma(t_0)$ for $t_0 \in (0,1)$, then*

$$\nabla u(z_0) = (Xu(z_0))X + (Yu(z_0))Y + (Zu(z_0))Z = -\frac{V(t_0)}{\|V\|_H}$$

where $V(t_0) \in T_{z_0}\mathbb{H}$ is the parallel translate of $V \in T_p\mathbb{H}$ along R, and $\|V\|_H$ the the norm of the horizontal part of V.

Proof. Let p, q be the upper and lower ends of R, respectively. Then $u(p) \geq u(z_0) \geq u(q)$ and $d_p(\cdot)$ and $d_q(\cdot)$ are smooth in a neighborhood of z_0. Since R is a minimizing geodesic and u is 1-Lipschitz then for any $z \in \mathbb{H}$,

$$d_q(z) \geq u(z) - u(q) \geq u(p) - u(q) - d_p(z) = d(p,q) - d_p(z).$$

Since R is a transport ray, this holds with equalities for $z = z_0$. Thus u is differentiable at z_0. In addition,

$$\nabla u(z_0) = \nabla d_q(z_0) = -\nabla d_p(z_0) = -\frac{V(t_0)}{\|V\|_H},$$

the last equality by equation (4.3.1). □

Definition 5.1.8. Let A^b be the set of points $p \in A$ (called "bad") which are the endpoint for at least two transport rays. Likewise define B^b. Denote (the "good" points) $A^g = A \setminus A^b$ and $B^g = B \setminus B^b$.

Definition 5.1.9. A subset $T' \subseteq T$ is called a *transfer subset* if for any $p \in T'$ there is a ray $R \in \mathcal{R}$ such that $p \in R \subseteq T'$.

5.2 DECOMPOSITION OF TRANSFER SETS

Following the development of Trudinger and Wang, we decompose T into the union of disjoint transfer subsets $\{T_j\}$ such that each T_j lies in a "geodesic cylinder" in the following sense.

Let $B_r(x_0), B_{r'}(y_0) \subset \Omega$ such that $\bar{B}_r \cap \bar{B}_{r'} = \emptyset$. Define the set:

$$T^* = \bigcup \{R_y^x \in \mathcal{R} : x \in A \cap B_r \text{ and } y \in B \cap B_{r'}\} \tag{5.2.1}$$

Lemma 5.2.1. T^* *is measurable.*

Proof. Let T' be the (transfer subset) which is the union of all transfer rays which intersect both closed balls $\bar{B}_r, \bar{B}_{r'}$. Let $\{R_{y_i}^{x_i}\}_{i\in\mathbb{N}} \subset \mathcal{R}$ such that $\{x_i\}_{i\in\mathbb{N}} \subset \bar{B}_r$, $\{y_i\}_{i\in\mathbb{N}} \subset \bar{B}_{r'}$, $x_i \to x$, and $y_i \to y$. Then the limiting geodesic connecting x and y must be a transport ray since

$$\begin{aligned} d(x,y) \geq |u(x)-u(y)| &\geq -|u(x)-u(x_i)| + |u(x_i)-u(y_i)| - |u(y_i)-u(y)| \\ &= |u(x)-u(x_i)| + d(x_i,y_i) + |u(y_i)-u(y)| \\ &\geq -|u(x)-u(x_i)| - d(x_i,x) + d(x,y) - d(y,y_i) - |u(y_i)-u(y)| \end{aligned}$$

and so $|u(x)-u(y)| = d(x,y)$. Therefore T' is closed.

Consider the transfer subset T'' which is the union of transport rays satisfying one of the following conditions:

a) the ray intersects both open balls B_r and $B_{r'}$ and both of its endpoints don't lie in the open balls,
b) the ray intersects one open ball and is tangent to the other, or
c) the ray is tangent to both balls.

T'' is similarly closed. It follows that $T' \setminus T''$ is measurable. Considering sequences of balls $B_{r_j}(x_0) \nearrow B_r(x_0)$ and $B_{r'_j}(y_0) \nearrow B_{r'}(y_0)$, we have corresponding sequences of closed sets $\{T'_j\}$ and $\{T''_j\}$. Hence $T^* = \bigcup\{T'_j \setminus T''_j\}$ is measurable. □

Remark 5.2.2. The set T^* defined in (5.2.1) is still measurable if B_r and $B_{r'}$ are replaced by open sets E and F such that $\bar{E} \cap \bar{F} = \emptyset$. Furthermore, the set T^* remains measurable if E and F are only measurable and $\bar{E} \cap \bar{F} = \emptyset$ since we can choose a sequence of open sets shrinking to these measurable sets.

Let $\{B_{r_j}(x_j), B_{r'_j}(y_j)\}$ be a sequence of pairs of balls such that $\mathcal{R}$ is covered by

$$\tilde{\mathcal{R}}_j = \{R_y^x \in \mathcal{R} : x \in B_{r_j}(x_j) \text{ and } y \in B_{r'_j}(y_j)\},$$

where we also require $r_j, r'_j \leq \frac{1}{16} d(x_j, y_j)$ and for each $R_1, R_2 \in \tilde{R}_j$, and each $p \in R_1$, $B_{\bar{r}_j} \cap R_2 \neq \emptyset$, where $\bar{r}_j = \max(r_j, r'_j)$. These may not be disjoint, so we let $\mathcal{R}_1 = \tilde{\mathcal{R}}_1$ and for $j > 1$

$$\mathcal{R}_j = \tilde{\mathcal{R}}_j \setminus \left\{ \bigcup_{k<j} \tilde{\mathcal{R}}_k \right\}$$

Then $\mathcal{R} = \bigcup \mathcal{R}_j$ and the $\mathcal{R}_j$ are mutually disjoint.

We will denote the following:

i) T_j is the transfer subsets consisting of the transport rays in $\mathcal{R}_j$,

ii) $A_j = T_j \cap X$ and $B_j = T_j \cap Y$,

iii) $A_j^b = T_j \cap A^b$ and $B_j^b = T_j \cap B^b$, and

iv) $A_j^g = T_j \cap A^g$ and $B_j^g = T_j \cap B^g$.

Lemma 5.2.3. *Fix some T_j in the above decomposition. For all $p \in A_j$ and $q \in B_j$,*

$$u(p) > u(q).$$

In particular, there is a level set of u intersecting each ray in $\mathcal{R}_j$.

Proof. Arguing by contradiction, suppose there exists a $p_1 \in X_j$ and a $q_2 \in Y_j$ such that $u(p_1) \leq u(q_2)$. Let $p_2 \in Y_j$ so that $R_{p_2}^{p_1} \in \mathcal{R}_j$. Since $u(p_1) - u(p_2) = d_c(p_1, p_2)$, then

$$d_c(p_1, p_2) \leq u(q_2) - u(p_2) \leq d_c(q_2, p_2) \leq 2r_j' \leq \frac{1}{8} d_c(p_1, p_2),$$

providing the contradiction. □

6.0 LIPSCHITZ ESTIMATES

Recall the definition

$$\mathcal{D}_p = \mathbb{R}^2 \times (-2\pi, 2\pi) \subset T_p\mathbb{H},$$

which is the domain of the geodetic exponential function $\exp_p(\cdot)$. Let

$$\mathcal{B}_p = \{(P,Q) \in \mathcal{D}_p \times \mathcal{D}_p : [\exp_p P] \neq [\exp_p Q]\}.$$

Define the following maps for $p \in \mathbb{H}$:

1. $F_p : \mathcal{B}_p \to T\mathbb{H}$, $F_p(P,Q) = \exp^{-1}_{(\exp_p P)}(\exp_p Q)$.
2. $\Pi_q(P,p)$, the parallel translate of $P \in T_q\mathbb{H}$ along a minimizing geodesic from q to p.
3. $\Phi_p : \mathcal{B}_p \to T_p\mathbb{H}$, $\Phi_p(P,Q) = \Pi_{(\exp_p P)}(F(P,Q), p)$.

6.1 INITIAL LEMMAS

In the following Lemmas, the norm $|V|_p = \|\pi_*(V)\|_p$ is the norm of the horizontal part of $V \in T_p\mathbb{H}$.

Lemma 6.1.1. *For any compact $K \subset \mathcal{B}\mathbb{H} = \cup_{p\in\mathbb{H}}\mathcal{B}_p$, there exists $C > 0$ such that for any $(p, P_1, P_2) \in K$, any multi-index $\beta = (\beta_1, \beta_2)$ with $k = |\beta|$, any $\xi_1, \ldots, \xi_k \in H_p\mathbb{H}$,*

$$\|D_\beta \Phi_p(P_1,P_2)(\xi_1,\ldots,\xi_k)\|_H \leq C \prod_{j=1}^{k} \|\xi_j\|_H. \tag{6.1.1}$$

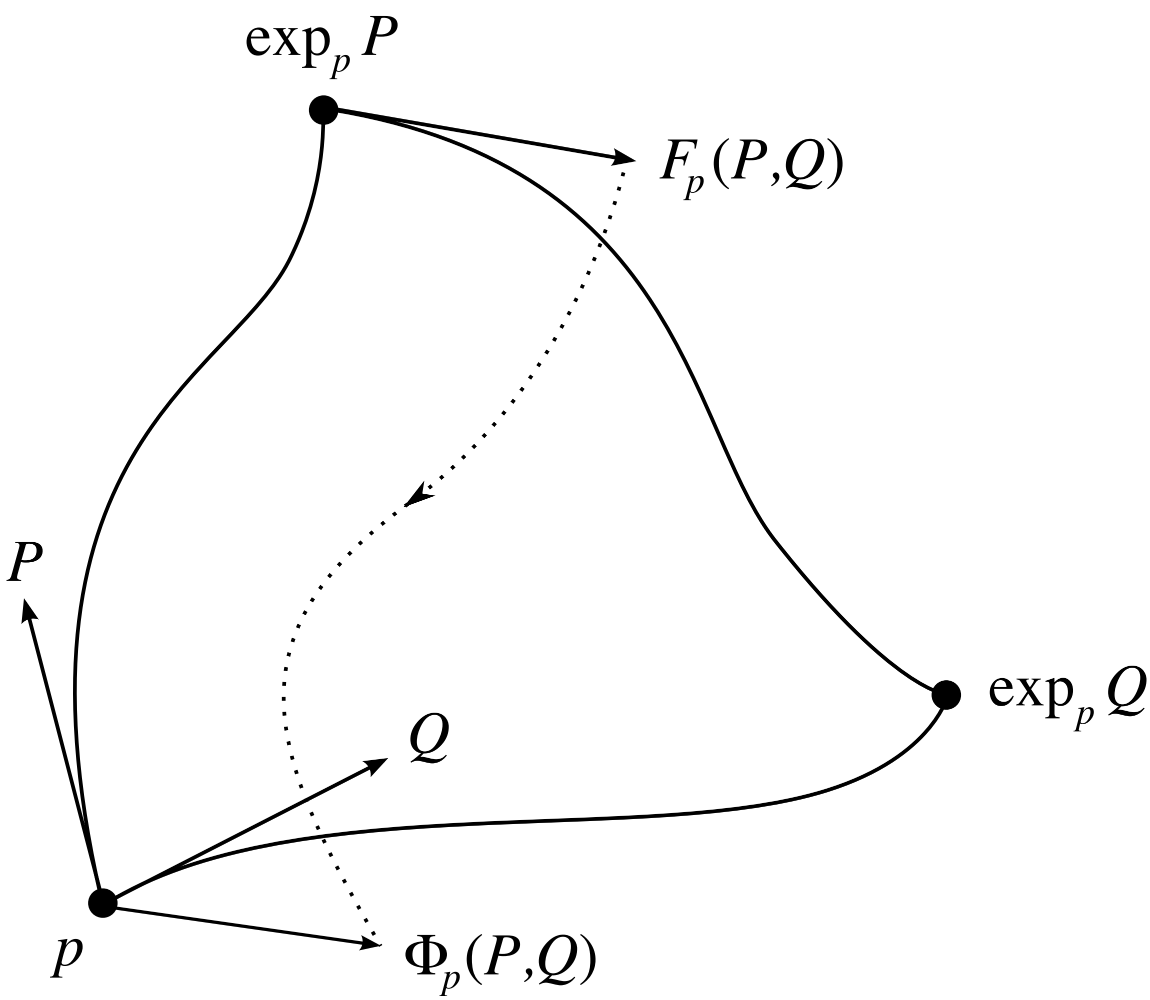

Figure 2: Functions F, Phi

Proof. In coordinates, Φ is a smooth mapping of the form $\Phi(p, P, Q) = (p, \varphi(P, Q))$ where $\varphi : \mathcal{B}_p \to \mathbb{R}^3$ is smooth. Indeed, since Φ is the parallel translate of $F = F(P_1, P_2) \in T_{\exp_p P_1}\mathbb{H}$,

$$\exp_{\exp_p P_1} F = \exp_p P_2 \qquad \text{or,}$$

$$(p * \exp_0 P_1) * \exp_0 F = p * \exp_0 P_2 \qquad \text{or, finally}$$

$$\exp_0 F = (-\exp_0 P_1) * (\exp_0 P_2).$$

By hypothesis, the right-hand side of this expression is not in the center of $\mathbb{H}$, so we may apply $\exp_0^{-1}$ to both sides. To show that this is a smooth function, consider:

$$(x, y, z) := \exp_0(v_1, v_2, v_3).$$

By Lemma (4.3.5), isolating v_1, v_2 gives:

$$v_1 = xh(v_3) + \frac{1}{2}yv_3,$$
$$v_2 = yh(v_3) - \frac{1}{2}xv_3,$$

where

$$h(v_3) = \frac{1}{2}\frac{v_3(1 + \cos v_3)}{\sin v_3} = \frac{1}{2}v_3 \cot\left(\frac{1}{2}v_3\right) = 1 - \frac{1}{12}v_3^2 - \cdots$$

Furthermore, a simple calculation shows:

$$\frac{2z}{x^2 + y^2} = \frac{1}{2}\frac{v_3 - \sin v_3}{1 - \cos v_3} = -h'(v_3) = \frac{1}{6}v_3 + \cdots .$$

Since h' is invertible, we can expand:

$$v_3 = \left(\frac{12z}{x^2 + y^2}\right) - \frac{1}{30}\left(\frac{12z}{x^2 + y^2}\right)^3 + \frac{3}{1400}\left(\frac{12z}{x^2 + y^2}\right)^5 - \frac{1}{6000}\left(\frac{12z}{x^2 + y^2}\right)^7 + \cdots$$

and so:

$$v_1 = x\left(1 - \frac{1}{12}\left(\frac{12z}{x^2 + y^2}\right)^2 + \frac{1}{240}\left(\frac{12z}{x^2 + y^2}\right)^4 - \cdots\right)$$
$$+ \frac{y}{2}\left(\left(\frac{12z}{x^2 + y^2}\right) - \frac{1}{30}\left(\frac{12z}{x^2 + y^2}\right)^3 + \cdots\right),$$

$$v_2 = y\left(1 - \frac{1}{12}\left(\frac{12z}{x^2 + y^2}\right)^2 + \frac{1}{240}\left(\frac{12z}{x^2 + y^2}\right)^4 - \cdots\right)$$
$$- \frac{x}{2}\left(\left(\frac{12z}{x^2 + y^2}\right) - \frac{1}{30}\left(\frac{12z}{x^2 + y^2}\right)^3 + \cdots\right).$$

Since $\exp_0^{-1}(x, y, z)$ is analytic in x, y, and $\frac{z}{x^2+y^2}$ (each of which are finite in K) we thus obtain existence of C_β such that for any $\xi_1, \dots, \xi_k \in \mathbb{R}^2$

$$|D_\beta \varphi(q, P_1, P_2)(\xi_1, \dots, \xi_k)| \leq C_\beta \prod_{j=1}^{k} |\xi_j|. \tag{6.1.2}$$

Indeed, the constant C_β is obtained by taking the supremum of the left-hand side of (6.1.2) over the compact set

$$\{(q, P_1, P_2, \xi_1, \dots, \xi_k) : (q, P_1, P_2) \in K, |\xi_1| = \cdots = |\xi_k| = 1\}.$$

Therefore we obtain (6.1.1). $\square$

Lemma 6.1.2. *Let K be a compact subset of $\mathcal{B}\mathbb{H}$. There exists C such that for any $p \in \pi_{\mathbb{H}}(K)$ and $P, Q_1, Q_2 \in K_p$*

$$\left\| \left[\Phi_p(P, Q_1) - \Phi_p(P, Q_2)\right] - (Q_1 - Q_2) \right\|_H \leq C\|P\|_H \|Q_1 - Q_2\|_H, \tag{6.1.3}$$

and for any $p \in \pi_{\mathbb{H}}(K)$ and $P, Q \in K_p$,

$$\left| d(\exp_p P, \exp_p Q) - \|P - Q\|_H \right| \leq C\|P - Q\|_H (\|P\|_H + \|P - Q\|_H). \tag{6.1.4}$$

Proof. Let $p \in \pi(K)$, $P, Q_1, Q_2 \in K_p$. For any $W \in T_p\mathbb{H}$ write $W = W^H + W^V$ for the horizontal-vertical decomposition. For $k = 1, 2$,

$$\begin{aligned} \Phi_p(P, Q_k) &= \Phi_p(0, Q_k) + \int_0^1 D_1\Phi_p(P(t), Q_k)P^H \, dt \\ &= Q_k + \int_0^1 D_1\Phi_p(P(t), Q_k)P^H \, dt, \end{aligned}$$

where $tP \in \mathcal{D}_p$ for $0 \leq t \leq 1$. Then we estimate using (6.1.1):

$$\begin{aligned} &\left\| D_1\Phi_p(tP, Q_1)P - D_1\Phi_p(tP, Q_2)P \right\|_H \\ &\qquad = \left\| \int_0^1 D_2 D_1 \Phi_p(tP, \tau Q_1 + (1 - \tau)Q_2)(P, Q_1 - Q_2) d\tau \right\|_H \\ &\qquad \leq C\|P\|_H \|Q_1 - Q_2\|_H, \end{aligned}$$

thus

$$
\begin{aligned}
&\Big\| \big[\Phi_p(P,Q_1) - \Phi_p(P,Q_2)\big] - (Q_1 - Q_2) \Big\|_H \\
&\qquad = \Big\| \int_0^1 D_1\Phi_p(tP,Q_1)P - D_1\Phi_p(tP,Q_2)P \ dt \Big\|_H \\
&\qquad \leq \int_0^1 \big\| D_1\Phi_p(tP,Q_1)P - D_1\Phi_p(tP,Q_2)P \big\|_H dt \\
&\qquad \leq C\|P\|_H \|Q_1 - Q_2\|_H.
\end{aligned}
$$

Now we prove (6.1.4). Let $p \in \pi(K)$ and $P, Q \in K_p$. Then

$$
\begin{aligned}
\Phi_p(P,Q) &= \Phi_p(P,P) - D_2\Phi_p(P,P)(P-Q) \\
&\qquad + \int_0^1 D_2^2\Phi_p(P,\tilde{Q}(t))(P-Q)(P-Q)tdt \\
&= -D_2\Phi_p(P,P)(P-Q) + \int_0^1 D_2^2\Phi_p(P,\tilde{Q}(t))(P-Q,P-Q)tdt,
\end{aligned}
$$

where $\tilde{Q}(t) := tP + (1-t)Q \in \mathcal{D}_p$ for $0 \leq t \leq 1$. Since $D_2\Phi_p(0,P) = Id_{T_p\mathbb{H}}$, we compute

$$
\begin{aligned}
D_2\Phi_p(P,P)(P-Q) &= D_2\Phi_p(0,P)(P-Q) + \int_0^1 D_1D_2\Phi_p(tP,P)(P,P-Q)dt \\
&= (P-Q) + \int_0^1 D_1D_2\Phi_p(tP,P)(P,P-Q)dt,
\end{aligned}
$$

where $tP \in \mathcal{D}_p$ for $0 \leq t \leq 1$. Thus using 6.1.1 with $C = C(T_0 \cup T)$, we have

$$
\big\| \Phi_p(P,Q) + (P-Q) \big\|_H \leq C\|P-Q\|_H \big(\|P\|_H + \|P-Q\|_H \big),
$$

or, finally

$$
\begin{aligned}
\big| d(\exp_p P, \exp_p Q) - \|P-Q\|_H \big| &= \big| \|\Phi_p(P,Q)\|_H - \|P-Q\|_H \big| \\
&\leq \big\| \Phi_p(P,Q) + (P-Q) \big\|_H \\
&\leq C\|P-Q\|_H (\|P\|_H + \|P-Q\|_H).
\end{aligned}
$$

□

6.2 KEY LEMMAS: LIPSCHITZ CONTROL OF DIRECTIONS OF TRANSPORT RAYS

Let $u : \mathbb{H} \to \mathbb{R}$ be 1-Lipschitz and let R_1, R_2 be transport rays. Suppose $y_k \in \left(R_k\right)^\circ$ satisfy $u(y_1) = u(y_2)$. Let $z_k, x_k \in \left(R_k\right)^\circ$ denote points at a distance σ above and below y_k, respectively, so that for $k = 1, 2$

$$u(z_k) = u(y_k) + \sigma, \tag{6.2.1}$$

$$u(x_k) = u(y_k) - \sigma, \tag{6.2.2}$$

$$d(x_k, z_k) = 2\sigma, \text{ and} \tag{6.2.3}$$

$$d(x_k, y_k) = d(y_k, z_k) = \sigma. \tag{6.2.4}$$

Lemma 6.2.1. *There exist real numbers C, σ_0 depending only on $T_0 \cup T$ such that if $0 < \sigma < \sigma_0$, then*

$$d(x_1, x_2), d(z_1, z_2) \leq C d(y_1, y_2).$$

The proof is adapted from [5]. Several key properties of Riemannian manifolds needed in that proof are not available in $\mathbb{H}$: namely, the Hopf-Rinow theorem and the existence of geodetically convex sets.

Proof. Without loss, we may assume

$$d(x_1, x_2) \leq d(z_1, z_2). \tag{6.2.5}$$

Otherwise we consider $-u$, reversing the direction of the rays. We may also assume that

$$d(y_1, y_2) < \sigma. \tag{6.2.6}$$

(Indeed, if $d(y_1, y_2) \geq \sigma$,

$$d(z_1, z_2) \leq d(z_1, y_1) + d(y_1, y_2) = d(y_1, y_2) + 2\sigma \leq 3d(y_1, y_2),$$

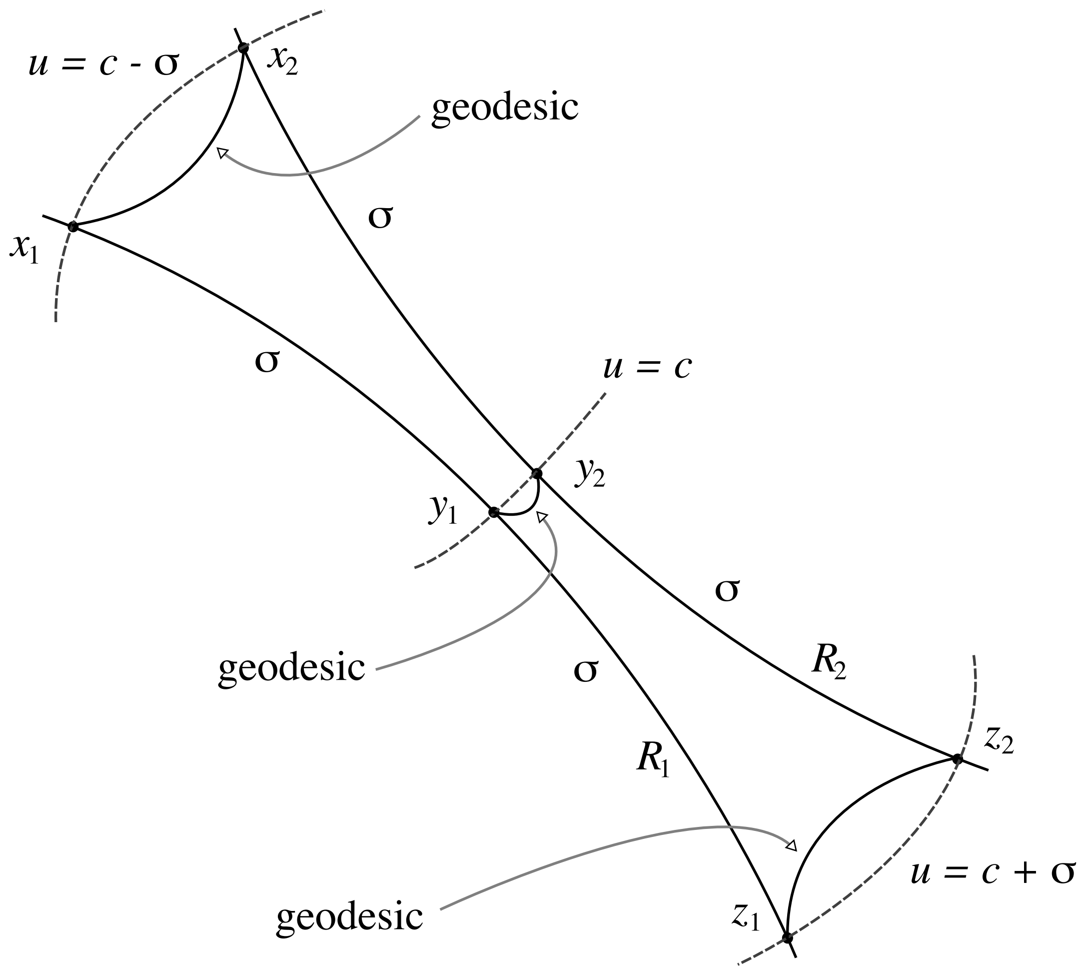

Figure 3: Diagram for Lemma 6.2.1

and the proof is complete.) Thus

$$d_c(x_1, z_2) \leq d_c(x_1, y_1) + d_c(y_1, y_2) + d_c(y_2, z_2) \leq 3\sigma, \tag{6.2.7}$$

$$d_c(z_1, x_2) \leq d_c(z_1, y_1) + d_c(y_1, y_2) + d_c(y_2, x_2) \leq 3\sigma \tag{6.2.8}$$

The Lipschitz bound will ultimately derive from statements about vectors in $T_{x_1}\mathbb{H}$. We will use $\exp_{x_1}^{-1}$ to map our problem into the tangent space at x_1. It is not clear, *a priori*, that this is possible; the domain of this mapping must be respected. Once this mapping is achieved, our proof follows the standard argument.

Since $x_k, y_k, z_k \in (R_k)^\circ$, we have by Lemma 4.3.6 that

$$[x_k] \neq [y_k] \text{ and } [x_k] \neq [z_k]. \tag{6.2.9}$$

Thus we may let

$$\begin{aligned} X_1 &= \exp_{x_1}^{-1} x_1 = 0 \in T_{x_1}\mathbb{H}, \\ Y_1 &= \exp_{x_1}^{-1} y_1 \in T_{x_1}\mathbb{H}, \text{ and} \\ Z_1 &= \exp_{x_1}^{-1} z_1 \in T_{x_1}\mathbb{H}. \end{aligned}$$

We may suppose $[x_1] \neq [x_2]$. Otherwise, we consider two cases: first, that we may shorten σ by some small $\varepsilon > 0$ so that $[x_1] \neq [x_2]$ (accomplishing our claim); or second, that such an ε cannot be found. In this second case we let $\alpha_k : [-\sigma, \sigma] \to R_k$ be the lenght parameterization of each curve between x_k and z_k. Then we have $[\alpha_1(t)] = [\alpha_2(t)]$ for $t \in [-\sigma, \sigma]$. By Lemma 3.2.4 there exists a $c \in \mathbb{R}$ so that $\alpha_1 = \alpha_2 + (0, 0, c)$, so that

$$\begin{aligned} d(x_1, x_2) = d(\alpha_1(-\sigma), \alpha_2(-\sigma)) &= \sqrt{4\pi c} \\ &= d(\alpha_1(\sigma), \alpha_2(\sigma)) = d(z_1, z_2) \\ &= d(\alpha_1(0), \alpha_2(0)) = d(y_1, y_1) \end{aligned}$$

and the lemma is proved.

Similarly, we may arrange so that

$$[x_1] \neq [y_2], [x_1] \neq [z_2], \tag{6.2.10}$$

$$[x_2] \neq [y_1], [x_2] \neq [z_1], \tag{6.2.11}$$

and in this case we may let

$$X_2 = \exp_{x_1}^{-1} x_2 \in T_{x_1}\mathbb{H},$$
$$Y_2 = \exp_{x_1}^{-1} y_2 \in T_{x_1}\mathbb{H}, \text{ and}$$
$$Z_2 = \exp_{x_1}^{-1} z_2 \in T_{x_1}\mathbb{H}.$$

Note that

$$\|X_k\|, \|Y_k\|, \|Z_k\| \leq 3\sigma, \tag{6.2.12}$$

where we drop the subscript H and use $\|\cdot\|$ for the norm of the horizontal projection.

Furthermore, we let:

$$\hat{Z}_k = \Phi(X_2, Z_k) \in T_{x_1}\mathbb{H},$$
$$\hat{Y}_2 = \Phi(X_2, Y_2) \in T_{x_1}\mathbb{H}, \text{ and}$$
$$\hat{X}_2 = \Phi(X_2, X_2) = 0 \in T_{x_1}\mathbb{H}.$$

From the foregoing definitions:

$$\begin{aligned}\|Z_1 - X_1\|_{x_1} = d_c(x_1, z_1) = u(z_1) - u(x_1) = u(z_2) - u(x_1)\\ \leq d(x_1, z_2) = \|Z_2 - X_1\|_{x_1}.\end{aligned}$$

Squaring this inequality yeilds

$$\begin{aligned}\|Z_1 - X_1\|_{x_1}^2 &\leq \|(Z_2 - Z_1) + (Z_1 - X_1)\|_{x_1}^2\\ &= \|Z_2 - Z_1\|_{x_1}^2 + 2\langle Z_2 - Z_1, Z_1 - X_1\rangle_{x_1} + \|Z_1 - X_1\|_{x_1}^2,\end{aligned}$$

from which

$$\langle Z_2 - Z_1, Z_1 - X_1\rangle_{x_1} \geq -\frac{1}{2}\|Z_2 - Z_1\|_{x_1}^2,$$

and finally, since

$$Y_1 = \frac{1}{2}(X_1 + Z_1),$$

we get

$$\langle Z_2 - Z_1, Z_1 - Y_1\rangle_{x_1} \geq -\frac{1}{4}\|Z_2 - Z_1\|_{x_1}^2. \tag{6.2.13}$$

Next we obtain a similar estimate of $\langle Z_1 - Z_2, Z_2 - Y_2\rangle$. To this end, note that

$$d_c(x_2, z_2) = u(z_2) - u(x_2) = u(z_1) - u(x_2) \leq d(x_2, z_1).$$

Furthermore,

$$\begin{aligned}
&\|\hat{Z}_k - \hat{X}_2\|_{x_1} = d_c(\exp_{x_1} Z_k, \exp_{x_1} X_2) = d_c(z_k, x_2), \text{ for } k = 1, 2 \text{ and} \\
&\|\hat{Y}_2 - \hat{X}_2\|_{x_1} = d_c(\exp_{x_1} Y_2, \exp_{x_1} X_2) = d_c(y_2, x_2),
\end{aligned}$$

thus

$$\|\hat{Z}_1 - \hat{X}_2\|_{x_1} \geq \|\hat{Z}_2 - \hat{X}_2\|_{x_1}.$$

Squaring this inequality, and using the fact that

$$\hat{Y}_2 = \frac{1}{2}(\hat{X}_2 + \hat{Y}_2)$$

we get

$$\langle \hat{Z}_1 - \hat{Z}_2, \hat{Z}_2 - \hat{Y}_2\rangle_{x_1} \geq -\frac{1}{4}\|\hat{Z}_2 - \hat{Z}_1\|_{x_1}^2. \tag{6.2.14}$$

In the subsequent estimate we use the inequality

$$-\|P\|^2 \geq -(1+\varepsilon)\|Q\|^2 - \frac{2}{\varepsilon}\|P - Q\|^2, \text{ for any } P, Q \in \mathbb{R}^n, \varepsilon \in (0,1], \tag{6.2.15}$$

where $\|\cdot\|$ is a norm in $\mathbb{R}^n$. This is easily checked: by expanding and rearranging terms, we rewrite this as

$$\left(1 - \frac{\varepsilon}{2}\right)\|P\|^2 + \left(1 + \frac{\varepsilon}{2} + \frac{\varepsilon^2}{2}\right)\|Q\|^2 \geq 2\langle P, Q\rangle,$$

and this is true since $(1 - \frac{\varepsilon}{2})(1 + \frac{\varepsilon}{2} + \frac{\varepsilon^2}{2}) = 1 + \frac{\varepsilon^2}{4} - \frac{\varepsilon^3}{4} \geq 1$ if $\varepsilon \in (0, 1]$.

In the left-hand side of (6.2.14) we put

$$\hat{Z}_1 - \hat{Z}_2 = (Z_1 - Z_2) + \big[(\hat{Z}_1 - \hat{Z}_2) - (Z_1 - Z_2)\big] \text{ and}$$
$$\hat{Z}_1 - \hat{Y}_2 = (Z_1 - Y_2) + \big[(\hat{Z}_1 - \hat{Y}_2) - (Z_1 - Y_2)\big].$$

We estimate the right-hand side of (6.2.14) from below using (6.2.15) with

$$P = \hat{Z}_1 - \hat{Z}_2 \text{ and } Q = Z_1 - Z_2.$$

Thus we get from (6.2.14)

$$\begin{aligned}
\langle Z_1 - Z_2, Z_2 - Y_2\rangle \geq & -\frac{1}{4}(1+\varepsilon)\|Z_2 - Z_1\|^2 \\
& - \frac{1}{2\varepsilon}\|(Z_1 - Z_2) - (\hat{Z}_1 - \hat{Z}_2)\|^2 \\
& - \|(Z_1 - Z_2) - (\hat{Z}_1 - \hat{Z}_2)\|\|Z_2 - Y_2\| \\
& - \|Z_1 - Z_2\|\|(Z_1 - Y_2) - (\hat{Z}_1 - \hat{Y}_2)\| \\
& - \|(Z_1 - Z_2) - (\hat{Z}_1 - \hat{Z}_2)\|\|(Z_1 - Y_2) - (\hat{Z}_1 - \hat{Y}_2)\|
\end{aligned} \tag{6.2.16}$$

for any $\varepsilon \in (0,1]$.

Now we estimate the error terms in (6.2.16). We use Lemma 6.1.2 with the compact set $K = T_0 \cup T$. In the calculations below C will denote different constants depending only on K.

From (6.1.4) it follows that if σ_0 is chosen small depending on K and $\sigma \leq \sigma_0$, then

$$\frac{1}{2}d(y_1, y_2) \leq \|Y_1 - Y_2\| \leq 2d(y_1, y_2), \tag{6.2.17}$$
$$\frac{1}{2}d(z_1, z_2) \leq \|Z_1 - Z_2\| \leq 2d(z_1, z_2). \tag{6.2.18}$$

This, together with our initial assumption assumption (6.2.5), we have

$$\|X_1 - X_2\| \leq 2\|Z_1 - Z_2\|.$$

Now, using this estimate with (6.1.3),

$$\|Z_2 - Y_2\| \le \|Z_2\| + \|Y_2\| \le 6\sigma$$
$$\big\|(\hat{Z}_1 - \hat{Z}_2) - (Z_1 - Z_2)\big\| \le C\|X_1 - X_2\|\|Z_1 - Z_2\| \le C\|Z_1 - Z_2\|^2$$
$$\big\|(\hat{Z}_2 - \hat{Y}_2) - (Z_2 - Y_2)\big\| \le C\|X_1 - X_2\|\|Z_2 - Y_2\| \le C\sigma\|Z_1 - Z_2\|.$$

These, together with (6.2.16) implies

$$\langle Z_1 - Z_2, Z_2 - Y_2\rangle \ge -\frac{1}{4}\left(1 + \varepsilon + C\sigma + C\frac{\sigma^2}{\varepsilon}\right)\|Z_2 - Z_1\|^2. \tag{6.2.19}$$

Choosing first $\varepsilon = \frac{1}{2}$, and then reducing if necessary $\sigma_0 > 0$ to achieve $C\sigma_0 + 2C\sigma_0^2 \le \frac{1}{2}$ where $C = C(T_0 \cup T)$, we obtain for $\sigma \le \sigma_0$

$$\langle Z_1 - Z_2, Z_2 - Y_2\rangle \ge -\frac{1}{2}\|Z_2 - Z_1\|^2. \tag{6.2.20}$$

Now, combining (6.2.13) and (6.2.20), we estimate

$$\begin{aligned}
\|Z_2 - Z_1\|\|Y_2 - Y_1\| &\ge \langle Z_2 - Z_1, Y_2 - Y_1\rangle \\
&= \langle Z_2 - Z_1, (Y_2 - Z_2) + (Z_2 - Z_1) + (Z_1 - Y_1)\rangle \\
&\ge -\frac{1}{2}\|Z_2 - Z_1\|^2 + \|Z_2 - Z_1\|^2 - \frac{1}{4}\|Z_2 - Z_1\|^2 \\
&= \frac{1}{4}\|Z_2 - Z_1\|^2.
\end{aligned}$$

Thus

$$\|Z_2 - Z_1\| \le 4\|Y_2 - Y_1\|.$$

Using (6.2.17) and (6.2.18), we conclude that if $\sigma \le \sigma_0$,

$$d(z_1, z_2) \le 2\|Z_1 - Z_2\| \le 8\|Y_1 - Y_2\| \le 16 d(y_1, y_2).$$

□

Lemma 6.2.2 (Ray Directions Vary Lipschitz Continuously). *Let R_1 and R_2 be transport rays, with upper end a_k and lower end b_k for $k = 1, 2$ respectively. If there are interior points $y_k \in (R_k)^\circ$ with $u(y_1) = u(y_2)$, then the ray directions satisfy a Lipschitz bound*

$$d_{T\mathbb{H}}(\nabla u(y_1), \nabla u(y_2)) \leq \frac{C}{\sigma} d(y_1, y_2), \tag{6.2.21}$$

where

$$\sigma := \min_{k=1,2} \{\sigma_0, d(y_k, a_k), d(y_k, b_k)\},$$

and the constants C and $\sigma > 0$ depend only on the compact set $T_0 \cup T$.

Proof. We choose σ_0 from Lemma 6.2.1, and assume $\sigma \leq \sigma_0$ in (6.2.1-6.2.4). Then we also have (6.2.5 – 6.2.11). In particular, every pair of points from $y_k, z_k, k = 1, 2$, is connected by a unique minimizing geodesic.

Denote

$$v_k = \sigma \nabla u(y_k) \in T_{y_k}\mathbb{H}, \ \ k = 1, 2. \tag{6.2.22}$$

Then $\|v_k\| = \sigma$. Let $\tilde{v}_2 \in T_{y_1}\mathbb{H}$ be the vector obtained by parallel translation of v_2 along a minimizing geodesic from y_2 to y_1. Then $\|\tilde{v}_2\| = \|v_2\| = \sigma$. Also,

$$d_{T\mathbb{H}}(v_2, \tilde{v}_2) = d(y_1, y_2).$$

Since the map $\exp : \mathcal{D} \to \mathbb{H}$ is smooth, it is locally Lipschitz as a map between the metric spaces $(T\mathbb{H}, d_{T\mathbb{H}})$ and $(\mathbb{H}, d)$. Let $C = C(T_0 \cup T)$ be its Lipschitz constant on the compact set $\mathcal{U} = \{(p, v) | p \in T_0 \cup T, v \in T_p\mathbb{H}, \|v\| \leq 2\sigma_0\} \subset T\mathbb{H}$. Then

$$d(\exp_{y_1} \tilde{v}_2, \exp_{y_2} v_2) \leq C d_{T\mathbb{H}}(v_2, \tilde{v}_2) = C d(y_1, y_2).$$

Note that $\exp_{y_k} v_k = z_k$, $k = 1, 2$. Thus, using the above inequality and Lemma 6.2.1, we obtain

$$d(\exp_{y_1} v_1, \exp_{y_1} \tilde{v}_2) \leq d(z_1, z_2) + d(z_2, \exp_{y_1} \tilde{v}_2) \leq C d(y_1, y_2).$$

By our choice of σ, the map $F : T\mathbb{H} \to \mathbb{H} \times \mathbb{H}$ defined by $F(p, v) = (p, \exp_p v)$ is a diffeomorphism from $\mathcal{U}$ to $F(\mathcal{U}) = \{(p, q) | p \in T_0 \cup T, q \in \mathbb{H}, d(p, q) < 2\sigma_0\}$. Thus the map F^{-1} is Lipschitz on $\{(p, q) | p \in T_0 \cup T, q \in \mathbb{H}, d(p, q) < \sigma_0\} \subset F(\mathcal{U})$, and there exists a constant C depending only on $T_0 \cup T$ such that for any $p \in T_0 \cup T$, $\xi, \eta \in T_p\mathbb{H}$, $\|\xi\|, \|\eta\| \leq \sigma_0$ we have

$$d_{T\mathbb{H}}(\xi, \eta) \leq Cd(\exp_p \xi, \exp_p \eta).$$

Using this estimate

$$d_{T\mathbb{H}}(v_1, \tilde{v}_2) \leq Cd(\exp_{y_1} v_1, \exp_{y_1} \tilde{v}_2) \leq Cd(y_1, y_2).$$

Combining this with estimates calculated above,

$$d_{T\mathbb{H}}(\sigma\nabla u(y_1), \sigma\nabla u(y_2)) \leq d_{T\mathbb{H}}(\sigma\nabla u(y_1), \tilde{v}_2) + d_{T\mathbb{H}}(\tilde{v}_2, \sigma\nabla u(y_2)) \leq Cd(y_1, y_2). \quad (6.2.23)$$

Since, for any $\xi, \eta \in T\mathbb{H}$ and $\lambda \in \mathbb{R}$,

$$d_{T\mathbb{H}}(\lambda\xi, \lambda\eta) \leq \max(1, |\lambda|) d_{T\mathbb{H}}(\xi, \eta)$$

and $\sigma \leq \sigma_0$, we conclude

$$d_{T\mathbb{H}}(\nabla u(y_1), \nabla u(y_2)) \leq \frac{1}{\sigma} d_{T\mathbb{H}}(\sigma\nabla u(y_1), \sigma\nabla u(y_2)) \leq \frac{C}{\sigma} d(y_1, y_2) \quad (6.2.24)$$

□

The directions of transport rays in $\mathbb{H}$ therefore vary Lipschitz continuously. In order for the initial and goal measures in Monge's Problem in $\mathbb{H}$ to be decomposed along the transport rays, it remains to show that the ends of the rays A and B are Lebesgue-negligible. Once this result is obtained, the optimal mapping in $\mathbb{H}$ can be built by performing the 1-dimensional problem along the rays.

7.0 EXAMPLES

Here we present a few examples of optimal mass transport in the 3-dimensional Heisenberg group. Let (x, y, z) be the coordinates in $\mathbb{H}$. Let $p(t) : [0, 1] \to \mathbb{H}$ be a geodesic with initial data $(v_x, v_y; a)$. We recall from Lemma 4.3.5:

$$\text{for } a = 0: \qquad \begin{aligned} x(t) &= v_x t, \\ y(t) &= v_y t, \\ z(t) &= 0; \end{aligned}$$

$$\text{and for } a \neq 0: \qquad \begin{aligned} x(t) &= \frac{v_x \sin at - v_y(1 - \cos at)}{a}, \\ y(t) &= \frac{v_y \sin at + v_x(1 - \cos at)}{a}, \\ z(t) &= \frac{v_x^2 + v_y^2}{2a^2}(at - \sin at). \end{aligned}$$

We talk of $v_x X + v_y Y$ being the initial velocity and a the (initial) acceleration of the curve.

7.1 MASS TRANSFER LIFTED FROM THE PLANE

Let μ be a (probability) distribution in $\mathbb{H}$, with compact support Ω_μ. Let $\pi : \mathbb{H} \to \mathbb{R}^2$ be the projection into the $xy-$plane. Consider the measures $\alpha = \pi_\# \mu$ and β in $\mathbb{R}^2$, each with compact support. In $\mathbb{R}^2$ we obtain a function $\hat{T} : \mathbb{R}^2 \to \mathbb{R}^2$ solving Monge's problem optimally mapping α into β. This mapping is built out of a 1-Lipschitz function $u : \mathbb{R}^2 \to \mathbb{R}$ and a function $a : \mathbb{R}^2 \to \mathbb{R}$ so that

$$\hat{T}(x) = x - a(x)\nabla u(x).$$

We lift this solution: let $T : \mathbb{H} \to \mathbb{H}$ be defined by

$$T(x) = x * [a(x)\nabla_{\mathbb{H}} u(x)]^{-1}.$$

This function pushes μ optimally to $T_{\#}\mu$, which projects to

$$\pi_{\#}T_{\#}\mu = \hat{T}_{\#}\pi_{\#}\mu = \hat{T}_{\#}\alpha = \beta.$$

7.2 VERTICAL DISTRIBUTIONS

The simplest class of examples whose solutions do not follow from lifted solutions are those which project into point masses. In this section we consider distributions along vertical line segments (more precisely, translations of closed intervals of the center $Z \leq \mathbb{H}$.)

7.2.1 MASS TRANSFER ON THE CENTER

Let μ, ν be two probability measures supported on the center $Z \leq \mathbb{H}$ (that is, the $z-$axis). We seek a mapping $f : Z \to Z$ pushing μ into ν minimizing the Carnot-Carathéodory distance d_{cc} of the Heisenberg group. The distance d between the origin and the point $(0, 0, z)$ is obtained from the geodesic parameterization, with $t = 1, a = \pm 2\pi, v_x = d, v_y = 0$:

$$|z| = \frac{d^2}{8\pi^2}(2\pi)$$

so that

$$d = \sqrt{4\pi|z|}.$$

This transportation is then equivalent to the one-dimensional case with cost function:

$$c(z_1, z_2) = \sqrt{4\pi|z_1 - z_2|}.$$

Note that this one-dimensional cost is a concave function of displacement.

Let $f : [0, 1] \to [1, 2]$ and $|f'| = 1$ (a.e.), so that f pushes the uniform measure onto the uniform measure. We seek such an f which minimizes

$$C = \int_0^1 \sqrt{4\pi |x - f(x)|}\, dx = \sqrt{4\pi} \int_0^1 \sqrt{f(x) - x}\, dx.$$

since the cost function is concave in $|f(x) - x|$, the optimal mapping is $f(x) = 2 - x$, for which $C = \frac{4}{3}\sqrt{2\pi}$. This is better than $f(x) = x + 1$ by a factor of $\sqrt{\frac{8}{9}}$.

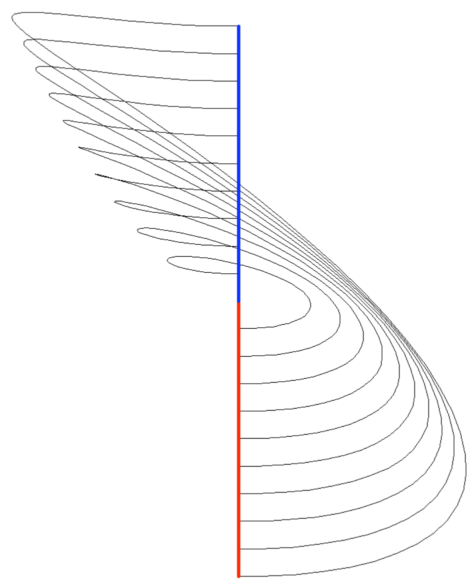

Figure 4: Optimal Transport Along Z

7.2.2 GENERAL UNIFORM DISTRIBUTIONS

Let $I = \{(0, 0, z) : 0 \le z \le L\}$ and $J = \{(r, 0, z) : r \in \mathbb{R}, b \le z \le b + L, b \in \mathbb{R}\} = p * I$, where $p = (r, 0, b)$. Consider uniform (probability) distributions on each. We seek a mapping $T : I \to J$ minimizing

$$C = \int_I d(q, T(q))\, d\lambda = \int_0^L d_0(r, 0, T_3(z) - z)\, dz,$$

where d_0 is the distance from the origin.

Fix $r > 0$. If $b = 0$, then the mapping is given as in section 7.1, and so $T(q) = p * q$ simply left translation. If b is large, then the situation becomes more like section 7.2.1, so we expect the origin to be mapped, not to p, but to $p * (0, 0, L)$. Let us then look for an optimal mapping amongst the family:

$$T_\alpha(0,0,z) = \begin{cases} (r, 0, b + L - z) & 0 \le z < \alpha \\ (r, 0, b - \alpha + z) & \alpha \le z \le L \end{cases}$$

For these functions,

$$C(\alpha) = \int_0^L d(r, 0, T_\alpha(z) - z)\, dz$$

where we abuse the notation for T. Separating the integral over the two pieces of I,

$$\begin{aligned} C(\alpha) &= \int_0^\alpha d(r, 0, b + L - 2z)\, dz + \int_\alpha^L d(r, 0, b - \alpha)\, dz \\ &= \int_0^\alpha d(r, 0, b + L - 2z)\, dz + (L - \alpha)\, d(r, 0, b - \alpha). \end{aligned}$$

We wish to find $0 \le \alpha \le L$ minimizing $C(\alpha)$, so differentiating:

$$\frac{\partial C}{\partial \alpha} = d(r, 0, b + L - 2\alpha) - d(r, 0, b - \alpha) - (L - \alpha)\, [Zd](r, 0, b - \alpha).$$

where the vector field $Z = \partial_z$.

Now, there exists numbers β_1, β_2 with $0 \le \beta_2 \le \beta_1 \le L - \alpha$ so that

$$\begin{aligned} \frac{\partial C}{\partial \alpha} &= (L - \alpha)\, [Zd](r, 0, b - \alpha + \beta_1) - (L - \alpha)\, [Zd](r, 0, b - \alpha) \\ &= (L - \alpha)\, \beta_1\, [Z^2 d](r, 0, b - \alpha + \beta_2) \end{aligned}$$

We then have critical values of $\alpha = 0, L$, and when $Z^2 d = 0$. Let a be the acceleration parameter for the geodesic from the origin to the point $(r, 0, b - \alpha + \beta_2)$. From equations (4.3.1), we have:

$$Z^2 d = Z\left(\frac{a}{d}\right) = \frac{dZa - aZd}{a^2} = \frac{a^3}{d^3}\left(\frac{1 + \cos a}{2\sin a - a(1 + \cos a)}\right)$$

This expression is zero (producing a critical value for C) when $|a| = \pi$.

In the case $\theta = \frac{\pi}{2}$,

$$(r, 0, b - \alpha + \beta_2) = \left(\frac{2d}{\pi}, 0, \frac{d^2}{2\pi}\right)$$

and

$$\alpha = b - \frac{\pi r^2}{8} + \beta_2.$$

Therefore, if $b \geq L + \frac{\pi}{8}r^2$, then $\alpha = L$ (wholly anti-parallel transportation) is optimal (as in section 7.2.1). Since $\beta_2 \leq L - \alpha$, if $b + L < \frac{\pi}{8}r^2$ then $\alpha = 0$ (wholly parallel transportation) is the optimal solution (as in section 7.1).

By a symmetric argument, in the case $\theta = -\frac{\pi}{2}$, if $b \geq L - \frac{\pi}{8}r^2$, then parallel transportation is optimal, and if $b + L < -\frac{\pi}{8}r^2$ then wholly anti-parallel transportation is the optimal solution.

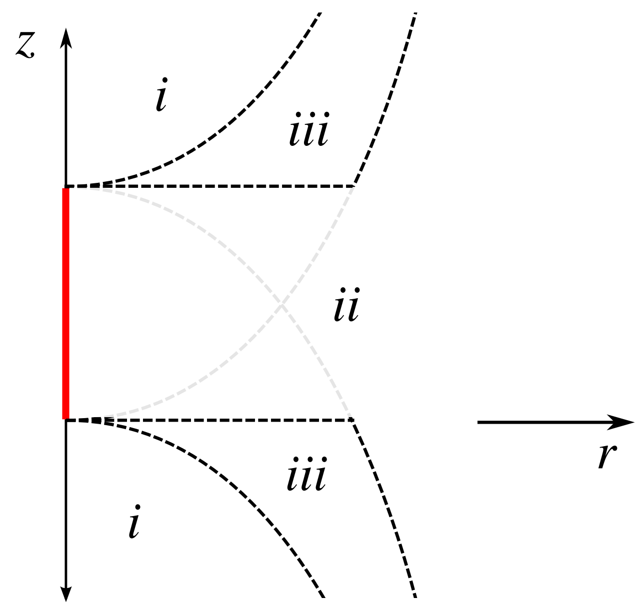

Figure 5: Regions for Transportation of Vertical Strips

If the destination segment is contained in (i) then the optimal solution is anti-parallel ($\alpha = L$), if it is contained in (ii) then the optimal solution is parallel ($\alpha = 0$), and if it touches (iii) then the optimal solution may be a mixture ($0 \leq \alpha \leq L$).

BIBLIOGRAPHY

[1] L. Ambrosio and S. Rigot, *Optimal mass transportation in the heisenberg group.*

[2] A. Bellaïche and J. Risler, *Sub-riemannian geometry*, Birkhäuser, 1996.

[3] L. Caffarelli, M. Feldman, and R.J. McCann, *Constructing optimal maps for monge's transport problem as a limit of strictly convex costs*, J. Amer. Math. Soc. **15** (2002), 1–26.

[4] L.C. Evans and W. Gangbo, *Differential equations methods for the monge-kantorovich mass transfer problem*, AMS Bookstore, 1999.

[5] M. Feldman and R.J. McCann, *Monge's transport problem on a riemannian manifold*, Trans. Amer. Math. Soc. **354** (2002), 1667–1697.

[6] N. Trudinger and X.-J. Wang, *On the mass transfer problem*, Calc. Var. Partial Differential Equations **13** (2001), no. 1, 19–31.

[7] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 54, American Mathematical Society, 2003.

[8] ______, *Optimal transport, old and new*, Springer, 2008.