May 19, 2019

by

J.H. Martel

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Mathematics
University of Toronto

Abstract

May 19, 2019

J.H. Martel
Doctor of Philosophy
Graduate Department of Mathematics
University of Toronto
2019

We describe new applications of optimal transportation to algebraic topology, and develop a method for building Spines/Souls using a definition of Singularity based on Kantorovich duality and semicouplings programs. Given an initial manifold-with-corners X, we construct continuous homotopy-reductions from X onto large-codimension subvarieties $\mathscr{Z} \hookrightarrow X$. The subvarieties $\mathscr{Z}$ are assembled from a contravariant functor $Z : 2^Y \to 2^X$, where $Z = Z(c, \sigma, \tau)$ is defined by a source (X, σ), target (Y, τ), and cost $c : X \times Y \to \mathbb{R}$. Our Theorems 3.1.1, 4.4.3, 4.4.4 describe criteria for inclusions $Z(Y_I) \hookrightarrow Z(Y_J)$ to be homotopy-isomorphisms and find a maximal index $J \geq 0$ for which $(\mathscr{Z} := Z_{J+1}) \hookrightarrow X$ is a codimension-J homotopy-isomorphism. Our Theorem 1.5.1 is an application which exhibits new small-dimensional $E\Gamma$ classifying spaces where Γ is finite-dimensional Bieri-Eckmann duality group, e.g. Spines of $\Gamma = PGL(\mathbb{Z}^2)$, $PGL(\mathbb{Z}^3)$, etc. The singularity functor Z is a new reduction theory for $E\Gamma$ models based on a convex interpretation of Bieri-Eckmann's homological duality and solutions to a homological subprogram we call "Closing the Steinberg symbol". The subprogram replaces X with a chain sum $\underline{F}$ having "well-separated gates", and on which Γ freely acts as shift operator. We also construct repulsion costs c and visibility costs v on $\underline{F}$ whose functors $Z(c, \sigma, \tau)$ are conjectured to satisfy sufficient conditions to produce large-codimension Γ-equivariant homotopy-reductions $\mathscr{Z}$ of X.

Acknowledgements

Contents

1 Introduction **1**

1.1 Kantorovich: the Bridge from Measure to Topology . . . 2

1.2 Reduction-to-Singularity Principle . . . 4

1.3 Cost Assumptions . . . 5

1.4 Deformation Retracts and Theorems 1.4.1, 1.4.2 . . . 8

1.5 Closing the Steinberg symbol and Theorem 1.5.1 . . . 11

1.6 Thesis Outline . . . 17

1.7 Conventions and Notations . . . 19

2 Background: Semicouplings and Kantorovich Duality **21**

2.1 Optimal Semicouplings and Cost . . . 22

2.2 Existence of c-Optimal Semicouplings . . . 23

2.3 Kantorovich Duality . . . 25

2.4 Uniqueness of Activated Domain . . . 30

2.5 Uniqueness of Optimal Semicouplings . . . 33

3 Deforming Source onto Active Domain $X \rightsquigarrow A$ **38**

3.1 Statement of Theorem 1.4.1 . . . 38

3.2 Averaged Gradients and Finite-Time Blow-Up . . . 40

3.3 Proof of Theorem 1.4.1 . . . 44

4 Kantorovich Singularity and Topological Theorem 1.4.2 **46**

4.1 Kantorovich's Contravariant Singularity Functor . . . 46

4.2 Local Topology and Local Dimensions of Z . . . 49

4.3 The Descending Filtration $Z_j, j = 0, 1, 2, \dots$ . . . 53

4.4 Local–Global Homotopy Reductions: (UHS) Conditions . . . 57

5 Repulsion Costs **63**

5.1 Chain sums and Well-Separated Gates . . . 63

5.2 Hypotheses and Practical Applications . . . 64
5.3 Repulsion Costs $c|\tau$, c_1 . . . 67
5.4 Two-Pointed Repulsion Costs c_2 . . . 71
5.5 Convex-Excisions . . . 73
5.6 Visibility . . . 74
5.7 ϵ-Regularizations . . . 77
5.8 Barycentres and Krein-Milman . . . 79
5.9 Visible Repulsion Costs . . . 80

6 Excisions and Cohomological Dimension **85**
6.1 Geometric Classifying Spaces $E\Gamma$ . . . 85
6.2 Background: Group Cohomology . . . 87
6.3 Excision versus Compactification . . . 90
6.4 Bieri-Eckmann Duality . . . 94
6.5 Arithmetic Groups: Excision . . . 97

7 Closing the Steinberg symbol **101**
7.1 Stitching Footballs from Regular Panels: Motivation . . . 101
7.2 Closing Steinberg: Definition and Consequence . . . 103
7.3 Closing Steinberg on $PGL(\mathbb{Z}^2)$: First Example . . . 106
7.4 Closing Steinberg on $GL(\mathbb{Z}^3)$: Example . . . 108

8 Conclusion **113**

Bibliography **114**

List of Figures

1.1 When $\partial(A+B+C)=0 \pmod 2$, we define $F=conv(A,B,C)$ 14
1.2 A chain sum $\underline{\mathrm{F}}=\sum_{i\in I} F_i$ with well-separated one-dimensional gates 15

2.1 Disconnected Active Domain 33
2.2 Connected Active Domain when $mass[\sigma]/mass[\tau] \approx 1^+$ 33

3.1 Halfspace Conditions fail and active domain is not homotopy-equivalent to source. 39

4.1 Singularity structure Z on active domain $A=Z_1$ 48
4.2 Singularity structure on Z_2 48
4.3 Horospherically truncated pair of pants, with active domain relative to a repulsion cost (See Definition 5.3.2 from Section 5.3). Halfspace Conditions fail. $P \rightsquigarrow Z_1 \rightsquigarrow Z_2$ are homotopy equivalent, where P is pair of pants. But Z_2 does not deformation retract to Z_3 62

5.1 Singularity structures of $\widetilde{c}$ on regular 2-simplex and regular square satisfy properties (D0)–(D5) 66
5.2 The yellow line designates the one-dimensional gate G. The blue line designates the interpolated singularity structure with respect to the gated-visibility cost v^*. 70
5.3 Graph of repulsion cost $x \mapsto c|\mathscr{H}_Y(x,0)$. The assymmetry reflects the "home-preference" of x to $y=0$. 70
5.4 Graph of repulsion cost $x \mapsto c|\mathscr{H}_Y(x,(1,0))$. The "home-preference" implies the pole at $y=(1,0)$ has smaller diameter than the other poles. . . . 71
5.5 Illustrating the projection map $proj: F \to \partial F$ defining two-pointed repulsion cost c_2 for ellipse F where $\partial F = \mathscr{E}[F]$. 73
5.6 Excision of convex ellipsoid 75
5.7 Excision domain $X=F[t]$, where (x,y) are visible in X and (x,w) is not visible in X 76

5.8 The outermost boundary is a $C^{1,1}$-regularization of Y. The inner hypersurface is a smooth ϵ-regularization Y_ϵ 78
5.9 The subdomain Ω is geodesically convex 82
5.10 Evaluating visibility cost $v(x, z)$ for pair $(x, z) \in \Omega \times Y_\epsilon$ 84

6.1 (a) Deep Zones are disjoint from Γ-orbit. (b) Excise the Deep Zones from X. (c) Optimal Semicoupling, with respect to repulsion cost, between target excision boundary (in blue) and activated source measure (in red). (d) Singularity structure (in yellow) of an optimal semicoupling between source and target with respect to repulsion cost. 92

7.1 Stitching a football F from identical regular hexagons or pentagons P_i . . 102
7.2 Adidas' "Telstar" design is football stitched from white hexagon and black pentagon panels. A football can also be stitched from black triangular and white pentagonal panels. 102
7.3 Convex excision $X[t]$ relative a Γ-equivariant excision parameter $t : \mathbb{Q}P^1 \to \mathbb{R}$ 108
7.4 Steinberg symbol in $X[t] \times \partial X[t]$ is represented as relative 1-cycle P with boundary ∂P equal to 0-sphere. 109
7.5 The translates by $I_0 = \{Id, \gamma, \gamma'\}$ successfully Close the Steinberg symbol. 109
7.6 Evaluating the repulsion cost relative to the repulsion-cost c_2 at various source points $x, x', x'', \ldots$ etc. in $X[t]$ 110
7.7 Active Domain for optimal semicoupling with respect to repulsion cost is homotopy-equivalent to the source, c.f. Theorem 1.4.1 110
7.8 Singularity Structure of optimal semicoupling. The active domain is homotopy-equivalent to one-dimensional tree, c.f. Theorem 4.4.4 111

Chapter 1

Introduction

Optimal transportation is a subject with a remarkably wide range of applications. Our thesis develops new applications of optimal transportation to algebraic topology. Everything proceeds from a definition of "Singularity" based on Kantorovich duality, which is the fundamental linear duality of mass transport. This thesis indicates a new construction which has the potential to the yield the spines long sought by geometers (e.g. [Ash84], [Thu86]), and such applications were the original motivation for this research.

Let X be an oriented Riemannian manifold-with-corners with boundary $Y = \partial X$. A topologist initially interprets Singularity as the "locus-of-discontinuity" of deformation retract maps $r : X \to Y$, but this definition is not topological since *continuous* retracts r from X to Y do not exist. However optimal transportation is useful setting to formally define Singularity in the category of Topology. This formalization is summarized by contravariant functors $Z : 2^Y \to 2^X$ between the categories of closed subsets of X, Y. The contravariant functor Z is defined via maximizers of the dual program to semicouplings programs $Z = Z(\sigma, \tau, c)$, introduced by [HS13, 1.(d)].

Mass transport motivates an economic definition: we say Singularity arises wherever there is competition for limited common resources. Formally we topologize this definition using duality of c-optimal semicouplings, where c is a choice of cost $c : X \times Y \to \mathbb{R}$ on a given source space (X, σ), target space (Y, τ), and defined whenever the source σ is *abundant* with respect to the prescribed target τ, i.e.

$$\int_X \sigma \geq \int_Y \tau. \tag{1.1}$$

Kantorovich duality characterizes the topology of Singularity in a contravariant functor $Z = Z(\sigma, \tau, c) : 2^Y \to 2^X$, defined by $Z(Y_I) = \cap_{y \in Y_J} \partial^c \psi(y)$ whenever equation (1.1) is satisfied. Here $\partial^c \psi(y)$ designates the c-subdifferential of a c-concave potential $\psi^{cc} =$

$\psi : Y \to \mathbb{R} \cup \{-\infty\}$. The formal definitions of c-concavity and $\partial^c\psi(y)$ is provided in 2.3.2–2.3.3. Briefly the c-subdifferential $\partial^c\psi(y)$ consists of all $x \in X$ for which

$$y \in argmax\{\psi(y') - c(x, y') + c(x, y) \mid y' \in Y\}.$$

The basic theory of c-concavity and c-subdifferentials is nonlinear analogue of the convexity and subdifferentials of convex lower semicontinuous functions à la Fenchel, Alexandrov, etc., for the choice of quadratic cost $c = dist^2/2$. See [ET99], [Vil09, Ch.5] for standard definitions.

Our end goal is to apply our general theory to concrete examples and build new explicit "Spines" and "Souls" of various spaces of interest to geometers and group theorists alike. As such, this thesis has two phases. The first phase is general, and will identify hypotheses on costs $c : X \times Y \to \mathbb{R}$ which ensures the existence of large-codimension homotopy reductions. Our main results are Theorems 3.1.1, 4.4.4, described below. In this first phase we thus introduce some general principles of Reduction-to-Singularity, as arising from duality of mass transport. The second phase is practical, and concerns the numerical applications of these general theorems. Our main practical results are Theorem 1.5.1 and the introduction of a subprogram we call "Closing the Steinberg symbol" (Chapter 7), and specially designed for constructing small-dimensional $E\Gamma$-classifying spaces when Γ is a finite-dimensional Bieri-Eckmann duality group. Essentially, Closing Steinberg is a combinatorial obstruction to applying our general theorems in specific instances. We provide some proofs-of-concept of all these ideas in §§7.3–7.4 for the groups $\Gamma = PGL(\mathbb{Z}^2)$, $PGL(\mathbb{Z}^3)$. Finally we include several conjectures, especially Conjecture 1.5.2, which assert that specific costs (namely visible repulsion costs $c = v$, §5.9.6) satisfy sufficient hypotheses of Theorems 3.1.1 and 4.4.4 to produce new examples of minimal spines of various $E\Gamma$ models.

1.1 Kantorovich: the Bridge from Measure to Topology

The symbols X, Y are reserved for finite-dimensional Riemannian manifolds-with-corners. We let σ, τ denote Radon measures on X, Y, called the "source" and "target" measures. Typically σ, τ are absolutely continuous with respect to the Riemannian volume measures on X, Y. Or equivalenty, absolutely continuous with respect to the Hausdorff measures $\mathscr{H}_X$, $\mathscr{H}_Y$.

The present thesis develops a bridge between measure theory and algebraic topology, as inspired from Kantorovich duality in optimal transport. The bridge is categorical, being a contravariant functor $Z = Z(c, \sigma, \tau)$ defined by a cost $c : X \times Y \to \mathbb{R}$ between source (X, σ) and target (Y, τ) measure spaces. If $2^X, 2^Y$ denote the category of closed subsets of X, Y respectively, then Z can be represented as a contravariant correspondance $Z : 2^Y \to 2^X$ between closed topological subsets of Y and X. The contravariance of Z means morphisms $Y_I \hookrightarrow Y_J$ between objects of 2^Y are mapped by Z to morphisms $Z(Y_I) \hookleftarrow Z(Y_J)$ between closed subsets (i.e. objects) $Z(Y_I), Z(Y_J)$ of 2^X. See Chapter 4 and Definition 4.1.1 for details.

Contravariance has concrete consequences, producing explicit equations describing singular chains. For X a topological space, one obtains the category 2^X whose objects X_I are the closed subsets of X, and whose morphisms are the inclusions $X_I \hookrightarrow X_J$ between closed subsets X_I, X_J of X when such inclusions exist. We use the functor Z to parameterize closed subsets $Z(Y_I) \hookrightarrow X$ according to closed subsets $Y_I \hookrightarrow Y$, and where inclusions $Y_I \hookrightarrow Y_J$ correspond to the reverse inclusions $Z(Y_I) \hookleftarrow Z(Y_J)$. Assembling these elementary inclusions, we find a new "cellular" decomposition $\{Z(Y_I)\}_{Y_I}$ of our source space X, and contravariantly parameterized by closed subsets Y_I of the target space Y. Thus we propose Kantorovich's bridge as a contravariant functor $Z : 2^Y \to 2^X$ and a new measure-theoretic tool for explicitly constructing topologically-nontrivial subvarieties. Our thesis illustrates the idea expressed in [Gro14a, §5.3] that "singular spaces" be replaced by contravariant functors between suitable categories. In short, a "singular space" is not a space but an object, namely a contravariant functor. These functors Z depend on the source σ, target τ, choice of continuous cost function $c : X \times Y \to \mathbb{R}$.

Our main theorems identify a local condition, which we call uniform Halfspace (UHS) conditions (see 4.4.2), and our main topological Theorems 3.1.1 and 4.4.4 identify an index $J \geq 0$ for which the source space X can be continously reduced via strong deformation retract to a codimension-J subvariety $Z_{J+1} \hookrightarrow X$. This is the main topological application of our thesis. The existence of effective homotopy reductions is a fundamental problem in algebraic-topology, and especially in homological computations on large-dimensional manifolds arising as $E\Gamma$-models, where Γ is an infinite discrete (torsion-free) group. We propose a reduction-to-singularity $X \rightsquigarrow \underline{Z}$ method constructing efficient small-dimensional classifying spaces, c.f. Chapters 6, 7, and §1.2 below. A general homotopy-reduction procedure is expressed in Theorem 1.5.1, which follows from our Closing Steinberg symbol construction (Definition 7.2.1 and Theorem 7.2.3).

1.2 Reduction-to-Singularity Principle

This section reinterprets some standard facts from algebraic topology. The author first learned the subject from [GJ81]. Standard references include [Bre93], [Bro82]. Algebraic topology is foremost based upon the singular homology functors $H_i(-;\mathbb{Z}) : TOP \to MOD_{\mathbb{Z}}$, $i \in \mathbb{Z}$, which is the covariant functor $X \mapsto H_*(X) = H_*(X;\mathbb{Z})$ from the category of nonpointed topological spaces to the category $MOD_{\mathbb{Z}}$ of $\mathbb{Z}$-modules (additive abelian groups). The homology groups are defined on topological spaces X according to the singular chain complexes $\{C_*^{sing}(X;\mathbb{Z}), \partial_*\}$, and the contravariant cohomology functors $H^*(-;\mathbb{Z}) : TOP \to MOD_{\mathbb{Z}}$ are defined via the cochain complexes $\{C^*(X;\mathbb{Z}),\mathbb{Z}), \delta^*\}$ where $C^*(X;\mathbb{Z}) = Hom(C_*^{sing}(X;\mathbb{Z}),\mathbb{Z})$. The first definitive computations distinguish the one-dimensional sphere S^1 from the one-dimensional line $\mathbb{R}^1$.

Next one deduces the nonexistence of continuous deformation retracts from the two-dimensional disk D to its boundary ∂D. It is useful to emphasize the formal negative nature of this previous sentence. The expression "there does not exist continuous deformation retracts $D \times I \to \partial D$" is of course true in TOP. It is a logical deduction based on the following algebraic observation: the identity morphism $Id : \mathbb{Z} \to \mathbb{Z}$, defined by $n \mapsto n$ of the additive abelian group $\mathbb{Z}$ is distinct from the zero morphism $0_{\mathbb{Z}} : \mathbb{Z} \to \mathbb{Z}$, $n \mapsto 0$. Yea $Id_{\mathbb{Z}} \neq 0_{\mathbb{Z}}$ in $Hom(\mathbb{Z},\mathbb{Z})$. More formally, recall the following useful definition.

Definition 1.2.1. Let X be a topological space, and $Y \hookrightarrow X$ be a subset. We say Y is a strong deformation retract of X (or, X deformation retracts onto Y) if there exists a continuous mapping $r : X \times [0,1] \to X$ with the following properties:

(i) for all $x \in X$, we have $r(x,1) \in Y$ and $r(x,0) = x$;

(ii) for all $y \in Y$ and $t \in [0,1]$, we have $r(y,t) = y$.

If X deformation retracts onto Y, then the inclusion $i : Y \to X$ is a homotopy isomorphism and induces an isomorphism $H_*(i) : H_*(Y;\mathbb{Z}) \to H_*(X;\mathbb{Z})$ of $\mathbb{Z}$-modules. A continuous retraction $r : X \times [0,1] \to X$ produces a continuous mapping $r(x) := r(x,1) : X \to Y$, and this mapping induces an isomorphism $H_*(r) : H_*(X) \to H_*(Y)$ which is inverse to $H_*(i)$.

Lemma 1.2.2. *Let X be a connected, oriented, aspherical topological space, and let $Y \subset X$ be a homologically nontrivial closed subspace. Then there exists no continuous deformation retracts from X onto Y.*

Proof. The hypotheses imply $H_*(X) = 0$ and $H_*(Y) \neq 0$. Suppose $r : X \to Y$ is a continuous retraction. Applying homology functors we find the composition $H_*(r \circ i)$

$= H_*(r) \circ H_*(i)$ defines an endomorphism (algebraic self-mapping) of $H_*(Y;\mathbb{Z})$. Since X is homologically trivial, the induced mapping $H_*(i)$ has zero image, and we find $H_*(r \circ i) = H_*(r) \circ H_*(i)$ coincides with the zero endomorphism of $H_*(Y;\mathbb{Z})$. However if r is a strong deformation retract of X onto Y, then the composition $r \circ i$ coincides with the identity mapping $Id_Y : Y \to Y$. Therefore $H_*(r \circ i) = H_*(Id_Y)$ coincides with the identity automorphism of $H_*(Y;\mathbb{Z})$ by functoriality. This is a formal contradiction unless Y is homologically trivial. □

The hypotheses are satisfied when X is contractible, and $Y = \partial X$. The topologist applies this reductio ad absurdum to deduce the following: if an oriented aspherical space X has homologically nontrivial boundary $Y = \partial X$, then there exists no continuous deformation retracts from X onto the boundary ∂X. Thus some essential obstruction exists. We propose the singularity loci constructed below are geometric forms of this obstruction.

Here is the motivating example. Let $X = D^2$ be two-dimensional unit disk, with boundary $Y = \partial X = S^1$. Suppose the topologist attempts to construct a simple deformation retract from X to the boundary S^1 and proposes the radial projection $r(x) = x \cdot |x|^{-1}$. Then r corresponds to a "mapping" $X \to Y$ which is almost continuous, but having locus-of-discontinuity a single point $\{o\}$. Observe the inclusion of the locus-of-discontinuity $\{o\} \hookrightarrow X$ is a homotopy-isomorphism, i.e. the singleton $\{o\}$ is homotopic to the disk X. We claim this homotopy-isomorphism is no coincidence, but rather indicates a general principle.

1.3 Cost Assumptions

Our strategy replaces continuous retracts $r : X \to \partial X$ with c-optimal semicouplings π from source (X, σ) to target $(\partial X, \tau)$. We present the basic theory and definitions of the semicoupling program in Chapter 2. Briefly, a semicoupling from a source (X, σ) to target $(Y := \partial X, \tau)$ is a Borel measure π on the Cartesian product $X \times Y$ whose marginals satisfy $proj_X \# \pi \leq \sigma$ and $proj_Y \# \pi = \tau$. Here $proj \# \pi$ denotes the pushforward measure of π by projection $proj : X \to X$, defined $proj \# \pi[U] = \pi[proj^{-1}(U)]$. The set of all semicouplings between source σ and target τ is a weak-$*$ compact subset $\Pi_{SC}(\sigma, \tau)$. A c-optimal semicoupling is a semicoupling which minimizes the total cost of transport with respect to c on $\Pi_{SC}(\sigma, \tau)$; see Theorems 2.3.5 and §2.2.

As we've seen, continuous retracts $r : X \to \partial X$ are nonexistant; but optimal semicouplings generally exist whenever $\sigma[X] \geq \tau[Y]$. We propose semicouplings as measure-theoretic alternatives to retractions from (X, σ) to $(\partial X, \tau)$. We measure the total cost of transport with respect to functions $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$. The function $c = c(x, y)$ represents the cost of transporting a unit mass at source x to target at y. That $\dim(X) > \dim(\partial X)$ has important consequences throughout our thesis, especially concerning (Twist) conditions 2.5.1.

Topology forces the nonexistence of continuous retracts. Similarly the geometry of the cost c controls the topology of the locus-of-discontinuity of c-optimal semicouplings. For topology to emerge from the measure theory, we require geometric assumptions on the cost. The proofs of our Theorems 3.1.1, 4.4.4, 1.5.1 below require cost functions $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ satisfying basic assumptions labelled (A0), ..., (A6). We abbreviate $c_y(x) := c(x, y)$. If $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is a function, we let $dom(f) := \{x \in X \mid f(x) \in \mathbb{R}\}$. We assume X is Riemannian such that functions $f : X \to \mathbb{R}$ have well-defined gradients $\nabla_x f$ with respect to the x-variable. The assumptions are the following:

(A0) The cost function $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is continuous throughout $dom(c) \subset X \times Y$ and uniformly bounded from below, e.g. $c \geq 0$. Moreover we assume the sublevels $\{x \in X \mid c(x, y) \leq t\}$ are compact subsets of X for every $t \in \mathbb{R}$, and $y \in Y$. That is, we assume $x \mapsto c(x, y)$ is *coercive* for every $y \in Y$.

(A1) The cost is twice continuously differentiable with respect to the source variable x, uniformly in y throughout $dom(c)$. So for every $y \in Y$, the Hessian function $x \mapsto \nabla^2_{xx} c(x, y)$ exists and is continuous throughout $dom(c_y)$.

(A2) The function $(x, y) \mapsto ||\nabla_x c(x, y)||$ is upper semicontinuous throughout $dom(c)$. So for every $t \in \mathbb{R}$ the superlevel set $\{||\nabla_x c(x, y)|| \geq t\}$ is a closed subset of $dom(c)$.

(A3) For every $y \in Y$, we assume $x' \mapsto \nabla_x c(x', y)$ does not vanish identically on any open subset of $dom(c_y)$.

(A4) The cost satisfies (Twist) condition with respect to the source variable throughout $dom(c)$. So for every x' the rule $y \mapsto \nabla_x c(x', y)$ defines an injective mapping $dom(c_{x'}) \to T_{x'} X$. See Definition 5.3.3.

(A5) For every $x \in X$, the function $y \mapsto c(x, y)$ is continuously differentiable; and for every $y \in Y$, the gradients $\nabla_y c(x, y)$ are bounded on compact subsets $K \subset X$.

A further assumption (A6) is convenient for the deformation retracts constructed in the present thesis. These retracts depend on the nonvanishing of certain "averaged" vector fields denoted $\eta_{avg}(x) \in T_xX$. The field η_{avg} is an average of a Y-parameter family of potentials $\eta(x,y)$ defined with respect to c-concave potentials $\psi^{cc} = \psi$, and parameters $\beta > 0$ by formulas

$$\eta(x,y) := |\psi(y_0) - \psi(y) - c_\Delta(x, y_0, y)|^{-\beta} \nabla_x (c(x, y_0) - c(x,y)) \tag{1.2}$$

for $y_0 \in \partial^c\psi(x)$, $y \in Y$. Compare (3.9), (3.12),4.4.2.

The formal definition of $\eta_{avg}(x)$ depends on the setting. Typically there is a Radon measure $\bar{\nu}_x$ on Y, depending on x, absolutely continuous with respect to $\mathscr{H}_Y$, and with average

$$\eta_{avg}(x) := (\bar{\nu}_x[Y])^{-1} \int_Y \eta(x,y) d\bar{\nu}(y). \tag{1.3}$$

(A6) The averaged vectors $\eta_{avg}(x)$ (1.3) are bounded away from zero, uniformly with $x \in Z'$, on the relevant subsets Z' of X. (See (3.9), (3.12), 4.4.2) and the hypotheses of Theorems 3.1.1, 4.4.3, 4.4.4).

In applications, the relevant subsets Z' are the nonactive domains $X - A$, or $Z_j - Z_{j+1}$, in the notation of Chapters 2–4. More previse formulations of (A6) are given in 3.2.2, (3.6), and the Uniform Halfspace (UHS) Conditions defined in 4.4.2. The assumption (A6) depends on properties of c-concave potentials defined on Y, and is not an absolute assumption on the geometry of the cost c like the previous (A0)–(A5).

Consequences of Assumptions (A0), (A1), ... will be clarified as our thesis progresses. In the simplest case where the source and target spaces X, Y are compact, and c is smooth and finite-valued throughout $X \times Y$, then Assumptions (A0)–(A2) are readily confirmed. Assumption (A3) forbids the cost $c(x,y)$ from being locally constant on any open subset of X. Assumption (A4) is more discriminate: the continuously differentiable functions $x \mapsto c_\Delta(x; y, y')$ necessarily have critical points if X is closed compact space. The Assumption (A4) implies the cross-differences $\nabla_x c_\Delta(x, y, y')$ are nonzero for distinct y, y', where $c_\Delta(x; y, y') := c(x,y) - c(x,y')$ is the two-pointed cross difference. In practice, costs c which have poles, e.g. $c(x,y) = +\infty$ when $x = y$, will more readily satisfy (A4). We will prove (A0)–(A4) implies the general uniqueness of c-optimal semicouplings. Peculiar to the semicoupling setting is (A4), which requires injectivity of mappings $y \mapsto \nabla_x c(x,y) : Y \to T_{x'}X$ for every $x \in X$, where $\dim(Y) < \dim(X)$. The Assumption (A5) is useful in the construction of our deformation retracts in Theorems 4.4.3–4.4.4 below.

1.4 Deformation Retracts and Theorems 1.4.1, 1.4.2

We continue to assume X, Y are Riemannian manifolds-with-corners, and suppose $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is a cost satisfying Assumptions (A0)–(A4). Let $\mathscr{H}_Y$ be the Hausdorff measure on the Riemannian manifold-with-corners Y.

For abundant source σ and target τ, there exists a unique c-optimal semicoupling π_{opt} from σ to τ. See §§2.4–2.5 below. Uniqueness results also established in [HS13], [CM10]. According to Kantorovich's duality theorem (§2.3), there exists c-concave potentials $\psi^{cc} = \psi : Y \to \mathbb{R} \cup \{-\infty\}$ whose c-subdifferential $\partial^c \psi : Y \to 2^X$ is uniquely prescribed by π_{opt}. Thus π_{opt} is supported on the graph of a measurable mapping $\partial^c \psi^c$ in $X \times Y$. This is important idea of optimal transport and developed further in §2.2 below.

The c-optimal semicoupling will not "activate" all the source measure for transport to τ whenever $\sigma[X] > \tau[Y]$. Equivalently, the domain $dom(\psi^c)$ is a nontrivial subset of the source X. Informally, c-optimal semicouplings will first allocate as much as possible from low-cost regions of the source. The union of all these activated low-cost regions defines a closed domain designated $A \hookrightarrow X$. Specifically we define $A := \cup_{y \in Y} \partial^c \psi(y)$.

Our first Theorem 1.4.1 describes a criteria to ensure the activated source $A \hookrightarrow X$ includes into the source as homotopy-isomorphism.

Theorem 1.4.1. *Let c be cost satisfying Assumptions (A0)–(A4). Suppose the source σ and target measure τ are absolutely continuous with respect to $\mathscr{H}_X$, $\mathscr{H}_Y$, respectively and satisfy (1.1). Let π be a c-optimal semicoupling from σ to τ, with dual c-concave potential $\psi^{cc} = \psi$ (2.3.5). Let A be the active domain (2.4.7). Let $\beta := \dim(Y) + 2$.*

Suppose every $x \in X - A$ has the property that

$$\eta_{avg}(x) := (\mathscr{H}_Y[Y])^{-1} \int_Y (c(x,y) - \psi(y))^{-\beta} \cdot \nabla_x c(x,y) \, d\mathscr{H}_Y(y), \tag{1.4}$$

is bounded away from zero, uniformly with respect to $y \in Y$. Then the inclusion $A \hookrightarrow X$ is a homotopy-isomorphism, and there exists explicit strong deformation retract $h : X \times [0,1] \to X$ of X onto $A = h(X, 1)$.

The Assumption (A6) is key technical hypothesis for constructing the deformation retract in Theorem 1.4.1. In the setting of 1.4.1, Assumption (A6) takes the following form. Let A be the active domain of a c-optimal semicoupling. Then (UHS) are satisfied if the average gradient $\eta_{avg}(x)$ is bounded away from 0, uniformly for $x \in X - A$. We remark that $\eta_{avg}(x)$ is certainly nonzero whenever the gradients $\nabla_x c(x,y)$, $y \in Y$, occupy a nontrivial halfspace for every x. Or equivalently, whenever the closed convex hull of $\nabla_x c(x,y)$, $y \in Y$, is disjoint from the origin $\mathbf{0}$ in $T_x X$.

Without the estimate $||\eta_{avg}(x)|| \geq C > 0$, our methods could only conclude that the source X deformation retracts onto ϵ-neighborhoods $A_\epsilon = \{x \in X | d(x, A) < \epsilon\} \hookrightarrow X$ of $A \hookrightarrow X$, where $\epsilon > 0$ is a sufficiently small real number.

Theorem 1.4.1 has the following application to quadratic costs

$$b(x, y) := d(x, y)^2/2 = ||x - y||^2/2 \tag{1.5}$$

for closed subsets X, Y in a Euclidean space $\mathbb{R}^N$. The gradients $\{\nabla_x c(x_0, y)\}_{y \in Y}$ are a subset of $T_{x_0}X$. Observe that the closed convex hull $conv\{\nabla_x c(x_0, y)\}_{y \in Y}$ contains the origin $\mathbf{0}$ in $T_{x_0}X$ if and only if x_0 lies in the convex hull $conv(Y)$ of Y in $\mathbb{R}^N$.

We illustrate with an example. Let $\sigma = \mathscr{L}$ be a Lebesgue measure on $\mathbb{R}^N$, and let $\tau = (100)^{-1} \sum_{i=1}^{100} \delta_{y_i}$ be an empirical measure (normalized sum of Dirac masses on $\mathbb{R}^N$). Evidently (1.1) is satisfied. Consider the restriction of b (1.5) to $\mathbb{R}^2 \times Y$. There exists unique b-optimal semicoupling from σ to τ, and let $A \subset \mathbb{R}^2$ be the active domain. The active domain A is a union of (possibly overlapping) Euclidean balls. The non-active domain $\mathbb{R}^N - A$ is an unbounded open subset of $\mathbb{R}^N$. Under the hypotheses of 1.4.1, we define an averaged potential $f_{avg} : \mathbb{R}^N - A \to \mathbb{R}$ which has property that both

$$f_{avg}(x_k) \to +\infty \quad \text{and} \quad ||\nabla_x f_{avg}(x_k)|| \to +\infty$$

whenever x_k is a sequence in $\mathbb{R}^N - A$ converging to $\lim_k x_k = x_\infty \in \partial A$. See (3.3) for definition of f_{avg}. Now the hypotheses of 1.4.1 require f_{avg} have no critical points on the open subset $\mathbb{R}^N - A$. The nonexistence of critical points can be achieved by a simple observation: if $A \supset conv(Y)$, then for every $x \in \mathbb{R}^N - A$ the gradients

$$\nabla_x c(x, y), \quad y \in Y$$

occupy a nontrivial Halfspace of $T_x\mathbb{R}^N$. This implies $\nabla_x f_{avg}$ is uniformly bounded away from zero.

On the other hand, the hypotheses of 1.4.1 are never satisfied when we restrict b (1.5) to a convex subset $X := F$ and its boundary $Y := \partial F$. If $\sigma = 1_F\mathscr{L}$, and if the target τ is supported on ∂F, and if (1.1) is satisfied with strict inequality, then the active domain A of the unique b-optimal semicoupling from σ to τ will not contain $conv(Y)$. Consequently η_{avg} will vanish somewhere on $X - A$, and the hypotheses of 1.4.1 are violated. In Chapter 5 we define a repulsion cost $c|\tau$ which will satisfy (UHS) conditions throughout the non active domains and satisfy hypotheses of 1.4.1.

Now Theorem 1.4.1 is but a first step. Our next Theorem 1.4.2 constructs further homotopy-reductions from the active domains $A \subset X$ to higher codimension closed subvarieties $\mathscr{Z} \hookrightarrow A$. We use "subvariety" to mean a subset described by the vanishing of a collection of twice-continuously differentiable functions. In applications the functions will even be smooth.

If σ, τ are source, target measures satisfying (1.1), then there exists c-optimal semi-couplings from σ to τ. Kantorovich duality 2.3.5 implies the existence of c-concave potentials $\psi^{cc} = \psi$ on the target Y. Let $Z : 2^Y \to 2^X$ be the corresponding singularity functor $Z(Y_I) = \cap_{y \in Y_I} \partial^c \psi^c(x)$. For integers $j \geq 1$, we define Z_j to be the subset of $x \in X$ where the local tangent cone is at least j-dimensional at x (the formal definition is provided in §4.3, c.f. 4.2.3, 4.3.4). The singularity structure is naturally cellulated by the cells

$$Z'(x) := Z(\partial^c \psi^c(x))$$

and admits a filtration

$$(X =: Z_0) \hookleftarrow (A =: Z_1) \hookleftarrow Z_2 \hookleftarrow Z_3 \hookleftarrow \cdots$$

of X by subvarieties Z_0, Z_1, etc.

Some further notation is necessary:

- Let d_Y be a metric distance on Y, and let $\mathscr{H}_Y$ be the Hausdorff measure on Y.

- We say ϕ is *δ-separated* if $\partial^c \phi(x)$ is a δ-separated discrete subset of Y for every $x \in dom(\phi)$, i.e. if $d_Y(y_0, y_1) \geq \delta$ for every $y_0 \neq y_1$ in $\partial^c \phi(x)$.

- Abbreviate $Y'(x) := dom(c_x)$, and let $1_{Y'(x)}$ be the indicator function.

- Under the assumptions (A0)–(A3), the cell $Z'(x)$ is has well-defined tangent space $T_x Z'(x)$. Let $pr_{Z'} : T_x X \to T_x Z'(x)$ be the orthogonal projection.

- For a real parameter $\beta > 0$ and $x \in X$, $y \in Y'(x) := dom(c_x)$, define the collection of tangent vectors

$$\eta(x,y) := |\psi(y_0) - \psi(y) + c_\Delta(x; y, y_0)|^{-\beta} \cdot pr_{Z'}(\nabla_x c_\Delta(x; y, y_0)).$$

- For every x, let $\bar{\nu} = \bar{\nu}_x$ be the Radon measure on Y defined by

$$d\bar{\nu}(y) := (1 - e^{d(y, \partial^c \psi^c(x))^2/\delta}).1_{Y'(x)}.d\mathscr{H}_Y(y),$$

and define

$$\eta_{avg}(x) := (\bar{\nu}[Y])^{-1} \int_Y \eta(x,y)\; d\bar{\nu}(y). \tag{1.6}$$

À priori, $\eta_{avg}(x)$ is a vector in $T_x Z'$. We say (UHS) Conditions are satisfied if η_{avg} is uniformly bounded away from zero (c.f. Definition 4.4.2).

Our next result identifies the maximal index $J \geq 1$ such that $(A = Z_1) \hookleftarrow Z_{J+1}$ is a homotopy-isomorphism, and indeed a strong deformation retract.

Theorem 1.4.2. *Let c be a cost satisfying Assumptions (A0)–(A6). Suppose σ, τ are source, target measures which are absolutely continuous with respect to $\mathscr{H}_X, \mathscr{H}_Y$, respectively and satisfying (1.1). Let $\psi^{cc} = \psi$ be a c-concave potential (2.3.5) dual to the c-optimal semicoupling from σ to τ, and suppose $\phi = \psi^c$ is a δ-separated c-convex potential for some $\delta > 0$. Let $j \geq 1$ be an integer.*

(a) *Suppose there exists a parameter $\beta > 0$ such that $\eta_{avg}(x')$ (1.6) is bounded away from zero, uniformly with respect to $x' \in Z_j - Z_{j+1}$; and*

(b) *$\partial^c \psi^c(x') \cap Z_{j+1} \neq \emptyset$ for every $x' \in Z_j - Z_{j+1}$.*

Then the inclusion $Z_{j+1} \hookrightarrow Z_j$ is a homotopy-isomorphism, and even a strong deformation retract. Furthermore if $J \geq 1$ is the maximal integer such that every $x' \in Z_J - Z_{J+1}$ satisfies conditions (a)–(b), then the inclusion $Z_{J+1} \hookrightarrow Z_1$ is a homotopy-isomorphism, and there exists an explicit strong deformation retract.

In practice the hypotheses (i)–(ii) of Theorem 1.4.2 may be difficult to verify. We elaborate further in Chapter 2.3.

Interesting applications arise when the target is the boundary $Y = \partial X$ of the source X. Combining Theorems 1.4.1, 1.4.2, we find cost functions $c : X \times \partial X \to \mathbb{R} \cup \{+\infty\}$ satisfying Assumptions (A0)–(A6) and sufficient (UHS) conditions produce codimension-J deformations of the initial source space X onto locally-lipschitz subvarieties Z_{J+1}, where J is an index ≥ 0. We describe topological applications in Chapters 5, 6, 7.

1.5 Closing the Steinberg symbol and Theorem 1.5.1

The second phase of our thesis relates to algebraic topology, and applies the above Reduction-to-Singularity to construct small-dimensional $E\Gamma$ classifying spaces, where Γ is an infinite discrete Bieri-Eckmann duality group of dimension ν and with dualizing module $\mathbf{D}$. See [BS73] for definitions. The setting implies Γ is finitely-generated, virtually-torsion free. E.g. $\Gamma = PGL(\mathbb{Z}^2)$, $PGL(\mathbb{Z}^3)$, $Sp(\mathbb{Z}^4)$, $G(\mathbb{Z})$ arithmetic groups, mapping class groups $MCG(\Sigma_g)$ of closed surfaces, knot-groups, etc. Our goal:

> Given a finitely generated group Γ satisfying Bieri-Eckmann's homological duality [BE73], to display an $E\Gamma$-model X with proper-discontinuous free action $X \times \Gamma \to X$ having topological dimension $\dim(X)$ equal to the virtual cohomological dimension $vcd(\Gamma)$ of the group Γ.

This problem has been intensively studied since Borel-Serre's computation

$$vcd(\mathbb{G}(\mathbb{Z})) = \dim(K\backslash^0\mathbb{G}(\mathbb{R})) - rank_{\mathbb{Q}}(\mathbb{G}),$$

where $\mathbb{G}$ is a $\mathbb{Q}$-split reductive linear algebraic group scheme and $rank_{\mathbb{Q}}(\mathbb{G})$ is the so-called $\mathbb{Q}$-rank of $\mathbb{G}$, c.f. [BS73], [Ser71]. For instance

$$vcd(GL(\mathbb{Z}^2)) = 2 - 1 = 1, \quad vcd(Sp(\mathbb{Z}^4)) = 5 - 2 = 3.$$

For algebraic groups $\mathbb{G}$ with $\mathbb{Q}$-rank equal to one, the above problem simplifies since numerous adhoc methods are available for continuously retracting an open manifold onto a codimension-one hypersurface, c.f. [Yas07]. But a general method has been apparently hidden from sight for higher codimensions, with the notable exceptions of [Gro91], [Ash84], [Sou78], [MM93]. We propose our Theorems 1.4.1 – 1.4.2 yield a new technique for constructing homotopy-reductions of large-codimension based on the reduction-to-singularity idea.

To implement the Theorems 1.4.1, 1.4.2 however require some further ideas. We assume a user first has an explicit geometric $E\Gamma$ model X available for sampling, e.g. (X, d) an finite-dimensional Cartan-Hadamard space with isometric group action $X \times \Gamma \to X$ which is proper discontinuous, free, and with finite volume quotient. Typically the space dimension $\dim(X)$ is much larger than the cohomological dimension $\nu := cd(\Gamma)$, $\nu << \dim(X)$. To effectively construct $E\Gamma$ models $\mathscr{Z}$ with $dim(\mathscr{Z}) = \nu$ is largely unsolved problem. Our thesis constructs new Γ-invariant closed subsets $\mathscr{Z}$ with $\dim(\mathscr{Z}) << \dim(X)$ and for which $\mathscr{Z} \hookrightarrow X$ is a homotopy-isomorphism and construct Γ-equivariant homotopy-reductions $X \rightsquigarrow \mathscr{Z}$. Our retractions are geometric-flows which continuously collapse X onto a large-codimension subvariety $\mathscr{Z}$. Given an initial $E\Gamma$-model X, our technique exhibits the large-codimension retract as the locus-of-discontinuity of a "retraction" (i.e. c-optimal semicoupling) from X to the boundary ∂X. *But which boundary, which cost?*

Our thesis studies these questions with our rational excision models, denoted $X[t]$, and repulsion costs. The excision models $X[t] := X - \cup_\lambda W_\lambda^t$ are obtained by equivariantly scooping-out/excising Γ-rational horoballs W_λ^t from X, with respect to a sufficiently "small" Γ-equivariant parameter t. The family of horoballs $\{W_\lambda^t\}$ are modelled on constant-curvature halfspaces far at-infinity. The Γ-rationality implies $X[t]$ and $\partial X[t]$ are Γ-invariant subsets of X. Crucially they inherit proper-discontinuous actions

$$X[t] \times \Gamma \to X[t], \quad \partial X[t] \times \Gamma \to \partial X[t].$$

The key property of our excisions $X[t]$ is that the reduced-homology of the boundary $\mathbf{D} := \tilde{H}_*(\partial X[t]; \mathbb{Z})$, with its natural $\mathbb{Z}\Gamma$-module structure, is explicit resolution of the Bieri-Eckmann dualizing module for Γ.

The homological modules $\mathbf{D} := \tilde{H}_*(\partial X[t])$ are called Steinberg modules. The $\mathbb{Z}\Gamma$-module $\mathbf{D}$ is principal and infinite cyclic, generated by a cycle $[B]$. I.e. $[B]$ is a basic "sphere-at-infinity" and are called Steinberg symbols. Contractibility of $X[t]$, the standard long-exact sequence of relative homology, and the $\kappa \leq 0$ geometry of X implies the boundary map $\partial : H_{*+1}(X[t], \partial X[t]) \to \tilde{H}_*(\partial X[t])$ is an isomorphism and with inverse given by "flat filling". Now the relative cycle $FILL[B]$ is also called a Steinberg symbol and is a disk. The dimension of this disk is precisely the maximal codimension of a spine. Homological duality implies this disk $P := FILL[B]$ is dual, with respect to intersection homology, to the spine fundamental class. Therefore minimal spines are transverse to Steinberg symbols and intersect precisely at a point. But we see how the Steinberg symbol retracts to a point $P \rightsquigarrow pt$. Our goal is to interpolate this retraction throughout $X[t]$ to obtain $X[t] \rightsquigarrow \mathscr{Z}$. We achieve this interpolation using the singularity functor arising from our two-pointed repulsion cost $\widetilde{c}$ defined in §5.4 and extended to visibility costs in §5.9.6.

Next we replace the excision $X[t]$ with a chain sum $\underline{F}$, where Γ acts on $\underline{F}$ as a type of shift operator. For formal definition of "chain sum" we refer the reader to [GJ81] or any reference of singular homology. The chain summands of $\underline{F}$ are a principal Γ-set, and the action $\underline{F} \times \Gamma \to \underline{F}$ is equivariantly isomorphic to $\Gamma \times \Gamma \to \Gamma$. These chain sums are defined by the user solving an elementary combinatorial subprogram we call Closing the Steinberg symbol, to be introduced in Chapter 7. The problem of Closing Steinberg is motivated by the following question. Suppose we have a countable collection of embedded isometric equilateral triangles $\{\blacktriangle_i\}_{i \in I}$, where each $\blacktriangle_i$ has a prescribed embedding $\blacktriangle_i \hookrightarrow \mathbb{R}^d$ into some d-dimensional affine space. Evidently each triangle $\blacktriangle$ is 2-dimensional with boundary $\partial \blacktriangle = \triangle$. The question is: can we determine a finite subset $I' \subset I$ for which the chain sum $\sum_{I'} \blacktriangle_i$ has chain boundary

$$\partial(\sum_{I'} \blacktriangle_i) = \sum_{I'} \partial \blacktriangle_i = \sum_{I'} \triangle_i$$

vanishing over $\mathbb{Z}/2$-coefficients? In otherwords, can we arrange the triangles $\blacktriangle_i$ to form the boundary of a cube, or regular platonic solid, or some other closed convex polyhedron?

There is important interpretation of Closing Steinberg in terms of algebraic topology and group-cohomology. Let C_q be the q-th singular homology groups. Algebraically we find the finite subset I of Γ, which Closes Steinberg and defines the convex base-chain

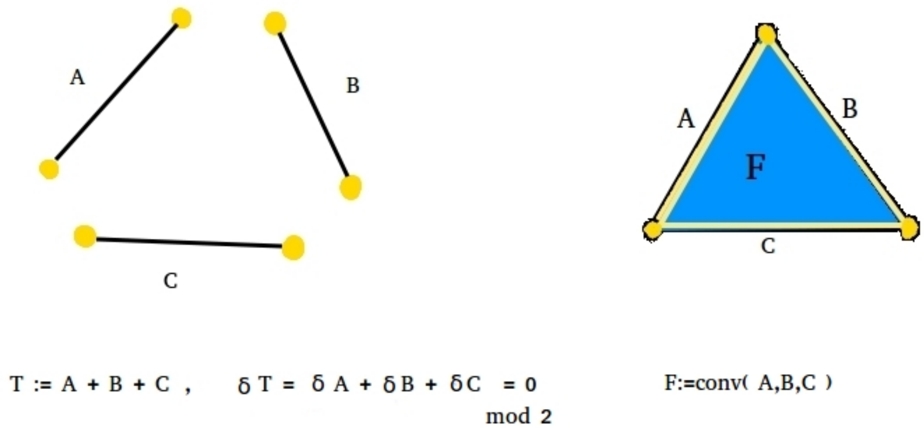

Figure 1.1: When $\partial(A+B+C)=0$ (mod 2), we define $F = conv(A,B,C)$

$F = F(I)$ in $\underline{F}$ corresponds to a symbol $\xi \in C_q(X, \partial X; \mathbb{Z}/2)$ satisfying $\partial_0 \xi = 0$ with respect to the formal boundary operator $\partial_0 : C_q(X, \partial X; \mathbb{Z}/2) \to C_{q-1}(\partial X; \mathbb{Z}/2)$. This algebraic-topological interpretation is further developed in Chapters 6.

Our method is defined for groups Γ satisfying Bieri-Eckmann's homological duality generalizing Poincaré duality [BE73]. Specifically, Closing Steinberg amounts to constructing a nontrivial 0-cycle $\xi \in H_0(\Gamma; \mathbb{Z}_2\Gamma \otimes \mathbf{D})$. Here $\mathbb{Z}_2\Gamma := \underline{\mathbb{Z}/2} \otimes_{\mathbb{Z}} \mathbb{Z}\Gamma$ is the induced $\mathbb{Z}\Gamma$-module with coefficients over $\mathbb{Z}/2$, considered as trivial $\mathbb{Z}\Gamma$-module. We remark that Bieri-Eckmann duality implies $H_0(\Gamma; \mathbb{Z}_2\Gamma \otimes \mathbf{D}) \approx H^\nu(\Gamma; \underline{\mathbb{Z}_2}\Gamma) \neq 0$ where $\nu = vcd(\Gamma)$.

In applications below the triangles $\blacktriangle$ are replace with a flat-filled relative chain $P \in C_q(X[t], \partial X[t]; \mathbb{Z})$ whose chain-boundary $\partial[P] = B$ is generator of Steinberg module $\mathbf{D}$. To Close the Steinberg symbol requires finding a finite subset I of Γ for which the translates $P.I$ have positions in $(X[t], \partial X[t])$ bounding a closed geodesically convex domain $F = conv[P.I]$. The symmetry group Γ acts isometrically on $X[t] \times \partial X[t]$, and we form the chain sum $\underline{F} := SUM[F(I).\Gamma]$, of the Γ-translates of the convex base chain $F = F(I)$. Our hypotheses ensure the chain sum $\underline{F}$ becomes a cubical fundamental class. The chain sum $\underline{F}$ can be interpreted as a "partition-of-unity" of the support $supp(\underline{F}) \subset X$. To successfully close the Steinberg symbol allows the user to replace a space X with $supp(\underline{F})$ and the chain sum $\underline{F}$. Our hypotheses of Closing Steinberg (see 7.2.1, 7.2.3) ensures the support $supp(\underline{F})$ is aspherical and homotopy-equivalent to X.

The above Theorems 1.4.1, 1.4.2 are general topological theorems obtained by our semicoupling methods. The theorems require costs c which satisfy the necessary hypotheses, and this is nontrivial. As we elaborated above, the quadratic costs are not sufficient, and we find best results obtained with our anti-quadratic repulsion costs.

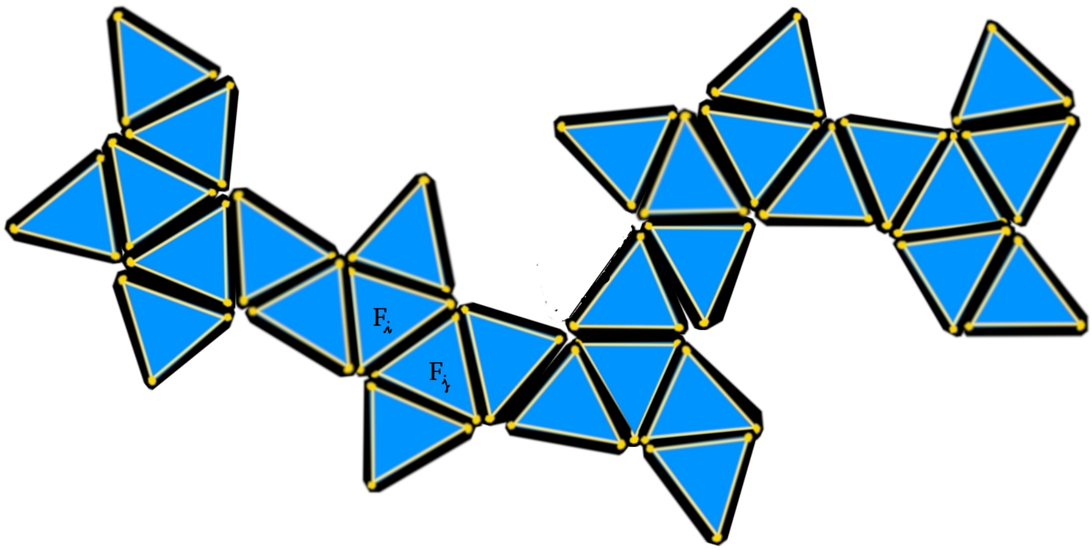

Figure 1.2: A chain sum $\underline{F} = \sum_{i \in I} F_i$ with well-separated one-dimensional gates

The following Theorem 1.5.1 summarizes our applications of the previous Theorems 1.4.1– 1.4.2. The theorem is a multi-stepped reduction program with input a geometric $E\Gamma$ model X, and outputs a Γ-equivariant homotopy reduction of X onto a codimension-J closed subvariety $\mathscr{X}$ where the index $J \geq 0$ is defined according to the hypotheses of Theorems 1.4.1, 1.4.2.

The reduction program runs as follows:

(a) Given an initial geometric $E\Gamma$ model X, equivariantly scoop out the Γ-rational horoballs W_λ^t and obtain a manifold-with-corners $X[t] := X - \cup_\lambda W_\lambda^t$. (6.3.3, 5.5.1)

(b) The boundary $Y := \partial X[t]$ is Γ-invariant, and the induced Γ-action is geometric (e.g., proper discontinuous). Choose the excision parameter t such that the natural $\mathbb{Z}\Gamma$-module $\mathbf{D} := \tilde{H}_q(\partial X[t]; \mathbb{Z})$ is the Steinberg module of Γ. (6.5.2, 6.4.3) Let B be a generator of $\mathbf{D}$ with flat-filling $P = FILL[B]$.

(c) Find a finite subset $I \subset \Gamma$ which successfully Closes the Steinberg symbol, and replace the excision model $X[t]$ with the chain sum $\underline{F}$, where

$$\underline{F} := \sum_{\gamma \in \Gamma} F[t].\gamma, \quad \text{and} \quad F[t] := conv[P.I] \cap X[t],$$

(7.2.1, 7.2.3).

(d) Let Y_ϵ be the ϵ-regularization of Y for small $\epsilon > 0$ (5.7.1), and let $\Omega \subset F[t]$ be the visibility domain (5.9.3). Define the chain sum

$$\underline{\Omega} := \sum_{\gamma \in \Gamma} \Omega.\gamma,$$

and let

$$v : \underline{\Omega} \times Y_\epsilon \to \mathbb{R} \cup \{+\infty\}$$

be the visible repulsion cost (5.9.6).

(e) Let τ be a Γ-invariant volume measure on Y (τ is canonical modulo scalars since Γ acts geometrically on Y). The construction of Y_ϵ is also Γ-invariant, and let τ_ϵ be the invariant volume measure on Y_ϵ (unique modulo scalars). Choose scalars satisfying

$$\rho := \int_{\underline{\Omega}/\Gamma} \sigma / \int_{Y_\epsilon/\Gamma} \tau_\epsilon > 1. \tag{1.7}$$

(f) For $\rho > 1$, let $Z : 2^{Y_\epsilon} \to 2^{\Omega}$ be the Kantorovich functor defined by v-convex potentials, which are dual maximizers to the v-optimal semicoupling from $1_{\underline{\Omega}}\sigma$ to τ_ϵ.

Following the above steps (a)–(f), we arrive at the following application of Theorems 1.4.1, 1.4.2.

Theorem 1.5.1. *Let $X \times \Gamma \to X$ be a geometric $E\Gamma$-model with Γ-invariant volume measure σ (6.1.2). Let $Z : 2^{Y_\epsilon} \to 2^{\underline{\Omega}}$ be defined by the above items (a)–(f). If $J \geq 0$ is the maximal index where the hypotheses of 1.4.1–1.4.2 are satisfied throughout Z_J, then $\mathscr{Z} := Z_{J+1}$ is a Γ-equivariant codimension-J strong-deformation retract of X.*

For arbitrary source and targets σ, τ, the maximal index J output by Theorem 1.5.1 is possibly $J = 0$. In this the hypotheses of Theorem 1.4.1 fail, and Theorem 1.5.1 is trivial. In case $J = 1$, then Theorem 1.5.1 reduces to a special case of Theorem 1.4.1. For the applications developed in Chapters 5, 6, 7, we expect the maximal index J is larger, say at least $J \geq 2$.

We emphasize that the primary obstruction to the reduction program of Theorem 1.5.1 is verifying the (UHS) conditions throughout the necessary subdomains. For general costs this appears difficult problem. The (UHS) conditions amount to requiring an average gradient vector $\eta_{avg}(x)$ (defined in (4.7)) be nonzero, and have a sequence of nonzero projections $proj_{Z'}\eta_{avg}(x) \neq 0$ for select subvarieties Z' containing x in X. See §4.4 and Definition 4.4.2 for details. A second obstruction is the hypothesis $Z'(x) \cap Z_{j+1} \neq \emptyset$ for $x \in Z_j$, which is necessary for the retraction defined in 1.4.2.

The definition of Closing Steinberg is specially adapted to geometric $E\Gamma$ models as described in §§7.3, 7.4 below when Γ is a standard arithmetic group. Thus we conjecture that Theorem 1.5.1 yields singularity structures which can be homotopy-reduced to the maximal codimension, and thus we propose new $E\Gamma$-models $\underline{Z}$ with space dimension equal to the cohomological dimension of Γ. This conjecture requires verifying several properties of the visibility cost v (5.2, 5.9.6).

Conjecture 1.5.2. *Under the hypotheses of Theorem 1.5.1, we conjecture that*

(i) *the visible repulsion cost v satisfies Assumptions (A0)–(A6); and*

(ii) *when the ratio ρ (1.7) is sufficiently close to 1^+, the activated source $A = Z_1$ of the unique v-optimal semicoupling is a continuous equivariant deformation retract of the source domain $\underline{\Omega} \approx X$; and*

(iii) *the maximal index $J \geq 1$ for which the hypotheses of 1.4.2 are satisfied is equal to $J = q + 1$, where q is the topological dimension of spheres generating the Steinberg module $\boldsymbol{D}$;*

(iv) *the inclusion $Z_{J+1} \hookrightarrow X$ is a Γ-equivariant homotopy-isomorphism, and even a strong deformation retract with $\dim(Z_{J+1}) = cd(\Gamma)$.*

The Conjecture 1.5.2 proposes the subvariety $\mathscr{Z} := Z_{J+1}$ is a minimal-dimension spine of $E\Gamma$, where the explicit retracts are given by our Theorems 1.4.1, 1.4.2. The Conjecture 1.5.2 requires several steps be verified. First we need verify the differential-geometric (Twist) condition for the visibility cost v, i.e. we require the function $dom(v_{x'}) \to T_{x'}\underline{F}$ defined by the rule $y \mapsto \nabla_x v(x', y)$ be injective for every choice of x' in $\underline{F}$. Next we need establish the homotopy-isomorphism between the activated source domain's inclusion $(Z_1 = A) \hookrightarrow (Z_0 = \underline{F})$. Finally we need verify (UHS) conditions are satisfied throughout the subvarieties $Z_1, Z_2, \ldots$ and their local "cells" $Z' = Z(\partial^v \psi^v(x'))$, where ψ^v is the v-convex potential output by Kantorovich duality. We postpone verification to future investigations.

1.6 Thesis Outline

Now we outline the contents of our thesis. Our thesis has two phases. The first phase is general, and develops our applications of optimal transport (and specifically the category of semicouplings) to algebraic topology. Chapter 2 is largely a survey of known results in the optimal transportation literature. We develop the principles of the semicoupling program: existence, uniqueness, and Kantorovich duality for costs satisfying Assumptions

(A0)–(A4). We conclude Chapter 2 with the proof of Theorem 1.4.1 (see 3.1.1), which is the base case for the larger-codimension retracts constructed in the next chapters.

Our Chapter 4 develops the basic geometric properties of Kantorovich's contravariant singularity functor $Z : 2^Y \to 2^X$. Section 4.1 is central to our thesis, especially Definition 4.1.1. Section 4.2 describes the local topology of the singularities and proves a useful codimension estimate. Assuming the cost satisfies Assumptions (A0)–(A6) and sufficient (UHS) conditions, we conclude Chapter 4 with the proof of our Theorem 1.4.2; see Theorems 4.4.3, 4.4.4 from §4.4. Thus Chapters 2–4 establish a general method for constructing strong deformation retracts based on Kantorovich duality and optimal transport.

The general results of the previous chapters require special costs for applications. Chapter 5 introduces a new class of cost functions called repulsion and visibility costs. These repulsion costs are distinct from the familiar "attraction" costs, e.g. quadratic costs $c = d^2/2$. To apply the homotopy-reductions of Theorems 1.4.1, 1.4.2 requires we verify the repulsion costs satisfy the necessary Assumptions (A0)–(A6). In this direction, our thesis admittedly achieves only partial results. Specifically we easily find the costs satisfy Assumptions (A0)–(A3) and (A5). But we only succeed in demonstrating (A4) for the repulsion cost denoted $c|\tau$ (Definition 5.3.2, Prop. 5.3.3). However we present simple heuristics suggesting the costs satisfy the Assumption (A4), i.e. (Twist). The remaining Assumption (A6) remains a conjecture.

The remaining Chapter 6 and Chapter 7 develop the applications of our repulsion costs and singularity functors to geometric $E\Gamma$ models. In Chapter 6 we let Γ designate a countable discrete group, and describe the background on $E\Gamma$-models X, and their Γ-equivariant excision models $X[t]$, which are manifolds-with-corners having Γ-invariant topological boundary $\partial X[t]$. The results of Chapter 6 are surely well-known to the experts, although our emphasis on excisions rather than bordifications has perhaps been unappreciated hitherto. This is the key to applying our semicoupling methods to $E\Gamma$ models, and we outline the basic ideas in §6.3. The excision construction, and its relation to Bieri-Eckmann's homological duality is described in §6.4 and summarized in Theorem 6.5.2. This theorem appears in various forms throughout the literature, e.g. it is effectively Borel-Serre's rational bordification model from [BS73], coupled with our own variation of Grayson's construction [Gra84].

The final Chapter 7 introduces the problem of Closing the Steinberg symbol, which is a homological subprogram we discovered to replace the excision model $X[t]$ with a cubical chain sum $\underline{F}$. This idea is defined and established in Definition 7.2.1 and Theorem 7.2.3, respectively. Successfully Closing Steinberg is key step towards the effective application of our semicoupling method to topological $E\Gamma$ models. The key feature of $\underline{F}$ is that Γ acts as shift-operator on the summands of $\underline{F}$. The summands of $\underline{F}$ are excisions of convex sets on which we install the repulsion costs from previous Chapter 5, thereby implementing the reduction-program detailed in Theorem 1.5.1. We conclude with some basic examples of Closing Steinberg in §§7.3–7.4.

1.7 Conventions and Notations

Throughout the thesis we adopt the following conventions: we let X, Y denote manifolds-with-corners, equipped with a Riemannian distance functions $d = dist_X$,$dist_Y$. See [BS73, Appendix] for formal definitions regarding "les variétés à coins". Briefly we recall a space X is a manifold-with-corners if X is locally modelled (via diffeomorphisms) to sectors $\mathbb{R}_+^k \times \mathbb{R}^{n-k}$ for various integers $0 \leq k \leq n$, where $n = dim(X)$ and $\mathbb{R}_+ := [0, +\infty)$. We let vol_X and vol_Y be the volume measures on X, Y. The Hausdorff measures are abbreviated $\mathscr{H}_X = \mathscr{H}^{\dim(X)}$. We reserve σ and τ for Radon measures on X and Y, called "source" and "target" measures, respectively. Typically σ, τ are mutually absolutely-continuous with respect to the Hausdorff measures $\mathscr{H}_X$, $\mathscr{H}_Y$. The support $supp(\mu)$ of a Radon measure μ is the minimal closed subset of full μ-measure. The domain $dom(f)$ of a function $f : X \to \mathbb{R} \cup \{+\infty\}$ consists of all points x where $f(x) < +\infty$.

The singularity structures arising from our thesis are assembled from closed subvarieties of the source space. We follow Poincaré 's original terminology of "varieties' and "subvarieties" from [Poi95, §§10–12]. In this thesis a subvariety Z of X is a closed subset defined by a collection of explicit equations (c.f. §4.2). If the equations are described by Lipschitz (or DC) functions, then we call Z a Lipschitz (or DC) subvariety. In applications the functions are usually smooth.

A topological space X is aspherical if X is connected and all homotopy-groups of the universal cover $\tilde{X}$ are trivial. If X has the structure of a locally finite cell complex, then X is aspherical if and only if $\tilde{X}$ is contractible. A continuous map $f : X \to X'$ between topological spaces is a homotopy-isomorphism if the induced maps $\pi_i(f) : \pi_i(X) \to \pi_i(X')$ are isomorphisms for all homotopy groups π_i, $i = 0, 1, 2, \ldots$. According to Whitehead's Theorem [Bre93, §VII.11], f is a homotopy-isomorphism if and only if the morphisms induced on homology $H_i(f) : H_i(X) \to H_i(X')$ are isomorphisms for all i. A space X deformation retracts onto the subspace A if there exists a continuous map $h : X \times [0, 1] \to X$ such that $h(x, 0) = x$, $h(x, 1) \in A$, $h(a, t) = a$ for all $x \in X$, $a \in A$, $t \in [0, 1]$. A deformation retract h defines homotopy-isomorphisms $x \mapsto h(x, 1)$. For the formal definitions of chain complexes, chain maps, cochain complexes, cochain maps, and the Koszul complex, we refer the reader to [Lan05, §§XX.1-2, XXI.1,2,4]. The singular chain groups $\{C_q^{sing}(X)|q = 0, 1, 2, \ldots\}$ on a topological space X are formally defined in [GJ81], or [Bre93, Chapter IV]. For the definition of simplicial chain groups and chain sums, see [Bre93, p. IV.21].

The identity mapping on whatever set is denoted Id. A category C is a collection of objects $Obj(C)$ (a set), and a collection of morphisms between objects $Hom_C(X, Y)$ (the set of morphisms in C between objects X, Y) with the property that compositions of morphisms is well-defined in C and associative, and for every object X the identity mapping $Id_X \in Hom_C(X, X)$. A subcategory C' of C is a category whose objects are a subset of the objects of C, and where $Hom_{C'}(X, Y) \subset Hom_C(X, Y)$ for every pair of objects in C'. A functor $F : C \to D$ between categories C, D is for every object X in C, and object $F(X)$ in D, and for every morphism $f : X \to Y$ in C, a morphism $F(f) : F(X) \to F(Y)$ in D. The functor F is contravariant if $F(f \circ g) = F(g) \circ F(f)$ for every pair of morphisms f, g in C with composition $f \circ g$ in C. The functor is covariant if $F(f \circ g) = F(f) \circ F(g)$. See [Lan05, p. 1.11] for complete definitions.

The symbol Γ usually designates an infinite torsion-free discrete group. Following standard notation, $E\Gamma$ is the universal cover $\tilde{X}$ of an Eilenberg-Maclane space $X = K(\Gamma, 1)$. The singular chain groups on $\tilde{X}$, with their natural Γ-action, are a topological model for the group-theoretic Γ-cohomology. We let $cd(\Gamma)$ denote the cohomological dimension of Γ, and equal to the unique integer $\nu \geq 0$ for which the group-theoretic cohomology group $H^\nu(\Gamma; \mathbb{Z}\Gamma)$ is nonzero, [Bro82]. A cubulation of a group Γ is an $E\Gamma$-model which has explicit cellular structure defined in terms of geometric identifications between cubes $I^n = [0, 1]^n$ for $n \geq 0$, with additional "wall-structures". Precise definitions are given in §7.2 in terms of chain sums and "gates", e.g. Definition 5.1.1. We let F denote a geodesically convex compact subset in a complete Riemannian manifold, and $\mathscr{E}$ denotes the extreme-point functor, $\mathscr{E}[F]$ consists of the extreme-points on F.

Chapter 2

Background: Semicouplings and Kantorovich Duality

The present chapter is largely review, especially §§2.1 –2.5 which assembles the basic facts of optimal semicouplings.

Our thesis relates algebraic topology to measure theory by replacing *"continuous deformation retracts* $r : X \to Y$" (which are nonexistent according to Brouwer's theorem, §1.2), with "*c-optimal semicouplings π_{opt}*" between a source (X, σ) and target (Y, τ), and where $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is a cost function satisfying the various Assumptions (A0), (A1), ... from §1.3.

The main theme of our thesis is to replace the graphs of continuous retracts r, with semicouplings and specifically c-minimal semicouplings (see the minimization program (2.5) in Section 2.3 below). Following Kantorovich's duality theorem, we study the dual maximization program to obtain Kantorovich's contravariant singularity functor $Z : 2^Y \to 2^X$ defined by the rule

$$Y_I \mapsto Z(Y_I) = \cap_{y_* \in Y_I} \partial^c \psi(y_*)$$

for closed subsets $Y_I \hookrightarrow Y$. The maximizers correspond to c-concave potentials $\psi : Y \to \mathbb{R} \cup \{-\infty\}$, see §2.3. The functor Z produces subvarieties $Z(Y_I) \hookrightarrow X$ described explicitly by the dual potentials ψ, ψ^c and c, see §4.2 for explicit local equations. The functor $Z : 2^{\partial X} \to 2^X$ produces locally DC-subvarieties $\mathscr{Z}$ of X for which the inclusions $\mathscr{Z} \hookrightarrow X$ are homotopy-isomorphisms. See Theorems 4.4.3, 4.4.4 for details.

We need remark on the definition of "probability" and its relation to optimal transportation methods. The standard methods of couplings and Monge-Kantorovich duality [Vil09], [San15] are contingent on the hypothesis that source and target measures have identically equal masses. That is, source σ and target τ must satisfy

$$\int_X \sigma - \int_Y \tau = 0. \tag{2.1}$$

The present thesis develops in the *semicoupling category*, and replaces (2.1) with the hypothesis that source measures σ be "abundant" with respect to a target measure τ, namely:

$$\int_X \sigma - \int_Y \tau \geq 0. \tag{2.2}$$

Semicouplings are abundant, while couplings need not exist (especially when inequality (2.1) is strict). However the existence of a renormalization factor $(\int_Y \tau)^{-1} < +\infty$ is a useful hypothesis. For instance, we prove the à priori existence of c-optimal semicouplings in §2.2, and our proofs shall assume the target measure τ has been renormalized to a probability measure. But the practical construction of the dual Kantorovich potentials does not require the renormalization, and nor does our proof of the existence of Kantorovich minimizers in §2.3. This illustrates the logical convenience of assuming an inequality $\int_X \sigma < +\infty$, rather than exactly evaluating some real number $(\int_X \sigma)^{-1}$.

There are advantages in ignoring the normalization condition (2.1), as studied by some authors, notably [HS13], [CM10], [Fig10b]. In this thesis, our applications to arithmetic groups (see §6) would be immediately obstructed if we required users to explicitly normalize the relevant Haar measures. This requires exactly calculating the volumes of polyhedral fundamental domains, and this calculation is practically impossible, e.g. compare [Lan66]. So from the beginning, our applications do not require precise normalization factors.

2.1 Optimal Semicouplings and Cost

Now we introduce the optimal semicoupling program. We continue with the notations from the Introduction. We reserve X, Y for complete finite-dimensional manifolds-with-corners. Let σ, τ be Radon measures on the source and target X, Y. Let $proj_X : X \times Y \to X$, $proj_Y : X \times Y \to Y$ be the canonical continuous projections. We typically assume σ, τ are absolutely-continuous with respect to $\mathscr{H}_X$, $\mathscr{H}_Y$.

Definition 2.1.1 (Semicoupling)**.** A *semicoupling* between source (X, σ) and target (Y, τ) is a Borel measure π on the product space $X \times Y$ with target-marginal satisfying $proj_Y \# \pi = \tau$, and source-marginal satisfying $proj_X \# \pi \leq \sigma$.

The inequality $proj_X\#\pi \leq \sigma$ holds if for every Borel subset O, the numerical inequality $(proj_X\#\pi)[O] \leq \sigma[O]$ is satisfied. We remark that π is a coupling between σ and τ when $proj_X\#\pi = \sigma$.

Next let $SC(\sigma,\tau) \subset \mathscr{M}_{\geq 0}(X \times Y)$ denote the set of all semicoupling measures π between source σ and target τ. One finds $SC(\sigma,\tau)$ is empty unless is satisfied, in which case we say "the source σ is abundant relative to the target τ". Informally a semicoupling $\pi \in SC(\sigma,\tau)$ describes an allocation of some activated source particles which fill (or saturate) a prescribed target.

A standard argument using Prokhorov's compactness criterion implies the following lemma, c.f. [Vil09, Lemma 4.4, pp.44]. We recall that a sequence of semicouplings $\{\pi_k\}_{k=1,2,...}$ converges to π_∞ in the narrow-topology if $\lim_{k\to+\infty}\int_{X\times Y} f(x,y).d\pi(x,y)$ for every bounded continuous function $f \in BC(X)$.

Lemma 2.1.2. *The set of semicouplings $SC(\sigma,\tau)$ is compact convex subset of $\mathscr{M}_{\geq 0}(X \times Y)$ with respect to the narrow topology.*

Proof. The convexity of semicouplings is clear. For every $\epsilon > 0$ both σ and τ admit compact subsets K_ϵ, L_ϵ such that $\sigma[X - K_\epsilon] < \epsilon$ and $\tau[Y - L_\epsilon] < \epsilon$. But for any semicoupling $\pi \in SC(\sigma,\tau)$, we find

$$\pi[(X-K_\epsilon)\times(Y-L_\epsilon)] \leq \pi[(X-K_\epsilon)\times Y] + \pi[X \times (Y-L_\epsilon)] < 2\epsilon,$$

since $\pi[(X-K_\epsilon)\times Y] = \sigma[X-K_\epsilon]$ and $\pi[X\times(Y-L_\epsilon)] = \tau[Y-L_\epsilon]$. Therefore $SC(\sigma,\tau)$ is precompact with respect to the weak-$*$ topology by Prokhorov's theorem. But it's immediate that $SC(\sigma,\tau)$ is weak-$*$ closed, and therefore the set is weak-$*$ compact. $\square$

2.2 Existence of c-Optimal Semicouplings

Now we instroduce costs. There is no canonical semicoupling without, say, selecting a linear functional on $SC(\sigma,\tau)$ and then minimizing. In §1.3 we described several assumptions, labelled (A0), ..., (A5). The Assumptions (A0)–(A3) are rather generic. The Assumption (A4) implies the general uniqueness of semicouplings, see Proposition 2.5.8. The uniqueness of such optimal semicouplings is important for our topological applications, since we are proposing the singularity structure of optimal semicouplings as canonical topological model. The final basic assumption our thesis requires is Assumption (A6) which is a "small-cancellation" hypothesis on local tangent vectors and ensures certain averages are nonzero. The nonvanishing of these averages is necessary for the continuity of the deformation retracts constructed in Chapter 2.3.

The existence of c-optimal couplings for costs satisfying $(A0) - -(A2)$ is a standard consequence of Fatou's lemma and Prokhorov's precompactness theorem. If π is a Radon measure on $X \times Y$, then we define

$$C[\pi] := \int_{X \times Y} c(x, y) d\pi(x, y).$$

Proposition 2.2.1. *Let σ, τ be source and target measures with $\int_Y 1 d\tau = 1$. If c is continuous cost, then c-optimal semicouplings π_{opt} exist such that*

$$C[\pi_{opt}] = \inf_{\pi \in SC(\sigma,\tau)} C[\pi].$$

Closer inspection reveals that Proposition 2.2.1 only requires c be lower semicontinuous, c.f. [Vil09], [San15].

We can leverage the existence of optimal couplings to the case of semicouplings. Indeed the transportation literature finds two different approaches to questions of existence and uniqueness of optimal semicouplings. The method of [CM10] interprets semicouplings as conventional couplings by formally adjoining a graveyard point $\{\dagger\}$ to the target, enlarging Y to $Y_+ = Y \coprod \{\dagger\}$. The target measure τ is then extended (relative to the source σ) to the measure $\tau_+ := \tau + \alpha.\delta_\dagger$, where α is the positive scalar $\alpha := \int_X 1.\sigma] - \int_Y 1.\tau$ and $\delta_\dagger$ is the Dirac measure supported at the graveyard point $\{\dagger\}$. The cost is extended to $c_+ : X \times (Y \cup \{\dagger\}) \to \mathbb{R}$ by declaring $\{\dagger\}$ a "tariff-free reservoir". Concretely we assume $c_+(x, y) > 0$ whenever $y \in Y$, and $c_+(x, \dagger) = 0$, for every $x \in X$. There is then a natural correspondance between semicouplings $\pi \in SC(\sigma, \tau)$ and couplings π_+ between σ and τ_+. We observe that c_+ is continuous if and only if c is continuous.

An alternative approach to semicouplings and uniqueness is developed in [HS13], wherein a different reduction to the coupling theory is described. Recall that the support of a measure space (X, σ) is denoted $spt(\sigma)$, and defined to be the smallest closed subset of X of full measure. The argument in [HS13] regarding uniqueness is two-stepped. First one determines conditions on the cost for the activated source domain $A = spt(proj_X\#\pi_{opt})$ to be uniquely determined. In the second step, the semicoupling is restricted to the activated source, and the restriction defines a coupling between $1_A.\sigma$ and target τ. Thus we are reduced to standard coupling theory. By standard arguments, one finds (Twist) condition the main hypothesis controlling uniqueness of the optimal coupling. We elaborate further in the next section.

2.3 Kantorovich Duality

In the following sections we prove that c-optimal semicouplings satisfying Assumptions (A0)–(A4) have σ-a.e. uniquely defined active domain $A \hookrightarrow X$. Restricting to the active domain, we obtain a coupling $1_{A\times Y}.\pi$ between $1_A.\sigma$ and τ which is optimal with respect to the restricted cost

$$c|A(x,y) = \begin{cases} c(x,y), & \text{if } x \in A, \\ +\infty, & \text{if } else. \end{cases} \tag{2.3}$$

Uniqueness of the optimal semicoupling now reduces to the question of whether the $c|A$-optimal coupling is unique. In the following sections we describe how so-called (Twist) conditions on the cost implies a general uniqueness of optimal couplings. The meaning of (Twist) is best illustrated through Kantorovich duality which we introduce below. Standard references for Kantorovich duality with respect to continuous costs include [Vil09, Ch.5], or [San15]. The following definitions are exceedingly useful.

Definition 2.3.1 (c-transforms)**.** If $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ is any function on the target Y, then the c-Legendre transform $\psi^c : X \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$\psi^c(x) := \sup_{y\in Y}[\psi(y) - c(x,y)],$$

for $x \in X$.

If $\phi : X \to \mathbb{R} \cup \{+\infty\}$ is any function on the source X, then the c-Legendre transform $\phi^c : Y \to \mathbb{R} \cup \{-\infty\}$ is defined by the rule

$$\phi^c(y) = \inf_{x\in X}[c(x,y) + \phi(x)],$$

for $y \in Y$.

Definition 2.3.2 (c-concavity)**.** A function $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ is c-concave if $\psi^{cc} = (\psi^c)^c$ coincides pointwise with ψ. Equivalently ψ is c-concave if there exists a lower semicontinuous function $\phi : X \to \mathbb{R} \cup \{+\infty\}$ such that $\phi^c = \psi$ pointwise.

The above definitions imply ψ, ψ^c satisfy the pointwise inequality

$$-\psi^c(x) + \psi(y) \leq c(x,y) \tag{2.4}$$

for all $x \in X$, $y \in Y$. The inequality (2.4) and especially the case of equality is very important for this thesis.

Definition 2.3.3 (c-subdifferential)**.** Let $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ be c-concave potential $\psi^{cc} = \psi$. Select $y_0 \in Y$ where $\psi(y_0)$ is finite-valued. The subdifferential $\partial^c\psi(y_0) \subset X$ consists of those points $x' \in X$ such that

$$-\psi^c(x') + \psi(y_0) = c(x', y_0).$$

Or equivalently such that for all $y \in Y$,

$$\psi(y) - c(x', y) \leq \psi(y_0) - c(x', y_0).$$

Assumptions $(A0), \ldots, (A4)$ on the cost c imply various properties of c-convex potentials and c-subdifferentials. The first useful property is that c-subdifferentials are nonempty wherever the potentials $\phi(x)$ or $\psi(y)$ are finite, see Lemma 2.3.4 below. Recall the domain of ϕ is defined $dom(\phi) := \{x \in X \ |\phi(x) < +\infty\}$.

Lemma 2.3.4. *Let $c : X \times Y \to \mathbb{R}$ be a cost satisfying Assumptions (A0)–(A2). Let $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ be a c-concave potential $(\psi^c)^c = \psi$. Abbreviate $\phi = \psi^c$. Suppose there exists $y' \in Y$ such that $\psi(y) \neq -\infty$. Then:*

(i) ψ is an upper semicontinuous function; and

(ii) $\partial^c\psi(y)$ is a nonempty closed subset of X for every $y \in dom(\psi)$; and

(iii) ϕ is lower semicontinuous function; and

(iv) $\partial^c\psi(y)$ is a nonempty closed subset of Y for every $x \in dom(\phi)$.

Proof. The Assumption (A0) implies $\psi(y) = \inf_{x\in X}[c(x, y) + \psi^c(x)]$ for every $y \in Y$ is an upper semicontinuous function of y. Indeed ψ is equal to the pointwise infimum of a family of continuous functions, namely the X-parameter family of continuous functions $y \mapsto c(x, y) + \psi^c(x)$. Likewise the c-transform $\psi^c(x)$ is a lower semicontinuous function of the source variable x. This proves (i), (iii).

The inequality (2.4) is an equality precisely when $x \in \partial^c\psi(y)$, or equivalently $\psi(y) \geq c(x, y) + \psi^c(x)$. So the subdifferential $\partial^c\psi(y)$ coincides with a sublevel set of $x \mapsto \psi^c(x) + c(x, y)$ and is therefore closed according to Assumption (A2). Thus $\partial^c\psi(y)$ is a closed subset of X for every $y \in Y$. Likewise $\partial^c\psi^c(x)$ is closed subset of Y for every $x \in dom(\psi^c)$.

It remains to show the c-subdifferentials are nonempty on the appropriate domains. Observe that $\phi = \psi^c$ is bounded from below on X unless ψ is identically $-\infty$. Indeed if $\{x_j\}_{j=1,2,\ldots}$ is a sequence in X such that $\lim_{j\to+\infty} \phi(x_j) = -\infty$, then $\psi \equiv -\infty$. This is clear from the definition $\phi(x) = \sup_{y\in Y}[\psi(y) - c(x, y)]$, and the Assumption (A0) that c is uniformly bounded below on $X \times Y$.

Moreover the Assumptions (A0)–(A1) imply the infimum defining $\psi(y)$ can be restricted to a compact subset of X. Indeed (A1) includes the hypothesis that the sublevels $\{x \in X | c_y(x) \leq t\}$ are compact subsets of X for every $t \in \mathbb{R}$. If $\psi(y)$ is finite, then we claim the infimum defining $\psi(y)$ can be restricted to a sublevel set. But observe that $\{x_k\}_k$ cannot be a minimizing sequence with $c(x_k, y)$ diverging to $+\infty$ when $\psi(y)$ is finite. So there exists $t \in \mathbb{R}$ such that

$$\psi(y) = \inf_{\{x \ | \ c(x,y) \leq t\}} [\phi(x) + c(x, y)].$$

But lower semicontinuous functions restricted to compact subsets attain their minima. Therefore $\partial^c\psi(y)$ is nonempty whenever $\psi(y) < +\infty$. Since $x \in \partial^c\psi(y)$ if and only if $y \in \partial^c\phi(x)$ whenever $\psi^{cc} = \psi$ and $\phi = \psi^c$, we find $\partial^c\phi(x)$ nonempty whenever $x \in dom(\phi)$, as follows from the definition 2.3.3 and the arguments above. This establishes (ii), (iv).

□

Further properties of c-convex potentials are developed in Section 2.5 below, c.f. Lemmas 2.5.2, 2.5.5, and Proposition 2.5.7.

For the remainder of this section, we suppose the unique active domain A has been specified (Proposition 2.4.7) and we set $c = c|A$. The semicoupling program then reduces to the coupling program. Both the semicoupling and coupling programs are driven by the "cost" of transporting a unit source mass to a unit target mass. The standard interpretation imagines some industrialist having source $(A, 1_A.\sigma)$ and prescribed target measure (Y, τ). The industrialist looks to activate a source domain in order to transport measure to the target – and all the while minimizing the total transit cost. As we've seen, this is a linear minimization program over a convex compact set.

On the other hand, Kantorovich's dual program is defined in terms of "prices". And here one imagines an autonomous transporter who negotiates prices with the industrialist. The transporter offers to purchase units of source measure at price $\phi(x)$, and then sells these units at various target locations at the price $\psi(y)$. The industrialist knows the cost of direct transport from source x to target y is $c(x, y)$, so the transporter must propose competitive prices to the industrialist. These competitive prices imply a constraint on prices, namely

$$-\phi(x) + \psi(y) \leq c(x, y), \qquad \text{for all } x \in A, y \in Y.$$

Now the transporter is seeking to maximize his/her own total surplus, namely the maximization program

$$\sup_{(\phi,\psi)} [-\int_A \phi(x) d\sigma(x) + \int_Y \psi(y) d\tau(y)],$$

the supremum taken over all pairs of functions $\phi : A \to \mathbb{R} \cup \{+\infty\}$, $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ satisfying the pointwise constraint $-\phi(x) + \psi(y) \leq c(x, y)$.

Now suppose the transporter has competitive prices (ϕ, ψ). How may the diligent transporter improve these prices to a pair (ϕ', ψ') having greater profit? For given source point x', the transporter is obliged to satisfy $\psi(y) - c(x', y) \leq \phi(x')$ for all $y \in Y$. This says $\sup_{y\in Y}[\psi(y) - c(x', y)] \leq \phi(x')$. But to minimize purchase price, the observant transporter replaces ϕ with $\phi'(x') = \sup_{y\in Y}[\psi(y) - c(x', y)]$. Similarly for target point y', the transporter wants to maximize the retail price subject to the constraint, and this maximum price is $\psi'(y') := \inf_{x\in A}[\phi'(x) + c(x, y')]$. One readily sees the prices (ϕ', ψ') are at least as profitable than the original (ϕ, ψ). So the maximization program can be restricted to those pairs of functions (ϕ, ψ) which are maximally-competitive with respect to the cost c. This leads to the fundamental definitions of the c-Fenchel-Legendre transform, c-concavity, and the c-subdifferential as defined above.

We denote the weak-$*$ convex compact subset of couplings between $1_A\sigma$ and τ by $\Pi_C(1_A.\sigma, \tau)$. The pointwise inequality (2.4) implies the inequality

$$\sup_{\psi\ c\text{-concave}} [-\int_A \psi^c(x).d\sigma(x) + \int_Y \psi(y).d\tau(y)] \leq \inf_{\pi\in\Pi_C(1_A.\sigma,\tau)} C[\pi]. \tag{2.5}$$

Kantorovich duality says the inequality (2.5) is an equality "sup = inf", and says there is "no duality gap" between the primal minimization and the dual maximization program. There are two basic questions to be addressed regarding (2.5):

(a) Is the supremum realized by c-concave potentials ψ?,

(b) Is the infimum realized by a c-optimal semicoupling π?

The Assumption (A0) that our costs c are continuous implies the answers to (a) and (b) are well-known. The following Theorem is quoted from [Vil09, Theorem 5.10, pp.57].

Theorem 2.3.5. *Let $c : A \times Y \to \mathbb{R}$ be a bounded nonnegative continuous cost. Then:*

- There is no duality gap between the primal program and the dual program, and

$$\sup_{\psi\ c\text{-concave}} [\int_A -\psi^c(x)d\sigma(x) + \int_Y \psi(y)d\tau(y)] = \min_{\pi\in\Pi_C(1_A.\sigma,\tau)} c[\pi].$$

- The dual program is solvable, and there exists a $c|A$-concave potential $\psi_* : Y \to \mathbb{R} \cup \{-\infty\}$ such that

$$\int_A -\psi_*^c(x)d\sigma(x) + \int_Y \psi_*(y)d\tau(y)] = \sup_{\psi\ c\text{-concave}} [\int_A -\psi^c(x)d\sigma(x) + \int_Y \psi(y)d\tau(y)].$$

To prove the existence of Kantorovich potentials dual to c-optimal semicouplings, it is again convenient to follow [CM10]. We first adjoin a tariff-free reservoir $\{\dagger\}$ to Y, obtaining an auxiliary target space $Y_+ := Y \coprod \{\dagger\}$ with target measure τ_+ and c_+ as defined above.

The hypotheses of 2.3.5 are satisfied. Therefore there exists maximizers (ϕ_+, ψ_+) to the dual program

$$\sup_{\substack{-\phi_+(x)+\psi_+(y)\leq c_+(x,y)\\ x\in X,\ y\in Y_+}} J_+[\phi_+, \psi_+], \tag{2.6}$$

where

$$J_+[\phi_+, \psi_+] := -\int_X \phi_+(x)d\sigma(x) + \int_{Y_+} \psi_+(y)d\tau(y).$$

Let (ϕ_+, ψ_+) be maximizers for the program (2.7). Then $\phi_+ = (\psi_+)^{c_+}$ is c_+-convex and $(\psi_+^{c_+})^{c_+} = \psi_+$. Now we apply a standard restriction argument to (ϕ_+, ψ_+) to obtain a c-convex potential $\phi_0 = \psi_0^c$ on the subset $A := \cup_{y\in Y}\partial^{c_+}\psi_+(y)$ of X. We refer the reader to [Vil09, Lemma 5.18, pp.75] for details.

Lemma 2.3.6. *In the above notation, let (ϕ_+, ψ_+) be c_+-dual potentials maximizing Kantorovich's dual program (2.7) for the extended cost c_+. Restricting ψ_+ to $Y \hookrightarrow Y_+$, we obtain a c-concave potential maximizing Kantorovich's dual program on the subdomain $A := \cup_{y\in Y}\partial^c\psi_0(y)$ in X.*

Proof. We replace the c_+-convex potential $\phi_+ = (\psi_+)^{c_+}$ with a c-convex potential ϕ_0 by the following construction. Define $\psi_0(\dagger) := -\infty$ and $\psi_0(y) = \psi_+(y)$ for $y \in Y$. Then

$$\phi_0(x) := (\psi_0)^c(x) = \sup_{y\in Y}[\psi_0(y) - c(x, y)]$$

is c-convex. Moreover we see:

- $\psi_0 \leq \psi_+$ pointwise throughout Y; and
- $\phi_0 \geq \phi_+$ pointwise throughout X; and
- $\phi_0(x) = \phi_+(x)$ whenever there exists $y \in Y$ with $x \in \partial^{c_+}\psi_+(y)$; and
- $\partial^{c_+}\phi_+(x) \subset \partial^c\phi_0(x)$ whenever there exists $y \in Y$ with $x \in \partial^{c_+}\psi_+(y)$.

Thus restricting to $A := \cup_{y\in Y}\partial^c\psi_0(y)$ in X, we obtain a c-convex potential ϕ_0 supported on A. □

The restrictions (ϕ_0, ψ_0) are c-dual potentials, and we find (ϕ_0, ψ_0) are maximizers to the restricted dual program

$$\sup_{\substack{-\phi(x)+\psi(y)\leq c(x,y)\\ x\in A,\ y\in Y}} J_A[\phi, \psi], \tag{2.7}$$

where

$$J_A[\phi, \psi] := -\int_A \phi(x) d\sigma(x) + \int_Y \psi(y) d\tau(y).$$

We see $-\phi_0, \psi_0$ are c-concave with $\phi_0 = \psi_0^c$.

2.4 Uniqueness of Activated Domain

The previous Section described the existence of c-optimal semicouplings for costs c satisfying Assumptions (A0)–(A4). Henceforth our discussion shall assume c-optimal semicouplings exist, and with existence given we next turn to uniqueness. Following the approach of [HS13], our basic uniqueness result for optimal semicouplings is two-stepped. First there is a monotonicity condition, namely Assumption (A3) from Section 1.3 which ensures uniqueness of activated source domains. The further asymmetric (Twist) condition (Assumption (A4)) then proves uniqueness of the optimal coupling, following a standard argument, e.g. [Vil09, Ch.12].

Lemma 2.4.1. *Let π be a semicoupling between abundant source σ and target τ. Then $proj_X \# \pi$ is absolutely continuous with respect to the source measure σ, and there exists a measurable function $f : X \to [0, 1]$ for which $f.\sigma = proj_X \# \pi$ and $\int_U f(x).d\sigma(x) = \pi[U \times Y]$ for every Borel subset U of X.*

Proof. The semicoupling π is a Borel measure and $proj_X$ is evidently Borel measurable, so $proj_X \# \pi$ is a Borel measure on X. The definition of $\pi \in SC(\sigma, \tau)$ implies $proj_X \# \pi$ is absolutely continuous with respect to σ. So Radon-Nikodym theorem implies $d(proj_X \# \pi)(x) = f(x) d\sigma(x) + d\nu(x)$, where $f(x) d\sigma(x)$ is absolutely-continuous part with respect to σ, and $d\nu(x)$ is the singular part.

Now we use Lebesgue's density theorem, which says: for σ-almost all $x \in X$, the limit

$$\lim_{r \to 0^+} (proj_X \# \pi)[B(x, r)] / \sigma[B(x, r)] =: f(x)$$

exists and is finite. Thus we obtain a Borel measurable function $f : X \to [0, 1]$ such that $f.\sigma = proj_X \# \pi$. For further details we refer the reader to [Vil03, Proposition 4.7, pp.132]. □

Definition 2.4.2 (Active Domain)**.** For given $\pi \in SC(\sigma, \tau)$, let $f = f_\pi$ be the Radon-Nikodym derivative of $proj_X \# \pi$ with respect to σ as in Lemma 2.4.1. Then $A := \{f > 0\} \subset X$ is the activated source domain of the semicoupling.

Now we formulate the BangBang principle, which characterizes the activated source of an optimal semicoupling. BangBang is classical and we refer the reader to [HL69, §II.12.1, pp.46], or [HS13, Prop 6.3, pp.2471], or [CM10, Prop 3.1, Thm 3.4].

Proposition 2.4.3 (BangBang Principle). *Let π be semicoupling in $SC(\sigma, \tau)$ and $(x', y') \in spt[\pi]$. For a real number $t < c(x', y')$, consider the Low-cost and High-cost regions*

$$L := \{c(-, y') < t\} \cap spt[\sigma], \quad \textit{and } H := \{c(-, y') \geq t\} \cap spt[\sigma]$$

in X. If $\sigma[L] > 0$ and $\sigma[H] > 0$, and if the restricted density $1_L \cdot f$ is not measurably identical to 1_L, then we can immediately construct an improved semicoupling $\tilde{\pi}$ such that $c[\tilde{\pi}] < c[\pi]$, and therefore π is not c-optimal.

Proof. Trivially we have $1_L.\sigma = 1_L.f.\sigma + 1_L.(1-f).\sigma$. If $1_L.(1-f).\sigma$ is not identically zero, then we can replace the semicoupling π with a semicoupling $\tilde{\pi}$ of strictly lower cost. Indeed mass is then more efficiently transport to y' from L rather than from H. Any active mass supported on H at density no greater than $\sigma[\{1_L.(1-f) > 0\}]$ is more efficiented routed out of L. Rerouting the mass defines a semicoupling $\tilde{\pi}$ with total cost strictly less than π. □

Corollary 2.4.4. *The marginal source density $f : X \to [0, 1]$ defined in 2.4.2 and Lemma 2.4.1 of an optimal semicoupling is measurably identical to the constant unit function $f = 1$ throughout the support $\{f > 0\}$. Therefore $proj_X\#\pi_{opt} = 1_A \cdot \sigma$ for every c-optimal semicoupling π_{opt} and some active domain $A \subset X$.*

Proof. The BangBang principle says c-efficient semicouplings π_{opt} draw from high-cost source regions only after the lower-cost resources have been totally exhausted. The active domain A therefore admits no nontrivial Low-cost/High-cost partitions as in 2.4.3. □

Next we clarify the role of Assumption (A3) from §1.3, which equivalently says the function $x \mapsto c(x, y)$ is non-constant on every open subset of $dom(c_y)$, for every $y \in Y$.

Lemma 2.4.5 ((Mono)). *Let $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be cost satisfying Assumptions (A0)–(A3), and let σ be a Radon measure on the source X. Then for every $y \in Y$, the single-variable function $t \mapsto \sigma[\{c_y < t\}]$ is strictly monotone-increasing for $t \in spt(c_y\#\sigma) \subset \mathbb{R}$.*

Proof. Since c_y is continuous, we find $\{t_1 < c_y < t_2\}$ is an open subset of X and $dom(c_y)$ for every $t_1, t_2 \in \mathbb{R}$. The subset $\{t_1 < c_y < t_2\}$ is nonempty for $t_1, t_2 \in spt(c_y\#\sigma)$. It is sufficient to prove

$$\sigma[\{t_1 < c_y < t_2\}] > 0 \tag{2.8}$$

for every connected interval $[t_1, t_2] \subset spt(c_y \# \sigma)$. But by definition of support, the strict-positivity of (2.8) follows. Thus the function $t \mapsto \sigma[\{c_y < t\}]$ as desired. □

We say a cost $c : X \times Y \to \mathbb{R}$ is monotone with respect to a source measure σ if the conclusion of Lemma 2.4.5 holds. Equivalently a cost is monotone with respect to σ if for every $t \in \mathbb{R}, y \in Y$, we have $\sigma[\{c_y = t\}] = 0$ whenever $\sigma[\{c_y \leq t\}] > 0$.

Lemma 2.4.6. *Let cost c satisfy Assumption (Mono) with respect to source measure σ. If $A \hookrightarrow X$ is the active domain of an c-optimal semicoupling, then there exists a unique measurable function $t : Y \to \mathbb{R}$ such that A can be expressed as the union of closed c_y-sublevel sets*

$$A := \cup_{y \in Y} \{x \mid c(x, y) \leq t(y)\}. \tag{2.9}$$

Proof. This is direct consequence of Proposition 2.4.3. For every $y \in Y$ we define $t(y)$ as the supremum of all $t \in \mathbb{R}$ for which $1_A.\sigma[\{c_y > t\}] > 0$. □

Remark. The measurable function $t : Y \to \mathbb{R}$ from Lemma 2.4.6 can be identified with the negative of a c-concave potential $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ satisfying the dual maximization program from Theorem 2.3.5, c.f. [CM10]. In otherwords $t(y) = -\psi(y)$ for all $y \in Y$. This explicit identification will be useful in §3.

Proposition 2.4.7. *[Unique Activation] Let (X, σ) be source and (Y, τ) target. Suppose the cost $c : X \times Y \to \mathbb{R}$ is monotonic with respect to σ (c.f. Lemma 2.4.5). Then the activated source domain $A = A_\pi$ of a c-optimal semicoupling $\pi \in SC(\sigma, \tau)$ is unique modulo sets of vanishing σ-measure.*

Proof. If A, A' are active source domains of c-optimal semicouplings π, π', then $1/2.\pi + 1/2.\pi'$ is also c-optimal, and with source-marginal $1/2.1_A + 1/2.1_{A'}$. Let $A \Delta A'$ be the set-theoretic symmetric difference. Then we have trivial identity

$$1/2.1_A + 1/2.1_{A'} = 1_{A \cap A'} + 1/2.1_{A \Delta A'},$$

where $A \cap A'$ and $A \Delta A'$ are disjoint. Now suppose $\sigma[A \Delta A'] > 0$. Then $A \Delta A'$ is nonempty. Selecting some $y \in A \Delta A'$, we consider the marginal cost $x \mapsto c_y(x) := c(x, y)$. If c satisfies (Mono), then c_y restricted to $A \Delta A'$ is nonconstant and $A \Delta A'$ can be partitioned into Low- and High-cost regions L, H satisfying the hypotheses of BangBang 2.4.3. But this contradicts the c-optimality of $\frac{1}{2}\pi + \frac{1}{2}\pi'$. So $\sigma[A \Delta A'] = 0$ and the active domains A, A' coincide σ-a.e. □

Figure 2.1: Disconnected Active Domain

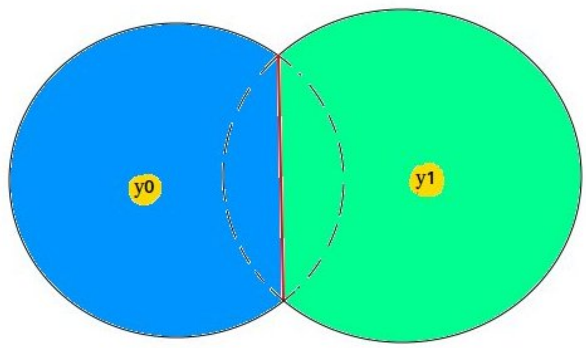

Figure 2.2: Connected Active Domain when $mass[\sigma]/mass[\tau] \approx 1^{+}$

2.5 Uniqueness of Optimal Semicouplings

Thus far we have established the existence of optimal semicouplings and uniqueness of active domains. Now we describe the (Twist) hypothesis and the uniqueness of optimal couplings when the source measure σ is absolutely continuous with respect to the reference source measure $\mathscr{H}^d_X$ in X.

The next definition elaborates Assumption (A4) from §1.3.

Definition 2.5.1 (**(Twist)**)**.** Let $c : X \times Y \to \mathbb{R}$ be cost function satisfying Assumptions (A0)–(A1). Then c satisfies (Twist) condition if for every $x' \in X$ the rule

$$y \mapsto \nabla_x c(x', y)$$

defines an injective mapping $\nabla_x c(x', \cdot) : dom(c_{x'}) \to T_{x'}X$.

Observe that (Twist) condition is equivalent to the function

$$x \mapsto c_\Delta(x; y_0, y_1) := c(x, y_0) - c(x, y_1)$$

admitting no critical points on X, whenever $y_0, y_1 \in Y$ are distinct. If X is compact closed manifold without boundary, then the standard Morse theory applied to $x \mapsto c_\Delta(x; y_0, y_1)$

implies the existence of critical points, and thus violates (Twist). Our settings assume X is a manifold-with-corners with nontrivial boundary $\partial X \neq \emptyset$. The (Twist) condition requires c_Δ admit no critical points on the interior of X, and all maxima/minima exist on the boundary. For instance, the repulsion cost constructed in Chapter 5 have the property that $c_\Delta(x; y_0, y_1)$ converges to $-\infty$ when $x \to y_1$, and converges to $+\infty$ when $x \to y_0$, and all other level sets $c_\Delta(-; y_0, y_1)^{-1}(s) \subset X$ are topologically connected and separating X into two components, for every $s \in \mathbb{R}$,

The Kantorovich duality yields a useful heuristic by which (Twist) condition ensures uniqueness of optimal couplings. The following lemmas are adapted from [GM96, Appendix C]. Recall the source X is equipped with a Riemannian distance function d_X, and $\dim(X)$-dimensional Hausdorff measure $\mathscr{H}_X$.

Lemma 2.5.2. *Let c be cost function satisfying Assumptions (A0), (A1), and (A2). Let D be compact geodesic disk in X, and V a compact subset of Y such that $c(x,y) < +\infty$ for every $x \in D$, $y \in V$. Define*

$$L(y) := \sup_{x,x' \in D} \frac{|c(x,y) - c(x',y)|}{d_X(x,x')}$$

for every $y \in V$. Then:

(i) the Lipschitz constant $L(y)$ is finite for every $y \in V$; and

(ii) the Lipschitz constant $y \mapsto L(y)$ is upper semicontinuous function of $y \in V$.

Proof. According to (A1), for every fixed $y \in V$ the function $x \mapsto c(x,y)$ is twice-continuously differentiable. So the supremum defining $L(y)$ is attained on the compact D. Moreover the convexity of D and the mean value theorem implies $L(y) = \sup_{x \in D} ||\nabla_x c(x,y)||$, where the supremum again exists and is finite after (A1). This proves (i).

Suppose $\{y_i | i = 1, 2, \ldots\}$ is sequence in $V \subset Y$ converging to limit $y_\infty \in V$. For each y_i, select some x_i' for which $L(y_i) = ||\nabla_x c(x_i', y_i)||$. But D is compact, so there exists convergent subsequence of $\{x_i'\}$. Extracting a convergent subsequence and relabelling indices, we find $\lim_{i \to +\infty} x_i' = x_\infty$ for some limit x_∞. Now Assumption (A2) says the function $(x,y) \mapsto ||\nabla_x c(x,y)||$ is upper semicontinuous, and therefore

$$||\nabla_x c(x_\infty, y_\infty)|| \geq \limsup_{i \to +\infty} ||\nabla_x c(x_i', y_i)||.$$

But $||\nabla_x c(x_i', y_i)|| = L(y_i)$ for every index i, and $L(y_\infty) \geq ||\nabla_x c(x_\infty, y_\infty)||$ according to the definition of L. Therefore $L(y_\infty) \geq \limsup_{i \to +\infty} L(y_i)$. This proves (ii).

□

Proposition 2.5.3. *Let $c : X \times Y \to \mathbb{R}$ be a cost satisfying Assumptions (A0)–(A5). Let $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ be a c-concave potential. Then ψ is locally Lipschitz on its domain $\psi : dom(\psi) \to \mathbb{R}$.*

Proof. The definition of c-concavity says $\psi(y) = \inf_{x \in X}\{\phi(x) + c(x, y)\}$ for some function $\phi : X \to \mathbb{R} \cup \{+\infty\}$. Thus ψ is the pointwise infimum of the X-family of functions $y \mapsto \phi(x) + c(x, y)$. Now we observe the infimum defining $\psi(y)$ can be restricted to a compact subset $K \subset X$, where K depends on y. Indeed the assumption (A1) implies $c(x, y)$ diverges to $+\infty$ whenever x diverges in X. Moreover if ϕ is c-convex, then ϕ is lower semicontinuous and attains its maximum on any compact subset K of X. Therefore $\phi(x) + c(x, y)$ is bounded on compact subsets of X, and the infimum defining ψ can be restricted to a compact subset K wherever $\phi(y) \neq -\infty$. So $\psi(y) = \inf_{x \in K}\{\phi(x) + c(x, y)\}$, and Assumption (A5) implies this family is uniformly Lipschitz. Therefore $\psi(y)$ is locally Lipschitz function for $y \in dom(\psi)$. □

Recall the definition of semiconvexity [Vil09, Definition 10.10, pp.228]:

Definition 2.5.4 (Semiconvexity)**.** A function $\phi : X \to \mathbb{R} \cup \{+\infty\}$ is semiconvex on an open subset U of X with modulus $C > 0$ at $x_0 \in X$ if for every constant-speed geodesic path $\gamma(t)$, for $0 \leq t \leq 1$ whose image is included in U, the inequality

$$\phi(\gamma(t)) \leq (1 - t)\phi(\gamma(0)) + t\phi(\gamma(1)) + t(1 - t)C dist(\gamma(0), \gamma(1))^2 \tag{2.10}$$

is satisfied for $0 \leq t \leq 1$. The function is locally semiconvex if ϕ is semiconvex at every $x_0 \in U$, with respect to a modulus $C > 0$ depending uniformly on $\gamma(0)$, $\gamma(1)$ varying in compact subsets K of U.

Lemma 2.5.5. *Let $c : X \times Y \to \mathbb{R}$ be cost satisfying Assumptions (A0)–(A2). Then every c-convex potential$\psi^c : X \to \mathbb{R} \cup \{+\infty\}$ is $\mathscr{H}_X$-almost everywhere locally-Lipschitz on its domain $dom(\psi^c) \subset X$. Furthermore every c-convex potential is locally-semiconvex on $dom(\psi^c)$.*

Proof. The definition of c-convexity implies $\phi(x) = \psi^c(x) = \sup_{y \in Y}\{\psi(y) - c(x, y)\}$ for every $x \in X$. Assumption (A0) implies the cost $(x, y) \mapsto c(x, y)$ is bounded with bounded sublevels. So for every x such that $\psi^c(x) < +\infty$, the supremum defining $\psi^c(x)$ can be restricted to a compact subset $K \subset Y$, where $K = K(x)$ varies with x. From Lemma 2.5.2 the family of functions $\{x \mapsto \psi(y) - c(x, y) \mid y \in dom(c_x)\}$ are locally Lipschitz, with respect to some finite Lipschitz constant L and independant of y. Indeed the upper semicontinuity of $y \mapsto L(y)$ implies $\sup_{y \in K} L(y)$ is attained and finite over the compact

K. This implies ψ^c is locally Lipschitz, with finite Lipschitz constant as satisfied by the family $\{\psi(y) - c(x,y) \mid y \in K\}$. Furthermore, from the definition of c-concavity, for every $x \in dom(\psi^c) \subset X$ we find $\psi^c(x)$ inherits the same local semiconvexity constants as the family of functions $\{\psi(y) - c(x,y) \mid y \in K\}$. $\square$

Under Assumptions (A0)–(A2) the graph of $\partial^c\psi^c$ is a closed subset of $X \times Y$, as the following Lemma shows.

Lemma 2.5.6. *Let $c : X \times Y \to \mathbb{R}$ be cost satisfying Assumptions (A0)–(A2). Let $\phi = \psi^c$ be a c-convex potential on X. Then the c-subdifferential $\partial^c\phi(x)$ is lower semicontinuous with respect to $x \in dom(\phi)$. So if $x_1, x_2, \ldots$ is a sequence in $dom(\phi)$ converging to $x_\infty \in dom(\phi)$, then the Gromov-Hausdorff limit $\lim_{k\to+\infty} \partial^c\phi(x_k)$ is contained in $\partial^c\phi(x_\infty)$.*

Proof. Lemma 2.3.4 implies ϕ and $\psi = \phi^c$ are lower semicontinuous and upper semicontinuous, respectively. Let (x_k, y_k) be a sequence in $dom(\phi) \times dom(\psi)$ with $y_k \in \partial^c\phi(x_k)$ for $k = 1, 2, \ldots$. Then $-\phi(x_k) + \psi(y_k) = c(x_k, y_k)$ for all k. But semicontinuity implies

$$\liminf_{k\to+\infty} \phi(x_k) \geq \phi(x_\infty), \quad \limsup_{k\to+\infty} \psi(y_k) \leq \psi(y_\infty).$$

Therefore

$$\limsup_{k\to+\infty} -\phi(x_\infty) + \psi(y_\infty) \geq c(x_\infty, y_\infty),$$

which implies $y_\infty \in \partial^c\phi(x_\infty)$, as desired.

$\square$

Proposition 2.5.7. *Let c be cost satisfying Assumptions (A0)–(A2). Then c-convex potentials ψ^c are $\mathscr{H}_X$-almost everywhere differentiable on $dom(\psi^c) \subset X$. Thus $dom(D\psi^c)$ is a full $\mathscr{H}_X$-measure subset of $dom(\psi^c)$.*

Proof. According to Lemma 2.5.5, the c-convex potentials ψ^c are locally Lipschitz on their domains $dom(\psi^c) \subset X$. Rademacher's theorem, [Vil09, Thm 10.8, pp.222], says locally Lipschitz functions are almost-everywhere differentiable on $dom(\psi^c)$ with respect to $\mathscr{H}_X$ on their domains. Therefore $\nabla_x\psi^c$ exists almost everywhere on $dom(\psi^c)$ as desired. $\square$

These preliminaries lead to the following standard uniqueness result, c.f. [Vil09, Thms 10.28, 10.42]:

Theorem 2.5.8. *Suppose cost c satisfies Assumptions (A0)–(A4). Let source σ be absolutely continuous with respect to $\mathscr{H}_X$. Suppose σ is abundant with respect to target τ and (2.1) holds. Then there exists a unique c-optimal semicoupling modulo sets of measure* *zero between σ and τ, and this c-optimal semicoupling is supported on the graph of a measurable map $T : A \to Y$, where $A \subset X$ is the active domain.*

Proof. By Theorem 2.4.7 we know there exists a unique active domain A for the c-optimal semicouplings. We assume this unique active domain A has been identified, and restrict ourselves to the $c|A$-optimal coupling problem between $1_A\sigma$ and τ. From Theorem 2.3.5 we know the dual program admits $c|A$-concave maximizers $\psi, \psi^{c|A} =: \phi$ on Y, A, respectively. By Lemma 2.5.7 the $c|A$-convex potential ϕ is almost everywhere differentiable on its domain $dom(\phi) \subset A$. Let $dom(D\phi)$ denote the domain of differentiability in $dom(\phi)$. Then $dom(D\phi)$ is a full-measure subset of $dom(\phi)$ by (2.5.7). By the definition of $c|A$, we find $dom(D\phi)$ is full measure subset of A.

Under the (Twist) condition, the rule

$$T(x) := \nabla_x c(x, \cdot)^{-1}(-\nabla_x \phi(x)) \tag{2.11}$$

is a well-defined map $T : dom(D\phi) \to Y$.

Claim #1: Under the above hypotheses, T defines a Borel-measurable map $T : A \to Y$ which pushes forward the restricted source $1_A\sigma$ to τ. This observation is due to Gangbo [Gan95][pp.8-9], c.f. [McC01, Proof of Thm. 9].

Assuming Claim #1, the pushforward $(Id \times T)\#1_A\sigma$ defines a semicoupling π^* in $SC(\sigma, \tau)$. But π^* is $c|A$-optimal, since $-\psi^c(x) + \psi(y) = c(x, y)$ holds almost everywhere (according to (2.11)) on the support of π^*, and therefore "sup=min" in (2.5). Thus π^* is supported on the graph of T. And in fact a standard argument proves that every c-optimal semicoupling is necessarily supported on the graph of some map of the form (2.11), c.f. [Vil09][Ch.10, pp.216]. But the property that optimal semicouplings are supported on the graphs of measurable maps T implies the semicouplings are unique, modulo sets of zero measure. Indeed if $\pi^*, \pi^{*\prime}$ are $c|A$-optimal, then their convex combination $\frac{1}{2}\pi^* + \frac{1}{2}\pi^{*\prime}$ is again c-optimal. But this convex combination cannot be supported on the graph of a measurable function, unless the graphs supporting π^*, $\pi^{*\prime}$ coincide almost-everywhere and then π^*, $\pi^{*\prime}$ coincide almost-everywhere. □

Chapter 3

Deforming Source onto Active Domain $X \rightsquigarrow A$

The present Chapter establishes Theorem 1.4.1 from the Introduction.

3.1 Statement of Theorem 1.4.1

The previous chapter described the background on c-optimal semicouplings and Kantorovich duality. The present chapter 3 is topological and contains our first result, namely Theorem 1.4.1 from the Introduction, c.f. Theorem 3.1.1. A cost satisfying (A0)–(A3) has a uniquely defined closed active domain $A \subset X$ (recall Lemma 2.4.6 and equation (2.9)). Our goal is to identify conditions for which the inclusion $A \hookrightarrow X$ is a homotopy-isomorphism. We denote the source space $X =: Z(\emptyset)$, and define the activated source $A =: Z_1$. We prove below that sufficient Halfspace conditions imply $A \hookrightarrow X$ is a homotopy-isomorphism when cost satisfies Assumptions (A0)–(A4) and (A6). These deformations are generalized in the next chapters, where we introduce Kantorovich's contravariant singularity functor $Z : 2^Y \to 2^X$, and describe homotopy-reductions $Z(Y_I) \rightsquigarrow Z(Y_J)$ between various cells for closed subsets $Y_I \hookrightarrow Y_J$.

For simplicity, we shall assume $dom(c_x) = Y$ for every x. The argument can be extended without difficulty to the case where $dom(c_x)$ is a proper subset of Y for $x \in X$.

Theorem 3.1.1. *Let c be cost satisfying Assumptions (A0)–(A4). Suppose the source σ and target measure τ are absolutely continuous with respect to $\mathscr{H}_X$, $\mathscr{H}_Y$, respectively and satisfy* (1.1)*. Let π be a c-optimal semicoupling from σ to τ, with dual c-concave potential $\psi : Y \to \mathbb{R} \cup -\infty\}$, $\psi^{cc} = \psi$ (2.3.5). Let $A = \{x | c(x,y) - \psi(y) \leq 0\}$ be the active domain (2.4.7). Let $\beta := \dim(Y) + 2$.*

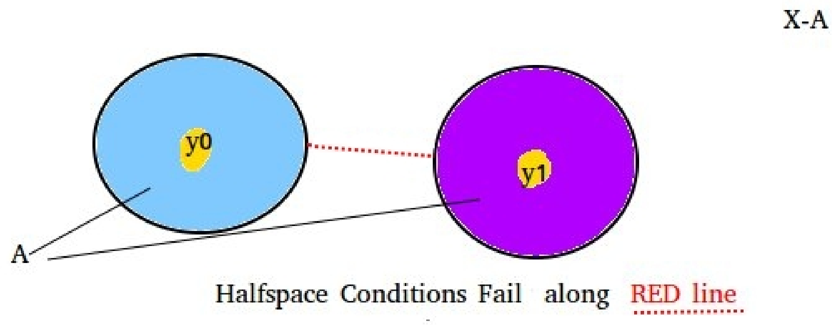

Figure 3.1: Halfspace Conditions fail and active domain is not homotopy-equivalent to source.

Suppose every $x \in X - A$ has the property that

$$\eta_{avg}(x) := (\mathscr{H}_Y[Y])^{-1} \int_Y (c(x,y) - \psi(y))^{-\beta} \cdot \nabla_x c(x,y) d\mathscr{H}_Y(y), \tag{3.1}$$

is bounded away from zero, uniformly with respect to $y \in Y$. Then the inclusion $A \hookrightarrow X$ is a homotopy-isomorphism, and there exists explicit strong deformation retract $h : X \times [0,1] \to X$ of X onto $A = h(X,1)$.

The arguments of 3.1.1 demonstrate that the disactivated domain $X - A$ deformation retracts onto every ϵ-neighborhood of A in X. I.e., X deformation retracts onto

$$(A)_\epsilon := \{x \in X \mid dist(x, A) \leq \epsilon\}$$

for every $\epsilon > 0$. The purpose of 3.1.1 is to obtain continuous deformations at $\epsilon = 0$. Our strategy studies the flow generated by an averaged vector field $\eta_{avg}(x) = \nabla_x f_{avg}$, which is the gradient of an averaged potential f_{avg}. C.f. [Nee85, §3].

The hypothesis that $\eta_{avg}(x)$ is uniformly bounded away from zero is a weak version of the following condition.

Definition 3.1.2. A collection $E = \{\eta_i \mid i \in I\} \subset T_x X$ of tangent vectors satisfies the **Halfspace condition** if there exists a nonzero linear functional $\ell : T_x X \to \mathbb{R}$ with $\ell(\eta_i) > 0$ simultaneously for all $i \in I$.

Equivalently, Halfspace condition says the convex hull $conv[E] = conv[\{\eta_i \mid i \in I\}] \subset T_x X$ does not contain the origin $0 \in T_x X$.

3.2 Averaged Gradients and Finite-Time Blow-Up

Suppose we have a Radon measure $\bar{\nu}$ absolutely continuous with respect to $\mathscr{H}_Y$, and we take the average gradient with respect to $\bar{\nu}$. The following lemma formally establishes that this averaged-gradient is indeed the gradient field of a continuously differentiable "averaged" potential.

Lemma 3.2.1. *Let $\beta > 0$. Let $\nu_1, \nu_2, \nu_3, \ldots$ be a sequence of empirical probability measures, i.e. renormalized sums of Dirac masses, which converge as $N \to +\infty$ in the weak-$*$ topology to the renormalized probability measure $(\bar{\nu}[Y])^{-1} \cdot \bar{\nu}$ on Y. Then:*

(i) For $x \in X - A$, the limit

$$\lim_{N \to +\infty} \int_Y (c(x,y) - \psi(y))^{-\beta} d\nu_N(y) \tag{3.2}$$

exists and converges to the finite integral

$$f_{avg}(x) := (\bar{\nu}[Y])^{-1} \int_Y (c(x,y) - \psi_y))^{-\beta}.d\bar{\nu}(y). \tag{3.3}$$

(ii) The rule $f_{avg} : X - A \to \mathbb{R}$ defines a continuously differentiable function with gradient

$$\nabla_x f_{avg} = (\bar{\nu}[Y])^{-1} \int_Y \nabla_x (c(x,y) - \psi(y))^{-\beta}.d\bar{\nu}(y).$$

Proof. If c satisfies (A0)–(A4), then the limit defining f_{avg} converges uniformly on compact subsets of $X - A$. So the limit (3.2) exists and is finite. Moreover the uniform convergence on compact subsets implies (ii), since the approximants are continously differentiable on $X - A$. Therefore $\nabla_x f_{avg}$ is the average of $\nabla_x (c(x,y) - \psi(y))^{-\beta}$ with respect to $\bar{\nu}$, as desired. $\square$

For $y \in Y$, $x \in X - A$, and $\beta > 0$, we abbreviate

$$f_y(x) := (c(x,y) - \psi(y))^{-\beta}. \tag{3.4}$$

According to 3.2.1 we define

$$favgdeff_{avg} : X - A \to \mathbb{R}, \quad f_{avg}(x) = (\bar{\nu}[Y])^{-1} \int_Y f_y(x).d\bar{\nu}(y). \tag{3.5}$$

Definition 3.2.2 (**Property (C)**)**.** The collection of functions $\{f_y | \ y \in Y\}$ satisfies Property (C) throughout $X - A$ with respect to the uniform probability measure $\frac{1}{\bar{\nu}[Y]} \ \bar{\nu}$

if there exists constant $C > 0$ such that

$$||\nabla_x f_{avg}|| \geq C \int_Y ||\nabla_x f_y||.d\bar{\nu}(y) \tag{3.6}$$

pointwise throughout $X - A$.

When Y is finite, $\#(Y) < +\infty$, the estimate (3.6) requires the ratio

$$||\nabla_x f_{avg}|| / \max_{y \in Y} ||\nabla_x f_y||$$

be uniformly bounded away from zero throughout $X - A$. In this case, Property (C) and Halfspace Condition 3.1.2 ensures the divergence of the average $\nabla_x f_{avg}$ whenever a gradient summand $\nabla_x f_y$ diverges.

When Y is infinite with positive dimension, the pointwise divergence of an integrand $||f_y|| \to +\infty$ need not imply the divergence of the average f_{avg} for every choice of $\beta > 0$. The rate at which f_y diverges must be sufficiently large. When cost c satisfies Assumptions (A0)–(A5), Lemma 2.5.3 proves that c-concave potentials $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ are locally Lipschitz throughout $dom(\psi)$. This implies $\beta = \dim(Y) + 2$ is sufficient. If x_{kk} is a countable sequence in $X - A$, then $f_{avg}(x_k)$ diverges to $+\infty$ if and only if $f_y(x_k)$ diverges for y belonging to some subset $V \subset Y$.

Lemma 3.2.3. *Let $c : X \times Y \to \mathbb{R}$ be cost satisfying (A0)–(A5), as above, and $A \subset X$ the active domain of a c-optimal semicoupling. Fix $y_0 \in Y$, and abbreviate $f_0(x) := f_{y_0}(x) = (c(x, y_0) - \psi(y_0))^{-\beta}$ for some $\beta > 0$, $x \in X$. Suppose:* ***(a)*** *$\nabla_x c(x, y_0)$ is uniformly bounded away from the origin; and*

b *$\nabla^2_{xx} c(x, y_0)$ is uniformly bounded above with respect to $x \in X - A$.*

Then for every $K > 0$, there exists $\epsilon > 0$ such that $\nabla^2_{xx} f_0 \geq K.Id > 0$ in the direction of $\nabla_x c(x, y_0)$ throughout the ϵ-neighborhood of $\{f_0 = +\infty\}$ in $X - A$.

Proof. The function f_0 is well-defined on $\{c(x, y_0) > \psi(y_0)\} \subset X$. If x converges to $x_\infty \in \{c(x, y_0) \leq \psi(y_0)\}$, then both f_0, $\nabla_x f_0$ diverge to infinity. We find $\nabla^2_{xx} f_0$ is equal to

$$\beta(1+\beta)(c(x, y_0) - \psi(y_0))^{-2-\beta} \nabla_x c(x, y_0) \otimes \nabla_x c(x, y_0) - \beta(c(x, y_0) - \psi(y_0))^{-1-\beta} \nabla^2_{xx} c(x, y_0).$$

By Proposition 3.2.2 the gradients $\nabla_x c(x, y_0)$ are uniformly bounded away from zero in neighborhoods of $\{f_0 = +\infty\}$. Assumption (A2) implies $\nabla^2_{xx} c(x, y_0)$ is uniformly bounded above on superlevel sets $\{f_0 \geq T\}$ for all $T > 0$. Factoring out the term

$\beta(c(x,y_0)-\psi(y_0))^{-1-\beta}$, we find $\nabla^2_{xx}f_0$ is positively proportional

$$(1+\beta)(c(x,y_0)-\psi(y_0))^{-1}\nabla_x c(x,y_0)\otimes\nabla_x c(x,y_0)-\nabla^2_{xx}c(x,y_0). \tag{3.7}$$

We observe (3.7) diverges to $+\infty$ with $(c(x,y_0)-\psi(y_0))^{-1}$. This implies $\nabla^2_{xx}f_0$ is positive semidefinite when $c(x,y_0)-\psi(y_0)>0$ is sufficiently small, and strongly convex

$$\nabla^2_{xx}f_0 \geq K > 0$$

in the direction of $\nabla_x c(x,y_0)$.

By Assumption (A1) the sublevels $\{x \mid \psi(y_0)\leq c(x,y_0)\leq \psi(y_0)+\epsilon'\}$ are compact subsets of $X-A$ for every $\epsilon'>0$. This implies a sufficiently small $\epsilon>0$ exists for which $\nabla^2_{xx}f_0 \geq K.Id > 0$ throughout the ϵ-neighborhood of $\{f_0=+\infty\}$ in the direction $\nabla_x c(x,y_0)$. □

Example. To illustrate Lemma 3.2.3, consider the $f(x)=x^{-\beta}$ for $x\geq 0$, $\beta>0$. The gradient flow $x'=(-\beta)x^{-1-\beta}$ is bounded away from zero in neighborhoods of the pole at $x=0$, and indeed diverges to $+\infty$. Moreover $f''(x)$ is obviously bounded away zero and diverging to $+\infty$ as $x\to 0^+$.

For initial condition $x_0>0$, the integral curve of the negative gradient flow is equal to $x(s)=(x_0^{2+\beta}-\beta(\beta+2)s)^{1/(2+\beta)}$, which converges in finite time to the pole at $x=0$ over the interval $0\leq s\leq \frac{1}{\beta(\beta+2)}x_0^{2+\beta}$. Thus we find $\omega(x_0)=\frac{1}{\beta(\beta+2)}x_0^{2+\beta}$ varies continuously with respect to $x_0>0$, and is even Lipschitz. Compare Lemma 3.2.6 below.

The blow-up in finite time is typical property of the gradient flow defined by the potentials f_0. Actually our applications require verifying these same properties for the averaged potential f_{avg} and its gradient $\nabla_x f_{avg}$.

Lemma 3.2.4. *Let f_{avg} be the average defined in Lemma 3.2.1, equation (3.3), with exponent $\beta=\dim(Y)+2$. If the distance from $x\in X-A$ to the boundary ∂A is sufficiently small, then f_{avg} is strongly convex in the direction of $\nabla_x f_{avg}$.*

Proof. Let x_k be a sequence in $X-A$ converging to a point $x_\infty\in\{f_{avg}=+\infty\}$. The choice of β says f_{avg} diverges if and only integrands f_y diverge, and there exists a subset $V\subset Y$ such that f_y diverges to $+\infty$ for every $y\in V$. The divergence of $f_y, y\in V$ also implies the divergence of the gradients $\nabla_x f_y$ and Hessians $D^2_{xx}f_y$, (see (3.7) in proof of 3.2.3). Moreover the Hessians $D^2 f_{avg}|_{x_k}$ are positive semidefinite when k is sufficiently large, being the asymptotic to the average rank-one quadratic forms $\langle\nabla_x f_y,-\rangle^2$. So

$D^2 f_{avg}[\nabla_x f_{avg}]$ is asymptotic to the average of

$$\langle \nabla_x f_y, \nabla_x f_{avg} \rangle^2 \tag{3.8}$$

for $y \in V$. Now we claim

$$\lim_{k \to +\infty} \langle \nabla_x f_y|_{x_k}, \nabla_x f_{avg}|_{x_k} \rangle^2 = +\infty, \text{ unless } \nabla_x f_y, \nabla_x f_{avg} \text{ are orthogonal.}$$

The (UHS) conditions imply $\nabla_x f_{avg}$ is uniformly bounded away from zero, and therefore the inner products (3.8) are not identically zero for all $y \in V$. This implies the divergence of $D^2 f_{avg}[\nabla_x f_{avg}]|_{x_k}$ as $k \to +\infty$.

□

Lemma 3.2.5 (Finite-time Blow-up)**.** *Suppose the functions $\{f_y \mid y \in Y\}$ satisfy (3.6) as above. Then for every initial value $x_0 \in dom(f_{avg})$, the gradient flow defined by the average gradient $x' = \nabla_x f_{avg}$ diverges to infinity in finite time.*

Proof. The estimate (3.6) shows the gradient $\nabla_x f_{avg}$ is uniformly bounded away from zero in the neighborhoods of the poles $\{f_{avg} = +\infty\}$ in $X - A$. Moreover f_{avg} is asymptotically convex in neighborhoods of the poles using Lemma 3.2.4 along directions of $\nabla_x f_{avg}$. This implies the gradient flow $x' = \nabla_x f_{avg}$, $x(0) = x_0$ blows-up in finite-time for every initial value $x_0 \in X - A$. □

Informally the estimate (3.6) implies every step in the discretized gradient flow (e.g., Euler scheme) has a definite size. Meanwhile the asymptotic convexity of Lemma 3.2.3 implies the discretized gradient flow well approximates the continuous gradient flow. But if step-sizes have a definite magnitude, then we definitely approach the poles after a finite number of steps and the integral curves blow-up in finite time.

The blow-up in finite time (Lemma 3.2.5) implies the maximal forward-time interval of existence for the gradient flow is a bounded interval $I(x_0) := [0, \omega(x_0)) \subset [0, +\infty)$. For general ordinary differential equations, it's known that $\omega(x_0)$ is lower semicontinuous as a function of x_0: for a sequence of initial values $x_0, x_1, \ldots$ converging to some x_∞, we have $\omega(x_\infty) \leq \liminf_{k \to +\infty} \omega(x_k)$. See [Har64, Theorem 2.1, pp.94]. In our particular setting, it is further necessary to establish the continuity of this maximal interval of existence. This is established in Lemma 3.2.6 below.

Lemma 3.2.6. *Let c satisfy Assumptions (A0)–(A5). Then the maximal intervals of existence $I(x_0) = [0, \omega(x_0))$ of solutions to the initial value problem 3.10 vary continuously* *with respect to the initial point* x_0*. In otherwords,* $x_0 \mapsto \omega(x_0)$ *varies continuously with* $x \in X - A$*.*

Proof. Our assumptions imply the domain $X - A$ is a complete open set. Furthermore $\eta_{avg}(x)$ being uniformly bounded away from zero implies the trajectories $s \mapsto \Psi(x_0, s)$ are finite on compact subsets of $X - A$. Therefore $\omega(x_0)$ is characterized by the two limits

$$\lim_{s \to \omega(x_0)^-} \Psi(x_0, s) \in \partial A, \qquad \lim_{s \to \omega(x_0)^-} ||\Psi(x_0, s)|| = +\infty,$$

and here we take advantage of the divergence $||\eta_{avg}(x)|| \to +\infty$ as $x \to \partial A$.

Moreover the asymptotic concavity (Proposition 3.2.4) of the average potential defining $\eta_{avg}(x)$ implies the flow defined by 3.10 is asymptotically contracting. This implies $\omega(x_0)$ is actually a Lipschitz function of x_0, i.e. satisfying $|\omega(x_1) - \omega(x_0)| \leq C.||x_1 - x_0||$ for some constant $C = C(x_0)$ depending on $x_0 \in X - A$. Hence ω is a continuous function, as desired.

$\square$

3.3 Proof of Theorem 1.4.1

Now we establish Theorem 3.1.1, which is Theorem 1.4.1 from the Introduction.

Proof of Theorem 3.1.1. According to Lemma 2.4.6 the active domain A can be expressed as $\cup_{y \in Y}\{x \mid c(x,y) \leq \psi(y)\}$ for the Kantorovich potential $\psi : Y \to \mathbb{R}$. If Halfspace Condition is satisfied at $x \in X - A$, then the collection of gradients $\{\nabla_x c(x,y) \mid y \in Y\}$ are all nonzero vectors and occupy some common nontrivial halfspace of $T_x X$. If we define

$$\eta(x,y) := (c(x,y) - \psi(y))^{-1-\beta} \cdot \nabla_x c(x,y),$$

then likewise $\{\eta(x,y) \mid y \in Y\}$ is a collection of vectors satisfying Halfspace condition. We divide the argument into two cases.

(Case I) Assume Y is finite with $N = \#(Y) < +\infty$. Define

$$\eta_{avg}(x) := N^{-1} \sum_{y \in Y} [(c(x,y) - \psi(y))^{-1-\beta} \cdot \nabla_x c(x,y)]. \tag{3.9}$$

Evidently when Y is finite, the sum (3.9) is finite vector. Then $x \mapsto \eta_{avg}(x)$ is a well-defined nonvanishing vector field on $X - A$ which diverges whenever a denominator converges $c(x,y) \to \psi(y)^+$. Thus $\eta(x, avg)$ is finite if and only if $x \in X - A$.

We propose integrating the vector field $x \mapsto \eta_{avg}(x)$ throughout $X - A$ to obtain the desired retraction. For every initial point $x_0 \in X - A$, there exists a unique solution $\Psi(x_0, s)$ to the ordinary differential equation

$$\frac{d}{ds}|_{(x_0,s)}\Psi = \eta_{avg}(\Psi(x_0, s)), \quad \Psi(x_0, 0) = x_0, \tag{3.10}$$

and defined over a maximal interval of existence $I(x_0) = [0, \omega(x_0))$.

According to Lemma 3.2.6 the maximal interval $I(x_0)$ varies continuously with respect to $x_0 \in X - A$. Moreover orbits $\{\Psi(x_0, s) \mid s \in I(x_0)\}$ converge in finite-time to the boundary $\partial(X - A)$ for every initial value $x_0 \in X - A$.

Finally, Lemma 3.2.6 below proves the orbits can be continuously reparameterized to obtain a continuous mapping $\Psi' : (X - A) \times [0, 1] \to X$ defined by

$$\Psi'(x_0, s) = \Psi(x_0, s\omega(x_0)). \tag{3.11}$$

(Case II) Suppose Y is infinite set, with uniform measure $\mathscr{H}_Y$. We define $\eta_{avg}(x)$ according to the vector-valued Bochner integral

$$\eta_{avg}(x) := (\int_{dom(c_x)} d\mathscr{H}_Y)^{-1} \int_{dom(c_x)} (c(x, y) - \psi(y))^{-1-\beta} \cdot \nabla_x c(x, y) d\mathscr{H}_Y(y), \tag{3.12}$$

where $dom(c_x)$ is closed compact subset of Y for every $x \in X$ by Assumption (A0). Assumption (A5) implies the exponent $\beta = \dim(Y) + 1$ is suitable according to Proposition 2.5.3. Lemma 3.2.1 implies the vector field $\eta_{avg}(x) = \nabla_x f_{avg}$ is the gradient of a continuously differentiable potential f_{avg} defined on $X - A$.

The proof proceeds as in (Case I). The vector field $x \mapsto \eta_{avg}(x)$ is well-defined nonvanishing vector field on $X - A$ which diverges to $+\infty$ whenever some denominator converges $c(x, y) \to \psi(y)^+$. We integrate the gradient fields and obtain the retraction of $X - A$ onto the poles ∂A. The flow converges in finite-time by Lemma 3.2.5. We reparameterize the flow according to equation (3.11), and obtain a continuous deformation using Lemma 3.2.6. $\square$

Chapter 4

Kantorovich Singularity and Topological Theorem 1.4.2

The present chapter develops these ideas further, describing topological properties of the singularities of optimal semicouplings. The main result is Theorem 1.4.2 from the Introduction, c.f. Theorems 4.4.3–4.4.4. Here we include the central definition of our thesis, namely Kantorovich's contravariant singularity functor $Z : 2^Y \to 2^X$ defined $Z = Z(c, \sigma, \tau)$ with respect to a choice of cost c satisfying Assumptions (A0)–(A4), and source and target measures σ, τ on X and Y, respectively.

4.1 Kantorovich's Contravariant Singularity Functor

Our thesis proposes a bridge between measure and algebraic topology. The bridge is realized by the contravariant functor $Z : 2^Y \to 2^X$, where 2^Y designates the category of closed subsets Y_I of Y, and where morphisms are the inclusions $Y_I \subset Y_J$ between closed subsets Y_I, Y_J whenever they exist. Likewise for 2^X. The singularity functor $Z : 2^Y \to 2^X$ is defined relative to a cost c on $X \times Y$ satisfying Assumptions (A0)–(A4). Let σ, τ be absolutely continuous with respect to the Hausdorff measures $\mathscr{H}_X$, $\mathscr{H}_Y$ on X, Y respectively. By Theorems 2.4.7 and 2.5.8 there exists unique c-minimizing measures, and c-concave potentials $\psi^{cc} = \psi$ on Y. The c-subdifferential $\partial^c \psi$ is uniquely determined, but the choice of potentials – i.e. the maximizers in dual maximization program (2.5) – are generally nonunique.

Definition 4.1.1 (Kantorovich Singularity)**.** The Kantorovich singularities in X are the closed subvarieties $Z(Y_I)$ functorially assembled from the closed subsets Y_I of Y by the rule

$$Y_I \mapsto Z(Y_I) := \cap\{\partial^c\psi(y)|y \in Y_I\}.$$

We declare

$$Z(\emptyset) := X$$

for the empty subset $\emptyset$ of Y.

The definition $Z(Y_I) = \cap_{Y_I} Z(y)$ yields a *contravariant functor* $Z : 2^Y \to 2^X$ where morphisms $Y_I \hookrightarrow Y_J$ correspond to $Z(Y_I) \hookleftarrow Z(Y_J)$ in X. The contravariant functor Z is uniquely prescribed by the choice of (c, σ, τ) under the above assumptions. For (c, σ, τ) satisfying the assumptions above, the singularity $Z = Z(c, \sigma, \tau)$ will generally admit many closed subsets $Y_I \subset Y$ for which the cells $Z(Y_I) \subset X$ are empty $Z(Y_I) = \emptyset$. It is useful to restrict ourselves to the nontrivial image of Z and formally define the support.

Definition 4.1.2. The support of the contravariant functor $Z : 2^Y \to 2^X$ is the subcategory of 2^Y, denoted *spt*(Z), whose objects are the closed subsets Y_I of Y for which $Z(Y_I)$ is nonempty subset of X. So

$$spt(Z) = \{Y_I \subset Y | Z(Y) \neq \emptyset\}.$$

Note that $\emptyset \subset Y$ is object in subcategory *spt*(Z), since $Z(\emptyset) = X$ according to Definition 4.1.1.

For given $x \in X$, we interpret $Z'(x) := Z(\partial^c\psi^c(x))$ as a "cellular neighborhood" of x in the active domain $A =: Z_1 = \cup_{y \in Y} Z(y)$. The c-concavity $\psi^{cc} = \psi$ provides explicit equations describing $Z'(x)$. See Section 4.2 and (4.2).

We conclude this section with the observation that the graph of the singularity functor $Z(\sigma, \tau, c)$ is closed, and hence "upper semicontinuous" with respect to Gromov-Hausdorff limits of (σ, τ, c). We refer the reader to any standard textbook on category theory for definition of "natural transformations" of functors, e.g. [Lan05, Ch.1]. Recall the graph of the subdifferential $Id \times \partial\phi$ of a lower semicontinuous convex function ϕ is a closed subset of $\mathbb{R}^N \times \mathbb{R}^N$.

Proposition 4.1.3. *Let X, Y be Riemannian manifolds-with-corners. Let σ_k, τ_k, $k \geq 1$, be a sequence of measures which converge in weak-$*$ topology to σ, τ, respectively. Let $c_k : X \times Y \to \mathbb{R}$ be a sequence of costs which converge pointwise to a cost $c : X \times Y \to \mathbb{R}$. Suppose c_k, c satisfy Assumptions (A0)–(A3).*

Then the correspondance

$$(\sigma_k, \tau_k, c_k) \mapsto Z(\sigma_k, \tau_k, c_k)$$

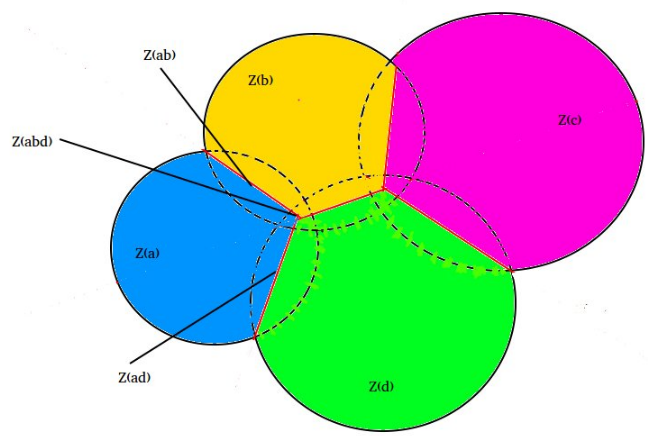

Figure 4.1: Singularity structure Z on active domain $A = Z_1$

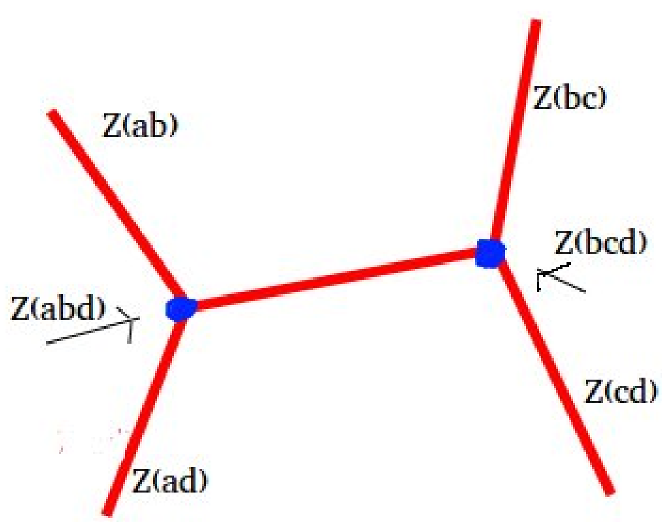

Figure 4.2: Singularity structure on Z_2

varies upper semicontinuously, and there exists an injective natural transformation between the functors

$$Z(\sigma, \tau, c) \hookrightarrow \lim_{k \to +\infty} Z(\sigma_k, \tau_k, c_k).$$

In particular, for every closed subset Y_I of Y, the cell $Z(Y_I)$ is an embedded subset of the Gromov-Hausdorff limit $\lim_{k\to+\infty} Z_k(Y_I)$.

Proof. See Lemma 2.5.6, c.f. [Vil09, Thm. 28.9, pp.780–790]. □

The point of Proposition 4.1.3 is that the singularities of the limit (σ, τ, c) are no more complicated than the approximant singularities of (σ_k, τ_k, c_k). In fact the singularity often simplifies in various limits. Proposition 4.1.3 is familiar property of lower semicontinuous convex functions: if $\phi_k : X \to \mathbb{R}$ is a sequence of lower semicontinuous convex functions which converge pointwise to a limit $\lim_{k\to+\infty} \phi_k = \phi_0$, then for every $x \in X$, the subdifferential $\partial\phi_0(x)$ is a subset of the Hausdorff limit $\lim_{k\to+\infty} \partial\phi_k(x)$.

4.2 Local Topology and Local Dimensions of Z

The present section examines the local differential topology and dimensions of Kantorovich's contravariant functor $Z : 2^Y \to 2^X$ with respect to costs c satisfying Assumptions (A0)–(A6). Our goal is to describe the basic differential topology of the cells

$$Z'(x) := Z(\partial^c\psi^c(x))$$

of given c-concave potentials $\psi : Y \to \mathbb{R}\cup\{-\infty\}$ having c-transform $\psi^c : X \to \mathbb{R}\cup\{+\infty\}$. Our main results are Propositions 4.2.3 and 4.3.4. The proof of 4.2.3 is essentially a local adaptation of the recent work of [KM18], and 4.3.4 is obtained as corollary to theorem of Alberti [Alb94]. We are grateful to Prof. R. J. McCann for many useful conversations.

Let the reader recall the definitions of c-concavity (Section 2.3) and c-subdifferentials (Definition 2.3.3). The c-concavity $\psi^{cc} = \psi$ represents a pointwise inequality on Y, namely

$$-\psi^c(x) + \psi(y) \leq c(x, y)$$

for all $(x, y) \in X \times Y$. The case of equality $\psi(y) - \psi^c(x) = c(x, y)$ is most important, and occurs if and only if $y \in \partial^c\psi^c(x)$, and if and only if $x \in \partial^c\psi(y)$. We recall that c-optimal semicouplings π are supported on the graphs of c-subdifferentials, and hence $-\psi^c(x) + \psi(y) = c(x, y)$ π-a.e.

Now we make some observations. Abbreviate $c_\Delta(x; y, y') := c(x, y) - c(x, y')$ for the two-pointed cross difference. If ψ^c is differentiable at $x \in X$, then $\partial^c\psi^c(x)$ is a singleton, say $\{y_0\}$. Then π-a.e. we have equality $\psi^c(x) + c(x, y) = \psi(y)$. On the other hand, if x is such that $\partial^c\psi^c(x)$ is not a singleton, say including two distinct points y_0, y_1, then $\psi^c(x) + c(x, y_0) = \psi(y_0)$ and $\psi^c(x) + c(x, y_1) = \psi(y_1)$, which implies

$$0 = \psi(y_1) - \psi(y_0) - c_\Delta(x, y_0, y_1).$$

Likewise $x \in \partial^c\psi(y)$ if and only if

$$y \in argmax[\{\psi(y_*) - c(x, y_*) + c(x, y) \mid y_* \in Y\}]. \tag{4.1}$$

For $x_0 \in X$, abbreviate $Z'(x_0) := Z(\partial^c\psi^c(x_0))$. We interpret $Z'(x_0)$ as the "local cell" containing x_0 in the activated domain A of the c-optimal semicoupling. If $x_0 \in X$ is a "singular point", where $\psi^c(x_0)$ is not differentiable, then $\partial^c\psi^c(x_0)$ is not a singleton and the cell $Z'(x_0)$ is described by the system of equations

$$Z'(x_0) = \{x \in X \mid 0 = \psi(y_0) - \psi(y_1) - c_\Delta(x; y_0, y_1),\ y_1 \in \partial^c\psi^c(x_0), y_1 \neq y_0\} \tag{4.2}$$

according to (4.1). From the expression (4.2) we obtain the following:

Lemma 4.2.1. *Under Assumptions (A0)–(A3), the cell $Z'(x) = Z(\partial^c\psi^c(x))$ is a closed locally DC-subvariety in X for every $x \in dom(\psi^c)$.*

Proof. The cell $Z'(x)$ is the intersection of sets of the form $\partial^c\psi(y)$, which are closed by Lemma 2.3.4. Assumption (A1) implies $x \mapsto \nabla^2_{xx}c(x, y)$ is locally bounded on X, uniformly in y. Therefore every x admits a neighborhood U and a constant $C \geq 0$ such that $||\nabla^2_{xx}c(x, y)|| \leq C.Id$ uniformly with respect to y throughout U. This implies $x \mapsto c(x, y)$ is locally semiconvex function on X, uniformly in y. But then the cross-differences $x \mapsto c_\Delta(x; y, y')$ are locally DC-functions, uniformly in y, y' in Y. This observation and equation (4.2) implies $Z'(x)$ is locally-DC subvariety of X. □

We caution the reader that Lemma 4.2.1 does not the specify the exact-dimension of $Z'(x)$. This question of dimension is addressed below, c.f. Proposition 4.2.3. Again $Z'(x)$ is viewed as a local "cell" on the active domain $x \in Z_1 = A$. Our next step is to describe the space of directions $T_xZ'(x)$ (compare [Vil09, Def.10.4.6, pp.257]). The definition of the vector field $\eta_{avg}(x) \in T_xZ'$ involves the orthogonal projection $pr_{Z'} : T_xX \to T_xZ'$ of T_xX onto the space of directions.

Lemma 4.2.2. *Let $x \in X$ be a singular point, where $\partial^c\psi^c(x)$ is not a singleton. Then* *the space of directions $T_x Z'(x)$ is a subset of the orthogonal complement*

$$orthog[\{\nabla_x c_\Delta(x; y_0, y_1) \mid y_1, y_0 \in \partial^c \psi^c(x)\}]$$

in $T_x X$.

Proof. Abbreviate $A(x,y) := A(x,y,y_0) = c_\Delta(x,y,y_0) - (\psi(y) - \psi(y_0))$. We examine the case of equality $A = 0$ of the inequality $A \leq 0$. A first-order deformation η at x and tangent to $Z'(x)$ must preserve the system of equations $\{A(x,y,y') = 0 | \ y, y' \in \partial^c \psi^c(x)\}$. But this only if $\eta \in T_x X$ satisfies the homogeneous linear equations $\eta \cdot \nabla_x c_\Delta(x,y,y') = 0$ for every $y, y' \in \partial^c \psi^c(x)$. □

The Lemma 4.2.2 indicates the *expected* Hausdorff dimension of $Z'(x)$, namely the dimension of the orthogonal complement $\{\nabla_x c_\Delta(x,y,y') \mid y, y' \in \partial^c \psi^c(x)\}$. With Halfspace conditions and ideas from [KM18], we confirm this dimension estimate in Proposition 4.2.3 below.

Some preliminary notation is convenient. Recall X is a Riemannian manifold-with-corners, equipped with Riemannian exponential mapping $exp_x : T_x X \to X$. If (X,d) is complete Cartan-Hadamard space, then exp_x is diffeomorphism between $T_x X$ and the universal covering space $\tilde{X}$. On general Riemannian manifold-with-corners X the exponential map is a local diffeomorphism between sufficiently small open neighborhoods U of $x_0 \in X$ with open balls U' in the tangent space $T_{x_0} X$. So we have C^1 diffeomorphisms between local neighborhoods of $x_0 \in X$ with neighborhoods of 0 in euclidean space $\mathbb{R}^n$, with $n = \dim(X) = \dim(T_{x_0} X)$. Thus for every $x_0 \in X$ there exists a local diffeomorphism splitting $B_\epsilon(x_0)$ as a product $B_\epsilon(x_0') \times B_\epsilon(x_0'')$, viewed as subset of $\mathbb{R}^{n-j} \times R^j$ with $x_0 = (x_0', x_0'')$ for $j \geq 0$.

Recall c satisfies Assumptions (A0)–(A4). Let $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ be a c-concave potential. Let $Z_1 := A$ be the active domain of the semicoupling defined by ψ, and let $x \in Z_1$.

Proposition 4.2.3. *Suppose the gradients $\{\nabla_x c_\Delta(x'; y_0, y_1) \mid y_0, y_1 \in \partial^c \psi^c(x)\}$ satisfy Halfspace conditions and are bounded away from zero in $T_{x'} X$, uniformly with respect to $x' \in Z'(x) = Z(\partial^c \psi^c(x))$. Then $Z'(x)$ is a codimension-j local DC-subvariety of the active domain Z_1, where*

$$j := \dim[span\{\nabla_x c_\Delta(x; y_0, y_1) \mid y_0, y_1 \in \partial^c \psi^c(x)\}].$$

When the gradients $\{\nabla_x c_\Delta(x'; y_0, y_1) \mid y_0, y_1 \in \partial^c \psi^c(x)\}$ satisfy Halfspace Conditions and Definition 3.2.2, then

$$codim_{Z_1} Z_j = j - 1 \text{ for every } j \geq 1. \tag{4.3}$$

Observe that (HS) Conditions fail whenever Z_j is empty, and evidently the formula (4.3) no longer applies.

Proof. Abbreviate $A(x, y) = c(x, y) - c(x, y_0)$. Consider the mapping $G : B_\epsilon(x_0) \to \mathbb{R}^d$ defined by

$$G(x) = (A(x, y_1), A(x, y_2), \dots A(x, y_j))$$

for $y_0, y_1, \dots, y_j \in \partial^c \psi^c(x_0)$. Assume

$$\{\nabla_x A(x_0, y_i, y_0)\}_{1 \leq i \leq j}$$

is a linearly independant subset of $T_{x_0} X$ which satisfies Halfspace conditions. The map $G : B_\epsilon(x_0) \to \mathbb{R}^j$ is local DC-function according to Lemma 4.2.1.

We use exponential mapping to obtain local C^1-diffeomorphism between an open neighborhood of x_0 in X. Next we apply the DC-implicit function theorem as stated in [KM18, Thm 3.8], and conclude there exists $\epsilon > 0$ and a biLipschitz DC-mapping ϕ from $B_\epsilon(x_0') \subset \mathbb{R}^{n-j}$ to $B_\epsilon(x_0'') \subset \mathbb{R}^j$ such that, for all $x = (x', x'') \in B_\epsilon(x_0') \times B_\epsilon(x_0'') \subset \mathbb{R}^{n-j} \times \mathbb{R}^j$ we have $G(x) = G(x', x'') = 0$ if and only if $x'' = \phi(x')$.

The basic subdifferential inequalities (4.2) imply $G(x) = 0$ if and only if $x \in Z'(x_0) \cap B(x_0, \epsilon)$. Now because $Z'(x)$ can be covered by countably many sufficiently small open balls, we conclude that $Z'(x)$ is a local DC-subvariety with Hausdorff dimension $\dim_{\mathscr{H}} Z'(x) = n - j$. ☐

N.B. The description of $T_x Z'(x)$ is symmetric with respect to y_1, y_0. There is further symmetry from the additive relations between cross-costs

$$c_\Delta(x; y_0, y_1) + c_\Delta(x; y_1, y_2) = c_\Delta(x; y_0, y_2).$$

This implies the obvious estimate $codim Z'(x) \leq \#(\partial^c \psi^c(x)) - 1$ under general conditions.

Recently [KM18] obtained an explicit parameterization of singularities arising from Euclidean quadratic costs, employing a hypothesis of affine independance between the connected components of the subdifferentials $\partial^c \psi^c(x)$. Their parameterization requires a global splitting $X = X_0 \times X_1$ of the source domain to express singularities ("tears" in their terminology) as the graphs of DC-functions $G : X_0 \to X_1$ as above. However further hypotheses on the convexity of source and target are required in their arguments.

4.3 The Descending Filtration $Z_j, j = 0, 1, 2, \ldots$

The Kantorovich functor $Z : 2^Y \to 2^X$ leads to a filtration of the source X. We follow an idea of Prof. D. Bar-Natan [Bar02a] and "skewer the cube" 2^X, or more accurately the support $spt(Z)$, according to local codimensions.

Definition 4.3.1. For integers $j = 0, 1, 2, \ldots$, let

$$Z_{j+1} := \{x \in X| \quad \dim[span\{\nabla_x c_\Delta(x; y, y')|y, y' \in \partial^c \psi^c(x)\}] \geq j\}$$

where $Z'(x) = Z(\partial^c \psi^c(x))$ for a c-concave potential $\psi^{cc} = \psi$ on target space Y.

According to the definition, Z_j is supported on the subcategory $spt(Z)$ of 2^Y for every $j \geq 1$. In this notation the activated support A coincides with the support of Z_1. From the definition we find Z_{j+1} is a union of cells $Z(Y_I)$ for which the local dimension is at least j. The cells $Z(Y_I)$ are closed (2.3.4). If the support $spt(Z)$ of the functor is finite, then the unions Z_{j+1} are closed. However if $spt(Z)$ is infinite, then the union Z_2, Z_3, ... is possibly not closed and depends on the cost c. The Euclidean quadratic cost $c = d^2/2$, for instance, often has Z_2 not closed when the target Y has $\dim(Y) > 0$. This motivates our study of "repulsion" costs in Chapter 5, where we expect Z_2 is always closed in practice.

Now we replace the c-subdifferentials with an important "localized" version using (Twist) condition (A4). The c-subdifferential $\partial^c \psi^c$ of a c-convex function $\psi^c : X \to \mathbb{R} \cup \{+\infty\}$ is non-local subset of Y. The subsets $\partial^c \psi^c \subset Y$ depends on global datum, namely the values of $c(x, y)$ for all $y \in dom(c_x)$. Hence $\partial^c \psi^c(x_0)$ depends on the values of $\psi(y)$ and $\psi^c(x)$ for every $y \in Y, x \in X$ and not simply the local behaviour of ψ^c near x_0. It is useful to introduce a local subdifferential, namely the so-called "subgradients" of a function $\psi^c : X \to \mathbb{R} \cup \{+\infty\}$.

Definition 4.3.2 (Local Subgradients $\partial_\bullet \phi$)**.** Let U be open set in X, sufficiently small such U is C^1-diffeomorphic to an open subset of $\mathbb{R}^n$ where $n = \dim(X)$. Let $\phi : U \to \mathbb{R}$ be a function. Then ϕ is subdifferentiable at x with subgradient $v^* \in T_x^* X$ if

$$\phi(z) \geq \phi(x) + v^*(z - x) + o(|z - x|) \text{ for all } z \text{ near } x.$$

Let $\partial_\bullet \phi(x) \subset T_x^* X$ denote the set of all subgradients to ϕ at x. Here o is the "little-oh" notation.

Evidently the subgradient $\partial_\bullet \phi(x)$ is local, and depending on the values of ϕ near x. Moreover $\partial_\bullet \phi(x)$ is a closed convex subset of $T_x^* X$ for every $x \in dom(\phi)$.

The Assumptions (A0)–(A3) imply $\psi^c : X \to \mathbb{R} \cup \{+\infty\}$ is locally semiconvex over its domain $dom(\psi^c) \subset X$ (Lemma 2.5.5). In otherwords, every $x \in dom(\psi^c)$ admits an open neighborhood U of x and a constant $C \geq 0$ such that $D^2\psi^c \geq -CId$ throughout U, where $D^2\psi^c = D^2_{xx}\psi^c$ is the distributional Hessian in the source variable x ([Vil09, Theorem 14.1, pp.363]). So given a c-convex potential ψ^c on X, for every $x_0 \in X$ there exists open neighborhood U and $C \geq 0$ such that $\psi^c|_{U_x} + C||x||^2/2$ is strictly-convex throughout U. This implies the local subdifferential and subgradients $\partial_\bullet \psi^c(x)$ coincide with the convex-analytic subdifferential of the local function $\psi^c|_U + C||x||^2/2$, when restricted to the sufficiently small neighborhood x in X.

Having introduced the local subdifferential, there is important comparison between $\partial^c\psi^c(x_0) \subset Y$ and $\partial_\bullet\psi^c(x_0)$, assuming (Twist) condition Assumption (A4). This relation is the inclusion

$$\partial^c\psi^c(x_0) \subset \{y \in Y | -\nabla_x c(x,y) \in \partial_\bullet \psi^c(x_0)\}. \tag{4.4}$$

Observe $\emptyset \neq \partial^c\psi^c(x_0)$ whenever $x_0 \in dom(\psi^c)$. We abuse notation and denote $\nabla_x c(x,y)$ for the canonical covector in $T^*_x X$, with the tangent vector in $T_x X$ using the ambient Riemannian structure. The inclusion (4.4) allows us to replace the global c-subdifferential $\partial^c\psi^c$ with the local convex set of subgradients $\partial_\bullet\psi^c$. The inclusion is generally strict. However it produces basic upper bounds on the Hausdorff dimension of $\partial^c\psi^c$. We quote the following theorem of G. Alberti:

Theorem 4.3.3 ([Alb94])**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be proper lowersemicontinuous convex function. For every $0 < k < n$, let $S^k(f) \subset \mathbb{R}^n$ be the subset defined by*

$$S^k(f) := \{x \in \mathbb{R}^n | \dim_{\mathscr{H}}(\partial_\bullet(f)) \geq k\}.$$

Then $S^k(f)$ can be covered by countably many $(n-k)$-dimensional DC-manifolds.

In otherwords $S^k(f)$ is countably $(n-k)$-rectifiable and has Hausdorff dimension $\leq (n-k)$. If f is $+\infty$-valued on $\mathbb{R}^n$, then Alberti's method shows $S^k(f)$ can be covered by countably many $(n'-k)$-dimensional manifolds where $n' = \dim_{\mathscr{H}}(dom(f))$. The domain $dom(f)$ is a closed convex subset of $\mathbb{R}^n$ having well-defined Hausdorff dimension.

Proposition 4.3.4. *[Dimension Estimate] Let $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be cost satisfying Assumptions (A0)–(A3). Let $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ be c-concave potential $\psi^{cc} = \psi$ with singularity functor $Z : 2^Y \to 2^X$. Define $n' := \dim_{\mathscr{H}}(dom(\psi^c))$. Then for every integer $j \geq 1$, the subvariety Z_j has Hausdorff dimension*

$$\dim_{\mathscr{H}}(Z_j) \leq n' - j + 1.$$

Proof. Consider the inclusion (4.4). For given x_0, let U be the open neighborhood and $C > 0$ such that $\psi^c|_U + C||x||^2/2$ is strictly-convex throughout U. For $x \in U$ the c-subdifferentials $\partial^c\psi^c(x)$ are contained in the closed convex local subdifferentials $\partial_\bullet\psi^c|U(x)$. Now we have equality

$$\partial_\bullet(\psi^c|_U(x) + C|x|^2/2) = \partial_\bullet\psi^c|_U(x) + C\langle -, x\rangle. \tag{4.5}$$

So the local subdifferential $\partial_\bullet(\psi^c_U + C|x|^2/2)$ in T^*_xX is an affine translate of $\partial_\bullet\psi^c|_U$ by the linear functional $C\langle -, x\rangle$.

Next we apply Alberti's theorem to the localized convex function $\psi^c|_U + C|x|^2/2$. Thus $S^k(\psi^c|_U + C|x|^2/2)$ and $S^k(\psi^c|_U)$ can be covered by countably many $(n'-k)$-dimensional manifolds where $n' := \dim_{\mathscr{H}}(dom\psi^c|_U)$.

Finally we relate $\dim_{\mathscr{H}}(\partial_\bullet\psi^c|_U(x))$ to the Hausdorff dimension of $Z'(x) = Z(\partial^c\psi^c(x))$ $= \cap_{y\in\partial^c\psi^c(x)}\partial^c\psi(y)$. From the definition of subgradients, we have

$$conv(\{\nabla_x c(x,y) | y \in \partial^c\psi^c(x)\}) \subset \partial_\bullet\psi^c(x).$$

Moreover the closed convex hull $conv(\{\nabla_x c(x,y)|y \in \partial^c\psi^c(x)\})$ has dimension

$$j = \dim(span\{\nabla_x c_\Delta(x, y, y_0) \mid y \in \partial^c\psi^c(x)\}).$$

For every $x \in Z_j$ we conclude $Z_j \cap U \subset S^j(\psi^c|_U + C|x|^2/2)$, which according to Alberti's theorem 4.3.3 yields the upper bound $\dim_{\mathscr{H}}(Z_j \cap U) \leq n' - j$. To conclude, we observe that Z_j can be covered by countably many open neighborhoods U of points $x \in Z_j$.

□

We supplement 4.3.4 with the following application of Clarke's Implicit Function Theorem, which gives a criterion for the singularities $Z_2, Z_3, \ldots$ to be *closed* subsets of X.

Proposition 4.3.5. *Let $c : X \times Y \to \mathbb{R}$ be a cost satisfying Assumptions (A0)–(A4). For $x \in X$, let $K_x := \{conv(\nabla_x c_\Delta(x, y_0, y_1) \mid y_0, y_1 \in \partial^c\psi^c(x)\} \subset T_xX$ be the closed convex hull. If $j \geq 2$ is an integer such that the convex sets K_x are disjoint from the origin $\boldsymbol{0}$ in $T_xZ'(x)$, uniformly with respect to $x \in Z_j$, then Z_j is a closed subset of X.*

Proof. Consequence of Clarke's Implicit Function Theorem, c.f. [Vil09, Proof of Theorem 10.50, pp.262–264]. □

We conclude this section with some remarks on the literature. The term "singularity" is evidently overburdened, and having various interpretations within the literature. Our thesis imagines "singularity" as referring to a locus-of-discontinuity of certain extremal measures (the c-optimal semicouplings). The key features of our formulation of Kantorovich singularity are the following: Definition 4.1.1 is *categorical* and Definition 4.1.1 is *topological*. Indeed as we described in our introduction, the idea of "singularity" as locus-of-discontinuity is not a topological definition.

Our topological Kantorovich singularity also yields an alternative perspective on the so-called "regularity theory" of optimal transportation. Typically regularity in optimal transport focuses on the mapping T defined in (2.11) in Chapter 2. There is large volume of research concerning the $C^{1,\alpha}$ or C^2, C^∞ regularity of T under various hypotheses on c, σ, τ. We refer the reader to [Vil09, Ch.12] for a survey. Our thesis however is interested in the continuity, and especially the discontinuities of T. Indeed we describe the topology of points x where c-convex potentials are not uniquely differentiable, where $\partial^c\psi^c(x) \subset Y$ is not a singleton. This thesis passes silently over questions of the type "How continuous is the map T away from the singularities?".

Several results concerning singularities of optimal transports have been attained in the literature. In [Fig10a] the singularities of optimal transports between two probability measures supported on bounded open domains in the plane $\mathbb{R}^2$ with respect to the quadratic euclidean cost $c(x, y) = ||x-y||^2/2$ (equivalently $c = -\langle x, y\rangle$), was studied. The main result of [Fig10a, §3.2] is that the singularity (Z_2 in our notation) has topological closure $\overline{Z_2}$ in $\mathbb{R}^2$ with zero two-dimensional Hausdorff measure, $\mathscr{H}^2(\overline{Z_2}) = 0$. The work of Figalli was extended in [FK10], where the determination $\mathscr{H}^n(\overline{Z_2}) = 0$ was established for singularities of optimal couplings under the hypothesis that probability measures are supported on bounded open domains of $\mathbb{R}^n$ with respect to euclidean quadratic cost. In [PF] a similar result is established with respect to costs on $\mathbb{R}^n$ satisfying more general nondegeneracy conditions, namely:

(i) $c \in C^2(X \times Y)$ with $||c||_{C^2} < +\infty$,

(ii) $x \mapsto \nabla_y c(x, y)$ injective map for every y,

(iii) $y \mapsto \nabla_x c(x, y)$ injective map for every x, and

(iv) $det(D^2_{xy}c) \neq 0$ for all (x, y).

These previous works suggest that under general assumptions on the cost, the singularity Z_2 has Hausdorff dimension $\dim_{\mathscr{H}}(Z_2) \leq \dim(X) - \dim(Y)$. The recent work [KM18] confirms this estimate under particular conditions, namely euclidean quadratic cost, convex source, and convexity and affine-independance of the disjoint target components.

4.4 Local–Global Homotopy Reductions: (UHS) Conditions

This section generalizes the homotopy-reduction constructed in 3.1.1, and establishes a retraction procedure

$$Z_1 \rightsquigarrow Z_2 \rightsquigarrow \cdots \rightsquigarrow Z_{J+1},$$

defined up to some maximal index $J \geq 1$ for which the inclusions

$$Z_1 \hookleftarrow Z_2 \hookleftarrow \cdots Z_{J+1}$$

are simultaneously homotopy isomorphisms. These retractions require our cost c satisfy Assumptions (A0)–(A6) and certain further hypotheses, namely "Uniform Halfspace Conditions" (Definition 4.4.2) and a further separation property (Definition 4.4.1). Our retraction will integrate a vector field, denoted $\eta_{avg}(x)$, whose flow under the above Assumptions yields continuous deformation retracts (Theorem 4.4.4). We view our retractions as interpretations in the category of semicouplings and Kantorovich duality to the "well-rounded retracts" of [Ash84], [Sou78]. Discussions on the "geometry-of-numbers" applications of the well-rounded retracts are found in [PS08], [Ste07, §A.6.4].

There is an important hypothesis necessary for the continuity of our retracts.

Definition 4.4.1 (Separated c-subdifferentials)**.** Let $c : X \times Y \to \mathbb{R}$ be a cost satisfying Assumptions (A0)–(A4). Let $\phi = \psi^c$ be a c-convex function, $\phi : X \to \mathbb{R} \cup \{+\infty\}$. Let $d_Y : Y \times Y \to \mathbb{R}$ a distance function. Then ϕ is δ-*separated* if $\partial^c\phi(x)$ is a δ-separated discrete subset of Y, i.e. $d_Y(y_0, y_1) \geq \delta$ for every $y_0 \neq y_1$.

Separation is a property of the singular points in $x \in Z_2 \subset dom(\phi)$, and is trivially satisfied when Z_2 is empty.

If the target Y is discrete, then every c-convex potential ϕ is separated, with say $\delta = \frac{1}{3}\max_{y,y' \in Y} d(y, y')$.

If Y is connected, $\dim(Y) > 0$, then the hypotheses of 4.4.1 are nontrivial. For example, when $c(x, y) = ||x - y||^2/2$ is the quadratic Euclidean cost, then c-convex potentials ϕ are often not separated when Z_2 is nonempty. This is related to the fact that medial axes $M(A)$ of bounded open subsets $A \subset X$ are not closed when A has, say, $C^{1,\alpha}$-regular boundary. This fact motivates our study of "anti-quadratic" repulsion costs in Chapter 5, where we expect potentials to be generally separated.

The retraction procedure requires some notation. Let c be a cost satisfying Assumptions (A0)–(A6). Abbreviate $Y'(x) := dom(c_x)$. Let the source σ and target measure τ

be absolutely-continuous with respect to the Hausdorff measures $\mathscr{H}_X$ and $\mathscr{H}_Y$ on X, Y, respectively. Let $\psi^{cc} = \psi$ be a c-concave potential maximizing Kantorovich's dual problem (§2.3) relative to a source σ and target τ. Let $Z = Z(c, \sigma, \tau)$ be the corresponding Kantorovich functor. For $x \in Z'$, let $pr_{Z'} : T_xX \to T_xZ'$ denote the orthogonal-projection mapping, where T_xZ' is the space of directions of Z' at x.

Definition 4.4.2. In the above notation, let ψ be a separated c-concave potential for some $\delta > 0$ (Definition 4.4.1). Let $x_0 \in dom(\psi^c)$, and abbreviate $Z' := Z(\partial^c\psi^c(x_0))$. Select any $y_0 \in \partial^c\psi^c(x_0)$.

For parameter $\beta > 2$, $x \in Z'$, $y \in Y$, define tangent vectors $\eta(x, y) \in T_xX$ by the equation

$$\eta(x, y) := |\psi(y_0) - \psi(y) + c_\Delta(x; y, y_0)|^{-\beta} \cdot pr_{Z'}(\nabla_x c_\Delta(x; y, y_0)). \tag{4.6}$$

Relative to x_0, τ, let $\bar{\nu}$ be the Radon measure on Y defined by

$$d\bar{\nu}(y) := (1 - e^{d(y, \partial^c\psi^c(x_0))^2/\delta}).1_{Y'(x_0)}.d\tau(y).$$

Then Uniform Halfspace (UHS) conditions are satisfied at x in Z' with respect to the parameter β if:

(UHS1) the Bochner integral $\eta_{avg}(x)$ defined by

$$\eta_{avg}(x) := (\bar{\nu}[Y])^{-1} \int_Y \eta(x, y).d\bar{\nu}(y) \tag{4.7}$$

is nonzero finite vector in $T_xZ' - \{\mathbf{0}\}$; and

(UHS2) there exists a constant $C > 0$, uniform with $x \in Z'$, for which the estimate

$$||\eta_{avg}(x)|| \geq C \int_Y ||\eta(x, y)|| d\bar{\nu}(y) > 0 \tag{4.8}$$

holds.

We make some remarks. First, we observe (UHS2) basically implies (UHS1). Second, the definition of $\eta(x, y)$ is independant of the choice of $y_0 \in \partial^c\psi^c(x')$. Third, in practice the parameter $\beta > 2$ is taken sufficiently large to ensure the divergence of $\eta_{avg}(x)$ whenever x converges into $\partial Z'$. If c satisfies Assumption (A5), then $\psi(y)$ is locally Lipschitz (Prop.2.5.3) and the exponent $\beta = \dim(Y) + 2$ is sufficient. Compare §3, Case (II) in the proof of Theorem 3.1.1.

Theorem 4.4.3. *Let $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be cost satisfying Assumptions (A0)–(A6). Suppose the source and target measures are absolutely continuous with respect to volume measures on X, Y. Let $Z : 2^Y \to 2^X$ be the contravariant singularity functor. Let $x' \in Z_1$ be a point supported on the activated domain of an optimal semicoupling π_{opt}, and abbreviate $Z' := Z(\partial^c \psi^c(x')) = \cap_{y_0 \in \partial^c \psi^c(x')} \partial^c \psi(y_0)$.*

If (UHS) Conditions are satisfied for all points $x \in Z'(x') \cap Z_j$ (Definition 4.4.2), and $Z'(x') \cap Z_{j+1} \neq \emptyset$ is nonempty, and ψ is separated (Definition 4.4.1),

then there exists a continuous map

$$\Psi : (Z' \cap Z_j) \times [0,1] \to Z' \cap Z_j$$

such that:

(i) $\Psi(x, s) = x$ for all $x \in Z' \cap Z_{j+1}$; and

(ii) $\Psi(x, 0) = x$ for all $x \in Z'$; and

(iii) $\Psi(x'', 1) \in Z' \cap Z_{j+1}$ for all $x'' \in Z' \cap Z_j$.

In addition, if the above hypotheses are satisfied, then $Z' \cap Z_{j+1}$ is strong deformation retract of $Z' \cap Z_j$ and the inclusion

$$Z' \cap Z_{j+1} \hookrightarrow Z' \cap Z_j$$

is a homotopy-isomorphism.

The vector field $\eta_{avg}(x)$ defined in Theorem 4.4.3 is tangent to the cell $Z'(x)$, and therefore the flow is constrained to the cells $Z'(x)$. For x varying over Z_j, the vector field $\eta_{avg}(x)$ varies continuously with respect to x according to Lemma 2.3.4. The hypothesis $Z' \cap Z_{j+1} \neq \emptyset$ cannot be ignored, and definitely necessary for continuity of the gradient flow of η_{avg}.

Next we claim the mappings

$$\{\Psi : Z'(x) \cap Z_j \to Z'(x) \cap Z_j\}$$

assemble to a continuous retraction $Z_j \times [0,1] \to Z_j$, for $x \in Z_j$, which establishes that $Z_j \hookleftarrow Z_{j+1}$ is a homotopy-isomorphism.

Theorem 4.4.4. *Let $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be cost satisfying Assumptions (A0)–(A6). Let $Z : 2^Y \to 2^X$ be the contravariant singularity functor defined by source and target measures absolutely continuous with respect to volume measures on X, Y.*

If $j \geq 1$ is an integer such that (UHS) Conditions are satisfied for all points $x \in Z_j$, and $Z'(x) \cap Z_{j+1} \neq \emptyset$ for every $x \in Z_j - Z_{j+1}$, and ψ is separated (Definition 4.4.1), *then the local homotopy-equivalences*

$$\{\Psi : (Z' \cap Z_j) \times [0,1] \to Z' \cap Z_{j+1}\}$$

constructed in Theorem 4.4.3 assemble to a continuous deformation retract $Z_j \times [0,1] \to Z_{j+1}$.

Furthermore if $J \geq 1$ *is that maximal integer where (UHS) conditions are satisfied throughout* Z_J*, then composing the retractions* $\{\Psi\}$ *produces a codimension-*J *homotopy-isomorphism* $Z_1 \simeq Z_{J+1}$.

The proof of Theorems 4.4.3–4.4.4 is analogous with the "base case" retraction of $X(= Z(\emptyset))$ onto the activated domain $A = Z_1$, see Theorem 3.1.1. The formal proof of 3.1.1 required Lemmas 3.2.5, 3.2.6. Likewise the formal proofs of 4.4.3–4.4.4 require analogous lemmas, and summarized in:

Lemma 4.4.5. *Let the hypotheses of 4.4.3–4.4.4 be satisfied. Let* $j \geq 1$ *be such that (UHS) conditions are satisfied for all* $x \in Z_j$*, and* $Z'(x) \cap Z_{j+1} \neq \emptyset$ *for every* $x \in Z_j - Z_{j+1}$*, and* ψ^c *is separated.*

For $x_0 \in Z_j - Z_{j+1}$*, consider the initial value problem*

$$x' = \eta_{avg}(x), \quad x(0) = x_0. \tag{4.9}$$

(a) For every initial value $x_0 \in Z_j - Z_{j+1}$*, the flow* Ψ *defined by* (4.9) *diverges to infinity in finite time.*

(b) The maximum interval of existence $I(x_0) := [0, \omega(x_0))$ *of solutions to* (4.9) *varies continuously with respect to the initial value* $x_0 \in Z_j - Z_{j+1}$.

Proof. We follow the arguments of 3.2.5. First (UHS) implies the flow (4.9) is well-defined and extendible throughout the interior of the cells $Z'(x_0) \subset Z_j$. At the ω-limit point

$$\bar{x} := \lim_{t \to \omega(x_0)^-} x(t)$$

we claim $\bar{x} \in Z'(x_0) \cap Z_{j+1}$. The key point is to verify $\bar{x} \in Z_{j+1}$. At the ω-limit point, we find

$$\partial^c \phi(\bar{x}) \supset \{\bar{y}\} \coprod \partial^c \phi(x_0),$$

and the subdifferential at $\bar{x}$ contains at least one new target point $\bar{y} \in Y$. The (UHS) condition implies the average of the projections $\nabla_x c_\Delta(\bar{x}, y_0, \bar{y})$ (with respect to $\bar{\nu}$) is a nonzero vector of $T_{\bar{x}} Z'$, and linearly independant from the nonzero tangent vectors

$\nabla_x c_\Delta(\bar{x}, y_1, y_0)$, $y_0, y_1 \in \partial^c\phi(x_0)$ (which are orthogonal to $T_{\bar{x}}Z'$). This proves $\bar{x} \in Z_{j+1}$. This proves (a).

Next we claim the flow defined by Ψ is asymptotically contracting, and therefore ω is a Lipschitz continuous function. The same arguments from 3.2.3, 3.2.4 shows the vector field η_{avg} is the gradient of an averaged potential f_{avg}, and this potential is asymptotically convex in the direction of $\nabla_x f_{avg}$. This proves (b).

□

Sketch of proof for Theorem 4.4.3. We follow the proof of Theorem 3.1.1, and construct a continuous vector field $\eta_{avg}(x')$ on $Z' \cap Z_j$ which blows-up precisely on $Z' \cap Z_{j+1} \subset Z' \cap Z_j$, which we assume is nonempty. For initial points x', the maximal interval of existence $I(x') = [0, \omega(x'))$ varies continuously with respect to x' (see Lemma 4.4.5). The field $\eta_{avg}(x')$ will generate a global forward-time continuous mapping

$$\Psi : (Z' \cap Z_j) \times I \to Z' \cap Z_j$$

satisfying the usual ordinary differential equation $\frac{d}{ds}[\Psi(x', s')] = \eta_{avg}(\Psi(x', s'))$ for all $s' \in I(x')$. The flow Ψ is reparameterized according to the parameter $\omega(x_0)$ to obtain a continuous deformation retract $\Psi' : (Z' \cap Z_j) \times [0, 1] \to Z' \cap Z_j$ as desired. □

Proof of Theorem 4.4.3. The (UHS) conditions ensure the cross-differences c_Δ have non-vanishing gradient $\nabla_x c_\Delta \neq 0$ throughout the domain Z', and the gradients $\nabla_x c_\Delta$ vary continuously over Z_j. Moreover the uniform Halfspace condition (UHS2) ensures $||\eta_{avg}(x)||$ is uniformly bounded away from zero in neighborhoods of the poles $\{||\eta_{avg}(x)|| = +\infty\} = Z' \cap Z_{j+1}$ in $Z' \cap Z_j$.

The vector field $\eta_{avg}(x')$ integrates to a global forward-time flow $\Psi : Z' \cap Z_j \times I \to Z' \cap Z_j$, where as usual we have $d/ds[\Psi(x', s')] = \eta_{avg}(\Psi(x', s'))$ for all $s' \in I(x')$. According to Lemma 4.4.5, for every choice of $x = x(0)$ initial value on $Z'(x) \cap Z_j$, the trajectory $x = x(t)$ converges in finite time to $Z' \cap Z_{j+1}$ with respect to the flow (3.10). Our Assumptions (A0)–(A6) imply continuous dependance on the choice of initial value. The argument from Proposition 3.2.4 with Lemma 4.4.5 prove the flow Ψ can be reparameterized according to the parameter $\omega(x_0)$ to obtain a continuous deformation retract Ψ' of $Z' \cap Z_j$ onto $Z' \cap Z_j$, as desired.

□

Finally we observe the "local" homotopy-isomorphisms $\{\Psi : Z'(x) \cap Z_j \rightsquigarrow Z'(x) \cap Z_{j+1}\}$ assemble to a continuous deformation retract of $Z_j \simeq Z_{j+1}$ and establish Theorem 4.4.4:

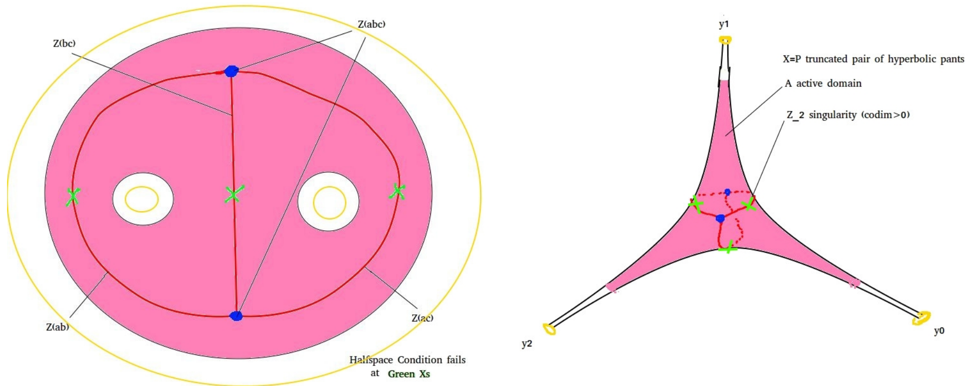

Figure 4.3: Horospherically truncated pair of pants, with active domain relative to a repulsion cost (See Definition 5.3.2 from Section 5.3). Halfspace Conditions fail. $P \rightsquigarrow Z_1 \rightsquigarrow Z_2$ are homotopy equivalent, where P is pair of pants. But Z_2 does not deformation retract to Z_3

Proof of Theorem 4.4.4. The proof of Theorem 4.4.3 constructs homotopy deformations $Z'(x) \cap Z_j \to Z'(x) \cap Z_{j+1}$, where $Z'(x) = Z(\partial\psi^c(x)) \subset Z_1$ for a given x in Z_j. When x varies over Z_j, the corresponding $Z' = Z'(x)$ are either coincident or intersect along a subset of Z_{j+1}. So the local strong deformation retracts $\{\Psi\}$ constructed in Theorem 4.4.4 assemble to a continuous deformation retract $Z_j \times [0,1] \to Z_j$. Therefore $Z_j \hookleftarrow Z_{j+1}$ is a homotopy-isomorphism. Composing the retractions $Z_j \rightsquigarrow Z_{j+1}$ for $j = 1, \ldots, J$ yields the deformation retract $Z_1 \rightsquigarrow Z_{J+1}$, as desired. $\square$

Chapter 5

Repulsion Costs

The previous Chapters have assumed c is a general cost satisfying the Assumptions (A0), …, (A6), etc. Practical applications require explicit costs which satisfy all these conditions simultaneously. This chapter introduces various costs $c_1, c_2, \ldots$, with their own peculiar features which, at least partially satisfy these conditions. We interpret these costs as "sign posts", leading us to the visibility cost v on convex excisions as candidate to satisfy all hypotheses required in Theorems 4.4.3 and 4.4.4.

5.1 Chain sums and Well-Separated Gates

We begin with general definitions, and then specialize to a more symmetric setting involving an isometric group action $X \times \Gamma \to X$. This will serve our applications in Chapters 6, 7. The figure 1.5 from the Introduction gives the idea. Let $\{F_i\}_{i\in I}$ be countable collection of compact convex subsets of complete Cartan-Hadamard source space (X, d). Given such $\{F_i\}_{i\in I}$ we let $\underline{F}$ denote the chain sum

$$\underline{F} = \underline{F}_I = \sum_{i\in I} F_i,$$

in the sense of singular chains and singular homology, e.g [GJ81].

Now suppose $\{F_i\}_{i\in I}$ consists of distinct compact convex sets, and

$$\mathscr{E}[F_i \cap F_j] \subset \mathscr{E}[F_i] \cap \mathscr{E}[F_j]$$

for all indices $i, j \in I$. Abbreviate $F_{ij} := F_i \cap F_j$. For every index i, assume F_{ij} is nonempty for only finitely many indices j. Each F_{ij} is compact convex subset. Let F'_{ij} denote the relative-interior $F'_{ij} := F_{ij} - \partial F_{ij}$.

Definition 5.1.1. In the above notation, let $\{G^-\} = \pi_0(\cup_{i,j} F'_{ij})$ be the set of connected components of $\cup_{i,j} F'_{ij}$. The set of **gates** $\{G\}$ of the chain sum $\underline{F} = SUM[\{F_i\}_{i \in I}]$ is the set $\{G\}$ of closures $G := \overline{G^-}$ of the connected components of $\cup_{i,j} F'_{ij}$.

In otherwords we consider the set $\{G^-\} := \pi_0(\cup_{i,j} F'_{ij})$, and define the set of gates $\{G\}$ as the topological closures $G = \overline{G^-}$ of the connected components $\{G^-\} = \pi_0(\cup_{i,j} F'_{ij})$. The hypotheses of Definition 5.1.1 imply $\cup_{i,j} F_{ij}$ is a closed subset of X.

Definition 5.1.2. The gates $\{G\}$ of a chain sum $\underline{F} = \sum_i F_i$ are well-separated if the components G, G' of the set $\{G\}$ are pairwise isometric.

Now we specialize via group symmetries. Suppose Γ is countable group acting by isometries on a complete Cartan-Hadamard space (X, d). If F is compact convex subset of X, then the set of Γ-translates $F.\Gamma$ determines a chain sum $\underline{F} = \sum_\gamma F.\gamma$. The gates $\{G\}$ of the chain sum $\underline{F}$ form a Γ-set, i.e. the set of gates $\{G\}$ is Γ-invariant, and therefore supports a Γ-action. Indeed gates correspond to nontrivial intersections $F.\gamma \cap F.\delta \neq \emptyset$ for $\gamma, \delta \in \Gamma$, $\gamma \neq \delta$.

The convex chain sums arising from isometric Γ-actions will feature in our applications below. There is a further useful hypothesis which ensures the gates are as Γ-symmetric as possible. Recall that a Γ-set is principal if Γ acts simply transitively (equivalently, there exists unique orbit and orbit map is a bijection).

Definition 5.1.3. The chain sum $\underline{F} = \sum_\gamma F.\gamma$ has Γ-well-separated gates if the Γ-set of gates $\{G\} = \{F.\gamma \cap F.\delta \neq \emptyset\}$ is a principal Γ-set. Or equivalently, the gates are well-separated with respect to Γ if there exists some fixed gate G' for which all other gates are uniquely Γ-isometric.

The applications we consider in Chapters 6, 7 below, are mainly concerned with F semiregular convex polyhedra. Specifically we will find a finite collection of regular polytope "panels" $G, G', \ldots$ and define $F = conv(G, G', \ldots)$ such that $\mathscr{E}[F] = \mathscr{E}[G] \cup \mathscr{E}[G'] \cup \cdots$.

5.2 Hypotheses and Practical Applications

The topological results of our thesis address two issues. The first is abstract. We identify hypotheses on costs c which guarantee that activated source domains of c-optimal semi-couplings admit large-codimension retracts. See Theorems 4.4.3, 4.4.4. The second issue addressed by this thesis is practical, and concerns the application of our general theory to particular costs.

The reader may recall that (Twist) ensures the uniqueness of c-optimal semicouplings, c.f. Proposition 2.5.8. (Twist) condition, i.e. Assumption (A4), is weaker condition than the usual (Twist) from the coupling theory. In our settings the target space $Y = \partial_* F[t]$ is of strictly smaller dimension than the source $X = F$, with $\dim(Y) < \dim(X)$. Here is advantage of semicouplings setting, and motivating the Conjecture 5.9.7.

Next we remark on the general hypotheses necessary for the practical applications of our costs in Chapters 6, 7. If $\underline{F}$ is a chain sum with well-separated gates $\{G\}$ (see §5.1), then the costs $c : \underline{F} \times \partial_*[\underline{F}] \to \mathbb{R} \cup \{+\infty\}$ best suited for our applications have the following properties:

(D0) the cost c is repulsive with $c(x, y) = +\infty$ whenever $x, y \in \partial_*[\underline{F}]$;

(D1) $dom(c)$ consists of pairs (x, y) which occupy a commmon chain summand F' of $\underline{F}$, hence $c(x, y) = +\infty$ if x, y occupy disjoint chain summands;

(D2) if G is a gate of $\underline{F}$ and $x \in G$, then $dom(c_x) \subset G$;

(D3) the cost c satisfies Assumptions (A0)–(A6), c.f. §1.3;

(D4) restricting c to a gate G yields a cost $c|G : G \times \mathscr{E}[G] \to \mathbb{R} \cup \{+\infty\}$ which satisfies sufficient (UHS) conditions;

(D5) if $\mathscr{H}_G$ is the Hausdorff measure on G, with canonical measure $\mathscr{H}^{can}_{\mathscr{E}[G]}$ on $\mathscr{E}[G]$, then the homotopy-reductions of Theorems 4.4.3, 4.4.4 with respect to $c|G$-optimal semicouplings from source $\mathscr{H}_G$ to target $\mathscr{H}^{can}_{\mathscr{E}[G]}$ yield deformation retracts of the gates G to points $G \rightsquigarrow \{pt\}$.

It is nontrivial problem to define a cost c satisfying (D0)–(D5). However we propose the visible repulsion costs defined in 5.9 satisfy these properties. In low dimensions we have successfully verified the properties (D0)–(D5) hold for special cases of the visibility costs. In the following sections we define costs $c = c_1, c_2, v$ which are readily seen to satisfy properties (D0)–(D2). However properties (D3)–(D5) are more difficult to establish, and indeed fail to hold in certain cases.

Example. Let $X = \blacktriangle^2$ denote a two-dimensional regular simplex, equipped with two-dimensional Hausdorff measure σ. The target $Y = \mathscr{E}[\blacktriangle] = \{y_0, y_1, y_2\}$ is a discrete three-point set, with target measure τ positively proportional to $\delta_{y_0} + \delta_{y_1} + \delta_{y_2}$. Suppose σ, τ satisfy

$$\rho := \int_X \sigma / \int_Y \tau \geq 1.$$

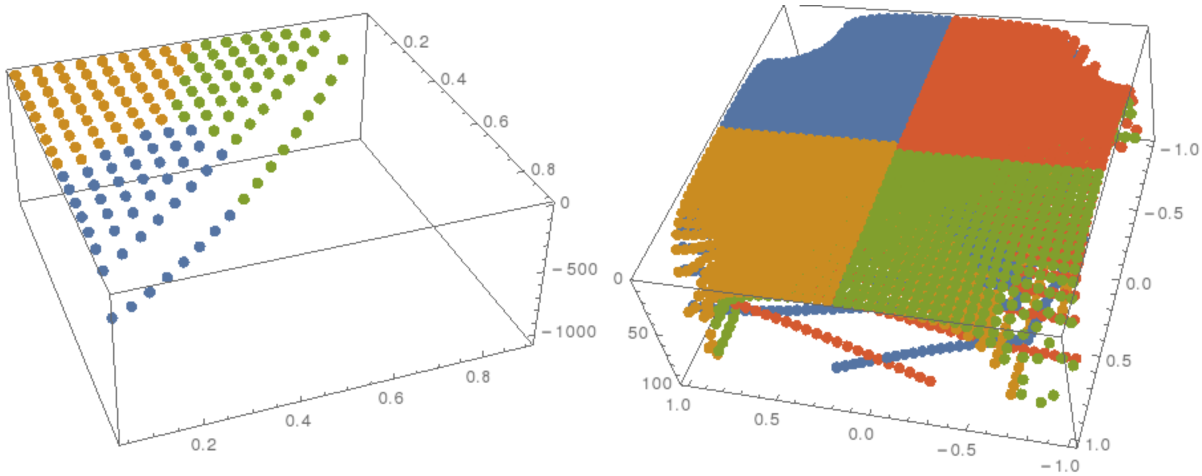

Figure 5.1: Singularity structures of $\widetilde{c}$ on regular 2-simplex and regular square satisfy properties (D0)–(D5)

We define

$$v(x, y_0) := \frac{1}{2}||x - y_0||^{-2} + \frac{\lambda_1}{1-\lambda_0}||x - y_1||^{-2} + \frac{\lambda_2}{1-\lambda_0}||x - y_2||^{-2}, \tag{5.1}$$

where $0 \leq \lambda_i \leq 1$ are the unique scalars satisfying

$$\lambda_0 + \lambda_1 + \lambda_2 = 1, \quad bar(\lambda_0 \delta_{y_0} + \lambda_1 \delta_{y_1} + \lambda_2 \delta_{y_2}) = x.$$

Now suppose I is a countable set, and $\{\blacktriangle_i | i \in I\}$ is a countable collection of simplices with chain sum $\underline{F} = \sum_{i \in I} \blacktriangle_i$ having well-separated gates $\{G\}$ isometric to a given one-dimensional simplex (i.e. fixed compact interval). The visibility cost v defined in equation (5.1) extends to a repulsion cost $\underline{v}$ on $\underline{F} \times \mathscr{E}[\underline{F}]$ satisfying properties (D0)–(D5) as above.

We comment on the conditions (D0)–(D5). The condition (D0) says transporting mass between points on the excision boundary are infinitely prohibitive. (D1) says mass can only be transported within a given chain summand F' of $\underline{F}$. (D2) is a strengthening of (D1): if a source point x occupies a gate $x \in G$, then mass at x can only be transported at finite cost to targets within the gate. In other words mass supported on the gate G is confined to the gate in any finite cost transport plan. (D4)-(D5) concern the restrictions $c|G$. The goal is to ensure the restricted transport problem along the gates is sufficiently well-behaved. In our case this means the Theorems 3.1.1, 4.4.4 can be applied to homotopy-reduce the gates G to points, $G \rightsquigarrow \{pt\}$. Of course the gates G are contractible, since they are basically convex, but the point is that the retractions of 4.4.4 successfully retract G to the maximal codimension.

Applying 4.4.4 to costs c on $\underline{F} \times \partial_*[\underline{F}]$, we then obtain continuous homotopy reductions which are interpolations of the local "gated" retracts $G \rightsquigarrow \{pt\}$ defined by $c|G$. The goal is to obtain a retraction $\underline{F} \rightsquigarrow \underline{Z}$ satisfying

$$\dim(\underline{Z}) + \dim(G) = \dim(\underline{F}).$$

5.3 Repulsion Costs $c|\tau$, c_1

Let F be a compact convex subset of the Euclidean space $\mathbb{R}^N$ and equipped with standard distance $d(x, y) = ||x - y||$. Recall:

- $\dim_{\mathscr{H}}(F)$ is the dimension of the minimal affine subspace containing F.

- The boundary $Y = \partial F$ is closed subset, and F, ∂F have integral Hausdorff dimensions $\dim_{\mathscr{H}}(F) = \dim_{\mathscr{H}}(\partial F) + 1$.

-The distance d on F restricts to a distance $dist = d|Y \times Y$ on Y.

- A point $x \in F$ is an extreme point if x is not the midpoint of any pair of distinct points $x_0, x_1 \in F$, $x_0 \neq x_1$, and where $x = [x_0, x_1]_{1/2}$ implies $x_0 = x_1 = x$.

- Alexandrov's construction ([Ale06], [Oli07]) of Gauss curvature using "spherical images" is a Radon measure ω on ∂F, defined as follows: let $y \in \partial F$ be a boundary point, and consider the set of all normal hyperplanes $N(y)$ supporting F at y. The subset $N(y)$ is a closed convex subset of the linear dual space $(\mathbb{R}^N)^*$. The measure ω on ∂F is defined to be the spherical-measure of $N(y)$.

- The extreme point set $E := \mathscr{E}[F]$ is a subset of ∂F, possibly not closed and with irrational Hausdorff dimension. For example Cantor's middle-third construction and the so-called Cantor staircase function readily leads to compact convex sets with extreme points $\mathscr{E}[F]$ homeomorphic to a Cantor set.

- $\mathscr{E}[F]$ and ∂F may coincide. If F is strictly convex, then ∂F contains no nontrivial affine segments, and $\mathscr{E}[F] = \partial F$.

All our constructions apply to compact convex subsets F with the following (IDE) property:

Definition 5.3.1. A convex compact subset F has integral-dimension extreme points **(IDE)** if there exists subsets $E^{(0)}, E^{(1)}, E^{(2)} \ldots$ of ∂F partitioning the extreme pointset

$$\mathscr{E}[F] = E^{(0)} \coprod E^{(1)} \coprod \cdots ,$$

and where $E^{(j)}$ is either empty or has constant local Hausdorff dimension

$$\dim_{\mathscr{H}}^{loc} E^{(j)} = j$$

for $j = 0, 1, 2, \ldots$.

Recall a subset D has local Hausdorff dimension d at $x \in D$ if the intersection $D \cap B_r(x)$ has Hausdorff dimension d for sufficiently small r-balls $B_r(x)$ centred at x. In otherwords, if

$$\liminf_{r \to 0^+} \dim_{\mathscr{H}}(D \cap B_r(x)) = d.$$

for every $x \in D$. The subset D has constant local Hausdorff dimension if the local Hausdorff dimension is constant with respect to $x \in D$.

If F has property (IDE), then the extreme points support a sufficiently canonical Hausdorff-type measure

$$\mathscr{H}_E^{can} := \mathscr{H}_{E^{(0)}} + \mathscr{H}_{E^{(1)}} + \cdots .$$

Plainly F has (IDE) if either $\mathscr{E}[F] = \partial F$, or if $\mathscr{E}[F]$ is discrete and F is polyhedral.

Definition 5.3.2. Let F be compact convex set with property (IDE) and extreme points $E := \mathscr{E}[F]$. Let τ be a Radon measure on E absolutely-continuous with respect to $\mathscr{H}_E^{can}$. For $y \in E$, let $\mathbf{e}(y)$ be the local Hausdorff dimension.

The repulsion cost $c|\tau : F \times E \to \mathbb{R} \cup \{+\infty\}$ is defined

$$c|\tau(x, y_0) := [\int_E d(x,y)^{-2-\mathbf{e}(y)} d\tau(y)] - \frac{1}{2} d(x, y_0)^{-2-\mathbf{e}(y_0)}.$$

The exponent $\mathbf{e}$ in Definition 5.3.2 ensures that $c(x, y)$ diverges to $+\infty$ whenever x converges to any point $x \to y$ in E. Indee our choice of $\mathbf{e}$ is motivated by the observation that $\int_{x \in \mathbb{R}^N | \ ||x|| < 1} ||x||^{-p}$ diverges to $+\infty$ whenever $p \geq N$ in Euclidean space. The term $-1/2.d(x, y_0)^{-2-e}$ imparts a definite "home-preference" to $c|\tau$ on $F \times E$. When a source point x is close to an extreme point y_0 in E, then the various pairings $c|\tau(x, y_*)$, $y_* \neq y_0$, dominates $c|\tau(x, y_0)$. So when a point x is activated there is no confusion in the optimization program: the point x goes to the lowest-cost target y_0.

Example. In case $E = \partial F$ and $\dim_{\mathscr{H}}(E) = N$, we directly define c_1 with respect to the Hausdorff measure $\tau = \mathscr{H}_{\partial F}$, and obtain

$$c_1(x, y_0) := (\int_{\partial F} d(x,y)^{-2-N} d\mathscr{H}_{\partial F}(y)) - \frac{1}{2} d(x, y_0)^{-2-N}.$$

Example. Let $Y = \{0,1\}$ be subset of $X = conv[Y] = [0,1]$ and $\tau = \delta_0 + \delta_1$. If $x \in X$, then

$$c|\mathscr{H}_Y(x,0) = \frac{1}{2}|x|^{-2} + |x-1|^{-2}$$

and

$$c|\mathscr{H}_Y(x,1) = |x|^{-2} + \frac{1}{2}|x-1|^{-2}.$$

The graph is modelled in Figure 5.3.

Example. Let $Y := \{y_0, \ldots, y_5\}$ be extreme points of a closed hexagon $X := conv[y_0, \ldots, y_5]= conv[Y]$ in $\mathbb{R}^2$, defined by $y_k = e^{2\pi k/6}$ for $k = 0, 1, \ldots, 5$. Represent the distribution on $Y = \mathscr{E}[X]$ by atomic measures $\mathscr{H}_Y = \delta_{y_0} + \cdots + \delta_{y_5}$. If x is contained in the convex hull X, then the cost of transporting a unit mass from x to $y_0 = (1,0)$ is

$$c|\mathscr{H}_Y(x, y_2) := \frac{1}{2} \cdot d(x,y_0)^{-2} + d(x,y_1)^{-2} + \cdots + d(x,y_5)^{-2}.$$

The graph is modelled in Figure 5.3.

Example. Let F be a three-dimensional ellipsoid with extreme points $\mathscr{E}[F] = \partial F =: Y$. Let $\tau = \mathscr{H}_Y$ be uniform Hausdorff measure on the extreme points. Then we have

$$c(x,y_0) := (\int_Y d(x,y)^{-4} d\mathscr{H}_Y(y)) - \frac{1}{2} d(x,y_0)^{-4}.$$

Having defined the repulsion-costs, we now consider which Assumptions (A0), (A1), etc., are satisfied.

Proposition 5.3.3. *Let $X = F$ be a compact geodesically-convex set with property (IDE) as above. Let $Y = E = \mathscr{E}[F]$ be the extreme-point set with canonical measure $\mathscr{H}_E^{can}$ and τ a Radon measure on E absolutely continuous with respect to $\mathscr{H}_E^{can}$.*

Then $c|\tau$ defined in Definition 5.3.2 satisfies Assumptions (A0)–(A5) throughout $dom(c) = (X - Y) \times Y$.

Proof. Evidently $x \mapsto d(x,y)^{-2-\mathbf{e}(y)}$ is smooth and strictly positive for $x \neq y$, and diverging to $+\infty$ when $x \to y$. Now examine the integral defining $c(x,y_0)$ in Definition 5.3.2. If $x \in X - Y$, then the integrand is smooth and finitely-valued with respect to $y \in E$. Now E is relatively-compact, and integrating over E we find $c|\tau$ is uniformly-continuous on compact subsets of $dom(c) = (X - Y) \times Y$. This proves (A0). By similar arguments, applied to $\nabla_x d(x,y_0)^{-2-e}$ and $\nabla^2_{xx} d(x,y_0)^{-2-e}$, we find $c|\tau$ is twice-continuously differentiable on $dom(c)$. Again, since Y is relatively compact we deduce $\nabla^2_{xx} d(x,y_0)^{-2-e}$ varies uniformly with respect to $y \in Y$. Likewise we find the sublevels of $c_y : dom(c_y) \to \mathbb{R}$

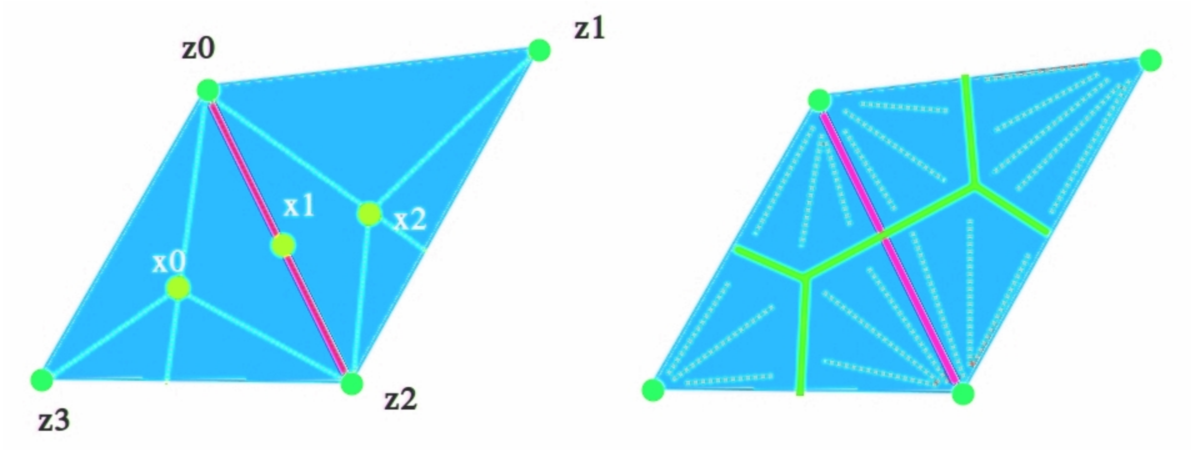

Figure 5.2: The yellow line designates the one-dimensional gate G. The blue line designates the interpolated singularity structure with respect to the gated-visibility cost v^*.

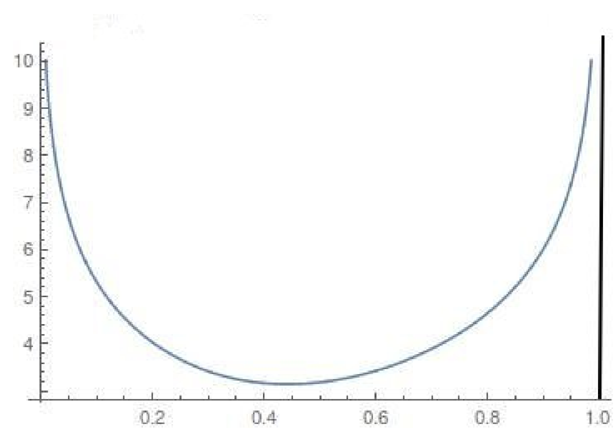

Figure 5.3: Graph of repulsion cost $x \mapsto c|\mathscr{H}_Y(x, 0)$. The assymmetry reflects the "home-preference" of x to $y = 0$.

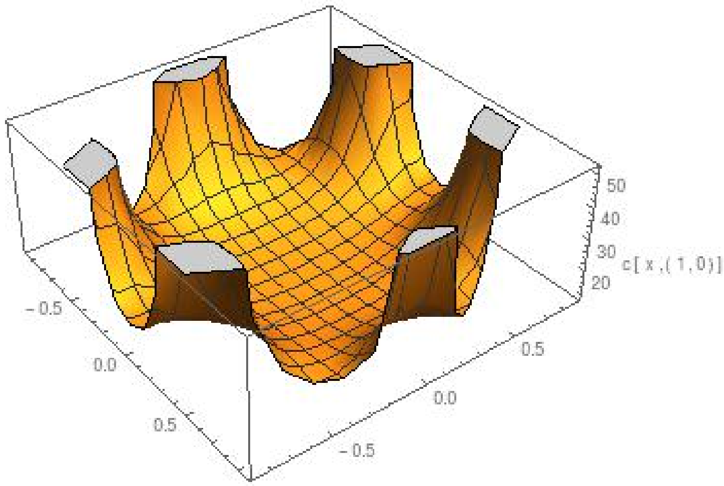

Figure 5.4: Graph of repulsion cost $x \mapsto c|\mathscr{H}_Y(x,(1,0))$. The "home-preference" implies the pole at $y=(1,0)$ has smaller diameter than the other poles.

are compact subsets of X. This proves (A1). Now we also see $||\nabla_x c(x,y)||$ varies continuously with respect to y. So (A2) is satisfied. Finally the reader finds that variations in $d(x,y_0)^{-2-e}$ prove the cost is nowhere locally constant, and (A3) is satisfied. The verification of $(A5)$ is likewise left to the reader. Now we address Assumption (A4), the important (Twist) condition. Abbreviate $dy := d\mathscr{H}_Y(y)$. Fix $x \in F$. Then (Twist) requires the rule

$$y_0 \mapsto \int_E \nabla_x q(x,y)^{-1} dy - \frac{1}{2}\nabla_x q(x,y_0)^{-1}$$

be an injective mapping $E \to T_{x'}F$. The left summand is independant of y_0. Moreover convexity of the domain F implies the injectivity of the rule $y_0 \mapsto \nabla_x q(x,y_0)^{-1}$. This establishes (Twist).

□

5.4 Two-Pointed Repulsion Costs c_2

The simplest case of the repulsion cost $c|\tau$ occurs for F a compact one-dimensional interval. This section modifies the previous definition to define a type of "two-pointed" repulsion cost abbreviated c_2 and defined throughout $F \times \partial F$.

Let F be one-dimensional compact convex set with $\mathscr{E}[F] = \partial F = \{y_0, y_1\}$, and define

$$c_2(x,y_0) := \frac{1}{2}dist(x,y_0)^{-2} + dist(x,y_1)^{-2}$$

as in previous Definition 5.3.2. Next suppose F is a compact convex set, with $x \in F$, $y \in \partial F$. Consider the directed geodesic ray $\rho(y, x)$ issuing from y towards x. If we extend the ray indefinitely, then $\rho(y, x)$ intersects ∂F at some unique point denoted $y' := proj(y, x)$. We say $proj(y, x)$ is the unique point opposite y with respect to x. For $y \in \partial F$ this defines a projection-type map

$$proj_y : (F - \{y\}) \to \partial F.$$

We use the projection map to extend c_2 to higher-dimensional convex compact sets F. For $x \in F$, $y \in \partial F$, we define

$$c_2(x, y) := \frac{1}{2} dist(x, y)^{-2} + dist(x, proj_y(x))^{-2}. \tag{5.2}$$

The formula (5.2) immediately extends to a convex chain sums $\underline{F}$ with well-separated gates $\{G\}$ as defined in Section 5.1, and this yields our definition of two-pointed costs.

Definition 5.4.1 (Two-Pointed Repulsion Cost)**.** Let $\underline{F}$ be convex chain sum with well-separated gates $\{G\}$ (Definition 5.1.2). For every $y \in \mathscr{E}[\underline{F}]$ the projection map $proj_y(x)$ is defined whenever x, y occupy at least one chain-summand F', and the formula

$$c_2(x, y) := \frac{1}{2} dist(x, y)^{-2} + dist(x, proj_y(x))^{-2}$$

defines the two-pointed repulsion cost $c_2 : \underline{F} \times \mathscr{E}[\underline{F}] \to \mathbb{R} \cup \{+\infty\}$.

The projection $proj$ varies continuously with x, y, and we find c_2 is continuous throughout its domain, where $dom(\widetilde{c}) = \{(x, y) \mid proj_y(x) \neq x \neq y\}$. Our hypotheses on the well-separated gates $\{G\}$ of $\underline{F}$ implies the gates G are closed subsets of the boundaries $G \subset \partial F$. If $y \notin G$ and $x \in G$, then $proj_y(x) = x$ and $c_2(x, y) = +\infty$.

This leads to the following observation, that c_2-optimal semicouplings π satisfying $\int c_2(x, y) d\pi(x, y) < +\infty$ will restrict to $c_2|G$-optimal semicouplings $1_G.\pi$ on the gates G of $\underline{F}$. Therefore c_2 satisfies Hypothesis (D2) from Section 5.2.

Now we consider the various cost assumptions.

Lemma 5.4.2. *Let $\underline{F} = \sum_{i \in I} F_i$ be convex chain sum with gates $\{G\}$ satisfying Definition 5.1.3. Then the two-pointed repulsion cost c_2 defined in (5.2) satisfies Assumptions (A0)–(A3) and (A5) throughout $dom(c_2)$.*

Proof. Assumptions (A0)–(A3), (A5) are readily verified, and we leave formal details to the reader.

∎

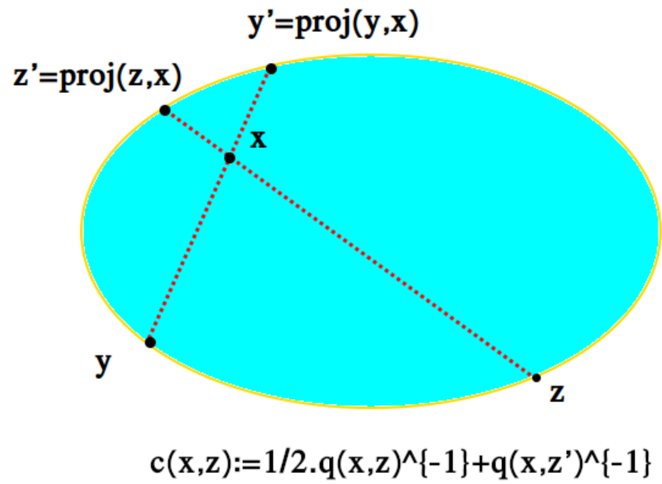

Figure 5.5: Illustrating the projection map $proj : F \to \partial F$ defining two-pointed repulsion cost c_2 for ellipse F where $\partial F = \mathscr{E}[F]$.

The cost c_2 will often fail to satisfy (Twist). Discussions with Professor R.J. McCann readily demonstrate the existence of convex sets F for which c_2 violates (Twist). Indeed symmetry arguments prove the existence of distinct y_0, y_1 in $\mathscr{E}[F]$ for which

$$x_i := argmin_{x \in X}\{c(x, y_i)\}, \quad for\ i = 0, 1,$$

coincide $x_0 = x_1$. In other words, the cost to y_0 and y_1 can be minimized at the same source point. For example, the regular right-angled square will fail (Twist) along an 'X'.

5.5 Convex-Excisions

Let (X, d) be a finite-dimensional Cartan-Hadamard space, [BGS85], and let F be a totally-geodesic compact subset of X. This implies the unique existence of geodesics between pairs of points x, y in X. Antipodal and focal points are nonexistant.

Our constructions are based on analogies between extreme points $\mathscr{E}[F]$ of F and the visual sphere at infinity $X(\infty)$ of X. Such analogies were developed by Thurston [Thu02] and Gromov [Gro82] in their applications of ideal simplices in X with vertices supported on $X(\infty)$.

The idea of excision, which is familiar from Eilenberg-Steenrod axioms of singular homology [GJ81], [Bre93], and which formalizes the idea of "scooping out" convex subsets.

Definition 5.5.1 (Convex-Excision Parameter t)**.** Let F be compact geodesically-convex subset of X. Let I be a discrete subset of $\mathscr{E}[F]$. An excision-parameter consists of a function $t : I \to \mathbb{R}$ and the collection of open subsets $\{W_\lambda^t\}$ for every $\lambda \in I$ having the following properties:

(i) the boundaries ∂W_λ^t are smooth submanifolds for every $\lambda \in I$;

(ii) the boundaries ∂W_λ^t pairwise intersect transversally;

(iii) if $K \subset X$ is compact subset, then $W_\lambda^t \cap K = \emptyset$ except for finitely many λ.

(iv) The excision is called strictly-convex if the subsets W_λ^t are strictly-convex.

We illustrate the above definition on the nonpositively curved (X, d). C.f. [BGS85, §I.3], §§6.3, 6.5. For every point-at-infinity $\lambda \in X(\infty)$ and basepoint $x_0 \in X$, the horofunction $h_{\lambda,x_0} : X \to \mathbb{R}$ is a geodesically-convex function on X. For every $t(\lambda) \in \mathbb{R} \cup \{-\infty, +\infty\}$, the sublevel

$$W_\lambda^t := \{x \in X |\ h_\lambda(x) < t(\lambda)\}$$

is a convex subset of X and named the "horoball centred at λ with radius t_λ".

Definition 5.5.2 (Convex-Excision Model)**.** Let t be an excision parameter (5.5.1). Then the convex-excision model is the complement

$$F[t] := F - \cup_{\lambda \in I} W_\lambda^t.$$

Notice $F[t]$ is a closed subset of F, generally nonconvex with topological boundary $\partial F[t] \subset F$. Our applications are especially concerned with the "excision-boundary" $\partial_* F[t]$ defined as follows:

Definition 5.5.3 (Excision Boundary)**.** Let t be an excision parameter, with excision model $F[t]$. The excision boundary $\partial_* F[t]$ is defined

$$\partial_* F[t] := \cup_{\lambda \in I} (\partial F[t] \cap \overline{W_\lambda^t}).$$

In otherwords $\partial_* F[t]$ is that subset of $F[t]$ which intersects some boundary component ∂W_λ^t.

5.6 Visibility

The convex-excisions $F[t]$ are not geodesically-convex subsets of F, and this is important observation. But this nonconvexity is no obstruction, and managed by introducing the

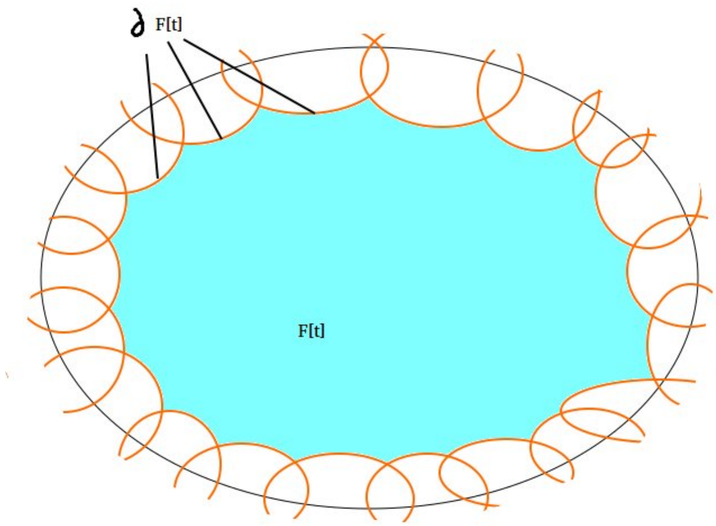

Figure 5.6: Excision of convex ellipsoid

definition of a visibility relation V (Definition 5.6.1). We use the visibility relation to define a visibility factor $k(x, y)$ which rescales the integrands in the repulsion costs introduced above. The formal definitions require some auxiliary lemmas.

Definition 5.6.1 (Visibility Relation V)**.** Let $F[t] \subset F$ be a convex excision (Definition 5.5.2). The pair of points $(x, x') \in F[t] \times F[t]$ are *visible*, and in relation xVx', if the unique geodesic segment γ joining x to x' in F is contained in $F[t]$.

Lemma 5.6.2. *[[McC01, Proposition 6]] Let (X, d) be Cartan-Hadamard space and $F \hookrightarrow X$ a compact geodesically convex subset. let x, y be arbitrary points in F and γ the unique unit-parametrized geodesic $\gamma : [0, T] \to F$ having $T = dist(x, y)$, and $x = \gamma(0)$, $y = \gamma(T)$. Then $\nabla_y d(x, y) = \gamma'(T) \in T_y F$, and $\nabla_x d(x, y) = -\gamma'(0)$.*

If $x \in F[t]$ and $y \in \partial_* F[t]$ are geodesically visible in $F[t]$, then Lemma 5.6.2 implies $\nabla_y d(x, y)$ is equal to the direction of impact of the geodesic ray from x to y. Therefore the cosine of the angle-of-impact at y (relative to an outward unit normal) is represented by the dot-product $\langle \nabla_y d(x, y), \mathbf{n}_y \rangle$ in $T_y F[t]$.

Lemma 5.6.3. *Let $X := F[t]$ be a convex excision of F, with excision-boundary $Y := \partial_* F[t]$. If $(x, y) \in X \times Y$ are geodesically visible in $F[t]$, then $\langle \nabla_y d(x, y), \mathbf{n}_y \rangle \geq 0$ where $\mathbf{n}_y$ is any outward unit normal vector at $y \in Y$ in $F[t] \subset X$.*

Proof. By hypothesis the initial domain F is convex, so x, y are visible along some geodesic γ in F. Consider the possible intersections of γ with X and Y. The convexity of the excised horoballs W_λ^t defining X implies the following: we have $\langle \nabla_y d(x, y), \mathbf{n} \rangle < 0$

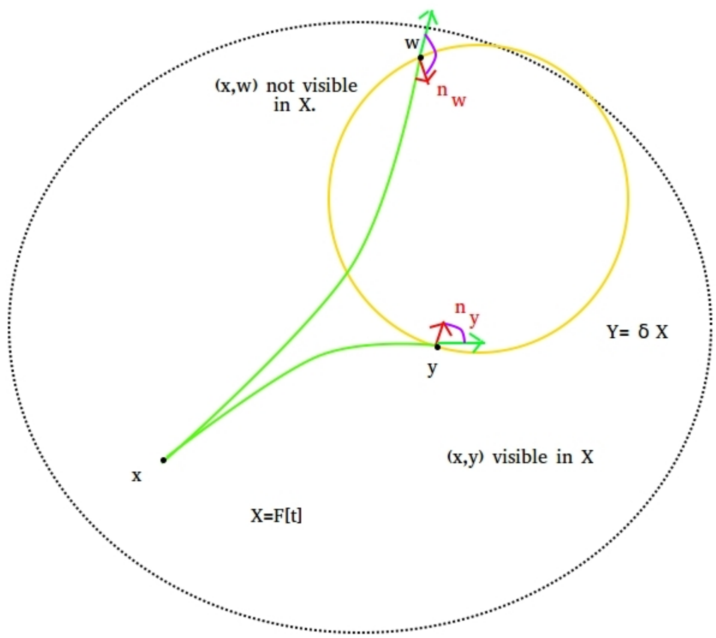

Figure 5.7: Excision domain $X = F[t]$, where (x, y) are visible in X and (x, w) is not visible in X

only if the geodesic γ from x impacts y from within the locally convex subdomain $F - F[t]$ containing y. Or equivalently, we find (x, y) are not visible in X only if the geodesic α exits $F[t]$ at some secondary point $y' \in \partial_* F[t]$. □

Informally, one imagines the cost of transmission from a source point $x \in F[t]$ to a visible target point $y \in \partial F[t]$ is measured by the angle-of-impact (and the quadratic distance) at $\partial F[t]$. We posit that a directed geodesic ray enters the target point most efficiently at $y_0 \in \partial F[t]$ when the angle-of-impact is orthogonal to the tangent space of the boundary, i.e., when the incoming ray arrives from x at a right angle to $T_{y_0} \partial F[t] \hookrightarrow T_{y_0} F[t]$. Conversely, if the boundary $\partial F[t]$ has outward pointing unit normal vector $\mathbf{n}$, then we say the cost of transmitting rays which impact $\partial F[t]$ orthogonally to $\mathbf{n}$ are infinitely prohibitive. Thus we augment the data "directed ray from x to y" , measured by magnitude and direction, with the "angle-of-impact" visibility factor $k(x, y)$.

Definition 5.6.4 (Visibility factor)**.** Let $X := F[t]$ be convex excision, and for $\epsilon > 0$ let Y_ϵ be the ϵ-regularization of the boundary $Y := \partial F[t]$ defined in 5.7.1. The visibility factor is the function $k : X \times Y_\epsilon \to \mathbb{R}_{\geq 0} \cup \{+\infty\}$ defined by the formula

$$k(x,y) := \begin{cases} \langle \nabla_y d(x,y), \mathbf{n}_y \rangle^{-1}, & \text{if } x,y \text{ are visible },\\ +\infty, & \text{if } x,y \text{ not visible }. \end{cases} \tag{5.3}$$

Evidently the Definition 5.6.4 represents a numerical function valued in $[1,+\infty]$, and diverging to $+\infty$ when y fails to be visible from x within $F[t]$ by Lemma 5.6.3.

5.7 ϵ-Regularizations

Definition 5.5.1 produces a manifold-with-corners $F[t]$, having an excision-boundary $Y := \partial_* F[t]$ which is generally not everywhere smooth and not having unique outward normals. The present section introduces a basic regularization Y_ϵ of Y with smoothly varying outward unit normal vectors. This implies the visibility factor $k : X \times Y_\epsilon \to \mathbb{R}_{>0} \cup \{+\infty\}$ is continuous throughout its domain.

The excision boundary Y is cellulated ("divided into cells") by the excision parameter t and the horoballs W^t_λ. The convex horoballs $W' := W^t_\lambda$ have well-defined inward normal vectors $n_y \in T_y W'$ for every $y \in W'$. However a given point $y' \in Y$ can occupy multiple horoballs W^t_λ, and therefore the outward unit normal vector at $T_y F[t]$ is not uniquely defined. We restore uniqueness by replacing Y with Y_ϵ as defined in Proposition 5.7.1 below.

Remark that the boundary $\partial((F[t]))_{\leq \epsilon}$ of the closed ϵ-neighborhood of $\partial F[t]$ defines a $C^{1,1}$-regularization of Y. There are different techniques for C^∞-regularizations, c.f. [Gro91, §3, pp.53], [Gro14b, §3.4, 5.7], and [BS73, Appendix, §6], but we propose the following:

Proposition 5.7.1 (C^∞ Regularization of manifold-with-corners)**.** *Let X be Cartan-Hadamard space, F a closed geodesically-convex subset of X, and t an excision parameter and $F[t]$ a convex excision, c.f. 5.5.1, 5.5.2. Define $Y = \partial_* F[t]$. Then for every $\epsilon > 0$, we can replace Y with a smooth submanifold Y_ϵ of F such that:*

(i) Y_ϵ is contained within the open ϵ-neighborhood of Y; and

(ii) Y_ϵ converges to Y in Gromov-Hausdorff topology to Y as $\epsilon \to 0^+$; and

(iii) there exists a degree-one 1-Lipschitz map $p : Y_\epsilon \to Y$.

Sketch of Proof 5.7.1. The following approach was suggested by Professor R.J. McCann. Recall the Definition 5.5.1 includes three hypotheses: boundaries ∂W_i are smooth, boundaries are pairwise transverse, and intersect local-finitely on compacta. This implies the excision boundaries Y, Y_ϵ are locally modelled on "standard orthogonal sectors" in $\mathbb{R}^n$.

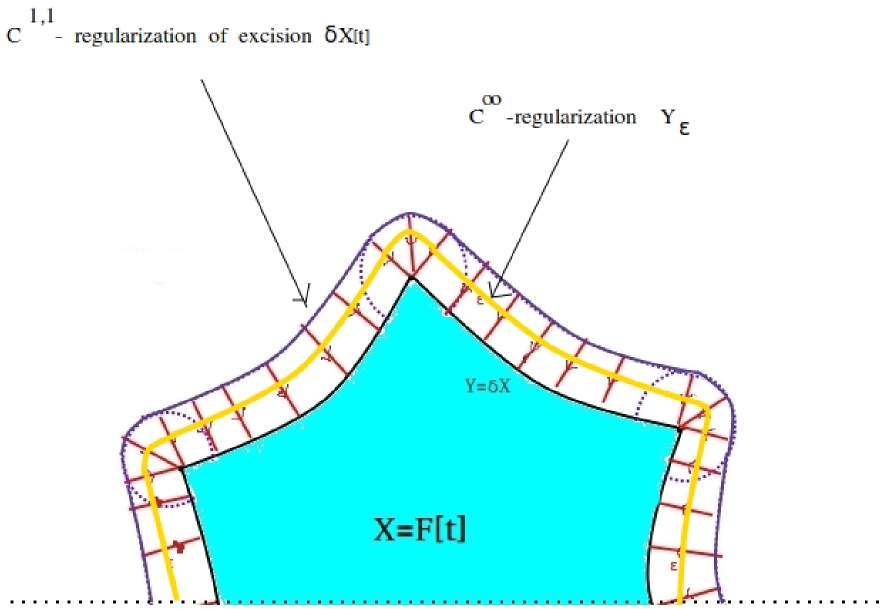

Figure 5.8: The outermost boundary is a $C^{1,1}$-regularization of Y. The inner hypersurface is a smooth ϵ-regularization Y_ϵ

Let $((\partial F[t]))_\epsilon$ be the open ϵ-neighborhood of $\partial_* F[t]$ in X. Next let $u : X \to \mathbb{R}$ be a harmonic function satisfying $u|_{F[t]} \equiv 0$ and $u|_{X-((\partial F[t]))_\epsilon} \equiv 1$. Now define

$$Y_\epsilon := u^{-1}(1/2), \qquad X_\epsilon := \{u \leq 1/2\}. \tag{5.4}$$

Then Y_ϵ is indeed a smooth submanifold satisfying conditions (i)–(iii). For instance, the harmonic function u will have nonvanishing gradient $\nabla_x u \neq 0$ on the open ϵ-neighborhood $((F[t]))_{<\epsilon}$, and therefore the standard "gradient-flow" argument constructs a deformation retract from the sublevels $\{u \leq 1/2\}$ onto $u = 0$. □

For example, consider the unit square $X = [0,1] \times [0,1]$ in $\mathbb{R}^2$. Then X is a manifold-with-corners, and having "sharp" corners at the four extreme points. For every $\epsilon > 0$, the closed ϵ-neighborhood $((X))_\epsilon$ in $\mathbb{R}^2$ is a $C^{1,1}$-manifold-with-boundary. Constructing the harmonic function u as in the proof of Proposition 5.7.1, we find $u^{-1}(1/2)$ is smooth regularization of the boundary $\partial((X))_\epsilon$. The gradient flow produces a degree-one continuous covering map $Y_\epsilon \to Y$.

The regularization from Proposition 5.7.1 replaces a pair of manifolds-with-corners (X, Y) with a pair of smooth manifolds (X_ϵ, Y_ϵ), c.f. (5.4). Visibility between pairs (x, y)

in $X \times \partial X$ and pairs $(x, y') \in X \times Y_\epsilon \subset X_\epsilon \times Y_\epsilon$ are basically equivalent via the map p from Proposition 5.7.1. Indeed if V_ϵ denotes the visibility relation on X, Y_ϵ per Definition 5.6.1, then $xV_\epsilon y'$ if and only if $xVp(y)$. So Y_ϵ has unique outward unit normal vectors $\mathbf{n}_y \in T_yX$ varying smoothly with $y \in Y_\epsilon$. Thus $k(x,y)$ is smooth with respect to visible pairs $(x,y) \in X_\epsilon \times Y_\epsilon$ for sufficiently small $\epsilon > 0$.

5.8 Barycentres and Krein-Milman

The present section describes a useful probabilistic "coordinate system" on the convex hull of a collection E of extreme points. The references [Phe89, Ch.I] and [Bar02b, pp. II.3-4] are useful background on Krein-Milman's theorem from convex geometry: if F is a convex compact body, then every point mass δ_x at a point $x \in F$ can be represented as the centre-of-mass of some mass-distribution supported λ at the extreme-points $\mathscr{E}[F]$ of F. Thus points on F can be coordinatized by probability measures on $\mathscr{E}[F]$. The "centre-of-mass" can be generalized to Riemannian geometry with the following alternative definition, see [Jös97].

Definition 5.8.1 (Riemannian barycentre)**.** Let (X, d, σ) be a complete finite-dimensional metric-measure space. Let λ be Radon measure on (X, d) and absolutely-continuous with respect to σ. Then $x_0 \in X$ is the barycentre of λ with respect to σ if

$$\int d^2(x_0, x) d\lambda(x) = \inf_{p \in X} \int d^2(p, x) d\lambda(x).$$

We abbreviate $x_0 = bar(\lambda|\sigma)$. Using the Riemannian geodesic exponential function $exp_x : T_xX \to X$, we find $q \in X$ is a barycentre of λ if the following critical point condition holds:

$$\int_X exp_q^{-1}(x) d\lambda(x) = 0 \text{ in } T_xX. \tag{5.5}$$

The following standard result asserts the unique existence of barycentres for complete nonpositively curved spaces, c.f. [Jös97, Theorem 3.2.1].

Lemma 5.8.2. *Let (X, d, μ) be a complete finite-dimensional metric-measure length space with nonpositive sectional curvature $\kappa \leq 0$. Let λ be a Radon measure on X and absolutely continuous with respect to μ, with bounded support and $\lambda[X] < +\infty$. Then there exists unique barycentre for λ and unique $x_0 \in X$ with*

$$\int_X d^2(x_0, x) d\lambda(x) = \inf_{p \in X} \int_X d^2(p, x) d\lambda(x).$$

If the background measure μ is Hausdorff-type, e.g. $\mu = \mathscr{H}_Y$, then we abbreviate $bar(-) = bar(-|\mu)$. Now let $E \subset X$ be a closed subset of a finite-dimensional complete space X, and let $\Delta(E) \hookrightarrow \mathscr{M}_{\geq 0}(E)$ denote the weak-$*$ compact subset of Radon probability measures supported on E. The barycentre map defines a weak-$*$ continuous mapping

$$bar : \Delta(E) \to conv[E],$$

and this mapping surjects onto the convex compact hull $conv[E]$ of E in X according to Krein-Milman theorem.

Definition 5.8.3. Let F be closed convex subset. For $x \in F$, let S_x consist of those probability measures λ in $\Delta(\mathscr{E}[F])$ with barycentre $bar(\lambda)$ equal to x.

For every $x \in F$ the subset S_x is nonempty compact convex subset of $\Delta(\mathscr{E})$. For general F and $x \in F$, the subset S_x is often not a singleton. Indeed Choquet's Theorem [Phe89] says F is a simplex if and only if the barycentre mapping $bar : \Delta(\mathscr{E}[F]) \to F$ is injective, i.e. if and only if the barycentre mapping is a weak-$*$ isomorphism and S_x is a singleton for every $x \in F$. Thus we face the problem of selecting a canonical choice of $\lambda_x^* \in S_x$ varying continuously with the point $x \in F$. So let F be a compact convex space satisfying (IDE) conditions (5.3.1). Let $E := \mathscr{E}[F]$ have canonical measure $\mathscr{H}_E^{can}$.

Definition 5.8.4 ([KP18])**.** In the above notation, for every $x \in F$ define λ_x^* to be the unique probability measure of S_x satisfying

$$\{\lambda_x^*\} = argmin\{\lambda \mapsto W_2^2(\lambda, \mathscr{H}_Y^{can}) \mid \lambda \in S_x\},$$

where W_2^2 denotes Wasserstein 2-distance with respect to the quadratic transport costs $c = d^2/2$ (c.f. [Vil09]).

It's well-known that W_2^2-minimizers are unique, and especially for subsets of a convex set F. We refer the reader to [KP18] for further details.

5.9 Visible Repulsion Costs

Throughout this section we presume F is a compact convex polyhedra such that the set of extreme points $\mathscr{E}[F]$ is discrete and finite subset. We continue with our convex excisions of F, but insist that the excised convex horoballs W_λ^t, $\lambda \in \mathscr{E}[F]$, defining the excision $F[t] := \cap_{\{\lambda\}}(F - W_\lambda^t)$ are horoballs centred at the extreme points $\mathscr{E}[F]$. Recall $\Delta(\partial_* F[t])$ consists of all probability measures supported on $\partial_* F[t]$. The following definition is convenient:

Definition 5.9.1. Let $\Delta^v(\partial_* F[t])$ be the set of probability measures λ on $\partial_* F[t]$ such that:

(i) $bar(\lambda) \in F[t]$; and

(ii) the support of λ is a subset of $\partial_* F[t]$ which is visible from $bar(\lambda)$ along geodesics contained in $F[t]$.

Now the excision $F[t]$ is generally a nonconvex subset of F. Our first step is to identify a convenient geodesically convex subset Ω of $F[t]$. We define Ω via the visibility relation $V \subset F[t] \times F[t]$ defined in 5.6.1.

Lemma 5.9.2. *Let F be a convex compact polyhedra and $F[t]$ a strictly convex excision centred on the extreme points $\mathscr{E}[F]$. Then the restricted barycentre map*

$$bar : \Delta^v(\partial_* F[t]) \to F[t]$$

is a continuous surjection.

Proof. Krein-Milman theorem implies $\Delta(\mathscr{E}[F]) \to F$ is a continuous surjection. If $F[t]$ is a convex excision of F centred at the extreme points $\mathscr{E}[F]$, suppose $\lambda \in \Delta(\mathscr{E}[F])$ is such that $bar(\lambda) \in F[t]$. Then we find there exists a measure $\lambda' \in \Delta^v(\partial_* F[t])$ with $bar(\lambda) = bar(\lambda')$. Indeed the geodesics joining the support of λ to x will intersect the excision horoballs W_t defining $F[t]$ at points $\{y'\}$. And a suitable convex combination of the $\{y\}$ will define a measure λ' having barycentre coincident with $bar(\lambda)$. We observe here that convexity of the excised W_t's is necessary hypothesis. $\square$

For $z \in \partial_* F[t]$, abbreviate $V_z := \{x \in F[t] \mid xVz\}$. Then we define

$$\Omega := (\cap_{z \in \partial_* F[t]} V_z). \tag{5.6}$$

The subset Ω consists of all $x \in F[t]$ which are simultaneously visible to the excision boundary $\partial_* F[t]$ by geodesics contained in $F[t]$. See Figure 5.9.

Lemma 5.9.3. *Under the above hypotheses, the subset Ω defined in (5.6) is a geodesically convex subset of $F[t]$.*

Proof. Consider the inclusion $F[t] \hookrightarrow F$. If $y \in \partial_* F[t]$, then the subset $V'_z := \{x \in F[t] \mid xV'_z\}$ is a geodesically convex subset of F for every $z \in \partial_* F[t]$. But observe $\cap_{z' \in \partial_* F[t]} V'_z$ is a subset of $F[t]$ and coincident with $\cap_{z' \in \partial_* F[t]} V_z =: \Omega$. Thus Ω is geodesically convex subset of $F[t]$. $\square$

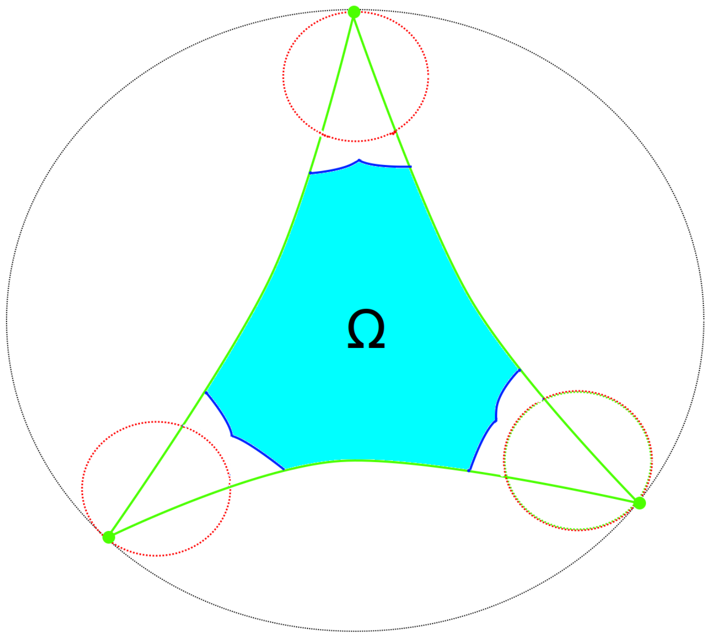

Figure 5.9: The subdomain Ω is geodesically convex

The second step concerns the projection $proj_y$ defined in §5.4. Recall $proj_y : (F - \{y\}) \to \partial F$ was the point of intersection of the unique geodesic issuing from y through x with ∂F. This same construction yields a projection map

$$proj_y : \Omega \to \partial\Omega. \tag{5.7}$$

Moreover suppose Y_ϵ is the ϵ-regularization of $Y = \partial_* F[t]$ as above. For $\epsilon > 0$ sufficiently small, and every $y \in Y_\epsilon$, the projection map (5.7) is well-defined map.

Lemma 5.9.4. *In the above notation, for every $y \in Y_\epsilon$, the projection map (5.7) is continuous.*

Proof. We find $proj_y$ is discontinuous at points x for which the geodesic segment $\alpha(s)$, $s > 0$ with $\alpha(0) = y$, $\alpha(dist(y, x)) = x$, intersects Y_ϵ and makes angle-of-impact exactly $\pi/2$ with respect to the normal vector $\mathbf{n}$ at the point of intersection. But such points x are nonvisible to a nontrivial portion of Y and Y_ϵ, and therefore $x \notin \Omega$. □

According to Lemma 5.9.2 every $x \in proj_y(\Omega)$ is the barycentre of some probability measure $\lambda_x \in \Delta^v(Y)$, $bar(\lambda_y) = y$. Before defining the visibility cost below, we need adapt this observation to Y_ϵ.

Lemma 5.9.5. *In the above notation, let Y_ϵ be the ϵ-regularization of Y. Let $\Delta^v(Y_\epsilon)$ consist of all probability measures λ on Y_ϵ for which* *(a) $bar(\lambda) \in \Omega$; and*

(b) $spt(\lambda)$ is a subset of Y_ϵ which is simultaneously visible from $bar(\lambda)$ along geodesics contained in X_ϵ.

Then:

(i) for every $x \in \Omega$, there exists at least one $\lambda \in \Delta^v(Y_\epsilon)$ such that $bar(\lambda) = x$; and

(ii) there exists a unique W_2^2-minimizer $\lambda_x^* \in \Delta^v(Y_\epsilon)$ such that

$$\{\lambda_x^*\} = argmin_\lambda\{W_2^2(\lambda, \mu_\epsilon)\},$$

where the minimum is taken over all λ satisfying (i) and μ_ϵ is the renormalized Hausdorff probability measure on Y_ϵ.

Proof. Consequence of the proof of 5.9.2 and 5.7.1. We leave the details to the reader. □

Finally with $F[t]$, Y, Y_ϵ, Ω, $proj_y$, λ_x^* as defined above, we present the main definition of this chapter.

Definition 5.9.6 (Visibility Cost)**.** Under the above hypotheses, the visibility cost $v : \Omega \times Y_\epsilon \to \mathbb{R} \cup \{+\infty\}$ is defined:

$$v(x, y_0) := \frac{1}{2}k(x, y_0).q(x, y_0)^{-1} + \int_{Y_\epsilon} k(x, y).q(x, y)^{-1} d\lambda^*_{proj_{y_0}(x)}(y). \tag{5.8}$$

Here k denotes the visibility factor defined in 5.6.4, and $q(x, y) := dist(x, y)^{2+\mathbf{e}}$ for a suitable integer $\mathbf{e} \geq 0$ (recall Definition 5.3.2 from Section 5.3). Definition 5.9.6 is our generalization of equation (5.1) on the simplex to convex-excisions $F[t]$. Following the above definitions we see $v(x, y)$ varies continuously with respect to $(x, y) \in \Omega \times Y_\epsilon$. We expect further properties hold, as summarized in the following:

Conjecture 5.9.7. *Under the above hypotheses, the visibility cost $v : \Omega \times Y_\epsilon \to \mathbb{R} \cup \{+\infty\}$ defined by (5.8) satisfies the hypotheses (D0)–(D5) from Section 5.2.*

If F is compact and strictly convex with $\partial F = \mathscr{E}[F]$, then the formula 5.8 defining $v(x, y)$ reduces to a modified "visual two-pointed" repulsion cost, namely

$$v_0(x, z) := \frac{1}{2}k(x, z).q(x, z)^{-1} + \int k(x, proj_z(x)).q(x, proj_z(x))^{-1} d\lambda^*_{proj_z(x)}. \tag{5.9}$$

Omitting the visibility factor k obviously yields c_2 (Definition 5.4.1). It would be useful to establish Conjecture 5.9.7, as well as establish the following

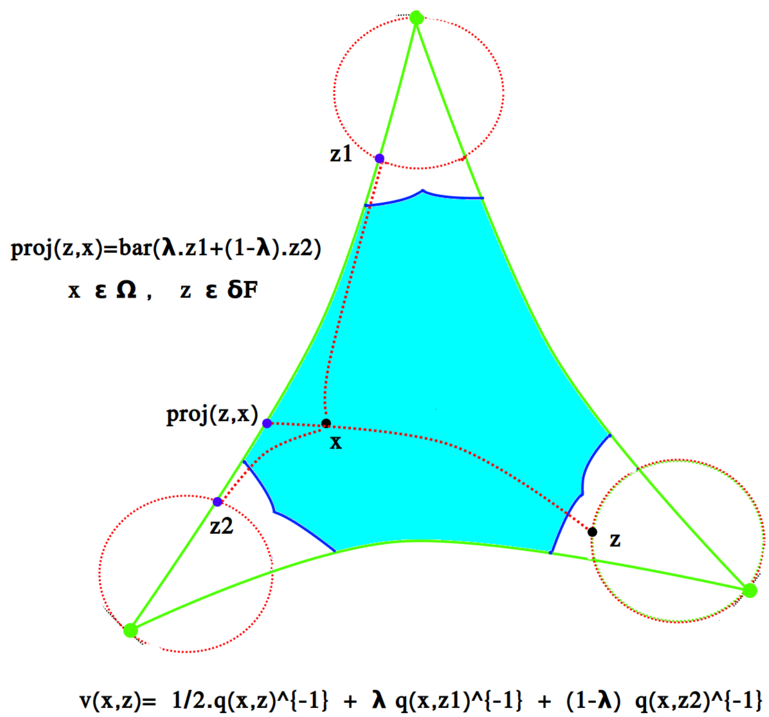

Figure 5.10: Evaluating visibility cost $v(x,z)$ for pair $(x,z) \in \Omega \times Y_\epsilon$

Conjecture 5.9.8. *Under the above hypotheses, every v_0-convex potential $\phi : \Omega \to \mathbb{R}$ has separated subdifferentials (Definition 4.4.1).*

Chapter 6

Excisions and Cohomological Dimension

This chapter begins the second phase of our thesis with the goal of establishing Theorem 1.5.1 from §1.5, and large codimension homotopy-reductions $X \rightsquigarrow \underline{Z}$, when $X \approx E\Gamma$ is a geometric classifying space model for infinite discrete groups Γ satisfying Bieri-Eckmann duality, §1.5.2.

Theorem 1.5.1 is a reduction program drawing together the semicoupling and singularity methods of Chapters 2, 4, 5. We apply these methods to source spaces $X[t]$ obtained by convex excision, with target $Y = \partial X[t]$ the excision boundary, as defined in §5.5. In practice we imagine a user has an infinite discrete group Γ with some standard geometric $E\Gamma$ model $X \times \Gamma \to X$. Given this initial data, the user can follow the excision construction §6.3 and obtain a manifold-with-corners $X[t] \times \partial X[t]$. Next, if the user successfully Closes the Steinberg symbol §1.5, 7), then we replace the excision $X[t]$ with the more convenient chain sum $\underline{F}$. The summands of the chain sum $\underline{F}$ are excisions $F[t]$, and $\underline{F}$ inherits a proper Γ-action where Γ acts as shift-operator on the chain summands. Next, the user needs construct the visibility cost v on the visible chain sum $\underline{\Omega}$ on $\underline{F}$. The chain sum $\underline{\Omega}$ is defined in 5.9.3, see Figure 5.9. The homotopy-reductions from Theorems 1.4.1,1.4.2 deformation-retract $\underline{F}$ onto closed singularities $\underline{Z}$. Everything is Γ-equivariant and the retracts $\underline{Z}$ are small-dimensional $E\Gamma$ classifying spaces.

6.1 Geometric Classifying Spaces $E\Gamma$

The purpose of this section is to introduce geometric $E\Gamma$ models. Useful references include [Bro82], [Bre93]. Let Γ be a finitely-generated infinite group. Poincaré's fundamental group functor $X \mapsto \pi_1(X, pt)$, originally defined in [Poi95, §12], is a bridge to topology. To display Γ as the fundamental group of a connected topological space means constructing the so-called proper classifying space $E\Gamma$. The $E\Gamma$ models can be characterized as universal covering spaces of the Eilenberg-Maclane space $K(\Gamma, 1)$.

Definition 6.1.1. Let Γ be abstract group with discrete topology. An $E\Gamma$ model is a topological space X equipped with a continuous map $\alpha : X \times \Gamma \to X$ satisfying the following properties:

(i) the topological space X is homologically-trivial, so all reduced homology groups vanish $\tilde{H}_i(X; \mathbb{Z}) = 0$ for $i \geq 0$;

(ii) the continuous map α is group action satisfying $\alpha(x, \gamma\delta) = \alpha(\alpha(x, \gamma), \delta)$ for all $x \in X$, $\gamma, \delta \in \Gamma$, and $\alpha(x, 1_\Gamma) = x$. We abbreviate $\alpha(x, \gamma) = x.\gamma$;

(iii) the action is proper-discontinuous, so for every $x \in X$ and bounded open neighbourhood U of x, there exists only finitely many $\gamma \in \Gamma$ such that $U \cap U.\gamma$ is nonempty;

(iv) the action is free, so for all $x \in X$, $\gamma \in \Gamma$, we have $x.\gamma = x$ if and only if $\gamma = Id$.

Examples of $E\Gamma$ spaces abound, and every pair X, X' of $E\Gamma$-models are homotopic.

- The universal covering space $\mathbb{R} \to \mathbb{R}/\mathbb{Z} \approx S^1$ is $E\mathbb{Z}$ model, where $\mathbb{R} \times \mathbb{Z} \to \mathbb{R}$ is defined by additive translation $(x, n) \mapsto x + n$.

- The n-dimensional torus $\mathbb{R}^n \to T^n = (S^1)^n$ defines $E\mathbb{Z}^n$ model.

- If $K \subset \mathbb{S}^3$ is a 1-dimensional knot, then excising the open ϵ-neighborhoods $N_\epsilon(K)$ from $\mathbb{S}^3$ produces a three-dimensional manifold-with-boundary $X = \mathbb{S}^3 - N_\epsilon(K)$. The universal covering $\tilde{X}$ is known to be an $E\Gamma$ model for $\Gamma = \pi_1(X, pt)$.

- The Poincaré disk $\mathbb{H}^2$ is an $E\Gamma$-model for every torsion-free finite-index subgroup Γ of $PGL(\mathbb{Z}^2)$.

- The quotient $X = K\backslash Sp(\mathbb{R}^4)$ of the group of linear symplectomorphisms of the standard four-dimensional symplectic space $(\mathbb{R}^4, \omega)$ by a maximal compact subgroup $K \approx U(2) = SO(4) \cap Sp(\mathbb{R}^4)$ admits a proper discontinuous right-action by the arithmetic group $Sp(\mathbb{Z}^4)$. One knows that X is a $E\Gamma$-model for every finite-index torsion-free subgroup Γ of $Sp(\mathbb{Z}^4)$.

- Teichmueller's space $\mathscr{T}_g$ is an $E\Gamma$ model for the mapping class group $\Gamma = MCG(\Sigma_g)$ of a closed orientable genus g surface Σ_g.

- Further examples include braid groups, right-angled Artin groups, etc., and almost all the groups that arise from geometric group theory and three-dimensional topology.

Our applications are especially concerned with *geometric* $E\Gamma$ models, per the following definition:

Definition 6.1.2. An $E\Gamma$ model X is geometric if X has a complete Cartan-Hadamard Riemannian metric g for which the action $X \times \Gamma \to X$ is isometric, and if the volume measure vol_X has finite covolume with respect to X (i.e. the quotient X/Γ has finite volume with respect to vol_X).

Constructing $E\Gamma$ models is basic to practical computations. For instance, effective $E\Gamma$ models solve the word problem on the abstract group Γ insofar as a group element $\gamma \in \Gamma$ can act by translations on X. The hypotheses of Definition 6.1.1 means we can distinguish γ from the identity element Id by finding some (any) point $x \in X$ which is displaced some positive (> 0) distance, and then $x.\gamma \neq x$ implies $\gamma \neq Id$ in Γ. This observation is basic to Fricke-Klein's ping-pong argument (see [Tit72]).

The $E\Gamma$ models are connected spaces X with action $X \times \Gamma \to X$. Viewing Γ as discrete topological space, the group action $\Gamma \times \Gamma \to \Gamma$ defined by $(\delta, \gamma) \mapsto \delta \cdot \gamma$ almost satisfies properties of Definition 6.1.1 with the exception of (i), namely the connectivity hypothesis that the reduced homology groups $\tilde{H}_*$ simultaneously vanish. But of course a discrete group Γ is generally disconnected with respect to the discrete topology. So $E\Gamma$ models $X \times \Gamma \to X$ are maximally-connected interpolations of the principal action $\Gamma \times \Gamma \to \Gamma$.

The proper-discontinuity hypothesis has important measure-theoretic consequences regarding so-called Radon measures. Recall that a Radon measure is a Borel measure which gives finite measure to compact subsets. A given point orbit $x.\Gamma$ is discrete in X, and fixed point free. Naturally we interpret the orbit $\sum_{\gamma \in \Gamma} \delta_x.\gamma$ of the Dirac atomic mass δ_x at x as representing a unit Dirac measure on the quotient. The proper discontinuity hypothesis ensures the correspondance between Radon measures on the topological quotient X/Γ and Γ-equivariant Radon measures on X defines a weak-* homeomorphism $\mathscr{M}_{\geq 0}(X)_\Gamma \approx \mathscr{M}_{\geq 0}(X/\Gamma)$.

6.2 Background: Group Cohomology

The following section reviews the basic facts of group-cohomology, i.e. the study of projective and free resolutions of $\mathbb{Z}\Gamma$-modules. Our treatment follows [Bro82]. These formalities are necessary for the definition of Bieri-Eckmann duality (§6.4) and Closing the Steinberg symbol (Ch.7).

Effectively computing the topological invariants of a group Γ is practically impossible without explicit $E\Gamma$ models. Recall $\mathbb{Z}\Gamma$ denotes the integral group-ring, consisting of finitely-supported $\mathbb{Z}$-valued distributions on the discrete group Γ. If X is $E\Gamma$ model, then the topologists' standard projective resolution of $\mathbb{Z}\Gamma$-modules over $\underline{\mathbb{Z}}$ is obtained via the singular chain complex $\{C_n(X)^{sing}, \partial_n\}_n$ on X. Here $\underline{\mathbb{Z}}$ denotes the additive abelian group $\mathbb{Z}$ equipped with trivial Γ action, where $n.\gamma = n$ for all $n \in \mathbb{Z}$, $\gamma \in \Gamma$.

More formally, let Γ be a discrete group. The category of linear representations of Γ is equivalent to the category of $\mathbb{Z}\Gamma$-modules. When M, N are Γ-modules, then $M \otimes N$ inherits Γ-module action via the diagonal action $m \otimes n.\gamma = m.\gamma \otimes n.\gamma$, and we set $M \otimes_\Gamma N := (M \otimes N)_\Gamma$ the coinvariant module, or quotient of $M \otimes N$ (tensor product as $\mathbb{Z}$-modules) by the Γ-action. If M, N are $\mathbb{Z}\Gamma$-modules, then $Hom(M, N)$ (which is $\simeq N \otimes M^*$) inherits $\mathbb{Z}\Gamma$-module structure via $(f.\gamma) : m \mapsto f(m.\gamma^{-1}).\gamma$. Thus we identify $Hom(M, N)^{\mathbb{Z}\Gamma}$ with $Hom_{\mathbb{Z}\Gamma}(M, N)$, i.e. the Γ-invariant homomorphisms correspond exactly to the $\mathbb{Z}\Gamma$-module morphisms $M \to N$. If F, C are two chain-complexes, then we declare their tensor product $F \otimes C$ to be a chain complex with dimension n part equal to

$$(F \otimes C)_n = \oplus_{p+q=n} F_p \otimes C_q,$$

and having a differential $D(f \otimes c) = df \otimes c + (-1)^{deg(f)} f \otimes d'c$. When we reduce our coefficients to $\mathbb{Z}/2$, then we forget signs and have $D(F \otimes c) = df \otimes c + f \otimes d'c$.

For a $\mathbb{Z}\Gamma$-module M, the homology groups $\{H_n(\Gamma; M)\}_{n\ geq0}$ with coefficients in M are defined as homology of the chain complex $H_n(F \otimes_{\mathbb{Z}\Gamma} M)$, where $F = \{F_n, \partial\}_n$ is a projective resolution of the $\mathbb{Z}\Gamma$-module $\underline{\mathbb{Z}}$ over $\mathbb{Z}\Gamma$. Here we see $\underline{\mathbb{Z}}$ as the additive integer group with trivial Γ-action, $\gamma.n = n$ for all $n \in \mathbb{Z}$, $\gamma \in G$. The topologists favourite coefficient group $\mathbb{Z}$ or $\mathbb{Z}/2$ are formally defined as trivial $\mathbb{Z}\Gamma$-modules, and denoted $\underline{\mathbb{Z}}$ or $\underline{\mathbb{Z}/2}$ when we wish emphasize the trivial $\mathbb{Z}\Gamma$-structure.

To define cohomology-with-coefficients, let $\{F_n, \partial_n\}$ be a projective resolution of $\mathbb{Z}$ over $\mathbb{Z}\Gamma$, and M the coefficient $\mathbb{Z}\Gamma$-module. There is a cochain-complex $\{Hom_{\mathbb{Z}\Gamma}(F_n, M)\}_n$, with coboundary $\delta = \{\delta_n\}$ defined adjointly by $\delta_n : Hom(F_n, M) \to Hom(F_{n+1}, M)$, $f \mapsto \delta z$, where $\delta z(f) = z(\partial_n f)$ for all homomorphisms $z : F_n \to M$ and $f \in F_n$. The cohomology of this cochain complex defines $H^*(\mathbb{Z}\Gamma; M)$. The cohomology modules $H^m(\Gamma; \mathbb{Z}\Gamma)$ have the following definition. Let $\{P_n, \partial_n\}_n$ be a projective resolution of $\mathbb{Z}$ over $\mathbb{Z}\Gamma$. The cochain complex $Hom_{\mathbb{Z}\Gamma}(P_n, \mathbb{Z}\Gamma)$ with coboundary defined adjointly has cohomology describing $H^*(\Gamma; \mathbb{Z}\Gamma)$.

Now we define homology groups with coefficients in a chain-complex. If $\{C_n, \partial_n\}_n$ is a chain complex, then we set $H_n(\Gamma; C) = H_n(F \otimes_\Gamma C)$, where $F \otimes_\Gamma C$ is the tensor product of the chain complexes F, C, graded appropriately. The homology groups with coefficients in the chain complex C is a chain-homotopy invariant, and hence determined by the homology groups of the chain complex C. If the homology of C concentrates to a single dimension $H_*(C) \simeq H_q(C) =: D$ for some integer $q \geq 0$, then the homology groups $H_n(\Gamma; C)$ reduce to $H_n(\Gamma; D)$.

When X is aspherical topological space supporting a continuous free properly discontinuous action $X \times G \to X$, then the singular chain groups $\{C_n^{sing}(X; \mathbb{Z})\}_n$ are abelian groups possessing a $\mathbb{Z}\Gamma$-module structure $C_n^{sing}(X;\mathbb{Z}) \times \Gamma \to C_n^{sing}(X;\mathbb{Z})$ arising from the geometric action. When Γ acts proper discontinuously on X with quotient X/Γ supporting a finite-equivariant measure, then the action of $\mathbb{Z}\Gamma$ on C_n^{sing} turns the chain groups into finitely-generated $\mathbb{Z}\Gamma$-modules.

An augmentation map $\epsilon : C_0(X) \to \mathbb{Z}$ is defined $\epsilon(v) = +1$ for every 0-cell on X. If we augment the complex by the above augmentation mapping, then we obtain a projective resolution (actually free resolution) of $\mathbb{Z}$ over $\mathbb{Z}\Gamma$. That is, the sequence

$$\cdots \to C_q^{sing}(X) \to C_{q-1}^{sing}(X) \to \cdots \to C_0^{sing}(X) \to \mathbb{Z} \to 0$$

is exact. The resolution is homologically trivial if we ignore the Γ-action. But Γ acts naturally on everything $C_*^{sing}(X) \times \Gamma \to C_*^{sing}(X)$, and the homology of the $\mathbb{Z}\Gamma$-modules become *topological* invariants of Γ. The comparison to the singular homology with coefficients in $\underline{\mathbb{Z}}$ arises from another augmentation mapping $\epsilon_0 := \mathbb{Z}\Gamma \to \underline{\mathbb{Z}}$ defined by

$$\epsilon_0(\sum_\gamma n_\gamma \gamma) = \sum_\gamma n_\gamma.$$

The following lemma is useful for relating the above homology groups to topology, c.f. [Bro82, VIII.7.4, pp.208].

Lemma 6.2.1. *Let M be a $\mathbb{Z}\Gamma$-module. There is natural $\mathbb{Z}\Gamma$-module isomorphism between $Hom_{\mathbb{Z}\Gamma}(M, \mathbb{Z}\Gamma)$ and $Hom_c(M, \mathbb{Z})$, where $Hom_c(M, \mathbb{Z}) \subset Hom(M, \mathbb{Z})$ consists of $\mathbb{Z}$-linear homomorphisms $f : M \to \mathbb{Z}$ satisfying: for every $m \in M$, there exist only finitely many $\gamma \in \Gamma$ for which $f(m.\gamma) \neq 0$. We call $Hom_c(M, \mathbb{Z})$ the module of Γ-compactly-supported homomorphisms.*

Proof. Let $F : M \to \mathbb{Z}\Gamma$ be a $\mathbb{Z}\Gamma$-linear morphism. Then F has the form $F(m) = \Sigma_{\gamma\in\Gamma} f_\gamma(m).\gamma$, where $f_\gamma : M \to \mathbb{Z}$ is a $\mathbb{Z}$-linear morphism for every γ. For fixed m, we see that only finitely many terms $f_\gamma(m)$ are nonzero, since the group ring $\mathbb{Z}\Gamma$ consists of finitely supported $\mathbb{Z}$-distributions over Γ.

As F is $\mathbb{Z}\Gamma$-linear, we have $F(m.\delta) = \delta^{-1}F(m)$, and so $\Sigma f_\gamma(m.\delta).\gamma = \Sigma f_{\delta\gamma}(m)\gamma$ for all m. Thus $f_\gamma(m) = f_{Id}(m.\gamma^{-1})$ for all m, γ, and we conclude that the coefficients f_γ determining $F = \Sigma_\gamma f_\gamma$ are uniquely determined by a particular γ, say, $\gamma = Id \in \Gamma$. The assignment $F \mapsto f_{Id}$ yields the correspondance $Hom_{\mathbb{Z}\Gamma}(M, \mathbb{Z}\Gamma) \to Hom_c(M, \mathbb{Z})$, which can immediately seen to be natural equivariant isomorphism with inverse $f_{Id} \mapsto \Sigma_\gamma f_{Id}(-.\gamma^{-1})$. $\square$

The lemma gives natural equivalence between the cohomology of cochain complexes

$$\{Hom_{\mathbb{Z}\Gamma}(C_n^{sing}(X), \mathbb{Z}\Gamma)\}_n \text{ and } \{Hom_c(C_n^{sing}(X), \mathbb{Z})\}_n.$$

The latter cochain complex is familiar as the compactly-supported cohomology on X consisting of cochains $z : C^{sing}(X) \to \mathbb{Z}$ for which $z(\sigma) = 0$ for almost all cells σ in X, and with only finitely many exceptions. When Γ acts cocompactly on the contractible space X, then compactly-supported cohomology on X can be identified by a Poincaré-Lefschetz duality

$$H_c^m(X) \simeq H_{d-m}(X, \partial X; \mathbb{Z}) \otimes O \simeq H_{d-m-1}(\partial X; \mathbb{Z}) \otimes \Omega,$$

where O is the orientation module on X, i.e. the $\mathbb{Z}\Gamma$-module supported on the abelian group $\{-1, +1\}$ and which measures whether a given element $\gamma \in \Gamma$ preserves the orientation of X or not. See [BS73][§11]. This duality is generalized to finite-volume quotients (usually noncompact) in Bieri-Eckmann's duality. See Section 6.4 below.

6.3 Excision versus Compactification

In the previous Chapter 5, we emphasized excisions $F[t]$ of convex bodies F where the excision parameters were supported on the extreme points $\mathscr{E}[F]$. We generalize this idea further to excisions supported on Γ-rational points at-infinity in the present section. Thus the convex excisions and visibility costs defined earlier have applications to the geometric $E\Gamma$ models.

In applications we find $E\Gamma$ models X arising from nonpositive curvature, namely from Cartan-Hadamard spaces, i.e. finite-dimensional complete nonpositively-curved spaces satisfying the triangle comparison inequalities of Alexandrov, Toponogov, and Cartan. See [BGS85] for basic definitions and compare Definition 6.1.2 from Section 6.1.

For general complete metric spaces Gromov defined a universal compactification, c.f. [BJ06], [BGS85]. In the nonpositive curvature, the compactification has direct geometric interpretation. For every point $x \in X$ in a d-dimensional space, the exponential map $exp_x : T_xX \to X$ determines a homeomorphism from the unit tangent sphere $S^{d-1} \subset T_xX$ to the visual boundary at-infinity $X(\infty)$ of X. Adjoining the visual boundary provides a topological compactification $\overline{X} = X \cup X(\infty)$. The compactification $\overline{X}$ is topologically a large-dimensional closed disk. The visual boundary $X(\infty)$ inherits a natural metric (so-called spherical Tits metric) and supports a uniform Lebesgue measure.

If X is a geometric $E\Gamma$ model, then the isometric action $X \times \Gamma \to X$ extends to a continuous action by homeomorphisms $\overline{X} \times \Gamma \to \overline{X}$. However there is some difficulty: Γ acts by homeomorphisms on $X(\infty)$, and this action is neither free nor proper discontinuous. In fact, following the argument of Thurston [**thurston1988geometry**], we observe that Brouwer's fixed point theorem implies every such continuous homeomorphism has at least one fixed point in $\overline{X}$. Therefore Radon measures on $X(\infty)$ do not descend to Radon measures on the topological quotient $X(\infty)/\Gamma$.

For example, the standard action of $PGL(\mathbb{Z}^2)$ on the Poincaré disk $\mathbb{H}^2$ is proper-discontinuous and virtually free. The boundary-at-infinity of $\mathbb{H}^2$ is a topological circle $\mathbb{H}^2(\infty) \approx S^1$, and it's well-known that $PGL(\mathbb{Z}^2)$ acts ergodically on this circle at-infinity with respect to the uniform Haar measure θ on S^1. Therefore every $PGL(\mathbb{Z}^2)$-equivariant θ measurable function on $\mathbb{H}^2(\infty)$ is constant. This says all equivariant Radon measures on the boundary circle are trivial.

The previous discussion indicates that conventional compactifications are not suitable for our semicoupling method, which is based on transporting Radon measures. Instead we pursue a general excision procedure, implicit in the literature and defined for arbitrary geometric $E\Gamma$ models. The basic idea is this: in a Cartan-Hadamard space $X \times \Gamma \to X$, there are "deep dark zones" which a given orbit $x.\Gamma$ will strongly avoid. These dark zones are Γ-equivariant collections $\{W_\lambda \mid \lambda \in I\}$ of convex horoballs with centres at-infinity and small radii. The dark zones W_λ are disjoint from all Γ-accumulation points at visual infinity. Now we excise, or "scoop-out", these halfspaces, and obtain a manifold-with-corners

$$X_0 := X - \cup_\lambda W_\lambda.$$

The excision X_0 has topological boundary $\partial X_0 \subset X$. The boundary ∂X_0 is naturally cellulated by the boundaries ∂W_λ for $\lambda \in I$. If Γ furthermore translates the halfspaces $\{W_\lambda\}_{\lambda \in I}$ such that $W_\lambda.\gamma \subset W_\lambda$ only if $W_\lambda.\gamma = W_\lambda$, then the excision boundary ∂X_0 is set-theoretically Γ-invariant. In these cases we obtain free and proper-discontinuous actions

$$\partial X_0 \times \Gamma \to \partial X_0, \quad X_0 \times \partial X_0 \times \Gamma \to X_0 \times \partial X_0,$$

where the action is diagonal $(x, y).\gamma = (x.\gamma, y.\gamma)$. Proper-discontinuity ensures ∂X_0 supports nontrivial Γ-equivariant Radon measures.

The excision procedure is summarized in the Figure 6.3. Below we give formal construction, and to be applied to arithmetic groups in §6.5.

Recall that a point λ on the visual sphere $X(\infty)$ can be characterized as the "asymptotic class" of a geodesic ray $s : [0, \infty] \to X$ diverging to some "point" $s(\infty)$ at visual-infinity. Let $\lambda \in X(\infty)$ be point at-infinity. For every choice of $x_0 \in X$, we define the

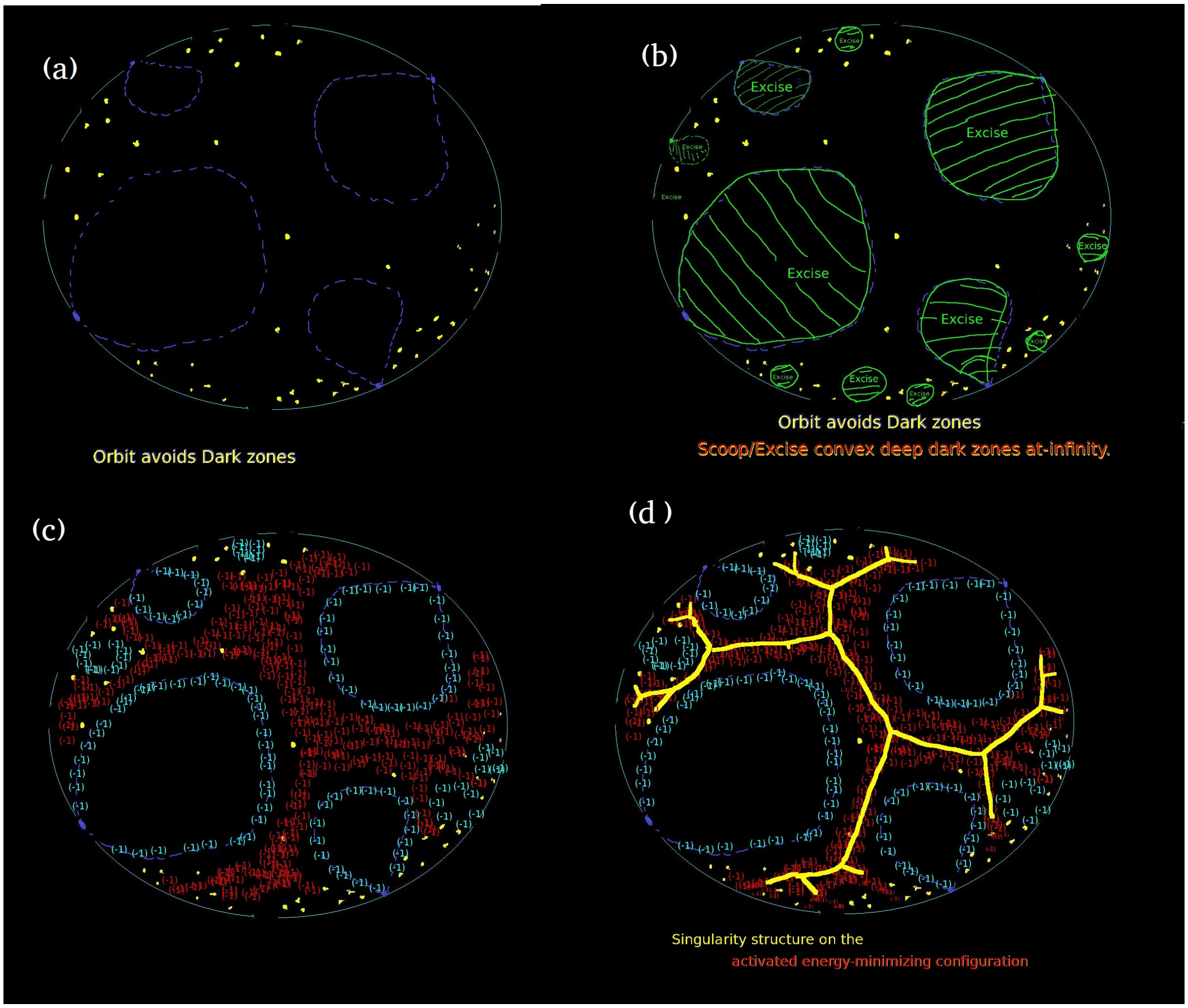

Figure 6.1: (a) Deep Zones are disjoint from Γ-orbit. (b) Excise the Deep Zones from X. (c) Optimal Semicoupling, with respect to repulsion cost, between target excision boundary (in blue) and activated source measure (in red). (d) Singularity structure (in yellow) of an optimal semicoupling between source and target with respect to repulsion cost.

horofunction $h_{\lambda,x_0}: X \to \mathbb{R}$ by the usual formula (c.f. [BGS85, §3]).

Definition 6.3.1 (Horofunctions)**.** Let $s : [0, +\infty] \to X$ be geodesic ray diverging to $s(+\infty) = \lambda$ with $s(0) = x_0$. Then the horofunction centred at λ is defined

$$h_{\lambda,x_0}(x) := \lim_{t \to +\infty} d(x, s(t)) - t.$$

There are several equivalent characterizations of horofunctions on complete Cartan-Hadamard spaces. A continuous function $h : X \to \mathbb{R}$ is a horofunction if and only if h is geodesically convex, h is 1-Lipschitz $|h(x) - h(y)| \leq |x - y|$, and for every $x \in X$ and $r > 0$, there exists two points x_1, x_2 with $d(x, x_1) = d(x, x_2) = 2r$. Equivalently, h is a horofunction on a complete Cartan-Hadamard space (X, d) if and only if h is a geodesically convex C^1 function with $|\nabla_x h| = 1$, c.f. [BGS85, Lemma 3.4].

For any $\lambda \in X(\infty)$, let

$$\Gamma_\lambda := \{\gamma \in \Gamma \mid \lambda.\gamma = \lambda\}$$

be the isotropy-group (i.e. stabilizer group) of the point at infinity. The hypothesis that Γ acts proper discontinuously on X implies every fixed point is necessarily a point at-infinity $\lambda \in X(\infty)$. Notice when γ acts isometrically, then any accumulation point in $X(\infty)$ of the orbit $\{x.\gamma^n \mid n \in \mathbb{Z}\}$ is a fixed point of γ. If γ is isometry fixing $\lambda \in X(\infty)$, then γ maps horoballs centred at λ to horoballs centred at λ. And since γ is distance-preserving between pairs of points, γ also preserves signed-distance between any two horospheres. This suggests the following definition.

Definition 6.3.2. For every $\lambda \in X(\infty)$, let $T_\lambda : \Gamma_\lambda \to \mathbb{R}$ be the group homomorphism defined by the signed-distance between successive γ-translates of the horospheres centred at λ.

So T_λ is defined by the identity $\{h_{\lambda,x} \leq t\}.\gamma = \{h_{\lambda,x} \leq t + T_\lambda(\gamma)\}$ for every $t \in \mathbb{R}$, and where x is an arbitrary basepoint in X. We find $T_\lambda = 0$ is trivial if and only if Γ_λ preserves the horospheres centred at λ setwise, i.e.

$$\{h_{\lambda,x} = t\}.\gamma \subset \{h_{\lambda,x} = t\}.\gamma$$

for every $t \in \mathbb{R}$, $x \in X$. Equivalently Γ_λ preserves the level sets of every horofunction $h_{\lambda,x}$ centred at λ with respect to any $x \in X$.

Definition 6.3.3 (Γ-rationality)**.** A Γ-invariant subset $J \subset X(\infty)$ is Γ-rational if the group homomorphisms $\{T_\lambda : \Gamma_\lambda \to \mathbb{R} \mid \lambda \in J\}$ are simultaneously trivial.

For example, if $\lambda \in X(\infty)$ is an accumulation point of an orbit $x.\Gamma$ in X, then γ (and its powers γ^2, γ^3 etc.) will not preserve horoballs centred at λ, i.e. the group homomorphism $T_\lambda : \Gamma_\lambda \to \mathbb{R}$ will be nontrivial.

The hypothesis that the homomorphisms $T_\lambda : \Gamma_\lambda \to \mathbb{R}$ are trivial for the collection $\{\lambda\}$ of $X(\infty)$ is necessary to ensure the excision boundary is Γ-invariant. In our applications below, this hypothesis is satisfied for every $\mathbb{Q}$-reductive group G whose derived group ${}^0G$ admits no nontrivial $\mathbb{Q}$-rational multiplicative homomorphisms

$$Hom_{/\mathbb{Q}}({}^0G, \mathbb{G}_m) = \{1\}.$$

This implies that horospheres centred at the ends of $\mathbb{Q}$-split tori in the symmetric-space model will be invariant under $G(\mathbb{Z})$ translates.

There is another hypothesis in addition to the triviality of the homomorphisms $\{T_\lambda \mid \lambda \in J\}$ necessary for the topological applications to small-dimensional $E\Gamma$-models. Namely the reduced singular homology of the excision boundary ∂X_0 must be concentrated in a unique dimension and be torsion-free. However because we are excising convex horoballs W (contractible connected subsets of X) the boundary ∂X_0 is homotopic to $\cup_{\lambda \in J} W_\lambda$, which is a union of contractible convex sets. According to Weil's nerve covering theorem the homotopy-type of this union is equal to the nerve of the aspherical covering $\{W_\lambda \mid \lambda \in J\}$. See [BS73, Theorem 8.2.1] for details.

For arithmetic groups $\Gamma := \mathbb{G}(\mathbb{Z})$, the excision is defined by convex horoballs corresponding to $\mathbb{Q}$-rational parabolic subgroups of $\mathbb{G}$, and the nerve of the covering is well-known to be identical to the rational Tits complex, so $\cup_{\lambda \in J} W_\lambda$ will be homotopic to simplicial complex $\mathscr{B}(\mathbb{G}, T)$ constructed in Section 6.5. The Solomon-Tits theorem identifies the simplicial complex $\mathscr{B}(\mathbb{G}, T)$ as homotopic to a countable wedge of spheres $\vee_{i \in I} \mathbb{S}_i^q$ of some dimension q. Evidently the reduced singular homology of ∂X_0 is concentrated in a single dimension and torsion-free, as desired. We elaborate these hypotheses in the §6.4 with Bieri-Eckmann homological duality.

6.4 Bieri-Eckmann Duality

Let the user produce a discrete matrix group Γ. Then by standard constructions we find $E\Gamma$ models X often having an isometric action $X \times \Gamma \to X$, where X is equipped with a complete proper nonpositively curved distance d. But there is further problem obstructing the user's computation of homological invariants of Γ. Namely the apparent space dimension $\dim(X)$ may fail to coincide with the virtual cohomological dimension

$\nu := vcd(\Gamma)$ of Γ. Informally, the "vcd" is the essential dimension at which nontrivial topological invariants of Γ are supported. For instance, a three-dimensional ball appears to have three-space dimensions, but the cohomology of the ball is zero-dimensional since the ball is homotopic to a point. We recommend [Bro82] or [Ser, Proposition 3] for background. By formal arguments, several equivalent characterizations are possible, e.g.

$$vcd(\Gamma) = \max\{ \ n \mid H^n(\Gamma'; M) \neq 0\},$$

for a $\mathbb{Z}\Gamma'$-module M, and where $\Gamma' \leq \Gamma$ is a finite index torsion free subgroup, which exist abundantly acording to Selberg's lemma [Alp87].

One optimistically expects $vcd(\Gamma)$ to coincide with the minimal geometric dimension of an $E\Gamma'$ model X. From the definitions, it is clear $vcd(\Gamma)$ is no greater than any $\dim(X)$. There is famous theorem of Eilenberg-Ganea which proves: if $3 \leq vcd(\Gamma) \leq n$, then there exists an n-dimensional $E\Gamma'$ model with the structure of a simplicial complex. Numerous references are available, e.g. [Bro82, p. VIII.7], [Ser, Proposition 10]. However the proofs of Eilenberg-Ganea's theorem are non-constructive, and abstract cellular inductive processes. Firstly, the proof requires the precise presentation of the group Γ' from which one builds the 2-complex of generators (1-cells) and relations (2-cells attached for every relator). This produces an abstract two-dimensional complex Y^2. Taking the universal cover $X^2 = \tilde{Y^2}$, one finds a simplicial complex whose homology groups vanish in dimensions ≤ 2. If one can identify the nontrivial $H_3(X^2)$ groups, then one may attach 3-cells (using Hurewicz theorem) to systematically annul all the nontrivial three-dimenionsal homology. Thus one obtains a 3-complex Y^3 obtained from X^2 by attaching 3-cells. Taking the universal cover $X^3 := \tilde{Y^3}$, the induction process continues where possibly some four-dimensional homology has arised from the attached 3-cells, which must be annulled by attaching 4-cells, etc.

The construction (as sketched above) is practically impossible to implement. Our thesis provides new general method for displaying the small-dimensional models according to the Reduction-to-Singularity principles of the previous chapters. The above "external" construction is replaced by the explicit reduction of an initial $E\Gamma$ model X to a closed subvariety $Z \subset X$.

Numerous large-dimensional $E\Gamma$ models are available. These models have space dimension much greater than the cohomological dimension. A precise determination of the dimension-gap is achieved in Borel-Serre's formula

$$vcd(G(\mathbb{Z})) = \dim(K/{}^0\mathbb{G}(\mathbb{R})) - rank_{\mathbb{Q}}({}^0\mathbb{G})$$

, c.f. [BS73, §8.6], whenever $\mathbb{G}$ is a $\mathbb{Q}$-reductive linear algebraic group. Their method is very general, c.f. Theorem 6.4.4 below. The argument of Borel-Serre is based on the construction of rational bordification models denoted $\overline{X}^{BS,/\mathbb{Q}}$, and the fact that $\Gamma = \mathbb{G}(\mathbb{Z})$ satisfies a homological duality generalizing Poincaré duality as discovered by Bieri-Eckmann [BE73].

Definition 6.4.1. A finitely generated group Γ is a duality group of dimension $\nu \geq 0$ with respect to a $\mathbb{Z}\Gamma$-module $\mathbf{D}$, if there exists an element $e \in H_\nu(\Gamma; \mathbf{D})$ with the following property: for every $\mathbb{Z}\Gamma$-module A, the "cap-product with e" defines $\mathbb{Z}\Gamma$-module isomorphisms $H^d(\Gamma; A) \approx H_{\nu-d}(\Gamma; A \otimes \mathbf{D})$, $f \mapsto f \cap [e]$.

The basic properties of duality groups are summarized in the following

Proposition 6.4.2 (Bieri-Eckmann duality, [BE73])**.** *Let Γ be duality group of dimension ν, with dualizing module $\boldsymbol{D}$. Then*

(i) we have $\mathbb{Z}\Gamma$-isomorphism $\boldsymbol{D} \approx H^\nu(\Gamma; \mathbb{Z}\Gamma) \neq 0$, so $\boldsymbol{D}$ is a torsion-free additive abelian group;

(ii) the homology group $H_\nu(\Gamma; \boldsymbol{D})$ is infinite cyclic generated by $[e]$ as additive abelian group;

(iii) the group Γ has cohomological dimension $cd(\Gamma)$ equal to ν.

Proof. The statements are direct consequences of duality. (i) We see $H^\nu(\Gamma; \mathbb{Z}\Gamma) \approx H_0(\Gamma; \mathbf{D}) \approx \mathbf{D}$. (ii) Duality implies $H^0(\Gamma; \underline{\mathbb{Z}})$ is isomorphic to $H_\nu(\Gamma; \underline{\mathbb{Z}} \otimes_{\mathbb{Z}\Gamma} \mathbf{D})$, which in turn is canonically isomorphic to $H_\nu(\Gamma; \mathbf{D})$ since $\underline{\mathbb{Z}} \otimes_{\mathbb{Z}\Gamma} \mathbb{Z}\Gamma \approx \mathbf{D}$. But $H^0(\Gamma, \underline{\mathbb{Z}})$ is canonically isomorphic to $\underline{\mathbb{Z}}$. (iii) The duality isomorphism implies for every $\mathbb{Z}\Gamma$-module A that $H^*(\Gamma; A)$ is isomorphic to $H_{\nu-*}(\Gamma; A \otimes \mathbf{D})$ which reduces to 0 whenever $\nu - * < 0$. □

Dualizing modules $\mathbf{D}$, which are unique up to $\mathbb{Z}\Gamma$-isomorphism, can be constructed for various groups Γ arising in practice. Whereas the Bieri-Eckmann duality produces canonical $\mathbb{Z}\Gamma$-isomorphism between $\mathbf{D}$ and the cohomology group $H^\nu(\Gamma; \mathbb{Z}\Gamma)$ supported at the cohomological dimension ν, this cohomological presentation of $\mathbf{D}$ is generally insufficient. We emphasize excisions X_0 whose topological boundary ∂X_0 produces homological (i.e. projective) resolutions of the dualizing module $\mathbf{D}$. This is better suited for constructing homology with coefficients in the dualizing module $\mathbf{D}$, as arising in the statement of Bieri-Eckmann duality above. In our chapter on Closing the Steinberg symbol, we effectively construct nontrivial cycles $\xi \in H_0(\Gamma; \mathbb{Z}_2\Gamma \otimes \mathbf{D})$ using the homological resolution above.

Theorem 6.4.3 ([BE73], [BS73])**.** *Let $(X[t], \partial X[t])$ be a Γ-equivariant rational excision model, where $X[t], \partial X[t]$ support invariant Radon measures σ, τ having finite Γ-covolume.*

Suppose there exists an integer $q \geq 0$ such that the reduced homology $\tilde{H}_*(\partial X; \mathbb{Z})$ of the topological boundary is concentrated at dimension $* = q$,

$$\tilde{H}_*(\partial \overline{X}^{BS/\mathbb{Q}}; \mathbb{Z}) = \begin{cases} 0, & \text{if } * \neq q, \\ \boldsymbol{D}, & \text{if } * = q. \end{cases} \tag{6.1}$$

where $\boldsymbol{D}$ is a nonzero torsion-free additive abelian group. Then for every finite-index torsion-free subgroup $\Gamma' < \Gamma$, the associated $\mathbb{Z}\Gamma'$-module $\boldsymbol{D'} := \boldsymbol{D} \otimes_{\mathbb{Z}} \mathbb{Z}\Gamma'$ is a dualizing-module for Γ' of dimension $\nu := \dim(X[t]) - q + 1$. So

$$H^\nu(\Gamma'; \mathbb{Z}\Gamma') \simeq H_0(\Gamma'; \boldsymbol{D'}) \approx \boldsymbol{D'}$$

is nonzero and $vcd[\Gamma] = cd[\Gamma'] = \nu$.

Proof. We refer the reader to [BE73, §6.2] and [BS73, §8.4, §11] for details. □

Consequently Borel-Serre's formula can be restated as follows:

Theorem 6.4.4 (Borel-Serre)**.** *Let Γ be a discrete infinite group with finite virtual cohomological dimension $vcd(\Gamma) < +\infty$ and satisfying Bieri-Eckmann homological duality. Suppose X is a Cartan-Hadamard manifold and $E\Gamma$ model such that X has finite covolume modulo Γ. Let q equal the spherical-dimension of Bieri-Eckmann's dualizing module. Then we have*

$$vcd(\Gamma) = dim(X) - (q + 1).$$

6.5 Arithmetic Groups: Excision

The subject of linear algebraic matrix groups and their arithmetic groups is extensive topic. In this section we describe the basic excision construction which enables the applications of our semicoupling methods to small-dimensional $E\Gamma$ classifying spaces, for $\Gamma := \mathbb{G}(\mathbb{Z})$ an arithmetic group. The basic examples are the arithmetic subgroups of the standard higher-rank $\mathbb{Q}$-reductive groups $\mathbb{G} = GL(V)$, e.g. the symplectic groups $Sp(\mathbb{R}^4, \omega)$, $Sp(\mathbb{R}^6, \omega)$, ..., and the split-orthogonal groups $O(V^{p,q})$ for $p, q \geq 2$.

Constructing small-dimensional classifying spaces is an old topic, originating in Minkowski's "geometry of numbers". The classic example is the reduction of the hyperbolic disk onto the so-called Farey tree, c.f. [Bro82, VIII.9, pp.215]. The discrepancy between space- and algebraic-dimensions was made precise in Borel-Serre's investigations [BS73], wherein the relation to Bieri-Eckmann duality was first discovered (summarized in 6.4.4

from §6.4). A general method for reducing the general linear group $GL(\mathbb{Z}^N)$ was discovered by the so-called "systolic well-rounded retract" introduced in [Sou78] and extended in [Ash84]. For instance, Soulé's method produces an interesting three-dimensional cube model for the codimension-two reduction of the five-dimensional symmetric space $\approx SO(3)\backslash PGL(\mathbb{R}^3)/PGL(\mathbb{Z}^3)$. But no general principle for constructing large codimension (≥ 2) deformation retracts was available, with the exception of $PGL(\mathbb{Z}^N)$. We propose the homotopy-reductions constructed in this thesis are effective generalizations of Ash-Soulé 's well-rounded retract.

So let the user choose an arithmetic group $\Gamma = \mathbb{G}(\mathbb{Z})$. With respect to our Reduction Program 1.5.1, first we construct excision models from an initial geometric $E\Gamma$ model. This key first step originates with [BS73]. The excision is obtained from the initial $E\Gamma$ model $X = K\backslash {}^0\mathbb{G}(\mathbb{R})^0$, and defined with respect to a choice of basepoint $[K]$ on X and equivariant excision parameters $t : \Phi_{co}^{\Gamma} \to \mathbb{R}$ defined below. The equivariant excision model $X[t] \times \partial X[t]$ has Γ-finite invariant Radon measure $\sigma \otimes \tau = vol_{X[t]} \otimes vol_{\partial X[t]}$, where Γ acts diagonally

$$X[t] \times \partial X[t] \times \Gamma \to X[t] \times \partial X[t], \quad (x, y).\gamma = (x\gamma, y.\gamma).$$

The main topological properties of the excision are summarized in Theorem 6.5.2, and surely well-known to the experts.

Our presentation of the excision models is generally described for discrete subgroups Γ of $\mathbb{Q}$-reductive linear algebraic matrix groups $\mathbb{G}$. The generality of the construction forces us to speak in terms of the Bruhat-Tits structure theory of $\mathbb{G}$. We assume $\mathbb{G}$ is totally split over $\mathbb{Q}$, and so maximal $\mathbb{Q}$-algebraic tori T in $\mathbb{G}$ are totally $\mathbb{Q}$-split and admit $\mathbb{Q}$-rational isomorphisms $T \approx \prod_{rank_{\mathbb{Q}}\mathbb{G}} \mathbb{G}_m$ onto a product of multiplicative groups. The excision construction involves the Γ-orbit of the set of $\mathbb{Q}$-coroots Φ^{co}, where $\Phi = \Phi(\mathbb{G}, T) \subset Hom_{/\mathbb{Q}}(T, \mathbb{G}_m)$ is the root system of $\mathbb{G}$ with respect to a maximally $\mathbb{Q}$-split algebraic torus. C.f. [BJ06] or [BT65] for terminology. The root system Φ is the conventional Bruhat-Tits-type Lie algebraic root system, and consists of the "eigenvalues" of the linear representation $T \to GL(\mathfrak{g}_{\mathbb{R}})$ where $\mathfrak{g}$ is the Lie algebra of $\mathbb{G}$.

If we fix a maximal connected compact subgroup K^0 in ${}^0\mathbb{G}(\mathbb{R})^0$, then Matsuomoto lemma ([BT65, §14]) allows us to represent elements of ${}_{\mathbb{Q}}W$ as orientation-preserving isometries of K^0. The choice of K^0 determines a ${}_{\mathbb{Q}}W$-invariant inner product on the Lie algebra of T, which is canonically diffeomorphic to a $rank_{\mathbb{Q}}({}^0\mathbb{G})$-dimensional Euclidean space. If $\Phi = \Phi(G, T) \subset Hom_{/\mathbb{Q}}(T, \mathbb{G}_m)$ denotes the set of roots of the adjoint action of T on the lie algebra of G, then the choice of ${}_{\mathbb{Q}}W$-invariant inner product allows us to represent the dual collection of so-called coroots. The coroots are a finite set of nonzero cocharacters

$$\Phi_{co} \subset Hom_{/\mathbb{Q}}(\mathbb{G}_m, T) \subset Hom_{/\mathbb{Q}}(\mathbb{G}_m, G).$$

Here $P(k^\gamma)$ corresponds to the maximal parabolic subgroup defined by the cocharacter k^γ, and $q.\gamma$ corresponds to the conjugate K^γ of the maximal compact, and $A_{P(k^\gamma),q.\gamma}$ corresponds to the split-central torus in $(P(k^\gamma), K^\gamma)$ Levi-Langlands coordinates.

Definition 6.5.1. Fix a basepoint $q = K$ in $K\backslash{}^0\mathbb{G}(\mathbb{R})^0$, and let $\Phi_{co} \subset Hom_{/\mathbb{Q}}(\mathbb{G}_m, T)$ be the coroots with respect to a maximal $\mathbb{Q}$-split selfadjoint torus T.

(i) An excision parameter is a Γ-equivariant function $t : \Phi_{co}^\Gamma \to \mathbb{R}_{>0}$.

(ii) For $k^\gamma \in \Phi_{co}^\Gamma$, let $W_{k^\gamma}^t$ be the convex horosphere consisting of all matrix elements $g \in \mathbb{G}(\mathbb{R})^0$ for which the $A_{P(k^\gamma),q.\gamma}$-coordinate is less than or equal to the scalar $exp(t(k^\gamma))$.

The excision model with respect to the basepoint q and excision parameter $t : \Phi_{co}^\Gamma \to \mathbb{R}_{>0}$ is then defined $X[t] := X - \cup_{\Phi_{co}^\Gamma} W_{k^\gamma}^t$.

Theorem 6.5.2. *Let $\mathbb{G}$ be the $\mathbb{Q}$-split form of a semisimple $\mathbb{Q}$-linear algebraic group for which $Hom_{/\mathbb{Q}}(\mathbb{G}, \mathbb{G}_m)$ is trivial. Let $\Gamma := \mathbb{G}(\mathbb{Z})$ be arithmetic group. Suppose $t : \Phi_{co}^\Gamma \to \mathbb{R}_{>0}$ is Γ-equivariant excision parameter. Then*

(i) the excision boundary $\partial X[t]$ is Γ-equivariant;

(ii) the uniform homogeneous measures σ, τ defined on $X[t]$, $\partial X[t]$ have finite volume Γ-quotients (actually the quotients are compact).

(iii) the excision $(X[t], \partial X[t])$ is diffeomorphic as manifold-with-corners to Borel-Serre's bordification

$$X[t] \times \partial X[t] \approx \overline{X}^{BS/\mathbb{Q}} \times \partial \overline{X}^{BS/\mathbb{Q}}.$$

(iv) the boundary $\partial X[t]$ has concentrated reduced homology nontrivial at dimension $q := rank_{\mathbb{Q}}(\mathbb{G}) - 1$, and $\tilde{H}_q(\partial X[t]; \mathbb{Z})$ considered as a $\mathbb{Z}\Gamma$-module is the Bieri-Eckmann dualizing module for Γ.

Proof. The hypothesis that $Hom_{/\mathbb{Q}}(\mathbb{G}, \mathbb{G}_m)$ is trivial descends by induction to all the semisimple parts ${}^0L$ of the various Levi factors $L = L_{P,q.\gamma}$ of $\mathbb{Q}$-parabolic subgroups P with respect to the Γ-orbits of the basepoint q. Together with the principle of "no accidental parabolics", we find the excision boundary is necessarily Γ-invariant subset of X. This proves (i). Item (ii) follows from standard argument of Borel-Harish-Chandra, see [BS73]. Rescaling the excision parameter t to 0^+ produces desired diffeomorphism $X[t] \approx X[0^+] = \overline{X}^{BS/\mathbb{Q}}$. This proves (iii). The collection of convex horospheres $\{W_k^t\}_k$ produces a covering of $\partial X[t]$ by contractible open sets whose nerve is isotopic to the spherical Tits building $\mathscr{B}(\mathbb{G}, \mathbb{Q})$. By Weil's nerve theorem, there is natural homotopy-isomorphism between $\partial X[t]$ and $\mathscr{B}(\mathbb{G}, \mathbb{Q})$. But the well-known Solomon-Tits theorem proves $\mathscr{B}(\mathbb{G}, \mathbb{Q})$ has the homotopy-type of a countable wedge of $(rank_{\mathbb{Q}}(\mathbb{G})-1)$-dimensional spheres. This proves (iv). □

Remark. The equivariant excision parameter $t : \Phi_{co}^{\Gamma} \to \mathbb{R}_{>0}$ is determined by its restriction to the initial coroot set Ψ_{co}. In practice we look to define t as symmetrically as possible, especially with respect to the natural action of the Weyl group ${}_{\mathbb{Q}}W$. However the roots of Ψ are not pairwise symmetric, since indeed ${}_{\mathbb{Q}}W$ does not act transitively on Ψ. E.g., the root system C_2 corresponding to the real split symplectic group $Sp(\mathbb{R}^4, \omega)$ is not totally regular, having roots of different lengths. We can either appeal to Minkowski's theorem [Ale06] to prescribe an excision parameter (unique modulo homothety) for which the codimension-one faces of $B \cap X[t]$ have given measures, or we can be satisfied with ${}_{\mathbb{Q}}W$-symmetry of the restricted excision parameter $t| : \Phi_{co} \to \mathbb{R}_{>0}$.

Chapter 7

Closing the Steinberg symbol

7.1 Stitching Footballs from Regular Panels: Motivation

This chapter introduces a subprogram we call ‘Closing the Steinberg symbol’. But before developing the formal definitions in Section 7.2 below, we offer some informal motivations. In low dimensions, our ideas relate to the problem of stitching footballs from uniform hexagonal panels, or uniform pentagonal panels, or combinations of both as in the Figures below. To stitch a football from panels $\{P_i \mid i \in I\}$ means finding a finite subset $I' \subset I$ for which the singular chain sum $\sum_{i\in I'} P_i$ has singular chain boundary which vanishes mod 2, so

$$\partial(\sum_{i\in I'} P_i) = \sum_{i\in I'} \partial P_i = 0$$

over $\mathbb{Z}/2$-coefficients. When P is two-dimensional hexagon or pentagon, the panels have singular boundary

$$\partial P = \sum_{e \text{ edge of } P} e.$$

We denote the closed convex hull of the football $F := conv\{P \mid \text{ panels}\}$. The panels then become closed subsets of the boundary ∂F.

For instance, since the 1960’s the standard football is stitched after Adidas’ “Telstar” design, having twenty white hexagon panels and twelve black pentagon panels. But in our applications we assume the patches $\{P_i\}_{i\in I}$ are pairwise isometric to some regular geodesically-flat polygon P. In its most elementary form, Closing the Steinberg symbol is the problem of assembling isometric translates of a fixed two-dimensional equilateral triangle into some two-dimensional sphere. Or, assembling isometric copies of some right-

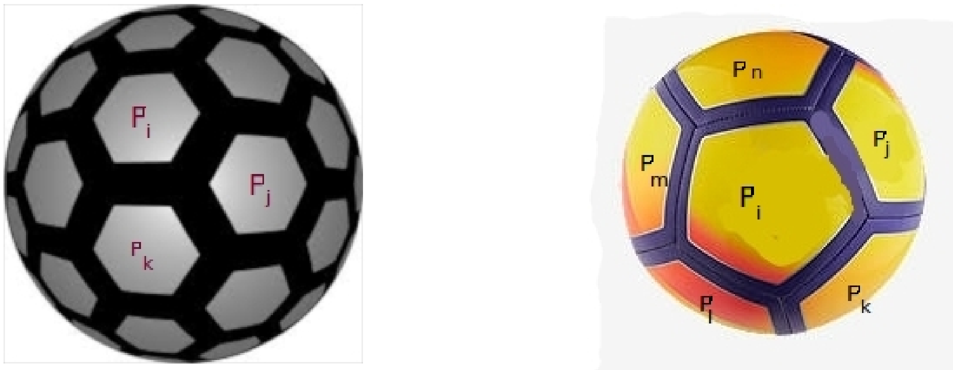

Figure 7.1: Stitching a football F from identical regular hexagons or pentagons P_i

Figure 7.2: Adidas' "Telstar" design is football stitched from white hexagon and black pentagon panels. A football can also be stitched from black triangular and white pentagonal panels.

angled cube into a three-dimensional sphere. Compare Figures 7.1–7.1.

We furthermore assume there is an isometric action by a discrete symmetry group Γ translating the polygon patches $P.\gamma$ for $\gamma \in \Gamma$. The Γ-symmetries lead to chain sums $\underline{F} := \sum_{\gamma \in \Gamma} F.\gamma$ of "footballs through space". We say the chain sum has "well-separated gates" if a pair of footballs $F.\gamma, F.\gamma'$ are either disjoint, identical, or intersect along a single panel P'. The support of a convex chain sum can have nontrivial topology, i.e. depending on the homotopy-type of the chain sum combinatorics. Since panels are contractible, a standard Mayer-Vietoris covering argument identifies the homotopy-type of the support of $\underline{F}$ with the nerve of the covering defined by the chain summands. We detail these ideas further in the sections below.

7.2 Closing Steinberg: Definition and Consequence

Let $X \times \Gamma \to X$ be a geometric $E\Gamma$ model (Definition 6.1.2). Suppose we define a Γ-equivariant family of convex horospheres $\{W^t_\lambda\}_\lambda$, producing an excision model

$$X[t] := X - \cup_\lambda W^t_\lambda$$

whose topological boundary $\partial X[t]$ is Γ-invariant and has reduced singular homology concentrated at some unique dimension $q \geq 0$, satisfying

$$\tilde{H}_q(\partial X[t]; \mathbb{Z}) = \mathbf{D} \neq 0, \quad \text{and } \tilde{H}_*(\partial X[t]; \mathbb{Z}) = 0 \text{ when } * \neq q. \tag{7.1}$$

In applications, the nonzero $\mathbb{Z}$-module $\mathbf{D}$ will be torsion-free. The symmetry action of Γ on X induces a natural $\mathbb{Z}\Gamma$-module structure on $\mathbf{D}$, and we view $\mathbf{D}$ as $\mathbb{Z}\Gamma$-module in the following.

The boundary operator defines a linear map between chain groups

$$\partial : C_{q+1}(X[t], \partial X[t]; \mathbb{Z}/2) \to C_q(\partial X[t]; \mathbb{Z}/2). \tag{7.2}$$

Consider a relative cycle $[P] \in H_{q+1}(X[t], \partial X[t]; \mathbb{Z})$, with $[P] \neq 0$. The long exact sequence of relative homology produces an isomorphism

$$H(\partial) : H_{q+1}(X[t], \partial X[t]; \mathbb{Z}) \simeq H_q(\partial X[t]; \mathbb{Z}), \tag{7.3}$$

so the boundary $\partial[P]$ represents a nontrivial cycle in $H_q(\partial X[t]; \mathbb{Z})$. The group Γ of symmetries flips, rotates, and translates the base cycle $[P]$ throughout the space, and every finite subset I of Γ produces a finite chain sum

$$\sum_{\gamma \in I} [P].\gamma,$$

with total chain boundary

$$\partial(\sum_{\gamma \in I} [P].\gamma) = \sum_{\gamma \in I} \partial[P].\gamma.$$

The basic problem of Closing Steinberg is to produce a finite subset $I \subset \Gamma$ for which the boundary of the nontrivial chain sum vanishes in the mod 2 homology group. Basically we seek nontrivial vectors in the kernel of the boundary operator $H(\partial)$. Formally we seek $I \subset \Gamma$ which defines nonzero elements

$$\xi = \sum_{\gamma \in I} P.\gamma \in C_{q+1}(X[t], \partial X[t]; \mathbb{Z}/2)$$

which are solutions to the following equation:

$$H(\partial)(\sum_{\gamma \in I} P.\gamma) = \sum_{\gamma \in I} H(\partial)(P).\gamma = 0 \quad (mod\ 2)$$

in the homology group $H_q(\partial X[t]; \mathbb{Z}/2)$, where $H(\partial)$ is defined as (7.3).

The complete definition of Closing Steinberg includes further geometric conditions on the Γ-translates $F.\Gamma$ of the the closed convex hull $F = conv[P.I]$ of the translates $B.I$. Let $X[t], \partial X[t]$ be a Γ-invariant excision model. Let $[P]$ be a flat-filled relative cycle representing a nonzero generator of $H_{q+1}(X[t], \partial X[t]; \mathbb{Z})$.

Definition: Closing Steinberg 7.2.1. *A finite subset I of Γ successfully Closes Steinberg if:*

-***(nontrivial mod** 2**)*** *the chain $\xi = \sum_{\gamma \in I} P.\gamma$ is nonvanishing over $\mathbb{Z}/2$ coefficients in the chain group $C_{q+1}(X[t], \partial X[t]; \mathbb{Z}/2)$;*

-***(vanishing boundary mod** 2**)*** *the boundary $\partial \xi = \sum_{\gamma \in I} \partial [P].\gamma$ vanishes over $\mathbb{Z}/2$-coefficients in the homology group $[\partial \xi] = 0$ in $H_q(\partial X[t]; \mathbb{Z})$;*

-***(well-defined convex hull)*** *the boundary-chain representing $\partial \xi$ is simultaneously visible from an interior point x in $X[t]$;*

-***(well-separated gates)*** *there exists a finite-index subgroup $\Gamma' < \Gamma$ such that the chain sum $\underline{F} = \sum_{\gamma \in \Gamma'} F.\gamma$ has nonempty well-separated gates structure precisely equal to the principal orbit $\{P.\gamma \mid \gamma \in \Gamma'\}$.*

In the above setting ∂P is coincident to $P \cap \partial X[t]$. The last hypothesis on well-separated gates means a pair of translates $F, F.\gamma$ are either disjoint or the gate $F \cap F.\gamma$ coincides with some translate $P.\gamma'$.

Our presentation (7.1) of $\mathbf{D}$ as the reduced homology of the boundary with $\mathbb{Z}\Gamma$-module structure means we can naturally view $\mathbf{D}$ as a chain complex. This chain complex is afforded by the singular chain groups on $\partial X[t]$. Homology groups with coefficients in a chain complex means we can interpret the 0th chain group $C_0(\partial X[t]; \mathbb{Z}/2) \otimes \mathbf{D}$ with qth chain group $C_q(\partial X[t]; \mathbb{Z}/2)$, and this interpretation has important consequences. Indeed the existence of chain sums ξ satisfying the first two hypotheses in Closing Steinberg is a consequence of homology.

Proposition 7.2.2. *Let Γ be a Bieri-Eckmann duality group, with dualizing module $\mathbf{D}$. Then there exists finite subsets I in Γ for which $\xi = \sum_{\gamma \in I} P.\gamma$ lies in the kernel of ∂_0 over $\mathbb{Z}/2$.*

Proof. The argument is homological. We interpret ξ as a chain sum representing a 0-cycle in $H_0(\Gamma; \mathbb{Z}/2\Gamma \otimes_{\mathbb{Z}\Gamma} \mathbf{D})$. The hypotheses of Closing Steinberg imply ξ is homologically nontrivial cycle. Bieri-Eckmann duality (Proposition 6.4.2) implies the kernel $\ker \partial_0$ is naturally isomorphic to the induced $\mathbb{Z}\Gamma$-module $\mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbf{D}$ which is nonzero. □

Our definition of Closing Steinberg was inspired by the author's study of [Cre84]. Cremona successfully Closes Steinberg in several cases for $\Gamma = GL(\mathscr{O}_{\sqrt{-d}})$, where $\mathscr{O}_{\sqrt{-d}}$ is the ring of integers of some Euclidean complex quadratic fields. In Cremona's terminology, the problem is to determine a "relation ideal $\mathscr{R}$" and construct a "basic polyhedron P whose transforms fill the space", c.f.[Cre84, pp.290]. We present some further examples in §§7.3–7.4 below.

Now suppose we find a finite subset I in Γ with $\partial_0[\sum_{\gamma \in I} P.\gamma] = 0$. Such subsets exist by Proposition 7.2.2. These subsets partially close the Steinberg symbol, except the orbit $P.I$ may not admit a simultaneously visible interior point and the gates of the chain sum $\underline{F} = SUM[F.\Gamma]$ may not be well-separated.

Our hypotheses regarding Closing Steinberg have useful consequences, which we summarize in the following theorem.

Theorem 7.2.3. *Suppose $I \subset \Gamma$ successfully Closes Steinberg (Definition 7.2.1). Define $F := conv[P.I]$ and $\underline{F} = \sum_{\gamma \in \Gamma} F.\gamma$. Then*

(i) the Γ-translates $F.\gamma$, $\gamma \in \Gamma$, form a chain sum

$$\underline{F} := \cdots [F]\gamma + [F]\gamma' + [F]\gamma'' + \cdots ,$$

and there exists finite-index subgroup $\Gamma' < \Gamma$ which acts as additive shift-operator on the summands of $\underline{F}$; and

(ii) the support of the chain sum $\underline{F}$ is a simply-connected subset of X, and $\underline{F}$ is a cubical $E\Gamma'$ model.

Proof. We can replace Γ with a finite-index torsion-free subgroup Γ' to ensure Γ' acts freely on X, and therefore $X[t], \partial X[t]$. Moreover we can ensure Γ' translates the flat-filled relative cycle $[P].\gamma$, for $\gamma \in \Gamma'$ freely. Then $[P].\gamma \neq [P]$ when $\gamma \neq id$. The definition of Closing Steinberg implies distinct translates F, F' are disjoint unless they intersect in a gate $G' = P.\gamma'$ for some $\gamma' \in \Gamma'$. So $F.\gamma$ equals F only if $\gamma = Id$ is trivial. This proves the summands $\{F.\gamma \mid \gamma \in \Gamma'\}$ of $\underline{F}$ form a principal Γ'-set, and establishes (i). The existence of an interior point $x \in F$ which is simultaneously visible to the translates $P.I$ in $X[t]$ proves $F = conv[P.I]$ is a compact convex set, and homeomorphic to some cube. Thus $\underline{F}$ is a chain sum of cubes, hence a cubical chain sum and therefore (ii). $\square$

7.3 Closing Steinberg on $PGL(\mathbb{Z}^2)$: First Example

Here we provide basic "proof-of-concept" by successfully Closing Steinberg with the finite subset I_0 defined below.

(Step 1: Construct Excision Model) Consider the Voronoi state model of 2-dimensional real states $Q \times PGL(\mathbb{Z}^2) \to Q$. The standard self-adjoint torus

$$A(s) := \{\begin{pmatrix} e^{-s} & 0 \\ 0 & e^{s} \end{pmatrix}\}$$

produces an orbit $q.A$ in $q.PGL(\mathbb{R}^2) \hookrightarrow Q$. The orbit $q.A$ has projective ends at $A(-\infty) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and $A(+\infty) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ on $Proj[\mathbb{R}^2]$. Within the excision model $Q[t]$ these orbits are truncated at

$$A[t] := A \cap Q[t] = \{\begin{pmatrix} e^{-s} & 0 \\ 0 & e^{s} \end{pmatrix} | -t \leq s \leq t\}.$$

Then

$$\partial A[t] = A \cap \partial Q[t] = \begin{bmatrix} e^{-t} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ e^{+t} \end{bmatrix}.$$

In the renormalized limits ([Fur76]) we get $\begin{bmatrix} e^{-s} \\ 0 \end{bmatrix} \to \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ e^{+s} \end{bmatrix} \to \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Compare with Figures 7.3, 7.3.

(Step 2: Close Steinberg. Obtain Cubical Model) Next we view $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as formal tensor element. And define a $\mathbb{Z}/2$-boundary operator

$$\partial_0([u] \otimes [v]) := [u] + [v],$$

which we view as valued in a boundary chain group $C_0(\partial Q[t]; \mathbb{Z}/2)$. The following subset

I_0 successfully closes the $PGL(\mathbb{Z}^2)$ Steinberg symbol $\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$:

$$I_0 = \{Id, \begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\-1 & 2\end{pmatrix}\}.$$

Observe that $\begin{pmatrix}0 & 1\\-1 & 2\end{pmatrix} \in U_{\begin{bmatrix}1\\1\end{bmatrix}}$.

Thus the chain sum

$$\xi := (\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}) + (\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\1\end{bmatrix}) + (\begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix})$$

represents a nontrivial 0-cycle $\xi \in H_0(PGL(\mathbb{Z}^2), \mathbb{Z}_2 \otimes \mathbf{D})$, where $\mathbf{D}$ is the dualizing $\mathbb{Z}PGL(\mathbb{Z}^2)$-module $\approx H_1(Proj[\partial Q[t]]; \mathbb{Z})$. Then $\partial_0 \xi = 0$. Compare Figure 7.3.

Observe that ξ corresponds to the image of $\begin{bmatrix}1 & 0 & 1\\0 & 1 & 1\end{bmatrix}$ under the boundary mapping

$$\partial_1(\begin{bmatrix}1 & 0 & 1\\0 & 1 & 1\end{bmatrix}) = \begin{bmatrix}0 & 1\\1 & 1\end{bmatrix} + \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} + \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}.$$

Here ∂_1 is the right-most differential ∂ in the resolution of [LS76, Theorem 3.1, pp.21]. Notice the symbol $\begin{bmatrix}1 & 0 & 1\\0 & 1 & 1\end{bmatrix}$ has all nonvanishing 2×2-minors, and thus represents nonzero element in the standard resolutions of the Steinberg module, e.g. [LS76, §4]. Thus

$$\partial_0 \xi = \partial_0 \circ \partial_1(\begin{bmatrix}1 & 0 & 1\\0 & 1 & 1\end{bmatrix}) = 0.$$

C.f. [Ste07, §A.5.2, pp.233], [AR79, §§2-5]. The convex hull in Q

$$\{\alpha\langle -, e\rangle^2 + \beta\langle -, f\rangle^2 + \gamma\langle -, e+f\rangle^2 | \alpha, \beta, \gamma \geq 0, \alpha + \beta + \gamma = 1\}$$

of the rank-one states associated to $e, f, e+f$ has barycentre at the hexagonal lattice $x^2 + xy + y^2$. Every quadratic state $p : \mathbb{R}^2 \to \mathbb{R}$ admits a $PGL(\mathbb{Z}^2)$-translate $p.\gamma$ which occupies $\underline{F}_0$. Taking the convex hull $F(I_0)$, and the global chain sum $\underline{F}_0 = \sum_{\gamma \in (\mathbb{Z}^2)} F(I_0).\gamma$, we recover the cubical model $\underline{F}_0 \times PGL(\mathbb{Z}^2) \to \underline{F}_0$.

(Step 3: Install repulsion costs. Construct Kantorovich Singularity) So we replace Q

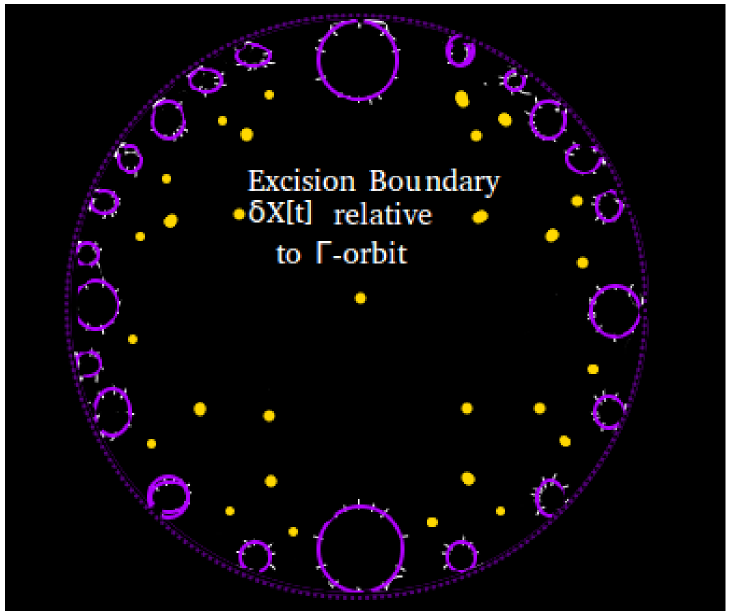

Figure 7.3: Convex excision $X[t]$ relative a Γ-equivariant excision parameter $t : \mathbb{Q}P^1 \to \mathbb{R}$

with an excision model $Q[t]$, and then a cubical chain sum $\underline{F}_0$. The boundary $\partial[F]$ of the chain summands F of $\underline{F}_0$ now coincide with the topological boundary $\partial[\underline{F}_0]$. Having constructed this chain sum, we are now ready to install $v : \underline{F}_0 \times \partial[\underline{F}_0] \to \mathbb{R} \cup \{+\infty\}$, and can proceed to studying the v-optimal semicoupling program between source σ on $X = \underline{F}_0$ and target τ on $Y = \partial\underline{F}_0$. Compare Figure 7.3

Suppose we compute the dual v-concave potential $\psi : \partial[\underline{F}_0] \to \mathbb{R} \cup \{-\infty\}$. Then we need verify Halfspace conditions to ensure the activated source is deformation retract of the initial excision $X[t]$, and finally we need ensure Halfspace conditions satisfied throughout the activated source to deformation retract $X[t] \rightsquigarrow Z_1 \rightsquigarrow Z_2$. These are the remaining Steps of Theorem 1.5.1 from Section 1.5. See Figure 7.3, 7.3

7.4 Closing Steinberg on $GL(\mathbb{Z}^3)$: Example

Recall [LS76] and [AGM]: to Close Steinberg means constructing a syzygy of the $\mathbb{Z}GL(\mathbb{Z}^3)$-module resolution:

$$\cdots \to C_1 \to C_0 \to \mathbf{D} \to 0,$$

of the Steinberg module $\mathbf{D} \approx St_3(\mathbb{Q})$ in (any) of the available resolutions in [LS76, §3], or [AGM, §§2-5]. In otherwords the Steinberg symbol $[B]$ is an element of C_0 which maps to a generator of $\mathbf{D}$. To Close Steinberg means finding Γ-translates $[B'] = [B].\gamma$

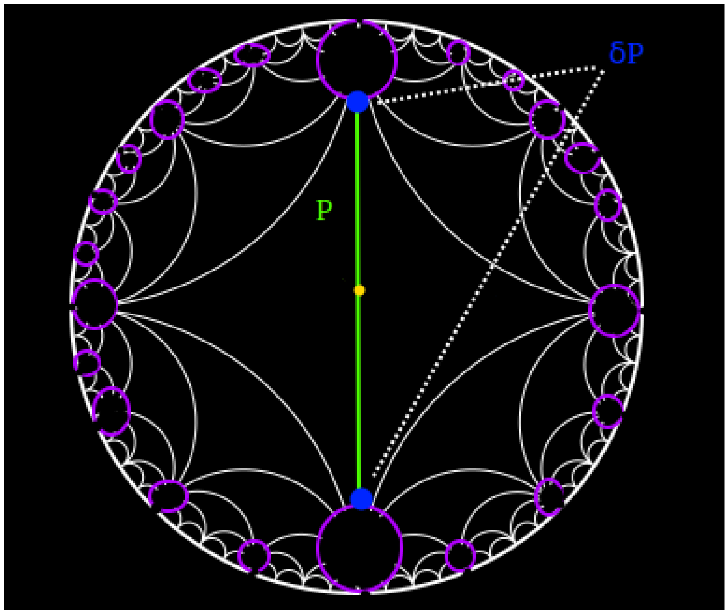

Figure 7.4: Steinberg symbol in $X[t] \times \partial X[t]$ is represented as relative 1-cycle P with boundary ∂P equal to 0-sphere.

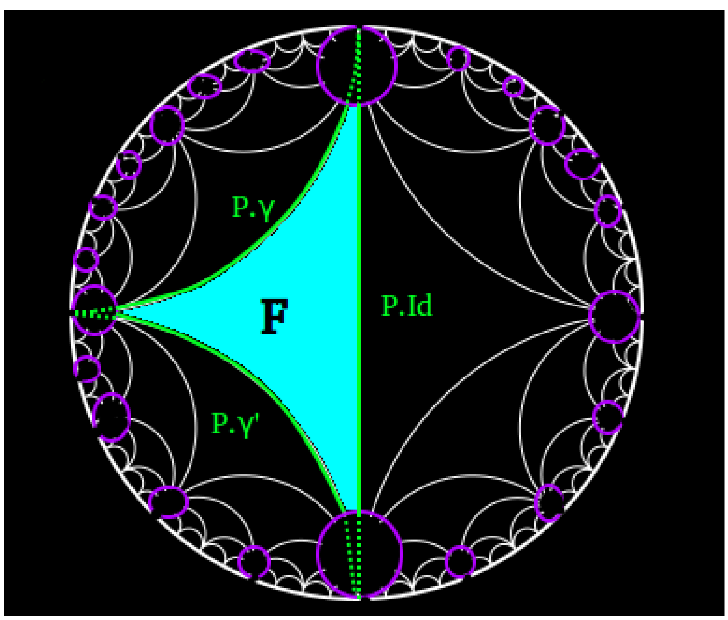

Figure 7.5: The translates by $I_0 = \{Id, \gamma, \gamma'\}$ successfully Close the Steinberg symbol.

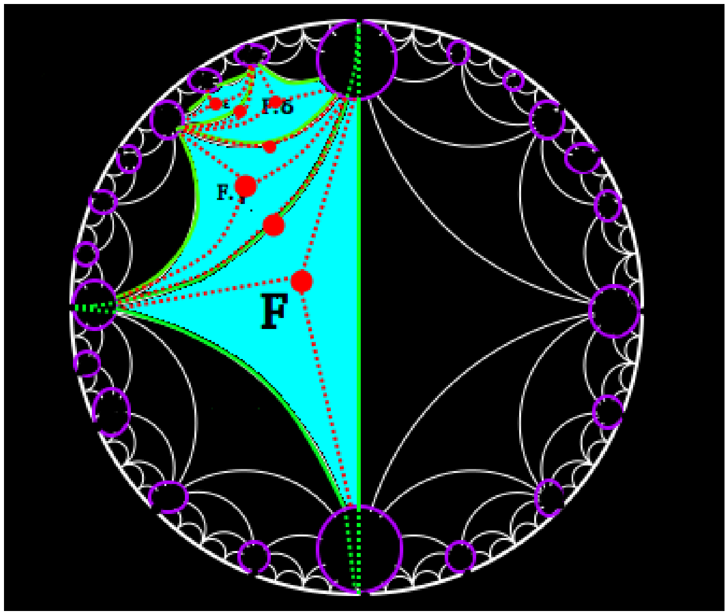

Figure 7.6: Evaluating the repulsion cost relative to the repulsion-cost c_2 at various source points $x, x', x'', \ldots$ etc. in $X[t]$

Figure 7.7: Active Domain for optimal semicoupling with respect to repulsion cost is homotopy-equivalent to the source, c.f. Theorem 1.4.1

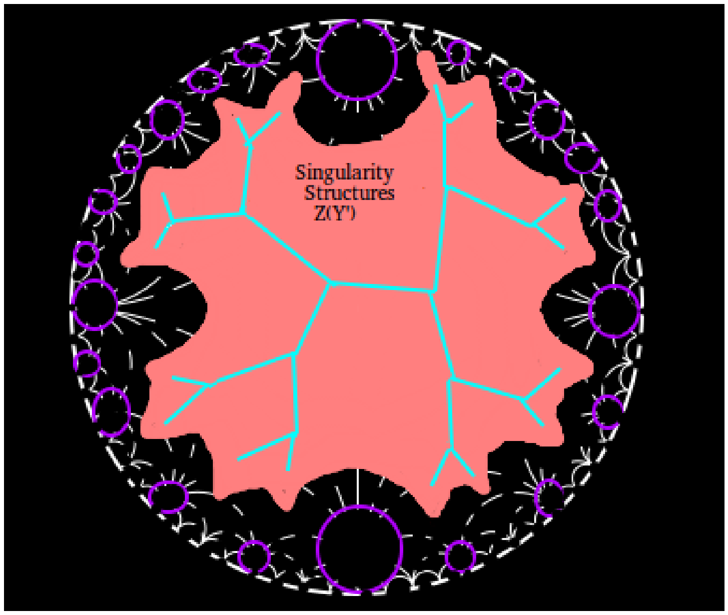

Figure 7.8: Singularity Structure of optimal semicoupling. The active domain is homotopy-equivalent to one-dimensional tree, c.f. Theorem 4.4.4

for which the chain sum $\sum_i B_i$ occupies the submodule

$$ker(\varphi : C_0 \to \mathbf{D}) = image(\partial_1 : C_1 \to C_0)$$

in C_0. Strictly speaking we work over $\mathbb{Z}/2$ with trivial $\mathbb{Z}\Gamma$-module structure, and replace $\mathbb{Z}\Gamma$ with the induced module $\mathbb{Z}_2\Gamma = \underline{\mathbb{Z}/2} \otimes_{\mathbb{Z}} \mathbb{Z}\Gamma$.

Now Closing Steinberg in $GL(\mathbb{Z}^3)$ means printing a nonzero element of C_1, e.g. something like

$$\xi := [a, b, c, d] + [c, d, e, f] + [e, f, a, b],$$

where $a, b, \ldots, f$ are all primitive integral vectors in $Proj[\mathbb{Q}^3]$. More concretely, consider

$$\xi' := \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Lemma 7.4.1. *Define $\partial_0[a, b, c] = [a] + [b] + [c]$. Then the image $\partial_0[\partial_1(\xi')]$ vanishes in $\boldsymbol{D}$ over $\mathbb{Z}/2$-coefficients.*

Proof. Compute according to the formulas. The columns of ξ' occur in $\partial_0[\partial_1(\xi')]$ with even multiplicity, and therefore vanish over $\mathbb{Z}/2$-coefficients. $\square$

We observe that the set of 3×3 minors of the summands of ξ' defines the finite subset

I' of $PGL(\mathbb{Z}^2)$ which Closes Steinberg.

As consequence of Theorem 7.2.3 we find

Proposition 7.4.2. *The convex hull of the rank-one states spanned by the columns of ξ', i.e.*

$$F := conv[\{x^2, y^2, z^2, (x+z)^2, (y+z)^2, (x+y+z)^2\}] \subset Q$$

forms a convex set F, whose translates $F.GL(\mathbb{Z}^3)$ tessalate Voronoi's cone Q of three-dimensional positive-semidefinite real states.

N.B. The above proposition is at least consistent with respect to dimensions, since $\dim[F] = 5$ coincides with $\dim Proj[Q] = \dim K\backslash {}^0GL(\mathbb{R}^3)$, where of course ${}^0GL(\mathbb{R}^3)$ is more commonly known as $SL(\mathbb{R}^3)$. Moreover it is not necessary that the orbit of F "fill" Q; in general the translates of F will tessalate a proper simply-connected subset of Q. There are further hypotheses in Closing Steinberg, namely the boundary summands

$$\partial_1[a,b,c,d] = [b,c,d] + [a,c,d] + [a,b,c]$$

must be Γ-translates of $[a,b,c]$. I.e., $\partial_1[a,b,c]$ is supported on the orbit $[a,b,c].\mathbb{Z}_2\Gamma$ in C_0. Next the translates $F.GL(\mathbb{Z}^3)$ assemble to chain sum $\underline{F} := \sum_{\gamma \in PGL(\mathbb{Z}^3)} F.\gamma$. Assembling all the constructions of our previous Chapters, we consider the visible repulsion cost $v : \underline{F} \times \partial[\underline{F}] \to \mathbb{R}$, and the v-optimal semicouplings from volume source measure σ on $\underline{F}$ to volume target measure τ on the excision boundary. We propose Kantorovich's functor $Z = Z(c_2, \sigma, \tau) : 2^{\partial[\underline{F}]} \to 2^{\underline{F}}$ realizes a spine for $GL(\mathbb{Z}^3)$.

Conjecture 7.4.3. *In the above notation with visibility cost $c = v$, the Kantorovich singularities produce codimension two deformation retracts $Q \approx \underline{F} \to Z_3$ onto those points $x' \in \underline{F}$ where $\partial^c\psi^c(x')$ has dimension ≥ 3.*

To practically construct the spine Z_3 requires the v-concave potentials $(\psi^v)^v = \psi$ arising from Kantorovich duality. The hypotheses of Closing Steinberg and the definition of v implies the (UHS) conditions are controlled by the gates. In this case, restricting the cost to a gate $v|G$, symmetry implies the two-dimensional gates deformation retract to a point. Thus we find $Q \approx \underline{F}$ retracts onto the codimension two subvariety $Z_3 \hookrightarrow Q$. It would be interesting to compare the above spine Z_3 with Soulé's cube [Sou78], [Ste07, Appendix], and the construction of [Gjo12]. We leave that to future investigations.

Chapter 8

Conclusion

So our thesis is concluded. Have we achieved our aims? Firstly we have developed a general method of Reduction-to-Singularity by which we construct continuous homotopy-reductions via the singularity structures of c-optimal semicouplings. Secondly we investigate concrete costs and settings which we propose as effective for applications. Our main results are Theorems 3.1.1, 4.4.4. For applications, our main result is Theorem 1.5.1. But admittedly our initial ambitions, namely Conjecture 1.5.2 remains partially unresolved and especially items (C1)–(C3). Moreover the problem of verifying that our visibility costs v satisfy (Twist) and sufficient (UHS) conditions remains open, c.f. Conjecture 5.9.7.

The present thesis is based on the keystone fact that "the disk $X = D$ admits no continuous retraction onto its boundary $Y = \partial D$" (recall Section 1.2), and attempts to reveal a new path forward. The algebraic-topology of Kantorovich's contravariant singularity functor $Z : 2^Y \to 2^X$, for $Z = Z(c, \sigma, \tau)$, has been developed §4.1.1. We use the functor Z to contravariantly parameterize closed subsets $Z(Y_I) \hookrightarrow X$ according to closed subsets $Y_I \hookrightarrow Y$. Assembling these inclusions, we find new cellular decompositions of a source space X, contravariantly parameterized by the target space Y. Given uniform Halfspace (UHS) conditions (Definition 4.4.2), our main Theorems 3.1.1, 4.4.4 identify an index $J \geq 1$ for which the source space X can be continously reduced via strong deformation retracts onto a codimension-J subvariety $Z_{J+1} \hookrightarrow X$. Our Theorem C expresses a general homotopy-reduction procedure based mainly on our Closing Steinberg symbol construction (Definition 7.2.1) and Theorem 7.2.3). Many new applications are possible, and to be developed in future investigations.

Bibliography

[Alb94] G. Alberti. "On the structure of singular sets of convex functions". In: *Calculus of Variations and Partial Differential Equations* 2.1 (1994), pp. 17–27. URL: https://doi.org/10.1007/BF01234313.

[Ale06] A.D. Alexandrov. *Convex Polyhedra*. Springer Monographs in Mathematics. Springer Berlin Heidelberg, 2006.

[Alp87] R.C. Alperin. "An elementary account of Selberg's lemma". In: *Enseign. Math.* 2. 33 (1987), pp. 269–273.

[Ash84] A. Ash. "Small dimensional classifying spaces for arithmetic subgroups of general linear groups". In: *Duke Math. J.* 51.2 (1984), pp. 459–468. URL: http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm (visited on 09/30/2010).

[AGM] A. Ash, P.E. Gunnells, and M. McConnell. "Resolutions of the Steinberg Module for $GL(n)$". In: (). URL: https://www2.bc.edu/avner-ash/Papers/Steinberg-AGM-V-6-23-11-final.pdf (visited on 05/20/2017).

[AR79] A. Ash and L. Rudolph. "The Modular Symbol and Continued Fractions in Higher Dimensions". In: *Inventiones math.* 55 (1979), pp. 241–250. URL: https://eudml.org/doc/186135.

[BGS85] W. Ballman, M. Gromov, and V. Schroeder. *Manifolds of Nonpositive Curvature*. Progress in Mathematics, Vol. 61. Birkhauser, 1985.

[Bar02a] Dror Bar-Natan. "On Khovanov's categorification of the Jones polynomial." In: *Algebraic Geometric Topology* 2 (2002), pp. 337–370. URL: http://eudml.org/doc/122177.

[Bar02b] A. Barvinok. *A Course in Convexity*. Graduate Studies in Mathematics 54. Springer-Verlag, 2002.

[BE73] R. Bieri and B. Eckmann. “Groups with homological duality generalizing poincare duality”. In: *Inventiones. Math.* 20 (1973), pp. 103–124. URL: https://eudml.org/doc/142208 (visited on 04/05/2017).

[BJ06] A. Borel and L. Ji. *Compactifications of symmetric and locally symmetric spaces*. Mathematical Theory and Applications. Birkhhauser, 2006.

[BS73] A. Borel and J.-P. Serre. “Corners and arithmetic groups”. In: *Comm. Math. Helv* 48 (1973), pp. 436–491. URL: https://eudml.org/doc/139559 (visited on 04/05/2017).

[BT65] A. Borel and J. Tits. “Groupes reductifs”. In: *Publications Mathematiques de l'IHES* 27 (1965), pp. 55–151. URL: http://eudml.org/doc/103858.

[Bre93] G. Bredon. *Topology and Geometry*. Graduate Texts in Mathematics no. 139. Springer-Verlag, 1993.

[Bro82] K.S. Brown. *Cohomology of groups*. Graduate Texts in Mathematics 87. Springer-Verlag, 1982.

[CM10] L. Caffarelli and R.-J. McCann. “Free boundaries in optimal transport and Monge-Ampere obstacle problems”. In: *Ann. of Math* 171.2 (2010), pp. 673–730.

[Cre84] J.E. Cremona. “Hyperbolic tessellations, modular symbols, and elliptic curves over complex quadratic fields”. In: *Compositio Mathematica* 51.3 (1984), pp. 275–324. URL: http://eudml.org/doc/89646.

[ET99] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*. Classics in Applied Mathematics, Vol. 28. SIAM, 1999.

[Fig10a] A. Figalli. “Regularity properties of optimal maps between nonconvex domains in the plane”. In: *Comm. Partial Differential Equations 35 (2010), no. 3* (2010), pp. 465–479.

[FK10] A. Figalli and Y.-H. Kim. “Partial regularity of Brenier solutions of the Monge-Ampère equation”. In: *Discrete Contin. Dyn. Syst. 28. no.2* (2010), pp. 559–565.

[Fig10b] Alessio Figalli. “The optimal partial transport problem”. In: *Archive for rational mechanics and analysis* 195.2 (2010), pp. 533–560.

[Fur76] H. Furstenberg. "A note on Borel's density theorem". In: *Proc. Amer. Math. Society* 55.1 (1976), pp. 209–212. URL: http://www.ams.org/journals/proc/1976-055-01/S0002-9939-1976-0422497-X/S0002-9939-1976-0422497-X.pdf (visited on 04/05/2017).

[Gan95] W Gangbo. "Habilitation thesis". In: *Universite de Metz* (1995). URL: https://www.math.ucla.edu/~wgangbo/publications/habilitation.pdf.

[GM96] W. Gangbo and R.-J. McCann. "Geometry of optimal transportation". In: *Acta Math.* 177 (1996), pp. 113–161. URL: https://projecteuclid.org/download/pdf_1/euclid.acta/1485890981.

[Gjo12] Oliver Gjoneski. "Contracting The Well-Rounded Retract". In: (2012). URL: https://arxiv.org/abs/1204.5218.

[Gra84] D. R. Grayson. "Reduction theory using semistability I". In: *Comm. Math. Helv.* 59 (1984), pp. 600–634. URL: math.u-bordeaux.fr/~fpazuki/index_fichiers/grayson1.pdf.

[GJ81] M.J. Greenberg and J.R.Harper. *Algebraic Topology: a first course*. ABP. Perseus Publishing, 1981.

[Gro] M. Gromov. "Systoles and intersystolic inequalities". In: *Actes de la Table Ronde de Géométrie Différentielle (Luminy, 1992). Sémin. Congr., 1, Soc. Math. France* (), pp. 291–362.

[Gro82] M. Gromov. "Volume and bounded cohomology". In: *Publications Mathematiques de l'IHES* 56 (1982), pp. 5–99. URL: http://eudml.org/doc/103988.

[Gro91] M. Gromov. "Sign and geometric meaning of curvature". In: *Rendiconti del Seminario Matematico e Fisico di Milano* 61.1 (1991), pp. 9–123. URL: https://doi.org/10.1007/BF02925201.

[Gro14a] M. Gromov. "Dirac and Plateau billiards in domains with corners". In: *Central European Journal of Mathematics* 12.8 (Aug. 2014), pp. 1109–1156.

[Gro14b] M. Gromov. "Plateau-Stein manifolds". In: *Central European Journal of Mathematics* 12.7 (2014), pp. 923–951.

[Har64] P. Hartman. *Ordinary Differential Equations*. John Wiley & Sons, 1964.

[HL69] H. Hermes and J.P. Lasalle. *Functional Analysis and Time Optimal Control*. Mathematics in Science and Engineering, Vol. 56. Academic Press, 1969.

[HS13] M. Huesman and K.-T. Sturm. "Optimal transport from Lebesgue to Poisson". In: *Ann. Prob. 41* no. 4 (2013), pp. 2426–2478. URL: http://projecteuclid.org/euclid.aop/1372859757.

[Jös97] J. Jöst. *Nonpositive curvature: geometric and analytic aspects*. Lecture notes in mathematics, ETH Zurich. Birkhauser, 1997.

[KP18] Y.-H. Kim and B. Pass. "A Canonical Barycenter via Wasserstein Regularization". In: *SIAM Journal on Mathematical Analysis* 50.2 (2018), pp. 1817–1828. DOI: 10.1137/17M1123055. URL: https://doi.org/10.1137/17M1123055.

[KM18] J. Kitagawa and R.-J. McCann. "Free Discontinuities in Optimal Transport". In: (2018). URL: http://www.math.toronto.edu/mccann/papers/KitagawaMcCann.pdf.

[Lan05] S. Lang. *Algebra*. Graduate Texts in Mathematics no. 211. Springer New York, 2005.

[Lan66] R.P. Langlands. "The Volume of the Fundamental Domain for Some Arithmetical Subgroups of Chevalley Groups". In: *in Algebraic Subgroups and Discontinuous Subgroups* (1966), pp. 143–148. URL: http://sunsite.ubc.ca/DigitalMathArchive/Langlands/pdf/chev-ps.pdf.

[LS76] R. Lee and R.H. Sczarba. "On the Homology and Cohomology of Congruence Subgroups." In: *Inventiones mathematicae* 33 (1976), pp. 15–54. URL: http://eudml.org/doc/142375.

[MM93] M.MacConnell and R. Macpherson. "Explicit reduction theory of Siegel modular threefolds". In: *Invent. Math* 111 (1993), pp. 575–625. URL: https://eudml.org/doc/144090.

[McC01] R.J. McCann. "Polar factorization of maps on Riemannian manifolds". In: *Geometric & Functional Analysis GAFA* 11.3 (2001), pp. 589–608.

[Nee85] A. Neeman. "Topology of quotient varieties". In: *Ann. Math* 122.2 (1985), pp. 419–459.

[Oli07] V. Oliker. "Embedding S^n into R^{n+1} with given integral Gauss curvature and optimal mass transport on S^n". In: (2007). eprint: math/0701398. URL: https://arxiv.org/abs/math/0701398.

[PS08] A. Pettet and J. Souto. "Minimality of the well-rounded retract". In: *Geometry and topology* 12 (2008).

[Phe89] R. R. Phelps. “Lectures on Choquet’s Theorem. Second Edition”. In: *Lecture Notes in Mathematics* 1757 (1989).

[PF] G. De Philippis and A. Figalli. “Partial Regularity for optimal transport maps”. In: (). URL: https://arxiv.org/abs/1209.5640.

[Poi95] H. Poincaré. “Analysis Situs”. In: *Journal de l’École Polytéchnique* 1 (1895), pp. 1–121.

[San15] F. Santambrogio. *Optimal Transport for Applied Mathematicians*. Progress in Nonlinear Differential Equations, Vol. 87. Birkhhauser, 2015.

[Ser] J.-P. Serre. “Cohomologie des groupes discrets”. In: *Prospects in Mathematics, Annals of Math. Studies 70, Princeton University Press, Princeton, 1970* (). URL: https://www.college-de-france.fr/media/jean-pierre-serre/UPL1711642281342131119_Serre_Cohom.Groupes_Discrets.pdf (visited on 11/07/2018).

[Ser71] J.-P. Serre. “Cohomologie des groupes discrets”. In: *Seminaire Bourbaki* 13 (1970-1971), pp. 337–350. URL: http://eudml.org/doc/109803.

[Sou78] C. Soulé. “The cohomology of SL(3,Z)”. In: *Topology* 17 (1978), pp. 1–22.

[Ste07] W. Stein. *Modular Forms, a Computational Approach (with an Appendix by P.E. Gunnells)*. Graduate Studies in Mathematics no. 79. Springer-Verlag, 2007.

[Thu02] W.P. Thurston. “Geometry and Topology of Three-Manifolds”. In: (2002). URL: http://library.msri.org/books/gt3m/.

[Thu86] William Thurston. “A spine for Teichmüller space”. In: *preprint* (1986).

[Tit72] J. Tits. “Free subgroups of linear groups”. In: *Journal of Algebra* 20.2 (1972), pp. 250–270. URL: https://www.sciencedirect.com/science/article/pii/0021869372900580.

[Vil03] C. Villani. *Topics in Optimal transportation*. Graduate Studies in Mathematics Vol. 58. Springer-Verlag, 2003.

[Vil09] C. Villani. *Optimal transport: old and new*. Grundlehren der mathematischen wissenschaften Vol. 338. Springer-Verlag, 2009.

[Yas07] Dan Yasaki. “On the existence of spines for Q-rank 1 groups”. In: *Selecta Mathematica* 12.3-4 (2007), p. 541.