February 23, 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

February 23, 2019

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

We describe some new applications of optimal transportation to algebraic topology, and our method for building Spines/Souls using a definition of "Singularity" based on Kantorovich duality. Given an initial manifold-with-corners X, we construct continuous homotopy-reductions from X onto large-codimension closed subvarieties $\mathscr{Z} \hookrightarrow X$. The subvariety $\mathscr{Z}$ is assembled from a contravariant functor $Z : 2^Y \to 2^X$, where $Z = Z(c, \sigma, \tau)$ is determined by a source (X, σ), target (Y, τ), and cost $c : X \times Y \to \mathbb{R}$. Our Theorems A, B describe criteria for inclusions $Z(Y_I) \hookrightarrow Z(Y_J)$ to be homotopy-isomorphisms and find a maximal index $J \geq 1$ for which $(\mathscr{Z} := Z_{J+1}) \hookrightarrow X$ is a codimension-J homotopy-isomorphism. Our Theorem C 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 $\Gamma = GL(\mathbb{Z}^2)$, $GL(\mathbb{Z}^3)$, etc. The singularity functor Z is a new reduction theory for $E\Gamma$ models based on convex excisions, and a subprogram we call "Closing the Steinberg symbol" which is a convex interpretation of Bieri-Eckmann's homological duality. "Closing Steinberg" replaces X with a chain sum $\underline{F}$ on which Γ freely acts as shift operator. Finally we construct "two-pointed repulsion" costs $\widetilde{c}$ and "visibility" costs v on $\underline{F}$ whose functors $Z(c, \sigma, \tau)$ 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 . . . 3

1.3 Cost Assumptions . . . 5

1.4 Deformation Retracts and Topological Theorems A,B . . . 7

1.5 Closing the Steinberg symbol and Theorem C . . . 9

1.6 Thesis Outline . . . 15

1.7 Conventions and Notations . . . 16

2 Semicoupling Programs and Topological Theorem A **19**

2.1 Optimal Semicouplings and Cost . . . 21

2.2 Existence of c-Optimal Semicouplings . . . 22

2.3 Kantorovich Duality . . . 23

2.4 Uniqueness of Activated Domain . . . 30

2.5 Uniqueness of Optimal Semicouplings . . . 33

2.6 Deforming $X \rightsquigarrow A$: Topological Theorem A . . . 37

3 Kantorovich Singularity and Topological Theorem B **45**

3.1 Kantorovich's Contravariant Singularity Functor . . . 45

3.2 Local Topology and Local Dimensions of Z . . . 47

3.3 The Descending Filtration $Z_j, j = 0, 1, 2, \ldots$ . . . 51

3.4 Local–Global Homotopy Reductions: (UHS) Conditions . . . 55

4 Repulsion Costs **60**

4.1 Hypotheses and Practical Applications . . . 60

4.2 Definition of Repulsion Costs . . . 63

4.3 Chain sums and Well-Separated Gates . . . 67

4.4 Two-Pointed Repulsion Costs . . . 68

4.5 Convex-Excisions 71
4.6 Visibility 72
4.7 ϵ-Regularizations 74
4.8 Barycentres and Krein-Milman 76
4.9 Visible Repulsion Costs 78

5 Symmetry and Dimension **83**
5.1 Proper Classifying Spaces $E\Gamma$ 83
5.2 Group Cohomology: Algebraic Background 85
5.3 Excision versus Compactification 88
5.4 Bieri-Eckmann Duality 92
5.5 Arithmetic Groups: Excision 95

6 Closing the Steinberg symbol **98**
6.1 Stitching Footballs from Regular Panels : Motivation 98
6.2 Closing Steinberg : Definition and Consequence 99
6.3 Closing Steinberg on $PGL(\mathbb{Z}^2)$: First Example 102
6.4 Closing Steinberg on $GL(\mathbb{Z}^3)$: Example 108

7 Conclusion **111**

Bibliography **112**

List of Figures

1.1 When $\partial(A+B+C)=0 \pmod 2$, we define $F = conv(A,B,C)$ 12
1.2 A chain sum $\underline{\mathrm{F}}{=}\sum_{i\in I} F_i$ with well-separated one-dimensional gates 13

2.1 Disconnected Active Domain 33
2.2 Connected Active Domain when $mass[\sigma]/mass[\tau] \approx 1^+$ 33
2.3 Halfspace Conditions fail and active domain is not homotopy-equivalent to source. 38

3.1 Singularity structure Z on active domain $A = Z_1$ 47
3.2 Singularity structure on Z_2 47
3.3 Horospherically truncated pair of pants, with active domain relative to a repulsion cost (See Definition 4.2.3 from Section 4.2). 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 59

4.1 Singularity structures of $\widetilde{c}$ on regular 2-simplex and regular square satisfy properties (D0)–(D5) 63
4.2 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^*. 64
4.3 Graph of repulsion cost $x \mapsto c|\mathscr{H}_Y(x,0)$. The assymmetry reflects the "home-preference" of x to $y=0$. 66
4.4 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. . . . 67
4.5 Illustrating the projection map $proj : F \to \partial F$ defining two-pointed repulsion cost $\widetilde{c}$ for ellipse F where $\partial F = \mathscr{E}[F]$. 70
4.6 Excision of convex ellipsoid 72
4.7 Excision domain $X = F[t]$, where (x,y) are visible in X and (x,w) is not visible in X 73

4.8 The outer boundary is the boundary of the closed outward ϵ-neighborhood $\partial((X[t]))_{\leq\epsilon}$ of the excision boundary $Y = \partial_* X[t]$. The inner hypersurface is the smooth ϵ-regularization Y_ϵ 75
4.9 The subdomain Ω is geodesically convex 80
4.10 Evaluating visibility cost $v(x, z)$ for pair $(x, z) \in \Omega \times \partial_* F$ 81

5.1 (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. 90

6.1 Stitching a football F from identical regular hexagons or pentagons P_i . . 98
6.2 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. 99
6.3 Convex excision $X[t]$ relative a Γ-equivariant excision parameter $t : \mathbb{Q}P^1 \to \mathbb{R}$ 105
6.4 Steinberg symbol in $X[t] \times \partial X[t]$ is represented as relative 1-cycle P with boundary ∂P equal to 0-sphere. 106
6.5 The translates by $I_0 = \{Id, \gamma, \gamma'\}$ successfully Close the Steinberg symbol. 106
6.6 Evaluating the repulsion cost relative to the repulsion-cost $\widetilde{c}$ at various source points $x, x', x'', \ldots$ etc. in $X[t]$ 107
6.7 Active Domain for optimal semicoupling with respect to repulsion cost is homotopy-equivalent to the source, c.f. Theorem A. 107
6.8 Singularity Structure of optimal semicoupling. The active domain is homotopy-equivalent to one-dimensional tree, c.f. Theorem B. 108

Chapter 1

Introduction

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 author is amazed by the subject of mass transportation, and has found it surprising tool for numerous topological applications. Especially surprising is the application of Singularity to explicitly construct new Spines long sought by geometers. A topologist initially interprets Singularity as the "locus-of-discontinuity" of, "deformation retract"-type set maps $r : X \to \partial X$, but this definition is not topological since the categories of Topology are defined by *continuous maps*. But *continuous* retracts r from X to ∂X do not exist!

However a topologist finds optimal transportation a useful setting to formally define Singularity in the category of Topology. This formalization is summarized by contravariant functors $Z : 2^{\partial X} \to 2^X$ between the categories of closed subsets of $X, \partial X$. The contravariant functor Z is defined via maximizers of the dual program to semicouplings programs $Z = Z(\sigma, \tau, c)$. 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.

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

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 basic theory of c-concavity and c-subdifferentials is beautiful 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] 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 high-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 general results are Theorems A, B, 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 C and the introduction of a subprogram we call "Closing the Steinberg symbol" (Chapter 6), 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 §§6.3–6.4 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 $c = v$, §4.9.5) satisfy sufficient hypotheses of Theorems A and B to produce new examples of minimal spines of various $E\Gamma$ models. We begin elaborating these ideas below.

1.1 Kantorovich: the Bridge from Measure to Topology

The present thesis develops a bridge between measure theory and algebraic topology, as inspired from Kantorovich duality in optimal transport. The bridge leads to a new general method for constructing explicit "souls" and "spines" in numerous geometric settings. 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 3 and Definition 3.1.1 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 yields a new response to Gromov's questions [Gro10], and 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 $c : X \times Y \to \mathbb{R}$, and Kantorovich duality applied to the c-optimal semicoupling program.

Our main theorems identify a local condition, which we call uniform Halfspace (UHS) conditions (see Definition 3.4.1), and our main topological Theorems A, B (see Theorems 3.4.2, 3.4.3) identifies an index $J \geq 1$ 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 5, 6, and §1.2 below. A general homotopy-reduction procedure is expressed in Theorem C, 1.5.1, which follows from our Closing Steinberg symbol construction (Definition 6.2.1 and Theorem 6.2.4).

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 $H_*(i)$. Now suppose X is connected and a homologically trivial topological space, e.g. contractible. Let Y be a homologically nontrivial closed subspace, so $H_*(X) = 0$ and $H_*(Y) \neq 0$. Let $r : X \to Y$ be 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 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.

The motivating example is this: 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. A further example: if X is closed disk with boundary $Y = \partial D$, then there exists no deformation retract $h : X \to Y$ which has singularity equal to two distinct points $\{a, b\}$. Our informal proof: the inclusion $\{a, b\} \hookrightarrow X$ is not a homotopy-isomorphism! We make these ideas rigorous in Chapters 2 and 3.

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 interpret semicouplings as measure-theoretic retractions from (X, σ) to $(\partial X, \tau)$, and 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 A, B, C below require cost functions $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ satisfying basic assumptions labelled (A0), …, (A5). Abbreviating $c_y(x) = c(x, y)$, 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. $c \geq 0$.

(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 $dom(c_y)$. Moreover we assume the sublevels $\{x \in X | c(x, y) \leq t\}$ are compact subsets of X for every $t \in \mathbb{R}$, and $y \in Y$.

(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 4.2.5.

(A5) The gradients $x \mapsto \nabla_x c(x, y)$ and $x \mapsto \nabla_x c_\Delta(x; y, y')$ are bounded away from zero on their relevant subdomains, uniformly in x and $y, y' \in dom(c_x)$. Here $c_\Delta(x; y, y') = c(x, y) - c(x, y')$ is the two-pointed cross difference.

The consequences of Assumptions $(A0), (A1), \ldots$ 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)–(A3) are readily verified. Assumptions (A4)-(A5) are more discriminate, and gradients $\nabla_x c$ and $\nabla_x c_\Delta$ will usually vanish over $X \times Y$. Indeed the continuously differentiable functions $x \mapsto c(x, y)$ and $x \mapsto c_\Delta(x; y, y')$ necessarily have critical points on compact spaces. The Assumption (A4) implies the cross-differences $\nabla_x c_\Delta(x, y, y')$ are nonzero for distinct y, y'. In applications below, the Assumption (A5) requires the gradients $\nabla_x c_\Delta(x, y, y')$ to furthermore have nonzero projections $proj_{Z'} \nabla_x c_\Delta(x, y, y') = \nabla_x^{Z'} c_\Delta(x, y, y') \neq 0$. See Property (C) as defined in 2.6.4 and (2.15), and this Property (C) plays key role in our proof of Theorem B, c.f. 3.4.3.

1.4 Deformation Retracts and Topological Theorems A,B

We continue with $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ a cost satisfying Assumptions (A0)–(A4). For every abundant source σ and target τ, there exists a unique c-optimal semicoupling π_{opt} from σ to τ. 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 being the main idea of our thesis, all this is developed further, c.f. §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 $dom(\psi^c) := \{x \in X | \psi^c(x) \neq +\infty\}$ is a nontrivial subset of the source X. Informally, when the source is abundant with respect to the target, a c-optimal transport 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 A describes a criteria to ensure the activated source $A \hookrightarrow X$ includes into the source as homotopy-isomorphism.

Theorem A. *Let c be cost satisfying Assumptions (A0)–(A4). Let A be the active domain of a c-optimal semicoupling from source (X, σ) to target (Y, τ). Suppose every $x_0 \in X - A$ has property that the closed convex hull*

$$\{\nabla_x c(x_0, y) \mid y \in dom(c_{x_0})\}$$

in $T_{x_0}X$ does not contain the origin $0 \in T_{x'}X$; and moreover suppose

$$||\nabla_x c(x_0, y)|| \geq C > 0$$

for a constant C uniform in $y \in dom(c_{x_0})$. Then the inclusion $A \hookrightarrow X$ is a homotopy-isomorphism and there exists a strong deformation retract.

In fact we prove a stronger result in Theorems 3.4.2, 3.4.3 below and replace the hypothesis on the convex hull of $\{\nabla_x c(x', y)\} \subset T_{x'}X$ with our Uniform Halfspace (UHS) condition. See Definition 3.4.1. Briefly, we say (UHS) conditions are satisfied throughout $X - A$ if there exists a constant $C > 0$ such that for every $x' \in X - A$, the average of $\{\nabla_x c(x', y) | y \in dom(c_{x'})\}$ with respect to the uniform Hausdorff measure $\mathscr{H}_Y$ on $dom(c_{x'})$

in $T_{x'}X$ is nonzero, and this average

$$\eta(x, avg) := (\mathscr{H}_Y[Y])^{-1} \int_{dom(c_{x'})} \nabla_x c(x', y) d\mathscr{H}_Y(y)$$

has norm bounded below

$$||\eta(x, avg)|| \geq C. \int_{dom(c_{x'})} ||\nabla_x c(x', y)|| d\mathscr{H}_Y(y).$$

The assumption (A5) is key technical hypothesis for constructing the deformation retract in Theorem A. Without the inequality, our methods could only conclude that the source X deformation retracts onto ϵ-neighborhood $A_\epsilon = \{x \in X | d(x, A) < \epsilon\} \hookrightarrow X$ of $A \hookrightarrow X$, where $\epsilon > 0$ is a sufficiently small real number. We strengthen the Assumption (A5) with our so-called uniform Halfspace (UHS) condition. These ideas are detailed in §§2.6, 3.4.

Remark. Theorem A has the following application to quadratic costs $c = d^2/2$. Specifically we consider quadratic costs $c(x, y) = ||x - y||^2/2$ for closed subsets X, Y in some Euclidean space $\mathbb{R}^N$. Observe the closed convex hull $\{\nabla_x c(x_0, y)\}_{y \in Y}$ contains the origin if and only if x_0 occupies the convex hull of Y. Therefore if the activated domain A of a $d^2/2$-optimal semicoupling contains the convex hull $conv(Y) \subset A$, then Theorem A implies the inclusion $A \subset X$ is a homotopy-isomorphism. The proof of Theorem A constructs an explicit strong deformation retract of X onto A. We find A contains $conv(Y)$ whenever the ratio $\int_X \sigma / \int_Y \sigma > 1$ is sufficiently close to 1^+, i.e. whenever the active domain A is a sufficiently large subset of X. However the applications of quadratic cost to closed convex sets X with target $Y = \mathscr{E}[X]$ supported on the extreme pointset is limited – indeed the active domain A of optimal *semicouplings* is a strictly proper subset of X, and every point $x_0 \in X - A$ in the complement fails to satisfy the hypotheses of Theorem A. For applications to source convex subsets X and targets $Y = \mathscr{E}[X]$, the observations of the previous paragraph motivate our introduction of anti-quadratic costs called "repulsion" costs. As we develop our applications below, we shall verify our repulsion costs satisfy the hypotheses of Theorem A for general semicouplings.

Now Theorem A is but a first step. Our next Theorem B constructs further homotopy-reductions from the active domains $A \subset X$ to higher codimension closed subvarieties $\mathscr{Z} \hookrightarrow A$. In practice one finds the singularity structure naturally cellulated by "local cells" and admits a filtration

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

of X into closed subvarieties (Z_1, Z_2, ... For every integer $j \geq 1$, we heuristically 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 §3.3, c.f. Propositions 3.2.4, 3.3.5. Our next result identifies the maximal index $J \geq 0$ such that $X \hookleftarrow Z_{J+1}$ is a homotopy-isomorphism, and indeed a strong deformation retract.

Theorem 1.4.1 (Theorem B). *Let c be cost satisfying Assumptions (A0)–(A5). Let $j \geq 1$ be integer. Suppose for every $x' \in Z_j - Z_{j+1}$ the collection of tangent vectors $\{\eta(x', y) | y \in dom(c_{x'}) - \partial^c \psi^c(x')\}$ defined in (3.5) satisfies (UHS) condition. Then the inclusion $Z_{j+1} \hookrightarrow Z_j$ is a homotopy-isomorphism.*

Furthermore if $J \geq 1$ is the maximal integer such that every $x' \in Z_J - Z_{J+1}$ satisfies (UHS) condition, 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 of Theorem 1.4.1 may be difficult to verify. It requires specific "average" vectors $\eta(x, avg)$ – and their projections $proj_{Z'}\eta(x, avg)$ – be simultaneously nonvanishing for a descending chain of subvarieties $\{Z'\}$. The reader will find details in Chapter 2.3 below.

Interesting applications arise when the target is the boundary $Y = \partial X$ of the source X. Combining Theorems A, B, we find cost functions $c : X \times \partial X \to \mathbb{R} \cup \{+\infty\}$ satisfying Assumptions (A0)–(A5) and sufficient (UHS) conditions produce codimension-J deformations of the initial source space X onto a closed locally-lipschitz subvariety Z_{J+1}, where J is an index ≥ 1. We propose the subvarieties Z_{J+1} produced by Theorems A and B explicitly construct "souls" generalizing the Cheeger-Gromoll-Perelman methods and illustrate examples in Chapters 4, 5, 6.

1.5 Closing the Steinberg symbol and Theorem C

The second phase of our thesis develops a method of constructing small-dimensional $E\Gamma$ classifying spaces, where Γ is an infinite discrete Bieri-Eckmann duality group of dimension ν and with dualizing module $\mathbf{D}$. The setting implies Γ is finitely-generated, virtually-torsion free. Ex. $\Gamma = GL(\mathbb{Z}^2)$, $GL(\mathbb{Z}^3)$, $Sp(\mathbb{Z}^4)$, $G(\mathbb{Z})$ 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 cohomological dimension $cd(\Gamma)$ of the group Γ.

This problem has been intensively studied since Borel-Serre's computation $vcd(G(\mathbb{Z})) = dim(K\backslash^0 G(\mathbb{R})) - rank_{\mathbb{Q}}(G)$, where $\mathbb{G}$ is a $\mathbb{Q}$-split reductive linear algebraic group scheme, see [BS73]. For instance $vcd(GL(\mathbb{Z}^2)) = 2 - 1 = 1$, $vcd(Sp(\mathbb{Z}^4)) = 5 - 2 = 3$. For groups G with $\mathbb{Q}$-rank equal to one, the problem simplifies since numerous adhoc methods are available for continuously retracting an open manifold onto a codimension-one hypersurface. 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 A, B yield a new technique for constructing homotopy-reductions of large-codimension based on the reduction-to-singularity idea. To implement the Theorems A, B 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 $X \times \Gamma \to X$ 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 and repulsion costs. The excision models $X[t] := X - \cup_{\{t\}} W_t$ are obtained by equivariantly scooping-out/excising Γ-rational horoballs W_t from X. The family of horoballs $\{W_t\}$ are 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 $FILL[B]$ 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 $FILL[B] \rightsquigarrow pt$. Our goal is to interpolate this retraction throughout $X[t]$ to obtain $X[t] \rightsquigarrow Z$. We achieve this interpolation using the singularity functor arising from our two-pointed repulsion cost $\widetilde{c}$ defined in §4.4 and extended to visibility costs in §4.9.5.

Next we replace the excision $X[t]$ with a convex chain sum $\underline{F}$, where Γ acts on $\underline{F}$ as a type of shift operator. That is, 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 introduced in Chapter 6. Informally the problem of Closing Steinberg is motivated by the following question probably known to Euclid school. 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? The reference [GS80] contains useful related background information.

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 $F = F(I)$ 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 5.

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

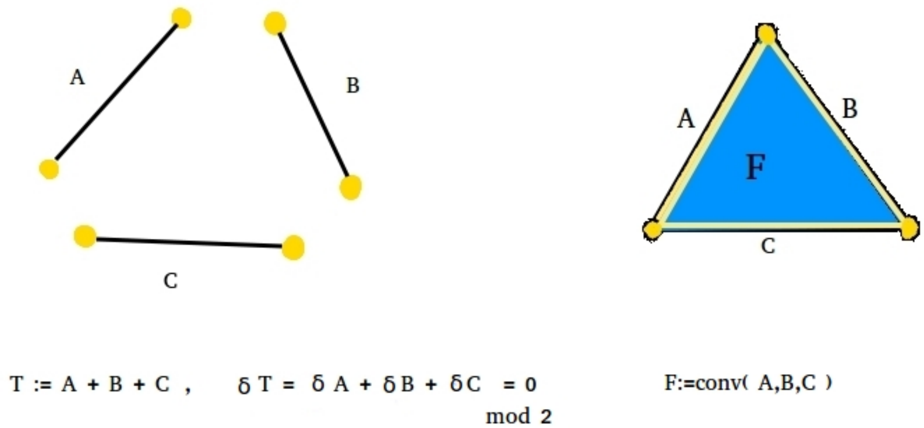

Figure 1.1: When $\partial(A+B+C)=0$ (mod 2), we define $F = conv(A,B,C)$

$\mathbb{Z}\Gamma$-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 ▲ are replace with a flat-filled relative chain $B \in C_q(X[t], \partial X[t]; \mathbb{Z})$ whose chain-boundary $\partial[B]$ is generator of Steinberg module $\mathbf{D}$. To Close the Steinberg symbol requires finding a finite subset I of Γ for which the translates $B.I$ have positions in $(X[t], \partial X[t])$ bounding a closed geodesically convex domain $F = conv[B.I]$. 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 6.2.1, 6.2.4) ensures the support $supp(\underline{F})$ is aspherical and homotopy-equivalent to X.

The above Theorems A, B are general topological theorems obtained by our semi-coupling methods. The theorems require costs c which satisfy the necessary hypotheses. As we elaborated above, the quadratic costs are not sufficient, and we find best results obtained with our anti-quadratic repulsion costs. The preferred cost for our applications are the "visibility costs" defined in Chapter 4 below. The following Theorem C summarizes our applications of the previous Theorems A, B. 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 $\underline{Z}$ where the index $J \geq 0$ is defined according to the hypotheses of Theorems A, B.

Theorem 1.5.1 (Reduction Program)**.** *Let $X \times \Gamma \to X$ be a geometric $E\Gamma$-model with*

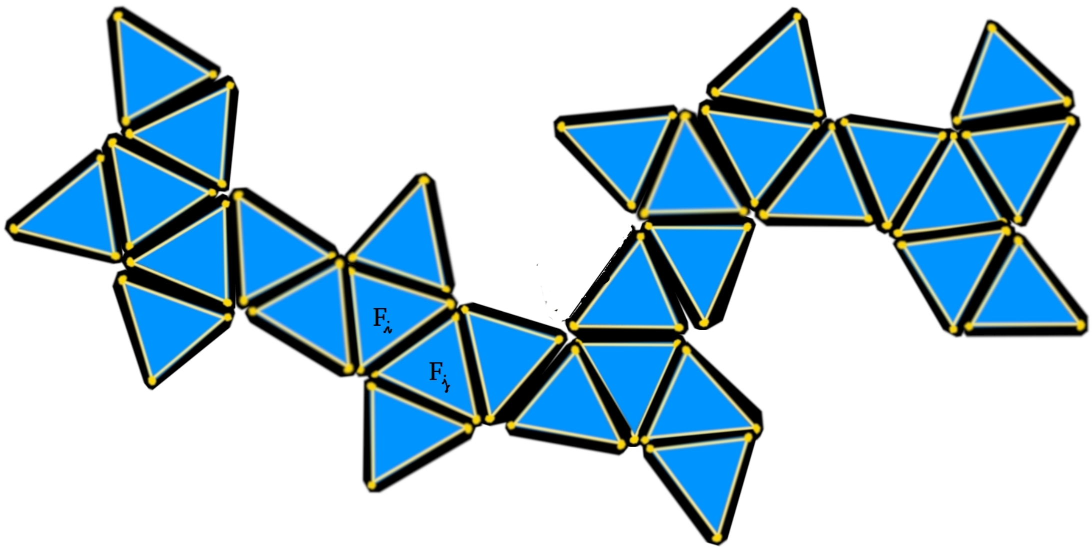

Figure 1.2: A chain sum $\underline{F} = \sum_{i \in I} F_i$ with well-separated one-dimensional gates

Γ-invariant volume measure σ. (c.f. Definition 5.1.2). Then we identify an index $J \geq 0$ and a codimension-J closed subvariety $\underline{Z}$ by executing the following steps:

(i) *Construct a Γ-equivariant excision $X[t]$ for which $\boldsymbol{D} := \tilde{H}_q(\partial X[t]; \mathbb{Z})$ is the Steinberg module of Γ. (c.f. Chapter 5). Let τ be a Γ-invariant volume measure on $\partial X[t]$ such that*

$$\int_{X[t]/\Gamma} \sigma / \int_{\partial X[t]} \tau > 1.$$

(ii) *Find a finite subset $I \subset \Gamma$ which successfully Closes the Steinberg symbol, and replace the excision model $X[t] \times \partial X[t]$ with the chain sum $\underline{F} \times \partial_*[\underline{F}]$, where $\underline{F} := SUM[F.\Gamma]$ and $F := conv[B.I]$. (c.f. Definition 6.2.1 and Theorem 6.2.4).*

(iii) *Let $v : \underline{F} \times \partial_*[\underline{F}]$ be the visible repulsion cost (§4.9), and solve the v-optimal semicoupling program between source σ and target τ. Kantorovich duality yields dual potentials ψ^v, ψ on $\underline{F}, \partial_*[\underline{F}]$, where $\psi = \psi^{vv}$ is v-concave potential; obtain Kantorovich's singularity functor $Z : 2^{\partial X[t]} \to 2^{X[t]}$ via Definition 3.1.1.*

(iv) *Using the singularity functor Z, filtrate the source $Z_0 \hookleftarrow Z_1 \hookleftarrow Z_2 \hookleftarrow \cdots$, where*

$$Z_j = \{x \in Z_1 | \dim_x Z(\partial^c \psi^c(x)) \geq j\}$$

for index $j \geq 1$. Construct homotopy-reductions $Z(Y_I) \rightsquigarrow Z(Y_J)$ for $Y_I \subset Y_J$, wherever sufficient Halfspace Conditions are satisfied according to Theorems A, B. (c.f. Theorems 2.6.2, 3.4.2, 3.4.3).

(Output) *Let $J \geq 1$ be maximal index where sufficient Halfspace conditions are satisfied throughout Z_J. Obtain the continuously assembled Γ-equivariant chain sum $\underline{Z} := Z_{J+1}$. Then $\underline{Z} \hookrightarrow X[t]$ is a Γ-equivariant codimension-J homotopy reduction of X.*

We emphasize that the primary obstruction to the reduction program of Theorem C is verifying the uniform Halfspace (UHS) conditions throughout the necessary subdomains. For general costs this appears difficult problem. The (UHS) conditions amount to requiring an average gradient vector $\eta(x, avg)$ (defined in (3.6)) be nonzero, and have a sequence of nonzero projections $proj_{Z'}\eta(x, avg) \neq 0$ for select closed subvarieties Z' containing x in X. See §3.4 and Definition 3.4.1 for details.

The definition of Closing Steinberg is specially adapted to geometric $E\Gamma$ models as described in §§6.3, 6.4 below when Γ is a standard arithmetic group. Thus we conjecture that Theorems A,B yield 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. We conjecture that our two-pointed repulsion cost $\widetilde{c}$ and visibility cost v satisfy the Conditions (D0)–(D4) from §4.1.

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

(i) *the visible repulsion cost v satisfies Assumptions (A0)–(A5); and*

(ii) *when the ratio*

$$\int_{X[t]/\Gamma} d\sigma(x) / \int_{\partial X[t]/\Gamma} d\tau(y)$$

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 $X[t] \approx X$;

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

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

Thus Conjecture 1.5.2 proposes the subvariety $Z_{J+1} \hookrightarrow X$ is a minimal-dimension spine of $E\Gamma$, and our Theorems A, B describes explicit deformation retract. The Conjecture 1.5.2 demands several steps be verified. First we need verify the differential-geometric (Twist) condition for v in (C1). This requires 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 $\underline{F}$. The second step (C2) establishes the homotopy-isomorphism between the activated source domains inclusion $(Z_1 = A) \hookrightarrow (Z_0 = \underline{F})$ of the v-optimal semicoupling. Step (C3) requires verifying that maximal (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. The author is presently incapable of establishing the steps (C1), (C2), ... We leave this 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. In Chapter 2 we develop the principles of the semicoupling program: existence, uniqueness, and Kantorovich duality for costs satisfying Assumptions (A0)–(A4), and we conclude Chapter 2 with the proof of Theorem A, which is the base case for the larger-codimension retracts constructed in the next chapters. The key technical hypothesis called uniform Halfspace (UHS) conditions is detailed in Definition 2.6.4 (especially estimate (2.15)). The (UHS) conditions gives a strong form of Assumption (A5) neccessary for the deformation retracts constructed throughout our thesis.

Our Chapter 3 develops the basic geometric properties of Kantorovich's contravariant singularity functor $Z : 2^Y \to 2^X$. §3.1 is central to our thesis, especially Definition 3.1.1. §3.2 describes the local topology of the singularities and proves a useful codimension estimate. Assuming the cost satisfies Assumptions (A0)–(A5) and sufficient (UHS) conditions, we conclude Chapter 3 with the proof of our Theorem B; see Theorems 3.4.2, 3.4.3 from §3.4. Thus with Chapters 2, 3 we establish 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 our applications. Chapter 4 introduces a new class of cost functions, so-called "repulsion", "(average) two-point repulsion", and "visibility" costs. These repulsion costs are contradistinct from the familiar "attraction" costs, e.g. quadratic costs $c = d^2/2$. Our repulsion costs have natural physical interpretations in terms of electron–electron interactions described below. To apply the homotopy-reductions of Theorems A, B however require we verify the repulsion costs satisfy the necessary Assumptions (A0)–(A5). In this direction, our thesis admittedly achieves only partial results. Specifically, we easily find the costs satisfy Assumptions (A0)–(A3). We only succeed in demonstrating (A4) for the repulsion cost denoted $c|\tau$ (Definition 4.2.3). However we present simple heuristics suggesting the costs satisfy the Assumption (A4), i.e. (Twist).

The remaining Chapter 5 and Chapter 6 develop the applications of our repulsion costs and singularity functors to constructing finite-dimensional $E\Gamma$ models. In Chapter 5 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 5 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 §5.3. The excision construction, and its relation to Bieri-Eckmann's homological duality is described in §5.4 and summarized in Theorem 5.5.2. 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]. Again we emphasize Excision rather than Bordification. The final Chapter 6 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 6.2.1 and Theorem 6.2.4, 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 closed convex bodies on which we install the repulsion costs from previous Chapter 4, and can therefore implement the reduction-program detailed in Theorem C. We conclude with some basic examples of Closing Steinberg in the remaining §§6.3 and 6.4.

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. Alternatively the Hausdorff measures $\mathscr{H}^n$ of dimensions $n = \dim(X), \dim(Y)$ respectively. We abbreviate $\mathscr{H}_X = \mathscr{H}^{\dim(X)}$. We reserve σ and τ for Radon measures on X and Y, called "source" and "target" measures, respectively. Typically σ, τ are 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]. In this thesis a subvariety Z of X is a closed subset defined by a collection of explicit equations (c.f. §3.2). If the equations are described by Lipschitz (or DC) functions, then we call Z a Lipschitz (or DC) subvariety.

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, 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. We let F denote a geodesically convex compact subset in a complete riemannian manifold, and $\mathscr{E}$ denotes the extreme-point functor, where $\mathscr{E}[F]$ consists of the extreme-points on F.

The symbol Γ designates an infinite torsion-free discrete group. Following standard notation, $E\Gamma$ is the universal cover $\tilde{X}$ of an Eilenberg-Maclane space $X = K(\Gamma, 1)$. 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. A $conv\mathscr{E}$-chain sum is a simplicial chain sum, where each summand has image isometric to a compact convex set. A cubulation of a group Γ is an $E\Gamma$-model which has explicit cellular structure defined in terms of isometric identifications between cubes $I^n = [0,1]^n$ for $n \geq 0$, with additional "wall-structures". Precise definitions are given in §6.2 in terms of convex chain sums and "gates". See Chapters 5, 6.

Chapter 2

Semicoupling Programs and Topological Theorem A

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), where $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is a cost function satisfying various Assumptions (A0), (A1), ... introduced in Section 1.3.

We begin with the definition of semicoupling. A semicoupling is a Radon measure π on the Cartesian product $X \times Y$, and c-optimal semicouplings π_{opt} are extremal semicoupling measures with precisely peculiar geometry. 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 linear duality, 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 dual Kantorovich potentials $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ are c-concave functions, 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 §3.2 for explicit local equations.

Example. If σ,τ are uniform measures on $X, Y = \partial X$, and if X is aspherical and ∂X is homologically nontrivial, then Kantorovich's singularity functor $Z : 2^{\partial X} \to 2^X$ produces closed locally DC-subvarieties Z' of X for which the inclusions $Z' \hookrightarrow X$ are homotopy-isomorphisms. See Theorems 3.4.2, 3.4.3 for details.

Remark. Before we proceed to the main line of this thesis, we must first remark on the definition of "probability" and its relation to optimal transportation methods. Hitherto 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

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

But our thesis obtains an important computational advantage by ignoring the normalization-condition (2.1). Indeed we argue (2.1) is totally nonphysical. From the perspective of symbolic computational mathematics, one can never verify equation (2.1) for a given pair of measures σ, τ on pair of spaces X, Y. That is, the renormalization factor $(\int 1\mu)^{-1}$ can be exceedingly difficult to evaluate. Except a space be finite or compact, one cannot generally evaluate this necessary factor. Especially when error and noise is persistent, effective coupling theory is not possible because there do not exist couplings between measures σ and τ with respective total masses 1.02 and 0.99. Again not unless the spaces are basically compact. In this thesis, our applications to arithmetic groups (see §5) 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.

The present thesis develops in the *semicoupling category*, being based on the hypothesis that source measures σ be "abundant" with respect to a target measure τ. That is, for every Borel subset Y' of Y with $\tau[Y'] > 0$, there exists Borel subset X' of X such that $\sigma[X'] > \tau[Y']$. Evidently source and targets with respective mass 1.02 and 0.99 have this property, and the set of semicouplings $SC(\sigma, \tau)$ is nontrivial convex subset. Briefly we replace the nonphysical equality (2.1) with the physically stable inequality

$$\int_X 1.d\mu(x) - \int_Y d\tau(y) \geq 0. \tag{2.2}$$

So where couplings are nonexistant, semicouplings are abundant. However we also admit that the existence of the renormalization factor $\int_X 1 d\mu(x) < +\infty$ is an exceedingly 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 again illustrates the logical convenience of assuming an inequality $\int_X 1.d\mu(x) < +\infty$, rather than exactly evaluating some real number $1/\int_X d\mu(x)$.

2.1 Optimal Semicouplings and Cost

Now we introduce the optimal semicoupling program. We reserve X, Y for complete finite-dimensional manifolds-with-corners of topological dimensions d, e respectively. We abbreviate $\mathscr{H}_X^d$ and $\mathscr{H}_Y^e$. In general d, e are distinct, and our main motivation is studying $Y = \partial X$, and especially where $X = X[t]$ is a convex excision model, see Chapters4, §4.5. Let $proj_X : X \times Y \to X$, $proj_Y : X \times Y \to Y$ be the canonical continuous projections. We exclusively study Radon measures σ, τ on the source and target X, Y, and usually assume they are absolutely-continuous with respect to the uniform measures $dvol_X, dvol_Y$, respectively.

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 "$\leq$" says that for every Borel subset O the numerical inequality $(proj_X \# \pi)[O] \leq \sigma[O]$ holds. 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 $\sigma[X] \geq \tau[Y]$, 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 $\{\pi_k\}_{k=1,2,\ldots}$ 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 our earlier §1.3 we enumerated several assumptions our cost functions should possess, labelled (A0)–(A5). The Assumptions (A0)–(A3) are rather generic. The Assumption (A4) implies the general uniqueness of semicouplings, see Proposition 2.5.6. 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 (A5), and more precisely described by Condition (C) as defined in §2.6, 2.6.4. The Condition (C) is a "small-cancellation" hypothesis on local tangent vectors which ensures that specific "average vectors" are nonvanishing. These nonvanishing of these "average vectors" is necessary for the continuity of the deformation retracts constructed in Chapter 2.3.

The existence of c-optimal couplings (not semicouplings!) for costs satisfying $(A0)-$ $-(A2)$ is a standard consequence of Fatou's lemma and Prokhorov's precompactness theorem. If π is a Radon measure $X \times Y$, 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 $\int_Y 1 d\tau = 1$. Then $SC(\sigma, \tau)$ is compact with respect to the narrow-topology. When c is continuous cost, then c-optimal semicouplings π_{opt} exists 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)**.** When $\psi : Y \to \mathbb{R} \cup \{-\infty\}$ is a 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 $x \in X$. When $\phi : X \to \mathbb{R} \cup \{+\infty\}$ is a function on the source, then the c-Legendre transform $\phi^c : Y \to \mathbb{R} \cup \{-\infty\}$ is defined by the rule $\phi^c(y) = \inf_{x \in X}[c(x, y) + \phi(x)]$.

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 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 fundamental.

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 $y \in Y$,

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

These definitions of c-convexity and c-subdifferentials are fundamental to this thesis. Our 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. Recall the domain of ϕ is defined $dom(\phi) := \{x \in X \ |\phi(x) < +\infty\}$.

Lemma 2.3.4. *Let c be 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 ψ is an upper semicontinuous function, and $\partial^c\psi(y)$ is a nonempty closed topological subset of X for every $y \in dom(\psi)$; and ϕ is lower semicontinuous function, and $\partial^c\psi(y)$ is a nonempty closed subset of Y for every $x \in dom(\phi)$.

Proof. The Assumptions (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 $y \mapsto c(x, y) + \psi^c(x)$. An analogous argument shows the c-transform $\psi^c(x)$ is a lower semicontinuous function of the source variable x.

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 $\{x_j\}_{j=1,2,...}$ 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 above and 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

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

However 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.

□

Several further properties of c-convex potentials are developed in Section 2.5 below, c.f. Lemmas 2.5.2, 2.5.4, and Proposition 2.5.5.

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

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

Kantorovich duality says the inequality (2.5) is an equality where "sup = inf". The equality "sup = inf" 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):

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

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

The Assumption (A0) that our costs c are continuous implies the answers to (i) and (ii) 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 nonnegatively-valued continuous cost. Then*

(1) there is no duality gap between the primal program and the dual program, and

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

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

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

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:

(i) $\psi_0 \leq \psi_+$ pointwise throughout Y; and

(ii) $\phi_0 \geq \phi_+$ pointwise throughout X; and

(iii) $\phi_0(x) = \phi_+(x)$ whenever there exists $y \in Y$ with $x \in \partial^{c_+}\psi_+(y)$; and

(iv) $\partial^{c_+}\phi_+(x) \subset \partial^c\phi_0(x)$ 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

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

Then $-\phi_0, \psi_0$ are c-concave with $\phi_0 = \psi_0^c$. As we assume c satisfies $(A0)--(A4)$, we can now develop a perturbation argument based on [McC01, Proof of Thm. 9] and prove the existence of c-optimal semicouplings between σ and τ. We recall that Assumptions (A0)–(A2) imply c-convex potentials ϕ are differentiable $\mathscr{H}_X$-almost everywhere on their domains. With further Assumptions (A3), (A4), we use (Twist) condition to characterize the transport as a measurable mapping

$$T : dom(\phi) \to Y, \quad T(x) := \nabla_x c(x,\cdot)^{-1}(-\nabla_x\phi(x)),$$

for $x \in dom(D\phi)$. The main identity implied by $(A0) - -(A4)$ is

$$T(x) \in \partial^c\phi(x)$$

for almost-every $x \in dom(\phi)$, where $\phi = \psi^c$ is c-convex.

The perturbation argument begins with a continuous function $h : Y \to \mathbb{R}$. For $\epsilon > 0$, define $\psi_\epsilon(y) := \psi(y) + \epsilon.h(y)$. Then

$$\phi_\epsilon(x) = \sup_{y\in Y}[\psi(y) + \epsilon.h(y) - c(x,y)] \tag{2.8}$$

for $x \in A$. If we presuppose that Y is compact and c is continuous, then the supremum defining $\phi_\epsilon(x)$ would obviously be attained by some $y_\epsilon = y_\epsilon(x) \in Y$. This remains true if c satisfies Assumptions (A0)–(A4), as demonstrated in the following

Lemma 2.3.7. *Let c satisfy Assumptions (A0)–(A4). Let $\phi = \psi^c$ be a c-convex potential.*

Then almost-every $x \in dom(\phi)$ has nonempty c-subdifferential $\partial^c\phi(x)$, and the supremum defining $\phi(x)$ in equation (2.8) is attained.

Proof. Lemma 2.5.4 and Proposition 2.5.5 proves $\phi = \psi^c$ is locally semiconvex on $dom(\phi)$ and almost-everywhere uniquely differentiable. But $dom(\phi)$ coincides with the image $Z_1 := \cup_{y\in Y}\partial^c\psi(y)$. The (Twist) condition ensures $T(x) \in \partial^c\psi^c(x)$ for almost-every $x \in dom(\phi)$, recall equation (2.13). □

The supremum defining $\phi_\epsilon(x)$ is attained under Assumptions (A0)–(A4), and we denote this maximizer $y_\epsilon = y_\epsilon(x)$. Indeed the supremum defining (2.8) can be restricted to a compact subset of Y and the function $y \mapsto \psi(y)+\epsilon.h(y)-c(x,y)$ is upper semicontinuous. These two conditions imply the supremum is attained.

At $x \in X$ where the supremum is uniquely attained, sufficiently small ϵ will imply

$$y_\epsilon = T(x) + o(1).$$

Thus

$$\phi_\epsilon(x) = \psi(y_\epsilon) + \epsilon.h(y_\epsilon) - c(x, y_\epsilon).$$

But

$$\psi(y_\epsilon) - c(x, y_\epsilon) \leq \psi(T(x)) - c(x, T(x)),$$

and

$$\psi(T(x)) + \epsilon.h(T(x)) - c(x, T(x)) \leq \phi_\epsilon(x) \leq \psi(T(x)) - c(x, T(x)) + \epsilon.h(y_\epsilon).$$

From this last inequality, we deduce

$$\phi_\epsilon(x) = \phi_0(x) + \epsilon.h(T(x)) + \epsilon.o(1).$$

Consider the one-parameter family of potentials $(\phi_\epsilon, \psi_\epsilon)$. If we define

$$J(\phi, \psi) := \int_A (-\phi(x))d\sigma(x) + \int_Y \psi(y)d\tau(y),$$

then (ϕ_0, ψ_0) maximizing Kantorovich's dual program implies

$$\lim_{\epsilon\to 0} \frac{J(\phi_\epsilon, \psi_\epsilon) - J(\phi_0, \psi_0)}{\epsilon} = 0.$$

Evaluating this single-variable derivative, we find $\lim_{\epsilon\to 0}(J(\phi_\epsilon,\psi_0\epsilon)-J(\phi_0,\psi_0)/\epsilon$

$$=\lim_{\epsilon\to 0}-\int_A\frac{\phi_\epsilon(x)-\phi_0(x)}{\epsilon}d\sigma(x)+\int_Y h(y)d\tau(y)=-\int_X h(T(x)).d\sigma(x)+\int_Y h(y)d\tau(y),$$

according to Lebesgue's dominated convergence theorem. Next the change-of-variables formula implies $\int_A h(T(x)).d\sigma(x)=\int_X h(y).d(T\#1_A.\sigma)(y)$, and therefore

$$-\int_X h(y).d(T\#1_A.\sigma)(y)+\int_Y h(y)d\tau(y)=0.$$

Since h is an arbitrary continuous function on Y, this implies $T\#1_A.\sigma=\tau$.

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*

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

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

$$\sigma[\{t_1 < c_y < t_2\}] > 0 \tag{2.9}$$

for every connected interval $[t_1, t_2] \subset spt(c_y \# \sigma)$. But by definition of support, the strict-positivity of (2.9) 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*

$$A := \cup_{y \in Y} \{x | c(x, y) \leq t(y)\}. \tag{2.10}$$

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 $1_A.\sigma[\{c_y > t\}] > 0$. □

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

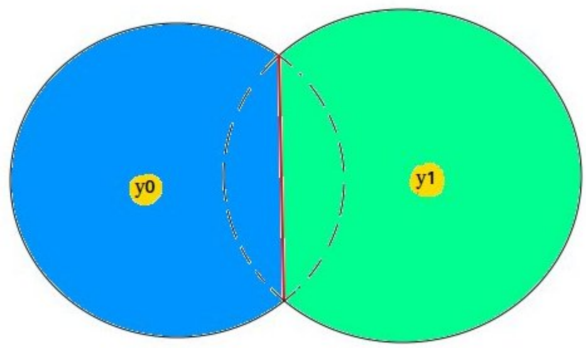

Figure 2.2: Connected Active Domain when $mass[\sigma]/mass[\tau] \approx 1^+$

All the ideas of this section are pre-Kantorovich duality, to which we turn next.

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 describes Assumption (A4) from Section 1.3:

Definition 2.5.1 ((Twist)). Let c be cost function satisfying Assumptions (A0) and (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 $\nabla_x c(x', \cdot) : dom(c_{x'}) \to T_{x'}X$.

We remark that (Twist) condition is equivalent to the function $x \mapsto c_\Delta(x; y_0, y_1)$ 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 4 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).

□

Let the reader recall the definition of semiconvexity, c.f. [Vil09, Definition 10.10, pp.228].

Definition 2.5.3 (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

$$\phi(\gamma(t)) \leq (1-t)\phi(\gamma(0)) + t\phi(\gamma(1)) + t(1-t)Cdist(\gamma(0), \gamma(1))^2 \tag{2.11}$$

is satisfied for $0 \leq t \leq 1$. The function is locally semiconvex if ϕ is semiconvex at every $x_0 \in U$ and the modulus $C > 0$ depending uniformly on $\gamma(0)$, $\gamma(1)$ which vary in a compact subset K of U.

Lemma 2.5.4. *Let $c : X \times Y \to \mathbb{R}$ be cost satisfying Assumptions (A0)–(A2). Then every c-convex potential ψ^c defined on X is $\mathscr{H}_X^n$-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 $\psi^c(x) = \sup\{\psi(y) - c(x,y) |\ y \in Y\}$ for every $x \in X$. Assumption (A0) implies 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 $Y' \subset Y$, where $Y'(x)$ varies with x. From Lemma 2.5.2 the family of functions $\{x \mapsto \psi(y) - c(x,y) | y \in dom(c_x)\}$ 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 $\sup_{y \in Y'(x)} L(y)$ is attained and finite over the compact $Y'(x)$. This implies ψ^c is locally Lipschitz, with finite Lipschitz constant as satisfied by the family $\{\psi(y) - c(x,y) |\ y \in Y'(x)\}$. 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 $\{\psi(y) - c(x,y) | y \in Y'(x)\}$. □

Proposition 2.5.5. *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 2.5.4, the c-convex potentials ψ^c are locally Lipschitz on their domains $dom(\psi^c) \subset X$. Then Rademacher's theorem, e.g. [Vil09, Thm 10.8, pp.222], says locally Lipschitz functions are almost-everywhere differentiable on $dom(\psi^c)$ with respect to volume measure $\mathscr{H}_X$ on their domains. Therefore $\nabla_x \psi^c$ exists almost everywhere on $dom(\psi^c)$ as desired. □

Given Lemmas 2.5.2, 2.5.4, 2.5.5, we can now describe how Assumption (A4), namely (Twist), leads to the uniqueness of optimal semicouplings. The Lemma 2.5.5 states ψ^c is $dvol_X$-almost everywhere differentiable on its domain $dom(\psi^c)$. Let $dom(D\psi^c)$ denote the domain of differentiability in $dom(\psi^c)$. So $dom(D\psi^c)$ is full-measure subset of $dom(\psi^c)$. Next consider the inequality

$$\psi^c(x) - \psi^c(x') \leq c(x', y) - c(x, y) \tag{2.12}$$

which holds almost-everywhere through the support of the optimal semicoupling. If $x = x_0$ belongs to $dom(D\psi^c)$, then inequality (2.12) becomes an equation in y, namely

$$\nabla_x \psi^c(x_0) + \nabla_x c(x_0, y) = 0.$$

The equation characterizes membership $y \in \partial^c \psi^c(x_0)$. Under the (Twist) condition, the rule

$$T(x) := \nabla_x c(x, \cdot)^{-1}(-\nabla_x \psi^c(x)) \tag{2.13}$$

is a well-defined map $T : dom(D\psi^c) \to Y$. So T is a map defined $\mathscr{H}_X$-almost everywhere on $dom(\psi^c)$.

Under the above hypotheses, we find T defines a Borel-measurable map $T : dom(\psi^c) \to Y$ which pushes forward the restricted source $1_{dom(\psi^c)}\sigma$ to τ. Thus the pushforward $(Id \times T)\#1_{dom(\psi^c)}\sigma$ defines a semicoupling π^* in $SC(\sigma, \tau)$ which minimizes the Monge-transportation program. Therefore the graph of T defines a semicoupling from σ to τ, namely $\pi := (Id, T)\#(1_{dom(\psi^c)}\sigma)$. But we find π is necessarily c-optimal, since $-\psi^c(x) + \psi(y) = c(x, y)$ holds almost everywhere on the support of π, and therefore "sup=inf" in (2.5). This discussion leads to standard uniqueness result, assuming the (Twist), c.f. [Vil09, Thms 10.28, 10.42].

Proposition 2.5.6. *Suppose cost c satisfies Assumptions (A0)–(A4). Let source σ be absolutely continuous with respect to the volume measure $\mathscr{H}_X^d$ on the source X. Then there exists unique c-optimal semicoupling (modulo sets of measure zero) between σ and τ.*

Proof. By Theorem 2.4.7 we know Assumptions (A0)–(A4) ensure there exists a unique active domain for 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-concave maximizers ψ, ψ^c on Y, X, respectively. By Lemma 2.5.5 the c-convex potential ψ^c on the source is almost-everywhere differentiable on its domain $dom\psi^c \subset X$. From Theorem 2.4.7 we know $dom\psi^c$ coincides, modulo sets of zero measure, with the active domain A.

Assuming (Twist) condition, the discussion above shows there exists measurable map $T : dom\psi^c \to Y$ whose graph $(Id, T)\#(1_{dom\psi^c}\sigma)$ defines a semicoupling between σ, τ, and this semicoupling is c-optimal. (See (2.13) for definition of T). 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 the above reasoning shows any c-optimal semicoupling is supported on the graph of a measurable map. So if π, π' are c-optimal, then their convex combination $\frac{1}{2}\pi + \frac{1}{2}\pi'$ is again c-optimal. But this convex combination cannot be supported on the graph of a measurable function, unless the graphs supporting π, π' coincide almost-everywhere and then π, π' coincide almost-everywhere.

□

2.6 Deforming $X \rightsquigarrow A$: Topological Theorem A

The present section is topological, and it's main result is Theorem 2.6.2. A cost satisfying (A0)–(A3) has a uniquely defined closed active domain $A \subset X$ (see Lemma 2.4.6 and equation (2.10) below). Our goal is to identify conditions for which the inclusion $A \hookrightarrow X$ is a homotopy-isomorphism. To anticipate some notation from §2.3, we consider the source space $X =: Z(\emptyset)$, and define the activated source $A =: Z_1$. Specifically we prove that sufficient Halfspace conditions imply $A \hookrightarrow X$ is a homotopy-isomorphism when cost satisfies Assumptions (A0)–(A5). 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 $Y_I \hookrightarrow Y_J$.

Definition 2.6.1. A collection $E = \{\eta_i | i \in I\} \subset T_xX$ of tangent vectors satisfies the **Halfspace condition** if there exists a nonzero linear functional $\ell : T_xX \to \mathbb{R}$ with $\ell(\eta_i) > 0$ simultaneously for all $i \in I$.

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

Theorem 2.6.2. *Let $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be cost satisfying Assumptions (A0)–(A4). Let $A \hookrightarrow X$ be the closed activated source of a c-optimal semicoupling. Suppose every*

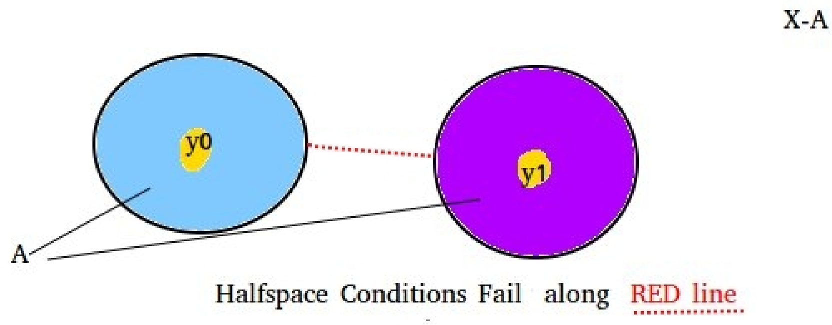

Figure 2.3: Halfspace Conditions fail and active domain is not homotopy-equivalent to source.

$x' \in X - A$ has the property that the collection

$$\{\nabla_x c(x', y_*) \mid y_* \in dom(c_{x'})\} \subset T_{x'}X$$

satisfies Halfspace Condition. Furthermore suppose the gradients

$$\{\nabla_x c(x', y_*) \mid y_* \in dom(c_{x'})\}$$

are uniformly bounded away from zero over $X - A$. 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 $A = h(X, 1)$.

Without reparameterizing the flows Ψ by arclength, our proof of Theorem 2.6.2 demonstrates that the disactivated domain $X - A$ is a deformation retract of every ϵ-neighborhood of A in X. That is X deformation retracts onto

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

for every $\epsilon > 0$. The interested reader may compare our method to [Nee85, §3]. The purpose of Lemma 2.6.8 and Assumption (A5) is to obtain continuous deformations at "$\epsilon = 0$", and the condition that the gradients be uniformly bounded away from zero is a form of Assumption (A5). Our strategy is to study the gradient flow of an "averaged" vector field $\eta(x, avg) = \nabla_x f_{avg}$ generated by an "averaged" potential f_{avg}. We establish Assumption (A5) by relating the rate-of-divergence of f_{avg} near a pole to the rate-of-divergence of a collection $\{f_* | y_* \in Y\}$ of related potentials.

We proceed to the proof of Theorem 2.6.2. Let Y be compact with volume measure $\mathscr{H}_Y$. Suppose c satisfies Assumptions (A0)–(A4). For simplicity we will assume

$dom(c_x) = Y$ for every $x \in X$. Let $t : Y \to \mathbb{R}$ be $\mathscr{H}_Y$-measurable function and $A := \cup_{y\in Y}\{c(x,y) \le t_y\}$. For every $y \in Y$, the function $x \mapsto \log(c(x,y) - t_y)$ is finite on $X - A$, and diverges to $-\infty$ as x converges to a limit $x_\infty \in \{c(x,y) \le t_y\} \subset A$. Likewise the gradient $\nabla_x \log(c(x,y) - t_y)$ diverges to point-at-infinity when $x \in X - A$ converges to a limit $x_\infty \in \{c(x,y) \le t_y\} \subset A$. Next we take the average gradient with respect to $\mathscr{H}_Y$, and claim this averaged-gradient is indeed the gradient field of a continuously differentiable "averaged" potential.

Proposition 2.6.3. *Let $\nu_1, \nu_2, \nu_3, \ldots$ be a sequence of empirical measures, i.e. rescaled sums of atomic Dirac measures, which converges as $n \to +\infty$ in the weak-$*$ topology to the renormalized Hausdorff measure $(\mathscr{H}_Y[Y])^{-1} \cdot \mathscr{H}_Y$ on Y.*

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

$$\lim_{N\to+\infty} \int_Y \log(c(x,y) - t(y)) d\nu_N(y)$$

exists and converges to the finite integral

$$f_{avg}(x) := (\mathscr{H}_Y[Y])^{-1} \int_Y \log(c(x,y) - t_y).d\mathscr{H}_Y(y).$$

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

$$\nabla_x f_{avg} = (\mathscr{H}_Y[Y])^{-1} \int_Y \nabla_x \log(c(x,y) - t_y).d\mathscr{H}_Y(y).$$

Proof. The cost c satisfies (A0)–(A4) implies the limit defining f_{avg} converges uniformly on compact subsets of $X - A$. So the limit in (i) 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 averaged-gradient of $\nabla_x \log(c(x,y) - t_y)$ with respect to $\mathscr{H}_Y$. $\square$

Next we fix $y_0 \in Y$, abbreviating $t_0 = t(y_0)$ and $f_0(x) := \log(c(x,y_0) - t_0)$. Suppose $\{x_j\}$ is sequence in $X - A$ which converges to $x_\infty \in \{f_0 = -\infty\} \subset A$. Note $\{f_0 = -\infty\} = \{x | c(x,y_0) \le t_0\}$. We want to compare the asymptotics of f_{avg} and f_0 in neighborhoods of $\{f_0 = -\infty\}$. It is elementary to upperbound the averaged gradient

$$||\nabla_x f_{avg}|| \le (\mathscr{H}_Y[Y])^{-1} \int_Y ||\nabla_x f_*|| d\mathscr{H}_Y(y_*), \tag{2.14}$$

using triangle-inequality. The inequality (2.14) says the averaged gradient diverges only if there is a measurable subset of Y where the integrand diverges. Our applications require the converse: the divergence of at least one gradient $\nabla_x f_*$ should imply the divergence of the average gradient. This converse motivates our Assumption (A5). Informally we require that there is small cancellation between the terms $\{\nabla_x f_* | y_* \in dom(c_x)\}$. Then the "large" gradients $\nabla_x f_*$ will have "large" average with respect to $\mathscr{H}_Y$. Here is formal definition.

Definition 2.6.4 (Property (C)). The collection of functions $\{f_* | y_* \in Y\}$ satisfies Property (C) throughout $X - A$ with respect to the uniform probability measure $(\mathscr{H}_Y[Y])^{-1}\mathscr{H}_Y$ if there exists constant $C > 0$ such that

$$||\nabla_x f_{avg}|| \geq C \int_Y ||\nabla_x f_*||.d\mathscr{H}_Y(y_*) \tag{2.15}$$

pointwise throughout $X - A$.

The constant $C > 0$ in the definition must depend independantly of x. When Y is finite, the estimate (2.15) requires the ratio

$$||\nabla_x f_{avg}||/\max ||\nabla_x f_*||$$

be uniformly bounded away from zero throughout $X - A$. Together Property (C) 2.6.4 and Halfspace Condition 2.6.1 ensure the divergence of any gradient $\nabla_x f_*$ implies the divergence of the average $\nabla_x f_{avg}$. Hence whenever a sequence $\{x_j\}$ converges in $X - A$ to $\{f_0 = -\infty\}$, we find f_{avg} and f_0 diverging asymptotically $f_{avg} \asymp f_0$ to $-\infty$.

Proposition 2.6.5. *Abbreviate $t_* = t(y_*)$, $f_*(x) := \log(c(x, y_*) - t(y_*))$, and*

$$M(x) := \max\{||\nabla_x f_*|| \mid y_* \in Y\}.$$

Suppose the collection $\{f_* | y_* \in Y\}$ satisfies Property (C) and (HS) Conditions throughout $X - A$. Then $||\nabla_x f_{avg}|| \asymp M$ in neighborhoods of the poles of f_{avg}, and there exists two positive constants $c, C > 0$ such that

$$0 < c \leq ||\nabla_x f_{avg}||/M(x) \leq C$$

in neighborhoods of ∂A in $X - A$.

Proof. The proposition is consequence of (2.15). If x is a sequence of points converging to A, then there exists an open subset V of Y such that f_0 and $\nabla_x f_0$ diverge along the sequence for every $y_0 \in V$. Then inequality (2.15) proves the average-gradient and $\nabla_x f_0$ have asymptotic norms. $\square$

The above Proposition 2.6.5 shows the average f_{avg} and f_0 diverge asymptotically ($f_{avg} \asymp f_0$) near the poles $\{f_0 = -\infty\} \subset \{f_{avg} = -\infty\}$. Notice $\{f_{avg} = -\infty\} = A$. Thus we find the average is approximated by its maximizing terms in neighborhoods of the poles. This ensures a strong form of Assumption (A5) holds true, i.e. the average gradient $\nabla_x f_{avg}$ is surely bounded away from zero in neighborhoods of the poles $\{f_{avg} = -\infty\}$.

The following lemma proves f_0 is strongly concave in neighborhoods of the poles $\{f_0 = -\infty\}$.

Lemma 2.6.6. *Suppose $\nabla^2_{xx} c(x, y_0)$ is uniformly bounded above and below throughout $X - A$. Then for every $K < 0$, there exists $\epsilon > 0$ such that $\nabla^2_{xx} f_0 \leq K.Id < 0$ throughout the ϵ-neighborhood of $\{f_0 = -\infty\}$ in $X - A$.*

Proof. The standard differentiation formulas show $\nabla^2_{xx} f_0$ is equal to

$$\frac{-1}{(c(x,y_0) - t_0)^2} \nabla_x (c(x, y_0) - t_0) \otimes \nabla_x (c(x, y_0) - t_0) + \frac{1}{c(x, y_0) - t_0} \nabla^2_{xx} (c(x, y_0) - t_0).$$

By Proposition 2.6.4 the gradients $\nabla_x (c(x, y_0) - t_0)$ are uniformly bounded away from zero in neighborhoods of $\{f_0 = -\infty\}$. The assumption (A2) says $\nabla^2 c(x, y_0)$ is uniformly bounded above and below, and implies

$$\frac{-1}{(c(x,y_0) - t_0)} \nabla_x (c(x, y_0) - t_0) \otimes \nabla_x (c(x, y_0) - t_0)$$

diverges to $-\infty$ pointwise. Therefore $\nabla^2 f_0 \leq K.Id < 0$ when $c(x, y_0) - t_0 > 0$ is sufficiently small. But the level sets of $\{x | t_0 \leq c(x, y_0) \leq t_0 + \epsilon'\}$ are compact subsets of $X - A$ for every $\epsilon' > 0$. This implies a sufficiently small $\epsilon > 0$ exists for which $\nabla^2 f_0 \leq K.Id < 0$ throughout the ϵ-neighborhood of $\{f_0 = -\infty\}$ as desired. $\square$

Example. The function $x \mapsto \log(x)$ satisfies conditions of the Lemma 2.6.6. The gradient flow $x' = -1/x$ has $\nabla_x(-\log x) = -1/x$ uniformly bounded away from zero in neighborhoods of the pole at $x = 0$. Indeed, we plainly see the gradient diverges to $+\infty$ when $x \to 0^+$. Explicitly for initial value $x_0 > 0$, we find $x(s) = \sqrt{x_0^2 - 2s}$ defines the integral curve for the gradient flow which blows-up in finite-time on $0 \leq s < x_0^2/2$.

The blow-up in finite time is typical property of the gradient flow defined by the potentials f_{avg} and f_0.

Lemma 2.6.7 (Finite-time Blow-up)**.** *Suppose the functions $\{f_* | y_* \in Y\}$ satisfy (2.15) 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 (2.15) 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 concave in neighborhoods of the poles. This implies the gradient flow blows-up in finite-time and every integral curve has a finite interval of existence. □

Informally the estimate (2.15) implies every step in the discretized gradient flow (e.g., Euler scheme) has a definite size. The asymptotic concavity of Lemma 2.6.6 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. In fact the points x_0 approach the poles with accelerating velocity.

Now we establish Theorem 2.6.2.

Proof of Theorem 2.6.2. According to Lemma 2.4.6 the active domain A can be expressed as $\cup_{y_* \in Y}\{x \mid c(x, y_*) \leq t(y_*)\}$ for a measurable function $t : Y \to \mathbb{R}$. 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

$$\eta(x, y_*) := (c(x, y_*) - t_*)^{-1} \cdot \nabla_x c(x, y_*),$$

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

$$\eta(x, avg) := N^{-1} \sum_{y_* \in Y} [(c(x, y_*) - t_*)^{-1} \cdot \nabla_x c(x, y_*)]. \tag{2.16}$$

Evidently when Y is finite, the sum (2.16) is finite vector. Then $x \mapsto \eta(x, avg)$ is a well-defined nonvanishing vector field on $X - A$ which diverges to the point-at-infinity whenever a denominator converges $c(x, y_*) \to t_*^+$. Thus $\eta(x, avg)$ is finite if and only if $x \in X - A$.

We propose integrating the vector field $x \mapsto \eta(x, avg)$ throughout $X - A$ to define the desired retraction. Thus we obtain a 1-parameter family of diffeomorphisms

$$\Psi : (X - A) \times [0, +\infty) \to X - A,$$

where $\Psi(x, s)$ is the unique solution of the ordinary differential equation

$$\frac{d}{ds}|_{(x', s')}\Psi = -\eta(\Psi(x', s'), avg), \quad \Psi(x', 0) = x' \tag{2.17}$$

for all $(x', s') \in (X - A) \times [0, +\infty)$. According to Lemma 2.6.7, the flow Ψ defined by the gradient $\eta(x, avg)$ blows-up in finite-time, and therefore every initial value $x_0 \in X - A$ determines a unique maximal interval of existence $I(x_0)$. Therefore orbits $\{\Psi(x', s) | s \in \mathbb{R}_{\geq 0}\}$ converge in finite-time to the boundary $\partial(X - A)$ for every choice of initial value $x_0 \in X - A$.

The standard results on continuous dependance on initial conditions and parameters proves the convergence-time, and the maximum interval of existence $I(x_0)$, varies continuously with respect to x_0. According to Lemma 2.6.8 below, the orbits can be continuously reparameterized according to arclength to obtain a continuous mapping

$$\Psi : (X - A) \times [0, +\infty] \approx (X - A) \times [0, 1] \to X.$$

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

$$\eta(x, avg) := (\int_{dom(c_x)} d\mathscr{H}_Y)^{-1} \int_{dom(c_x)} (c(x, y_*) - t_*)^{-1} \cdot \nabla_x c(x, y_*) d\mathscr{H}_Y(y_*), \tag{2.18}$$

where $dom(c_x)$ is closed compact subset of Y for every $x \in X$ by Assumption (A0). According to Lemma 2.6.3 above the vector field $\eta(x, avg) = \nabla_x f$ is the gradient of a continuously differentiable potential f defined on $X - A$.

The proof proceeds as in (Case I). The vector field $x \mapsto \eta(x, avg)$ is well-defined nonvanishing vector field on $X - A$ which diverges to $+\infty$ whenever some denominator converges $c(x, y_*) \to t_*^+$. We integrate the gradient fields and obtain the retraction of $X - A$ onto the poles ∂A. The flow converges in finite-time by Lemma 2.6.7, and we can reparameterize according to arclength to obtain a continuous deformation retract by Lemma 2.6.8 below. $\square$

Strictly speaking, the definition of deformation retract requires we reparameterize the flows according to arclength. So the proof of Theorem 2.6.2 formally requires Lemma 2.6.8 below to obtain a continuous retract

$$\Psi' : (X - A) \times [0, 1] \to X - A.$$

The continuous deformation retract Ψ' is defined by the identity

$$\Psi'(x_0, s) = \Psi(x_0, S)$$

where $S \in [0, +\infty)$ is the unique real parameter satisfying the equation

$$\int_0^S ||\frac{d}{dt}\Psi(x_0, t)||dt = s.\ell(x_0).$$

Lemma 2.6.8. *Assume the cost c satisfies Assumptions (A0)–(A5). For initial value $x_0 \in X - A$, let $\Psi : (X - A) \times [0, +\infty) \to X - A$ be the flow with $\Psi(x_0, 0) = x_0$ satisfying (2.17). Define the arclength $\ell(x_0)$ of the forward-time trajectory*

$$\ell(x_0) := \int_0^{+\infty} ||\frac{d}{dt}\Psi(x_0, s)||ds. \tag{2.19}$$

Then

(i) the arclength is finite $\ell(x_0) < +\infty$ for every $x_0 \in X - A$; and

(ii) the arclength $x_0 \mapsto \ell(x_0)$ varies continuously with respect to the initial value $x_0 \in X - A$.

Proof. First we prove (i). In both (Case I) and (Case II) the vector field $\eta(x, avg)$ is an averaged sum of gradient fields (see (2.16)). For brevity we write $\eta(x, avg) = \nabla f_{avg}$.

Now Lemma 2.6.7 says the flow blows-up to infinity in finite time. Therefore the arclength $\ell(x_0)$ defined in (2.19) is supported over a finite interval $\int_0^{+\infty} = \int_{I(x_0)}$, where $I(x_0)$ is the maximal interval of existence for the integral curve of (2.17) with initial value $x(0) = x_0$. Since $||\frac{d}{dt}\Psi(x_0, s)||$ surely varies continuously with respect to the time parameter s, we conclude $\ell(x_0)$ is finite.

To prove (ii), we need recall the standard "continuous dependance" theorems of ordinary differential equations. The average-gradient $\nabla_x f_{avg} = \eta(x, avg)$ varies continuously over compact subsets of $X - A$. Our hypotheses ensure the maximal interval of existence $I(x_0)$ varies continuously with respect to x_0. Compare [Har64, Theorem 2.1, pp.94] for further details.

□

Chapter 3

Kantorovich Singularity and Topological Theorem B

The previous Chapter 2 developed the measure-theoretic background relevant to optimal semicoupling program, and we concluded with our topological Theorem A, see Theorem 2.6.2. Theorem A describes a topological homotopy-isomorphism $A \subset X$, where A is the unique open activated domain, see §2.4, and where so-called HalfSpace Conditions and Property (C) are satisfied, see Definition 2.6.4. The present chapter further develops these ideas, and describes large codimension deformation retracts and topological properties of the singularities of optimal semicouplings. This chapter presents the central definition of our thesis, namely Kantorovich's contravariant singularity functor. We establish the basic topological properties 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 and τ on X and Y, respectively. The Property (C) is generalized to our so-called Uniform Halfspace (UHS) condition.

3.1 Kantorovich's Contravariant Singularity Functor

Our thesis proposes a bridge between measure and algebraic topology. The bridge is realized by the contravariant functor $Z : 2^Y \to 2^X$, whose formal topological definition is the following. In the topological category we represent $2^X := Hom_{TOP}(X, 2)$, where $2 = \{0, 1\}$ is the two-pointed set equipped with non-Hausdorff topology of open sets $\{\emptyset, \{0\}, \{0, 1\}\}$. Then $Hom_{TOP}(X, 2)$ designates all continuous mappings $f : X \to 2$. In otherwords 2^X is the category of all closed subsets of X (rather than the set of all subsets of 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 $dx := \mathscr{H}^d_X$, $dy := \mathscr{H}^e_Y$ on X, Y respectively. By Theorems 2.4.7 and 2.5.6 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. maximizers in dual maximization program (2.5)) are generally nonunique.

Definition 3.1.1 (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 contravariant functor $Z : 2^Y \to 2^X$, in that morphisms $Y_I \hookrightarrow Y_J$ are transformed 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 establish formal definition of support.

Definition 3.1.2. The support of the contravariant functor $Z : 2^Y \to 2^X$ is the subcategory of 2^Y, denoted $spt(Z)$, whose objects are those 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 3.1.1.

For given $x \in X$, we view $Z'(x) := Z(\partial^c\psi^c(x))$ as defining a "cellular neighborhood" of x, or more specifically on the active domain $A = Z_1 = \cup_{y \in Y} Z(y)$. If the cell $Z'(x)$ is not empty, and $\partial^c\psi^c(x)$ belongs to $spt(Z)$, then $Z'(x)$ is a manifold-with-corners. The c-concavity $\psi^{cc} = \psi$ provides explicit equations describing $Z'(x)$. See Section 3.2 and (3.2).

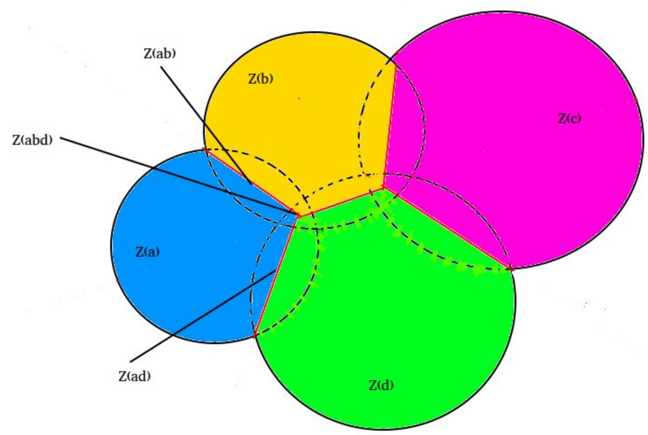

Figure 3.1: Singularity structure Z on active domain $A = Z_1$

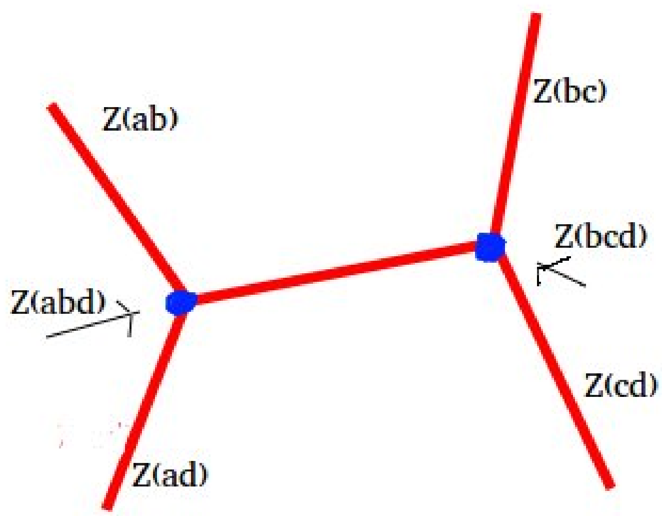

Figure 3.2: Singularity structure on Z_2

3.2 Local Topology and Local Dimensions of Z

The present section defines 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)–(A4). 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 3.2.4 and 3.3.5. The proof of 3.2.4 is essentially a local adaptation of the recent work of [KM18], and 3.3.5 is then obtained as corollary to theorem of Alberti [Alb94]. We are grateful to Prof. R. J. McCann for many useful conversations.

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

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

for all $(x, y) \in X \times Y$, with equality $\psi(y) - \psi^c(x) = c(x, y)$ if and only if $y \in \partial^c \psi^c(x)$, and if and only if $x \in \partial^c \psi(y)$. The pointwise inequality says $x \in \partial^c \psi(y)$ if and only if

$$y \in argmax[\{\psi(y_*) - c(x, y_*) + c(x, y) | y_* \in Y\}]. \tag{3.1}$$

Abbreviate $c_\Delta(x; y, y') := c(x, y) - c(x, y')$ for the two-pointed cross difference. For every $x_0 \in X$, let $\partial^c \psi^c(x_0) \subset Y$ be the c-subdifferential of ψ^c, and abbreviate $Z'(x_0) := Z(\partial^c \psi^c(x_0))$ be the closed subvariety defined by the singularity functor Z. The singularity $Z'(x_0)$ is described by the system of equations

$$Z'(x_0) = \{x \in X \mid 0 = \psi(y_0) - \psi(y) - c_\Delta(x; y_0, y), \ y \in \partial^c \psi^c(x_0), y \neq y_0\} \tag{3.2}$$

according to (3.1).

From the expression (3.2) we obtain the following

Lemma 3.2.1. *Under Assumptions (A0)–(A3), the singularity $Z'(x) := Z(\partial^c \psi^c(x))$ is local DC-subvariety in X for every $x \in X$.*

Proof. 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 $C > 0$ such that $||\nabla^2_{xx} c(x, y)|| \geq C > 0$ 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 (3.2) implies $Z'(x)$ is locally-DC subvariety of X.

□

The reader should observe Lemma 3.2.1 does not the specify the dimension of $Z'(x)$. In the remaining section, we study this question of dimension.

Definition 3.2.2. Abbreviate $A(x, y) := A(x, y, y_0) = c_\Delta(x, y, y_0) - (\psi(y) - \psi(y_0))$.

A coarse dimension estimate follows from a calculus computation describing the space-of-directions to $Z'(x)$ at x.

Lemma 3.2.3. *Let $x \in X$ be supported on the active domain of a c-optimal semicoupling, with $Z'(x) := Z(\partial^c\psi^c(x))$. Then the space of directions $T_xZ'(x)$ can be characterized as the orthogonal complement*

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

in T_xX.

Proof. Again we examine the equality-case of $A \leq 0$. 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 if and only if $\eta \in T_xX$ 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 result follows. □

The Lemma 3.2.3 indicates the expected Hausdorff dimension of $Z'(x)$, namely the dimension of the orthogonal complement $\{\nabla_x c_\Delta(x, y, y')|y, y' \in \partial^c\psi^c(x)\}$. With Halfspace conditions and ideas from [KM18], we confirm this dimension estimate in Proposition 3.2.4 below. Some preliminary notation will be convenient. Recall X is a Riemannian manifold-with-corners, equipped with riemannian exponential mapping $exp_x : T_xX \to X$. If (X, d) is complete Cartan-Hadamard space, then exp_x is diffeomorphism between T_xX 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. Abbreviate $Z_1 := dom(\psi^c)$ and let $x \in dom(\psi^c)$.

Proposition 3.2.4. *Suppose the gradients $\{\nabla_x c_\Delta(x'; y_0, y_1) \mid y_0, y_1 \in \partial^c\psi^c(x)\}$ satisfy (HS) 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*

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

Thus whenever the gradients $\{\nabla_x c_\Delta(x'; y_0, y_1)\}_{y_0, y_1 \in \partial^c \psi^c(x)}$ satisfy Condition (C) and Definition 2.6.4, then for every $j \geq 1$ we have the following codimension formula:

$$codim_{Z_1} Z_j = j - 1.$$

Proof. Consider the mapping $G : B_\epsilon(x_0) \to \mathbb{R}^d$ defined by

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

Here we abbreviate $A(x, y_i) := A(x, y_i, y_0)$ for $y_0, y_1, \dots, y_j \in \partial^c \psi^c(x_0)$ such that

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

is linearly independant subset of $T_{x_0} X$ (which necessarily satisfies Halfspace conditions). The map $G : B_\epsilon(x_0) \to \mathbb{R}^j$ is local DC-function according to Lemma 3.2.1.

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 (3.2) 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.

Remark. The proof of 3.2.4 does not necessarily require the DC-implicit function theorem, since the cross-differences $x \mapsto c_\Delta(x, y, y')$ are assumed to be C^2 according to Assumption (A1). We could therefore have applied a standard C^1- or C^2-implicit function theorem (e.g. [Spi71, Thm. 2.12]) to conclude $Z'(x)$ is a C^1 (or C^2) subvariety of X. However we prefer the DC perspective, which appears the more natural category for optimal transport methods.

Recently [KM18] obtained an explicit parameterization of the 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 explicit parameterization requires a global splitting $X = X_0 \times X_1$ of the source domain to express the singularities ("tears" in their terminology) as the graph of DC-function $G : X_0 \to X_1$ as above.

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

The Kantorovich functor $Z : 2^Y \to 2^X$ can be further used to filtrate the source X according to codimensions. We "skewer" the cube 2^X, or more accurately the support $spt(Z)$, along the diagonal following an idea of Prof. D. Bar-Natan [Bar02a].

Definition 3.3.1. 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.

The lemmas below prove the subvarieties Z_j are well-defined for $j = 1, 2, \ldots$. 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. Our definition exhibits Z_j as a union of manifolds-with-corners $Z'(x)$.

Lemma 3.3.2. *For every $j \geq 1$, we find Z_j is a closed topological subset of X.*

Proof. We claim that whenever $\{x_k\}_k$ is countable sequence in Z_j converging to a limit x_∞ in X, then $x_\infty \in Z_j$. Along such a convergent sequence the subdifferentials $\partial^c\psi^c(x_k)$ vary upper semicontinuously. So $\partial^c\psi^c(x_\infty) \subset \partial^c\psi^c(x_k)$ for sufficiently large k. Therefore $x \mapsto \dim(Z'(x))$ is an upper semicontinuous numerical function and Z_j is topological closed subset of X.

□

Thus we filtrate the singularity according to codimension, and obtain a descending chain of closed subsets

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

of the source X. From Lemma 3.2.1 we found the equality $\dim(\partial^c\psi(y_0)) = \dim(X) - \dim(Y)$ under Assumptions (A0)-(A4) and Halfspace (HS) conditions.

At this point we need 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. That is, 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 3.3.3 (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

$$\phi(z) \geq \phi(x) + v^*(z - x) + o(|z - x|).$$

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 $dom(\psi^c) \subset X$, (c.f. Lemma 2.5.4). In otherwords, every $x \in dom(\psi^c)$ admits an open neighborhood U of x and a constant $C \geq 0$ such that $D^2 f > -C Id_n$ throughout U, where $D^2 f = D^2_{xx} f$ is the Hessian in the source variable x. So given a c-convex potential ψ^c on X, for every $x_0 \in X$ there exists open neighborhood U and $C > 0$ 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 $\psi^c|U$, 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

$$\partial^c\psi^c(x_0) \subset \{y \in Y | -\nabla_x c(x, y) \in \partial_\bullet\psi^c(x_0)\}. \tag{3.3}$$

Observe $\emptyset \neq \partial^c\psi^c(x_0)$ whenever $x_0 \in dom(\psi^c)$. We abuse notation and simultaneously denote $\nabla_x c(x, y)$ for the canonical covector in T_x^*X, with the tangent vector in T_xX using the ambient Riemannian structure. The inclusion (3.3) 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 3.3.4 ([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 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 3.3.5. *[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 (3.3). 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

$$\partial_\bullet(\psi^c|_U(x) + C|x|^2/2) = \partial_\bullet\psi^c|_U(x) + C\langle -, x\rangle. \tag{3.4}$$

So the local subdifferential $\partial_\bullet(\psi^c_U + C|x|^2/2)$ in $T^*_x X$ 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

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

For every $x \in Z_j$ we conclude

$$\dim_{\mathscr{H}}(Z_j \cap U) \subset S^j(\psi^c|_U + C|x|^2/2),$$

which according to Alberti's theorem 3.3.4 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 $x \in Z_j$. This follows from the fact that Z_j is topologically closed subset of X, and therefore every covering by open sets can be reduced to a countable covering. □

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 (semicouplings). An economic definition of singularity is this: singularity arises wherever there is competition for limited common resources. The key features of our formulation of Kantorovich singularity are the following: Definition 3.1.1 is *categorical* and Definition 3.1.1 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 (2.13) 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 those points x where c-convex potentials are not uniquely differentiable, where $\partial^c\psi^c(x) \subset Y$ is not a singleton. Our thesis thus passes silently over questions of the type "how differentiable is the map T on its domain of continuity?".

Several results concerning singularities of optimal transports have been attained in the literature. In [Fig10] 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 [Fig10, §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 $\mathscr{H}^2_X(\overline{Z_2}) = 0$. 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

$$\dim_{\mathscr{H}}(Z_2) \leq \dim(X) - \dim(Y) - 1.$$

The recent work [KM18] confirms this estimate under particular conditions, namely euclidean quadratic cost and affine-independance of the disjoint target components. We developed the method of [KM18] in our Proposition 3.3.5. The Proposition 3.3.5 confirms this popular expectation of the dimensions of singularities of optimal transports.

3.4 Local–Global Homotopy Reductions: (UHS) Conditions

The final section 2.6 of Chapter 2 constructed a continuous deformation retract of a source space X onto an active A when so-called Condition (C) and Halfspace Conditions were satisfied. See Theorem 2.6.2. This section generalizes the homotopy-reduction constructed in 2.6.2, and describes a retraction procedure $Z_1 \rightsquigarrow Z_2 \rightsquigarrow \cdots Z_{J+1}$, 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. The retraction requires our cost c satisfy Assumptions (A0)–(A5), and a further hypothesis of "uniform Halfspace conditions", see Theorem 3.4.2. Briefly, our retraction is defined by integrating a vector field $\eta(x', inv)$ which – under sufficient Assumptions – can be reparameterized according to arclength to obtain a continuous deformation. The retraction is guaranteed only when our so-called (UHS) condition is satisfied (see Definition 3.4.1), and this is the primary obstruction in practice.

Remark. Our deformations are effectively generalizations of the well-rounded retracts of [Ash84], [Sou78] as interepreted in the optimal semicoupling category in the Kantorovich singularities. We refer the reader to [PS08], [Ste07, §A.6.4], for discussions on the lattice-theoretic and number-theory applications of the well-rounded retracts. That our retractions are generally large-codimension ($J \geq 2$) should prove useful for applications, and new constructions of small-dimensional classifying spaces.

Now we describe our retraction procedure, which requires some notations. Let c be a cost satisfying Assumptions (A0)–(A5), and $\psi^{cc} = \psi$ a c-concave potential maximizing Kantorovich's dual problem (§2.3) relative to a source σ and target τ. We let $Z = Z(c, \sigma, \tau)$ be the corresponding Kantorovich functor. Suppose $j \geq 1$ and consider the inclusion $Z_j \hookrightarrow Z_{j+1}$, as defined in Definition 3.3.1. We select some x' from $Z_j - Z_{j+1}$, evaluate $\partial^c \psi^c(x') \subset Y$, and abbreviate $Z' := Z(\partial^c \psi^c(x'))$. Moreover let

$$pr_{Z'} : T_{x'} X \to T_{x'} Z'(x)$$

denote the orthogonal-projection mapping, where $T_{x'} Z'(x)$ is the space of directions of $Z'(x)$ at x'. Our retraction constructs an "averaged" Bochner integral $\eta(x, avg)$. See [Yos68, §V.5] for formal properties of Bochner integrals.

Definition 3.4.1 (**(UHS) Conditions**). Let $x \in Z' = Z(\partial^c \psi^c(x'))$ and select some $y_0 \in \partial^c \psi^c(x')$. For choice of parameter $0 < \alpha < 1$, we define tangent vectors $\eta(x, y_*) \in T_x X$, $y_* \in Y$, by the equation

$$\eta(x, y_*) := |\psi(y_0) - \psi(y_*) + c_\Delta(x; y_*, y_0)|^{-\alpha} \cdot pr_{Z'}(\nabla_x c_\Delta(x; y_*, y_0)). \tag{3.5}$$

We say Uniform Halfspace (UHS) conditions are satisfied at x in Z' with respect to parameter α if

(UHS1): the Bochner integral $\eta(x, avg)$ defined by

$$\eta(x, avg) := (\mathscr{H}_Y[Y'(x) - \partial^c \psi^c(x)])^{-1} \int_{Y'(x) - \partial^c \psi^c(x)} \eta(x, y_*).d\mathscr{H}_Y(y_*) \tag{3.6}$$

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

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

$$||\eta(x, avg)|| \geq C \int_{Y'(x) - \partial^c \psi^c(x)} ||\eta(x, y_*)|| d\mathscr{H}_Y(y_*) > 0 \tag{3.7}$$

holds.

The definition 3.4.1 is key to our thesis, and we remark on this definition and then present our main Theorem 3.4.2. First the definition of $\eta(x, y_*)$ is independant of the choice of $y_0 \in \partial^c \psi^c(x')$. Secondly, in practice the parameter $0 < \alpha < 1$ is taken sufficiently small to ensure convergence of the integral defining $\eta(x, avg)$. The choice $\alpha = 1/2$ is sufficient in our applications. Indeed the Bochner integral is an improper integral, e.g. whenever $Y'(x') - \partial^c\psi^c(x')$ is a connected open subset of Y. Evidently $\eta(x', y_*)$ diverges when $y_* \in Y'(x')$ converges towards the closed subset $\partial^c\psi^c(x')$ in $Y'(x')$. Our (UHS1) condition requires this divergence be integrable and finite with respect to the uniform measure $\mathscr{H}_Y$ restricted to $Y'(x) - \partial^c\psi^c(x')$.

Theorem 3.4.2. *Let $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be cost satisfying Assumptions (A0)–(A5). 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. Let $x' \in Z_1$ be a point supported on the activated domain of an optimal semicoupling π_{opt}, and abbreviate $Z'(x') := Z(\partial^c\psi^c(x')) = \cap\{\partial^c\psi(y_0) | y_0 \in \partial^c\psi^c(x')\}$.*

If (UHS) Conditions are satisfied for all points $x'' \in Z' \cap Z_j$, then we construct continuous map

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

such that

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

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

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

Therefore $Z' \cap Z_{j+1}$ 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.

For x' varying over Z_j, with correspondant $Z' = Z'(x') := Z(\partial^c\psi^c(x'))$, the mappings

$$\{\Psi : Z' \cap Z_j \to Z' \cap Z_j\}$$

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

Theorem 3.4.3. *Suppose uniform Halfspace (UHS) conditions 3.4.1 are satisfied throughout Z_j for some $j \geq 1$.*

Then the local homotopy-equivalences

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

constructed in Theorem 3.4.2 assemble to a continuous deformation retract $Z_j \times [0, 1] \to Z_{j+1}$.

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 $Z_1 \simeq Z_{J+1}$.

Sketch of proof for Theorem 3.4.2. We construct a continuous vector field $\eta(x', avg)$ on $Z' \cap Z_j$ which blows-up precisely on $Z' \cap Z_{j+1} \subset Z' \cap Z_j$. The field $\eta(x', avg)$ will generate a global forward-time continuous mapping

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

satisfying the usual ordinary differential equation $\frac{d}{ds}[\Psi(x', s')] = \eta(\Psi(x', s'), inv)$ for all $s' \geq 0$. But given our hypotheses, the flow generated by $\eta(x', avg)$ will converge in finite-time to the poles $Z' \cap Z_{j+1}$. Therefore every initial point x' defines a trajectory $x = x(s)$, with $x(0) = x'$ and existing on a finite interval $I(x')$, and where everything varies continuously with respect to x'. The flow Ψ is then reparameterized according to arclength to obtain a continuous deformation retract $\Psi' : (Z' \cap Z_j) \times [0, 1] \to Z' \cap Z_j$ as desired.

$\square$

The interested reader should compare the following proof of Theorem 3.4.2 with our earlier "base case" retraction of $X (= Z(\emptyset))$ onto the activated domain $A = Z_1$, see Theorem 2.6.2. Now without further ado, we proceed to the construction of Ψ.

Proof of Theorem 3.4.2. The (UHS) conditions ensure the cross-differences c_Δ have non-vanishing gradient $\nabla_x c_\Delta \neq 0$ throughout the domain Z'. The gradients $\nabla_x c_\Delta$ then vary continuously over Z_j. Moreover, the uniform Halfspace condition (UHS2) ensures $||\eta(x, avg)||$ is uniformly bounded away from zero in neighborhoods of the poles $\{\eta(x, avg) = \infty\} = Z' \cap Z_{j+1}$ in $Z' \cap Z_j$. Indeed, if $\{x_k\}$ is sequence of points in $Z' \cap Z_j$ converging to a point $x_\infty \in Z' \cap Z_{j+1}$, then there exists at least one y_* and open neighborhood V of y_* such that $\eta(x, y)$ diverges to infinity for all $y \in V$. So (UHS2) implies the divergence of $||\eta(x, avg)||$ to $+\infty$. So $||\eta(x, avg)||$ is surely uniformly bounded away from zero near the poles. This establishes a strong form of Assumption (A5).

The vector field $\eta(x, avg)$ integrates to a global forward-time flow $\Psi : Z' \cap Z_j \times [0, +\infty) \to Z' \cap Z_j$, where as usual we have $d/ds[\Psi(x', s')] = \eta(\Psi(x', s'), inv)$ for all $s' \geq 0$. According to Lemma 2.6.7, 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 (2.17). Our Assumptions (A0)–(A4) imply continuous dependance on the choice of initial value. The argument from Lemma 2.6.8 then proves lengths of trajectories will also vary

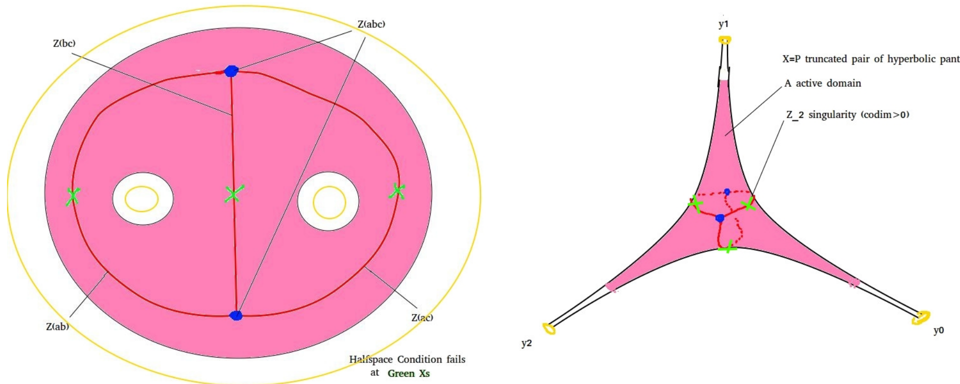

Figure 3.3: Horospherically truncated pair of pants, with active domain relative to a repulsion cost (See Definition 4.2.3 from Section 4.2). Halfspace Conditions fail. $P \rightsquigarrow Z_1 \rightsquigarrow Z_2$ are homotopy equivalent, where P is pair of pants. But Z_2 does not deformation retract to Z_3

continuously with respect to initial value. Therefore the flow Ψ can be reparameterized into a continuous deformation retract Ψ' of $Z' \cap Z_j$ onto $Z' \cap Z_j$, as desired.

$\square$

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 3.4.3:

Proof of Theorem 3.4.3. The proof of Theorem 3.4.2 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 3.4.3 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 4

Repulsion Costs

The previous Chapters have assumed c is a general cost satisfying the Assumptions (A0), ..., (A5). The present chapter introduces a new class of costs denoted $c|\tau$, $\widetilde{c}$,, $\widetilde{c}^*$, v, and derived from a highly simplified model of "electron–electron" interactions. These "electron–electron" costs are contradistinct from "electron–positron" costs, which are modelled by the usual squared-distance costs $c = d^2/2$. These new costs are defined specifically in our repulsion costs $c|\tau$ (Definition 4.2.3), two-pointed repulsion costs $\widetilde{c}$ (Definition 4.4.1), and their generalizations to $\widetilde{c}^*$ (Definition 4.9.7) and visibility cost v (Definition 4.9.5). In Section 4.1 we describe properties (D0)–(D4) which are useful for practical applications. We conjecture that our costs $\widetilde{c}^*, v$ satisfy these properties (D0)–(D4), c.f. Conjectures 4.9.8, 4.9.6 in certain settings, as we elaborate below.

4.1 Hypotheses and Practical Applications

The topological results of our thesis address two issues. The first is abstract and very general. We identify hypotheses on costs functions c which guarantee that the activated source domains of c-optimal semicouplings admit large-codimension strong deformation retracts. See Theorems A, B, C from Introduction, 1.4, 1.4.1, 1.5.1 and Theorems 3.4.2, 3.4.3. Thus we identify closed subvarieties $Z \subset X$, of the source domain (X, σ), for which the inclusions $Z \hookrightarrow X$ are homotopy isomorphisms. The second issue addressed by this thesis is practical, and concerns the application of our general theory to some particular costs defined in the present chapter. Sections 4.4, 4.5 introduce some conjectures, e.g. 4.9.8, 4.9.6 regarding the (Twist) conditions which are necessary for applications. The reader may recall that (Twist) ensures the uniqueness of c-optimal semicouplings, c.f. Proposition 2.5.6. Observe that (Twist) condition, i.e. Assumption (A4), is weaker condition than the usual (Twist) conditions of the conventional coupling theory. Indeed in our settings the target space $Y = \partial_* F[t]$ is presumed to have strictly smaller dimension than the source $X = F$, with $\dim(Y) < \dim(X)$. Here is an advantage of our semicoupling setting, and motivation for the conjectures 4.9.8, 4.9.6. Despite our expectations that these conjectures are true, it is useful to make some remarks on the general hypotheses necessary for the practical applications of our costs to Chapters 5, 6. If $\underline{F}$ is a chain sum with well-separated gates $\{G\}$ (see §4.3 for definitions, 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 a type of repulsion cost, with $c'(x,y) = +\infty$ whenever $x, y \in \partial_*[\underline{F}]$;

(D1) the domain $dom(c')$ consists of pairs (x,y) which occupy a commmon convex 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)–(A5), c.f. §1.3;

(D4) the restriction of c' to a gate G yields a restricted 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 $\mathscr{H}^{can}_{\mathscr{E}[G]}$, then the homotopy-reductions of Theorems 3.4.2, 3.4.3 with respect to c'-optimal semicouplings from source $\mathscr{H}_G$ to target $\mathscr{H}^{can}_{\mathscr{E}[G]}$ yield deformation retracts of the gates G to points, $G \rightsquigarrow \{pt\}$.

We define costs $c = c|\tau, \widetilde{c}, \widetilde{c}^*, v$ in the following sections which are readily seen to satisfy properties (D0)–(D2). However the properties (D3)–(D5) are more difficult to verify, and in fact we cannot verify these properties for the costs of interest. The visible repulsion cost v is the basic tool developed for our applications in Chapters 5, 6, but the verification of (D3)–(D5) is omitted and summarized in the Conjectures 4.9.8, 4.9.6. We further expect the hypotheses of Theorem A (see 1.4 and Theorem 3.4.2) to be sufficiently satisfied: when the active domain $A \hookrightarrow X$ of v-optimal semicouplings is sufficiently large, then we expect the inclusion $A \hookrightarrow X$ to be a homotopy-isomorphism. We formalize this expectation in:

Conjecture 4.1.1. *Let $X = F[t]$ be a convex excision with excision boundary $Y = \partial_* F[t]$ (see §4.3 for definitions). Let σ be source measure on X and τ a target measure on Y with $\int_X \sigma > \int_Y \tau$. Let v be the visible repulsion cost constructed in Definition 4.9.5. Then the active domains $A = Z_1$ of v-optimal semicouplings are strong deformation retracts of the source X when the target mass $\int_Y \tau$ is sufficiently large.*

The heuristic motivating Conjecture 4.1.1 is that the necessary Halfspace conditions (recall Theorem 3.4.2) fail on a particular compact subdomain of X. As we increase the target mass $\int_Y \tau$ sufficiently close to $\int_X \sigma$, the active domains $A \subset X$ expand and eventually "fill" X in the limit. Thus the hypotheses of Theorem A weaken as the target mass increases. The idea motivating Conjecture 4.1.1 is that the necessary Halfspace Conditions are satisfied in sufficiently small tubular neighborhoods of Y in X. However the repulsion costs $\widetilde{c}$ diverge to $+\infty$ uniformly throughout these tubular neighborhoods. Since active domains are necessarily domains where v is bounded, we find Halfspace conditions failing on compact subsets interior to X which are eventually disjoint from $X - A$.

We remark that Conjecture 4.1.1 definitely fails for "attractive" costs (i.e. quadratic distance $c = d^2/2$) with X, Y as above. Informally we expect the active domains A of $d^2/2$-optimal semicouplings between X, Y are deformation retracts of the target Y when $\int_Y \tau$ is sufficiently large, but this is only a guess.

Furthermore our applications of Theorem B (c.f. Theorem 3.4.3) require sufficient (UHS) conditions be satisfied by the gradients of $\widetilde{c}, v$. In practice we expect (UHS) conditions of $\widetilde{c}, v$ are controlled by restriction to the gates $\widetilde{c}|G$, $v|G$. Moreover we construct gates G possessing large symmetries, c.f. Definition 5.5.1 and Theorem 5.5.2, and we expect these symmetry conditions ensure sufficient Halfspace conditions are satisfied. This heuristic underlies our Conjecture 1.5.2, and is application of our Theorems A, B, C to the construction of Spines for $E\Gamma$ models X admitting additional geometric stucture, e.g. complete and nonpositively curved. In the remaining chapters we apply these methods to $E\Gamma$-classifying spaces X, where Γ is an infinite discrete group satisfying Bieri-Eckmann homological duality, e.g. when $\Gamma = \mathbf{G}(Z)$ is any standard arithemetic group. See Sections 5.5, 6.3, 6.4 for details.

To illustrate the above hypotheses, we include Figure 4.1 which illustrates the singularity structure of the averaged two-pointed repulsion costs $\widetilde{c}^*$ on a two-simplex and a two-dimensional square. In these elementary cases the properties (D0)–(D5) are satisfied.

For brevity, we illustrate the above ideas with the "visibility cost" v in the simplest case of a two-dimensional regular simplex $\blacktriangle^2$. Here we find $X = \blacktriangle^2$ is the source, 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 some positive scalar multiple of $\delta_{y_0} + \delta_{y_1} + \delta_{y_2}$ satisfying $\int_X \sigma > \int_Y \tau$. Then we have

$$v(x, y_0) = \frac{1}{2}||x - y_0||^{-2} + \frac{\lambda_1}{1 - \lambda_0}||x - y_1||^{-2} + \frac{\lambda_2}{1 - \lambda_0}||x - y_2||^{-2}, \tag{4.1}$$

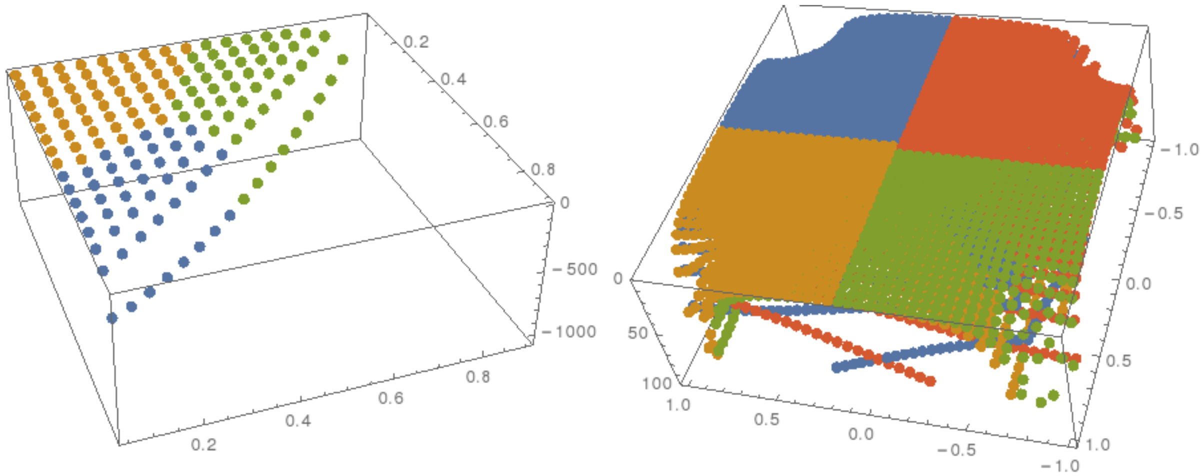

Figure 4.1: Singularity structures of $\widetilde{c}$ on regular 2-simplex and regular square satisfy properties (D0)–(D5)

where $0 \leq \lambda_i \leq 1$ are the unique scalars satisfying $\lambda_0 + \lambda_1 + \lambda_2 = 1$ and $bar(\lambda_0 \delta_{y_0} + \lambda_1 \delta_{y_1} + \lambda_2 \delta_{y_2}) = x$.

Now suppose $\blacktriangle_i$, $i \in I$, 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. interval). Then the visibility cost v defined in equation (4.1) extends to a repulsion cost v^* on $\underline{F}$ satisfying properties (D0)–(D5) as above. In particular, the singularity structures $Z(\sigma, \tau, v^*)$ are continuous interpolations of the singularity structures $Z(\sigma|G, \tau|G, v^*|G)$ restricted to every gate G. Compare Figure 4.1.

4.2 Definition of Repulsion Costs

Let F be a compact geodesically-convex finite-dimensional subset of some euclidean $\mathbb{R}^N$ and equipped with standard distance $d(x, y) = ||x - y||$. Now suppose Y is closed subset, say the boundary of a convex set $Y := \partial F$. The Riemannian distance d on F then restricts to a geodesic distance $dist = d|Y \times Y$ on Y. Recall $\dim_{\mathscr{H}}(F)$ coincides with the dimension of the minimal affine subspace containing F. The boundary $Y = \partial F$ is well-defined topological space, and both F, ∂F have integral Hausdorff dimensions satisfying $\dim_{\mathscr{H}}(F) = \dim_{\mathscr{H}}(\partial F) + 1$.

Definition 4.2.1 (Extreme points)**.** Let F be closed convex. Then $x \in F$ is an extreme-point if x is not the midpoint of any pair of distinct points $x_0, x_1 \in F$, $x_0 \neq x_1$. Thus $x = [x_0, x_1]_{1/2}$ implies $x_0 = x_1 = x$.

The topological boundary ∂F possesses a canonical Hausdorff measure and with in-

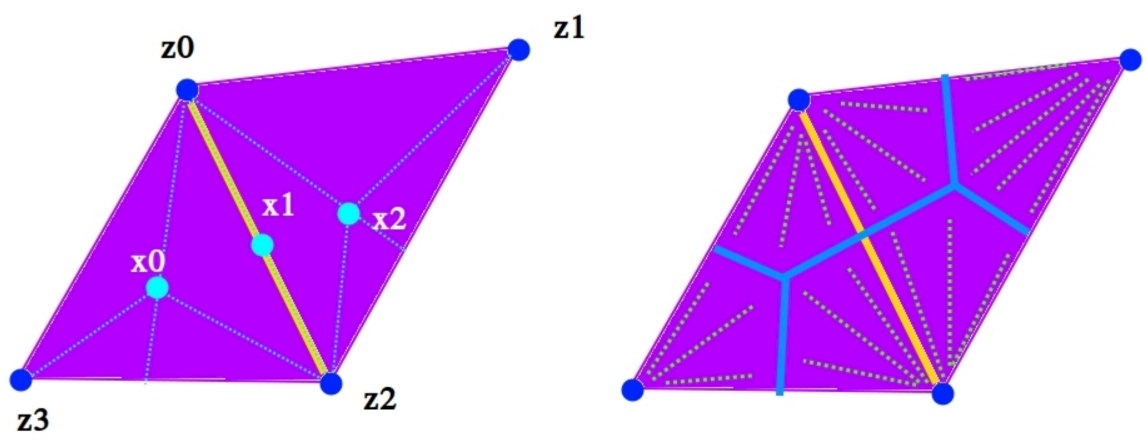

Figure 4.2: The yellow line designates the one-dimensional gate G. The blue line designates the interpolated singularity structure with respect to the gated-visibility cost v^*.

tegral dimension. Alexandrov's spherical image construction, or Gauss curvature ω, is another interesting Radon measure defined on ∂F. We recall the definition: 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 naturally found to be a closed spherically-convex subset of the linear dual space. The measure ω on ∂F is defined to be the spherical-measure of $N(y)$. We refer the reader to [Ale06], [Oli07] for further details.

However the extreme pointset $E := \mathscr{E}[F]$ is a subset of ∂F, possibly not closed and with possibly irrational Hausdorff dimension. Indeed 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. Frequently $\mathscr{E}[F]$ and ∂F may coincide, e.g. when F is strictly convex such that ∂F contains no nontrivial affine segments.

Our constructions below are readily generalized to compact convex subsets F satisfying the following property:

Definition 4.2.2 (Property (IDE))**.** A convex compact subset F has integral-dimension extreme points, called property (IDE), if there exists relatively-closed 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 Hausdorff dimension $\dim_{\mathscr{H}} E^{(j)} = j$ for every integer $j = 0, 1, 2, etc.$.

Property (IDE) says the convex set F has extreme points with an "elementary" structure. The property (IDE) implies the extreme points $E := \mathscr{E}[F]$ has a sufficiently canonical Hausdorff-type measure $\mathscr{H}_E^{can} := \mathscr{H}_{E^{(0)}} + \mathscr{H}_{E^{(1)}} + \cdots$. One finds F has property (IDE) if $\mathscr{E}[F] = \partial F$ or F is a polyhedra. Henceforth we presume our convex bodies F have property (IDE) and typically $\mathscr{E}[F] = \partial F$. For simplicity, the reader is free to presume that F is a convex and compact polyhedra for the duration of the thesis, but is not strictly necessarily.

For the following definition, let F be an $(N+1)$-dimensional compact convex set with property (IDE), with extreme points $E := \mathscr{E}[F]$ and canonical measure $\mathscr{H}_E^{can}$ on E. For $y \in E$, let $\mathbf{e}(y) := j$ be the unique integer such that $y \in E^{(j)}$.

Definition 4.2.3. Let τ be a Radon measure supported on the extreme points $E := \mathscr{E}[F]$ and absolutely-continuous with respect to $\mathscr{H}_E^{can}$. We define the repulsion cost $c|\tau : F \times E \to \mathbb{R} \cup \{+\infty\}$ by the rule

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

In the simplest case when $E = \partial F$ and $\dim_{\mathscr{H}}(E) = N$, we directly define c with respect to the Hausdorff measure $\tau = \mathscr{H}_{\partial F} =: dy$, and obtain

$$c(x, y_0) := (\int_{\partial F} q(x,y)^{-2-N} dy) - \frac{1}{2} d(x, y_0)^{-2-N}.$$

We illustrate with some examples.

Example. Let $Y = \{0, 1\}$ be subset of $X = conv[Y] = [0, 1]$ and $\tau = 1_Y = \delta_0 + \delta_1$. 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 4.2.

Example. Let $Y := \{y_0, y_1, y_2, \ldots, y_5\}$ 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 4.2.

Example. Let F be smooth compact 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

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

The exponent $\mathbf{e}$ in the definition 4.2.3 of $c|\tau$ 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

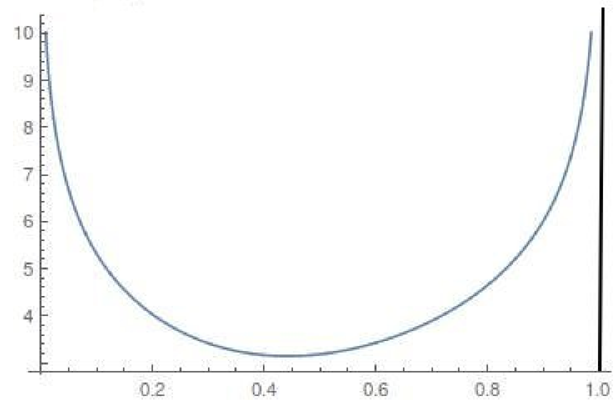

Figure 4.3: Graph of repulsion cost $x \mapsto c|\mathscr{H}_Y(x,0)$. The assymmetry reflects the "home-preference" of x to $y=0$.

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 $\frac{-1}{2}d(x,y_0)^{-2-e}$ imparts a definite "home-preference" to $c|\tau$ on $conv\mathscr{E}\times\mathscr{E}$. When a source point x is close to an extreme point y_0 in $E=\mathscr{E}[conv[E]]$, then the various pairings $\{c|\tau(x,y_*)|y_*\neq y_0\}$ becoming definitely greater than $c|\tau(x,y_0)$. So when a point x is activated there is no confusion in the optimization program: the point x goes home to the nearest point y_0.

Having defined the repulsion-costs, we now verify which Assumptions (A0), (A1), etc., are satisfied. We verify Assumptions (A0)–(A3) in the following

Proposition 4.2.4. *Let $X=F$ be 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 4.2.3 satisfies Assumptions (A0) – (A3) throughout $dom(c)=(X-Y)\times Y$*

Proof. Evidently $x\mapsto d(x,y_*)^{-2-e}$ 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 4.2.3. 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 $c_y: dom(c_y)\to\mathbb{R}$ are compact subsets of X. This proves (A1). Now we also see $||\nabla_x c(x,y)||$ varies continuously with respect to y. So indeed (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.

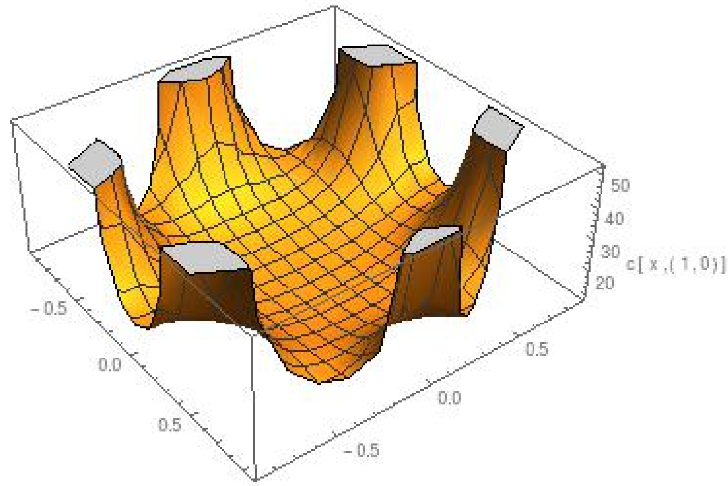

Figure 4.4: Graph of repulsion cost $x \mapsto c|\mathscr{H}_Y(x,(1,0))$. The "home-preference" implies the pole at $y=(1,0)$ has smaller diameter than the other poles.

We leave the details to the reader.

□

Proposition 4.2.5. *Let F be a compact geodesically-convex set with extreme points $E = \mathscr{E}[F]$. Then the repulsion cost $c|\mathscr{H}_E$ satisfies (Twist) condition throughout dom$(c|\mathscr{H}_E)$.*

Proof. We abbreviate $dy := d\mathscr{H}_E(y)$, and $q(x,y) = d(x,y)^{2+e}$ where $e = \dim_{\mathscr{H}}(E)$. Fix $x \in F$. Then (Twist) condition 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, so (Twist) follows from injectivity of the rule $y_0 \mapsto \int_Y \nabla_x q(x,y_0)^{-1}$. □

Thus we have defined our repulsion costs $c|\tau$. Our applications require extending the above definitions to nonpositively-curved complete finite-dimensional Riemannian spaces (X,d), e.g. Cartan-Hadamard manifolds. The reader is referred to [BGS85] for standard definitions. We elaborate below.

4.3 Chain sums and Well-Separated Gates

We begin with general definitions, and then specialize to a more symmetric setting involving isometric group action $X \times \Gamma \to X$. Recall Figure 1.5 from the Introduction. Let

$\{F_i\}_{i\in I}$ be countable collection of compact convex subsets of complete Cartan-Hadamard source space (X,d), e.g. $X=\mathbb{R}^N$ with isometric embeddings $F_i \hookrightarrow X$ for each $i\in I$. Given such $\{F_i\}_{i\in I}$ we let $\underline{F}$ denote the chain sum $\underline{F}=\underline{F}_I=SUM[\{F_i\}_{i\in I}]$.

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 4.3.1. In the above notation, let $\{G^-\}=\pi_0(\cup_{i,j}F'_{ij})$ be the set of topological connected components of $\cup_{i,j}F'_{ij}$. The gates $\{G\}$ of the chain sum $\underline{F}=SUM[\{F_i\}_{i\in I}]$ is the set $\{G\}$ of closed subsets $G:=\overline{G^-}$ 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 4.3.1 imply $\cup_{i,j}F_{ij}$ is a locally closed subset of X, and a closed subset of X.

Definition 4.3.2. The gates $\{G\}$ of a convex chain sum $\underline{F}=$ are well-separated if the components G,G' of the setl $\{G\}$ 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 convex chain sum $\underline{F}=SUM[F.\Gamma]$. The gates $\{G\}$ of the chain sum $\underline{F}=SUM[F.\Gamma]$ form a Γ-set, i.e. the set of gates $\{G\}$ is Γ-invariant and therefore supports 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 4.3.3. We say the convex chain $\underline{F}=SUM[F.\Gamma]$ 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.

4.4 Two-Pointed Repulsion Costs

The previous Section 4.3 introduced a type of repulsion costs $c|\tau$ defined on closed convex bodies F where either $\mathscr{E}[F]=\partial F$ or $\mathscr{E}[F]$ is a closed discrete set. The definition of $c|\tau$

is generally defined for convex sets satisfying the (IDE) condition (Definition 4.2.2). We think $c|\tau$ – which satisfies (A0)–(A4) – is an interesting cost for study, but has a limited role in our applications. The simplest case of this cost leads to the two-pointed repulsion cost $\widetilde{c}$ defined in the present section. We warn the reader that our definition of $\widetilde{c}$ is temporary, and will not recur in Chapters 5, 6. However $\widetilde{c}$ will be modified to an "averaged" two-pointed repulsion cost in Definition 4.9.7 in §4.9. The definition is inductive. If F is one-dimensional compact convex set, with $\mathscr{E}[F] = \partial F = \{y_0, y_1\}$, then we define

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

as in previous Definition 4.2.3. Next suppose F is a compact convex set, with $x \in F$ and $y \in \partial F$. Consider the directed geodesic ray $\rho(y, x)$ issuing from y towards x. If we extend the ray indefinitely, we find $\rho(y, x)$ first intersects ∂F at some unique point $y' := proj(y, x)$ opposite y with respect to x. So for $y \in \partial F$ we obtain a projection-type map $proj_y : (F - \{y\}) \to \partial F$. We use the projection map to inductively define $\widetilde{c}$ to higher-dimensional convex compact sets F. Indeed if $x \in F$, $y \in \partial F$, then we define

$$\widetilde{c}(x, y) := \frac{1}{2} dist(x, y)^{-2} + dist(x, proj_y(x))^{-2}. \tag{4.2}$$

The formula (4.2) immediately extends to a convex chain sums $\underline{F}$ with well-separated gates $\{G\}$ as defined in Section 4.3, and this yields our definition of two-pointed costs.

Definition 4.4.1 (Two-Pointed Repulsion Cost)**.** Let $\underline{F}$ be convex chain sum with well-separated gates $\{G\}$ (Definition 4.3.2). 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

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

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

The projection $proj$ varies continuously with x, y, and we find $\widetilde{c}$ is continuous throughout its domain, where $dom(\widetilde{c}) = \{(x, y) \mid proj_y(x) \neq xmy \neq x\}$. Our hypotheses on the well-separated gates $\{G\}$ of $\underline{F}$ implies the gates G are closed subsets of the boundaries ∂F of chain summands F with $G \subset F$. If $y \notin G$ and $x \in G$, then $proj_y(x) = x$ and $\widetilde{c}(x, y) = +\infty$. This observation implies that $\widetilde{c}$-optimal semicouplings π having finite total cost, will restrict to $\widetilde{c}|G$-optimal semicouplings $\pi.1_G$ on the gates G of $\underline{F}$. Therefore $\widetilde{c}$ satisfies Hypothesis (D2) from Section 4.1.

Now we consider the Assumptions (A0), (A1), (A2), etc..

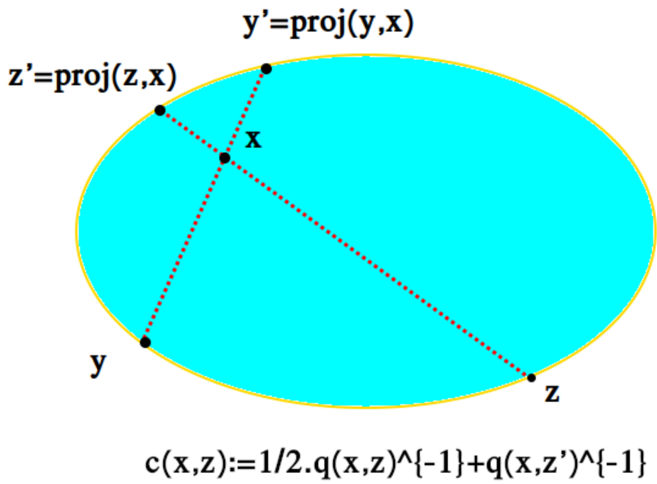

Figure 4.5: Illustrating the projection map $proj : F \to \partial F$ defining two-pointed repulsion cost $\widetilde{c}$ for ellipse F where $\partial F = \mathscr{E}[F]$.

Lemma 4.4.2. *Let $\underline{F} = \sum_{i \in I} F_i$ be convex chain sum with gates $\{G\}$ satisfying Definition 4.3.3. Then the two-pointed repulsion cost $\widetilde{c}$ defined in (4.2) satisfies Assumptions (A0)–(A3) throughout $dom(\widetilde{c})$.*

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

□

The key property necessary for applications is the Assumption (A4) known as (Twist) condition. However discussions with Professor R.J. McCann indicate the existence of convex sets F for which $\widetilde{c}$ – as defined above – fails to satisfy (Twist). Indeed "general" convex sets are difficult to control, and it may happen that distinct points $y_1, y_2 \in \mathscr{E}[F]$ have coincident critical points at some $x \in F$, with $\nabla_x \widetilde{c}(x, y_1) = 0 = \nabla_x \widetilde{c}(x, y_2)$. But such equality violates (Twist).

The applications we consider in Chapters 5, 6 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 $\mathscr{E}[F] = \mathscr{E}[G] \cup \mathscr{E}[G'] \cup \cdots$. We don't expect semiregularity is sufficient to ensure $\widetilde{c}$ satisfies (Twist). However the following sections will introduce an application of Krein-Milman theorem, and will replace $\widetilde{c}$ with an "averaged two-pointed repulsion cost" $\widetilde{c}^*$ (see Definition 4.9.7). We then expect semiregularity is sufficient to ensure $\widetilde{c}^*$ satisfies the necessary (Twist) conditions, c.f. Conjecture 4.9.8.

4.5 Convex-Excisions

Henceforth we assume F is a totally geodesic subset of a Cartan-Hadamard space (X, d). Therefore we presume there exists unique geodesics between any pair of points x, y in X, and antipodal points are nonexistant. The present section formalizes our idea of "scooping out" convex subsets. Recall the idea of excision, which is familiar from Eilenberg-Steenrod axioms of singular homology. Let X be complete riemannian manifold and F a closed convex body. We propose extending the Krein-Milman theory of F and $E = \mathscr{E}[F]$ to nonconvex closed subsets of F obtained by excising convex neighborhoods of the extreme points E.

Definition 4.5.1 (Convex-Excision Parameter t)**.** Let F be compact geodesically-convex subset of a finite-dimensional Cartan-Hadamard space X. Let $\{\lambda\}$ be a countable collection subset of $\mathscr{E}[F]$. An excision-parameter t is a function $t : \mathscr{E}[F] \to \mathbb{R}_{>0}$, and defines a countable collection of open subsets $\{W_\lambda^t\}$ for every subset $\{\lambda\}$ of $\mathscr{E}[F]$ having the following properties:

(i) the boundaries ∂W_λ^t are smooth manifolds for every $\lambda \in \{\lambda\}$;

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

(iii) for every compact subset K of X, only finitely many convex subsets W_λ^t have nontrivial intersection.

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

The basic example we pursue involves a complete Cartan-Hadamard space (X, d). See [BGS85, §I.3] for formal definitions and Sections 5.3, 5.5. For every point-at-infinity λ and basepoint $x_0 \in X$, the horofunction $h_{\lambda,x_0} : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ is a geodesically-convex function on X. For every choice of $t_\lambda \in \mathbb{R} \cup \{-\infty, +\infty\}$, the sublevels

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

are convex subsets of X, namely "horoballs centred at λ with radius t_λ.

Definition 4.5.2 (Convex-Excision)**.** Given an excision parameter t, Definition 4.5.1, we designate the relative complement $F[t] := F - \cup_{\lambda \in \mathscr{E}[F]} W_\lambda^t$ a convex-excision model.

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

Definition 4.5.3 (Excision Boundary)**.** Let t be an excision parameter with excision $F[t]$. The excision-boundary $\partial_* F[t]$ is defined by $\partial_* F[t] := \cup_{\lambda \in \mathscr{E}[F]} (\partial F[t] \cap \overline{W_\lambda^t})$.

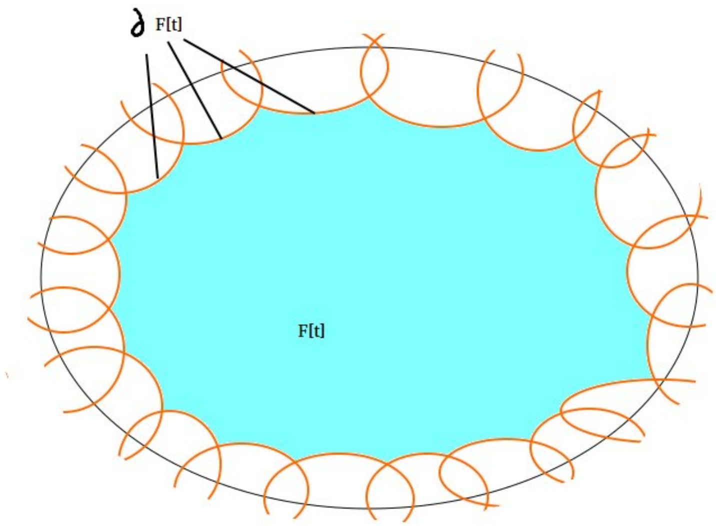

Figure 4.6: Excision of convex ellipsoid

In otherwords $\partial_* F[t]$ is that subset of $F[t]$ which intersects some boundary component ∂W^t_λ. Thus we obtain $F[t]$ by excising the t-family of open convex subsets W^t_λ from F. Our thesis exclusively studies locally-finite excision parameters. Thus for any λ we suppose there exists only finitely many λ' for which $W^t_\lambda \cap W^t_{\lambda'}$ is nonempty.

4.6 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 is managed by introducing the definition of a visibility relation V (Definition 4.6.1). For costs, we use the visibility relation to define a visibility factor $k(x, y)$ which rescales every term of the type $q(x, y)^{-1}$ in the repulsion costs introduced above. The formal definitions require some prepatory lemmas.

Definition 4.6.1 (Visibility Relation V)**.** Let $F[t] \subset F$ be a convex excision (Definition 4.5.1). Then a 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 4.6.2. *[[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 $\nabla_x d(x, y) = -\gamma'(0)$.*

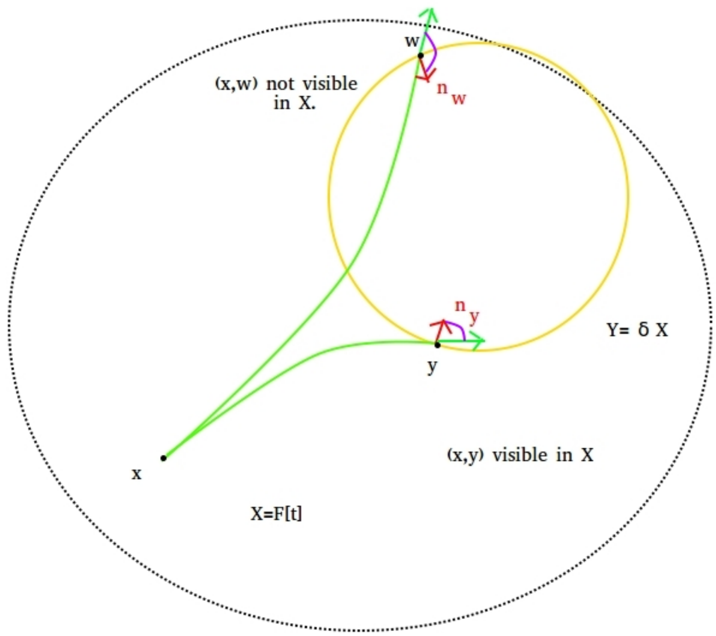

Figure 4.7: Excision domain $X = F[t]$, where (x, y) are visible in X and (x, w) is not visible in X

If $x \in F[t]$ and $y \in \partial_* F[t]$ are geodesically visible in $F[t]$, then Lemma 4.6.2 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 4.6.3. *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 the outward unit normal vector at $y \in Y$ in $X[t] \subset X$.*

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 $\langle \nabla_y d(x, y), \mathbf{n} \rangle < 0$ 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 coss 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 4.6.4 (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 4.7.1. The visibility factor is the function $k : X \times Y_\epsilon \to \mathbb{R}_{\geq 0} \cup \{+\infty\}$ defined by the formula

$$k(x,y) := \begin{cases} \langle \nabla_y d(x,y), \mathbf{n}_y \rangle^{-1}, & \text{if } x, y \text{ are visible },\\ +\infty, & \text{if } x, y \text{ not visible }. \end{cases} \tag{4.3}$$

The definition 4.6.4 represents a numerical function valued in $[1, +\infty]$, and diverging to $+\infty$ when y fails to be visible from x within $F[t]$, according to Lemma 4.6.3.

4.7 ϵ-Regularizations

The definition 4.5.1 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 manifold-with-corners which will have smoothly varying outward normal vectors. This implies the corresponding visibility factor $k : X \times Y_\epsilon \to \mathbb{R}_{>0} \cup \{+\infty\}$ will be continuous throughout its domain.

The boundary of $\partial_* F[t]$ is cellulated by the convex excision parameter t and $\{W[t_y] | \, y \in Y\}$. The individual convex halfspaces $W' := W[t]$ have well-defined inward normal vectors $n_y \in T_y W'$ for every $y \in W'$. However a given point $y' \in \partial_* F[t]$ can occupy multiple subsets of $\{W[t_y] | \, y \in Y\}$, 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 4.7.1 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. With more effort one may obtain a C^∞-regularization, c.f. [Gro91, §3, pp.53], [Gro14b, §3.4, 5.7], and Douady-Hérault's "Arrondissement des variétés à coins", [BS73, Appendix, §6].

Proposition 4.7.1 (C^∞ Regularization of manifold-with-corners)**.** *Let X be complete riemannian manifold, F a closed geodesically-convex subset of X, and t an excision*

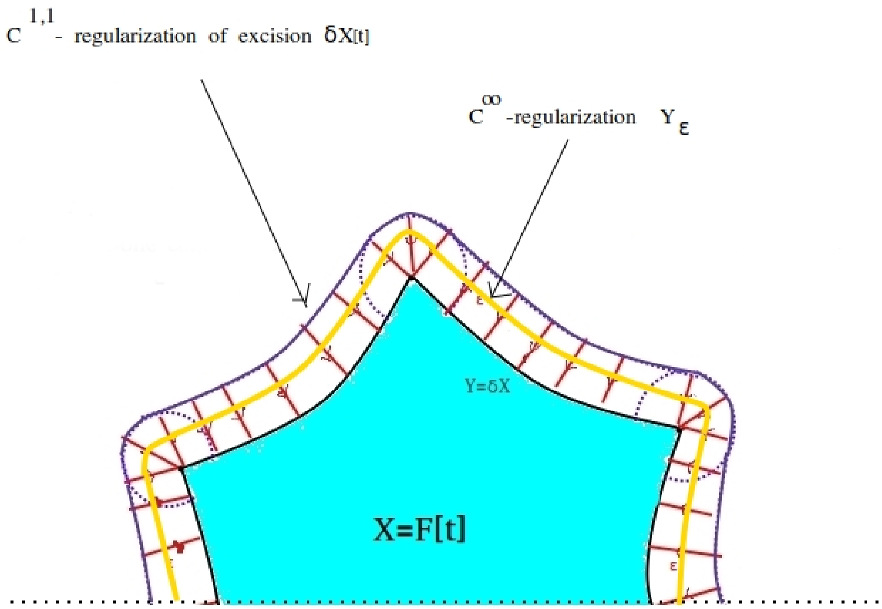

Figure 4.8: The outer boundary is the boundary of the closed outward ϵ-neighborhood $\partial((X[t]))_{\leq\epsilon}$ of the excision boundary $Y = \partial_* X[t]$. The inner hypersurface is the smooth ϵ-regularization Y_ϵ

parameter (see Definition 4.5.1) and $F[t]$ a convex excision. Define $Y = \partial_* F[t]$. Then for every $\epsilon > 0$, we can replace Y with a smooth manifold Y_ϵ such that:

(1) Y_ϵ is contained within the open ϵ-neighborhood of $\partial_* F[t]$; and

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

(3) there exists a degree-one continuous map $p : Y_\epsilon \to Y$.

Sketch of Proof 4.7.1. The following approach was suggested by Professor R.J. McCann. Recall the Definition 4.5.1 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 $\mathbb{R}^n$. 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 defining $Y_\epsilon := u^{-1}(1/2)$, we see Y_ϵ is indeed smooth manifold satisfying conditions (1), (2), (3). 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 4.7.1, 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 4.7.1 replaces a manifold-with-corners $(X, \partial X)$ with a pair (X, Y_ϵ) where Y_ϵ is the smooth regularization of the closed ϵ-neighborhood $Y = \partial((X))_{\leq \epsilon}$. We remark that the visibility between pairs (x, y) in $X \times \partial X$ and pairs $(x, y') \in X \times Y_\epsilon$ are basically equivalent via the degree-one continuous map p constructed in Proposition 4.7.1. Indeed if V_ϵ denotes the visibility relation on X, Y_ϵ per Definition 4.6.1, then we find $xV_\epsilon y'$ if and only if $xVp(y)$, where p is the degree-one map from Proposition 4.7.1.

Thus Y_ϵ is a smooth manifold having uniquely-defined outward unit normal vector $\mathbf{n}_y \in T_y X$ varying with $y \in Y_\epsilon$. Therefore $k(x, y)$ is smooth with respect to visible pairs (x, y) for all sufficiently small $\epsilon > 0$. For given $x \in F[t]$, the function $y \mapsto k(x, y)$ will possibly have nontrivial domain in Y_ϵ. Evidently $dom(k_x)$ coincides with the set of points y visible from x in $F[t]$.

4.8 Barycentres and Krein-Milman

The previous Section 4.2 introduced the definition of "extreme points" in anticipation of the present section, which 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] provide 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 $\mathscr{E}[F]$.

Definition 4.8.1 (Linear barycentre)**.** Suppose λ is a probability measure supported on a nonempty compact subset Y of some topological vector space X. A point x in X is the barycentre of λ, and denoted $bar(\lambda)$ if $f(x) = \int_X f(w) d\delta_x(w) dw = \int_X f(w) d\lambda(w)$ for every continuous linear functional f on X.

The linear barycentre can be generalized to Riemannian geometry with the following alternative definition, see [Jös97].

Definition 4.8.2 (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 $exp_x : T_x X \to X$, we find $q \in X$ is a barycentre of λ if the following critical point condition holds:

$$\int_X exp_q^{-1}(x) d\lambda(x) = 0 \text{ in } T_x X. \tag{4.4}$$

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

Proposition 4.8.3. *Let (X, d, σ) 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 σ, 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 $\sigma = \mathscr{H}$ is Hausdorff-type, e.g. $\mathscr{H}_Y = \mathscr{H}$, then we abbreviate $bar(-) = bar(-|\sigma)$. 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 4.8.4. 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 not typically 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 $\lambda_x^0 \in S_x$ varying continuously with the point $x \in F$. In our applications below, the convex set F is typically not a simplex.

Now suppose F is a compact and geodesically convex space satisfying (IDE) conditions of Definition 4.2.2. Then $Y := \mathscr{E}[F]$ has canonical Radon measure $\mathscr{H}_Y^{can}$. Following the recent work of Y.-H.Kim and B. Pass [KP18] we define:

Definition 4.8.5 ([KP18])**.** In the above notation, for every $x \in F$ we define the following probability measure λ_x^* to be the canonical element of S_x:

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

where W_2^2 denotes Wasserstein 2-distance with respect to the quadratic transport costs $c = d^2/2$.

With respect to the quadratic cost $c = d^2/2$, it is well-known that minimizers of W_2^2 are unique, and especially for subsets of a convex set F. We refer the reader to [KP18] for further details. We call λ_x^* the "canonical probability measure" with barycentre x.

4.9 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, $\{\lambda\} \subset \mathscr{E}[F]$, defining the excision $F[t] := \cap_{\{t\}}(F - W_\lambda^t)$ 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 4.9.1. Let $\Delta^v(\partial_* F[t])$ be the collection of probability measures λ on $\partial_* F[t]$ which satisfy the following two conditions:

(I) the barycentre $bar(\lambda)$ belongs to $F[t]$; and

(II) the support of λ is a subset of $\partial_* F[t]$ which is simultaneously 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 in this section is to identify a convenient geodesically convex subset Ω of $F[t]$. We define Ω via the visibility relation $V \subset F[t] \times F[t]$ (Definition 4.6.1).

Lemma 4.9.2. *Under the above hypotheses, 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. ☐

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

$$\Omega := (\cap_{z \in \partial_* F[t]} V_z). \tag{4.5}$$

See Figure 4.9. 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 $F[t]$.

Lemma 4.9.3. *Under the above hypotheses, the subset Ω 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 $F[t]$. ☐

The next second step concerns the restriction $proj_z|\Omega : \Omega \to \partial F[t]$ of the projection $proj_z$. Recall $proj_y(x)$ is the projection defined in Section 4.4, and we restrict to $z \in \partial_* F[t]$.

Lemma 4.9.4. *Under the above hypotheses, the restriction $proj_z|\Omega : \Omega \to \partial F[t]$ is a continuous map for every $z \in \partial_* F[t]$.*

Proof. We find $proj_z$ is discontinuous at points x for which the geodesic segment $\alpha(s)$, $s > 0$ with $\alpha(0) = z$, $\alpha(dist(z, x)) = x$, intersects $\partial_* F[t]$ 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 then nonvisible to a nontrivial portion of $\partial_* F[t]$, and therefore $x \notin \Omega$. ☐

The restricted projection $proj_z|\Omega$ will not surject onto $\partial F[t]$, and the image of $proj_z|\Omega$ is some closed connected subset of $\partial F[t]$. Moreover according to Lemma 4.9.2 every $y \in (proj_z|\Omega)$ is the barycentre of some probability measure $\lambda_y \in \Delta^v$, $bar(\lambda_y) = y$. Adapting the Definition 4.8.5, we let λ_y^* denote the canonical probability measure. This unique canonical measure varies continuously with respect to $y \in \partial F[t]$.

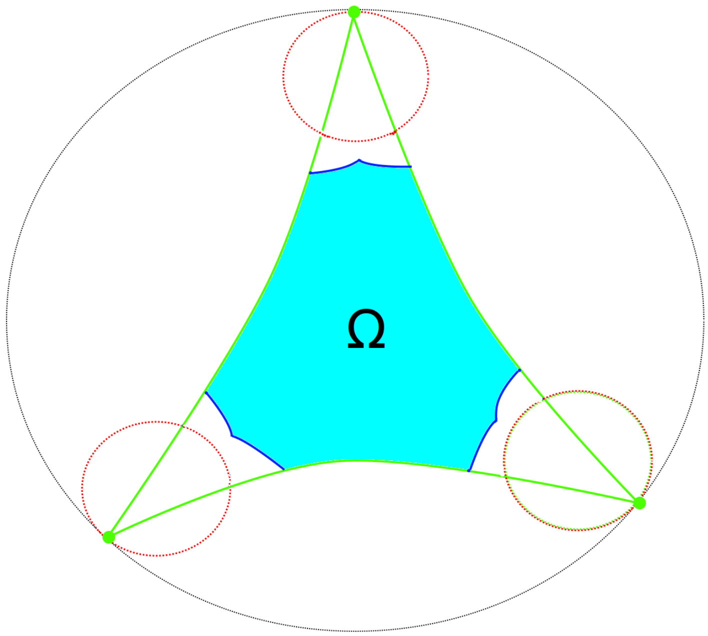

Figure 4.9: The subdomain Ω is geodesically convex

In the third step, we define our visibility cost v with the explicit formula (4.6). Let $\Omega \subset F[t]$ be the subset defined in (4.5). If $Y = \partial_* F[t]$ is the excision boundary (Definition 4.5.3), then let Y_ϵ be the ϵ-regularization from Proposition 4.7.1 for some small $\epsilon > 0$.

Definition 4.9.5 (Visibility Cost)**.** Under the above hypotheses, let

$$v(x, y_0) := \frac{1}{2} k(x, y_0).q(x, y_0)^{-1} + \int k(x, y).q(x, y)^{-1} d\lambda^*_{proj(y_0,x)}(y). \tag{4.6}$$

be the visibility cost $v : \Omega \times Y_\epsilon \to \mathbb{R} \cup \{+\infty\}$.

Here k denotes the visibility factor defined in 4.6.4, and $q(x, y) := dist(x, y)^{2+e}$ for some suitable integer $e \geq 0$ (recall Definition 4.2.3 from Section 4.2).

Under the above hypotheses, one finds the visibility cost $v = v(x, y)$ defined by equation (4.6) varies continuously with $(x, y) \in \Omega \times Y_\epsilon$. In fact we propose v has many further properties throughout $\Omega \times Y_\epsilon$, namely the hypotheses (D0)–(D4) from Section 4.1 which are expedient for our topological applications. We summarize our expectations in the following conecture:

Conjecture 4.9.6. *Under the above hypotheses, the visibility cost $v : \Omega \times Y_\epsilon \to \mathbb{R} \cup \{+\infty\}$ defined in Formula (4.6) satisfies the hypotheses (D0)–(D5) from Section 4.1.*

The reader will find applications of the above constructions in Chapters 5, 6 below.

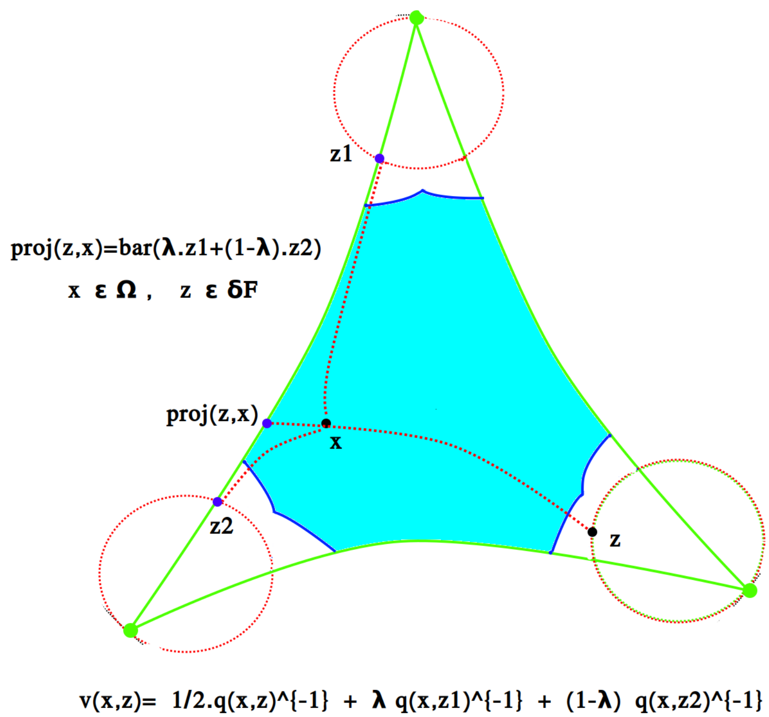

Figure 4.10: Evaluating visibility cost $v(x,z)$ for pair $(x,z) \in \Omega \times \partial_* F$

The definition 4.9.5 has useful limit case when the "radii" t_λ at $\lambda \in Y$ vanish $t_\lambda \to 0^+$. This limit produces the generalization of $\widetilde{c}$ to $\widetilde{c}^*$:

Definition 4.9.7 (Averaged Two-Pointed Repulsion $\widetilde{c}^*$)**.** Let F be a compact convex set. Then the formula

$$\widetilde{c}^*(x, y_0) := \frac{1}{2} q(x, y_0)^{-1} + \int_Y q(x,y)^{-1} d\lambda^*_{proj(y_0,x)(y)}$$

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

We consider the above definition of $\widetilde{c}^*$ as an "averaged two-pointed" repulsion type cost, and defer the detailed study of this generalized cost to future investigations. However we expect that the hypotheses of semiregularity of F are sufficient to ensure $\widetilde{c}^*$ satisfies (Twist) condition. We formalize this expectation in the following

Conjecture 4.9.8. *Let $F = conv[G, G', \ldots]$ be the compact convex hull of a finite collection of pariwise-isometric regular polyhedral gates $G, G', \ldots$. Then the averaged two-pointed repulsion cost $\widetilde{c}^* : F \times \mathscr{E}[F] \to \mathbb{R} \cup \{+\infty\}$ satisfies Assumptions (D0)–(D5)) throughout its domain.*

Remark. If F is compact and strictly convex with $\partial F = \mathscr{E}[F]$, then the formula 4.6 defining $v(x,y)$ reduces to a modified "visual two-pointed" repulsion cost, namely

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

Omitting the visibility factor k obviously yields $\widetilde{c}$ (Definition 4.4.1). This visual two-pointed cost v is a topic for future investigations, e.g. it would be useful to establish Conjecture 4.9.6.

Chapter 5

Symmetry and Dimension

This chapter begins the second phase of our thesis with the goal of establishing Theorem C from Section 1.5. Our Theorem C is a reduction program drawing together the semi-coupling and singularity methods of Chapters 2, 3, 4, and applies these methods to source spaces $X[t]$ obtained by convex excision, with target $Y = \partial_* X[t]$ the topological excision boundary. For applications we imagine a third party user having an infinite discrete group Γ and some standard geometric $E\Gamma$ model $X \times \Gamma \to X$. Given this initial data, we outline an excision construction in §5.3 to obtain manifold-with-corners $X[t] \times \partial X[t]$. If the user can successfully Close the Steinberg symbol §6, then we replace the excision $X[t]$ with a chain sum $\underline{F}$. The summands of chain sum $\underline{F}$ are excised convex sets $F[t]$, and $\underline{F}$ inherits a proper Γ-action, where Γ acts like shift-operator on the summands. We install the two-pointed repulsion cost $\widetilde{c}$ and visibility cost v on $\underline{F}$, and propose the homotopy-reductions from Theorems 2.6.2 and 3.4.2 to 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, where "small" depends on sufficient (UHS) conditions being satisfied.

5.1 Proper Classifying Spaces $E\Gamma$

Let Γ be a finitely-generated infinite group. It is difficult to speak further in the abstract category of groups, but 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 $E\Gamma$. Recall the topological definition:

Definition 5.1.1. 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 $\gamma = Id$.

The $E\Gamma$ models can be characterized as universal covering spaces of the Eilenberg-Maclane space $K(\Gamma, 1)$. One finds every pair X, X' of $E\Gamma$-models are homotopic. Useful references include [Bro82], [Bre93]. Examples of $E\Gamma$ spaces abound.

1. 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 $(x, n) \mapsto x + n$.

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

3. 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 $E\pi_1(X, pt)$ model.

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

5. 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 $Sp(\mathbb{Z}^4)$.

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

Further examples include the braid groups, and mapping class groups, and almost all the groups that arise from geometric group theory and three-dimensional topology.

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

Definition 5.1.2. An $E\Gamma$ model X is geometric if X has a complete Cartan-Hadamard metric d 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).

The standard $E\Gamma$ models are usually geometric.

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 5.1.1 means we can distinguish γ from the identity element Id_Γ 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 5.1.1 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, especially when the group is infinite. 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 $\mathscr{M}_{\geq 0}(X)_\Gamma \approx \mathscr{M}_{\geq 0}(X/\Gamma)$.

5.2 Group Cohomology: Algebraic Background

The following section is primarily notation and review of the basic facts of group-cohomology, i.e. the study of projective and free resolutions of $\mathbb{Z}\Gamma$-modules. The formalities are necessary for the definition of Closing the Steinberg symbol (Chapter 6) and Bieri-Eckmann duality (Section 5.4). 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.

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.

If we augment the complex by the augmentation mapping $\epsilon : C_0(X) \to \mathbb{Z}$ (defined by $\epsilon(v) = +1$ for every 0-cell on X) then we obtain a projective resolution (free) 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. Forgetting the Γ-action, the resolution is homologically-trivial. But Γ acts on everything, and the $\mathbb{Z}\Gamma$-modules become essential topological invariants of Γ. The comparison to the singular homology with coefficients in $\underline{Z}$ arises from the augmentation mapping $\mathbb{Z}\Gamma \to \underline{\mathbb{Z}}$ defined by

$$\epsilon : \sum_\gamma n_\gamma \gamma \mapsto \sum_\gamma n_\gamma.$$

Now we require basic lemma (c.f. [Bro82, VIII.7.4, pp.208]):

Lemma 5.2.1. *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 $f_{id} \mapsto \Sigma_\gamma f_{id}(-.\gamma^{-1})$. □

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 only finitely many nonzero.

When Γ acts cocompactly on the contractible space X, then compactly-supported cohomology on X can be identified by a Poincaré-Lefschetz duality

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

where Ω 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. This duality is generalized to finite-volume quotients (usually noncompact) in Bieri-Eckmann's duality. See Section 5.4 below.

5.3 Excision versus Compactification

This section emphasizes the distinction between excisions and compactifications. This is key to our studying the semicoupling programs associated to various $E\Gamma$-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 5.1.2 from Section 5.1.

M.Gromov defined a universal compactification for general complete metric spaces, 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 a large-dimensional closed topological disk. The isometric action $X \times \Gamma \to X$ extends to a continuous action by homeomorphisms $\overline{X} \times \Gamma \to \overline{X}$. Brouwer's fixed point theorem implies every such continuous homeomorphism has at least one fixed point in $\overline{X}$. The visual boundary $X(\infty)$ inherits a natural metric (so-called spherical Tits metric) and supports a uniform Lebesgue measure. However there is serious difficulty: the Γ-action on $X(\infty)$ is neither free nor properly discontinuous. Therefore Radon measures on $X(\infty)$ will not descend to Radon measures on the topological quotient $X(\infty)/\Gamma$. In fact it appears to be general principle that compactifications $\overline{X}$ of geometric $E\Gamma$ models X cannot support nontrivial Γ-equivariant Radon measures, c.f. Ahlfor's theorem, [Thu02, §8.4.2, pp.180]. 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 $\mathbb{H}^2(\infty) = S^1$, and it's well-known that $PGL(\mathbb{Z}^2)$ acts ergodically on this circle at-infinity with respect to the one-dimensional uniform measure $d\theta$. Therefore every $PGL(\mathbb{Z}^2)$-equivariant $d\theta$ measurable function on $\mathbb{H}^2(\infty)$ is constant. This says all equivariant Radon measures on the boundary circle are trivial.

So conventional compactifications are not suitable for our semicoupling method, which studies transport between Radon mesaures. Thus we pursue a general excision procedure, implicit in the literature, and defined for a discrete groups Γ and Cartan-Hadamard spaces (X, d, σ) augmented with a free and properly discontinuous isometric action $X \times \Gamma \to X$. The basic idea: 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 at-infinity. The dark zones W_λ are disjoint from all Γ-accumulation points at visual infinity. We excise these halfspaces, and obtain a manifold-with-corners

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

The excision X_0 has topological boundary $\partial X_0 \subset X$. Observe 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. Moreover the fact that Γ-orbits avoided the dark zones W_λ implies closed geodesic loops on the quotient X_0/Γ are everyone contained in a compact-core of X_0/Γ.

Here is formal description of the excision procedure, to be applied to arithmetic groups in Section 5.5. The construction is summarized in the Figure 5.3, where the formal definitions are described below.

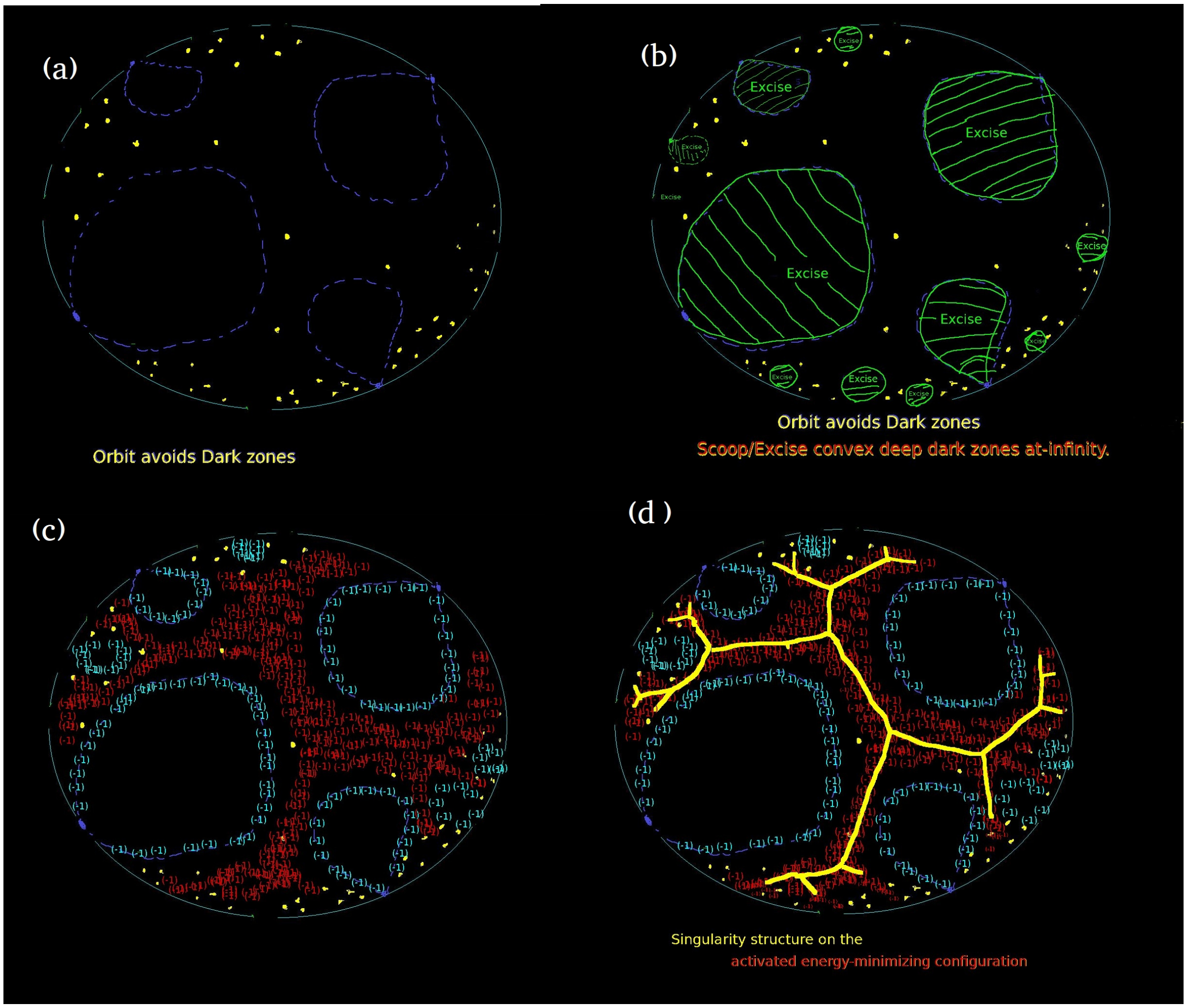

Figure 5.1: (a) Deep Zones are disjoint from Γ-orbit. (b) Excise the Deep Zones from X. (c) Optimal Semicoupling, with respect to repulsion cost, between target excision boundary (in blue) and activated source measure (in red). (d) Singularity structure (in yellow) of an optimal semicoupling between source and target with respect to repulsion cost.

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 $x \in X$, we define the horofunction $h_{\lambda,x} : X \to \mathbb{R}$ by the following formula (see [BGS85, §3]).

Definition 5.3.1 (Horofunctions)**.** Let $s : [0, +\infty] \to X$ be geodesic ray diverging to $s(+\infty) = \lambda$ with $s(0) = x$. Then the horofunction centred at λ is defined

$$h_{\lambda,x}(x') := \lim_{t\to+\infty} d(x', s(t)) - d(x, s(t)) = \lim_{t\to+\infty} (d^2(x', s(t)) - d^2(x, s(t))\frac{1}{2t}.$$

There are several equivalent characterizations of horofunctions. 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 $\{x.\gamma^n | n \in \mathbb{Z}\}$ 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 observation suggests the following definition.

Definition 5.3.2. 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 5.3.3 (Γ-rationality)**.** A Γ-invariant subset $J \subset X(\infty)$ is Γ-rational if the group homomorphisms $\{T_\lambda : \Gamma_\lambda \to \mathbb{R} | \lambda \in J\}$ are simultaneously trivial.

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 $\{T_\lambda | \lambda \in J\}$ 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 – which are contractible connected subsets of X – the homotopy-type and homology of ∂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 $\Gamma = G(\mathbb{Z})$, the excision is defined by convex horoballs corresponding to $\mathbb{Q}$-rational parabolic subgroups of 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 $\mathscr{B}(G, T)$ constructed in Section 5.5. The Solomon-Tits theorem identifies the simplicial complex $\mathscr{B}(G, T)$ 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 §5.4 with Bieri-Eckmann homological duality.

5.4 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 $\dim(X) = \dim(X, \sigma)$ 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 impossible to implement in reality. Our thesis provides new general method for displaying the small-dimensional models according to the Reduction-to-Singularity principles of the previous chapters. The "external" construction above is replaced by the reduction of an initial $E\Gamma$ model X to a closed subvariety $Z \subset X$.

Indeed, there are numerous large-dimensional $E\Gamma$ models available. These standard $E\Gamma$-models will have space dimension much greater than the cohomological dimension. A precise determination of the dimension-gap is achieved in Borel-Serre's formula $vcd(G(\mathbb{Z})) = \dim(K/{}^0G(\mathbb{R})) - rank_{\mathbb{Q}}({}^0G)$, c.f. [BS73, §8.6], whenever G is a $\mathbb{Q}$-reductive linear algebraic group. Their method is very general, c.f. Theorem 5.4.4 below. The argument of Borel-Serre is based on the construction of rational bordification models $\overline{X}^{BS,/\mathbb{Q}}$, and the fact that $\Gamma = G(\mathbb{Z})$ satisfies a homological duality generalizing Poincaré duality as discovered by Bieri-Eckmann [BE73].

Definition 5.4.1. 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 5.4.2 (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 $\nu - * < 0$. $\square$

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 5.4.3 ([BE73], [BS73]). *Let $(X[t],\partial X[t])$ be a Γ-equivariant rational excision model, where $X[t],\partial X[t]$ support invariant Borel 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 $* = q$,*

$$\tilde{H}_*(\partial\overline{X}^{BS/\mathbb{Q}};\mathbb{Z}) = \begin{cases} 0, & \text{if } * \neq q, \\ \boldsymbol{D}, & \text{if } * = q. \end{cases} \tag{5.1}$$

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 5.4.4 (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*

$$vcd(\Gamma) = dim(X) - (q+1).$$

5.5 Arithmetic Groups: Excision

For the applications of our optimal semicoupling methods, we recommend the patient reader focus on Theorem 5.5.2. 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 $\Gamma = G(\mathbb{Z})$ an arithmetic group. The main applications we pursue are for $\Gamma = G(\mathbb{Z})$ the arithmetic group of "standard" higher-rank $\mathbb{Q}$-reductive group $G = GL(V)$, the symplectic groups $Sp(\mathbb{R}^4, \omega)$, $Sp(\mathbb{R}^6, \omega)$, $\ldots$, and the split-orthogonal groups $O(V^{p,q})$ for $p, q \geq 2$.

The problem of constructing small-dimensional classifying spaces is 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 (c.f. Theorem 5.4.4 from §5.4). But no general principle for constructing large codimension (≥ 2) deformation retracts for the standard geometric models was available, with the exception of $GL(\mathbb{Z}^N)$. A general method for reducing the general linear group $GL(\mathbb{Z}^N)$ was discovered by the so-called systolic well-rounded retract argument introduced in [Sou78], and extended [Ash84]. For instance, Soulé's method produces an interesting three-dimensional cube model for the two-dimensional deformation retract of the five-dimensional symmetric space $\approx SO(3)\backslash SL(\mathbb{R}^3)$ predicted by Borel-Serre's formula. The well-rounded retraction is an elementary and remarkable construction, but the apparent generalizations for different symmetry groups, e.g. the lagrangian systoles or null systoles arising from Sp or $O_{p,q}$ geometries, do not enable large codimension retractions. The reader is referred to [Gro] for further references concerning systoles. Our program proposes new homotopy-reductions which appear to be genuine generalizations of Ash-Soulé 's well-rounded retract.

So let the user choose an arithmetic group $\Gamma = G(\mathbb{Z})$. With respect to our Reduction Program, as described in Theorem C (Section 1.5.1), we begin with construction excision models from a standard given $E\Gamma$ geometric model. This key first step originates with [BS73]. The excision is obtained from the initial $E\Gamma$ model $X = K\backslash^0 G(\mathbb{R})^0$, 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 $\sigma \otimes \tau = dvol_X \otimes dvol_Y$, 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 5.5.2, 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 G. The generality of the construction forces us to speak in terms of the Bruhat-Tits structure theory of G. We assume G is totally split over $\mathbb{Q}$, and so maximal $\mathbb{Q}$-algebraic tori T in G are totally $\mathbb{Q}$-split and admit $\mathbb{Q}$-rational isomorphisms $T \approx \prod_{rank_{\mathbb{Q}}(G)} \mathbb{G}_m$ onto a product of multiplicative groups. The excision construction involves the Γ-orbit of the set of $\mathbb{Q}$-coroots Φ^{co}, where $\Phi = \Phi(G, T) \subset Hom_{/\mathbb{Q}}(T, \mathbb{G}_m)$ is the root system of G with respect to a maximally $\mathbb{Q}$-split algebraic torus. 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 G.

Definition 5.5.1. Fix a basepoint $q = K$ in $K\backslash^0 G(\mathbb{R})^0$, 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 $V[k^{\gamma}, t]$ be the convex horosphere consisting of all matrix elements $g \in G(\mathbb{R})^0$ for which the $A_{P(k^{\gamma}),q.\gamma}$-coordinate is less than or equal to $exp(t(k^{\gamma}))$. 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.

The excision model with respect to the basepoint o and excision parameter $t : \Phi_{co}^{\Gamma} \to$

$\mathbb{R}_{>0}$ is then defined to be

$$X[t] := X - \cup_{\Phi^{\Gamma}_{co}} V[k^{\gamma}, t].$$

Theorem 5.5.2. *Let G be the $\mathbb{Q}$-split form of a semisimple $\mathbb{Q}$-linear algebraic group for which $Hom_{/\mathbb{Q}}(G, \mathbb{G}_m)$ is trivial. Let $\Gamma := G(\mathbb{Z})$ be arithmetic group. Suppose $t : \Phi^{\Gamma}_{co} \to \mathbb{R}_{>0}$ 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 $q := rank_{\mathbb{Q}}(G) - 1$, 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 $Hom_{/\mathbb{Q}}(G, \mathbb{G}_m)$ is trivial descends by induction to all the semisimple parts ${}^0L$ of the various Levi factors $L = L_{P,o.\gamma}$ of $\mathbb{Q}$-parabolic subgroups P with respect to the Γ-orbits of the basepoint o. 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 $\{V[k,t]\}$ produces a covering of $\partial X[t]$ by contractible open sets whose nerve is isotopic to the spherical Tits building $\mathscr{B}(G, \mathbb{Q})$. By Weil's nerve theorem, there is natural homotopy-isomorphism between $\partial X[t]$ and $\mathscr{B}(G, \mathbb{Q})$. But the well-known Solomon-Tits theorem proves $\mathscr{B}(G, \mathbb{Q})$ has the homotopy-type of a countable wedge of $(rank_{\mathbb{Q}}(G) - 1)$-dimensional spheres. This proves (iv). □

Remark. The equivariant excision parameter $t : \Phi^{\Gamma}_{co} \to \mathbb{R}_{>0}$ 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 6

Closing the Steinberg symbol

6.1 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 6.2 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.

Figure 6.1: Stitching a football F from identical regular hexagons or pentagons P_i

To stitch a football from panels $\{P_i | i \in I\}$ 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 $F := conv\{P|$ 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 $\{P_i\}_I$ are pairwise isometric to some regular geodesically-flat polygon P.

Figure 6.2: 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.

We furthermore assume there is an isometric action by a discrete symmetry group Γ translating the polygon patches throughout space $P.\gamma$ for $\gamma \in \Gamma$. The Γ-symmetries lead to convex 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". Indeed panels are contractible, and 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.

6.2 Closing Steinberg : Definition and Consequence

Let $X \times \Gamma \to X$ be a geometric $E\Gamma$ model (Definition 5.1.2). Suppose we define a Γ-equivariant family of convex horospheres $\{V[t]\}$, producing an excision model $X[t] :=$

$X - \cup_t V[t]$ whose topological boundary $\partial X[t]$ is Γ-invariant and has reduced singular homology concentrated at a dimension. So there exists some integer q such that

$$\tilde{H}_q(\partial X[t]; \mathbb{Z}) = \mathbf{D} \neq 0, \quad \text{and } \tilde{H}_*(\partial X[t]; \mathbb{Z}) = 0 \text{ when } * \neq q.$$

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 $\mathbf{D}$.

Imagine a relative cycle $[B] \in H_{q+1}(X[t], \partial X[t]; \mathbb{Z})$, with $[B] \neq 0$. The long exact sequence of relative homology produces isomorphism

$$\partial : H_{q+1}(X[t], \partial X[t]; \mathbb{Z}) \simeq H_q(\partial X[t]; \mathbb{Z}),$$

so the boundary $\partial[B]$ represents nontrivial cycle in $H_q(\partial X[t]; \mathbb{Z})$. The group Γ symmetries acts by "flips" and "translates" of the base cycle $[B]$ throughout the space . And every finite subset I of Γ produces a finite chain sum

$$SUM[[B].I] = [B] + [B].\gamma + \cdots ,$$

with total chain boundary

$$SUM[\partial[B].I] = \partial[B] + \partial[B].\gamma + \cdots .$$

The basic problem of Closing Steinberg is to produce a finite collection of flips $I \subset \Gamma$ for which the total boundary vanishes mod 2. I.e., for which the total chain sum of relative cycles $[B] + [B].\gamma + \cdots$ assembles to a closed chain with respect to $\mathbb{Z}/2$ coefficients. 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-angled cube into a three-dimensional sphere.

In algebraic terms, Closing Steinberg involves constructing a nonzero chain element $\xi \in C_0(X[t]) \otimes \mathbb{Z}/2\Gamma \otimes \mathbf{D}$ representing a nonzero element $0 \neq [\xi] \in H_0(\Gamma; \mathbb{Z}/2\Gamma \otimes_{\mathbb{Z}} \mathbf{D})$. To represent ξ in the chain group (with coefficients over $\mathbb{Z}/2$) requires determining a finite subset $I \subset \Gamma$ for which the flat-filled relative chain $[B] \in C_q(X, \partial X; \mathbb{Z}/2)$ has translates $\xi = SUM[B.I]$ in $C_0(X) \otimes \mathbb{Z}/2 \otimes \mathbf{D}$ representing a nontrivial 0-cycle over $\mathbb{Z}/2$-coefficients.

The complete definition of Closing Steinberg includes further geometric conditions on the Γ-translates $F.\Gamma$ of the the closed convex hull $F = conv[B.I]$ of the translates $B.I$. Let $X[t], \partial X[t]$ be a Γ-invariant excision model. Let $[B]$ be a flat-filled relative cycle representing a nonzero generator of $H_{q+1}(X[t], \partial X[t]; \mathbb{Z})$.

Definition: Closing Steinberg 6.2.1. *A finite subset I of Γ successfully Closes Steinberg if:*

(i) *(nontrivial mod 2) the chain $\xi = SUM[B.I]$ is nonvanishing over $\mathbb{Z}/2$ coefficients in the chain group $C_{q+1}(X[t], \partial X[t]; \mathbb{Z}/2)$;*

(ii) *(vanishing boundary mod 2) the restricted chain $\partial\xi = SUM[\partial[B].I]$ on $\partial X[t]$ vanishes over $\mathbb{Z}/2$-coefficients;*

(iii) *(well-defined convex hull) the boundary-chain representing $\partial\xi$ is simultaneously visible from an interior point x in $X[t]$;*

(iv) *(well-separated gates) the convex chain sum $\underline{F} = SUM[F.\Gamma]$ has nonempty well-separated gates structure precisely equal to the principal orbit $\{B.\gamma | \gamma \in \Gamma\}$.*

In the above setting ∂B is coincident to $B \cap \partial X$, and hypothesis (ii) says $SUM[\partial[B].I]$ is identically zero in $C_q(\partial X[t]; \mathbb{Z}/2)$. The hypothesis (iv) on well-separated gates structure means a pair of translates $F, F.\gamma$ are either disjoint or the gate $F \cap F.\gamma$ coincides with some translate $B.\gamma'$.

The next proposition establishes existence of nontrivial kernel vectors $\ker \partial_0 \cap B.\Gamma$, and proves the existence of finite subsets I which satisfy the first two conditions of Definition 6.2.1. In otherwords there exists no formal homological obstructions to successfully Closing Steinberg.

Proposition 6.2.2. *There exists finite subsets I in Γ for which $SUM[B.I]$ lies in the kernel of ∂_0 over $\mathbb{Z}/2$.*

Proof. The argument is homological. We interpret $\xi = SUM[B.I]$ as a chain sum representing a 0-cycle in $H_0(\Gamma; \mathbb{Z}/2\Gamma \otimes_{\mathbb{Z}\Gamma} D)$. The hypotheses of Closing Steinberg imply ξ is homologically nontrivial cycle. As consequence of Bieri-Eckmann duality (Proposition 5.4.2), the kernel $\ker \partial_0$ is canonically isomorphic to the induced $\mathbb{Z}\Gamma$-module $\mathbb{Z}/2 \otimes_{\mathbb{Z}} D$, which is nonzero. □

Our definition of Closing Steinberg was inspired by the author's study of [Cre84]. Indeed 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 related to determining a suitable "relation ideal $\mathscr{R}$" and constructing a "basic polyhedron B whose transforms fill the space". The interested reader should compare [Cre84, pp.290], where Cremona successfully Closes Steinberg with certain finite subgroups I' of Γ. We present some further examples in §§6.3–6.4 below.

Now suppose we find a finite subset I in Γ with $\partial_0[SUM[B.I]] = 0$, which after Proposition 6.2.2 is always possible. Then we have partially Closed Steinberg, except the the orbit $B.I$ may not admit a simultaneously visible interior point and the gates of the convex chain sum $\underline{F} = SUM[F.\Gamma]$ may not be well-separated. That a formal solution can always be made visible is the subject of the following conjecture, which we have been incapable of formally proving.

Conjecture 6.2.3. *Let I be finite subset of Γ realizing Proposition 6.2.2. Then there exists subset I' of I which satisfies Proposition 6.2.2 and has well-defined convex hull in $X[t]$. So there exists interior point $x' \in X[t]$ which is simultaneously visible from the translates $B.I$.*

All our hypotheses regarding Closing Steinberg has useful consequences, which we summarize in the following theorem.

Theorem 6.2.4. *Suppose $I \subset \Gamma$ successfully Closes Steinberg (Definition 6.2.1). Define $F := conv[B.I]$ and $\underline{F} = SUM[F.\Gamma]$.*

(i) The Γ-translates $F.\gamma$, $\gamma \in \Gamma$, form a chain sum

$$\underline{F} := \cdots [F]\gamma + [F]\gamma' + [F]\gamma'' + \cdots ,$$

and Γ virtually acts as additive shift-operator on the summands of $\underline{F}$.

(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 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 $[B].\gamma$, for $\gamma \in \Gamma$ freely. Then $[B].\gamma \neq [B]$ when $\gamma \neq id$. The definition of Closing Steinberg implies distinct translates F, F' are disjoint unless they intersect in a gate $G' = B.\gamma'$ for some $\gamma' \in \Gamma$. So $F.\gamma$ equals F only if $\gamma = id$ is trivial. This proves the summands $\{F.\gamma | \gamma \in \Gamma\}$ 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 $B.I$ in $X[t]$ proves $F = conv[B.I]$ is a compact convex set, and homeomorphic to some cube. Therefore (ii). □

6.3 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 projectivization limits 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 6.3, 6.3.

(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 6.3.

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 $\underline{F}_0 = SUM[F(I_0).PGL(\mathbb{Z}^2)]$, 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

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 6.3

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 C from Section 1.5. See Figure 6.3, 6.3

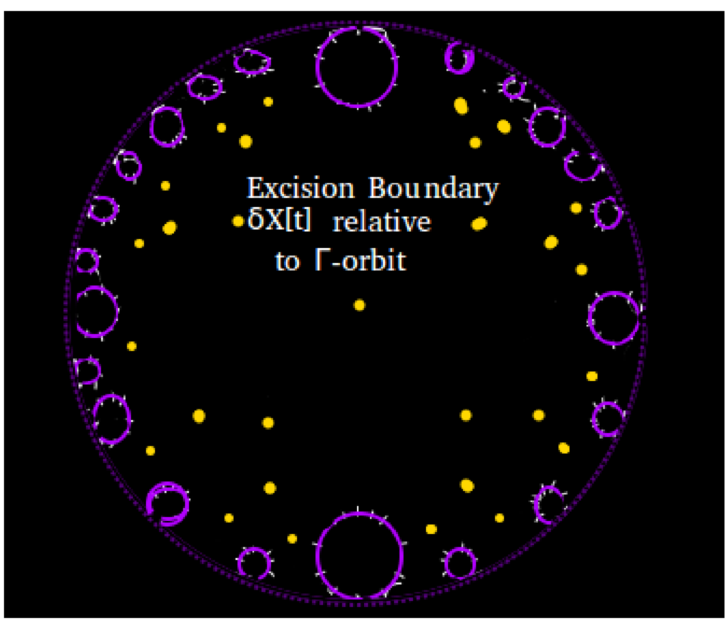

Figure 6.3: Convex excision $X[t]$ relative a Γ-equivariant excision parameter $t : \mathbb{Q}P^1 \to \mathbb{R}$

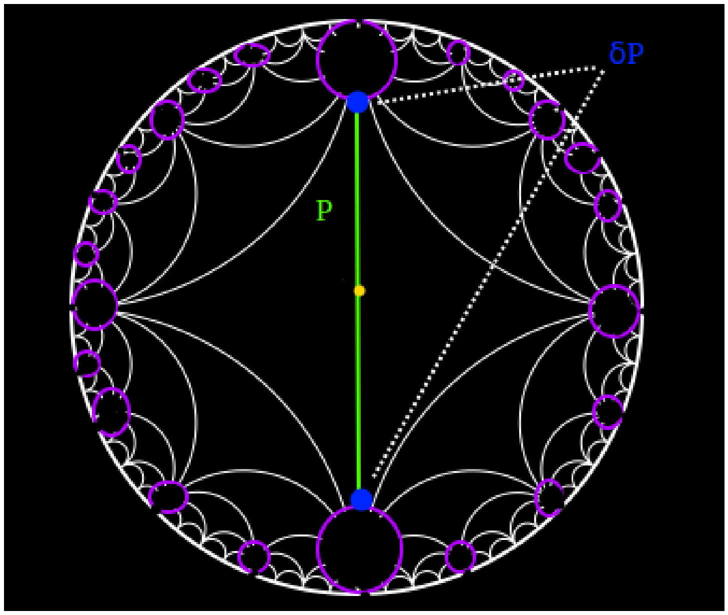

Figure 6.4: Steinberg symbol in $X[t] \times \partial X[t]$ is represented as relative 1-cycle P with boundary ∂P equal to 0-sphere.

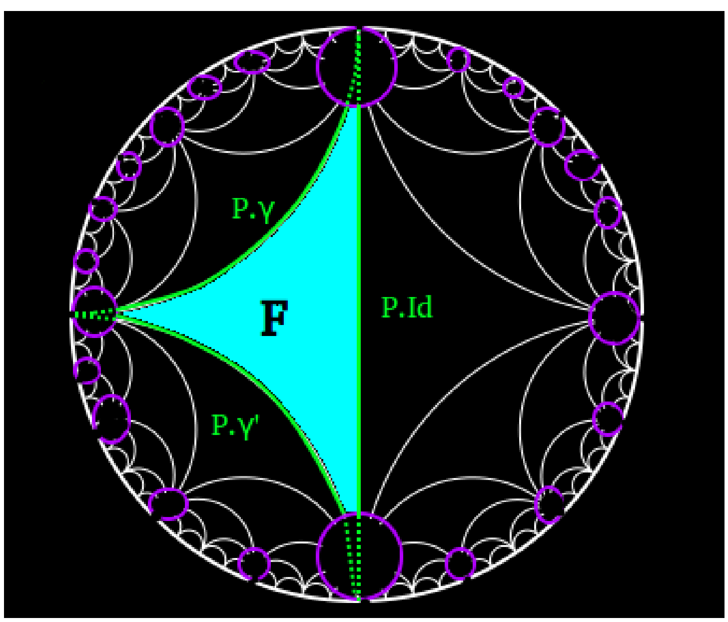

Figure 6.5: The translates by $I_0 = \{Id, \gamma, \gamma'\}$ successfully Close the Steinberg symbol.

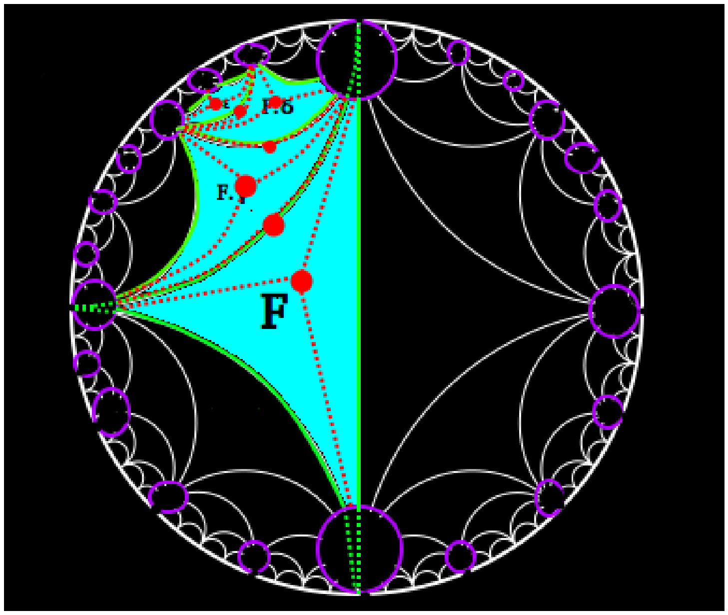

Figure 6.6: Evaluating the repulsion cost relative to the repulsion-cost $\widetilde{c}$ at various source points $x, x', x'', \ldots$ etc. in $X[t]$

Figure 6.7: Active Domain for optimal semicoupling with respect to repulsion cost is homotopy-equivalent to the source, c.f. Theorem A.

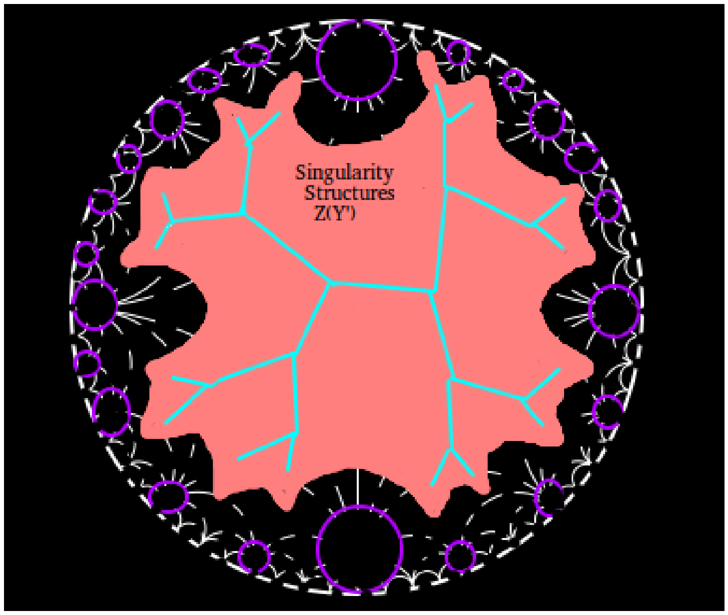

Figure 6.8: Singularity Structure of optimal semicoupling. The active domain is homotopy-equivalent to one-dimensional tree, c.f. Theorem B.

6.4 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 $[B'] = [B].\gamma$ for which the chain sum $SUM[\{B'\}']$ 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 6.4.1. *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. □

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 6.2.4 we find

Proposition 6.4.2. *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 $\underline{F} := SUM[F.GL(\mathbb{Z}^3)]$. 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 $Z = Z(\widetilde{c}, \sigma, \tau) : 2^{\partial[\underline{F}]} \to 2^{\underline{F}}$ realizes a spine for $GL(\mathbb{Z}^3)$.

Conjecture 6.4.3. *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]. We leave that to future investigations.

Chapter 7

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 general results are mainly the Theorems A,B, from the Introduction, namely Theorem 3.4.2, 3.4.3. For applications, our main result is Theorem C, 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 still stands, c.f. Conjecture 4.9.6.

The present thesis has developed the keystone fact that "the disk $X = D$ admits no continuous retraction onto its boundary $Y = \partial D$" (recall Section 1.2) and revealed 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 §3.1.1. 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 a new polyhedral decomposition of a source space X, and contravariantly parameterized by the target space Y. Given uniform Halfspace (UHS) conditions (Definition 3.4.1), our main Theorems A, B 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 6.2.1) and Theorem 6.2.4). Many new applications are possible, which we shall develop in our 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).

[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.

[Fig10] 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.

[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.

[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.

[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.

[Gro10] M. Gromov. "Manifolds: Where do we come from? What are we? Where are we going?" In: (2010). URL: www.ihes.fr/~gromov/PDF/manifolds-Poincare.pdf.

[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.

[GS80] Branko Grunbaum and G. C. Shephard. "Tilings with congruent tiles". In: *Bull. Amer. Math. Soc. (N.S.)* 3.3 (Nov. 1980), pp. 951–973. URL: https://projecteuclid.org:443/euclid.bams/1183547682.

[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).

[Sou78] C. Soulé. "The cohomology of SL(3,Z)". In: *Topology* 17 (1978), pp. 1–22.

[Spi71] M. Spivak. *Calculus On Manifolds: A Modern Approach To Classical Theorems Of Advanced Calculus*. Avalon Publishing, 1971.

[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/.

[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.

[Yos68] K. Yosida. *Functional Analysis*. 2nd ed. Die Grundlehren der mathematischen Wissenschaften. Band 123. Springer-Verlag, 1968.