The Structure of Groups with a Quasiconvex Hierarchy

Daniel T. Wise

ANNALS OF MATHEMATICS STUDIES

Annals of Mathematics Studies
Number 209

The Structure of Groups with a Quasiconvex Hierarchy

Daniel T. Wise

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD
2021

Published by Princeton University Press
41 William Street, Princeton, New Jersey 08540
6 Oxford Street, Woodstock, Oxfordshire OX20 1TR

press.princeton.edu

Library of Congress Cataloging-in-Publication Data

Names: Wise, Daniel T., 1971– author.
Title: The structure of groups with a quasiconvex hierarchy / Daniel T. Wise.
Description: Princeton : Princeton University Press, 2021. | Series: Annals of mathematics studies ; Number 209 | Includes bibliographical references and index.
Identifiers: LCCN 2020040293 (print) | LCCN 2020040294 (ebook) | ISBN 9780691170442 (hardback) | ISBN 9780691170459 (paperback) | ISBN 9780691213507 (ebook)
Subjects: LCSH: Hyperbolic groups. | Group theory.
Classification: LCC QA174.2 .W57 2021 (print) | LCC QA174.2 (ebook) | DDC 512/.2—dc23
LC record available at https://lccn.loc.gov/2020040293
LC ebook record available at https://lccn.loc.gov/2020040294

British Library Cataloging-in-Publication Data is available

Editorial: Susannah Shoemaker and Kristen Hop
Production Editorial: Nathan Carr
Production: Jacquie Poirier
Publicity: Matthew Taylor

This book has been composed in LaTeX

Printed on acid-free paper. ∞

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

to Yael

Contents

Acknowledgments

My research was supported by NSERC. The first version of this document was completed while on sabbatical at the Hebrew University in 2008-2009, and I am especially grateful to Zlil Sela and the Einstein Institute of Mathematics as well as the Lady Davis Foundation for its support. I am also grateful to Ian Agol, Macarena Arenas, Hadi Bigdely, Dave Futer, Mark Hagen, Frédéric Haglund, Jingyin Huang, Jason Manning, Eduardo Martinez-Pedroza, Denis Osin and his class, Piotr Przytycki, and Eric Swenson for many corrections. In addition, I wish to thank production editor Nathan Carr and illustrator Laurel Muller of Princeton University Press for their contributions.

This project would not have been possible without the continuing support of my darling wife Yael. I thank our children Tali, Steve, Ari, and Yoni for their patience. Finally, I thank my parents Batya and Mike for their continual confidence-building measures.

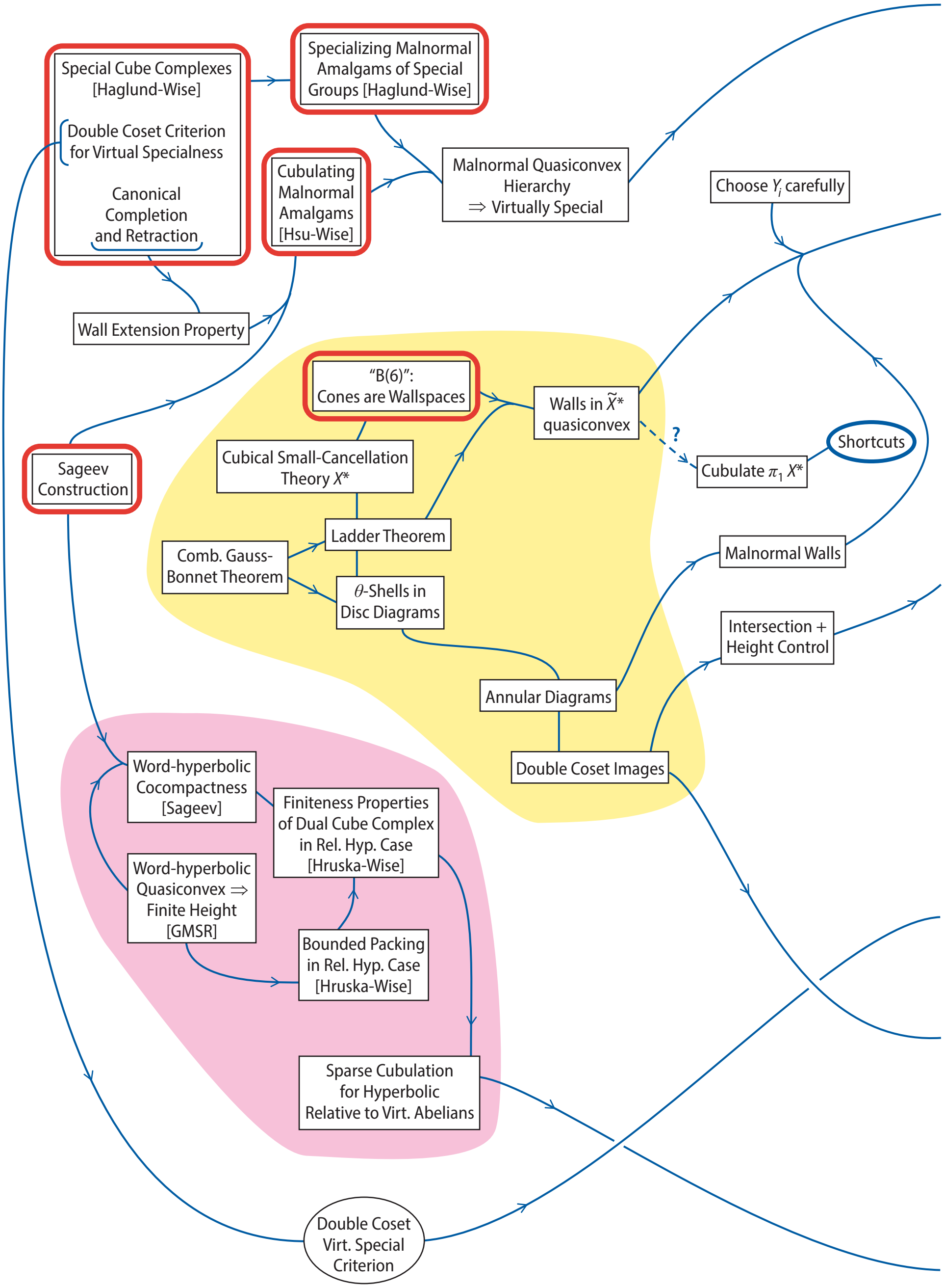

Figure 1. Flow chart (left side) indicating main points and some topical constellations.

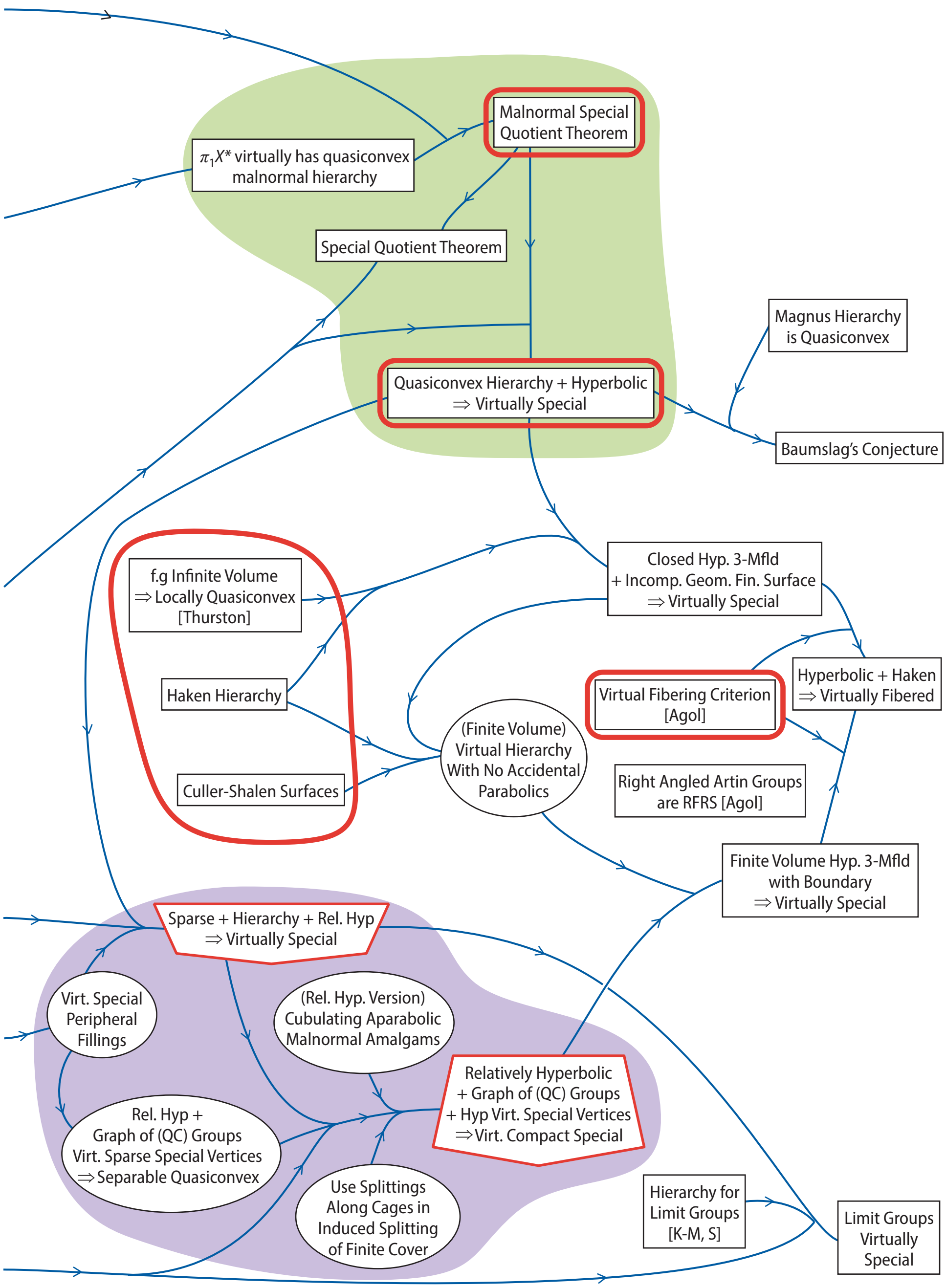

Figure 2. Flow chart (right side) indicating main points and some topical constellations.

The Structure of Groups with a Quasiconvex Hierarchy

Chapter One

Introduction

This text has several parts:

In the first part of the text we develop a small-cancellation theory over cube complexes. When the cube complex is 1-dimensional, we obtain the classical small-cancellation theory, as well as the closely related Gromov graphical small-cancellation theory.

It is hard to say what the main result is in the first part, since it seems the definitions are more important than the theorems. For this and the second part, the reader might wish to scan the table of contents to get a feel for what is going on. We give the following sample result to give an idea of the scope here. In ordinary small-cancellation theory, when $W_1, \ldots, W_r$ represent distinct conjugacy classes, the presentation $\langle a, b, \ldots \mid W_1^{n_1}, \ldots, W_r^{n_r} \rangle$ is "small-cancellation" for sufficiently large n_i. In analogy with this we have the following:

C6-sample. *Let X be a nonpositively curved cube complex. Let $Y_i \to X$ be a local-isometry with Y_i compact for $1 \leq i \leq r$ such that each $\pi_1 Y_i$ is malnormal, and $\pi_1 Y_i, \pi_1 Y_j$ do not share any nontrivial conjugacy classes. Then $\langle X \mid \widehat{Y}_1, \ldots, \widehat{Y}_r \rangle$ is a "small-cancellation" cubical presentation for sufficiently large "girth" finite covers $\widehat{Y}_i \to Y_i$.*

Many other general small-cancellation theories have been propounded. For instance two such graded theories directed especially towards Burnside groups were produced by Olshanskii and McCammond. Stimulated by Gromov's ideas of small-cancellation over word-hyperbolic groups, there have been later important works of Olshanskii, followed by more recent theories "over relatively hyperbolic groups" by Osin [Osi06] and Groves-Manning [GM08]. The theory we propose is decidedly more geometric, and arguably favors explicitness over scope. However, although it may be more limited by presupposing a nonpositively curved cube complex as a starting point, it has the advantage of not presupposing (relative) hyperbolicity—yet some form of hyperbolicity must lurk inside for there to be any available small-cancellation.

In the second part of the text we impose additional conditions that lead to the existence of a wallspace structure on the resulting small-cancellation presentation. We can illustrate the nature of the results with the following sample:

B6-sample. *Let G be an infinite word-hyperbolic group acting properly and cocompactly on a CAT(0) cube complex. Let $H_1, \ldots, H_k$ be quasiconvex subgroups that are not commensurable with G. And suppose that each H_i has separable hyperplane stabilizers. There exist finite index subgroups $H'_1, \ldots, H'_k$ such that the quotient $G/\langle\langle H'_1, \ldots, H'_k \rangle\rangle$ has a codimension-1 subgroup.*

Here $\langle\langle A, B, \ldots \rangle\rangle$ denotes the normal closure of $\{A \cup B \cup \cdots\}$ in the group.

In the third part of the text, we probe further and seek a virtually special cubulation.

We then prove the following:

Theorem A (Special Quotient Theorem)**.** *Let G be a word-hyperbolic group that is virtually the fundamental group of a compact special cube complex. Let $H_1, \ldots, H_r$ be quasiconvex subgroups of G. Then there are finite index subgroups $H'_i \subset H_i$ such that: $G/\langle\langle H'_1, H'_2, \ldots, H'_r \rangle\rangle$ is virtually special.*

We then prove the following:

Theorem B (Quasiconvex Hierarchy $\Rightarrow$ Virtually Special)**.** *Let G be a word-hyperbolic group with a quasiconvex hierarchy, in the sense that it can be decomposed into trivial groups by finitely many HNN extensions and amalgamated free products along quasiconvex subgroups. Then G is virtually special.*

There are two important applications of the virtual specialness of groups with a quasiconvex hierarchy: It is applied to hyperbolic 3-manifolds with a geometrically finite incompressible surface to reveal their virtually special structure. This resolves the subgroup separability problem for fundamental groups of such manifolds. It also completes a proof that Haken hyperbolic 3-manifolds are virtually fibered. It is also applied to resolve Baumslag's conjecture on the residual finiteness of one-relator groups with torsion.

The fourth part of the text deals with groups that are hyperbolic relative to virtually abelian subgroups, and provides similar structural results for many such groups when they also have quasiconvex hierarchies.

Chapter Two

CAT(0) Cube Complexes

2.a Basic Definitions

An *n-cube* is a copy of $[-\frac{1}{2},\frac{1}{2}]^n$, and a 0-*cube* is a single point. We regard the boundary of an n-cube as consisting of the union of lower dimensional cubes. A *cube complex* is a cell complex formed from cubes, such that the attaching map of each cube is combinatorial in the sense that it sends cubes homeomorphically to cubes by a map modeled on a combinatorial isometry of n-cubes. The *link* of a 0-cube v is the complex whose 0-simplices correspond to ends of 1-cubes adjacent to v, and these 0-simplices are joined up by n-simplices for each corner of an $(n+1)$-cube adjacent to v.

A *flag complex* is a simplicial complex with the property that each finite set of pairwise-adjacent vertices spans a simplex. A cube complex C is *nonpositively curved* if $\mathrm{link}(v)$ is a flag complex for each 0-cube $v \in C^0$.

Two-dimensional nonpositively curved complexes with one 0-cell, are a special case of the $C(4)$-$T(4)$ small-cancellation presentations that have old roots within combinatorial group theory. The nonpositively curved cube complexes were introduced to geometric group theory by Gromov in [Gro87] as a source of examples of high-dimensional metric spaces with nonpositive curvature. The supporting details of this theory were given by Moussong, Bridson, and Leary, in the locally finite, finite dimensional, and general cases. We refer to [Mou88, Lea13] but especially to [BH99] for a general account of CAT(0) geodesic metric spaces.

2.b Right-Angled Artin Groups

Let Γ be a simplicial graph. The *right-angled Artin group* or *raag* or *graph group* $G(\Gamma)$ associated to Γ is presented by:

$$\langle \ v : v \in \text{ vertices}(\Gamma) \ \mid \ [u,v] : (u,v) \in \text{ edges}(\Gamma) \ \rangle.$$

For our purposes, the most important example of a nonpositively curved cube complex arises from a right-angled Artin group. This is the cube complex $C(\Gamma)$ containing a torus T^n for each copy of the complete graph $K(n)$ appearing in Γ [CD95, MV95]. The cube complex $C(\Gamma)$ is sometimes called a *Salvetti complex.*

Each added torus T^n is isomorphic to the usual product $(S^1)^n$ obtained by identifying opposite faces of an n-cube. Note that $\pi_1 C(\Gamma) \cong G(\Gamma)$ since the 2-skeleton of $C(\Gamma)$ is the standard 2-complex of the presentation above.

To see that $C(\Gamma)$ is nonpositively curved we must show that link(a) is a flag complex where a is the 0-cube of $C(\Gamma)$. Each vertex of link(a) corresponds to an element of $\Gamma^0 \times \{\pm 1\}$. A set of vertices form an n-simplex precisely if they correspond to a corner of an $(n+1)$-cube of c, which holds precisely if they correspond to $n+1$ distinct generators oriented arbitrarily, that is, an $(n+1)$-clique of Γ with a ± 1 associated to each vertex. It is then clear that link(a) is simplicial as the intersection of simplices is a simplex. Moreover, link(a) is a flag-complex, since the 2^n different ways of orienting the vertices of an n-clique correspond to the 2^n different corners of the associated n-cube of c, and hence each collection of pairwise-adjacent vertices spans a simplex of link(a).

2.c Hyperplanes in CAT(0) Cube Complexes

Simply-connected nonpositively curved cube complexes are called *CAT(0) cube complexes* because they admit a CAT(0) metric where each n-cube is isometric to $[-\frac{1}{2}, \frac{1}{2}]^n \subset \mathbb{R}^n$; however we shall rarely use this metric.

The crucial characteristic properties of CAT(0) cube complexes are the separative qualities of their hyperplanes: A *midcube* is the codimension-1 subspace of the n-cube $[-\frac{1}{2}, \frac{1}{2}]^n$ obtained by restricting exactly one coordinate to 0. A *hyperplane* is a connected nonempty subspace of the CAT(0) cube complex C whose intersection with each cube is either empty or consists of one of its midcubes. The 1-cubes intersected by a hyperplane are *dual* to it. We will discuss immersed hyperplanes within a nonpositively curved cube complex in Section 6.a.

Remark 2.1. Hyperplanes in a CAT(0) cube complex C have several important properties [Sag95]:

(1) If D is a hyperplane of C then $C - D$ has exactly two components.
(2) Each midcube of a cube of C lies in a unique hyperplane.
(3) Regarding each midcube as a cube, a hyperplane is itself a CAT(0) cube complex.
(4) The union $N(D)$ of all cubes that D passes through is the *carrier* of D and is a convex subcomplex of C (see Section 2.d) that is isomorphic to $D \times I$.

Here $I = [-\frac{1}{2}, +\frac{1}{2}]$ is a 1-cube

2.d Geodesics and the Metric

Although we have defined the standard 1-cube to be a copy of $[-\frac{1}{2}, \frac{1}{2}]$, it will often be convenient to consider real intervals as 1-dimensional cube complexes whose vertices are the integer points. In particular, we let I_n denote the interval

$[0,n]$ subdivided so that all integers are vertices. A *length n path from x to y* in a cube complex X is a combinatorial map $I_n \to X$ where $0,n \mapsto x,y \in X^0$. A path is a *geodesic* if there is no shorter length path with the same endpoints. We emphasize that geodesics are almost never unique when $\dim(X) \geq 2$, indeed there are $n!$ distinct geodesics connecting vertices at opposite corners of an n-cube. We define the *distance* between 0-cubes in a connected nonpositively curved cube complex to be the length of a geodesic between them. As usual, this provides a genuine metric on the 0-cells of the 1-skeleton. Moreover we are then able to declare the distance $\mathsf{d}(A,B)$ between subcomplexes as the minimal distance $\mathsf{d}(a,b)$ where $a,b \in A^0, B^0$. We also define the *diameter* $\text{diam}(Y)$ of a connected complex to be the supremum of the lengths of geodesics in Y.

The combinatorial viewpoint we have adopted does not use the CAT(0) comparison metric, and we refer to [BH99] for an extensive account of that viewpoint—for cube complexes and many other spaces.

2.e Properties of Minimal Area Cubical Disk Diagrams

This section was motivated by lectures of Andrew Casson from the University of Texas at Austin in the '80s (apparently on generalized $C(4)$-$T(4)$ presentations related to Heegaard decompositions). I am grateful to Yoav Moriah who shared his notes with me and to Michah Sageev who encouraged me to take a look at this. Part of this material was explained using the alternate viewpoint of "pictures" in [Sag95, Sec 4.1]. While the results are easy, I had not previously considered the relevance of disk diagrams to cubical complexes of dimension ≥ 3. The viewpoint here, and in particular Lemma 2.3, is due to Casson. We note that the properties listed in Remark 2.1 can be deduced from this viewpoint.

A *disk diagram* D is a compact contractible combinatorial 2-complex with a chosen planar embedding $D \subset \mathbb{R}^2$. Its *boundary path* or *boundary cycle* $\partial_{\mathsf{p}} D$ is the attaching map of the 2-cell containing the point at ∞ where we regard $\mathbb{R}^2 \cup \infty$ to be the 2-sphere. The disk diagram D is *trivial* if it consists of a single 0-cell. A *spur* of D is an open edge in ∂D that ends on a valence 1 vertex of ∂D. Note that there is a spur for each backtrack in $\partial_{\mathsf{p}} D$. A 1-cell of D is *isolated* if it does not lie on the boundary of any 2-cell. A 0-cell v of D is *singular* if $\text{link}(v)$ is not isomorphic to a cycle, i.e., D does not look like $\mathbb{R}^2$ at v. The diagram D is *singular* if it has a singular 0-cell. Equivalently D is singular if it is not homeomorphic to a closed 2-ball, in which case D is either trivial, has a cut vertex, or consists of a single isolated edge.

We say D is a *square disk diagram* if it is a cube complex, i.e., all its 2-cells are squares. Many of the arguments below are by induction on $\mathsf{Area}(D)$ which equals the number of squares in D. (We note that there are instances where it is more natural to instead count the number of edges or even the number of cells in D.)

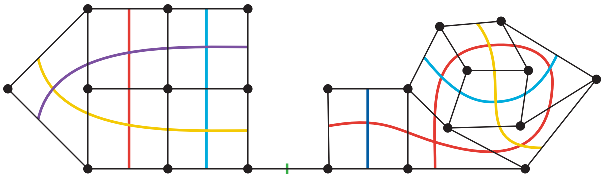

Figure 2.1. Dual curves in a square complex disk diagram.

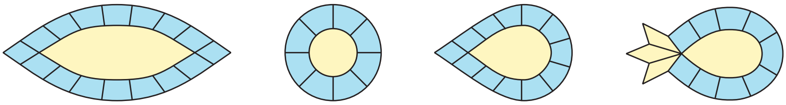

Figure 2.2. A bigon, nongon, monogon, and oscugon.

A *diagram in a complex* X is a combinatorial map $D \to X$ where D is a disk diagram. In this section, we study *cubical disk diagrams* which are disk diagrams in a nonpositively curved cube complex X. Of course, every cubical disk diagram is a square disk diagram.

We often use the following standard fact about the existence of disk diagrams (see [§2.2] [ECH$^+$92] or [LS77]):

Lemma 2.2 (van Kampen)**.** *A closed combinatorial path $P \to X$ is nullhomotopic if and only if there exists a diagram $D \to X$ with $P \cong \partial_{\mathsf{p}} D$ so that there is a commutative diagram:*

$$\begin{array}{ccc} \partial_{\mathsf{p}} D & \to & D \\ \| & & \downarrow \\ P & \to & X \end{array}$$

Let D be a square disk diagram. The *dual curves* in D are (noncombinatorial) paths that are concatenations of midcubes of squares of D. In addition, the midcube of an isolated edge of D provides a dual curve that is a trivial path. Note that when $D \to \widetilde{X}$ is a disk diagram in a CAT(0) cube complex, each dual curve maps to a hyperplane of $\widetilde{X}$.

The 1-cells crossed by a dual curve are *dual* to it. Note that each midcube lies in a unique maximal dual curve (or cycle). One simply extends outwards uniquely across dual 1-cells. A *bigon* is a pair of dual curves that cross at their first and last midcubes. A *monogon* is a single dual curve that crosses itself at its first and last midcubes.

An *oscugon* is a single dual curve that starts and ends at distinct dual 1-cells that are adjacent but don't bound the corner of a square. A *nongon* is a single dual curve of length ≥ 1 that starts and ends on the same dual 1-cell, so it corresponds to an immersed cycle of midcubes. We refer the reader to Figure 2.2.

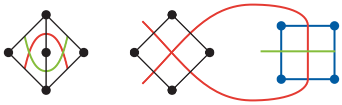

Figure 2.3. On the left is a smallest possible bigon. On the right is a monogon which must contain a smaller bigon.

Lemma 2.3. *Let $D \to X$ be a disk diagram in a nonpositively curved cube complex. If D contains a bigon, nongon, or oscugon, then there is a new diagram D' such that:*

(1) *D' and D have the same boundary path, so $\partial_{\mathsf{p}} D' \to X$ equals $\partial_{\mathsf{p}} D \to X$,*
(2) $\mathsf{Area}(D') \leq \mathsf{Area}(D) - 2$, *and*
(3) *pairs of edges on $\partial_{\mathsf{p}} D'$ that lie on the same dual curve of D' are precisely the same as pairs of edges on $\partial_{\mathsf{p}} D$ that lie on the same dual curve of D.*

Corollary 2.4. *No disk diagram contains a monogon.*

If D has minimal area among all diagrams with boundary path $\partial_{\mathsf{p}} D$, then D cannot contain a bigon, a nongon, or an oscugon.

Proof. The second statement follows immediately from Lemma 2.3. Consider a minimal area counterexample D to the first statement: So D contains an immersed rectangular strip $[-\frac{1}{2}, \frac{1}{2}] \times [0, n]$ of squares whose first and last square map to the same "cross-square," and this strip carries a dual curve σ at $\{0\} \times [0, n]$. We may assume $n > 2$ as if $n = 2$ then two adjacent edges at the corner of the cross-square are identified, and this violates the nonpositive curvature of the nonpositively curved cube complex X that D maps to, and if $n = 1$ then a square fails to embed in X. Choose m with $1 < m < n$. Then the 1-cube $\{\frac{1}{2}\} \times [m-1, m]$ is dual to a dual curve λ which must cross σ a second time. We can therefore apply Lemma 2.3 to replace D by D' and obtain a smaller area diagram that is still a counterexample by Condition 2.3.(3). $\square$

Proof of Lemma 2.3. **Reducing to the bigon case:** Consider a monogon, nongon, oscugon, or bigon within D that is *smallest* in the sense that the smallest subdiagram E containing it has minimal area. We first observe this smallest situation must arise from a bigon. Indeed, for a monogon, nongon, or oscugon, the associated dual curve α has length ≥ 1, for by the nonpositive curvature of the cube complex, squares locally embed, and so even for a monogon, the dual curve must pass through at least one more square besides its self-crossing square. Thus, as illustrated on the right in Figure 2.3, a second dual curve β crosses α and then must cross α a second time to exit. The pair α, β then provides a smaller situation. We next observe that α and β cannot contain a proper subpath that is a nongon or oscugon, for this would likewise lead to a smaller situation.

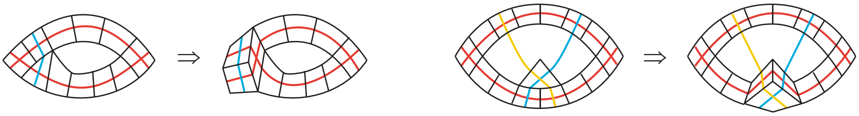

Figure 2.4. Some hexagon moves.

We are thus led to examine a *bigonal diagram* E which is a subdiagram with the property that each of its squares is either: one of the squares s_1, s_2 where α, β intersect; or a square with a midcube traversed by exactly one of α, β; or a square contained inside the bigon formed by α, β. Moreover, the bigonal diagram has the additional feature that the rectangles carrying α, β both embed.

Zipping a bigon: We now show that any bigonal subdiagram can be replaced by a disk diagram with the same boundary path but smaller area. Specifically, we will perform a slight modification to obtain a disk diagram with the same boundary but containing a smaller area bigonal diagram, and hence this smaller disk diagram itself can have its area reduced by 2.

The "base case" arises from two squares meeting along a corner as on the left in Figure 2.3. By nonpositive curvature, these two squares map to the same square in X, and hence we can remove this cancellable pair to decrease the area, by replacing the pair of squares by a pair of edges glued together at a point.

Observe that every dual curve in E other than α, β must pass through both α and β, since otherwise there would be a smaller bigon.

A *hexagon move* on a diagram D is the replacement of three squares forming a subdivided hexagon by an alternate three squares forming a subdivided hexagon. This corresponds to pushing a hexagon on one side of a 3-cube to obtain the hexagon on the other side.

The plan is to find a (certain type of) minimal triangle in the complement of the dual curves, and to then perform a hexagon move to obtain a new disk diagram with a smaller bigon as in Figure 2.4. The first type of minimal triangle has one side on α and one side on β and no dual curves passing through it. The second type has its base on α, and neither of its two other sides are subsegments of β.

If there is at least one crossing pair of dual curves as on the right of Figure 2.5, then we shall show below that the second type of triangle exists, and so we can perform a hexagon move of the second type. Hence by induction, the new diagram can have its area reduced by 2. If the bigon contains no crossing pair of dual curves as on the left in Figure 2.5, then the first type of triangle occurs, and so we can perform a hexagon move of the first type. We emphasize that a first type hexagon move can then be performed in either direction, i.e., at each corner (and this is the crucial point in obtaining Lemma 2.6 below).

Hexagon moves do not affect the boundary path (nor affect whether or not dual curves cross within the diagram); they simply adjust the route that dual curves take within a diagram, and hence Condition 2.3. (3) holds.

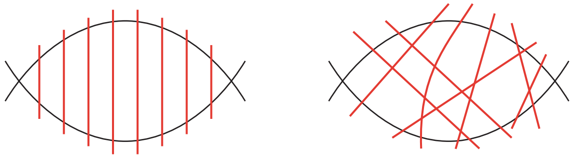

Figure 2.5. Dual curves of a bigonal diagram.

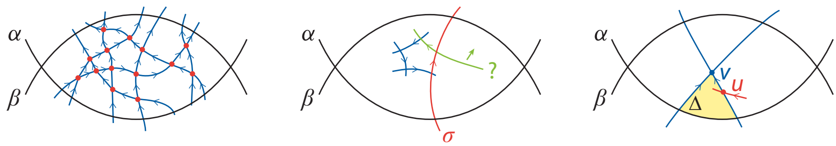

Figure 2.6. The directed graph Λ formed from dual curves cannot have a directed cycle.

A minimal triangle exists: The collection of dual curves within our bigon forms a graph Λ, and we make Λ into a directed graph by orienting all dual curves upwards from α to β, and thus orienting each edge of Λ (see the left of Figure 2.6). Observe that Λ has no directed cycle. Indeed, consider a directed cycle ξ. Suppose that ξ travels counterclockwise—as an analogous argument works in the clockwise case. Among the dual curves contributing edges to ξ, let σ denote the one having rightmost intersection with α. Let λ denote the next dual curve contributing an edge in the directed cycle ξ. Then λ would intersect α even further to the right which is impossible (see the middle of Figure 2.6). Here we use that each pair of dual curves intersect only once which follows from the minimality assumption on the bigon.

Each vertex of Λ (not on α, β) is the "top" of a triangle whose base is on α. Choose a vertex v that is minimal (excluding the leaf vertices on α) in the partial ordering induced by the directed graph with no directed cycles. Then the corresponding triangle Δ is our desired triangle of the second type. Indeed, if any other dual curve crosses either leg of Δ then there would be an even lower vertex u, which contradicts the minimality of v (as on the right of Figure 2.6). □

We shall later use the term *shuffle* to refer to an adjustment of a disk diagram obtained through a finite sequence of hexagon moves.

Definition 2.5 (Cornsquare)**.** Let $D \to X$ be a disk diagram. A *cornsquare* in D consists of a 2-cube c in D and dual curves α, β emanating from consecutive edges a, b of c that terminate on consecutive edges a', b' of $\partial_{\mathsf{p}} D$. The path $a'b'$ is the *outerpath of the cornsquare*.

We refer the reader to Figure 2.7. Note that we allow the possibility that there are squares in D between α, β. However, D can be shuffled so that there

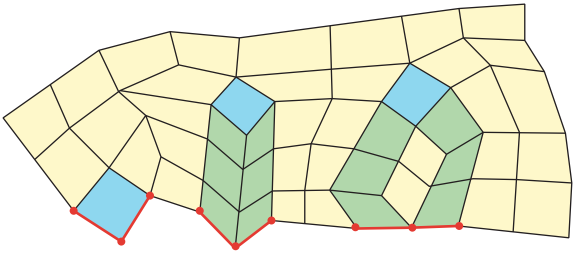

Figure 2.7. Three cornsquares in a disk diagram D. The outerpath of each is the corresponding red horizontal subpath of $\partial_{\mathsf{p}} D$. Note that D has other cornsquares besides these.

are no such squares, and moreover, following Lemma 2.6, D can be shuffled so that $a'b'$ actually forms the corner of a square.

The term "cornsquare" is short for "corner of generalized square" which captures the idea that there is a hidden square along $a'b'$, and although it might possibly be remote, it can be brought towards $\partial_{\mathsf{p}} D$ by shuffling. This notion arises again in Section 2.i and will play an important role in the more general context of Chapter 3.

The final part of the argument of Lemma 2.3 leads to the following useful point that we frequently employ.

Lemma 2.6 (Crossing pair has a square)**.** *Let $D \to X$ be a diagram in a nonpositively curved cube complex. Suppose D contains a cornsquare whose outerpath is $a'b'$. Then there is another diagram $D' \to X$ with $\mathsf{Area}(D') \leq \mathsf{Area}(D)$ such that D' contains a square whose boundary path contains $a'b'$.*

In particular, let $D \to X$ be a diagram containing dual curves α, β that are dual to consecutive edges a', b' of a square, and also dual to edges a, b with a common endpoint. Then the images of a, b in X bound the corner of a square of X.

Proof. The "zipping bigon" part of the proof of Lemma 2.3 only used that there was a square on *one* corner of the bigonal diagram. The sequence of moves either push squares outwards through the top or bottom dual curves, or they push hyperplanes past a', b' resulting in a shorter bigon. The final stage of this sequence is a diagram consisting of a single square on a, b. □

Remark 2.7. Let α, β be dual curves intersecting in a minimal area diagram $D \to X$. There is a cornsquare or spur in each of the four "quadrants" of D subtended by α, β. Indeed, let aPb be a subpath of $\partial_{\mathsf{p}} D$, with a dual to α and b dual to β, that does not contain a backtrack. Consider an innermost pair e_1, e_2 of edges in aPb whose dual curves σ_1, σ_2 are either equal or cross. Then e_1, e_2 must be consecutive and hence provide a cornsquare or spur. Indeed, an edge e_3

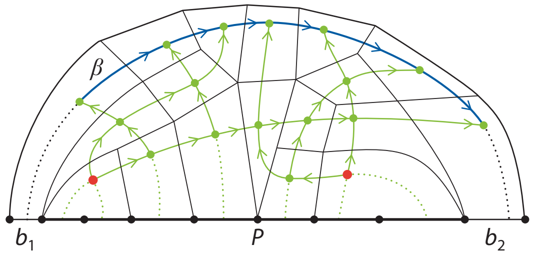

Figure 2.8. The digraph Λ consists of parts of the dual curves in D between b_1Pb_2 and β. The two vertices without ancestors correspond to squares with a corner on b_1Pb_2.

between e_1, e_2 would have a dual curve σ_3 that either returned to aPb between e_1, e_2 or that crosses σ_1 or σ_2, and this violates innermostness of e_1, e_2.

We now repeat the argument at the end of the proof of Lemma 2.3 to glean a bit more information:

Lemma 2.8 (Square or spur on each side)**.** *Let $D \to X$ be a minimal area disk diagram in a nonpositively curved cube complex. Let β be a dual curve in D that starts and ends on edges b_1, b_2 where b_1Pb_2 is a subpath of $\partial_{\mathsf{p}} D$. Then b_1Pb_2 contains a length 2 subpath e_1e_2 such that either e_1e_2 is a backtrack at a spur of D, or e_1e_2 bounds the corner of a 2-cube of D.*

Proof. It suffices to assume that β is innermost, in the sense that the curve dual to each edge of P must cross β. If P is trivial, then since D has no oscugon by Corollary 2.4, we see that b_1, b_2 traverse the same edge, which is a spur and we are done with $b_1, b_2 = e_1, e_2$. More broadly, innermostness allows us to assume that $b_1Pb_2 \to D$ embeds, for otherwise there would be a cutpoint in D, which would subtend a diagram containing a dual curve that violates the innermost assumption.

Consider the graph Λ whose vertices are centers of squares in the P-component of $D - \beta$, and whose edges are parts of dual curves joining centers of adjacent squares. We refer the reader to Figure 2.8. We direct Λ by directing β from b_1 to b_2, and directing all other dual curves from P to β. As in the end of the proof of Lemma 2.3, there is no directed cycle ξ in Λ. Indeed, ξ cannot have an edge on β, for then an edge of Λ is directed away from β which contradicts that dual curves are directed from P to β. Regard traveling up towards β and then in the direction of β as "clockwise." Suppose ξ were a cycle not containing an edge on β and assume it travels counterclockwise (a similar argument works in the clockwise case). Among the dual curves forming ξ, let σ denote the one which intersects P closest to b_2. As there are no bigons or monogons by Lemma 2.3, the next dual curve λ that ξ provides would have to intersect P closer to b_2, which is impossible.

Since Λ has no directed cycle, it has a vertex v with no ancestor. Then v is the center of a square s with consecutive edges e_1e_2 forming a subpath of P. Indeed, each dual curve (other than β) travels from P to β, for otherwise we would either contradict that β is innermost, or there would be a bigon. Hence the two incoming dual curves at v arrive from edges of b_1Pb_2 that lie on s, as otherwise v would have an ancestor. □

2.f Convexity

Although the basic properties of CAT(0) cube complexes and their hyperplanes have been explained many times in the literature, the reviewers have asked me to sketch some of these properties from the diagrammatic viewpoint elaborated upon in Section 2.e.

A subcomplex $Y \subset X$ of the CAT(0) cube complex is *convex* if for each pair of vertices $a, b \in Y^0$, each (combinatorial) geodesic joining a, b lies in Y.

Lemma 2.9. *The intersection of two convex subcomplexes is a convex subcomplex.*

Proof. This follows immediately from the definitions. □

Lemma 2.10 (Helly property)**.** *Let X be a CAT(0) cube complex. Let $Y_1, \ldots, Y_n$ be finitely many convex subcomplexes. Suppose $Y_i \cap Y_j \neq \emptyset$ for each i, j. Then $\cap_{i=1}^n Y_i \neq \emptyset$.*

Proof. We first show this is true in the base case when $n=3$. For $i \neq j$ let x_{ij} be a vertex in $Y_i \cap Y_j$. Let P_i be a geodesic in Y_i from x_{ki} to x_{ij}. Let D be a disk diagram for $P_1P_2P_3$. Finally, choose the above such that $\mathsf{Area}(D)$ is minimal.

Consider a square s in D. Since each P_i is a geodesic, no dual curve in D has both ends on the same P_i. Thus by the pigeon-hole principle, for some i, each dual curve through s has an end on P_i. We thus have a cornsquare on P_i and hence after shuffling we can reduce the area which is impossible. Thus D is a tripod, and its central point is an element of $Y_1 \cap Y_2 \cap Y_3$.

We now use the base case to help us prove the result by induction: For $1 \leq i < n$ let $Y_i' = Y_i \cap Y_n$. Since Y_i, Y_j, Y_n have pairwise nonempty intersection by hypothesis, the special case implies that $Y_i' \cap Y_j' = Y_i \cap Y_j \cap Y_n \neq \emptyset$. Thus $\cap_{i=1}^n Y_i = \cap_{i=1}^{n-1} Y_i' \neq \emptyset$ by induction. □

An *immersion* is a local injection. A map $\phi: Y \to X$ between nonpositively curved cube complex is a *local-isometry* if it is an immersion and for each $y \in Y^0$, whenever u, v are ends of 1-cubes at y, if $\phi(u), \phi(v)$ form a corner of a 2-cube in X at $\phi(y)$, then u, v form a corner of a 2-cube in Y. A subcomplex that embeds by a local-isometry is *locally-convex.* A connected locally-convex subcomplex

$\widetilde{Y}$ of a CAT(0) cube complex $\widetilde{X}$ is called *convex*. Equivalently, a connected subcomplex $Y \subset X$ of a CAT(0) cube complex is convex if: for each cube c of X with $\dim(c) \geq 2$, if an entire corner of c lies in Y then all of c lies in Y. It can be deduced from the viewpoint in Section 2.e that for a CAT(0) cube complex $\widetilde{X}$, a subcomplex $\widetilde{Y} \subset \widetilde{X}$ is convex if and only if firstly: an n-cube lies in $\widetilde{Y}$ precisely when its $(n-1)$-skeleton lies in $\widetilde{Y}$, and secondly: P lies in $\widetilde{Y}$ whenever $P \to \widetilde{X}$ is a geodesic path whose endpoints lie in Y^0.

The combinatorial notion of convexity we employ here is consistent with the usual notion of convexity one encounters for geodesic metric spaces. Indeed, a subcomplex $\widetilde{Y} \subset \widetilde{X}$ is "combinatorially convex" as defined above precisely when it is "metrically convex" (in the CAT(0) metric) in the sense that $P \subset \widetilde{Y}$ whenever $P \to \widetilde{X}$ is a (not necessarily combinatorial) geodesic with endpoints in $\widetilde{Y}$.

Lemma 2.11 (Locally-convex $\Rightarrow$ convex)**.** *Let X and Y be CAT(0) cube complexes. Let $Y \to X$ be a local-isometry. Then $Y \to X$ is an embedding, and its image is a convex subcomplex.*

In particular, a connected locally-convex subcomplex is convex.

Proof. Consider a geodesic $P \to X$ that is path-homotopic to a path $Q \to Y \to X$. Let D be a disk diagram between $P \to X$ and $Q \to X$. Assume $(\mathsf{Area}(D), |Q|)$ is minimal in the lexicographical order among all possible choices with P fixed. There is no cornsquare on Q for otherwise we could shuffle to obtain a smaller diagram D' between Q' and P. Thus each dual curve starting on Q ends on P and no two cross. Suppose D contains a square s. Then at most one end of one dual curve through s ends on Q, and so $|P| \geq |Q| + 2$, so P is not a geodesic. Thus D is a line and $P = Q$. Thus $Y \to X$ is an isometry, and in particular an embedding. Moreover, a geodesic P in X with endpoints in Y satisfies $P = Q \subset Y$. $\square$

Corollary 2.12 (Local-isometry π_1-injects)**.** *If $Y \to X$ is a local-isometry of nonpositively curved cube complexes, then $\pi_1 Y \to \pi_1 X$ is injective.*

Proof. If $\sigma \to Y$ represents a nontrivial element of $\pi_1 Y$ then its lift $\tilde{\sigma} \to \widetilde{Y}$ is not closed. Hence the image $\sigma \to X$ of σ represents a nontrivial element in $\pi_1 X$ since its lift $\tilde{\sigma} \to \widetilde{X}$ is also not closed, by Lemma 2.11. $\square$

2.g Hyperplanes and Their Carriers

Let M be the disjoint union of all midcubes of a cube complex X. Let $\bar{M}$ be the quotient of M obtained by identifying each midcube with the subcube of each larger midcube that it lies in. Note that the map $M \to X$ induces a continuous map $\bar{M} \to X$. An *(abstract) hyperplane* of X is a component U of $\bar{M}$.

The *(abstract) carrier* of U is defined to be $N(U)=U\times I$, and we define $N(U)\to X$ such that for each cube m of U, we have $m\times I$ maps isomorphically to the cube c containing m as a midcube.

Lemma 2.13 (Hyperplanes exist). *Let X be a CAT(0) cube complex. Every midcube of a cube of X lies in a hyperplane of X.*

Proof. Let U be a component of $\bar{M}$. Suppose that $U\to X$ is not injective. Then there is a pair of midcubes of U mapping to the same cube of X. We can assume these are 1-dimensional midcubes mapping to the same 2-cube s. Let a,b be the consecutive edges of the 2-cube that these midcubes end on. There is then a sequence of 1-midcubes joining them, and we thus obtain a corresponding rectangular strip R that starts with a and ends with b. Let P be a path along one side of this strip. By possibly extending R by adding a copy of s, we may assume that P is a closed path in X, and so P bounds a disk diagram $E\to X$. The union $D=E\cup_P R$ is another disk diagram that contains an oscugon associated to the dual curve carried by R. Applying Lemma 2.3, we obtain a new disk diagram D' with the same boundary path as D, but having no oscugons. Thus the initial and terminal edge of R are identified to a spur in D' under the composition $R\to D\to D'$. In particular they map to the same edge in X. This contradicts that a,b are not parallel in s. □

Lemma 2.14 (CAT(0) hyperplane properties). *Let X be a CAT(0) cube complex.*

(1) *The map $N(U)\to X$ is an embedding for each hyperplane U.*
(2) *$N(U)\subset X$ is a convex subcomplex.*
(3) *Each hyperplane is simply-connected.*
(4) *Let U,V be hyperplanes of X that* cross *in the sense that they contain distinct midcubes in a cube. If U,V are dual to 1-cubes a,b that share a 0-cube, then a,b lie in a common 2-cube.*

Proof. We first show that the map $\phi:N(U)\to X$ is a local-isometry. Let a,b be edges at a 0-cube v of $N(U)$, and suppose $\phi(a),\phi(b)$ bound the corner of a square in X. If one of a,b is dual to U then a,b form the corner of a square in $N(U)$. If neither is dual to U, then they each form a square with the edge c dual to U at v, and thus by nonpositive curvature of X, there is a 3-cube bounded by the three squares with corners at $\phi(a),\phi(b),\phi(c)$. This 3-cube contains a midcube m that is part of U, and a,b bound the corner of a square parallel to m in $m\times I$.

As we have verified that $N(U)\to X$ is a local-isometry, the convexity of $N(U)\subset X$ holds by Lemma 2.11.

We now show that $N(U)$ and hence U is simply-connected. Consider an essential closed path $P\to N(U)$, such that $P\to X$ is nullhomotopic. Let $D\to X$ be a disk diagram with boundary path D. We moreover choose the above such that $(\mathsf{Area}(D),|P|)$ is minimal. If D has a spur then we can shorten P. Otherwise, an innermost pair of edges on P with crossing dual curves yields a cornsquare in D. By Lemma 2.6 we can shuffle to obtain a new disk diagram with the same area but with a genuine square having a corner on P. By local convexity, we can remove it to obtain a smaller area counterexample D', P'.

Property (4) holds by Lemma 2.6. Indeed, let a', b' denote the edges at the square s' where the hyperplanes U, V cross. Choose rectangular strips R_a, R_b that start at the square s' and end at the edges a, b. Let E be a disk diagram between paths P_a, P_b along the bases of R_a, R_b. Let $D \to X$ be the diagram formed by $R_a \cup_{P_a} D \cup_{P_b} R_b$. Then Lemma 2.6 provides a square at a, b.

Alternatively, we sketch an explanation depending on convexity (which in turn depended on Lemma 2.6): Note that $N(U) \cap N(V)$ is a convex subcomplex which contains a square s' as well as a vertex $a \cap b$. Consider a length n geodesic joining them, and then verify that it extends to a product $I \times I \times [0, n]$ where $I \times I \times \{0\}$ maps to s' and $I \times I \times \{n\}$ maps to a square at a, b. □

Let U be a hyperplane in a CAT(0) cube complex X. Let $N^o(U)$ be the *open carrier* consisting of the open cubes intersecting U. We refer to each component of $N(U) - N^o(U)$ as a *frontier* of U. Each frontier is a subcomplex $U \times \{\pm\frac{1}{2}\}$ if we identify $N(U)$ with $U \times [-\frac{1}{2}, \frac{1}{2}]$.

The complement $X - U$ consists of two subspaces called *halfspaces*. Each halfspace is associated with two *combinatorial halfspaces*, namely, the smallest subcomplex containing it, and a largest subcomplex contained in it, and these are referred to as a *major halfspace* and *minor halfspace*. Note that the major and minor halfspaces meet along a frontier F, and are the closures of components of $X - F$.

Another convenient property of hyperplanes is that:

Corollary 2.15. *Let U be a hyperplane of the CAT(0) cube complex X. Each frontier of U is convex. Each major and minor halfspace is convex.*

Proof. Each subcomplex $U \times \{\pm\frac{1}{2}\}$ is a convex subcomplex of $N(U)$, and thus convex in X as it is a convex subcomplex of a convex subcomplex.

The closure of each component of $X - (U \times \{\pm\frac{1}{2}\})$ is convex since its boundary $U \times \{\pm\frac{1}{2}\}$ is convex. Indeed, it is sufficient to consider geodesics starting and ending on the boundary of a component, which is a frontier. □

Corollary 2.16. *Let γ be a combinatorial path in a CAT(0) cube complex. Then γ is a geodesic if and only if the hyperplanes dual to the edges of γ are distinct. Thus the distance between 0-cells equals the number of hyperplanes separating them.*

Proof. Suppose γ is a geodesic, and suppose $e_1\gamma' e_2$ is a subpath where e_1, e_2 are dual to the same hyperplane H. Then $e_1\gamma' e_2$ has endpoints on one of the subcomplexes $H \times \{\pm\frac{1}{2}\} \subset N(H)$, which is convex by Corollary 2.15, but e_1 and e_2 do not lie in this subcomplex, and so its convexity is contradicted.

Conversely, suppose γ has the property that the hyperplanes dual to its edges are distinct. Let γ' be a shortest subpath of γ that is not a geodesic. Let σ be a geodesic with the same endpoints as γ'. Then $|\sigma| \geq |\gamma'|$ since each hyperplane dual to an edge of γ' must separate the endpoints of σ and is hence dual to an odd number of its edges. $\square$

The *convex hull* of a subset $S \subset X$ is the subcomplex $\text{hull}(S) \subset X$ that is the smallest convex subcomplex of X containing S. Corollary 2.15 and Lemma 2.9 imply that $\text{hull}(S)$ lies in the intersection of all minor halfspaces containing S. We will see from Lemma 2.19 that $\text{hull}(S)$ equals this intersection.

For 0-cubes p, q in a CAT(0) cube complex X, the *interval* $\mathcal{I}(p,q)$ is defined by $\mathcal{I}(p,q) = \text{hull}(\{p,q\})$.

Lemma 2.17. *For each 0-cube $k \in \mathcal{I}(p,q)$ we have $\mathsf{d}(p,k) + \mathsf{d}(k,q) = \mathsf{d}(p,q)$. Equivalently, each 0-cell $k \in \mathcal{I}(p,q)$ lies on a geodesic from p to q.*

Proof. By the triangle inequality, it suffices to verify that $\mathsf{d}(p,k) + \mathsf{d}(k,q) \leq \mathsf{d}(p,q)$. By Corollary 2.16, $\mathsf{d}(a,b)$ equals the number of hyperplanes separating a, b. First observe that each hyperplane separating exactly one of p, k and k, q, must also separate p, q. Secondly, we verify that no hyperplane H separates both p, k and k, q. Indeed, then $\{p,q\}$ lies in one halfspace of H but k lies in the other. Hence K lies in the minor halfspace of H containing $\{p,q\}$, but k does not, so $k \notin \mathcal{I}(p,q)$. $\square$

Lemma 2.18. *Let P, Q be convex subcomplexes of the CAT(0) cube complex X. Consider all paths that start at a vertex of P and end at a vertex of Q, and let γ have minimal length among all such paths. Then every edge of γ is dual to a hyperplane that separates P, Q.*

Proof. Let p, q be the endpoints of γ in P, Q. Let $\mathcal{I} = \mathcal{I}(p,q)$ be the interval consisting of the convex hull of $\{p,q\}$. Let H be a hyperplane that separates p, q but intersects P. Since $\mathcal{I}, N(H), P$ pairwise intersect, they must triply intersect by Lemma 2.10, so there exists a 0-cube $k \in \mathcal{I} \cap N(H) \cap P$. However $\mathsf{d}(k,q) < \mathsf{d}(p,q)$ by Lemma 2.17, and since $k \in P$ this contradicts the choice of p, q. $\square$

As mentioned earlier, minor halfspaces play the following useful role:

Lemma 2.19. *A subcomplex Y of a CAT(0) cube complex X is convex if and only if Y is the intersection of minor halfspaces.*

Proof. The intersection of minor halfspaces is convex by Corollary 2.15 and Lemma 2.9. We now show that if $Y \subset X$ is convex, then Y is the intersection of minor halfspaces. Let p be a 0-cube in $X - Y$. Let γ be a geodesic fom Y to p.

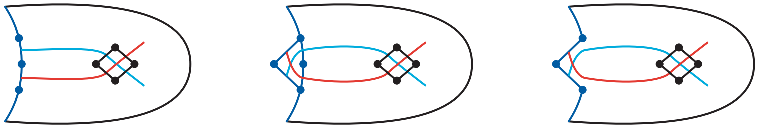

Figure 2.9. We first add a square along $Q \to Y$ to obtain a bigon, and are then able to reduce the area by 2.

By Lemma 2.18, each hyperplane U dual to an edge of γ separates Y, p. Thus a minor halfspace of U contains Y but not p. □

2.h Splaying and Rectangles

We now describe several related properties concerning the dual curves in minimal area cubical disk diagrams. We emphasize that our treatment focuses on subcomplexes and exclusively considers paths that are combinatorial, as discussed in Section 2.d.

The following is implicit in the proof of Lemma 2.11.

Lemma 2.20 (Splayed). *Let $Y \subset X$ be a convex subcomplex of a CAT(0) cube complex. Let P be a path whose endpoints lie on Y, and let D be a disk diagram between P and Y, so there is a [geodesic] immersed path $Q \to Y$ with the same endpoints as P, and D is a diagram for PQ^{-1}. Suppose D has minimal area among all possible such choices fixing P and Y.*

Then there is no intersection in D between dual curves starting on distinct 1-cells of Q.

The statement of Lemma 2.20 holds with Q allowed to vary either among all such immersed paths, or among all such geodesics. Indeed, the argument by contradiction given below provides a lower area diagram D without increasing the length of Q.

Proof. We first show that when a and b are consecutive 1-cells in Q, then the dual curve in D starting at a is disjoint from the dual curve starting at b.

Suppose a, b are parallel in D to 1-cells a', b' that meet at the corner of a square c' in D. Since X is CAT(0), by Lemma 2.14.(4), the 1-cells a, b must also meet at a square c. Since Y is convex, we see that $c \subset Y$.

We can thus adjust the diagram D to obtain a new diagram D' formed by attaching c to Q along a, b. Now $\mathsf{Area}(D') = \mathsf{Area}(D) + 1$. However, D' contains a bigon, and therefore by Lemma 2.3, its area can be reduced by two, to obtain a new diagram D'' with $\mathsf{Area}(D'') < \mathsf{Area}(D)$. This would contradict the minimality of D. See Figure 2.9.

Suppose there is a dual curve in D that starts on a and ends on b. Since $Q \to Y$ is an immersed path, the edges a, b in D are distinct. The dual curve

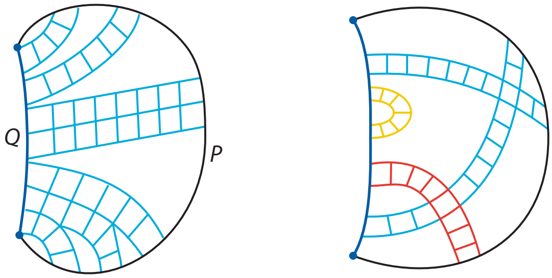

Figure 2.10. As on the left, dual curves emanating from edges of $Q \to D$ are splayed in the minimal area diagram D between P and the convex subcomplex $Y \subset X$. The configurations on the right cannot occur.

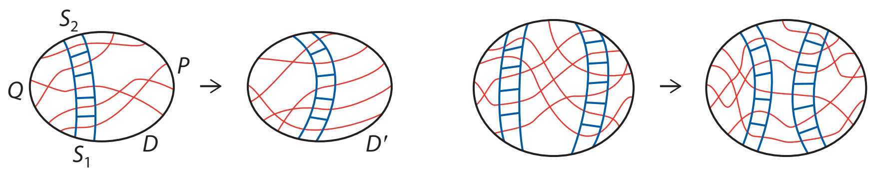

Figure 2.11.

thus provides an oscugon in D. By Lemma 2.3, there is a new diagram D' with the same boundary path, but $\mathsf{Area}(D') < \mathsf{Area}(D)$.

The general statement holds by considering an innermost pair of 1-cells whose dual curves are either equal or intersect. As proven above, these 1-cells cannot be adjacent. But any 1-cell on Q between them would give another dual curve, which either intersects one of these, or ends on another 1-cell of Q lying between them as in the right in Figure 2.10. This contradicts our innermost assumption. $\square$

Lemma 2.21 (Pushing beyond crossings)**.** *Let $D \to X$ be a minimal area disk diagram. Let S be a rectangular strip carrying a dual curve in D that starts and ends on 1-cells s_1, s_2 such that $\partial_{\mathsf{p}} D$ is of the form $s_1 P s_2 Q$. There exists a new diagram D' with $\partial_{\mathsf{p}} D = \partial_{\mathsf{p}} D'$ and $\mathsf{Area}(D') = \mathsf{Area}(D)$ such that s_1, s_2 are still connected by a strip S' but the dual curves emanating from S' to P are splayed: No two cross each other on the P side of S'.*

We refer to the left pair of diagrams in Figure 2.11 indicating the total transformation from D to D'.

Proof. This follows by repeatedly using hexagonal replacement moves. Consider an innermost pair of a, b of edges along ∂S whose dual curves cross on the side bounded by P. If they are not adjacent, then they are not innermost. Note that

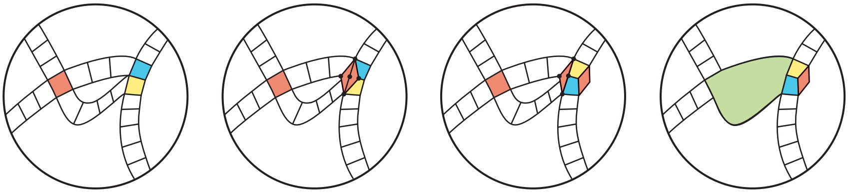

Figure 2.12.

the two dual curves cannot equal each other, or there would be a bigon with S, and thus the area can be reduced by Lemma 2.3.

Let c be the first square where the dual curves cross. Add a cancellable pair of copies c'', c' of c along a, b. This increases the area by 2, and increases the area between S and P by 2. Perform a hexagonal replacement along S and the contiguous copy c' of c to obtain S'; the area between P and S' is now one more than the area between P and S was. Finally the copy c'' of c has a bigon with c. We are able to reduce the area by 2. This area reduction is on the P side of S', and so the resulting diagram D' has the property that the area between P and S' has been reduced by one. See the sequence of pictures in Figure 2.12 for a single transformation. Performing this repeatedly yields a new diagram where S' has splayed strips on the P side, as claimed. □

Remark 2.22. We can apply Lemma 2.21 to understand the potential behavior between rectangular strips in disk diagrams. Let D be a diagram that has a pair of disjoint strips. Then we can replace it with a new diagram with the same boundary and at most as much area, such that the strips are moved inwards towards each other, but strips emanating from them are now splayed. See the transformation on the right in Figure 2.11.

This is particularly relevant when we consider a diagram between two convex subspaces Y_1, Y_2, and in particular, a diagram between a convex subspace and the carrier of a hyperplane. We are able to reach the conclusion of a "flat rectangle" between the rectangular strips.

A *grid* is a complex isomorphic to $I_m \times I_n$ for some m, n. We record the following easy observation. We revisit this later with pseudo-grids in Section 3.q.

Lemma 2.23. *A square disk diagram D is a grid if and only if:*

(1) *$\partial_{\mathsf{p}} D$ is a concatenation $\partial_{\mathsf{p}} D = V_1 H_1 V_2 H_2$.*
(2) *Each dual curve is either* vertical *and ends on V_1, V_2 or is* horizontal *and ends on H_1, H_2.*
(3) *Each horizontal dual curve crosses each vertical dual curve exactly once.*
(4) *Horizontal dual curves do not cross, and vertical dual curves do not cross.*

Corollary 2.24 (Grid with tails)**.** *Let $D' \to X$ be a minimal area disk diagram in a nonpositively curved cube complex. Suppose $\partial_{\mathsf{p}} D' = V_1' H_1' V_2' H_2'$, and D' has no spur or cornsquare with outerpath on V_1', H_1', V_2', or H_2'. Then D' is the union of a grid D together with a (possibly trivial) arc A_{ij} attached at the four corners of D, where $\partial_{\mathsf{p}} D = V_1 H_1 V_2 H_2$, and the corners are at the four concatenation points.*

Proof. Suppose a dual curve σ starts on V_i' and ends on H_j'. We claim σ is trivial, as otherwise, σ crosses another dual curve σ', and since σ, σ' intersect at a single point, σ' has an endpoint on V_i' or H_j', so there is a cornsquare on V_i' or H_j'. Thus D' has four (possibly trivial) arcs $A_{ij} = V_i' \cap H_j'$, and removing these four arcs yields a diagram D that is a grid by Lemma 2.23. □

2.i Annuli

This section can be postponed until annuli arise in Chapter 14 and more importantly Section 5.o and its sequels.

An *annular diagram* is a compact complex A with $\pi_1 A \cong \mathbb{Z}$ such that there is a chosen planar embedding $A \subset \mathbb{R}^2$. The annular diagram has two *boundary paths* or *boundary cycles* which correspond to the attaching maps of the two 2-cells that can be added to A to form a 2-sphere S^2.

An *annular diagram in a complex* X is a combinatorial map $A \to X$ where A is an annular diagram. It is natural to refer to A as an annular diagram *between* P_1 and P_2 as A indicates the homotopy between them as in the following standard analog of Lemma 2.2 (see [LS77]): Let $P_1 \to X$ and $P_2 \to X$ be closed paths in X. Then there is a homotopy between them if and only if there is an annular diagram $A \to X$, such that each $P_i \to X$ factors as $P_i \to A \to X$ where each $P_i \to A$ is a boundary path of A. Moreover, identifying the subdivided circles P_1 with $S^1 \times \{0\}$ and P_2 with $S^1 \times \{1\}$, the homotopy $S^1 \times [0,1] \to X$ factors as $S^1 \times [0,1] \to A \to X$.

We say A is *singular* if A is not homeomorphic to a cylinder, and we adopt the terminology used for disk diagrams: isolated 1-cell, singular 0-cell, spur, and so forth. We now turn to studying annular diagrams in a nonpositively curved cube complex X. The annular diagram is then a square complex, and we define its *dual curves* as we did for a disk diagram. We define a *cornsquare* in A as in Definition 2.5 except that we require that the dual curves emanating from the cornsquare do not enclose a boundary path of A—as for instance in the third annulus in Figure 2.13.

A *flat annulus* is an annular diagram A with the property that each dual curve is either closed or has an end on each boundary path of A, and that for each square, at least one of its dual curves is closed.

Lemma 2.25 (Flat annulus)**.** *Let $A \to X$ be an annular diagram. Suppose there is no spur and no cornsquare with outerpath on a boundary path of A. Then A is a flat annulus.*

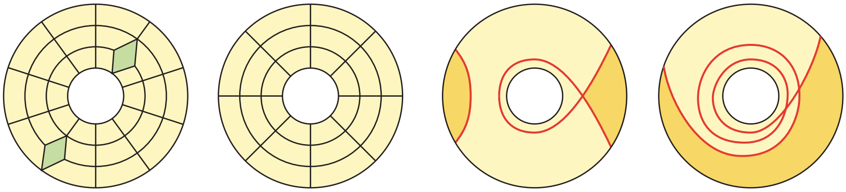

Figure 2.13. Two flat annuli are illustrated on the left. The second is a product. The first has only one closed dual curve, and it self-crosses several times. The third and fourth figure illustrate the situation where a dual curve has both ends on the same boundary path, in which case one can find a cornsquare.

We are especially interested in flat annuli that are minimal area, in which case they do not contain any bigons. We note however that a flat annulus A that is not a product will have a finite cover that contains bigons. The first annulus in Figure 2.13 illustrates a simple but typical example of the type of annulus examined in Lemma 2.25. This contrasts with the motivating case of a product, illustrated by the second annulus in Figure 2.13. The reader can imagine more elaborate examples.

Proof. Suppose d is a dual curve that starts and ends on the outer boundary path. We will show that there is a cornsquare. The analogous argument works when d starts and ends on the inner boundary path.

If d doesn't cross itself, then we choose the side of d not containing the inner boundary path. If d crosses itself, then we consider a minimal initial and terminal part of d that cross at some square, and choose the side of the diagram they subtend which does not contain the inner boundary path. The "chosen" sides are shaded in the third and fourth diagrams in Figure 2.13. In either case, our chosen side of the diagram contains an innermost pair of dual curves that cross each other, and within this lies the claimed cornsquare, as in Remark 2.7.

Now suppose there is a square s with two dual curves that end on the same boundary path of A. This forms a "triangle" whose top is in s and whose base is on ∂A. An innermost such triangle yields a cornsquare in A. □

Remark 2.26. The dual curves not ending on ∂A can travel around A, but we cannot always choose the annular diagram so that these dual curves do not self-cross. For instance, we refer the reader to the annular diagram at the very left of Figure 2.13. While minimal area of the diagram can help avoid some such self-crossing behavior, there is no way to avoid it in general, and we can only conclude that the "horizontal" dual curves travel "around" A, possibly multiple times. Of course, in the special case when immersed hyperplanes of X do not cross themselves, self-crossing of dual curves is impossible for any annular diagram mapping to X.